Physical Metallurgy

2

Robert W. Cahn and Peter Haasen (†), editors

FOURTH, REVISED AND ENHANCED EDITION

NORTH-HOLLAND

Prof. Robert W. Cahn, editor

PHYSICAL METALLURGY

VOLUME II

LIST OF CONTRIBUTORS

A. S. Argon
E. Arzt
H. K. D. H. Bhadeshia
H. Biloni
J. L Bocquet
W. J. Boettinger
G. Brebec
R. W. Cahn
G. Y. Chin†
T. W. Clyne
R. D. Doherty
H. E. Exner
R. Ferro
D. R. Gaskell
H. Gleiter
A. L. Greer
P. Haasen†
J. P. Hirth
S. Hofmann
E. D. Hondros
E. Hornbogen
G. Kostorz
C. Laird
P. Lejček
W. C. Leslie
Y. Limoge
J. D. Livingston
F. E. Luborsky
T. B. Massalski
J. R. Nicholls
A D. Pelton
D. G. Pettifor
D. P. Pope
M. Rühle
A. Saccone
S. R. J. Saunders
M. P. Seah
W. Steurer
J.-L. Strudel
R. M. Thomson
C. M. Wayman
M. Wilkens
A. H. Windle
H. J. Wollenberger

PHYSICAL METALLURGY

Fourth, revised and enhanced edition

Edited by

Robert W. CAHN
University of Cambridge

Peter HAASEN†
University of Göttingen

VOLUME II

1996

NORTH-HOLLAND
AMSTERDAM—LAUSANNE—NEW YORK—OXFORD—SHANNON—TOKYO

ELSEVIER SCIENCE B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

ISBN: 0 444 89875 1

This book is printed on acid-free paper.

Printed in The Netherlands

SYNOPSIS OF CONTENTS

Volume 1

Volume 2

Volume 3

CHAPTER 10

QUALITATIVE AND QUANTITATIVE SURFACE MICROSCOPY

H.E. EXNER

Technical University Darmstadt
Department of Materials Science
D-64287 Darmstadt, Germany

R. W. Cahn and P. Haasen†, eds.
Physical Metallurgy; fourth, revised and enhanced edition

1. Introduction

In technical materials, the microstructure develops during processing. Apart from rare cases where microstructural features persist unchanged in nature and geometry from the raw material to the final product (e.g., hard nonmetallic refractory particles), each individual processing step influences the amount, the composition and the geometric appearance of the constituents and of defects in a material. This is due to the fact that microstructures usually are far from the ideal thermodynamic and geometric equilibrium predicted by thermodynamic (or rather thermostatic) considerations (see ch. 5 on metallurgical thermodynamics and ch. 6 on phase diagrams). The usual route — casting, plastic forming, heat-treating — leads to microstructures vastly different from those obtained after powder-metallurgical production, for example. Vice versa, the mechanics of the individual processing techniques can be best studied by monitoring the microstructural changes as a function of processing conditions (compare, for example, ch. 8 on solidification; ch. 9, § 2; chs. 15 and 16 on diffusive and nondiffusive phase transformation, respectively, or ch. 28 on recrystallization). Knowledge of the details of the formation of microstructures is essential in order to understand the relationships between processing parameters and the behaviour of materials in practical application. Since the most important technological properties are strongly influenced by the microstructure (see, for example, ch. 25 on the mechanical properties of multiphase alloys) this understanding is important for the development of metallic (as well as non-metallic) materials.

Several definitions of the term *microstructure* have been proposed in the literature (see, for example HORNBOGEN and PETZOW [1970,1991], SCHATT [1991], HOUGARDY [1981], HORNBOGEN [1981, 1984, 1986a,b], LÜCKE [1984], METALS HANDBOOK [1985], HEROLD–SCHMIDT [1988] or JEGLITSCH [1989]. For the purpose of this book, the following seems appropriate: The microstructure of crystalline materials is defined by the type, the structure and the number of phases, by the number, the geometric appearance (size, shape etc.) and the topological arrangement of the individual phase regions and their interfaces, and by the type, structure and geometry of lattice defects (which are in most cases not part of the thermodynamic equilibrium structure). The experimental study of metallic microstructures, and their qualitative and quantitative description is termed *metallography*. (Sometimes, this term has also been used for the preparation of ceramic materials and polymers for microscopic inspection. Eventually, this improper use will be substituted by *materialography* or a similar term to be newly introduced to include metallography, ceramography and plastography.)

Metallography dates back to the 17th century when English, French and German scientists first studied metallic objects by means of simple optical devices (see, for example, SMITH [1960], TENSI [1968] or PUSCH [1979]). The birth of modern metallography took place 200 years later and is dated to 1863 when H. C. Sorby developed an incident-light microscope, or to 1865 when he first observed and described some microstructural elements of technical iron. Today, a large arsenal of devices and techniques for microstructural investigations has become available (see, for example, METALS HANDBOOK, Vol. 9 [1985] and Vol. 10 [1986], LIFSHIN [1992, 1994a], CAHN and LIFSHIN [1993], and the books and journals listed under *Further Reading* at the end of this chapter).

This chapter deals with the techniques of microscopy and metallography as means for microstructural investigation. It focuses on qualitative and quantitative methods of optical, scanning-electron and scanning tunneling surface microscopies with short reviews of the other imaging and compositional analyzing techniques. A special chapter (chapter 11) is devoted to transmission electron microscopy, including analytical TEM.

2. *Optical microscopy*

Metallic materials are usually opaque; therefore investigations of plane cross-sections by incident light prevail in metallography. However, the transparency of some metals and silicon to infrared light in thin sections has been effectively exploited. Optically, the individual components of a metallic alloy differ in their amplitude and phase characteristics. While amplitude objects become visible owing to differences in light absorption and thus appear in different grey shades or even colours, phase objects only differ in the refractive indices which cannot be recognized without additional provision. The preparation of cross-sections, the enhancement of contrast by etching and other methods, as well as the microscopic set-up must be carefully optimized for the material under investigation and adjusted to the purpose of the investigation in order to get maximum information from a microscopic study.

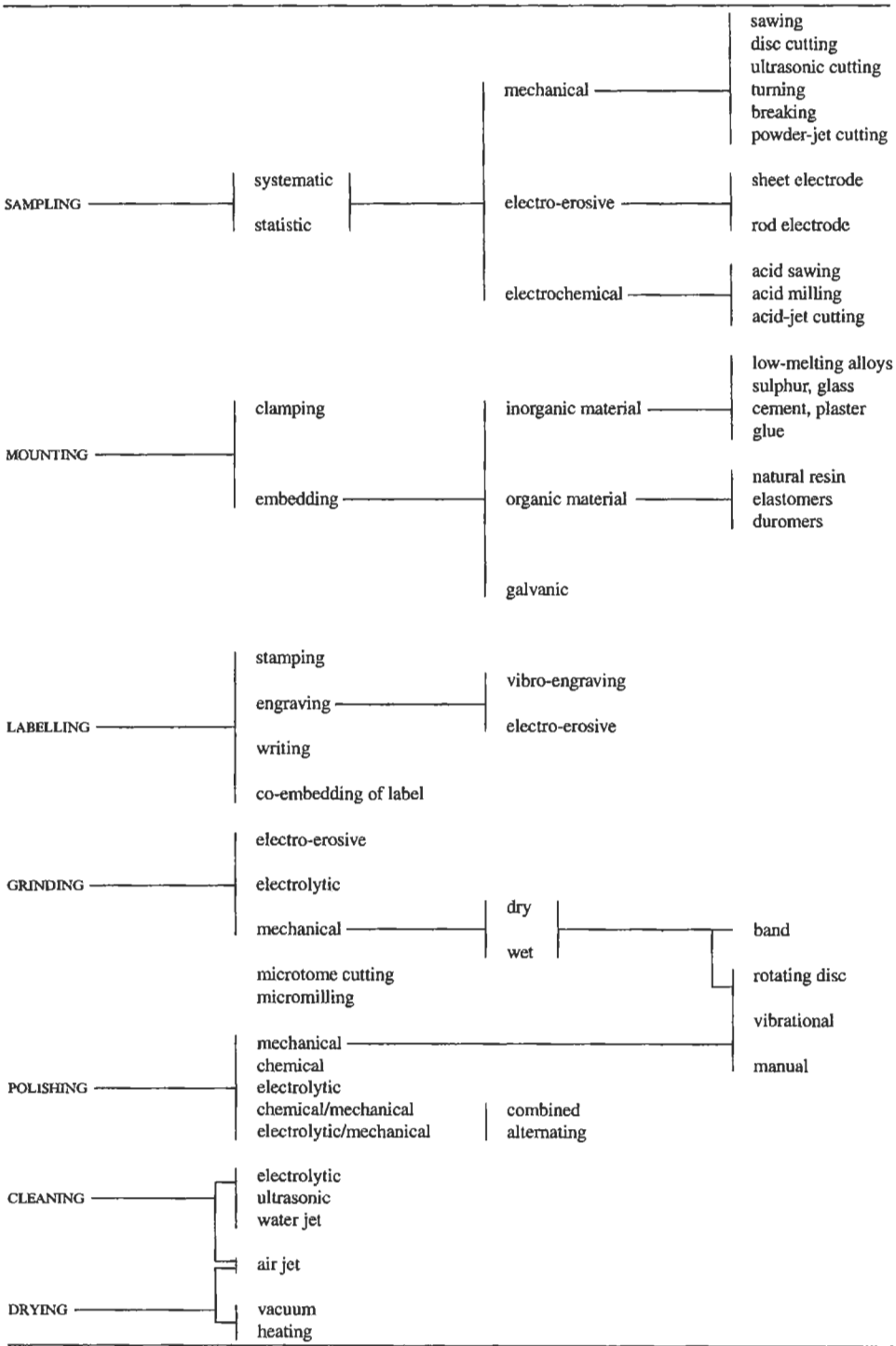
2.1. Metallographic specimen preparation

The essential steps and techniques of metallographic sample preparation are shown in table 1. This large variety of methods has been described in handbooks, monographs and review articles, e.g., by PETZOW and EXNER [1968], SAMUELS [1971], in METALS HANDBOOK [1985, 1986], by ELSSNER and KOPP [1984], VANDER VOORT [1984a], SCHUMANN [1990], LOUTHAN [1992], TELLE and PETZOW [1992], ASH HANDBOOK [1992], or PETZOW [1994]. Details are discussed in a multitude of original papers in a variety of journals and conference proceedings (see also *Further Reading* at the end of this chapter). Though some systematic studies of the construction of metallographic devices (e.g., WASCHULL [1985], KOPP and MÜLLER [1987] or FUNDAL and GROSS [1993]) and of the consecutive steps of sample preparation (e.g., NELSON [1989], MÜLLER and KOPP [1989], TELLE and PETZOW [1992] or WASCHULL [1993]) have been published, successful preparation of metallographic laboratories samples is still a matter of skill. Since accreditation of metallographic became a major issue, systematic evaluations of procedures and standard documentation are necessary additions to the empirically developed recipes (see, for example, RÜCKER and BJERREGARD [1993] or WIELAND [1993]). In the following, a few of the more basic aspects of the present state of the art are reviewed.

2.1.1. Sampling

The location from which a specimen is taken depends on whether the investigation is aimed (1) to give data for a specific area (systematic sampling), e.g., if the origin of a failure is clearly visible, (2) to characterize a larger piece (e.g., a laboratory sample) or (3) to characterize the quality of a large amount of material (as in quality control). In the

Table 1
Steps of metallographic preparation (after PETZOW and EXNER [1968]).



last two cases the statistical fluctuations due to unavoidable inhomogeneities must be considered, and usually more than one specimen is necessary to get a reliable result (statistical sampling). Since usually nothing is known as to the degree of homogeneity, statistical parameters (usually taking the arithmetic mean and the relative standard error, see, e.g., PETZOW and EXNER [1968]) should be determined from samples which are taken either at arbitrary or at specially defined locations. Furthermore, damaging the specimen during cutting it from a larger piece gives rise to erroneous results: Electro-erosive cutting ("spark-machining"), for example, changes the composition near the cut faces to an appreciable depth: e.g., 0.9 and 0.3 wt% carbon (stemming from the electrolyte) and 0.8 and 0.2 wt% copper (from the electrode) were found in pure iron at 50 and 150 μm depth, respectively, below the electro-eroded surface. Careful work (slow and interrupted cutting) reduces the depth of influence to 10 μm. Mechanical cutting (usually by water-cooled wheels) does not change the composition but introduces stresses to a depth of 100 μm and more (WAVER [1973], WELLNER [1980], KIESSLER *et al.* [1982], VANDER VOORT [1984a] or TELLE and PETZOW [1992]). In spite of the disadvantages of these commonly used techniques, others, like chemical cutting by a fast-moving endless wire wetted by an aggressive liquid, available commercially as "acid saws", are only used for special purposes (for single crystals, semiconductors, brittle intermetallics, etc.) because of the long cutting times needed (hours, instead of the minutes needed for mechanical cutting). The same considerations apply to ultrasonic erosion, electrochemical sectioning or laser cutting. In order to avoid artifacts, a careful choice of the sampling technique adjusted to the specific material and its conditions and a careful control of the result must be made.

2.1.2. Mounting

Embedding or clamping are relatively uncritical operations. Some resins reach a temperature up to 150°C during curing, which may lead to annealing effects in the specimen; others are cold-setting. Galvanic deposition of a thin copper or nickel layer reduces edge-rounding during preparation to an acceptable level even for oblique sectioning (see below). Smearing and edge-rounding of open porosity during polishing can be avoided by infiltration of a low-viscosity resin under vacuum or by a well-wetting melt (solder for metals, glass for ceramics).

2.1.3. Grinding

The surface of a cut cross-section usually shows a high degree of irregularity which is removed in successive steps of grinding with emery paper (paper covered with SiC particles closely graded from coarse to fine between 80 and 20 μm, see fig. 1). Heating can be limited to a tolerable degree using water-cooling, but deformation of the surface is unavoidable (SAMUELS [1971], PETZOW and EXNER [1968], WAVER [1973], KIESSLER *et al.* [1978], VANDER VOORT [1984a], TELLE and PETZOW [1992], and PETZOW [1994]). It was found empirically (LIHL and MEYER [1960]) that the deformation depth X_D is a square function of scratch depth X_S ($X_D = aX_S - bX_S^2$, where a and b are material constants). Figure 1 shows the depth of scratches, the deformation depth and the total depth influenced in grinding of steel. In an oblique taper section, the deformed layer becomes visible after etching (fig. 2).

References: p. 1016.

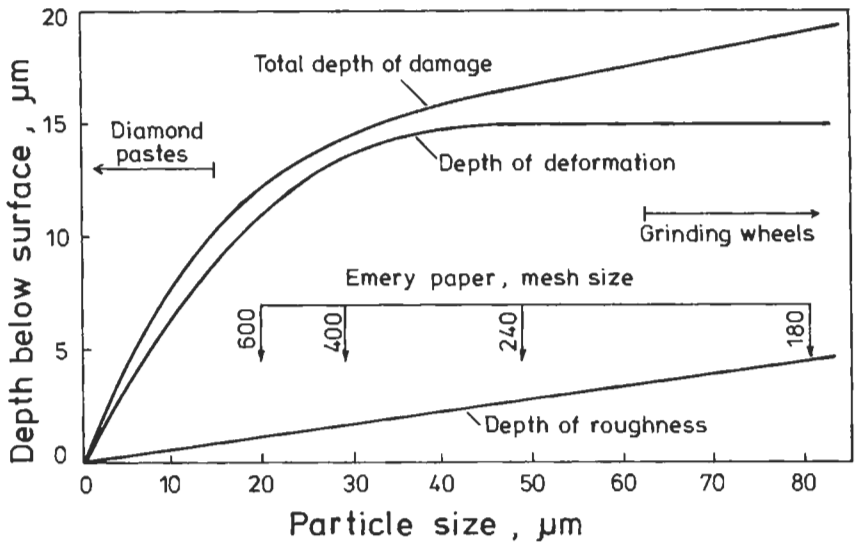


Fig. 1. Depth of roughness, depth of deformation and total depth of damage as a function of particle size of the grinding and polishing medium (after PETZOW and EXNER [1968]).

2.1.4. Polishing

In order to prepare a surface free of artefacts, the damaged layer is removed by polishing. While it is relatively easy to obtain scratch-free surfaces, preparation of an ideally undisturbed surface is difficult if not impossible by mechanical polishing. As shown repeatedly (e.g., by TURLEY and SAMUELS [1981] for copper) mechanically polished surfaces show all the signs of plastic deformation (shear-bands, recrystallization, subgrains, etc.) even after prolonged polishing with fine grades ($< 1\ \mu m$). The depth of the remaining disturbance is small ($< 1\ \mu m$) and is tolerable for most purposes. *Chemical* and *electrolytic polishing* techniques do not cause deformation. Another technique leaving little surface damage is ion-milling or ion-polishing (LEHTINEN and MELANDER [1980]), i.e., sputtering of the surface resulting from the impact of energetic argon ions (used more frequently for preparation of TEM foils). Removal of the layer damaged during grinding by polishing is time-consuming. If sample preparation is not carefully controlled, deformation from the coarse grinding steps may resist and may influence not only the microscopic appearance (BÜHLER and HOUGARDY [1979], SAMUELS [1971] and POKORNY [1980]) but also X-ray measurements (line broadening, blurring of reflections) or mechanical tests like indentation hardness and toughness (BERNST [1965] and EXNER [1969a]). Figure 3 shows the influence of various polishing methods on reflectivity, which is sensitive to surface damage. Investigations under polarized light suffer from surface strain as well. *Thermal polishing*, i.e., annealing in vacuo after mechanical polishing, was shown to produce smooth strain-free surfaces by uniform evaporation (IRANI and CAHN [1971]).

Disc-polishing using emulsions of carefully graded loose diamond or alumina

powders on cloth is the usual way of mechanical polishing. For hard materials, a drill-like set-up using a wooden stick with diamond paste has advantages. Material removal is adequate and most materials can be polished this way. Microcutting by diamond blades has been used as a one-step, time- and labour-saving preparation mode for soft materials. Such *ultramicrotoming* (for references see PETZOW and EXNER [1968] and KLOCKENKÄMPER *et al.* [1979]) has been shown to produce very large localized deformation (BÜCKLE [1964]); nevertheless, if the cutting parameters are optimized, a surprisingly perfect surface quality of microcut metals can be achieved (PETZOW and KNOSP [1973] and PETZOW and EXNER [1975]). By micromilling cross-sections and serial sections of medium-hard metals and alloys can be prepared (KIESSLER and ELSSNER [1980] and PETZOW and EXNER [1975]).

There is a nearly infinite number of recipes for chemical and electrochemical polishing. Reviews are to be found in early metallographic standard texts, e.g., TEGART [1957], PETZOW and EXNER [1968], BIGGS [1970] and SHIGOLEV [1974] while more modern references are scarce (METALS HANDBOOK [1985], PETZOW [1994]). Theoretical approaches are available (e.g., WAGNER [1954] and TOUSEK [1981], for reviews and early references see PETZOW and EXNER [1968] and BIGGS [1970]); however, they do not allow deduction of optimized polishing procedures for a given material. For electrolytic polishing, the form of the current density–voltage curve suggests which potentials should

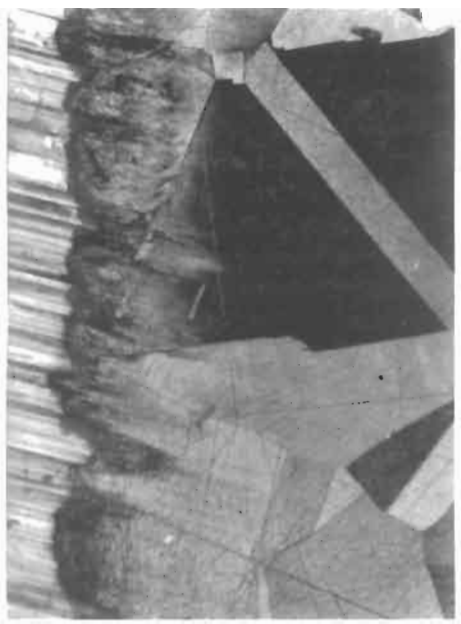


Fig. 2. Deformed layer in brass (CuZn 30) after cutting by a new diamond wheel. Oblique section (5° to surface which is on the left side). Etched. 200×. (From KIESSLER *et al.* [1982].)

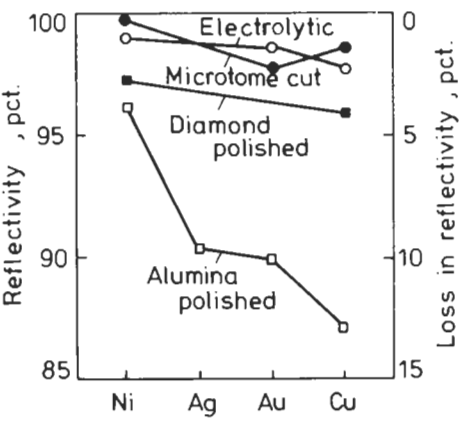


Fig. 3. Reflectivity (in percent of cleavage surface) of metal surfaces polished by different techniques (after PETZOW and KNOSP [1973]).

be used to avoid etching effects. (Usually a plateau is observed where polishing occurs.) It has been shown that these curves should be measured under practical conditions and not in special cells (RÖSCHENBLECK and WOLTER [1979]). In spite of the limited exact theoretical knowledge on the mechanisms of mechanical, chemical and electrolytic processes, preparation of metallographic cross-sections is not a limiting factor for microstructural investigations since the state of the art allows the investigator to deal with even the most difficult materials.

2.1.5. Replica techniques

Surface studies by transmission electron microscopy are possible by pressing a foil of a suitable substance tightly to the surface or to form a replicating foil by casting an organic replica. This technique was used extensively prior to the development of scanning electron microscopy for high-resolution surface studies. Sample preparation is well developed (see, for example, GOODHEW [1973]). By the so-called double-replica technique, fine precipitates can be extracted from a matrix by chemical or electrochemical etching and embedded in a resin foil. This technique is a quick and reliable standard procedure for investigating the morphology and the spatial distribution of dispersed particles and, in particular, is useful for EDX analysis of small features avoiding signals from the matrix (see, for example, SCHRÖDER *et al.* [1990] or CZYRSKA–FILEMONOWICZ *et al.* [1992]). Replicas can also be studied by optical microscopy, scanning electron microscopy (§ 3) and soft X-ray microscopy (§ 5.5). For remote sampling (e.g. for large engineering components which must not be destroyed, for taking samples at temperatures up to 120°C or for radioactive materials), replica techniques using movable preparation equipment are frequently the best if not the only way for microscopic inspection (see, for example, PELLOUX [1970], WENDLER and NEUBAUER [1979], BIRNER and LÖHBERG [1980], or LÖBERT [1982]). Replication techniques have also been applied in optical and SEM fractography, e.g., for the study of stress-corrosion cracking (CONOR [1972]) and fatigue crack initiation (BROWN and SMITH [1982], DIESER [1984], or BRÜGEL *et al.* [1988]). Serial sectioning of plastic replicas is much easier than that of the sample itself, and was proposed for three-dimensional reconstruction of fracture surfaces (BAUER and HALLER [1981]).

2.2. Etching and other contrasting techniques

The human observer is capable of distinguishing between different phases and lattice defects if these show a grey or colour contrast in microscopic viewing. Polished surfaces rarely provide sufficient contrast owing to the fairly similar reflectivity of metallic phases, and contrast enhancement is usually necessary. A number of metallographic techniques are available to reveal the microstructure.

2.2.1. Chemical and electrolytic etching

When a polished surface is attacked by an etching medium, different phases and different lattice orientations usually show differing rates of dissolution. Crystal imperfections and grain boundaries are locations of increased dissolution potential. These differences in chemical and electrolytic attack are the basis for the most frequently used

metallographic techniques for optical contrast enhancement. In the monographs by TEGART [1957], BECKERT and KLEMM [1984] and PETZOW [1978, 1994] or in handbooks as, for example, METALS HANDBOOK [1985], known recipes for technical metals and alloys are reviewed. Theoretical understanding of material removal (see, for example, ENGEL [1958], SCHAARWÄCHTER [1968], BIGGS [1970], HERBSLEB and SCHWAAB [1978]), though well-established in corrosion science, is rarely used for finding the optimum etching conditions for a new material as, for example, in potentiostatic etching (WORCH *et al.* [1994]). Usually these are established empirically, aided by educated guesses. Local changes in reflectivity and shadows produced by the rough topography of the specimen surface give rise to grey contrast when viewed in the microscope. Thus, grain boundary grooves and facets, height differences between grains of different orientation and between phases, or etch pits at points where dislocations penetrate the cross-section, are typical contrast features created by attack-etching. Sometimes, deep-etching by chemical or electrochemical attack may favourably be used to show the spatial geometry of microstructural features and to make them accessible to stereometric measurement as shown, for example, by PAUL and MÜRRLE [1981], FEIJOO and EXNER [1990], or FEIJOO *et al.* [1990] (see also § 2.4.1).

2.2.2. Thermal etching

The thermodynamic instability of a polished surface will lead to effects similar to those mentioned in § 2.2.1 when material transport is activated by heating. The basic mechanism is surface diffusion (rather than selective evaporation), and the kinetics of thermal grain-boundary grooving and facetting are well understood (see, for example, MULLINS [1961]). Thermal etching is advantageously used for chemically stable materials such as ceramics (compare fig. 7c below).

2.2.3. Ion-etching

The basis of the well known but infrequently used technique of ion-etching, reviewed by WECHSUNG [1977], POHL and BURCHARD [1980a,b], GRÄF *et al.* [1993], POHL [1994], PETZOW [1994], PECHENYAKOV and KOVACHEVA [1995], is cathodic atomization (sputtering) by bombardment of the surface with chemically neutral (e.g., argon) or reactive (e.g., oxygen) ions. The physics of sputtering has been surveyed by ÖCHSNER [1975] and by PIVIN [1983]. The rate of material removal depends on the atomic weights of the material and the ions (the highest rates being observed when these are approximately equal), on the energy and density of the ions hitting the surface, and on the atomic bonding in the material. By adjusting the sputtering parameters (voltage, gas pressure), selective material removal can be made to produce a clear topography and clean surfaces. Ion-etching is advantageously used for composite materials as shown in fig. 4 and coated metals (STAPF *et al.* [1986], GRÄF *et al.* [1993], POHL [1994]). However, artifacts are easily produced (fig. 5). Ion etching is also useful, if the microstructure is prepared for the application of surface-sensitive analytical techniques as well as interference-layer contrasting (see § 2.2.5).

2.2.4. Staining (tinting) and anodic oxidation

A large number of so-called etching techniques do not produce a surface relief by dissolving the surface but produce a surface layer by a chemical reaction (JEGLITSCH [1968]). These layers vary in thickness as a function of composition and orientation of the microstructural components. They are transparent and rather than having a specific colour themselves, produce interference colours varying with thickness, *d*. Light waves reflected at the surface and at the layer-substrate interface interact, causing extinction of a specific wavelength λ_e according to the equation (for normal incidence)

$$\lambda_e = \frac{2n}{m} d, \tag{1}$$

where *m* determines the order of interference ($m = 1, 3, 5$ ... corresponds to 0, 1, 2 ... order) and *n* is the refractive index of the deposited layer; for a more detailed discussion see § 2.2.5.

Reaction layers of which the thickness varies with composition of the substrate can be deposited by chemical attack, by electrolytic processes, e.g. by potentiostatic oxidation (anodizing), or by oxidation when heating a metallic specimen in air (thermal tinting). Tinting techniques have been extensively discussed in the literature (see, for example, JEGLITSCH [1968], GRÜTZNER and SCHÜLLER [1969], BERAHA [1970], YANKOVITH [1970], BERAHA and SPIGHLER [1977], HERBSLEB and SCHWAAB [1978], GAHM and JEGLITSCH [1981], GAHM *et al.* [1982], WECK and LEISTNER [1982–1986], VANDER VOORT [1984a,b, 1985a], ZHOU *et al.* [1993], PETZOW [1994], and many others). Anodic oxidation is another possibility to produce layers with a thickness varying with orientation and composition of the microstructural features resulting in an orientation- or phase-specific colour contrast. Anodic oxidation is carried out by immersing the sample in an acid solution of carefully adjusted pH and applying a voltage in the order of 100 V. Intermetallic phases have been identified in this way (SEEGER *et al.* [1990]) and the grain structure of aluminium can be revealed (YANG [1990]), among numerous other applicat-

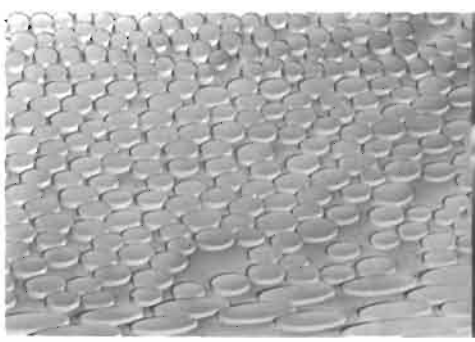


Fig. 4. Ion-etched cross section of a graphite fibre-reinforced polymer. Scanning electron micrograph, 300 × (courtesy I. Gräf).

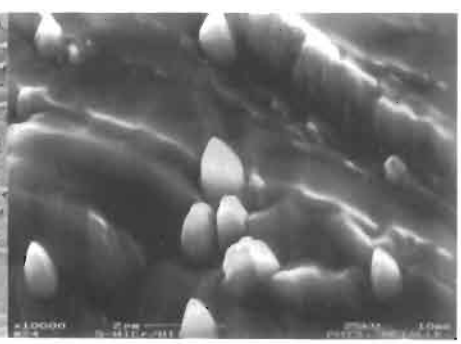


Fig. 5. Artefacts produced by ion-etching of a Ni-base superalloy. (Cones formed owing to the presence of non-conducting inclusions). Scanning electron micrograph, 10,000 × (courtesy I. Gräf).

ions (DANIELSON [1985]). In its early days, the colour contrast obtained after depositing interference layers in the presence of oxygen (reactive sputtering) was also attributed to varying thickness (ONDRACEK and SPIELER [1973]); in fact it depends on a different optical principle, as discussed in detail in the next section.

2.2.5. Interference-layer contrast

A plane-parallel layer of a non-absorbing or weakly absorbing substance acts like an optical reflection–interference filter which, by multiple reflection at the metal-layer and layer-air interfaces, causes contrast enhancement between neighbouring phase regions, provided these differ in their optical constants. The optical principles of these effects have been reviewed by PEPPERHOFF and ETTWIG [1970], ZOGG *et al.* [1977], BÜHLER and HOUGARDY [1979, 1980], BÜHLER [1981] or PETZOW [1994], and a large number of applications in metallographic practice have been published (for references see BÜHLER and HOUGARDY [1979, 1980], EXNER and ROTH [1980], GAHM and JEGLITSCH [1981], GRÄF [1981] and WU *et al.* [1982]). The important factor in interference-layer colour contrasting is the phase shift of the light wave reflected at the layer–metal interface. In normal bright-field microscopy, differences of this phase angle are much too small for metallic phases to be detectable. The transparent or semitransparent layers enhance these differences dramatically, revealing a pronounced colour contrast if two conditions are fulfilled:

(i) The *phase condition* relates thickness of the layer d, the optical properties of the metallic phase (phase shift of the reflected wave through the angle δ) and of the layer (refractive index n), and the order of interference (defined by m as above) to the wavelength λ_{min} for which interference causes maximum reduction in intensity:

$$\lambda_{min} = 4\pi nd/[\pi(m-1)+\delta]. \tag{2}$$

Usually, the first-order interference ($m=1$) yields the best results.

(ii) The *amplitude condition* describes the relative intensity R_{min} of the reflected light with wavelength λ_{min}. For non-absorbing layers,

$$R_{min} = \frac{\text{intensity of reflected wave}}{\text{intensity of incident wave}} = \left(\frac{q-q_m}{1-q\cdot q_m}\right)^2, \tag{3}$$

with

$$q = (n-1)/(n+1), \qquad q_m = \sqrt{(n_m-n)^2 + k_m^2/(n_m+n)^2 + k_m^2}.$$

n_m and k_m are the refractive index and the absorption coefficient of the metallic phase, respectively. Much effort has been undertaken (BÜHLER and HOUGARDY [1979], BÜHLER and KOSSEL [1981], AYDIN and BÜHLER [1981, 1984], AYDIN *et al.* [1983] and BÜHLER [1986]), to determine n_m and k_m values for phases occurring in metallic materials in order to be able to calculate the required n, λ_{min} and d values for maximum contrast between phases present in a material. The contrast is defined by $K=(R_1-R_2)/R_1$, where R_1 and R_2 ($R_1>R_2$) are the relative reflectivities of two phases, and reaches a maximum if $R_2=0$, i.e, when interference causes complete extinction of the colour under consideration in one

of the two phases ($K=1$). If, then, a filter for the corresponding wavelength λ is used, this phase appears black. If white light is used, maximum colour contrast will be obtained.

Plane-parallel interference layers can be deposited by evaporation or by sputtering. In order to fulfill the amplitude condition $R_{min}=0$, a non-absorbing layer must have a high refractive index if the substrate has a high reflectivity (as do all metallic phases). ZnS, ZnSe, TiO_2 and ZnTe ($n=2.4$–3.5) are used for evaporation. *Absorbing* layers can be deposited by reactive sputtering. This technique, first described by BARTZ [1973], uses a low gas pressure (~10^{-6} bar, usually oxygen) in the sputtering chamber. A commercially available device has proved very useful in practical application (BÜHLER and HOUGARDY [1979], EXNER and ROTH [1980]) and, at least in principle, allows layers to be produced with widely varying optical properties. A disadvantage is the fact that calculating the optimum contrast conditions becomes somewhat more difficult than for non-absorbing layers (ZOGG *et al.* [1977]). A general difficulty is that the quality of surface preparation before applying the interference layer is crucial for the result. For example, scratches or contamination by polishing liquids will produce artefacts. Ion polishing, therefore, is preferable compared to electropolishing or mechanical polishing techniques as shown, for example, for steels by GAUDIG and SCHECK [1984]. Also, the accuracy of measurement of the optical constants depends on the reproducibility of surface preparation (KRONER [1986]). Since calculation of optimum contrast conditions can be applied only if the optical constants are known to an accuracy of 2% (SCHRÖDER and HOUGARDY [1985]), the trial-and-error approach can often be quicker and more practicable. The reason that interference contrasting is described here more fully than other methods is that this technique is still not widely familiar. It has excellent reproducibility and versatility and could substitute many of the classical contrasting techniques. The extreme sensitivity of the human eye for colour hues and the possibility of using filters makes colour-contrasting highly attractive. In addition, quantitative evaluation with respect to phase composition and exact phase identification (see, for example, ZOGG *et al.* [1977]) are possible. Care must be taken in the latter cases in photographic reproduction which may change the original colours appreciably (CROUSE *et al.* [1977] and EXNER *et al.* [1980]).

2.3. Principles of light microscopy and optical contrast enhancement

The highly developed state of the mechanical and optical design of microscopes used in metallographic work makes it impossible to come near to an adequate description in the context of a book on physical metallurgy. A large number of monographs (e.g., PAYNE [1957], OETTEL [1959], MALIES [1959], FREUND [1960, 1969], KINGLAKE [1965], BIGGS [1970], LOVELAND [1970], PHILIPS [1971], GALOPIN and HENRY [1972], MODIN and MODIN [1973], BEYER [1974], ROST [1981], BRADBURY [1991], SCHADE [1993]) as well as articles in handbooks (e.g., VANDER VOORT [1985b], LOUTHAN [1987], SCHUMANN [1990], TELLE and PETZOW [1992] or HOLIK [1993]) are available which treat the basic as well as the practical aspects of optical microscopy and photography comprehensively. Accordingly, the optical fundamentals (e.g., wave optics, properties of lenses and correction for aberrations) or the various components of the optical microscope (illumi-

nation systems and light sources, objectives and eyepieces, polarizers, interferometric attachments, phase contrast equipment, stages etc.) need not be discussed in detail here. Important features are the resolution limit, depth of focus, and the different ways to enhance contrast by optical manipulations.

2.3.1. Resolution and depth of focus

Figure 6 shows the *limiting resolution* (minimum distance between two points and maximum number of lines per unit length, seen as separate features) and the *depth of focus* as a function of the objective's numerical aperture $n \sin\alpha$ (where n is the refractive index of the medium between the objective's front lens, e.g., 1 for air, 1.25 for cedar oil, and α is half the opening angle of the objective; thus, $n \sin\alpha$ is a quantitative measure for the amount of light reaching the objective). The resolution limit is, in theory, $d = 0.6\,\lambda / n\sin\alpha$, where λ is the wavelength used (Abbé limit). For green light (λ = 500 nm) using an immersion oil between specimen and high-power objective (resulting in a numerical aperture $n \sin\alpha = 1.25$), we get $d = 0.24\ \mu$m. In practice, however, the resolution limit is close to 1 μm even if the illumination system as well as all the lenses and apertures are optimally adjusted. Magnifications between 500 and 1000 times the numerical aperture, i.e., up to 1000×, are useful while higher magnifications yield no additional information and therefore are called empty magnification. The depth of focus is inversely proportional to the square of the numerical aperture and is extremely limited at high magnifications (approx. 0.1 μm).

2.3.2. Bright-field illumination

By far the majority of microstructural investigations by optical microscopy are carried out with vertical illumination, usually called *bright field*, and most optical micrographs of metallic materials are taken this way. Regularly reflected light is used and no additional manipulations of the light beam are necessary. Thus, high light intensities are obtained. The instrumentation is rather simple, and the use does not require great skill. The contrast is a result of locally varying light intensity due to one of the pretreatments of the specimen section discussed above. The human eye can differentiate between two grey levels if the contrast K is approximately 0.2, i.e., five grey levels between black and white are easily distinguishable. Photomultipliers and television cameras are more sensitive, and more than a hundred grey levels can be registered. If higher contrast is needed, one of the special techniques described in the next four sections (sometimes improperly called "optical etching") can be useful.

2.3.3. Oblique illumination, dark field and stop contrast

If the direction of the incident light is changed from vertical to oblique, the contrast can be reversed with a gain in contrast which, for suitable specimen surfaces, is often striking. This can be achieved by simply moving the condenser aperture slightly off the optical axis, which produces a shadow-like contrast. While such *oblique illumination* is only applicable for low magnifications (long-working-distance objectives), a very useful alternative is *dark-field illumination*: The light from the light source does not pass through the objective but is reflected to the surface by a ring-shaped mirror or lens

References: p. 1016.

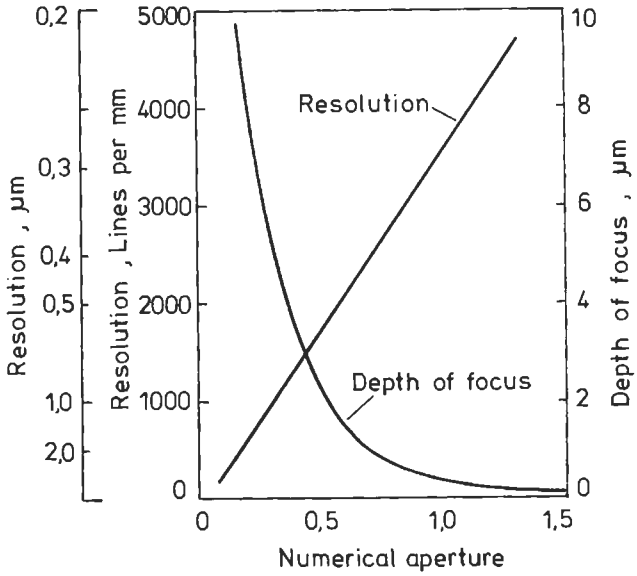


Fig. 6. Theoretical resolution (resolved lines per mm and distance between two distinguishable points) and depth of focus as a function of the objective's numerical aperture (theoretical for green light).

around the objective so that only stray (diffusely reflected) light reaches the objective. Rough surfaces, fissures, pores, grain boundaries and other surface irregularities are revealed, appearing bright on a dark background. This technique also lends itself for checking the quality of polish since scratches clearly show as bright lines. *Opaque-stop microscopy* produces images similar to dark field. Instead of changing the illumination, a ring stop is placed between the light source and the condenser lens. By moving the stop, different areas may be illuminated. Tilt angles with respect to flat portions of the specimen surface have been measured, for example during studying tilt and twist boundaries (BIGGS [1970]). This is possible to a high degree of accuracy (depending on magnification, of the order of 1′ to 60′ of arc). For qualitative inspection, opaque-stop microscopy provides a sensitive type of dark-field contrast at no loss in resolution.

2.3.4. Polarized-light microscopy

Plane-polarized light (produced by placing a polarizer in front of the condenser lens) vibrates in one plane only. When reflected from an optically isotropic surface, the direction of polarization does not change and will be transmitted by an analyzer placed behind the eyepiece and set parallel to the polarizer. If the analyzer is rotated, the transmitted light intensity is reduced and ideally drops to zero at crossed position of polarizer and analyzer. If the plane-polarized beam is reflected by an optically anisotropic surface it is sub-divided into two components vibrating at right angles to each other. The intensities of the components vary as a function of crystallographic orientation and its relation to the plane of polarization. Therefore, the amount of light transmitted by the analyzer is a

function of the orientation of a crystal which causes changes in brightness (*degree of extinction*) when the specimen stage is rotated or, for a polycrystalline material, for the various crystals when viewed with crossed polarizers. Furthermore, some optically anisotropic substances (e.g., nonmetallic inclusions such as cuprous oxide) show distinctly different tints in white polarized light due to an optical effect called *reflection pleochroism*, owing to a variation of reflectivity with wavelength or degree of extinction. The tint and its change when rotating the analyzer or the stage are characteristic for such materials.

Polarized light is particularly useful in metallography for differentiating between optically isotropic and anisotropic components of the structure and for revealing the grain structure and twins in anisotropic metals and alloys, such as Zn, Mg, Ti and V. Though it is difficult to achieve a polished surface which is wholly strain-free, some metals which are hard or impossible to etch can be effectively examined (IRANI and CAHN [1971]). Even the cubic crystals can become optically active if etch pits or grooves are produced by etching (for references see PHILLIPS [1971]) or by coating with an anisotropic film e.g., by anodizing or by other epitaxially grown films. Comprehensive reviews on the use of polarized reflected light as an aid in metallography (and mineralogy) are available (CONN and BRADSHAW [1952], HARTSHORNE and STUART [1952], MOTT [1952], PHILLIPS [1971], GALOPIN and HENRY [1972], MODIN and MODIN [1973], MCCRONE *et al.* [1978], SCHUMANN [1990], TELLE and PETZOW [1992], among others). The quantitative use of polarized light has been restricted mainly to transmitted light in the fields of petrography and biology, though various measurements are possible in reflection on metals as well. A special application is the imaging of magnetic domains in metal crystals making use of the Kerr effect, reviewed by BOWMAN and BOOTH [1971]. The contrast which results from a rotation of polarization direction by only 1′ to 20′ of arc can be improved by interference layers (see § 2.2.5).

2.3.5. Phase contrast and interference contrast

Before some of the modern sample preparation techniques (e.g., interference-layer deposition or reproducible electrolytic etching) were fully developed, optical techniques transforming phase-angle or height differences into grey or colour contrast gained some interest in metallography (JEGLITSCH and MITSCHE [1967], BEYER [1974], MODIN and MODIN [1973], VANDER VOORT [1985b], SCHUMANN [1990], TELLE and PETZOW [1992] or HOLIK [1993]). Though they are developed more than 30 years ago, they are now available with some higher-priced metal microscopes, and have been described in detail in numerous publications (e.g., OETTEL [1959], JEGLITSCH and MITSCHE [1967] and BEYER [1974]) their applications in studying metallic microstructures have remained few.

Phase contrast (used extensively for transmitted-light studies in biology, see, for example, YAMAMOTO and TAIRA [1983]) transforms the invisibly small phase-angle shift caused by a small difference in height (or in optical properties) of an object and its surrounding into an amplitude (light-intensity) difference visible to the human eye. This is effected by retarding (positive phase contrast) or accelerating (negative phase contrast) a portion of the directly reflected wave by half a wavelength by inserting a phase platelet. Interference of this modified reflected wave from the phase object with the

References: p. 1016.

unmodified wave yields a grey contrast, i.e. the image of the phase object, in the field of view. Extremely small height differences (1–5 nm) become visible whereas in bright field a step height of approximately 100 nm is necessary for detection.

Interference contrast uses polarized light and a prism arrangement which splits the polarized light into two beams of equal intensity, the directions of which diverge by a small angle. These two beams hit the surface at slightly different positions and, after reflection, interfere with each other after passing a crossed polarizer. Several technical arrangements are possible; those used in practice are the differential interference contrast due to Nomarski (generally known as *Nomarski contrast*) and Michelson interference (CONN and BRADSHAW [1952], JEGLITSCH and MITSCHE [1967], PHILLIPS [1971], ROSENBERGER [1977] or HOLIK [1993]). If height differences exist at a specimen surface, the different levels appear in different colour hues which can be changed by shifting the prism. JEGLITSCH and MITSCHE [1967] have demonstrated applications of interference contrast with steel and cast iron as examples; other applications covering a multitude of materials are interesting as well (BENESOVSKY and IHRENBERGER [1965], SPIESS [1965], BEYER [1974] and ROSENBERGER [1977] or SCHUMANN [1990]).

2.3.6. Filters

By the appropriate choice of illumination wavelengths the contrast between differently coloured constituents may be enhanced. Alternatively to coloured glass, a movable-wedge interference filter allows monochromatic light of any wavelength to be selected. The contrast obtained is especially suitable for black and white photography (where green light is usually used) and for automatic image-analyzing equipment with a black and white camera as the detector.

2.4. Special optical devices and accessories

2.4.1. Stereomicroscopy

Optical stereomicroscopes (see, for example, WATTS [1982]) are useful for rough surfaces up to 100× magnification. A recent version allows vertical as well as 45° side viewing without tilting the object. Today, the larger depth of focus and wide range of magnification available with scanning electron microscopy has nearly completely replaced optical stereomicroscopy for studies of rough surfaces. Optical stereomicroscopes have kept their place in the metallographic laboratory for quick inspection and quality control purposes. If a higher resolution or a larger depth of focus than available with optical stereomicroscopes is needed, scanning light macrography or confocal microscopy (see § 2.4.2) and particularly scanning electron microscopy (which is presently clearly the best choice, see § 3.4.1) can be used.

2.4.2. Laser scanning and confocal microscopy

The light scanning microscope dates back to an invention of YOUNG and ROBERTS [1951] called the *flying spot microscope*. A similar principle is used in scanning light macrography developed in 1968 and applied for imaging of fracture surfaces with a high depth of field (CLARKE [1987]). With the development of small gas lasers producing

monochromatic and virtually ideal parallel light rays, a light spot of high intensity can be focussed on tiny areas of a polished or rough surface. The first laser scanning microscopes (BRAKENHOFF *et al.* [1979], SHEPPARD [1987]) were soon followed by a commercial instrument built by Carl Zeiss, first in 1982. A major advantage is the application of the *confocal principle* first described and patented in 1961 (see MINSKY [1988]). By reducing the depth of focus to a very small dimension with an effective depth resolution of 50 to 20 μm, rough surfaces can be optically sectioned plane by plane. The consecutive images can then be put together to produce a micrograph with a high depth of resolution orders of magnitude larger than in the conventional microscope. The information can also be stored in a computer and, using available software, can be used for a three-dimensional reconstruction of rough surfaces. The optical principles and instrumental design have repeatedly been discussed (see, for example, WILSON and SHEPPARD [1984], BOYDE [1985, 1990], VANDER VOORT *et al.* [1989], or TURNER and SZAROWSKY [1993]). Applications of the laser scanning microscope in materials science and other fields have recently been reviewed by ELSSNER *et al.* [1991] and TURNER and SZAROWSKY [1993].

2.4.3. Scanning near-field optical microscopy

As discussed in § 3 and 4, the scanning principle is used in many ways to increase the resolution up to the atomic level. With the upcoming of nanoscale science and technology, nano-optics, i.e., microscopes which collect light from extremely localized sources by fibre optics and produce images with a resolution in the range of 20–100 nm with a theoretical limit of approx. 5 nm, were developed. Though the principle idea dates back to 1928, the systematic development started only in the 80's resulting in the so-called *scanning near-field optical microscope* (SNOM) which can be regarded as a member of the super-resolution microscopes (WICKMARASINGHE [1992a,b]) described in § 4. The resolution of this new type of microscope is not limited by the Abbé limit, i.e., it is not determined by the wave length of light but by the size of the probe aperture. A review on near-field microscopy, nano-optics and SNOM was recently published by POHL [1992].

2.4.4. High-temperature microscopy

Heating stages for high-temperature optical microscopy need a long-working-distance objective (e.g. a reflecting objective) that will offer the necessary oxidation protection. They are also on the way to obsolescence since stages for scanning electron microscopy are now available and used much more often than optical hot stages. Reviews of earlier work which often resulted in important new information on the transformation behaviour of metallic alloys, on melting and on other kinetic processes have been published by REINACHER [1965], LOZINSKIJ [1961], MITSCHE *et al.* [1964,1969], MODIN and MODIN [1973], KULMBURG *et al.* [1974], METALS HANDBOOK [1985] or SCHUMANN [1990]. Besides the technical problems which now seem under better control (MUGGLI and MC CRONE [1977], HOEKSTRA and BRUIS [1979]), the question remains unsettled whether processes observed on the polished surface take place in the bulk qualitatively or quantitatively in the same way.

References: p. 1016.

2.4.5. Television cameras

First attached to the optical microscope for demonstration purposes, television cameras, as was soon realized by an English company, could be adapted for quantitative analyses of microstructural images by adding a voltage threshold and a few electronic components. This started a rapid development of *quantitative television microscopes*. The modern versions are now the most powerful automatic devices for image analysis, details of which will be discussed below (§ 7.1).

2.4.6. Microphotometry and ellipsometry

Only twenty years ago, microscope *photometers* were introduced in metallography after having been used for some time in ore microscopy. Several highly sensitive commercial instruments are now available which use photomultipliers to register the intensity of reflected light (usually compared to a standard). Identification of small microstructural components (e.g. nonmetallic inclusions), detection of surface damage, orientation and concentration analysis, and sensitive detection of phase boundaries in quantitative image analysis were prominent early applications (MITSCHE and SCHEIDL [1964] and PETZOW and KNOSP [1973]). Measurement of optical constants, e.g., for contrast calculations in interference–layer contrasting, have become an important domain of photometry (see § 2.2.5). Differential reflectometry is an interesting but seldomly used surface analytical technique filling the gap between X-ray diffraction and the ion- and electron probes discussed in § 6. Numerous applications (study of ordering in solid solutions, of corrosion phenomena and of lattice defect) have been reported by HUMMEL [1983]. Another sensitive reflection technique is *ellipsometry* (see, for example, VISSCHER [1973] and AZZAM and BASHARA [1977]) which uses the phase shift of linearly polarized light for unambiguous determination of optical constants and for following the growth kinetics of thin surface films down to atomic dimensions.

2.4.7. Interferometry

One of the most useful techniques for measuring the height of surface steps and other topographic features with high resolution (routinely 10–100 nm in reflection) is interferometry (KELLER [1977]). For *double-beam interferometry*, the same set-up as for interference contrast (§ 2.3.5) is used. The prism arrangement is set in such a way that interference of the two beams after reflection at an oblique plane surface causes a linear sinusoidal fluctuation of brightness which to the human eye appears as a series of parallel black stripes. If the surface is not plane, the stripes are distorted or set off (at steps). The width of the stripes can be reduced by a more difficult *multiple-beam interference* arrangement and their distance from each other, corresponding to half the wavelength in height, can be suitably chosen by adjusting the prisms. (For interference contrast as described in § 2.3.5 infinite stripe distance, i.e., only one brightness minimum, is used for illumination.) A large variety of microinterferometers are available (see, for example PHILLIPS [1971], MODIN and MODIN [1973], BEYER [1974] and KELLER [1977]). By measuring stripe distances and set-offs, height differences are determined. Also, angles of surface tilt can be measured with high accuracy. Among the applications in metallography have been studies of crystal growth kinetics (e.g., NANEV [1981]), of

grain-boundary grooving (MYKURA [1955]) and of surface films and surface roughness (for references see BEYER [1974]). A recent application of interferometry is the characterization of targets for inertial-confinement fusion (FOREMAN *et al.* [1993]), and special devices have been developed for this purpose.

2.4.8. Microhardness

As a local mechanical test, indentation hardness measurements using low loads (5–500 mN) are carried out in microscopically small regions. Micrometer eyepieces or, more often, specially designed arrangements which keep the image and the scale sharp simultaneously are used. Detailed specifications for microhardness testing are found in literature (MOTT [1956], BÜCKLE [1965], SEMLITSCH and BERGMANN [1969], GAHM [1969], DENGEL [1973], MORNHEIM [1977], SCHUMANN [1990], TATE [1993], among others). Microhardness indentations are often used as markers for measuring distances and for accessing surface deformations induced subsequent to the indentations. More recently, ultra-hardness testing with loads in the range of 0.5 to 5 mN was developed. However, the resolution of optical microscopy is at its limits as the indentation size drops below 1 μm and the scanning electron microscope (see, for example, ZAHN [1992]) or depth measurements (as with instruments called mechanical microprobe or nanoindenter; see, for example, OLIVER [1993]) must be used.

3. *Scanning electron microscopy*

One of the most versatile instruments for microstructural investigations is the *scanning electron microscope* (SEM). Compared to the optical microscope it not only expands the resolution range by at least one order of magnitude (useful magnifications beyond $10^4\times$) and the depth of field by two orders of magnitude (ranging from 1 μm at $10^4\times$ to 2 mm at $10\times$) but also offers a series of possibilities for image formation which are more or less easy to interpret and reveal a clear picture of plane cross-sections as well as three-dimensional surfaces such as, for example, fracture surfaces, deep-etched or corroded surfaces and surfaces of porous materials.

3.1. Basic features of scanning electron microscopy

Excellent monographs and handbook articles on the physical fundamentals, the various techniques, the instrumental details and the application of scanning electron microscopy are available (e.g., SEILER [1968], HEYWOOD [1971], HEARLE *et al.* [1972], OATLEY [1972], REIMER and PFEFFERKORN [1977], LORETTO [1984], EXNER [1985], VERHOEVEN [1986], NEWBURY *et al.* [1986], GABRIEL [1987, 1992a], CHESCOE and GOODHEW [1990], LYMAN [1990], LYMAN *et al.* [1990], JOY [1992], WELLS [1993], LIFSHIN [1993b], GOLDSTEIN *et al.* [1994]), and a large number of various types of commercial instruments are in daily use in almost every materials research and development laboratory around the world.

The principle of the scanning electron microscope is rather simple: An electron beam scans the surface of the sample in the same way as in a cathode-ray tube which is used

to display the image. The electrons are emitted usually from a heated tungsten cathode. For higher electron currents (about 10 to 100 times that of a W electron gun) yielding a higher brightness, a lanthanum hexaboride (LaB_6) cathode may be used. The beam is focussed at the surface to a small diameter (approx. 10 to 100 nm). More recently, field-emission cathodes became commercially available which provide high-resolution images by reducing the probe diameter further and increasing the brightness by another 100 times needing a higher vacuum (10^{-7} Pa as compared to 10^{-5} Pa for the thermionic guns). The acceleration voltage ranges between 1 and 50 kV while the current through the surface is in the order of 10 pA.

The magnification is increased by reducing the current in the deflection coils; the normal range is 1–50 000×. The SEM image is written on a high-quality screen (cathode-ray tube) with an image resolution typically of 0.1 mm which, at 10 000 times magnification, corresponds to 10 nm resolution on the specimen. The brightness of each picture point is determined by the detector signal which is of the order of a few pA and must be amplified by direct electron multiplication or by conventional amplifiers. Noisy background due to high amplification is the major cause of unclear images. The dramatic improvement of image quality and resolution during the three decades of commercial development is due to improvement of all components of the microscope (more sensitive detectors, more powerful cathodes and more precise scanning devices) and to noise reduction in all electronic devices.

Figure 7 shows the interaction between the primary electron beam and the sample: Various kinds of radiation are emitted which, when collected in a suitable detector, can be used for image formation. Table 2 gives some data important for imaging metallic microstructures. By far the most popular techniques are the *secondary electron mode* (SE) revealing surface topography and, in advantageous situations, also atomic-number and crystal-orientation contrast with high resolution and information from a small depth below the surface, and the *backscattered electron mode* (BE) which gives topographic as well as pronounced materials contrast from a much larger depth and width and therefore reduced resolution. Figure 8a shows a typical SE image, fig. 8b a typical BE image of polished metal surfaces. Detector electronics can be adjusted so as to respond to either SE or BE electrons with their quite different energy distributions.

Back-scattered electrons (BE) are produced by single large-angle as well as by multiple small-angle elastic scattering events. In the 10–20 kV range of incident beam voltage, approximately 50% of each type leave the surface with a wide spectrum of energies in a range higher than 50 eV. As the atomic number of the material hit by the incident beam decreases, a smaller number of electrons are back-scattered and more energy is lost. In materials with high atomic number, a large number of electrons is back-scattered by atoms close to the surface, with little change in energy. Thus, yield, energy spectrum and depth of escape of back-scattered electrons are directly related to the atomic number of the material (fig. 9, see also § 3.3.2).

Low-energy *secondary electrons* (SE) are formed by interaction of the primary electrons with loosely bound atomic electrons. The energy distribution of secondary electrons depends on the primary energy of incident electrons, the number of outer-shell electrons, the atomic radius and, most pronouncedly, on the surface barrier of the

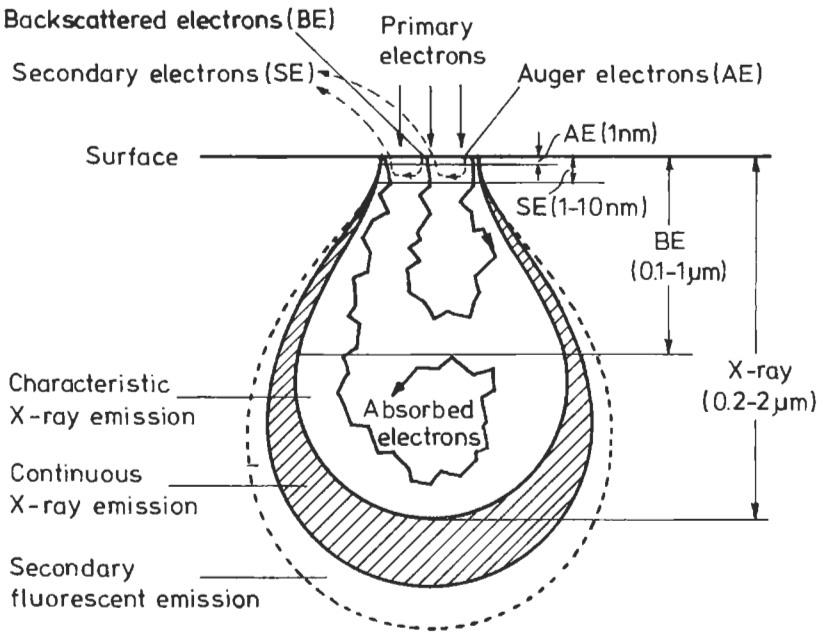


Fig. 7. Types of electron-beam-excited electrons and radiation (schematic) and depth of information in the scanning electron microscope. (The fluorescence emission can come from as deep as 100 μm.)

material. The energy of secondary electrons is in the range of 0 to 10 eV with a pronounced maximum at around 3 eV. At 50 eV the frequency of SEs approaches zero, i.e., there is no overlap with the energy spectrum of back scattered electrons. However, high-energy BEs excite SEs as well which contributes to noise and decreases lateral as well as depth resolution (see also § 3.3.1). The probability that low-energy secondary electrons will escape from the surface decreases exponentially with the depth of their generation. More than half of the total yield is emitted within a depth of about 0.5 nm. The yield is not strongly dependent on atomic number (for Au it is higher only by a factor of two than for C, see fig. 9). The main factor for secondary electron yield is the angle between incident beam and sample surface. Since the envelope of the excited volume (see fig. 7) moves closer to the surface when the beam hits the surface at a small angle, SE yield increases. Thus, variations of surface inclination cause pronounced changes of SE yield (fig. 9). This fact is used for obtaining information on surface topography (§ 3.3.1).

A series of techniques are available to modify the signals in order to obtain better or additional information, as, for example, *black-level suppression* (i.e. differential amplification which distributes the contrast over the full range of the cathode-ray tube or the photographic film) or *nonlinear amplification* (contrast enhancement by preferential contrast expansion at either end of the gray scale, improving the visibility of features in otherwise dark holes). In *Y-modulation*, the CRT (cathode-ray tube) beam is deflected proportionately to the detector

Table 2
Physical effects producing radiation, detector types and detected signals used for imaging and analyzing metal surfaces in the scanning electron microscope.

Detected signal	Type of detector	Information	Basic effects	Maximum resolution	Minimum depth of information	Remarks
Secondary electrons (SE)	Scintillator photomultiplier with Faraday cage	Surface topography, material contrast, crystal orientation constrast	SE yield depends strongly on surface tilt and weakly on atomic number and crystal orientation	5–20 nm	1–2 nm	Background due to SE excited by BE reduces resolution and enlarges depth of information. Material contrast can be suppressed by superimposing the inverted BE signal
Back-scattered electrons (BE)	Solid state or scintillator photomultiplier	Material composition, topography, crystal orientation	BE yield depends on atomic number, increasing for heavier elements	0.1–1 μm	0.1–1 μm depending on primary electron energy (acceleration voltage)	Topological contrast can be suppressed by a ring-shaped detector. Higher resolution can be obtained by using an energy filter
Specimen current (absorbed or target current)	No external detector necessary	Complementary contrast to BE	BE yield results in corresponding electrical current	0.1–1 μm	As with BE	Conventional amplification difficult and noisy. Used in early development of SEM, may revive with better amplifiers
X-rays	Semiconductor detector	Element distribution	Emission of characteristic radiation by electron bombardment	~1 μm	> 1 μm	Element analysis by spectrometers, X-ray intensity images with point density corresponding to element concentration
Cathode-luminescence	photomultiplier with mirror	Detection of non-metallic and semiconducting phases	Emission of photons by electron bombardment	0.5–100 μm	0.5–100 μm	Applicable to metallic materials in rare cases only

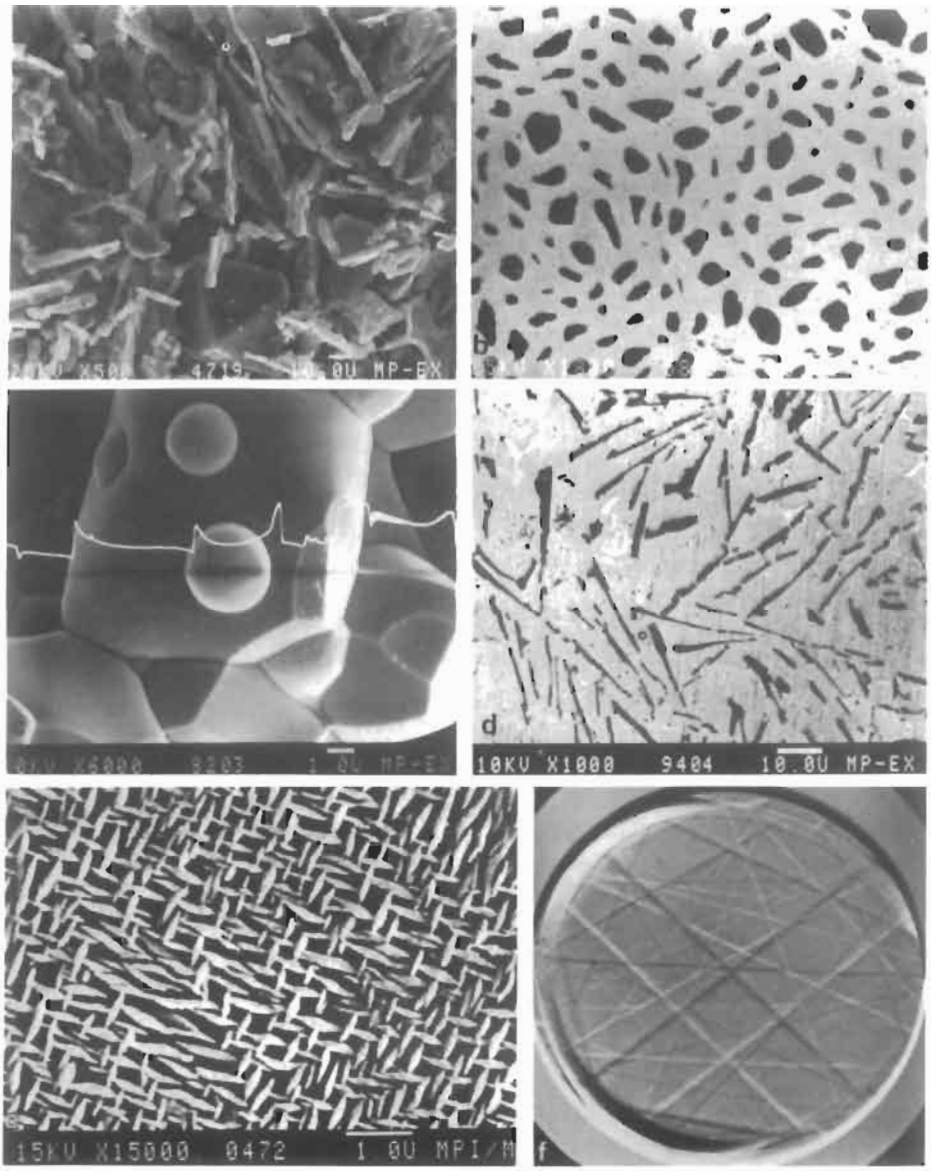


Fig. 8. Typical images produced in the scanning electron microscope (reduced in size for printing): (a) Topographic contrast by secondary electrons, aluminium–12 wt% silicon (aluminium matrix deep-etched). 300×. (From PAUL and MÜRRLE [1981].) (b) Atomic-number (material) contrast by backscattered electrons, silver-matrix/nickel-fibre composite. 1080×. (From MÜRRLE *et al.* [1980].) (c) Topographic contrast by secondary electrons, alumina–10 wt% zirconium oxide ceramic, thermally etched. The black line is a contamination trace of the line scan which produces the intensity profile shown by the white line (*y*-modulation). 3600×. (From BAUER and HALLER [1981].) (d) Secondary electron image revealing material roughness contrast. (The rough aluminium matrix appears brighter than the smoother silicon particles, due to polishing scratches.) Aluminium–12 wt% silicon alloy. 600×. (e) High-resolution backscattered electron image of partially stabilized zirconium oxide. (The martensitically transformed platelets are normal to the surface). 9000×. (f) Electron channelling pattern of tungsten grains in a tungsten–10 wt% nickel heavy metal. (From TAKAJO [1982].)

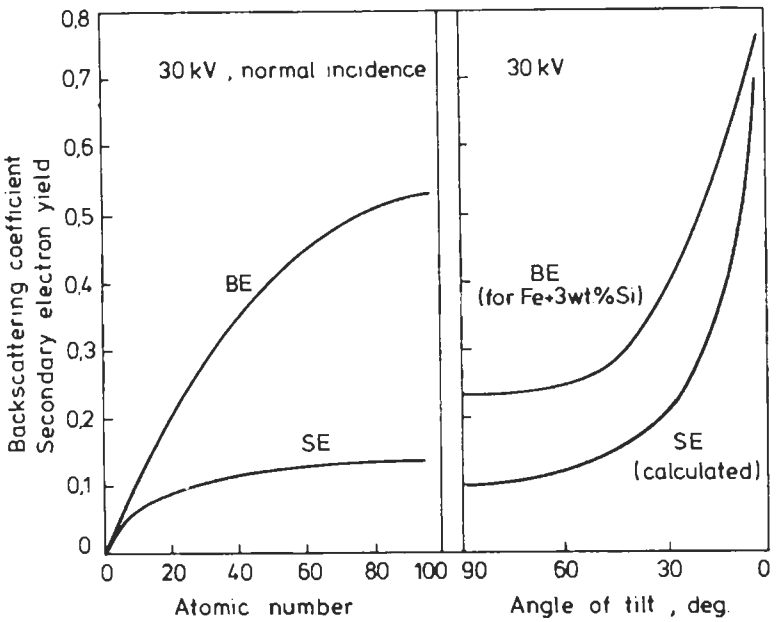


Fig. 9. Backscattering coefficient and secondary electron yield as a function of atomic number of the material hit by the electron beam (normal incidence), and as a function of inclination of surface with respect to the incident beam (schematic, after GOLDSTEIN and YAKOWITZ [1977]).

signal (fig. 8c) allowing detection of low contrast not apparent to the eye in the intensity-modulated image. If one of these signal transformations is used, its limitations and deficiencies must be considered in interpreting the image, otherwise artefacts may provide erroneous information. Some practical aspects of the more frequently used techniques for investigation in the materials laboratory have been reviewed by HILLMER [1979a], EXNER [1985], LYMAN *et al.* [1990] or GABRIEL [1990b], GOLDSTEIN *et al.* [1994], among others.

3.2. Specimen preparation

The major advantage of scanning electron microscopy is that in many cases little effort is necessary for specimen preparation. Practical experience has been reviewed for example by NEWBURY and YAKOWITZ [1977a], by HILLMER [1979b], EXNER [1985] or NEWBURY *et al.* [1986], and, more recently, by GABRIEL [1992b] and GOLDSTEIN *et al.* [1994]. Clean, deformation-free cross-sections and clean fracture surfaces of metallic materials can be investigated directly. Non-conducting materials must be coated by a thin conducting layer (10–100 nm, usually gold applied by sputtering or carbon and metals applied by high-vacuum evaporation) to avoid charging effects. Oxides caused by long-time exposure to the atmosphere or by high-temperature reaction can be removed electrolytically (YUZAWICH and HUGHES [1978]) or by hydrogen reduction (MADESKI [1980]). Cross-sections are prepared in exactly the same way as for light microscopy but

care must be taken in cleaning since residual polishing liquids or etchants trapped in pores or cracks cause contamination of the surface when putting the sample into the vacuum chamber. Organic mounting resins and any other hydrocarbons (grease etc.) must be removed in order to keep contamination by carbon and cracking products as small as possible. The shape of microstructural features in multiphase alloys can advantageously be studied after deep etching to remove some of the matrix metal (see, for example, McCALL [1973], FREMUNT *et al.* [1980], MADESKI [1980], HUN-DE and JING-YUN [1980], or PAUL and MÜRRLE [1981]). Figure 8a shows an aluminium-silicon alloy with part of the aluminium matrix removed by etching. The octahedral shape of the primary silicon crystals and the complicated arrangement of the eutectic silicon lamellae would not show up as clearly in a plane cross-section.

Etching is not necessary when *material* (*atomic-number*) *contrast* is utilized for image formation: Differences of atomic number show up as differences in brightness, the phase containing lighter elements appearing darker. Unavoidable height differences of hard and soft phases after mechanical polishing result in additional and usually unwanted topographic effects. However, differences in surface roughness after polishing have been shown recently to produce pronounced contrast between hard and soft phases if contamination is avoided (see § 3.3.1). Weak contrast mechanisms, as for example channelling electron contrast or magnetic contrast (see § 3.3.3 and § 3.3.4 respectively), are impossible to detect in the presence of a deformed layer or of topographic features. Therefore, deformation-free and plane cross-sections must be prepared by careful polishing when studying microstructures by means of these contrast mechanisms. Sample preparation by electron polishing or by ion polishing is advantageous for achieving maximum contrast and avoiding artefacts in these cases (NAKAGAWA [1987]).

3.3. Typical forms of contrast

3.3.1. Topographic contrast

The most pronounced contrast effects result from the dependence of secondary electron (SE) yield, back-scattering coefficient and detector collection efficiency on the angle between the surface element and the primary electron beam (see § 3.1) or the detector position, respectively. The resulting contrast (see fig. 8a) is analogous to an optical image where the light comes from the detector and the observation direction is that of the incident electron beam (NEWBURY [1977], REIMER and PFEFFERKORN [1977] or JOY [1992]). The stereoscopic effect is enhanced by shadows in regions which are hidden from the detector and by enhanced emission at edges. These latter effects are disadvantageous when deep holes and cracks or transparent edges are present. On the other hand, the edge effect can be advantageously be used to differentiate between phases of similar atomic weight for which material contrast (§ 3.3.2) is not pronounced. This has been demonstrated for an aluminium silicon alloy (fig. 8d), e.g., by PAUL and BAUER [1983], see also EXNER [1985].

The high spatial resolution (both transversal and in depth) obtained with secondary electrons excited by the incident beam is disturbed by secondary electrons excited by the back-scattered electrons coming from deeper in the specimen; thus the clear topographic

image can be obscured by atomic-number contrast and the resolution is reduced pronouncedly if special measures are not taken. There are several ways to reduce these effects (BLASCHKE [1979]). Superposition of the SE signal and the inverted BE signal seems to reduce interference from deep specimen layers almost completely (VOLBERT [1981] and PFEFFERKORN and BLASCHKE [1982]).

The vast majority of applications of the SEM in materials science (as well as other fields) rely on this mode while all other types of contrast together are used much less (estimated to less than 10 percent). The major reason for this fact is the high resolution obtained with SE micrography ranging from 5 to 10 nm in conventional instruments to 1 nm and below in the currently most advanced (ion-emission gun equipped) scanning electron microscopes.

3.3.2. Material (atomic number) contrast

As discussed in § 3.1, the yield of back-scattered electrons as well as that of secondary electrons is a function of the atomic number Z (fig. 9). Material contrast is useful for qualitative identification of phases and is especially suitable for quantitative evaluation of microstructural geometry by stereological techniques (see § 7). However, the atomic number contrast is usually obscured by topographic contrast (fig. 9) and, for secondary electrons, is sufficiently strong only in favourable cases, i.e., for ideally flat surfaces. For backscattered electrons, the topographic contrast can be greatly reduced even if the surface is rough by using ring-shaped detectors. Edge effects may still be a problem, and ideally flat polishing is necessary for difficult specimens in which the phases have similar compositions or consist of elements which are near neighbours in the periodic table. A difference of average atomic number, $\Delta Z < 1$ (i.e., even mixtures of nearest neighbours in the periodic table) is sufficient for slight material contrast, at least in the lower range of atomic numbers where the contrast is more pronounced (see fig. 9). Figure 8b shows an example of strong BE material contrast.

The limitation for many problems in physical metallurgy is the lack in resolution due to the large depth of information (see fig. 7) which, for highly dispersed phases, results in blurred edges and detection of features not intersected by the cross-section but close below it. However, if phase interfaces are normal to the cross-section, excellent resolution is obtained (fig. 8e).

By reducing the energy of the primary beam, disturbing effects can be reduced. Table 3 shows the depth of information as a function of primary voltage. Even at low voltages, the depth of information from metal surfaces is at least one order of magnitude larger for backscattered electrons than for secondary electrons. Material contrast is reduced at lower voltages and additional contrast-enhancement techniques by surface treatment have been worked out for individual cases.

A large number of problems in physical metallurgy and materials science have been solved using the BE mode, particularly since the early semiconductor-based detectors were substituted by optimized scintillator–light guide–photomultiplier (SLP) detectors (Robinson or YAG type). Some useful applications of these types of detectors were reported by BAUER and EGG [1984], for example.

Table 3
Information depth obtained with secondary and backscattered electrons for a few elements as a function of acceleration voltage of the primary electron beam, and with secondary electrons (according to SEILER [1968] and REIMER and PFEFFERKORN [1977]).

Element	Atomic number	Density (Mg/m^3)	Information depth (nm)					
			back-scattered electrons					secondary electrons
			5 keV	10 keV	20 keV	30 keV	50 keV	
C	6	2.3	330	970	2800	5300	11 600	10
Al	13	2.7	120	360	1050	1950	4290	1.2
Cu	29	8.9	40	110	320	590	1300	0.5
Ag	47	10.5	30	90	270	500	1100	1.0
Au	79	19.3	20	50	150	270	600	1.8

3.3.3. Electron channelling contrast, electron channelling and Kossel patterns

Electron channelling patterns (ECP) arise because of the fact that the primary electrons penetrate into the crystal to a depth which depends on the atomic packing density along different crystallographic directions. If the electrons follow the channels between rows of atoms, their re-escaping probability becomes low. A large number of Bragg conditions are satisfied with quite small angular variations. By rocking the primary electron beam about a point in the sample, diffraction lines from a small area ($< 10\ \mu m$) are obtained. These *selected-area electron channelling patterns* (SACP) are similar to Kikuchi patterns obtained in transmission electron microscopy (ch. 11, § 3.2) and are therefore often called pseudo-Kikuchi patterns. Information on the crystal orientation and crystal perfection, grain boundaries, twins and other crystallographic features is obtained from minute regions of a surface layer less than 50 nm thick. An example is shown in fig. 8f. Extensive discussions of electron channelling contrast show its usefulness for materials investigations (MCCALL [1973], NEWBURY and YAKOWITZ [1977b], PAYNE [1982], VERHOEVEN [1986], NEWBURY *et al.* [1986], JOY [1992], RIEDLE *et al.* [1994]). In the investigation of fracture surfaces, bend contours were observed by ECP (DAVIDSON [1974]). The study of orientation relationships of twins and of local textures are prominent examples of successful application for SACP. Though more often used in transmission or with higher primary energy in the electron microprobe analyzer, *Kossel patterns* (named after the explorer of this effect) are another efficient means to analyze the structure of small volumes in the order of μm^3 in the scanning electron microscope. Kossel lines arise by interference of X-rays excited by the focussed electron beam at the surrounding lattice planes. On the film, lines with bright-dark fine structure are obtained which are easily interpreted and give useful information on orientation, lattice parameter and lattice deformation after distortion (BRÜMMER and NIEBER [1975], DINGLEY *et al.* [1982] or NEWBURY *et al.* [1986]).

An interesting study of orientation relationships of new grains formed by recrystallization of aluminium, using the transmission Kossel technique in a modified electron microprobe analyzer by DOHERTY and CAHN [1972] is discussed in ch. 28, § 3.3.

References: p. 1016.

A new type of microscopy called "*orientation image microscopy*" (ADAMS *et al.* [1993] and WRIGHT [1993]) uses backscattered Kichuchi patterns to determine lattice orientation in localized regions (approximately 0.2 μm). Rapid indexing of the diffraction patterns by computer analysis is combined with computer-controlled movement of the electron beam or the sample stage. Thus, a computer image of the surface is obtained displaying orientation differences of the grains in polycrystalline material with a resolution of 1 degree misorientation. This technique was shown to yield information exceeding that by X-ray texture analysis and stereological grain size analysis as demonstrated for rolled aluminium and iron-silicon alloys (ADAMS *et al.* [1993]).

The above-mentioned methods have recently been discussed and critically compared by RANDLE [1993].

3.3.4. Magnetic contrast

Magnetic fields of ferromagnetic crystals can affect the interaction of the primary beam or the resulting emission. *Type-I magnetic contrast* uses the deflection of the highly directional electrons by the leakage field which may amount up to 20% for materials with strong fields such as cobalt. The resolution with which the boundaries of magnetic domains can be picked up is only of the order of several μm owing to the diffuse nature of leakage fields. *Type-II magnetic contrast* arises from the deflection of primary electrons by the Lorentz force inside the crystal. Magnetic domains appear in light–dark contrast due to differences in backscattering coefficient, with a strong tilt-dependence, and can be enhanced by filtering using high-energy BE only, with a resolution in the order of 100 nm. (For details and applications of magnetic contrast see NEWBURY and YAKOWITZ [1977b], REIMER and PFEFFERKORN [1977], ELSBROCK and BALK [1982], VANDER VOORT [1985c], NEWBURY *et al.* [1986], VERHOEVEN [1986] or JOY [1992]).

3.3.5. Charge collection microscopy and electron-beam-induced current measurements

Secondary electron emission is sensitive to surface potentials and electric field gradients. Therefore, in samples where local variations of potentials appear, regions of differing brightness will show up in the SE image. This effect can be enhanced by a positive or negative bias of a few volts which impede or enhance SE emission, respectively. Usually (though incorrectly) called voltage contrast, this technique provides another way to produce images in the SEM. EBIC (electron-beam induced current) and EMF (electromotive force) measurements depend on similar phenomena (creation of excess electron/hole pairs by the electron beam) and give useful information on diffusion length and lifetime of minority carriers in semiconductors. The term "charge collection microscope" (LEAMY [1982]) has been coined for these techniques which make the scanning electron microscope one of the most versatile tools for the characterization of semiconductor materials and failure analysis in the electronic industry (see, for example, NEWBURY *et al.* [1986], SCHAPER and BÖSEL [1985], or VERHOEVEN [1986]).

3.3.6. X-ray mapping

Emission of fluorescent X-rays can be picked up by a solid-state detector and

analysed with respect to its energy. For producing an image (i.e., a so-called X-ray map), an energy window is set which selects X-ray quanta typical for a specific element. The efficiency of the X-ray process is small, typically 10^{-4} photons/incident electron, and only a few thousand photons reach the detector per second. The beam of the display cathode-ray tube receives an intensity pulse whenever a certain preset number of X-ray quanta of preset energy reaches the detector. In this way, regions with a high concentration of the corresponding element appear bright while others appear dark. If this is done consecutively for several elements, X-ray maps of the distribution of these elements in the region scanned by the electron beam are obtained.

Owing to the large depth of X-ray emission (see fig. 7) and the large background noise due to Bremsstrahlung, the resolution of X-ray maps is limited to a few μm. Nevertheless, X-ray maps are used extensively for demonstrating element distributions in metallic materials allowing a quick and unambiguous identification of phases and concentration gradients caused, for example, by segregation during solidification or by diffusion at the interface between phases which are not at thermodynamic equilibrium.

3.3.7. Cathodoluminescence

The recombination of electron/hole pairs causes a release of energy which is emitted from the sample. Part of the emission is in the visible or infrared part of the spectrum and can be collected and amplified to form an image. The ability of some materials to emit this long-wave radiation is known as cathodoluminescence (CL). Only few examples have been reported with metallic materials where radiation is caused by surface plasmon effects (e.g., SIEBER [1982] or RICHARDS and TRIGGS [1982]). Ceramic materials and semiconductors have been occasionally studied by CL, for example Al_2O_3–ZrO_2, MgO and ZrO_2 ceramics. SiC fibres or diamond (CZERNUSKA and PAGE [1985], BROWN *et al.* [1988] and PAGE [1993] and GaAs (DAVIDSON [1974], see also PFEFFERKORN *et al.* [1980], HOLT and SABA [1985], NEWBURY *et al.* [1986] or JOY [1992]). Mineralogical and particularly biological applications are frequent (see, for example, REIMER and PFEFFERKORN [1977], GOLDSTEIN and YAKOWITZ [1977], GOLDSTEIN *et al.* [1994]).

3.4. Accessory equipment

The large specimen chambers of most commercial instruments allow special stages to be mounted by which various experiments can be carried out to yield additional information on the nature of metallic microstructures and their development.

3.4.1. Stereomicroscopy

Simple tilting stages or goniometer stages allow stereopair micrographs to be taken which give an excellent three-dimensional impression of rough surfaces when viewed in a stereoscope, by the "anaglyphe" method using a red and a green filter for the two images or by viewing with polarized light. The tilt angle and the viewing distance determine the subjective impression of depth. Quantitative evaluation of the parallax yields accurate data on the z coordinates at regularly spaced or arbitrarily chosen x–y positions. These data can then can be combined to construct height profiles or height

maps (e.g., BLASCHKE and WALTINGER [1971], BOYDE [1973], ARZT and FISCHMEISTER [1979], BAUER and EXNER [1981], BEATON and FILSHIE [1983], GABRIEL [1987, 1992a], LYMAN *et al.* [1990], LIENKAMP and KUNAVER [1994]). This technique can be used to obtain height differences destruction free and with high resolution which will often compare favourably with stylus-type instruments) and characteristic parameters (roughness indices, distribution of tilt angles for surface elements, etc.) for fracture surfaces and other rough surfaces. Figure 10 shows the profile map of a ceramic fracture surface together with the tilt-angle distribution. Further applications for solving problems in materials science have been reported for example by EXNER and FRIPAN [1985], SIGL and EXNER [1989], FEIJOO and EXNER [1991], LIENKAMP and EXNER [1994], and by many others.

3.4.2. Dynamic and non-ambient-temperature SEM

Large depth of focus and the possibility of rapidly changing the magnification in combination with mechanical or low- and high-temperature stages are prerequisites for continuous observation of specimens subject to applied stress, magnetic or electric fields, chemical reaction, and various effects due to cooling or heating. A multitude of phenomena in physical metallurgy have been studied, including fatigue crack growth in ferrous alloys (SCHAPER and BÖSEL [1985]), crack propagation and delamination in hard-layer coated steel during cyclic bending (WETZIG *et al.* [1984]), cracking of surface layers during heat treatment (GRABATIN *et al.* [1983]), deformation of superplastic lead–tin eutectic, the effect of stress and magnetic field on the configuration of magnetic domains in an iron–silicon alloy (NEWBURY and YAKOWITZ [1977b]), solid-state sintering (FULRATH [1972]), or liquid-phase sintering (RIEGGER *et al.* [1980]). Video-recording is an ideal way to register events of interest, and special devices have been developed to record fast processes such as cracking or martensitic transformation with much higher resolution than is possible with optical microscopy.

4. Scanning tunneling, atomic force and related microscopies

Only a little more than a decade ago, the application of the concept of electron tunneling known since the early 1920's led to the development of a new type of microscope, the scanning tunneling microscope (STM) by BINNIG *et al.* [1982], BINNIG and ROHRER [1982]. STM is not just another type of high-resolution electron microscopy. A whole class of novel techniques based on an astonishingly simple concept allows the detection of almost any kind of interaction on a nanometer or even subnanometer scale. Some of these techniques (scanning tunneling microscopy, scanning tunneling spectroscopy, atomic force microscopy and magnetic force microscopy) have rapidly evolved into routine tools for surface characterization, and new techniques based on similar principles are invented each year. In the decade of its development, dramatic impacts in fields as diverse as materials science, biology or electro-chemistry (among many others) have been made, and the papers published exceeded the thousand some time ago with a rapidly increasing rate of several hundreds each year. Review papers and monographs became available (e.g., HANSMA and TERSOFF [1987],

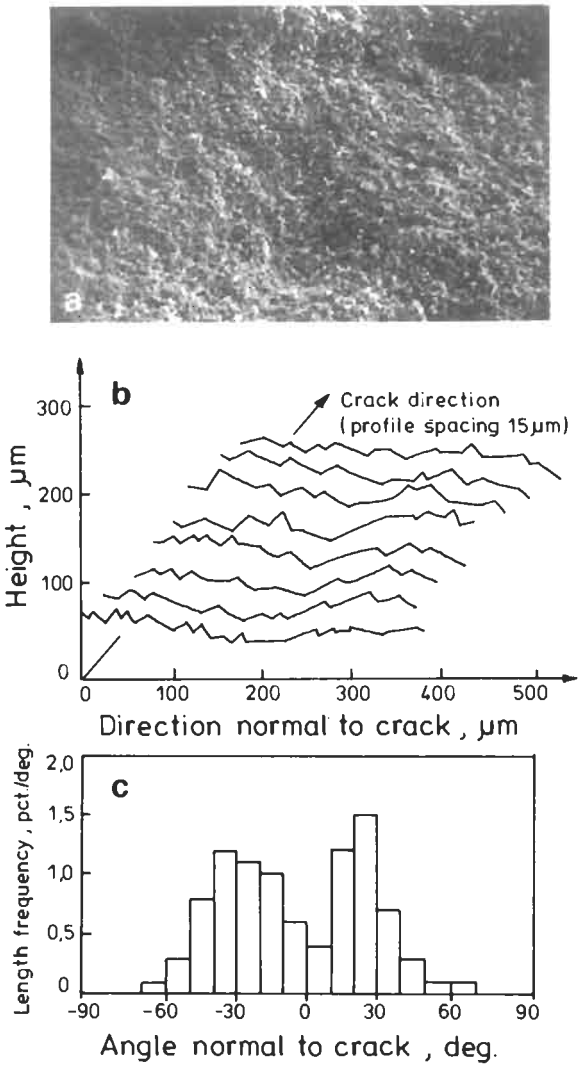


Fig. 10. Fracture-surface analysis by instrumented stereoscopy: (a) scanning electron micrograph evaluated by stereometry; (b) line profiles; (c) distribution of profile line length as a function of tilt angle of surface profiles perpendicular to crack direction. (After BAUER *et al.* [1982].)

BINNIG and ROHRER [1987], GARCIA [1987], WICKRAMASINGHE [1989, 1992a,b, 1993], BONNELL [1990, 1993a,b], SARID [1991], GÜNTHERODT and WIESENDANGER [1992], WIESENDANGER and GÜNTHERODT [1992], STROSCIO and KAISER [1993], DINARDO [1994a, b], see also Further reading). The rapid development makes it difficult to follow up all issues of interest, even in a single field like physical metallurgy, and only some basic principles and a few examples of typical applications will be presented in the following.

References: p. 1016.

4.1. Basic principles and capabilities

An infinite potential energy barrier exists between two surfaces separated by a large distance. This barrier becomes finite when the surfaces are in close proximity, typically in the order of a few nanometers, and, when a voltage is applied, a statistically significant number of electrons can penetrate the energy barrier. This effect was called electron tunneling and has been theoretically very well understood and quantitatively described on the basis of the classical wave particle dualism concept of quantum mechanics though a microscopic quantum-mechanical theory taking geometric effects into account is needed in the case of atomic resolution (see, for example, TERSOFF and HAMANN [1985], TERSOFF [1993], TERSOFF and LANG [1993]).

The overall scheme of a scanning tunneling microscope is shown in fig. 11. To produce images of a surface at very high resolutions, a sharp metal tip, often made of tungsten or platinum, is scanned over the surface of the sample under investigation. The motion of the tip is controlled by piezoelectronic elements (usually lead zirconium titanate) in x and y directions by applying simple linear voltage ramps. A third element controlling the z position of the tip can be operated in two modes: in "constant current" imaging, the tip follows a constant charge density contour representing the surface profile by keeping tunneling current constant. This is achieved by a computer-controlled feedback mechanism. The signal required to alter the vertical tip position produces the image. Altering the level of constant current (or the voltage applied between the tip and the surface) produces contours of different charge densities. In the second mode, "constant height" operation, the tip is held at a fixed z position and, while scanning the surface, changes its relative position to the sample surface. Thus, the sample/tip separation and, as a consequence, the tunneling current vary which is used as signal.

As the tunneling currents are in the range of 10^{-9} A, the signal must be amplified by 7 to 10 orders of magnitude. Effective shielding, avoiding mechanical vibrations by combining a series of damping strategies, careful sample preparation and a clean and sharp tip are critical for obtaining optimum resolution and unambiguous interpretation of the images. In contrast to conventional electron microscopy, scanning tunneling microscopy can not only be performed in vacuum but also in air or even in liquids.

4.2. Atomic force microscopy

Only a little later than STM, in 1986, BINNIG *et al.* [1986] developed the atomic force microscope (AFM) capable of measuring or imaging the forces between the sample surface and the tip. This is achieved by attaching the tip to a cantilever. Forces acting on the tip produce a deflection which can be measured down to 10^{-12} m corresponding to forces in the order of 10^{-13} N. (At the time this article was written, commercial instruments have force and depth resolutions of approx. 10^{-7} N and 4×10^{-11} m, respectively). Since these forces can be measured locally with lateral resolutions in the order of 0.1 nm, the forces between groups of atoms or even single atoms can be measured, and force microscopy has rapidly developed to a quantitative probe of all kinds of surface forces as electrostatic, capacitive, magnetic, adhesive, capillary, frictional or van der Waals (SARID [1991], MEYER and HEINZELMANN [1992], HUES *et al.* [1993], BURNHAM

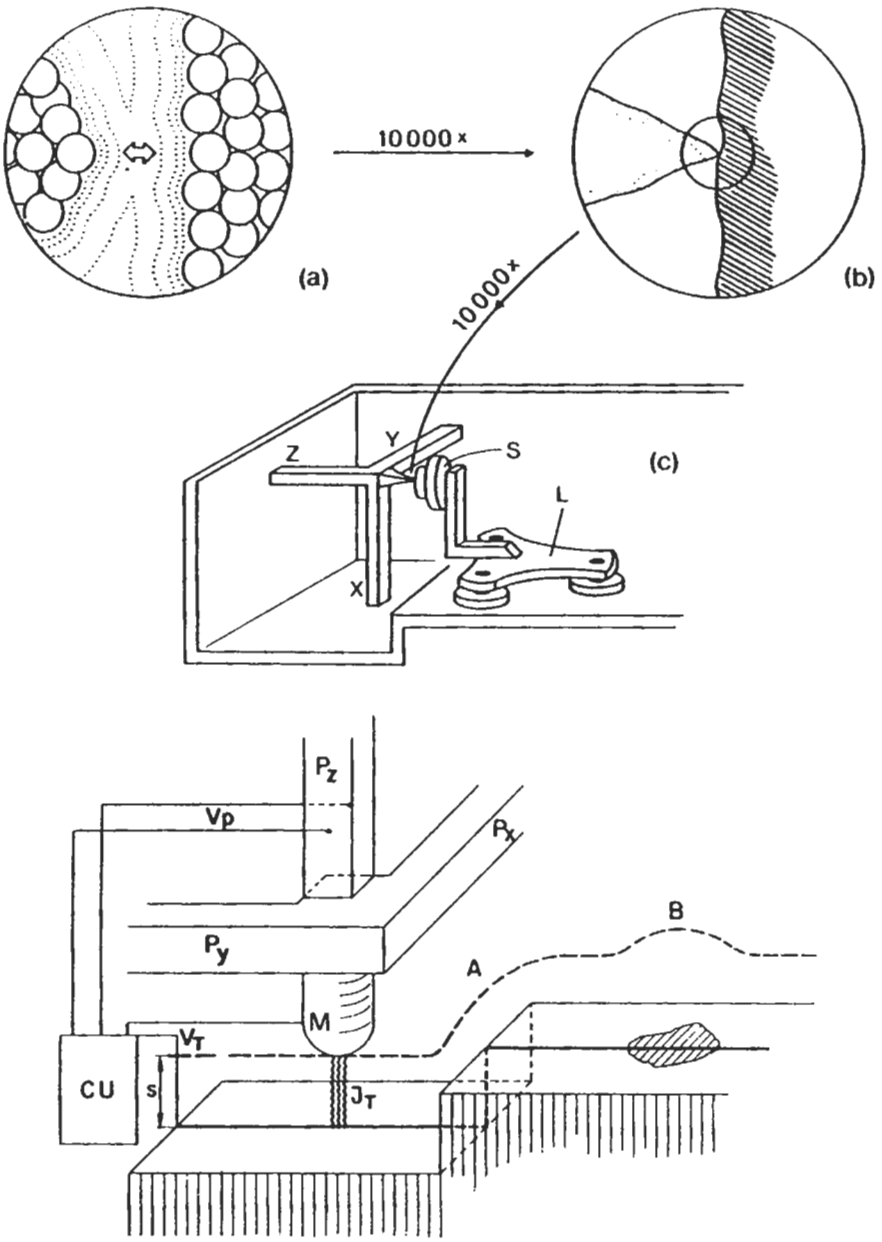


Fig. 11. Overall scheme of the scanning tunneling microscope (after BONNELL [1993c]).

and COLTON [1993]). The theoretical treatment of AFM is a complex problem and is still in progress involving empirical potentials, molecular dynamics (ABRAHAM *et al.* [1989]),

continuum elasticity theory (TOMANEK *et al.* [1989]) and ab-initio calculations (for a recent review see MEYER and HEINZELMANN [1992]). Though there are a number of open questions in interpreting the major features of force distance curves and the nature of the imaging mechanisms as discussed by BURNHAM and COLTON [1993], atomic force microscopes are now commercially available at reasonable prices, capable to measure surface forces and nanomechanical properties of materials as well as to modify surfaces.

4.3. Tunneling spectroscopy

Tunneling spectroscopy (TS) is another source of information available with the scanning tunneling microscope (see, for example, HAMERS [1993], STROSCIO and FEENSTRA [1993], BONNELL [1993b]) providing data relating to the local electronic structure of the surface. Compared with other surface spectroscopy techniques such as photoemission microscopy (UPS and IPS), electron energy loss spectroscopy (EELS) or infrared spectroscopy (IS and IRRAS) which provide information averaged over a large surface region (see § 6), with TS it is possible to measure the electronic structure on an atom-by-atom basis. Surface irregularities as impurities, steps, defects or electronically inequivalent atoms in ordered structures can not only be seen by STM but also be correlated with their effect on local changes of the electronic structure. Point spectroscopy involves moving the tip to a feature of interest, modulating the tip voltage and recording the resulting change of tunneling current. In a second method for obtaining spectroscopic information, the tip voltage is modulated at high frequency with respect to the time constant of the feedback controller and the current is recorded at discrete values of the tip voltage. Though well funded by bulk tunneling (see, for example, HANSMA [1982] or WOLF [1986]), proper analysis of the data is crucial to the use of this powerful technique and is presently still in the state of development.

4.4. Related scanning techniques

The same scanning and feedback principles as in scanning tunneling microscopy can also be applied to use other types of interactions between the sample and the tip for characterizing the properties of surfaces and to produce images. In addition to atomic force microscopy (AFM), *magnetic force microscopy* (MFM), after the first successful attempts in 1987, has been developed as a technique of its own with commercial instruments becoming available and interesting results being reported (for a recent review see GRÜTTNER *et al.* [1992]). The principle is the same as that of the atomic force microscope, but the tip is made by electrochemical etching of ferromagnetic foils or wires, most often nickel, but also iron, cobalt or amorphous FeBSiC, or by coating the tungsten tip with a ferromagnetic thin film. The probe senses the magnetic stray field (which is not easily related to sample magnetisation and, therefore, may be a disadvantage of this technique compared to electron probe based techniques like *scanning electron microscopic polarisation analysis* (SEMPA)). Also, lateral resolution is routinely limited to 40 to 100 nm with optimum values of 10 to 20 nm. On the other hand, MFM can be operated in air and requires little or no sample preparation.

Another modification of the force microscope is the electrostatic force microscope.

A family of techniques yielding information on the structure of buried interfaces, is *ballistic electron spectroscopic emission microscopy* (BEEM), useful for studying the transport and scattering of electrons and holes in multilayer structures (see reviews by KAISER *et al.* [1993] or BELL *et al.* [1993]). Furthermore, numerous STM related techniques have been invented and demonstrated to yield specific information on materials surfaces. Among those described in recent reviews (WICKRAMASINGHE [1992a,b, 1993], POHL [1992]), are the *scanning noise microscope*, the *scanning tunneling potentiometer*, the *scanning photon microscope*, the *inverse photoemission microscope*, *near field thermal microscopy*, *scanning optical* and *acoustical probe microscopes*, with more to come up in the future (for a list of more than 20 different techniques of scanning microscopy see WICKRAMASINGHE [1993], for details see the most recent publications on STM conferences).

4.5. Applications

The major reason for rapid acceptance of scanning tunneling microscopy is the fact that three-dimensional images of the real surface are obtained at unprecedented levels of lateral and vertical resolution. Under optimized sample and tip conditions, quantitative information about topography and electronic structure can be obtained for individual surface atoms as well as for groups of atoms, defects, etc. Thus, STM is not just another microscopic technique but, as discussed above, comprises a new class of surface analysis techniques not available when the last edition of this book was published. Applications are manifold, and, though the majority of research papers are aimed to demonstrate the opportunities of STM and related techniques under extreme conditions, applications to actual problems of physical metallurgy and materials science have been reported.

Obviously, scanning tunneling microscopy is the ideal technique to study the arrangement of atoms on the surface of crystalline solids. This arrangement influences processes like epitaxial growth of thin films or catalytic reactions as well as surface properties like surface energy or electronic behaviour (see the recent review by UNERTL [1993]). Since surface atoms have dangling bonds, they rearrange in order to minimize energy and form a structure which is said to be reconstructed (ZANGWILL [1988]). The reconstructions of the (111) and (100) surfaces of Si are probably the ones most often studied (see, for example, GIESSIBL [1995]. The (100) and the (110) surfaces of Ir, Au and Pt have similar reconstructions. Reduced atomic densities are observed for the (110) surfaces called missing row structure because every second row of atoms has been removed. Surprisingly, the atoms in the (111) surfaces show a higher density than in the bulk. These and many other results have been reviewed by UNERTL [1993], OPIELKA *et al.* [1993] or by STROSCIO and KAISER [1993]. An example of an STM image is shown in fig. 12. The local symmetry reflects the local electronic density of states. The defects in the periodic structure (upper left) are due to impurities.

For the study of amorphous materials or where large single crystals do not exist, STM and related techniques are presently the best structural methods available. Their major limitations are that the subsurface structure cannot be detected and the atomic position cannot be determined with the same degree of precision as with diffraction techniques. These and other limitations of STM and AFM in surface crystallography are

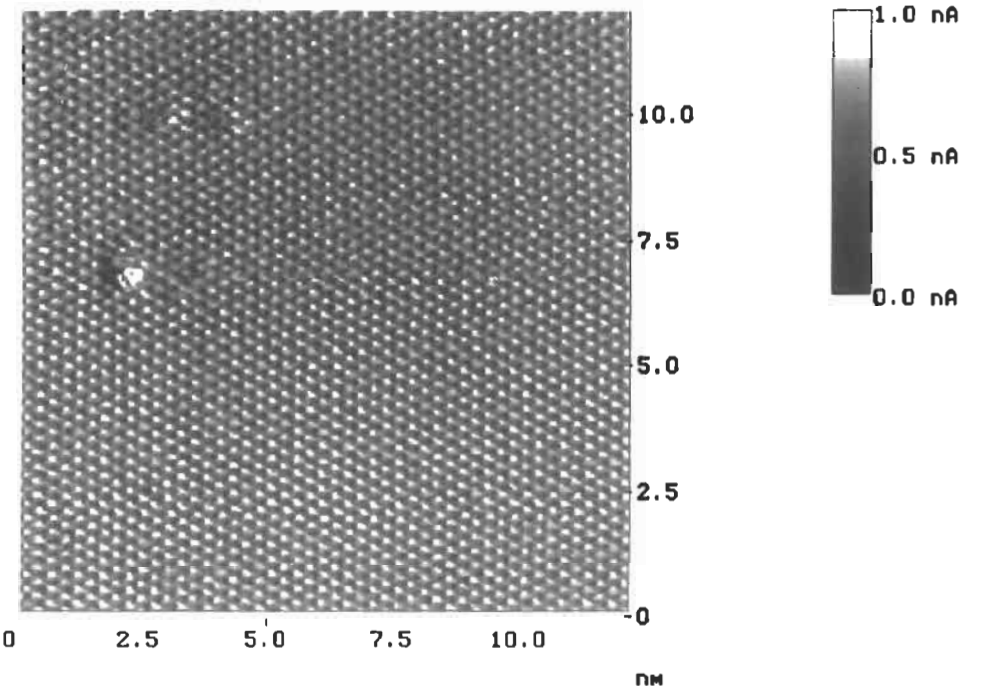


Fig. 12. Scanning tunneling micrograph of the (0001) surface of graphite (courtesy of M. GÖBEL and VON BLANCKENHAGEN, Institut für Materialforschung I, Kernforschungszentrum Karlsruhe). Scan size 12 × 12 nm, bias –60.8 mV; z range 1.0 nA (black: 0, white 1.0 nA). Note the scales are in nm at the axes and the grey levels indicate nA.

discussed in recent reviews (e.g., UNERTL [1993]).

Surface diffusion at low temperatures (down to room temperature for Au and other metals) was measured by STM and MFM with interesting correlaries on the possibility to extrapolate high temperature data. Hillock formation by surface diffusion, surface phases and their transformation with temperature, surface premelting (explaining why solids can not be superheated above their melting points) and many other phenomena (see, for example, KUK [1993], STROSCIO and KAISER [1993] and the reviews cited above and under Further reading at the end of this chapter) were studied by the techniques discussed in this paragraph.

The ability to apply scanned probe techniques to solid/liquid interfaces opens a wide field of applications in adhesion, friction, materials processing and, most widely studied, electrochemistry. Special instruments for electrochemical STM have been developed and a wealth of information on the evolution of surfaces and films during electrochemical treatments of surfaces has been reported (for reviews see SONNENFELD *et al.* [1990], SIEGETHALER [1992] or CATALDI *et al.* [1990], BARDA and FAN [1993]). For example, the development of smooth areas during electropolishing, preferential attack at dislocations, roughening during etching, the formation of islands in the early stages of film

deposition and long-range ordered structures on the surface of etched Si, like those obtained in ultra-high vacuum, were observed.

Images of magnetic recording media were produced by MFM and their recording behaviour was investigated. Laser-written domains in magnetically hard and soft materials as well as natural domains and Bloch wall segments in a wide range of magnetic materials were imaged, and magnetic surface forces were measured. Again, there are a number of limitations, some of which will certainly be overcome in future developments.

The various spectra, obtained with tunneling spectroscopy, particularly tunneling current/applied bias curves, have been analysed for their information on the correlation between the position and the electronic structure of surface atoms.

With the development of nanotechnology more and more questions will seek answers which can only be given on the basis of techniques reading atomic resolution. With the exception of atomic force microscopy, the use of these techniques is rather specialised at present, and it remains to be awaited which of them will contribute most to solve materials science problems.

5. *Other special techniques of surface microscopy*

In addition to light microscopy and transmission and scanning electron microscopy used routinely in all fields of materials research, development and control, microstructures can be investigated by several more exotic image techniques. While some of these, such as scanning tunneling, atomic force, photoemission or field-ion microscopy, are of high interest for various advanced studies of material surfaces, others are still in the stage of experimentation, have been substituted by other techniques or are more useful in other fields of application like biology or mineralogy.

Instruments capable of analyzing the chemical nature and the electronic state of surface atoms have been developed at a rapid rate in the last few years, utilizing all kinds of interaction with incident photons, electrons and ions. In § 6, a summary of these techniques in a classification with respect to exciting and emitted species is presented. The information obtained stems nearly exclusively from regions close to the surface and thus is decisively distinct from that furnished by bulk chemistry. If the lateral resolution is sufficiently good, the signals can also be used for surface imaging.

A full account of all available techniques cannot be given in this context but short surveys of those being used to some extent in the investigation of metal surfaces in physical metallurgy will be given in this section.

5.1. Scanning acoustic and thermal wave microscopy

Though based on early discoveries and inventions (for reviews on historical developments see WICKRAMASINGHE [1983], SOMEKH [1990] or BRIGGS [1992]), *acoustic* and *thermal-wave microscopy* are rather new developments (KORPEL *et al.* [1971], LEMONS and QUATE [1974], WICKRAMASINGHE *et al.* [1978], KOCH [1979], ASH [1980], ROSENCWAIG [1982] and MURPHY [1993]).

Various types of signals and detectors have been used to form magnified images and, depending on the physical principle, the instruments have been called acoustic, thermal-wave, photoacoustic, or electronacoustic microscopes or thermal-wave and thermographic microscopes, or thermoacoustic probes. Either a piezoelectric crystal is attached to the surface which emits acoustical signals, or alternatively, a laser or an electron beam are focussed on the surface resulting in periodic surface heating as the beam intensity is modulated and thermal and acoustic waves are sent through the sample. After reflection or scattering, the signals are picked up by a suitable acoustic or thermal detector (ultrasonic transducer, gas microphone, infrared detector) and transformed back into electrical signals. By applying the scanning principle, the efficiency of these types of microscopes can be fully utilized. The laser or electron beam is scanned across the field of view, and the signal received is used to modulate the brightness of a cathode-ray tube. The information furnished is different from that obtained with optical microscopes or with scanning electron microscopes in that it reveals sub-surface features such as, for example, pores and other defects, grain boundaries and microstructural components which differ from the surrounding matrix by their acoustic or thermal properties (density, specific heat, thermal conductivity).

5.1.1. Scanning laser acoustic microscopy

The *scanning laser acoustic microscope* (SLAM), also termed *scanning photoacoustic microscope* (SPAM) has now been developed to a commercially available instrument owing to its high versatility. In the range of high frequencies (1 to 10 GHz) the laser scans typically a field of 200 μm square, and the typical magnification is 1000×. The penetration depth is in the order of 100 μm. Since at these high frequencies the wavelength is similar to that of visible light (3 GHz corresponding to 520 nm), the maximum resolution of the acoustic and the optical microscopes are comparable. At the other hand of the spectrum, at 1 Hz to 100 MHz, the resolution is low but the depth of information increases to the size of the sample, e.g., to 50 mm at 10 MHz where, at a typical magnification of 2 to 3×, the scanning area is approx. 100 mm square.

With the possibility of changing the frequency, the depth from which the information is obtained can be varied and, as with the confocal principle, three-dimensional images can be produced. Also, using stereographic or holographic type of image processing is useful not only to get 3D images but also to reduce the blemishing effects of superimposed features. Quantitative information on the depth of features below the surface, the thickness and composition of surface layers or the depths of cracks can be extracted (for a recent review see BRIGGS [1994]).

Various useful applications of scanning acoustic microscopes have been reported in the field of materials as, for example, by WICKRAMASINGHE [1983], TAM [1986], DIESER and HERRMANN [1988], DIESER and MATTHAEI [1988], MATTHAEI *et al.* [1988a,b], OPIELKA *et al.* [1990], ARNOLD [1991], GILMORE [1993], MURPHY [1993] or BRIGGS [1994]. With the commercially available instrumentation and higher resolutions to be achieved in the future by decreasing the noise and wavelength with extensive work under way further, at research institutes as well as at industrial firms (GILMORE [1993]), acoustic microscopy is a promising alternative for materials characterization and product testing.

5.1.2. Thermal-wave microscopy

With the *electroacoustic principle* excitation by pulsing the beam in a scanning electron microscope, a *thermal-wave microscope* is available by just attaching an additional detector to the SEM. The spatial resolution of such a device (also called thermoacoustic probe) is determined by the spot size of the incident beam, the modulation frequency and the thermal conductivity ranging from one μm or less for insulating materials at high frequencies (several GHz) to a few mm at 100 Hz. Owing to the fact that thermal waves are more sensitive to local variations and have a better resolution than acoustic imaging, there are numerous potential applications for the thermal-wave microscope as well. Nevertheless, this technique has not been routinely applied in materials investigations. Some applications of thermal-wave microscopy were reported by WINSCHUH *et al.* [1989, 1993] and MURPHY [1993] among others.

5.2. Field-ion and field-electron microscopy

Very high gradients of electric fields at the surface of a metal cause emission of electrons and ions. This is the basis of field-ion and field-electron microscopy.

5.2.1. Field-ion microscopy

Extremely high resolutions, down to the atomic dimensions (< 0.15 nm) are achieved in *field-ion microscopy* (FIM) and the *imaging atom probe* (IAP). A large number of monographs (e.g., MÜLLER and TSONG [1969], BOWKETT and SMITH [1970], WAGNER [1982, 1985], SAKURAI *et al.* [1990], TSONG [1990] or MILLER *et al.* [1992] and handbook articles (e.g., EDELMANN [1980], SMITH [1986], SEIDMAN [1993] or CEREZO and SMITH [1994]) give a full account of these and the related spectrometric techniques. In field ion microscopy, noble gas atoms (usually helium, but also neon or argon) or hydrogen are ionized by tunneling of electrons at the cooled surface of a pointed metal tip. The ions are extracted from the surface by a high voltage and accelerated to hit a channel-plate converter. The secondary electrons are multiplied and emitted radially to a fluorescent screen. A high resolution image of the tip is obtained showing individual atoms and their arrangement. As a typical example, the image of a tungsten tip is shown in fig. 13. Ionization occurs more readily at terrace steps which appear bright and outline facets of low-surface-energy planes. Lattice defects cut by the tip surface such as dislocations, stacking faults, grain boundaries or antiphase boundaries in ordered structures are revealed. Vacancies and interstitials can be observed and their movement studied by taking photographs at certain time intervals.

At the field strength produced at a pointed tip in the field ion microscope (approx. 5×10^{10} V/m at an applied voltage of 10 kV) it is also possible that atoms of the specimen material are ionized and are emitted from the surface and accelerated toward a detector the light output of which is recorded photographically. This mode of image generation is called field ion evaporation microscopy (removal of specimen material) in contrast to the field desorption microscope (removal of absorbed gas ions) described above. Field evaporation can be controlled with great precision by applying the potential as short high-voltage pulses. By using pulse times in the order of a few milliseconds, two to three

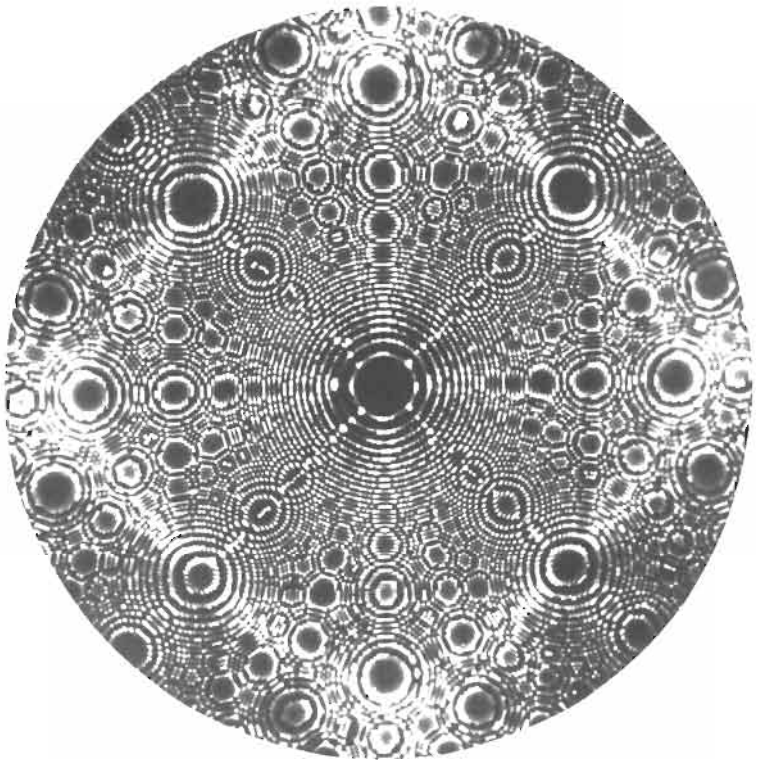


Fig. 13. Field-ion micrograph (reduced in size for printing) of a single-crystal platinum tip, (100)-oriented with facets of low-index planes. Magnification approx. 35×10^6. (From DOERR and OWNBY [1975].)

atoms can be removed at a time. The amount of the material analysed by the pulse technique is in the order of one million atoms per day, i.e. a volume of 10^{-23} m^3. The resolution is further increased by cooling the tip. With sharp tips, and at low temperatures, it is possible to image individual atoms but typically much larger areas with a higher ion output are observed in order to get a larger volume analysed, though at a decreased resolution.

5.2.2. Atom-probe field-ion microscopy

The chemical nature of the ions pulled off from the tip surface by the high-voltage pulse can be identified by attaching a *time-of-flight* (TOF) mass spectrometer. By positioning the tip with respect to an aperture hole in the fluorescent screen it is possible to focus each individual surface atom. In the detector system specially designed for this purpose, each ion arriving produces a single voltage pulse. A minicomputer controls a digital timer and analyses the data. Combining the time of flight information with the information on the location where the ion had been removed, it is possible to obtain an

atomic resolution image of microstructural features and to measure the mass-to-charge ratio of single atoms from preselected regions of the specimen. The lateral resolution of this instrument called *atom probe* is in the order of a few tenths of a nanometer while the depth resolution depends on the interplanar spacing which, for high index planes, is in the order of hundreds of a nanometer.

In another mode, called *imaging atom probe* (IAP) maps can be produced for different kinds of elements by switching the detector on during the time when the species of interest arrive at the screen after the pulse was applied. These maps are similar to the X-ray maps produced in the SEM with two decisive differences: The resolution of the atom probe maps in orders of magnitude better than that of X-rays. However, since in the imaging atom probe the surface atoms are removed by the pulsed voltage and time gating can only be performed for one species at a time, the maps observed for different elements can not be obtained from exactly the same parts of the specimen. For one element, three-dimensional information is obtained by analysing layer by layer, and the element distribution in the sample is obtained which offers unique opportunities in advanced studies of microstructure and its evolution.

Finally, the *pulsed-laser atom-probe field-ion microscope* or, in short, *pulsed laser atom probe* (PLAP) must be mentioned which uses a laser pulse instead of the voltage pulse on top of the steady-state dc voltage (KELLOGG and TSONG [1980]). Pulsed-laser atom-probe field-ion microscopy has been used to characterize semiconductors and conducting ceramics where a voltage pulse may lead to fracture due to thermal stresses.

5.2.3. Field-electron microscopy

Field-electron microscopes (FEM) are non-commercially made laboratory equipment in which an etched single-crystal tip is heated in high vacuum. The emitted electrons are accelerated by an anode and produce an image on a fluorescent screen. The intensity of electrons emitted (field emission current) depends on the voltage and the work of emission; the lattice structure and local geometric structure of surfaces can be studied with high resolution down to a few nanometers (MÜLLER [1936], EDELMANN [1980]).

The crystallographic structure of clean surfaces and (if by chance a grain boundary was located in the tip) the structure and the movement of grain boundaries as well as changes of the tip geometry and allotropic transformations of the tip material during heating have been studied; by measuring the energy distribution of the field electrons the electronic structure of the single-crystal tip was investigated (for references see EDELMANN [1980]). Absorption of gas from the vacuum chamber or of evaporated substances (metals or oxides) changes the image drastically, which has been used for studying the sites of adsorption, the migration of adsorbed species along grain boundaries and the formation of compounds.

It seems that this technique has been overrun and replaced by other developments since no mention is made in more recent reviews.

5.2.4. Applications of field-ion microscopy

In contrast to the statement made in the last edition of this book, the disadvantages connected with the high effort of preparing suitable tips and the extremely small volume

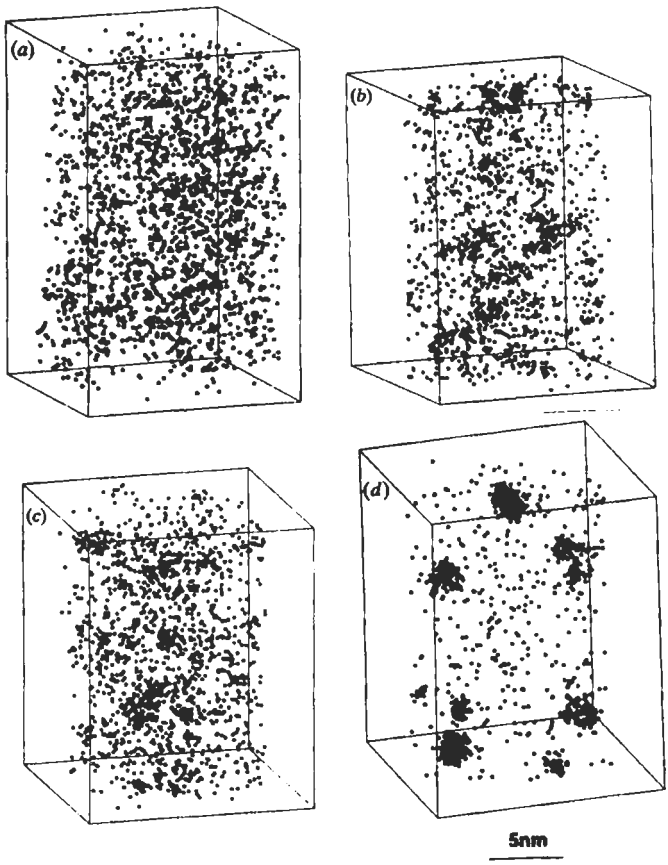


Fig. 14. Distribution of cobalt atoms in a very small volume (approx. 10^{-23} m^3) of a copper-1at.% cobalt alloy at different stages of annealing at 450°C. The annealing times are (a) 10 min, (b) 30 min, (c) 120 min and (d) 440 min (from CEREZO and SMITH [1994]).

analysed (requiring long analysis times and a large number of tips for obtaining statistically significant results) do not prevent the application of field ion microscopy, and many interesting studies of materials have been reported. The identification of small carbides in HSLA steels (KNEISSL and BRENNER [1990]) and particularly the study of the composition of carbides in secondary hardening in high speed steels (FISCHMEISTER *et al.* [1988]) are examples where field ion microscopy has contributed to the quantitative understanding of technically important materials. Extensive studies of radiation effects in metals have been carried out (for references see SEIDMAN [1993]).

In a recent review by CEREZO and SMITH [1994], a large number of case studies are discussed including secondary precipitates and segregation at γ/γ'-interfaces in superalloys, phases and Bloch walls in Alnico permanent magnets, ordering and site occupation in various alloys, spinodal decomposition of iron–chromium alloys, segregation at

grain boundaries and surfaces, oxidation processes and investigations of nanostructured and multilayer materials. Numerous field-emission microscopic studies have been made in the field of nucleation and growth where experimental results were needed to understand the morphology, crystallography and chemistry of particles *in statu nascendi* and growth for example of Guinier–Preston zones in various aluminium alloys (HONO *et al.* [1986, 1989]) and particularly of cobalt particles in copper cobalt alloys (WAGNER and BRENNER [1978], HAASEN and WAGNER [1980, 1985], WAGNER [1985], HAASEN [1985], and CEREZO *et al.* [1992]). Figure 13 shows, as an example for the kind of three-dimensional information obtained in such studies, the evolution of particles from a homogeneous Cu–1 at.%Co solid solution (CEREZO *et al.* [1992], CEREZO and SMITH [1994]), as analysed by the *position-sensitive atom probe field-ion microscope* (POSAP). For further examples of applications of the various types of field-ion microscopy as well as for details and references see the monographs by WAGNER [1982, 1985], MILLER and SMITH [1989], TSONG [1990], SAKURAI *et al.* [1990], MILLER *et al.* [1992], the overviews by BRENNER [1978], EDELMANN [1980], SMITH [1986], SEIDMAN [1993], or CEREZO and SMITH [1994] and the proceedings of various conferences (e.g., those of the International Field Emission Symposia) and current and special issues of journals devoted to surface analysis (e.g., those of Surface Science and Journal de Physique).

5.3. Photo-electron emission microscopy

Microscopes using electrons excited by ultraviolet light were developed in the late fifties and made commercially available at the end of the sixties (for reviews see WEGMANN [1972], SCHWARZER [1981]). In these *photo-electron emission microscopes* (PEEM), a high-intensity beam of ultraviolet light is focussed by means of quartz lenses and mirrors on a small area of a surface (area in the order of 0.2 mm^2, energy density several W/cm^2) which activates emission of relatively slow electrons. A high voltage is applied between the sample and a pierced anode facing the sample surface. These electrons, after passing the axial aperture of the anode and the magnetic lenses of a three-stage electron microscope, produce a highly magnified image on a fluorescent screen or a photographic film. Heating the sample up to temperatures of beginning thermionic emission allows *in-situ* high-temperature studies. Contrary to scanning electron microscopy, the light-beam cannot be focussed to a sufficiently small size, and integral image generation is used which has the advantage of small exposure times, varying with material and magnification from a tenth to a hundred seconds. The sample surface is cleaned by ion bombardment, but contamination is sometimes a problem in spite of the high vacuum in the chamber (typically 10^{-10} bar).

The intensity distribution of photo-electron emission is essentially determined by the variation of electron emission with material and orientation (material and orientation contrast), with the orientation of a surface element with respect to the optical axis (relief or topographical contrast) and with distortions of the magnetic field in the vicinity of tips and edges. The topographical contrast is stronger than in SEM images but the stereoscopic impression is usually less pronounced since the surface of the sample is always normal to the axis of projection and cannot be tilted. Material contrast is usually very

clear in unetched plane surfaces if the work-functions of different phases in a material are sufficiently different, which is usually the case. Orientation contrast is only pronounced in cubic metals but not in non-cubic metals and alloys, in oxides or in carbides. Thus, material- and orientation contrast can be easily separated in most practical cases.

Studies in materials research have provided much interesting information in all kinds of high-quality metallographic work (for early references see WEGMANN [1972] and SCHWARZER [1981]). Owing to the very small depth of information (10 nm), the high lateral resolution (< 10 nm), the excellent phase separation and the possibility for in-situ heating, photo electron microscopy is excellently capable for quantitative kinetic studies of changes in microstructural geometry (DANNÖHL *et al.* [1971], WEGMANN [1972] and MIDDLETON and EDMONDS [1977]). For example, a study by TAYLOR and POLLARD [1982] has revealed the bonding sequence (grain-boundary movement and annihilation) during diffusion-bonding of steel under load at temperatures up to 1000°C.

The major draw-back of PEEM has been the very high price of instrumentation. The production of the commercially built Metioscope was discontinued and recent application reports are scarce. The photo-electric effect using ultraviolet light excitation is now utilized in various kinds of spectrometers, like ESCA (*electron spectroscopy for chemical analysis*), and UPS (*ultraviolet light-excited photoelectron spectroscopy*), see § 6.

5.4. Scanning Auger-electron microscopy

So-called Auger electrons are emitted when an electron beam hitting the surface creates electron holes in a lower shell (K, L or M) and when this hole is filled by an electron from a higher shell. The excess energy of this jump is transferred to another electron which is emitted. An electron detector and an electron spectrometer are used to register the number of electrons as a function of kinetic energy which depends exclusively on the energy levels of the electron shells involved and thus is specific for the element. By scanning the electron beam, successive points and line profiles can be analysed and used to form an image by modulating the brightness of a cathode-ray tube by the signal intensity of the Auger electrons. *Auger maps* very similar in appearance as X-ray maps (see § 3.3.6) are obtained in this way, but, since emission of Auger electrons is limited to a depth of a few (< 5) nanometers (owing to the short mean free path of the low-energy Auger electrons) compared to a few micrometers of the characteristic X-rays, see fig. 7) with a much better depth resolution. Three-dimensional information can be obtained by consecutively sputtering off layers by argon (with sputtering rates in the order of micrometers per hour). Lateral resolution is also much below 1 μm, down to 10 nm in favourable cases depending mainly on the spot size of the beam. In experimental microscopes, 5 nm resolution has been obtained more recently (VENABLES and HEMBREE [1991]), and a resolution in the order of 1 nm seems possible (BAKER and CASTLE [1994]).

In principle, Auger electron microscopy could be performed in the scanning electron microscope with an Auger electron detector attached. However, since ultrahigh vacuum is usually not available in an SEM, the *scanning Auger microscopes* (SAM) have become instruments on their own. Recent reviews are available which give a full account of the basic principle and the electronic systems of the SAM (SIGNORELLI *et al.* [1988], RIVIÈRE

[1990], BAKER and CASTLE [1994]). Sample preparation is not critical as long as clean surfaces are produced. This is possible by sputtering off any adsorbed layers and other contaminants inside the UHV chamber or by fracturing the sample in the specimen chamber. An extension is *angular distribution Auger microscopy* (ADAM) in which the angles at which the Auger electrons leave the surface are measured and analysed to produce images providing information on the structure of thin films, single crystals and monolayers (FRANK and HUBBARD [1993]).

Applications of scanning Auger electron microscopy in physical metallurgy, microelectronics and corrosion are manifold, particularly in combination with Auger electron spectroscopy (AES) which is discussed in detail in Chapter 13 (see also § 6). The following case studies have been reviewed by BAKER and CASTLE [1994]: pit formation at inclusions in steel and pit propagation laterally under the passivating oxide film (BAKER [1993]), embrittlement by segregation (SEAH [1975]), in-situ fracture of spray-formed aluminium alloys (BAKER and TSAKIROPOULOS [1993]), wetting during vacuum brazing of a gold–nickel alloy (BROOKER *et al.* [1984]), among others. With some of the present problems relating to difficulties in imaging insulating materials or rough surfaces solved, SAM will increase its potential for materials characterization owing to its sensitivity for chemical composition, good lateral and extremely high depth resolution.

5.5. X-ray microscopy, topography and fluorescence

The concept of *X-ray microscopy* was developed early, at the beginning of this century (see reviews by COSSLETT and NIXON [1960] or DUKE and MICHETTE [1990]). The lack of sufficiently intense sources was overcome by the development of microfocus X-ray sources and plasma sources, and, particularly, by the availability of synchroton sources (FEDER *et al.* [1977], ASH [1980], DUKE [1981], KOCH [1983], MICHETTE [1990], BUCKLEY and RARBACK [1990]). Amplitude and phase contrast can be used for imaging (see, for example, RUDOLPH *et al.* [1990]). Scanning X-ray microscopes (BUCKLEY and RARBACK [1990]), devices for integration of the low-intensity projected or reflected X-rays by means of special TV cameras, microchannel plates or image intensifiers (CHENG *et al.* [1990]) and holographic techniques for three-dimensional imaging (HOWELLS *et al.* [1990]) have been developed. Owing to the high absorption of X-rays in crystalline materials, applications of *soft X-ray microscopy* has been more or less limited to biological objects (for details and references see, for example, DUKE and MICHETTE [1990]).

X-ray topography is a useful technique to study defects in single crystals such as dislocations, twins and stacking faults, subgrain boundaries, interfacial defects and strains, domain structures and other substructural entities (TANNER [1976], HÖCHE and BRÜMMER [1980], PANGBORN [1992]). The penetration depth of 5 μm and a lateral resolution of > 1 μm restricts application to relatively perfect crystals (defect density $< 10^5/cm^2$) but owing to its high selectivity for different types of defects and their location (subgrain boundaries, stacking faults, structure of ferromagnetic domains, dislocations) X-ray topography has become a standard technique for monitoring crystal quality, especially in the semiconductor industry. Imaging uses a Bragg reflection, the intensity of which varies with local variations of orientation, lattice parameter and

structural factor. It is not possible to magnify the image directly, owing to the lack of X-ray lenses. High-resolution film and photographic magnification has been widely used, typical exposure times ranging from 10 min to 2 h with a 1 kW X-ray source. More recently, digital image storage and accumulation have become available, providing better resolution and higher speed. X-ray topography has been applied mainly in the following principal areas of research (WEISSMAN *et al.* [1984], PANGBORN [1992]): development and quality control of electronic devices, kinetics of transformation, crystallisation and crystal growth, and deformation and fracture. Classical studies are the investigation of the internal magnetic structure of nontransparent ferromagnetic crystals (BOWMAN and BOOTH [1971]), and the investigation of damage due to ion implantation (TANNER [1976], DE RHODE and SMITS [1981]). Of particular interest are in-situ dynamic studies like the movement of dislocations in a silicon single crystal under external stress (CHANG *et al.* [1982]), the evolution of the plastic zone during crack propagation in a silicon crystal (MICHOT and GEORGE [1982]), and recrystallisation of deformed aluminium (GASTALDI and JOURDAN [1984]). With portable X-ray generators, advanced detection systems and computer analysis of results, X-ray topography is also useful in nondestructive evaluation of deformation states, microscopic flaws and cracks in parts outside the laboratory (PANGBORN [1992]).

Finally, *fluorescence microscopy* must be mentioned. If a fluorescing substance is irradiated by photons (X-rays or light of short wavelength, usually ultraviolet), some of the energy is re-emitted as light of longer wavelength which is typical for the substance. This effect is called *fluorescence* and is used in mineralogy for identification purposes and, after suitable staining with fluorescent substances, in biology (ROST [1981] and AXELROD *et al.* [1983]). Extremely small amounts of fluorescent nonmetallic phases can be detected. However, since very few phases in metallic alloys are fluorescent, this technique is rarely used for materials characterization.

5.6. Imaging by other types of spectroscopic information

There is a very large number of spectroscopic techniques in which photons, electrons or ions interact with surface near regions of the sample and some kind of radiation is analysed (see § 6). In addition to the microscopic techniques discussed up to now, there are presently a few more using the signals reflected or emitted from the surface to form images. Examples are nuclear magnetic resonance (NMR) microscopy (PETTEGREW [1990], MALLARD [1993], WACHTMAN [1993]) presently only used in biomedical applications, or imaging in secondary ion mass spectroscopy (SIMS). No doubt several other microscopic techniques will be developed in the future. It will be interesting to see to what extend these techniques will be useful in physical metallurgy.

6. *Topochemical techniques and surface spectroscopy*

Surface analytical methods have become available to materials researchers for nearly three decades. Two of the techniques which have become most popular, ESCA (electron spectroscopy for chemical analysis) and AES (Auger electron spectroscopy) were

described in the late sixties (SIEGBAHN *et al.* [1967], HARRIS [1968]). The number of additional spectroscopic techniques for chemical analysis increases at an even higher rate than that of microscopies and often combinations of topochemical and imaging methods are realized in a single instrument. Excellent reviews and books are available which present the progress and state of the art of the fundamentals and the applications of photon, electron and ion probes as well as electric field probes (for example CARLSON [1975], CZANDERNA [1975], IBACH [1977], MCINTYRE [1978], BRÜMMER *et al.* [1980], METALS HANDBOOK [1986], FISCHMEISTER [1988], SIBILIA [1988], WALLS [1992], RIVIÈRE [1990], BRIGGS and SEAH [1990, 1992], LIFSHIN [1990, 1994], ASM HANDBOOK [1992], CAHN and LIFSHIN [1993], among many others). Some of the techniques based on these physical principles are of special interest to materials scientists and engineers while others are used only occasionally. No attempt will be made here to give a full coverage of these methods. Rather, a recent overview by ORTNER and WILHARTITZ [1991] based on an earlier compilation by GRASSERBAUER [1989] is used here to present the relevant information on the large variety of methods in a systematic way (Tables 4 to 6). For references, the reader is referred to the books and reviews mentioned above. In addition, reviews on the application of surface analysis in materials science and technology, for example that by FISCHMEISTER [1988], and the relevant chapters in the encyclopedia edited by CAHN and LIFSHIN [1993] and in Volumes 2A and 2B of Materials Science and Technology edited by LIFSHIN [1992] and [1994], respectively, should be consulted. Combining the most effective techniques for characterizing the chemical and structural properties of materials (see fig. 15) will result in deeper insights into the nature of materials as well as improvements in the development of processing and properties of products needed for modern technologies.

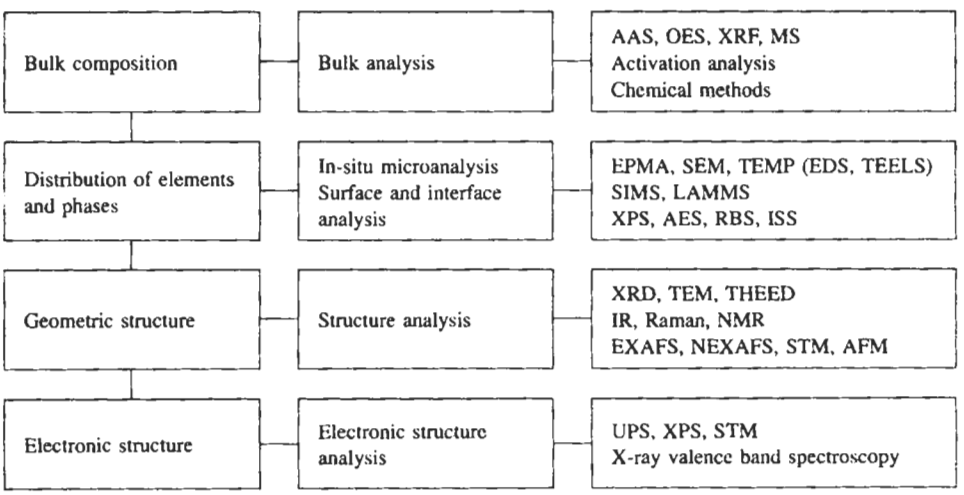


Fig. 15. Characterization of materials by chemical, spectroscopic and microscopic techniques (from ORTNER and WILHARTITZ [1991]). AAS = atomic absorption spectrometry, OES = optical emission spectrometry, XRF = X-ray fluorescence spectrometry, MS = mass spectrometry, XPD = X-ray diffraction, further acronyms see Tables 4 to 6).

References: p. 1016.

Table 4
Ion probe techniques (from ORTNER and WILHARTITZ [1991]).

Technique	Resolution		Process	Information content
	Lateral	Depth		
			Techniques based on sputtering	
SIMS (Secondary ion mass spectroscopy)	(a) Static [mm]	first monolayer	Sputtering with Ar^+, O_2^+, O^-, Ga^+, Cs^+ at 0.001 to 0.1 nm/s; TOF-SIMS (time-of-flight SIMS)	Surface studies in catalysis, gas-surface reactions, polymer analysis, only method with trace and ultratrace detection power at high lateral and depth resolution (detection limit from ppt to high ppb-levels)
	(b) Dynamic 500–1000 nm 50 nm (with Cs^+) 120 nm (with O)	1–5 AL (atomic layers)	As above, but with typically 1 nm/s Double focussing (mass spectrometry) with M/ΔM between 1000 and 100000	Isotopic specificity High resolution and high density depth profile possible Imaging resolution 1 μm Three-dimensional distribution analysis
LRI-SIMS (Laser resonance ionization SIMS)	1000 nm	1-5 AL	Very complex instrumentation, single element at a time method; selective ionization of one atomic species in sputtered atom cloud by double or triple photon absorption. TOF-MS (time-of-flight mass spectrometer)	Very high detection sensitivity at ppb- and sub-ppb-levels, e.g. 2 ppb ^{56}Fe in Si with consumation of 0.8 AL ≈ 40 Fe atoms sputtered!
LNRI-SIMS (Laser non-resonant ionization SIMS)	1000 nm	1–5 AL	Multiphoton ionization with intense pulsed UV-laser beam ($\geq 10^9$ W cm^{-2}) in sputtered atom cloud	As SIMS, but with detection limit presently still in μg/g-range
SNMS (Sputtered neutrals mass spectrometry)	1000 nm	1–5 AL	Secondary ionization of sputtered particles in HF-plasma above sample surface Quadruple MS-detection	As SIMS, but with detection limit presently still in μg/g-range
GDMS (Glow discharge mass spectrometry)	[mm]	10–100 nm	Ar-glow discharge for sputtering. Ionization in negative glow region Detection with double-focussing mass spectrometer, M/ΔM = 1000–10000	Advantages more on bulk-ultratrace analysis side than on topochemical analysis side Detection limit in the ppt to ppb range for pin samples only!

Technique	Resolution		Process	Information content
	Lateral	Depth		
GDOES (Glow discharge optical emission spectrometry)	[mm]	10–100 nm	As for GDMS, but with detection of atomic emission in UV/VIS-range	Detection limit in 1 ppm range Routine method for μm layer characterization in metallurgy
			Techniques based on scattering	
RBS (HEIS) (Rutherford back-scattering spectroscopy, or high energy ion scattering spectroscopy)	[mm]	2–3 nm Measurements up to 1000 nm	α or $^4He^+$-scattering E_0:1–2 MeV Measurement of energy distribution of back-scattered α or $^4He^+$-ions	Qualitative and quantitative elemental analysis; depth profiling by deconvolution of measured energy distribution with energy loss functions possible, however, generally poorer depth resolution than SIMS, but non-destructive depth profiling! Additional structural information by channeling and blocking phenomena: study of formation of epitaxial films
MEIS (Medium energy ion scattering spectroscopy)	[mm]	0.3–0.4 nm	200–600 keV	Same as above
ISS (LEIS) (Ion scattering spectroscopy, or low energy ion scattering spectroscopy)	[mm]	First monolayer analysis	500–600 eV He^+, Ne^+	Same as above. Adsorption results
ERD (Elastic recoil detection)	[mm]	First monolayer up to some μm	Elastic recoil detection of forward scattered light target atoms by bombardment with typically 3 MeV $^4He^{++}$ up to 30 MeV Si^+	Depth profiling for ^1H and ^2H with α, or H, C, N, O with Si^+

Table 5
Electron probe techniques (from ORTNER and WILHARTITZ [1991]).

Technique	Resolution		Process	Information content
	Lateral	Depth		
EPMA (EPXMA) (Electron probe microanalysis, or electron probe X-ray micro-analysis)	Primary electron-beam: 5–50 keV, diameter 1–100 nm Excited volume: 1–5μm^3 (pear shaped) 1–5 μm	1–5 μm	Scanning primary electron-beam excites characteristic X-rays in flat, polished samples. With conventional detectors the elements Na-U, with windowless proportional flow counters the elements B to U can be detected	Quantitative topochemical analysis: WDS (wavelength dispersive spectroscopy): 10–100 μg/g EDS (energy dispersive spectroscopy): 100–5000 μg/g Optimized for quantitative X-ray analysis after correction for absorption and enhancement (computer programmes) by use of standards possible with an accuracy of ~ 1% rel., standardless with an accuracy of 3–10% rel. as a function of composition
EELS (Electron energy loss spectrometry)	Resolution same as as for X-ray nano-analysis Detection limit: ~ 3–5% accuracy: ~ 10–20% relative Especially suitable for light elements due to lacking sensitivity of EDS. High electron yields with electron spectrometers beneath thin specimen. However low signal/noise ratios and high background due to multiple elastic scattering.		In connection with transmission electron microscope Imaging possible by beam scanning (STEM) or stigmatic electron spectrometer	Qualitative and semi-qualitative information of elemental composition Position and structure of absorption edges of valence band or near valence band core levels contain *bonding information* Fine structure evaluation in combination with high spectral resolution (0.1 eV at 100 keV) allows to gain *stoichiometric information* for extremely small domains Sensitive detection of structural changes in nano domains possible by study of position and width of plasmon peaks
HREELS (High resolution electron energy loss spectrometry)	poor (mm to cm)	monolayers	Monoenergetic beam focussed onto the surface of a solid, measurement of reflected electrons Two electron spectrometers necessary (one for monochromatization of primary electron beam, typical E_o = 500 meV, and one for energy analysis of reflected electrons)	Useful for studies of adsorbates on solid surfaces (surface coverage with adsorbates, information on bonding geometry, studies of catalytic phenomena, thin film characterization)

Technique	Resolution		Process	Information content
	Lateral	Depth		
AES (Auger electron spectrometry)	Primary electron beam: 1–10 keV poor (low signal to noise-ratio, very high continuum electron background due to multiple scattering requires registration of 1st derivative of spectrum, but very high surface sensitivity)	0.5–5 nm	Measurement of kinetic energy of emitted Auger-electrons UHV of 10^{-10} mbar mandatory In situ sample preparation: fracturing, annealing; thin film deposition	Micro surface analysis (elemental identification, quantative elemental analysis, direct compound identification by evaluation of chemical shifts and line shapes possible, depth profiling in combination with Ar^+-sputtering) Relative detection limit depends on beam diameter and current density, e.g. 10–20 mass% at 100 nm and 1nA, 1–2 mass% at 1 μm and 1 nA or 100 nm and 100 nA, or 0.1–0.2 mass% at 1μm and 100 nA Imaging resolution 200 nm
THEED (Transmission high energy electron diffraction)	$\lambda = 0.1$ nm at 150 eV $\lambda = 0.01$ nm at 15 keV (High energy of primary electron beam yields high resolution of imaging) Minimum phase size for structure analysis of precipitates: ~ 20 nm		Electron diffraction (coherent interference effects)	TEM diffraction patterns (point patterns for monocrystalline specimens, ring patterns for polycrystalline specimens, stripe formed patterns for large grained or textured specimens, and diffuse circular halos for amorphous materials) useful for determination of lattice constants and identification of phases
LEED (Low energy electron diffraction)	Primary electron beam 1–5 keV poor	monolayers	Diffraction at solid surfaces Measurement of reflected electrons of a mono-energetic PE-beam on a fluorescent screen	LEED patterns reveal the periodicity of atoms of the surface and the overall symmetry, but not the individual atomic position (periodicity is needed for reflection)
RHEED (Reflected high energy electron diffraction)	Primary electron beam > 10 keV poor	thin films		Useful for determination of crystalline order of the surface of a single crystal or study of adsorption and segregation processes on surfaces or overlayers
XNA (X-ray nano-analysis)	5–30 nm at beam diameter 3nm and sample thickness 100-20 nm Absolute detection limit: 10^{-20} g Elemental distribution imaging: lateral resolution ca. 1 nm (signal enhancement by digital image processing)		In connection with TEM Large area detectors (EDS) in small distance from specimen Principal disadvantages are low collection yields (10^{-2} to 10^{-4}) and low X-ray yield for light elements (B, C, N, O, ...)	Important for the identification of precipitates at grain boundaries or interfaces and for sub-μm phases and inclusions

Table 6
Photon probe techniques (from ORTNER and WILHARTITZ [1991]).

Technique	Resolution		Process	Information content
	Lateral	Depth		
LRMA (Laser Raman microanalysis)	1 μm	0.1 μm	Inelastic photon scattering (Stokes and anti-Stokes)	Molecular and structural microprobe; complementary info to IR about functional group vibrations, evaluation of polarization effects
SERS (Surface enhanced Raman spectroscopy)	1 μm	0.1 μm	As above	As above, but enhancement up to 10^6; limited to Ag, Cu, Au as substrates
NIR-FT-RS (Near infrared Fourier transform Raman spectrometry)	1 μm	0.1 μm	As above, but near IR	As above, but less fluorescence interference
FT-IR-M (Fourier transform infrared microscopy)	10 μm	1–10 μm	IR-absorption with interferometric recording of spectra	IR (0.75–1000 μm) absorption by excitation of vibrational and rotational states Molecular and functional group information
ATR-IR-S (Attenuated total reflection infrared spectrometry)	Several mm	1–10 μm	Internal IR-reflection	As above: SA of polymers, catalysis, electro-chemical reactions. By use of polarized light: information on orientation of adsorbates on surfaces
IRRAS (Infrared reflectance absorption spectrometry)	mm	1–10 μm	Angle resolved IR-reflectance measurements	Determination of complex optical properties and identification of chemisorbed species
EXAFS (Extended X-ray absorption fine structure spectrometry)	ca. 20 μm	μm	Absorption of monochromatic X-rays (synchrotron radiation), several keV	"Local atomic structure probe" Measurement of atomic distances between absorbing atom and neighbours (range 0.6 nm)
SEXAFS (Surface EXAFS)	ca. 20 μm	1 nm	As above, but up to 1 keV	Accurate determinations of bond lengths of adsorbates, surface molecular studies. The polarization of synchrotron radiation allows to probe atomic distances in specific directions
NEXAFS (Near edge X-ray absorption fine structure spectrometry)	ca. 20 μm	μm	As above, but keV	Determination of valence states of very fine precipitates, not measurable by TEM: studies of crystal growth, nucleation, formation of interfaces

Technique	Resolution		Process	Information content
	Lateral	Depth		
ESCA Electron spectrometry for chemical analysis			Two techniques, XPS and UPS, see below	
XPS (X-ray photoelectron spectroscopy)	100 μm (crystal focussing) $\leq$ 10 μm (ESCASCOPE of VG, FT-lens)	Typically 5–10 atomic layers	Absorption of monochromatic X-rays and measurement of the kinetic energy of ejected photo-electrons	Qualitative and quantitative elemental analysis of surfaces ($Z > 1$; sensitivity limit ca. 0.1 mass% but 0.01 monolayers! plus information on chemical bonding (by chemical shifts of binding energies) and binding energies. Depth profiles with sputtering installation possible Imaging resolution < 10 μm
UPS (Ultraviolet photoelectron spectroscopy)	In the order of mm	Several atomic layers	Absorption of monochromatic UV-radiation and measurement of the kinetic energy of ejected photoelectrons. (Excitation with noble gas discharge lamps or synchrotron radiation)	Study of electronic structures (of the valence band of solids or of bonding electrons in gas-molecules (even vibrational excited states can be resolved) Extremely good vaccuum ($p \leq 10^{-11}$ mbar) necessary. Study of chemisorption phenomena, adsorption, desorption (esp. with respect to catalysis)
ARUPS (Angle resolved ultraviolet photoelectron spectroscopy)	In the order of mm	Single first mono-layers	Same as for UPS, structural sensitivity enhancement by use of polarized radiation	Measurement of valence band emission spectra for "density of state"-calculations; very weak adsorption phenomena, e.g., of noble gases. Variation of detector angles allows to exploit directional effects in photoelectron-emission
LAMMS (Laser micromass spectroscopy)	NTLD: 1–2 μm	Several monolayers	Non-thermal laser desorption (NTLD $\leq 10^8$ W/cm^2)	Molecular information from adsorbed or chemisorbed species or from surface contaminants
	TVLE: 1–2 μm	In the order of μm	Thermal volume laser evaporation (TVLE $\geq 10^8$ W/cm^2) Detection by TOF-MS (time of flight mass spectroscopy)	Microanalysis for elements, some bonding and molecular information from fragmentation patterns. Detection limit in the order of 1–10 μg/g

7. *Quantitative interpretation of microstructural geometry*

Quantitative methods for characterizing the geometry of microstructures are an important means for understanding the relationship between production parameters and behaviour of technical materials in application as well as for providing the experimental basis for theoretical studies of microstructural genesis and its effect on mechanical and physical properties. The important elements of *quantitative microstructural analysis* (also termed *quantitative metallography*) are *image analysis*, i.e., the quantification of features visible in metallographic cross-sections or transmission images; *stereology*, i.e., estimation of three-dimensional geometry from two-dimensional image data; and the *interpretation* of the three-dimensional parameters.

Generally, more than one single geometric aspect is of relevance in characterizing the microstructural geometry of multiphase materials, and a combination of parameters will be needed to give a full description. Such combinations have been proposed for studying the evolution of microstructure or for empirical correlations with mechanical properties. A large number of techniques have been developed and employed for characterizing microstructural geometry. Table 7 shows an overview over important analysis methods. This section aims to give a short introduction to techniques of image analysis and stereology, focussing on applications to problems relevant in materials science.

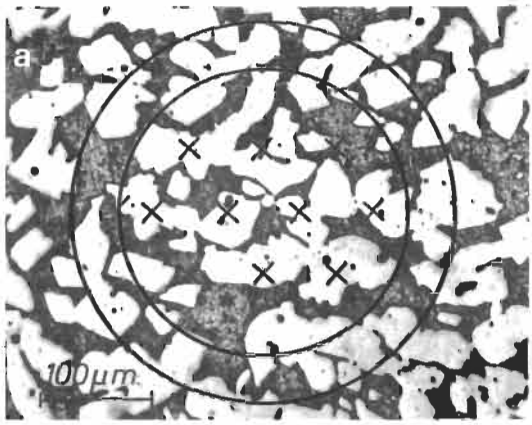
7.1. Image analysis

Simple comparison of microstructure with standard test charts is extensively used in quality control of materials (for example, ASTM standards for determination of grain size or of slag inclusions). The accuracy of these comparison methods is rather limited compared with counting techniques (EXNER [1969b, 1994] and EXNER and HOUGARDY [1988]). These semi-quantitative methods should not be used for research work and will not be discussed any further here.

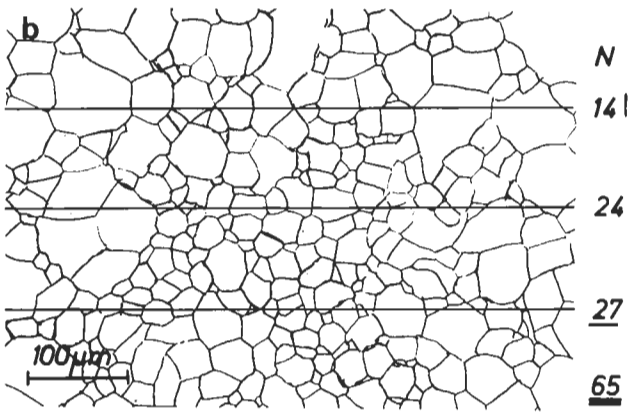
When scanning an image for collecting data, counting with point and line grids, measurements along straight lines, or area measurements can be used. Each of these strategies has its merits, according to the purpose of the investigation (parameters to be determined) and available instrumentation (manual, semi-automatic or fully automatic).

The manual acquisition of image data is carried out directly at the microscope, using inserts in the ocular or ground plates with arrays of points or lines, or with micrographs and suitable grids or rulers (GAHM [1971], EXNER and HOUGARDY [1988] or EXNER [1994]). Figure 16a shows a typical point grid superimposed over the microstructure of a two-phase material. The number of points falling within the black phase and the number of intersections between the outer circle and the perimeter lines of the black phase are counted (yielding, as discussed below, estimates for volume fraction of the black phase and its area of interface with the white phase). Figure 16b shows linear analysis of the grain size in a single-phase material. A number of fields must be evaluated in this way in order to get statistically significant numbers. Simple electronic counters and step stages are useful means to speed up these counting procedures. More complex data (size distribution and elongation of planar features) can be determined by

(a)



(b)



(c)

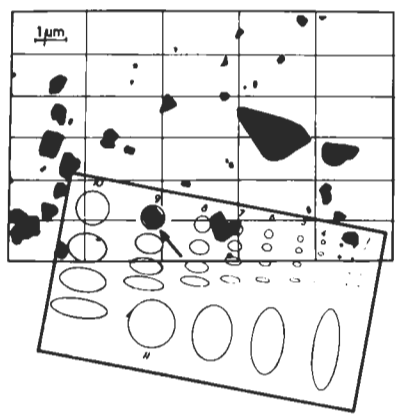


Fig. 16. Typical grid-, line- and graticule-counting procedures in manual image analysis: (a) point-counting for volume fraction analysis (circles can be used for interface density measurement, see text); (b) line-counting for determination of mean linear grain size of a single-phase material (numbers on the right margin show number of grain boundaries intersected); (c) graticule-counting for size distribution measurements (area of circles and ellipses of varying aspect ratio increase by a factor of two in each higher class).

Table 7
Techniques for the geometric characterization of three-dimensional microstructures and two-dimensional images ("fingerprinting").

	3D Description		2D Description
1.	Stereology Global parameters (volume-, surface- and line densities, orientation) Distributions (size, shape)	1.	Image comparison Standard charts, graticules
2.	Stereometry Nonplanar surfaces (fracture surfaces, deep-etched interfaces), Transparent materials (3D microscopy, TEM foils)	2.	Simple geometric parameters Perimeter, area, shape factor, size and shape distribution
3.	Serial sectioning Topological parameters (number density, genus) Shape and arrangement	3.	Mathematical morphology Manipulation of digital images (opening, closing, skeletonizing) Morphological parameters
4.	Model structures (simple shapes like spheres, cubes) Regular lattices, statistical arrangements	4.	Complex special techniques Fourier analysis (periodicity) Fractal analysis (selfsimilarity) Pair correlation functions, graph theory (arrangement)
5.	Computer generated structures (dendritic patterns, fractal structures etc.) Point processes and mosaiques	5.	Geometric abstraction Circles, hexagon, planar lattices (analytical modelling) Poisson point process, Voronoi mosaique (computer modelling)

simple graticules, an example of which is shown in fig. 16c. Manual methods for microstructural analysis have been reviewed in detail by FISCHMEISTER [1965, 1981], GAHN [1971], EXNER [1972a, 1994], UNDERWOOD [1985] or EXNER and HOUGARDY [1988]. Exact procedures for practical assessment of the most important stereological parameters (volume fraction, interface density, size and distance, orientation, elongation, proximity and contiguity, see below) have been worked out recently by a committee of the German Society for Materials including estimates of experimental accuracy (EXNER [1993]).

The first step towards automation was realized in lineal scanners which were popular during the sixties and seventies (EXNER [1972a] and GAHM [1975]). More recently, semi-automatic instruments have been provided with digitizers as input device to a calculator or computer (see, for example, HILLJE and REDMANN [1976], EXNER [1978b] and EXNER and HOUGARDY [1988]). Digitizers are tablets by which the exact coordinates of points touched with a special pencil are registered at a high rate (up to 10^4 per second) with high precision (0.1 mm). The operator indicates the intersections of the scanning line

with the perimeter or outlines the perimeters of the microstructural features using his ability to differentiate between features of different kind (composition, structure) on the basis of grey or colour shading and of shape. Thus, the operator keeps steady track of what is measured and is able to exclude artefacts by simple judgement. Modern instruments compute all relevant microstructural parameters from these data (see, for example, RUSS [1995]). Disadvantages are the high time-expenditure and concentration required, in addition to some danger of introducing human bias. Nevertheless, this type of instruments is often the best choice for practical work due to their reliability and versatility.

Fully automatic instruments are available with a wide efficiency range (FISCHMEISTER [1981] and EXNER and HOUGARDY [1988]). Most of them use videocamera (CCD) type scanning modules with electronic phase detection. Simple low-price instruments allow determination of a limited number of parameters for high-contrast situations. High priced instruments, usually modular devices with fast hardware modules or, more recently, flexible software modules and full grey-level image-storage with up to 1024×1024 picture-point lateral resolution and 256 steps of grey-level resolution combine high speed, versatility and convenience for any type of automatic image data acquisition. Scaling, shading correction, scanning-table steering and focussing are all carried out automatically, and all kinds of image manipulation (filtering, edge contrast enhancement, halo correction, etc. and image modification techniques based on mathematical morphology the principles of which are outlined in § 7.3 below, etc.) can be implemented if required. Figure 17a shows the grey-level histogram and fig. 17b demonstrates the effect of a special filter (pseudoplast filter) on the appearance of a light micrograph of a three-phase alloy. Phase discrimination (either by dedicated hardware modules or by software algorithms), parameter selection and evaluation, statistical treatments and many other routines can be activated by push-button commands or by a simple computer program designed individually for any specific problem. As an example, fig. 17c shows the appearance of a microstructural evaluation on the monitor or screen of an automatic image analyzer. "Interactive" devices — usually using a mouse or a light-pen to be used on the display monitor — allow the operator to use his judgment for excluding artefacts or adding missing details (e.g., parts of a grain-boundary network).

In principle, any image-forming instrument can be used as input device; only light microscopes and scanning electron microscopes have been used in on-line configurations. Transmission electron micrographs and other printed images have been evaluated only occasionally for solving problems in materials science and technology. Software for automatic image analysis is now offered with most highly computerized scanning instruments mentioned in this chapter, particularly with energy dispersive X-ray analysers (electron microprobe). The combination of image analysis and energy-dispersive X-ray analysis was used for determining geometric parameters and composition of individual particles simultaneously (EKELUND and HERTZMANN [1981]).

Instrumentation for image analysis has developed rapidly over the past two decades (for historical and recent reviews see FISCHMEISTER [1965], GAHM [1971], EXNER [1972a], HOUGARDY [1976], FISCHMEISTER [1981], EXNER and HOUGARDY [1988], or EXNER [1994]) and is still advancing at a rapid rate. In this context, the dramatic

References: p. 1016.

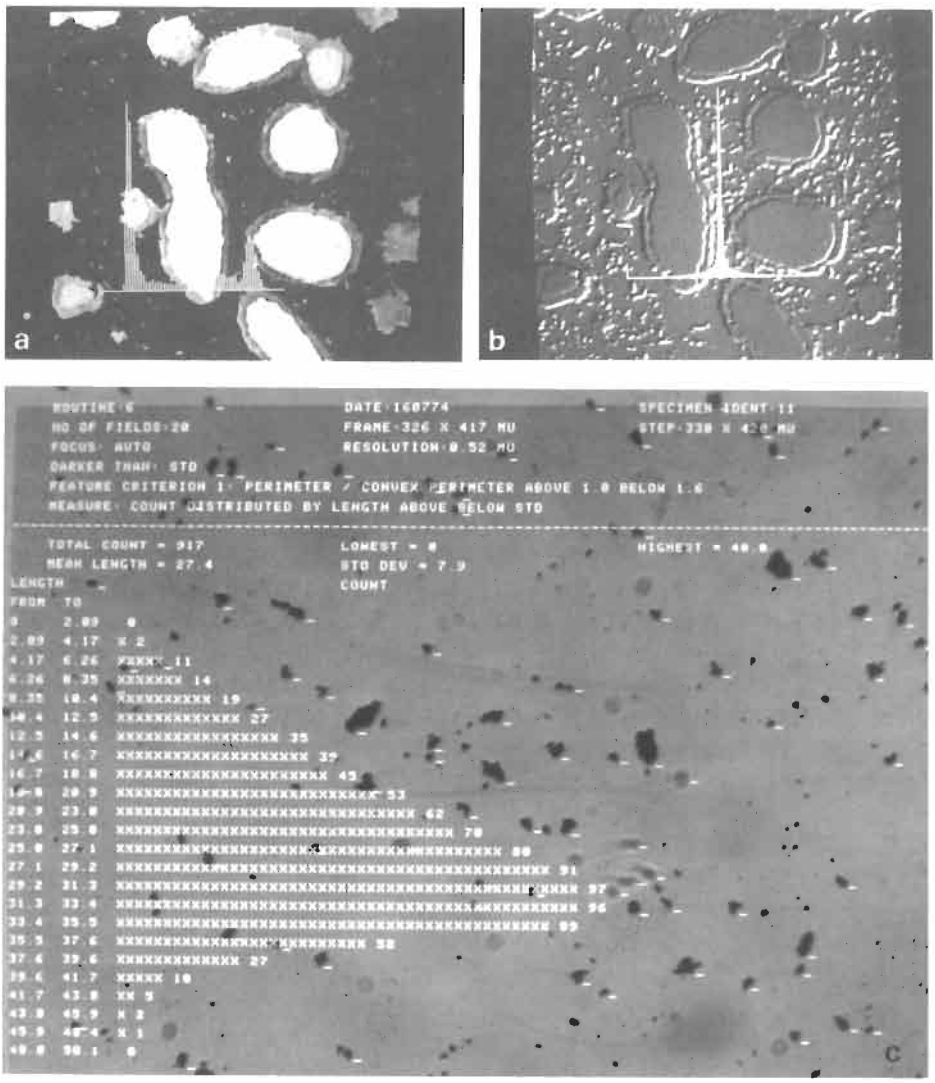


Fig. 17. Typical procedures of television-equipped fully automatic image analyzers: (a) Microscopic image of a three-phase copper-cadmium alloy with superimposed grey-level distribution; (b) Pseudoplast filter producing a plastic appearance by showing gradient variation only; (c) Size-distribution analysis of nonmetallic inclusions in a dirty steel with superimposed counting flags indicating proper detection and area distribution histogram.

developments in digital image processing must be mentioned, which not only are useful for quantitative image analysis but also for contrast enhancement and image restoration (RUSS [1990], JAHNE [1993] or TALBOT [1994]). The state of the art has now reached a level which allows only minor improvements with respect to practical applications.

Though recent developments (e.g., processing of real colours and grey levels) may widen the scope to the treatment of even more complex problems in image analysis, the present rapid decrease of price levels at a given performance is certainly the best reason for a much more general utilisation of automatic image analysers.

7.2. Planar characteristics and stereology

Image analysis yields a number of geometric characteristics describing the geometry of planar images of cross-sections or thin-film projections, the most important of which are listed in table 7. They usually can be qualitatively interpreted in terms of the three-dimensional microstructures from which they originate. Quantitative three-dimensional characteristics are obtained by a body of mathematical methods based on integral geometry and statistics which is called stereology. There is an enormous number of original papers and formulae spread over the scientific literature of fields as far apart as mathematics, geology, materials science and biology. Useful monographs and standard textbooks (DE HOFF and RHINES [1968], UNDERWOOD [1970], SALTYKOV [1974], WEIBEL [1979, 1980], RHINES [1986], RUSS [1986, 1990, 1992], EXNER and HOUGARDY [1988] or COSTER and CHERMANT [1989]) review the literature and should be consulted whenever a quantitative evaluation of spatial microstructures is attempted, in order to avoid misinterpretations and redundant derivations. A few of the more important aspects of microstructural geometry (table 8), their experimental evaluation from surface images and some applications in physical metallurgy are discussed in the following sections.

Standard nomenclature for the stereological terms as outlined in the classical textbooks by DE HOFF and RHINES [1968], UNDERWOOD [1970] and WEIBEL [1980], see also EXNER [1994] and as standardized by the International Society for Stereology (EXNER [1987]) is used in the following and is recommended for general use.

7.2.1. Volume-fraction analysis

Simple plausibility proofs as well as sophisticated mathematical treatments show that, for a homogeneously dispersed phase, area fraction, lineal fraction and point fraction are statistically unbiased estimates for the volume fraction, i.e., in standard nomenclature (UNDERWOOD [1970], WEIBEL [1980] and EXNER and HOUGARDY [1988]):

$$V_V = A_A = L_L = P_P. \tag{4}$$

The experimental error of this estimate decreases with an increasing number of points counted or lines and areas measured and, at a 5% level of error probability, for point-counting using a wide-spaced regular point grid, is approximately (HILLIARD [1976]):

$$\pm\Delta V = 2\left[P_P(1 - P_P)/P\right]^{1/2}, \tag{5}$$

where P is the total number of points of the test grid. A more stringent evaluation of errors includes the variation of area fractions between different test fields (NICHOLSON [1978]).

Volume fraction has been the parameter most frequently used in quality control of

Table 8
Some important geometric characteristics of planar images and three-dimensional microstructural features [a].

Planar characteristic [b]	Example	Corresponding three-dimensional characteristic [c]	Microstructural example
Area fraction, A_A	Particle intersects	*Volume fraction, V_V	Amounts of phases in multi-phase material
Line density, L_A	Intersect perimeters	*Surface fraction, S_V	Grain-boundary or phase-interface area
Point density, P_A	Triple points in intersected grain-boundary network	*Line density, L_V*	Length of dislocation lines or grain-boundary triple lines
Number density, N_A	Number of intersected particles	Numerical density, N_V	Number of particles
Distribution by area $f(a)$	Intersect area distribution	Distribution by number and linear size $f(l)$	Number of spherical particles as a function of diameter
Distribution by length $f(l)$	Chord length distribution	Distribution by volume $f(v)$	Number of grains as a function of volume
Mean size	Mean intersect area or mean linear intercept	*Mean linear size $\bar{L}$ Mean volume, V	Mean linear grain size Mean particle volume
Distances	Nearest-neighbour distance	*Mean linear distance D Nearest-neighbor distance	Surface-to-surface distance of particles in one direction Distance of particles in dispersion-hardened materials
Shape characteristics	Area/perimeter ratio, length/width ratio	Shape parameters	Shape changes of particles or pores in processing
Planar curvature	Local and mean curvature	*Mean curvature	Curvature of pore-solid interfaces
Arrangement in plane	Random, regular, clustered	Arrangement in space Connectivity *Contiguity	Clustering of carbides in high-speed steel Number of closed loops in the pore space of a sintered material Amount of surface shared with other grains of the same phase in a multiphase material
Orientation	Preferred orientation of intersection lines	*Orientation of grain boundaries and interfaces	Anisotropy of pore/solid interface, grain boundary area in rolled sheets or drawn wires

[a] Three-dimensional parameters marked with an asterisk can be calculated from data obtained with planar images without any assumptions but randomness of sampling. All other three-dimensional characteristics are restricted to model geometries, to known shapes or known size distribution or both, or can only be assessed by three-dimensional measurements (stereometry, serial sectioning or chemical desintegration).
[b] Subscript A indicates that quantity is per unit area.
[c] Subscript V indicates that quantity is per unit volume.

materials and has found many useful applications in materials research. It is obviously closely related to composition and exerts a decisive effect on the properties of multiphase materials. In equilibrium phase diagrams, the lever rule allows phase boundaries to be determined to a high accuracy with a minimum number of samples. For example, LINDÉN [1972] has redetermined the two-phase region in the copper–aluminium system, and ALDINGER [1969] has shown excellent agreement of results obtained by classical and by image-analyzing methods in the four-phase region of the Ag–Cu–Cd–Zn diagram. Modern phase-diagram studies employ computer calculations to a high degree which are facilitated if accurate quantitative data are available.

A typical field for volume-fraction analysis is the determination of nonmetallic inclusions, of carbides, and of the time–temperature dependence of phase transformations in steels. Figure 18 shows a typical isothermal time–temperature–transformation (TTT) diagram for a chrome–molybdenum steel which makes it possible to read not only the nature of phases obtained after a certain holding time at constant temperature but also the fraction of austenite transformed (LAFOND and MOLIEXE [1975]). Studies of transformation kinetics of discontinuous precipitation or recrystallization are further examples. A very interesting aspect of volume-fraction analysis is the possibility of determining local composition, which is straightforward if the equilibrium compositions of the individual phases present in the microstructure are known. Thus, segregation and inhomogeneity phenomena and other systematic compositional variations can be quantitatively determined, sometimes with as good or better accuracy than by microchemical methods.

Good examples can be found in the investigation of surface and heat treatment of steels, e.g., the analysis of the carbide distribution in the surface of a hardened steel part

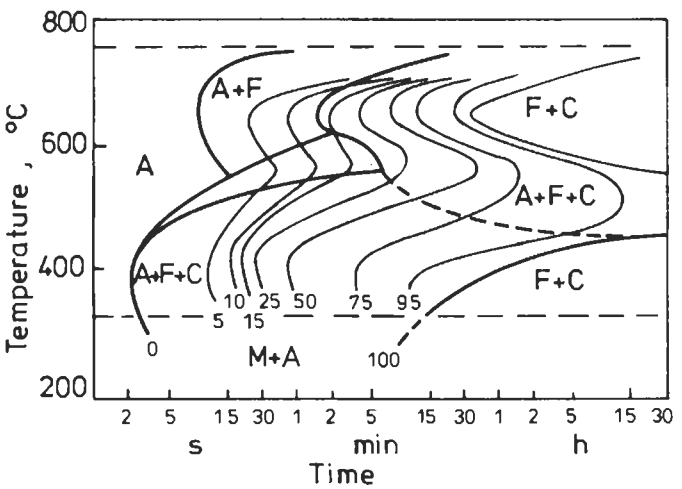


Fig. 18. Isothermal time temperature transformation diagram for a chromium-molybdenum steel. (A = austenite, F = ferrite, C = carbide, M = martensite). The heavy lines give start and end of a reaction, the numbers at the thin curves give the volume fraction of austenite transformed. (After LAFOND and MOLIEXE [1975].)

References: p. 1016.

(FISCHMEISTER [1972]), or the scanning electron microscopy study of the influence of heat treatment on the composition of the matrix in high-speed steels (FISCHMEISTER *et al.* [1986], KARAGÖZ and FISCHMEISTER [1987]).

7.2.2. Interface density

Dispersity of a microstructural component (isolated particles as well as complex-shaped continuous or semi-continuous phases) can best be quantitatively described with a single number by measuring the density of its interface area per unit volume. Stereology (SALTYKOV [1974], UNDERWOOD [1970] and WEIBEL [1980]) tells us that:

$$S_V = 2P_L, \tag{6}$$

where S_V is the interface area per unit volume of the material and P_L is the number of intersections (per unit length) between a scanning line and the intersections of a plane cross-section with the interface as seen in a metallographic image. If we relate the interface of a microstructural component (phase α) to the volume of this particular component, we get the *specific interface density* of the α-phase:

$$S_L^\alpha = S_V/V_V = 4N_\alpha/L_\alpha = 2P_L/P_P, \tag{7}$$

where L_α is the length of the line inside the α-phase ($L_\alpha = LV_V$, L is the total length of the scanning line) and N_α is the number of α-regions transversed by the scanning line. Of course, the density of grain-boundary area can be determined in the same way.

An analysis of statistical errors shows (HILLIARD [1976]) that the standard error, ΔS_V, is approximately

$$\Delta S_V = kS_V/\sqrt{P}, \tag{8}$$

where P is the total number of intersections and k is approximately 1.2 for contiguous grains and 4 for randomly dispersed particles. The range $S_V \pm \Delta S_V$ corresponds to 5% error probability (95% significance level).

Prominent examples of application of surface-density measurements are found in powder metallurgy, where the pore–solid interface provides the driving force for sintering processes (ch. 33), in the study of coarsening processes (grain growth and particle growth, ch. 28) and in finding correlations between microstructure and properties. Figures 19a and b show two simple examples of the latter, where linear relationships are observed between the hardness of various metals and their grain-boundary density (SALTYKOV [1974], RHINES [1976, 1986] and ONDRACEK [1978]) and between the coercive force and the interface per unit volume of magnetic cobalt phase in WC–Co hard alloys used in cutting and wear applications (FISCHMEISTER and EXNER [1966]).

The concept of surface density is less popular for characterizing microstructural geometry in physical metallurgy than might have been expected considering its versatility combined with extreme ease of evaluation. While for the separate determination of number and size of particles simplifying assumptions about shape are necessary (see below), the only requirement for obtaining significant numbers for S_V is statistical sampling, while shape effects as well as variations of the size of the individual features

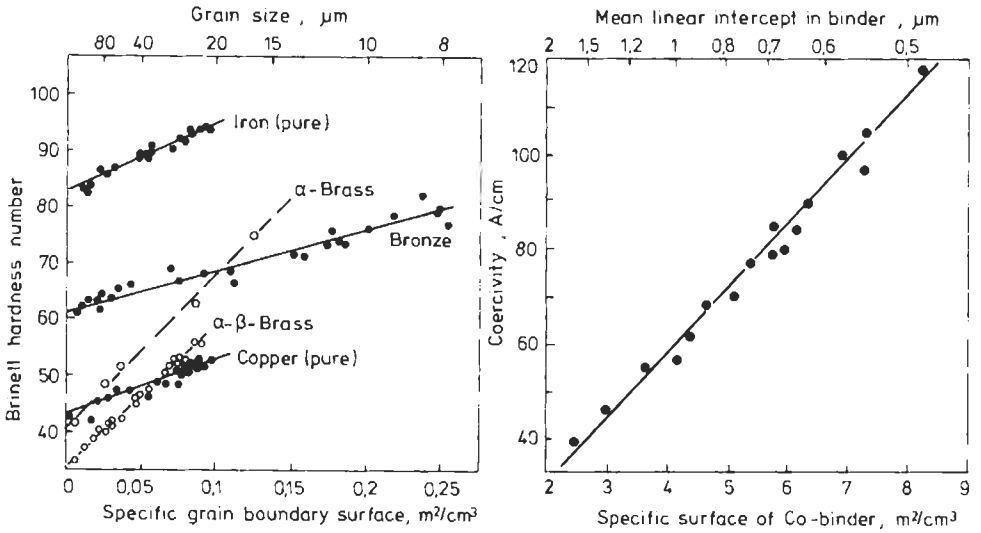


Fig. 19. Simple relationships between properties and microstructural geometry: (a) hardness of some metals as a function of grain-boundary density (data from SALTYKOV [1974] and RHINES [1976]); (b) coercive force of the cobalt phase in tungsten carbide–cobalt hardmetals as a function of interface density (after FISCHMEISTER and EXNER [1966]).

have no influence. Surface densities can also be used for describing the contiguity, i.e., the degree to which grains of a phase are connected in space or the degree of preferred orientation (see below). In modelling microstructure–property relationships or processes of microstructural genesis during casting, heat-treatment, deformation, etc., interface densities are most useful quantities for averaging the complex geometric appearance of phases in most technical materials. The use of interface densities (instead of size) was suggested in theoretical modelling of coarsening by EXNER and FISCHMEISTER [1966a] and by GLICKSMAN *et al.* [1992], of sintering by AIGELTINGER and EXNER [1977], EXNER and GIESS [1989] and LIU and PATTERSON [1993], and of recrystallisation by DE HOFF [1972, 1989]. The latter author developed the framework for stereological theories of kinetic and dynamic evolution of microstructures (DE HOFF [1972]).

7.2.3. Size and distance

In contrast to volume fraction and surface density or specific surface, three-dimensional size and separation (distance) are not related to two-dimensional data in a simple manner. Only if the shape of an object is known can its size be defined by a single parameter. The assessment of nearest-neighbour distances requires knowledge of the spatial arrangement of objects, which is often not available from planar images. Thus, models are needed to estimate size and distance and their distribution.

The most useful model relates to specific surface and surface density. The size and distance can be described by the average linear distances between the surface points inside (mean linear size) and outside (mean linear distance) the volume covered by a

phase α. For multiphase materials we get the *mean linear intercept* (*mean linear particle size* for a dispersed phase):

$$\bar{L}_\alpha = L_\alpha / N_\alpha = 4 \frac{V_V}{S_V}, \tag{9}$$

and the *mean linear distance* between the surface elements of phase α:

$$D_\alpha = \left(L - L_\alpha\right) / N_\alpha = 4 \frac{1 - V_V}{S_V}. \tag{10}$$

For space-filling grains ($V_V = 1$), the *mean linear grain size* (mean linear grain intercept, Heyn's grain size) is:

$$\bar{L} = 1 / N_L = 2 S_V. \tag{11}$$

This concept makes it possible to characterize size and distance independently of shape, arrangement, or size distribution. *Mean intersect area* (Jeffries' grain size) $\bar{A}$ is an indirect measure related to the length of grain edges per unit volume (FISCHMEISTER and ARZT [1981]). $\bar{A}$ has no concise meaning with respect to the size of isolated particles if shape is not uniform (and simple). Rather, as can be shown using derivations by DE HOFF [1967, 1980] and CAHN [1967], see also EXNER [1994], there is a direct correlation between the mean intersect area and the integral mean curvature C_V

$$C_V = 2\pi N_A = 2\pi V_V / \bar{A} \tag{12}$$

where N_A is the number of intersect area per unit area of the cross section. C_V thus is another important global stereological quantity useful for describing microstructural changes (see fig. 21 below, and EXNER [1994] for references).

There are numerous other measures for characterising mean size and/or distance. For example, a measure of size frequently used is $\bar{R} = 3\bar{L}/4$, derived by simply relating the radius of a sphere to its mean linear intercept. It is obvious that the result obtained does not give much more useful information than do $\bar{L}$ or S_V since $\bar{R}$ is simply the radius of equally-sized spheres with the same specific surface and volume fraction (but not average volume or number) as the (irregular) particles or grains. For space-filling polyhedra, similar size measures can be derived (THOMPSON [1972]) which, however, are similarly restricted in describing the geometry of real grain structures. Other size parameters discussed in literature (for references and derivations see RHINES and DE HOFF [1968], THOMPSON [1972], UNDERWOOD [1970], WEIBEL [1980], FISCHMEISTER and ARZT [1981], COSTER and CHERMANT [1989] or OHSER and LORZ [1994]) rely on well-defined geometries as well. The assumptions made are usually clearly stated in the original papers but often neglected by users of the final results, which may lead to erroneous conclusions.

Measures of distance other than that defined in eq. (10) usually are based on model assumptions concerning the spatial arrangement. For example, nearest-neighbour distances in three dimensions can be calculated from two-dimensional counts if the shape and size of all particles is identical and if the volume fraction is not too high (see, for

example, BANSAL and ARDELL [1972]) or if parallel platelets are considered (ROOSZ *et al.* [1980] and TÖRRÖNEN [1980]). However, if the nearest-neighbour distances of irregularly shaped and distributed particles is to be estimated, no straightforward methods are available. As discussed recently by OHSER and LORZ [1994], the transformation of pair correlation functions (see e.g., STOYAN *et al.* [1987]) from two to three dimensions seems to be a possibility which, however, may prove too complicated for practical applications. Thus, computer simulations of sections through three-dimensional arrangements (for example based on the Poisson–Voronoi mosaique (see OHSER and LORZ [1994]) may be a viable alternative to serial sectioning.

The situation becomes even more complex when information on the three-dimensional size distributions of grains or particles is required. Direct stereological procedures to derive the 3D distributions from distributions of linear intercepts or from intercept areas are available for idealised models in which the assumption is made that all features have constant shape and that this shape is extremely simple (sphere, cube, ellipsoid). Since this situation is rare or at least not typical in real microstructures, these models are not useful in most practical cases. The same is true for other analytical approaches.

For example, KING [1982] has derived a method for assessing three-dimensional size distributions (cumulative particle mesh size distribution) for irregular particles from planar data, i.e., area distribution of sections or projections, or from linear intercept distributions. No *a priori* assumptions on regular geometric shape, convexity, congruency or particular functional form of the size distributions are made in this widely neglected paper. However, a shape-dependent constant must be known or determined experimentally for a uniformly sized sample. Due to this complication, this or other unfolding procedures have not found much interest outside the stereological literature. Nevertheless, solutions to the so-called "tomato salad problem" have been published continuously over the years. A recent approach concerns polyhedric grains and may prove useful under some conditions for grain size characterization in single-phase materials (ZHAO and CHEN [1993], OHSER and MÜCKLICH [1995]). As a general rule, rather than standard text books or original papers, special reviews (e.g., EXNER [1972b], SIMPSON and STANDISH [1977], CRUZ-ORIVE [1978], FISCHMEISTER and ARZT [1981], EXNER [1994] or OSHER and LORZ [1994]) should be consulted for representative conversion procedures, their applicability and restrictions, and the related derivations should be followed up before deriving (or even worse, publishing) more of the already abundant solutions.

In reality, microstructural elements vary not only in size but also in shape. The shape variation may be determined by growth processes and is often limited by crystallographic conditions. In such cases, simplifying assumptions may be justified and special solutions may be found to solve a particular problem. An interesting example was published recently (MÜCKLICH *et al.* [1994], OHSER and LORZ [1994]) where a numerical solution was presented for establishing the size and shape distribution of hexagonal prisms as present in silicon nitride from the size and shape of the intersect areas.

In view of these difficulties, serial sectioning has been used for establishing three-dimensional grain and particle size distributions. Alternatively, chemical disintegration or removing of the embedding matrix can be applied to separate grains (e.g., by attacking the grain boundaries of aluminium by liquid gallium as shown by RHINES and PATTER-

References: p. 1016.

SON [1982]) and to size the isolated features microscopically or by weighing. It is interesting to note that logarithmic normal size distributions have been observed in most cases (see, for example, EXNER [1972b] and RHINES and PATTERSON [1982] or RHINES [1986]) which make possible to describe the size distribution by only two parameters (e.g., the arithmetric mean and the standard deviation).

There are a large number of problems in materials research where size, distance and size-distribution are relevant. One of the best-known relationships between microstructural geometry and properties is the Hall–Petch relationship predicting a linear dependence of the yield stress of polycrystalline materials on the square root of reciprocal grain size. However, size distribution has a significant influence on this relationship (KÜHLMEYER [1978], GOKHALE and DE HOFF [1984], RHINES [1986], see also EXNER [1994]) since deformation starts at the largest rather than the average grain (ESSMANN *et al.* [1968]). Experimental results confirm the simple form of the equation owing to the fact that the shape of grain-size distributions in most annealed single-phase materials is fairly similar. Examples for the predicted linearity between the yield strength and the reciprocal square root of the grain size or the particle size are shown in figs. 20a and b. It is interesting to note that prestraining changes the parameters but not the general form of the Hall–Petch relationship (WANG and MURR [1980], see fig. 20a).

However, geometric similarity is not a general rule and can not be assumed *a priori* but needs experimental proof for every material and processing schedule. Large deviations from geometric similarity are frequent. For example, the width of grain-size distributions in recrystallized metals depends strongly on deformation prior to annealing, even at long annealing times (OKAZAKI and CONRAD [1972] and RHINES and PATTERSON [1982]). This has consequences for the relationships between properties and microstructure which must not be neglected.

For aluminium samples, it has been reported by GOKHALE and RHINES [1993] (see also DE HOFF [1984], RHINES [1986] and EXNER [1994]) that the variation in size distribution has a much larger influence on the yield strength than the variation in mean grain size by a factor of 2 and that the same is true for the tensile strength. This example clearly demonstrates the importance of a full appreciation of microstructural parameters in theoretical modelling and the ambiguity of well-accepted relationships between microstructure and properties. Another typical example is the dependence of transverse rupture strength of WC–Co alloys on carbide particle size, spacing and contiguity (see EXNER [1994]).

Other useful equations involving mean size and distance are derived for simplified geometries, often taking a model in which grains or particles are assumed to be spheres of uniform size. For instance, the Zener relationship for dispersed particles limiting the grain size (ch. 28),

$$\bar{L} = k\bar{L}_\alpha / V_V \tag{13}$$

(where $\bar{L}$ is the equilibrium grain size in a material with particles of mean linear size $\bar{L}_\alpha$ and volume fraction V_V), is usually derived in this way (PORTER and EASTERLING [1981]). k is a constant depending on the shape, the size distribution and the arrangement of the particles. The value usually taken ($k = 4/3$) relies on the assumption that the

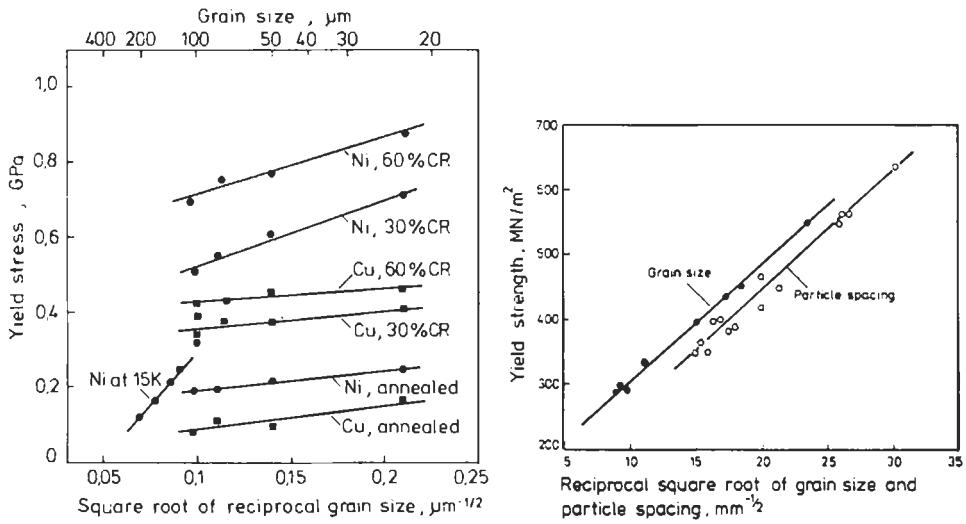


Fig. 20. Experimental correlations between size and distance of microstructural features and the yield strength of metallic alloys, (a) yield strength of some metals as a function of square root of mean linear grain size. CR = cold reduction in percent (data from WANG and MURR [1980] and RAZA [1982], (b) yield strength of unalloyed low-carbon steels as a function of the reciprocal square root of the mean particle spacing and of the mean linear grain size (data from LIU and GURLAND [1968] and ANAND and GURLAND [1976] according to EXNER and HOUGARDY [1988]).

particles are regularly dispersed single-size spheres. Therefore, deviating numbers for k are obtained experimentally.

A more sophisticated treatment is needed including size distributions and it seems that Monte Carlo simulations provide a means for assumption-free modelling. A very promising approach for a generalisation of the Zener equation is to treat the interaction between grain boundaries and matrix/particle interfaces by stereological modelling (AIGELTINGER and EXNER [1977], EXNER [1994].

Equations describing the kinetics of grain growth and particle growth by Ostwald ripening or coalescence, the dependence of dendrite arm spacing and lamellar spacing on freezing rate, and several other processes determining the size of microstructural features have been derived on the basis of simplified model geometries some of which are discussed in more detail in other chapters (e.g., ch. 9, § 15). Experimental proofs given on the basis of quantitative microscopy will be statistically significant and free of systematic errors only if the geometric details of the theoretical models are compatible with the real microstructure. As an example, fig. 21 shows the comparison of size distributions of vanadium carbide particles of spherical shape with predictions made by the Wagner–Lifshitz theory for Ostwald ripening (ch. 15, § 2.4.2). Perfect agreement for reaction-controlled growth is obtained for long annealing times while significant deviations exist for the as-sintered (i.e., briefly annealed) state (EXNER [1973]). This comparison is made on the basis of linear intercepts (EXNER and LUKAS [1971]) rather

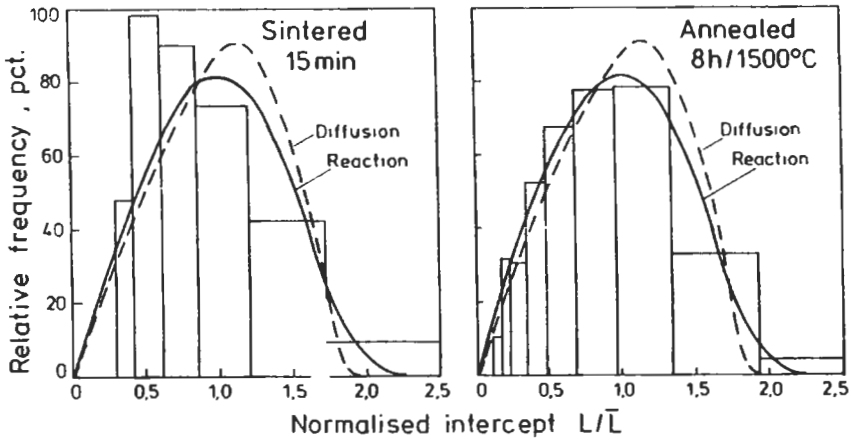


Fig. 21. Size distribution of spherical VC particles in a nickel matrix sintered and heat-treated at 1500°C (liquid matrix). Histograms: experimental results; solid curves: theoretical prediction (after EXNER [1973]). (a) Sintered 1 h, mean linear intercept = 17.5 μm; (b) annealed 8 h, mean linear intercept = 45.5 μm.

than of sphere diameters, since unfolding of the experimental data generally increases the experimental scatter pronouncedly (NICHOLSON [1978]).

7.2.4. Orientation, contiguity, shape and other complex parameters

In addition to volume fraction and interface density (and the related parameters size and distance for separate features in a microstructure) there is a variety of other, more complex aspects of microstructural geometry. Two of these aspects, both rather important in many practical problems, can be quantitatively characterised using measurements made in plane cross sections, namely orientation and contiguity (degree of neighbourship and skeleton formation) using the concept of weighted interface densities.

The degree of orientation is quantified by relating the oriented part of the interface to the total interface of a microstructural component yielding

$$F_{or} = S_V \text{ (oriented)}/S_V \text{ (total)} \tag{14}$$

which, as outlined in early literature (e.g., UNDERWOOD [1970], SALTYKOV [1974], see also EXNER [1994]) can be calculated from directed point counts along perpendicular lines for linearly as well as planarly oriented interfaces (for counting recipes see EXNER [1993]).

Similarly, the degree of skeleton formation of a microstructural component can be derived by relating the grain boundary density to the total interface, i.e.,

$$C^{\alpha\alpha} = S_V \text{ (grain boundary)}/S_V \text{ (total)} \tag{15}$$

as suggested by GURLAND [1958, 1966], and the degree of neighbourship between two microstructural components, α and β, by relating the interface density between component α and β (in a two- or a multiphase alloy) to the total interface of component α, i.e.,

$$C^{\alpha\beta} = S_V \text{ (interface } \alpha\beta)/S_V \text{ (total } \alpha) \quad (16)$$

Useful extensions of the concept of contiguity are the continuous volume and the separated volume (LEE and GURLAND [1978]). Recent reviews of these aspects, their assessment including practical recipes, and a review of applications in materials science are available (EXNER [1993, 1994], FAN *et al.* [1993]). Particularly for characterizing the microstructure of cemented carbides, the concept of contiguity proved indispensible (see, for example, EXNER and FISCHMEISTER [1966], GURLAND [1966], EXNER and GURLAND [1970], or LEE and GURLAND [1978]).

Most recently, a new parameter for characterising the degree of connectiveness in two-phase microstructures was proposed by POECH and RUHR [1994]. Based on the skeletonizing procedure provided by mathematical morphology (see § 7.3) and available with most modern image analysers, the "matricity" of the two phases is defined by the length of the skeleton line in one of the phases divided by the sum of the length of skeleton lines in both phases. The application of this parameter was successfully demonstrated for nickel–silver and martensite–austenite model alloys and for modified aluminium–silicon castings (POECH and RUHR [1994] and POMPE [1994]).

The most difficult problem in quantification of microstructural geometry is shape. Though numerous attempts have been made to find parameters for "average" shape in three dimensions, a general solution is not available, and some proposed parameters (see, for example UNDERWOOD [1970, 1976, 1980], FISCHMEISTER [1974], MEDALIA [1980], EXNER and HOUGARDY [1988], or EXNER [1994]) are very limited, while others are of not much use at all since they either cannot be assessed experimentally or are size-sensitive. There are two ways out: one is to use planar shape descriptors for sections or projections as a "fingerprint" for three-dimensional shape, the other is serial sectioning in order to get the three-dimensional reconstructions or to isolate individual grains and particles chemically (RHINES and PATTERSON [1982], DE HOFF [1982]). While the latter techniques are time consuming and only useful for demonstration purposes or very specific applications, e.g., for proving or disproving a theoretical concept, planar shape characteristics are in widespread use.

Classical shape parameters to describe various aspects such as elongation, bulkiness, waviness, or symmetry can be defined on the basis of combining metric properties like area, perimeter, Ferret's diameter, moments of inertia, etc. (BEDDOW *et al.* [1980], SCHWARZ [1980], EXNER and HOUGARDY [1988], COSTER and CHERMANT [1989], see also EXNER [1994]). Applications of planar shape descriptors to problems in physical metallurgy have remained scarce, but now — with the availability of shape parameters in most automatic instruments — they can be used for descriptive purposes or for establishing empirical correlations for shape-sensitive properties of materials. As an example, a close correlation between fracture strength of eutectic aluminium–silicon alloys and shape of the silicon precipitates has been demonstrated (PAUL *et al.* [1982]).

Other concepts of shape characterization are based on clear mathematical concepts, for example those using Fourier or Walsh coefficient (see, for example, MELOY [1980] or FLOOK [1987]), those based on mathematical morphology (COSTER *et al.* [1981], COSTER and CHERMANT [1989], see § 7.3) or, as a derivative of the presently fashionable

fractal analysis, fractal dimension. Provided that the experimental procedure for the assessment of fractals is sufficiently precise, a unique value characterising the irregularity of interfaces and nonplanar surfaces is obtained. Applications of this concept to problems in materials science are abundand and some interesting examples have been published including fracture surfaces, grain boundaries in deformed metals, phase interfaces after martensite transformations, among other (see reviews by WRIGHT and KARLSSON [1982], HORNBOGEN [1986a,b, 1987], or EXNER [1994]). Considering the fact that these concepts are rather abstract from human perception and even more their limitation to planar geometry, it seems rather doubtful that they will be successful in the fields of materials science and technology (see also EXNER [1994]).

Serial sectioning has been used to evaluate *topological parameters*, mainly in connection with studying the grain shape in metals (RHINES *et al.* [1974] and RHINES and PATTERSON [1982]) and the development of microstructure during sintering (DE HOFF *et al.* [1972] and AIGELTINGER and EXNER [1972]). In this way, the number of separate parts (or, in other words, the number of pores or particles) per unit volume can be assessed. Stereological equations relating numbers per unit area or unit length to numbers in three dimensions are usually not reliable since they depend on information concerning shape- and size-distribution. Another useful topological parameter is connectivity (or genus) which describes the number of connections (channels) between nodes (larger regions). Figure 22 shows the results of a study of sintered copper on which the geometric properties of the pore space have been followed up as a function of porosity, i.e., pore volume fraction (AIGELTINGER and EXNER [1972]). If chemical disintegration for isolating microstructural features is not feasible, serial sectioning, though extremely time-consuming even with automatic sample preparation, automatic instrumentation and computer processing, may be the ultimate answer to geometric analysis of irregularly shaped microstructural features in three dimensions (DE HOFF [1982]).

Another important aspect closely related to both shape and size is curvature. Variations of local curvature provide the driving forces for shape changes and coarsening of microstructural features (e.g., fibers, pore systems, precipitations, etc.) but are difficult to quantify. Integral mean curvature, on the other hand, can be measured quickly and unambiguously from cross-sections using equation (12), see § 7.2.3. As a sensitive parameter for monitoring the path of microstructural changes during annealing (DE HOFF and ISAWARAN [1982]), it should find much more widespread application. In fig. 22, the variation of curvature during the sintering process can be clearly related to the qualitative and quantitative changes of pore geometry. Curvature distribution, on the other hand, cannot be assessed from planar cross sections, rather, stereometric techniques, must be applied (FEIJOO *et al.* [1990]). Though very tedious, the distribution of curvatures of dendrite surfaces was measured with interesting results on the development of dendrites during solidification of aluminium alloys, and a comparison with theoretical predictions was made (FEIJOO and EXNER [1991]).

Finally, there remains the problem of characterizing the *arrangement* of microstructural features. Again, three-dimensional characterization must rely on simplifying assumptions or on serial sectioning. The classification (random, regular, clustered) and quantitative description of planar arrangement of features defined by a characteristic

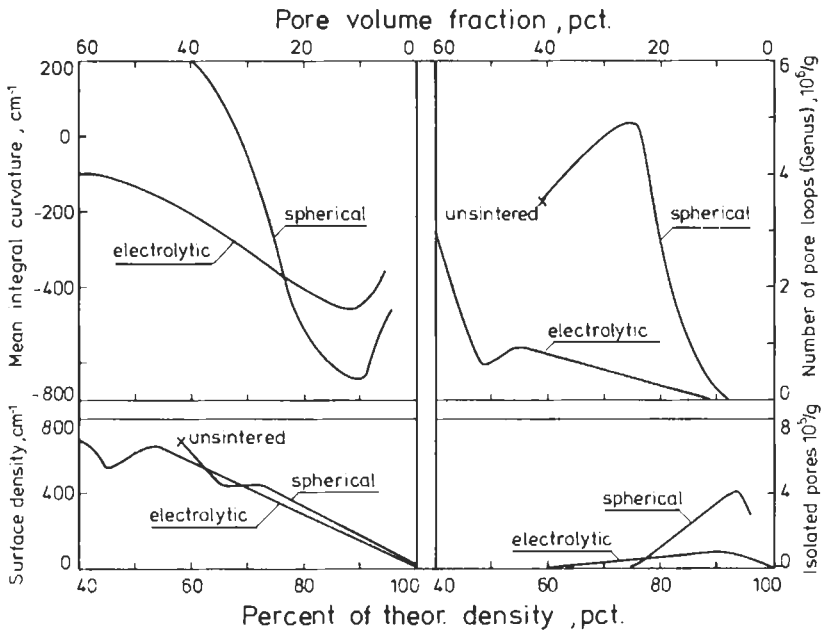


Fig. 22. Changes of metric and topological properties of pore space in sintered copper powder as a function of porosity varied by sintering at 1000°C. The spherical and the electrolytic (dendrite-shaped) powders show different pore structures at comparable porosity. The right part of the figure clearly shows the formation of new contacts (increasing genus) at the early sintering stage and the closure of pore channels (sharply decreasing genus) and formation of isolated pores and their disappearance by shrinking in the later stages. These topological changes are also reflected by the metric properties shown in the left part of the figure: Surface density and integral mean curvature decrease due to smoothening of the pore-solid interface but increase for a short period when the channels start to close and at the very late stages when isolated pores shrink. (After AIGELTINGER and EXNER [1972].)

point (e.g., centre of gravity) on the basis of nearest-neighbour distance distributions has been proposed (SCHWARZ and EXNER [1983]) and a number of statistical procedures became available from other fields, e.g., from pattern recognition and cluster analysis. Up to recently, application of arrangement parameters to materials is near to non-existent. A very interesting exception is a study on the arrangement of carbides in high-speed steels, which has shown the feasibility of simple statistical concepts and the value of arrangement analysis (WERLEFORS *et al.* [1979]). Mathematical morphology (see § 7.3) and, particularly, techniques based upon pair correlation functions and other statistical approaches based on stochastic geometry (see, for example, STOYAN *et al.* [1987], STOYAN [1990], OHSER and LORZ [1994] or EXNER [1994], BERTRAM and WENDROCK [1995] or WENDROCK *et al.* [1995]) supply a wealth of concise descriptors which, however, are only developed for planar arrangement. As mentioned above (§ 7.2.3), most of these approaches seem not a practicability for 3D arrangements and seem rather insensitive to changes of arrangements which are often easily detected by visual

inspection, as, for example, lamellar patterns or particles lined up in rows which have a certain degree of irregularity.

7.3. Mathematical morphology

In addition to classical geometric parameters, a special methodology based on "structuring elements", termed *mathematical morphology* has been developed, largely in France, into a consistent and practically useful framework for image analysis. A full theoretical account of this field, including many pertinent references, has been published by SERRA [1982, 1988]. There is no space here to describe these sophisticated concepts in detail. Attempts to give a comprehensible treatment of image analysis and of image manipulation based on mathematical morphology which, for the materials scientist, are easier to understand than Serra's mathematical theory, have been published by FISCHMEISTER [1981], COSTER and CHERMANT [1989], RUSS [1990, 1992], CHERMANT and COSTER [1994] or CHERMANT [1994], among others. Some procedures which have been proven useful in practical work are outlined in the following.

The microstructure or, rather, its binary image is treated as a set of points which can be transformed by various operations in order to extract characteristic parameters of this set, which then are used to describe the geometry of the microstructure. The effects of the most useful operations, erosion and dilatation, opening and closing are demonstrated in fig. 23. Erosion is achieved by moving the structuring element from picture-point to picture-point and removing all those points which do not fulfill a specified criterion, e.g., that all neighbour points are inside the dispersed features shown in fig. 23a, which, in other words, means removing all points at the edge of the features. By this operation, the two small features 1 and 2 and the bridge connecting two parts of the large particle are eliminated. The area of the large feature has obviously become smaller as shown by the dashed line in fig. 23a. By *dilatation*, i.e., by adding all those matrix points which have at least one neighbouring point inside the features, the original size is very closely reconstructed but the two small features are still gone as is the bridge between the two large features (fig. 23b). Therefore, this combined operation is called *opening*. Similarly, if dilatation is carried out first, followed by erosion (figs. 23c and 23d) the result is reversed and a bridge is formed by this combined operation called *closing*.

It is obvious that this manipulation very effectively allows the elimination of small features or contacts arising from detection noise in automatic instruments or to separate relevant information from the image. Another operation for segmenting connected features is the *grassfire transformation* (NAWRATH and SERRA [1979]) allowing identification and counting of separate particles and grain corners by alternating or repeated erosions and dilatations. Another transformation procedure called skeletonizing reduces each feature (e.g., irregular intersect area) to a pattern of lines which then can be analysed with respect to the number of branches and nodes. The ratio of line lengths of the skeletons of two interdispersed phases was suggested by POECH and RUHR [1994] to characterise the character of two-phase materials. The concepts of *covariograms* and *star function*, and a large number of other parameters were suggested for characterizing size, spacing, shape, arrangement and many other aspects of microstructural geometry by

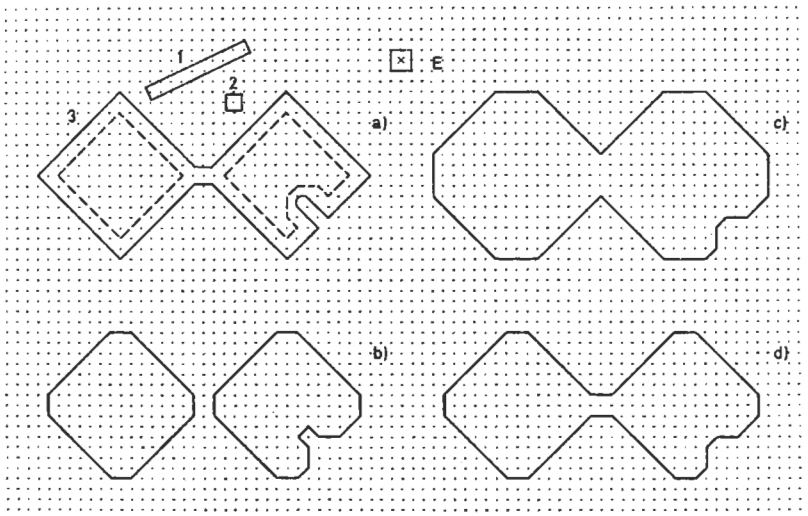


Fig. 23. Two-dimensional image transformations by operations of mathematical morphology (EXNER and HOUGARDY [1988]). E is the element of 9 picture-points (with the reference point in the centre) by which the following operations are carried out: (a) Original features and eroded outline (dashed) of feature 3; (b) dilated outline of eroded fig. 23a; (c) dilated outline of fig. 23b; (d) eroded outline of fig. 23c. Combined operations are called opening (a and b) and closing (c and d).

SERRA [1982, 1988] and COSTER and CHERMANT [1989] and a large number of authors which tried to apply mathematical morphology to practical problems.

Most information obtainable in this way should be strictly considered as fingerprint-type information on microstructural geometry (FISCHMEISTER [1981]) owing to the lack of quantitative relationship with three-dimensional characteristics of the microstructure and to uncontrolled effects of image distortion. The stereological aspects of mathematical morphology have been discussed by SERRA [1982, 1988] and COSTER and CHERMANT [1989], among others. In general, mathematical morphology is very useful in all those cases where changes of microstructural geometry need to be monitored without seeking 3D information. In cases where physical modelling of microstructural changes is attempted, stereological parameters should be preferred. This statement seems to remain valid in spite of the later attempts to clarify the stereological aspects of mathematical morphology culminating in the mathematically correct but unpracticable fact that all stereological equations could be derived from mathematical morphology and thus are a sub-set of this concept.

More recently, mathematical morphology has been extended to the analysis of grey-tone pictures and certainly will further proceed to colour images and other applications not directly relating to microstructures. As mentioned above (§ 7.1), image modification by morphological processing (filtering, cleaning and many other amendments) are indispensible applications of mathematical morphology used in most modern image analysers.

References: p. 1016.

7.4. Further aspects

Generally, more than one single geometric aspect is of relevance in characterizing the microstructural geometry of multiphase materials, and a combination of parameters will be needed to give a full description. Such combinations have been proposed for studying the evolution of microstructure (for example, during sintering, see fig. 22 and ch. 31) or for empirical correlations with mechanical properties (for example, those of cast and annealed aluminium alloys (PAUL *et al.* [1982]), of cemented carbides (EXNER and GURLAND [1970]) or of sintered steel (EXNER and POHL [1978]). In addition to these and other examples cited above, successful applications of quantitative characterization of microstructural geometry have been reported in the literature; some of these have been reviewed in overviews, for example by FISCHMEISTER [1972], SALTYKOV [1974], HOUGARDY [1975], LAFOND and MOLIEXE [1975], RHINES [1976, 1986], MATHY [1977], EXNER [1978a, 1994, 1995], ONDRACEK [1978, 1982], RUSS [1986, 1990, 1992], EXNER and HOUGARDY [1988], or COSTER and CHERMANT [1989], among others. Quite often, however, data are published not so much to aid the solution of a materials problem but rather to demonstrate the feasibility of a new technique or a new instrument. Considering the high standard of methodology now achieved, it seems timely to utilize quantitative analysis of microstructure more directly as a clue for understanding the behaviour of materials during processing and in application.

References

ABRAHAM, F. F., I. P. BATRA and S. CIRACI, 1989, Phys. Rev. Lett. **60**, 1314.

ADAMS, B. L., S. I. WRIGHT and K. KUNZE, 1993, Metall. Trans. **24A**, 819.

AIGELTINGER, E. H., and H. E. EXNER, 1972, Z. Werkstofftechn. (J. Mater. Tech.) **3**, 425.

AIGELTINGER, E. H., and H. E. EXNER, 1977, Metall. Trans. **8A**, 421.

ALDINGER, F., 1969, Metallography **2**, 363.

ANAND, L., and J. GURLAND, 1976, Metall. Trans. **7A**, 191.

ARNOLD, W., 1991, Europhys. News **22**, 171.

ARZT, E., and H. F. FISCHMEISTER, 1979, Pract. Metallogr. **16**, 547.

ASH, E. A., ed., 1980, Scanned Image Microscopy (Academic, London).

ASM HANDBOOK, 1992, Vol. 10, Materials Characterization (ASM International, Metals Park).

AXELROD, D., N. L. THOMPSON and T. P. BURGHARDT, 1983, J. Microsc. **129**, 19.

AYDIN, I., and H. E. BÜHLER, 1981, in: Metallographie — Techniken der Gefügebeschreibung bei metallischen and keramischen Werkstoffen, Pract. Metallogr. Special Issue **12**, eds. H.E. Bühler, K. Dieser and W.U. Kopp (Dr. Riederer Verlag, Stuttgart) p. 35.

AYDIN, I., H. E. BÜHLER, G. ELSSNER and I. STAPF, 1983, Pract. Metallogr. **20**, 105.

AYDIN, I., and H. E. BÜHLER, 1984, Pract. Metallogr. **21**, 243.

AZZAM, R. M. A., and N. M. BASHARA, 1977, Ellipsometry and Polarized Light Microscopy (North-Holland, Amsterdam).

BAKER, M. A., 1993, Surf. Interf. Anal. **20**, 535.

BAKER, M. A. and P. TSAKIROPOULUS, 1993, Surf. Interf. Anal. **20**, 589.

BAKER, M. A., and J. E. CASTLE, 1994, in: Materials Science and Technology, Vol. 2B, ed. E. Lifshin (VCH, Weinheim) p. 219.

BANSAL, R. P., and A. J. ARDELL, 1972, Metallography **5**, 97.

BARDA, J., and F. F. FAN, 1993, in: Scanning Tunneling Microscopy and Spectroscopy, Theory, Techniques and Applications, ed. D.A. Bonnell (VCH Publ., New York) p. 287.

BARTZ, G., 1973, Pract. Metallogr. **10**, 311.
BAUER, B., and A. HALLER, 1981, Pract. Metallogr. **18**, 327.
BAUER, B., and H.E. EXNER, 1981, Stereolog. Jugosl. **3**, Suppl. 1, 255.
BAUER, B., M. FRIPAN and V. SMOLEJ, 1982, in: Fracture and the Role of Interfaces, vol. 2, Fatigue, eds. K.L. Maurer and F.E. Matzer (Engineering Materials Advisory Services, Warley) p. 591.
BAUER, B., and B. EGG, 1984, Pract. Metallogr. **21**, 461.
BEATON, C., and B. FILSHIE, 1983, Pract. Metallogr. **20**, 31.
BECKERT, M., and H. KLEMM, 1984, Handbuch der metallographischen Ätzverfahren (Deutscher Verlag für Grundstoffindustrie, Leipzig, Stuttgart).
BEDDOW, J.K., M.D. NASTA and G.C. PHILIP, 1980, in: Testing and Characterization of Fine Particles, eds. J.K. Beddow and T. Meloy (Heyden, London) p. 44.
BELL, L.D., W.J. KAISER, M.H. HECHT and L.C. DAVIS, 1993, in: Scanning Tunneling Microscopy, eds. J.A. Stroscio and W.J. Kaiser (Academic Press, Boston) p. 307.
BENESOVSKY, F., and A. IHRENBERGER, 1965, Pract. Metallogr. **2**, 31.
BERAHA, E., 1970, Pract. Metallogr. **7**, 131 and 242.
BERAHA, E., and B. SPIGHLER, 1977, Color Metallography (ASM, Metals Park, OH).
BERNST, R., 1965, Pract. Metallogr. **2**, 162.
BERTRAM, M., and H. WENDROCK, 1995, in: Fortschritte in der Metallographie, Pract. Metallogr., Spec. Issue **26**, ed. A. Kneissl (Carl Hornser Verlag, München) p. 375.
BEYER, H., 1974, Theorie und Praxis der Interferenzmikroskopie (Akademische Verlagsgesellschaft Geest und Portig, Leipzig).
BIGGS, W.D., 1970, in: Physical Metallurgy, 2nd Ed., ed. R.W. Cahn (North-Holland, Amsterdam) p. 655.
BINNIG, G., H. ROHRER, C. GERBER and E. WEIBEL, 1982, Phys. Rev. Lett. **40**, 178, and **49**, 57.
BINNIG, G., and H. ROHRER, 1982, Helv. Phys. Acta **55**, 726.
BINNIG, G., C.F. QUATE and C. GERBER, 1986, Phys. Rev. Lett. **56**, 930.
BINNIG, G., and H. ROHRER, 1987, Rev. Mod. Phys. **59**, 615.
BIRNER, E., and R. LÖHBERG, 1980, Pract. Metallogr. **17**, 14.
BLASCHKE, R., and H. WALTINGER, 1971, Beiträge elektronenmikroskopische Direktabbildung von Oberflächen, vol. 4/2 (Verlag R.A. Remy, Münster) p. 425.
BLASCHKE, R., 1979, Beiträge elektronenmikroskopische Direktabbildung von Oberflächen, vol. 12/1 (Verlag R.A. Remy, Münster) p. 359.
BONNELL, D.A., ed., 1990, Scanning Tunneling Microscopy, Theory and Practice (VCH, Publ., New York).
BONNELL, D.A., ed., 1993a, Scanning Tunneling Microscopy and Spectroscopy (VCH Publ., New York).
BONNELL, D.A., 1993b, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon, Oxford) p. 427.
BONNELL, D.A., 1993c, in: Scanning Tunneling Microscopy and Spectroscopy, Theory, Techniques and Applications, ed. D.A. Bonnell (VCH Publ., New York) p. 7.
BOWKETT, K.M. and D.A. SMITH, 1970, Field Ion Microscopy (North Holland, Amsterdam).
BOWMAN, M.J., and A.D. BOOTH, 1971, Metallography **4**, 103.
BOYDE, A., 1973, J. Microsc. **98**, 452.
BOYDE, A., 1985, Science **230**, 1270.
BOYDE, A., 1990, in: Modern Microscopies: Techniques and Applications, eds. P.J. Duke and A.G. Michette (Plenum Press, New York, London) p. 185.
BRADBURY, S., 1991, An Introduction to the Optical Microscope (Oxford University Press, Oxford).
BRAKENHOFF, G.J., P. BLOM, and P. BARENDS, 1979, J. Microsc. **117**, 219.
BRENNER, S.J., 1978, Surf. Sci. **70**, 427.
BRIGGS, D., and M.P. SEAH, 1990, Practical Surface Analysis (Wiley, Chichester).
BRIGGS, A., 1992, Acoustic Microscopy (Oxford University Press, Oxford).
BRIGGS, A., 1994, in: Materials Science and Technology, Vol. 2B, ed. E. Lifshin (VCH, Weinheim) p. 241.
BROOKER, A.D., J.E. CASTLE, J.M. COHEN and M.B. WALDRON, 1984, Met. Technol. **2**, 66.
BROWN, L.M., M. YOUNIS KHAN and M.M. Chaudri, 1988, Philos. Mag. **A57**, 187.
BROWN, R., and G.C. SMITH, 1982, Metallography **15**, 269.
BRÜGEL, H., J.R. ROSS and G. KÖNIG, 1988, Pract. Metallogr. **25**, 578.

BRÜMMER, O., and J. NIEBER, 1975, Microchim. Acta, Suppl. **6**, 345.
BRÜMMER, O., J. HEYDENREICH, K. H. KREBS and H. G. SCHNEIDER, eds., 1980, Handbuch Festkörperanalyse mit Elektronen, Ionen und Röntgenstrahlen (Vieweg, Braunschweig).
BÜCKLE, H., 1964, Berg- und Hüttenmänn. Monatsh. **109**, 72.
BÜCKLE, H., 1965, Mikrohärteprüfung und ihre Anwendung (Verlag Berlin Union, Stuttgart).
BUCKLEY, C. J. and H. RARBACK, 1990, in: Modern Microscopies, Techniques and Applications, eds. P.J. Duke and A.G. Michette (Plenum Press, New York) p. 69.
BÜHLER, H. E., and H. P. HOUGARDY, 1979, Atlas der Interferenzschichten-Metallographie (DGM Informationsgesellschaft, Oberursel).
BÜHLER, H. E., and H. P. HOUGARDY, 1980, Atlas of Interference-Layer Metallography (DGM Informationsgesellschaft, Oberursel, and The Metals Society, London).
BÜHLER, H. E., 1981, in: Microstructural Science, vol. 9, eds. G. Petzow, R. Paris, E.D. Albrecht and J.L. McCall (Elsevier North-Holland, New York) p. 19.
BÜHLER, H. E., and D. KOSSEL, 1981, Pract. Metallogr. **18**, 385.
BÜHLER, H. E., 1986, Pract. Metallogr. **23**, 309.
BURNHAM, N. A., and R. J. COLTON, 1993, in: Scanning Tunneling Microscopy and Spectroscopy, Theory, Techniques and Applications, ed. D.A. Bonnell (VCH Publ., New York) p. 191.
CAHN, J. W., 1967, Trans. Metall. Soc. AIME **239**, 610.
CAHN, R. W., J. E. EVETTS, J. PATTERSON, R. E. SOMEKH and C. K. JACKSON, 1980, J. Mater. Sci. **15**, 702.
CAHN, R. W., and E. LIFSHIN, eds., 1993, Concise Encyclopedia of Materials Characterization (Pergamon, Oxford).
CALVO, F. A., A. J. CRIADO, J. M. GOMEZ DE SALAZAR and F. MOLLEDA, 1987, Metallography **20**, 213.
CARLSON, T. A., 1975, Photoelectron and Auger Spectroscopy (Plenum, New York).
CATALDI, T. R. I., I. G. BLACKHAM, G. A. D. BRIGGS, J. B. PETHICA and H. A. O. HILL, 1990, J. Electroanal. Chem. **290**, 1.
CEREZO, A., J. M. HYDE, M. K. MILLER, S. C. PETTS, R. P. SETNA, G. D. W. SMITH, 1992, Phil. Trans. Royal Soc. **A341**, 313.
CEREZO, A., and G. D. W. SMITH, 1994, in: Materials Science and Technology, Vol. 2B, ed. E. Lifshin (VCH, Weinheim) p. 513.
CHANG, S. L., H. J. QUEISSER, H. BAUMGART, W. HAGEN and W. HARTMANN, 1982, Phil. Mag. **46**, 1009.
CHENG, P. C., S. P. NEWBURRY, H. G. KIM and I. S. HWANG, 1990, in: Modern Microscopies, Techniques and Applications, eds. P.J. Duke and A.G. Michette (Plenum Press, New York) p. 87.
CHERMANT, J. L., 1994, Rev. Métallurgie **91**, 201.
CHERMANT, J. L., and M. COSTER, 1994, Acta Stereol. **13**, 125.
CHERNUSKA, J. T., and T. F. PAGE, 1985, J. Am. Ceram. Soc. **68**, C196.
CHESCOE, D., and P. J. GOODHEW, 1990, The Operation of Transmission and Scanning Electron Microscopes (Oxford University Press, Oxford).
CLARKE, T. M., in: Metals Handbook, 9th ed., Vol. 12, Fractography (American Society for Metals, Metals Park) p. 78.
CONN, G. K. T., and F. J. BRADSHAW, 1952, Polarized Light in Metallography (Butterworths, London).
CONOR, P. C., 1972, Metallography **5**, 301.
COSSLETT, V. E., and W. C. NIXON, 1960, X-ray Microscopy (Cambridge University Press).
COSTER, M., and J. L. CHERMANT, 1989, Précis d'Analyse d'Images (Presses du Centre National de la Recherches Scientifique, Paris).
CROUSE, R. S., R. J. GRAY and B. C. LESLIE, 1977, in: Interpretive Techniques in Microstructural Analysis, eds. J.L. McCall and P.M. French (Plenum, New York) p. 43.
CRUZ-ORIVE, L. M., 1978, J. Microsc. **112**, 153.
CZANDERNA, A. W., ed., 1975, Methods of Surface Analysis (Elsevier, Amsterdam).
CZYRSKA-FILEMONOWICZ, A., S. GORCYCA and K. SPYRADEK, 1992, Pract. Metallogr. **29**, 35.
DANIELSON, P. E., 1985, in: Metals Handbook, Vol. 9, Metallography and Microstructures (American Society for Metals, Metals Park) p. 142.
DANNÖHL, H. D., R. GRABER and L. WEGMANN, 1971, Microstructures **2**, 15.
DAVIDSON, D. L., 1974, in: SEM 1974, Proc. 7th Annual SEM Symp. (IIT Research Institute, Chicago) p. 927.

DE HOFF, R.T., 1967, Trans. Metall. Soc. AIME **239**, 617.
DE HOFF, R.T., and F.N. RHINES, eds., 1968, Quantitative Microscopy (McGraw-Hill, New York).
DE HOFF, R.T., 1972, in: Treatise on Materials Science and Technology, Vol. 1, ed. H. Herrmann (Academic Press, New York, London) p. 247.
DE HOFF, R.T., E. AIGELTINGER and K.R. CRAIG, 1972, J. Microsc. **95**, 69.
DE HOFF, R.T., 1980, Mikroskopie **37**, 32.
DE HOFF, R.T., 1982, paper presented at Stereology 82, Symp. Int. Soc. for Stereology, Sheffield.
DE HOFF, R.T., and C.V. ISAWARAN, 1982, Metallurg. Trans. **13A**, 1389.
DE HOFF, R.T., 1984, in: Practical Applications of Quantitative Metallography, eds. J.L. McCall, J.H. Steele (American Society for Testing and Materials, Philadelphia) p. 146.
DE HOFF, R.T., 1989, in: Science of Sintering: New Directions for Materials Processing and Microstructure Control, eds. D.P. Uskovic, M. Palmour and R.M. Spriggs (Plenum Press, New York, London) p. 55.
DE RHODE, W.H., and J.W. SMITS, 1981, J. Appl. Phys. **52**, 3969.
DENGEL, D., 1973, Z. Werkstofftechn. (J. Mater. Techn.) **4**, 292.
DIESER, K., 1984, Z. Werkstofftechn. **15**, 177.
DIESER, K., and E. HERRMANN, 1988, Pract. Metallogr. **25**, 374.
DIESER, K., and E. MATTHAEI, 1988, in: Microstructural Science, Vol. 16, Metallography of Advanced Materials, eds. H.J. Cialone, G.W.E. Johnson, M.E. Blum and G.F. Vander Voort (ASM International, Metals Park) p. 305.
DINARDO, N.J., 1994a, Nanoscale Characterization of Surfaces and Interfaces (VCH, Weinheim).
DINARDO, N.J., 1994b, in: Materials Science and Technology, Vol. 2B, ed. E. Lifshin (VCH, Weinheim) p. 1.
DINGLEY, D.J., C. HARPER and S. LONG, 1982, in: Electron Microscopy and Analysis 1981, Conf. Series 61, ed. M.J. Goringe (The Institute of Physics, Bristol and London) p. 63.
DOERR, R.M., and P.D. OWNBY, 1975, Pract. Metallogr. **12**, 78.
DOHERTY, R.D., and R.W. CAHN, 1972, J. Less-Common Met. **28**, 279.
DOYEN, G., and D. DRAKOVA, 1995, The Physical Principles of STM and AFM Operation (Akademie-Verlag, Berlin).
DUKE, P.J., 1981, Proc. Roy. Microsc. Soc. **16**, 186.
DUKE, P.J., and A.G. MICHETTE (eds.), 1990, Modern Microscopies: Techniques and Applications (Plenum Press, New York, London).
EDELMANN, C., 1980, in: Handbuch Festkörperanalyse mit Elektronen, Ionen und Röntgenstrahlen, eds. O. Brümmer, J. Heydenreich, K.H. Krebs and H.G. Schneider (Vieweg, Braunschweig) p. 263.
EKELUND, S., and S. HERTZMANN, 1981, in: Computers in Materials Technology, ed. T. Encson (Pergamon Press, Oxford) p. 153.
ELSBROCK, J.B., and L.J. BALK, 1982, Beiträge elektronenmikroskopische Direktabbildung von Oberflächen, vol. 15 (Verlag R.A. Remy, Münster) p. 17.
ELSSNER, G., G. KIESSLER and H.G. KAPITZA, 1980, Pract. Metallogr. **28**, 443.
ELSSNER, G., and W.U. KOPP, 1984, Pract. Metallogr. **21**, 633.
ENGEL, H.J., 1958, Arch. Eisenhüttenw. **29**, 73.
ESSMANN, U., M. RAPP and M. WILKENS, 1968, Acta Metall. **16**, 1275.
EXNER, H.E., and H.F. FISCHMEISTER, 1966, Z. Metallkde. **57**, 187.
EXNER, H.E., and H.F. FISCHMEISTER, 1966b, Arch. Eisenhüttenw. **37**, 417.
EXNER, H.E., 1969a, Trans. Met. Soc. AIME **245**, 677.
EXNER, H.E., 1969b, Pract. Metallogr. **6**, 639.
EXNER, H.E., and J. GURLAND, 1970, Powder Metallurg. **13**, 13.
EXNER, H.E., and H.L. LUKAS, 1971, Metallography **4**, 325.
EXNER, H.E., 1972a, in: Stereology and Quantitative Metallography, ASTM Spec. Tech. Publ. 504 (American Society for Testing and Materials, Philadelphia).
EXNER, H.E., 1972b, Int. Metallurg. Rev. **17**, 111.
EXNER, H.E., 1973, Z. Metallk. **64**, 273.
EXNER, H.E., and D. POHL, 1978, Powder Metallurg. Int. **10**, 193.
EXNER, H.E., 1978a, in: Quantitative Metallography (Associazione Italiana di Metallurgia, Milano) p. 313.
EXNER, H.E., 1978b, Pract. Metallogr. **15**, 15.

EXNER, H. E., and J. ROTH, 1980, Pract. Metallogr. **17**, 365.
EXNER, H. E., H. BACK and J. ROTH, 1980, Pract. Metallogr. **17**, 344.
EXNER, H. E., and H. P. HOUGARDY, 1983, Quantitative Image Analysis of Microstructures (DGM Informationsgesellschaft, Oberursel).
EXNER, H. E., 1985, in: Metals Handbook, 9th ed., Vol. 9; Metallography and Microstructures (American Society for Metals, Metals Park) p. 89.
EXNER, H. E., and M. FRIPAN, 1985, J. Microscopy **139**, 161.
EXNER, H. E., 1987, Acta Stereol. **6**, Suppl. II, 179.
EXNER, H. E., and E. A. GIESS, 1989, in: Science of Sintering: New Directions for Materials Processing and Microstructure Control, eds. D.P. Uskovic, M. Palmour and R.M. Spriggs (Plenum Press, New York, London) p. 55.
EXNER, H. E., ed., 1993, Pract. Metallogr. **30**, 216, 287, 322 and 389.
EXNER, H. E., 1994, in: Materials Science and Technology, Vol. 2B, ed. E. Lifshin (VCH, Weinheim) p. 281.
EXNER, H. E., 1995, in: Fortschritte der Metallographie, Pract. Metallogr., Spec. Issue 26, ed. A. Kneissl (Carl Hanser Verlag, München) p. 49.
FAN, Z., A. P. MIODOWNIK and P. TSAKIROPOULOS, 1993, Mater. Sci. Technol. **9**, 1094.
FEDER, R., E. SPILLER, J. TOPALIAN, A. N. BRORS, W. GUDAT, B. J. PANESSA, Z. A. ZADUNAISKY and J. SEDAT, 1977, Science **197**, 259.
FEIJOO, D., B. BAUER and H. E. EXNER, 1990, J. Comp.-Assist. Microsc. **2**, 3.
FEIJOO, D., and H. E. EXNER, 1991, J. Cryst. Growth. **113**, 449.
FEIJOO, D., and H. E. EXNER, 1992, Pract. Metallogr. **27**, 605.
FISCHMEISTER, H., 1965, Pract. Metallogr. **2**, 257.
FISCHMEISTER, H., and H. E. EXNER, 1966, Arch. Eisenhüttenw. **37**, 499.
FISCHMEISTER, H., 1972, J. Microsc. **95**, 25.
FISCHMEISTER, H. F., 1974, Z. Metallk. **65**, 558.
FISCHMEISTER, H., 1981, in: Computers in Materials Technology, ed. T. Ericson (Pergamon Press, Oxford) p. 109.
FISCHMEISTER, H. F., and E. ARZT, 1981, in: Gefüge der Metalle (DGM Informationsgesellschaft, Oberursel) p. 177.
FISCHMEISTER, H. F., S. KARAGÖZ, I. LIEM, J. PAUL and M. H. POECH, 1986, Acta Stereol. **5**, 287.
FISCHMEISTER, H. F., 1988, Fresenius Z. Anal. Chem. **334**, 421.
FISCHMEISTER, H. F., S. KARAGÖZ and H. D. ANDRÉN, 1988, Acta Metall. **36**, 817.
FLOOK, A. G., 1987, Acta Stereol. **6**, Suppl. III, 1009.
FOREMAN, L. R., N. E. ELLIOTT and D. M. STUPIN, 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 238.
FRANK, D. G. and A. T. HUBBARD, 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 4.
FREMUNT, P., J. SVEJCAR and J. VARHANICEK, 1980, Pract. Metallogr. **17**, 497.
FREUND, H., ed., 1960, Handbuch der Mikroskopie in der Technik (Umschau Verlag, Frankfurt) Band 1, Die optischen Grundlagen, die Instrumente und Nebenapparate für die Mikroskopie in der Technik, Teil 2, Allgemeines Instrumentarium der Auflichtmikroskopie.
FREUND, H., ed., 1969, Handbuch der Mikroskopie in der Technik (Umschau Verlag, Frankfurt) Band 3, Mikroskopie der metallischen Werkstoffe, Teil 2, Qualitative und quantitative Untersuchungsverfahren in der Metallkunde.
FULRATH, R. M., 1972, in: SEM 1972, Proc. 5th Annual SEM Symp. (IIT Research Institute, Chicago) p. 17.
FUNDAL, E., and S. GROSS, 1993, Pract. Metallogr. **30**, 294.
GABRIEL, B. L., 1987, in: Metals Handbook, Vol. 12: Fractography (American Society for Metals, Metals Park) p. 166.
GABRIEL, B. L., 1992a, SEM: A User's Manual for Materials Science (American Society for Metals, Metals Park).
GABRIEL, B. L., 1992b, in: SEM: A User's Manual for Materials Science (American Society for Metals, Metals Park) p. 137.
GAHM, J., 1969, Zeiss-Mitteilungen **3**, 40.

GAHM, J., 1971, Zeiss-Mitteilungen **5**, 249.
GAHM, J., 1975, in: Quantitative Gefügeanalyse in Medizin, Biologie und Materialentwicklung, Pract. Metallogr., Special Issue 5, ed. H.E. Exner (Dr. Riederer Verlag, Stuttgart) p. 29.
GAHM, H., and F. JEGLITSCH, 1981, in: Microstructural Science, vol. 9, eds. G. Petzow, R. Paris, E.D. Albrecht and J.L. McCall (Elsevier North-Holland, New York) p. 65.
GAHM, H., F. JEGLITSCH and H. HORL, 1982, Pract. Metallogr. **19**, 369.
GALOPIN, R., and N.F.M. HENRY, 1972, Microscopic Study of Opaque Minerals (Heffer, Cambridge,).
GARCIA, N., ed., 1987, Proc. First Intern. Conf. on Scanning Tunneling Microscopy, Surf. Sci. **181**, 1.
GASTALDI, J. and C. JOURDAN, 1984, in: Applications of X-Ray Topographic Methods, eds. S. Weissmann, F. Balibar and J.F. Petroff (Plenum Press, New York) p. 273.
GAUDIG, W., and R. SCHECK, 1984, Pract. Metallogr. **21**, 624.
GIESSIBL, J., 1995, Science **267,** 68.
GILMORE, R., 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 4.
GLICKSMAN, M.E, R.N. SMITH, S.P. MARSH and R. KULIK, 1992, Metall. Trans. A, **A23,** 659.
GOKHALE, A.M., and F.N. RHINES, 1983, in Microstructural Science, Vol. 11, eds. R.T. De Hoff, J.D. Braun and J.L. McCall (Elsevier Publ., Amsterdam) p. 3.
GOLDSTEIN, J.I., and H. YAKOWITZ, eds., 1977, Practical Scanning Electron Microscopy (Plenum, New York).
GOLDSTEIN, J.I., D.E. NEWBURY, P. ECHLIN, D.C. JOY, A.D. ROMIG, C.E. ZYMAN, C. FIORI and E. LIFSHIN, 1994, Scanning Electron Microscopy and X-Ray Microanalysis (Plenum Press, New York).
GOODHEW, P.J., 1973, Specimen Preparation in Materials Science (North-Holland, Amsterdam, and American Elsevier, New York).
GRABATIN, H., H.D. KUNZE and J. RICKEL, 1983, Pract. Metall. **20**, 609.
GRÄF, I., M. POHL and W.G. BURCHARD, 1993, Pract. Metallogr. **30**, 469.
GRÄF, I., 1981, in: Microstructural Science, vol. 9, eds. G. Petzow, R. Paris, E.D. Albrecht and J.L. McCall (Elsevier North-Holland, New York) p. 45.
GRASSERBAUER, M., 1989, Surface and Interface Analysis (Technical University, Vienna), cited from ORTNER and WILHARTITZ [1991].
GRÜTTER, P., H.J. MAMIN and D. RUGAR, 1992, in: Scanning Tunneling Microscopy II, eds. R. Wiesendanger and H.J. Güntherodt (Springer, Berlin) p. 151.
GRÜTZNER, G., and H.J. SCHÜLLER, 1969, Pract. Metallogr. **6**, 346.
GÜNTHERODT, H.J. and R. WIESENDANGER, eds., 1992, Scanning Tunneling Microscopy I, (Springer-Verlag, Berlin).
GURLAND, J., 1958, Trans. Metall. Soc. AIME **212**, 452.
GURLAND, J., 1966, Trans. Metall. Soc. AIME **236**, 642.
HAASEN, P., 1985, Metall. Trans. A, **16A**, 1173.
HAASEN, P. and R. WAGNER, 1985, Ann. Rev. Mater. Sci. **15**, 43.
HAMERS, R.J., 1993, in: Scanning Tunneling Microscopy and Spectroscopy, Theory, Techniques and Applications, ed. D.A. Bonnell (VCH Publ., New York) p. 51.
HANSMA, P.K., ed., 1982, Tunneling Spectroscopy: Capabilities, Applications, and New Techniques (Plenum Press, New York).
HANSMA, P.K., and J. TERSOFF, 1987, J. Appl. Phys. **61**, R1.
HARRIS, L.A., 1968, J. Appl. Phys. **39,** 1419.
HARTSHORNE, N.H., and A. STUART, 1952, Crystals and the Polarising Microscope (Edward Arnold, London).
HEARLE, J.W.S., J.T. SPARROW and P.M. CROSS, 1972, The Use of Scanning Electron Microscopy (Pergamon Press, Oxford).
HERBSLEB, G., and P. SCHWAAB, 1978, Pract. Metallogr. **15**, 213.
HEROLD-SCHMIDT, U., 1988, Pract. Metallogr. **25**, 3.
HEYWOOD, V.H., ed., 1971, Scanning Electron Microscopy — Systematic and Evolutionary Applications (Academic, London).
HILLIARD, J.E., 1976, in: Fourth International Congress for Stereology, NBL Spec. Techn. Publ. **413**, eds. E.E. Underwood, R. De Wit and G.A. Moore (National Bureau of Standards, Gaithersburg) p. 59.
HILLJE, G., and G. REDMANN, 1976, Pract. Metallogr. **12**, 629.

HILLMER, T., 1979a, Pract. Metallogr. **16**, 476.
HILLMER, T., 1979b, Pract. Metallogr. **16**, 521.
HILLMER, T., 1982, Pract. Metallogr. **19**, 509.
HÖCHE, H.R., and O. BRÜMMER, 1980, in: Handbuch Festkörperanalyse mit Elektronen, Ionen und Röntgenstrahlen, eds. O. Brümmer, J. Heydenreich, K.H. Krebs and H.G. Schneider (Vieweg, Braunschweig) p. 57.
HOEKSTRA, S., and W.H.J. BRUIS, 1979, Pract. Metallogr. **16**, 583.
HOFMANN, S., 1980, Surf. Interf. Anal. **2**, 148.
HOLIK, A.S., 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 286.
HOLT, D.B. and F.M. SABA, 1985, Scann. Electr. Microsc. **3**, 1023.
HONO, K., T. HASHIZUME, Y. HASEGAWA, K. HIRANO and T. SAKURAI, 1986, Scripta Metall. **20**, 487.
HONO, K., T. SAKURAI and H.W. PICKERING, 1989, Metall. Trans. A, **20A**, 1585.
HORNBOGEN, E., 1968, Pract. Metallogr. **2**, 51.
HORNBOGEN, E., and G. PETZOW, 1970, Z. Metallkde. **61**, 81.
HORNBOGEN, E., 1981, Z. Metallk. **72**, 739.
HORNBOGEN, E., 1984, Acta Metall. **32**, 615.
HORNBOGEN, E., 1986a, Pract. Metallogr. **23**, 258.
HORNBOGEN, E., 1986b, J. Mater. Sci. **21**, 3737.
HORNBOGEN, E., 1987, Z. Metallkde. **78**, 622.
HORNBOGEN, E., 1989, Intern. Mater. Rev. **34**, 277.
HORNBOGEN, E., and G. PETZOW, 1991, Pract. Metallogr. **28**, 320 and 383 (reprint of HORNBOGEN and PETZOW [1970] with English translation).
HOUGARDY, H.P., 1975, Research Film **8**, 444.
HOUGARDY, H.P., 1976, in: 4th Int. Congr. for Stereology, NBL Spec. Techn. Publ. **431**, eds. E.E. Underwood, R. De Wit and G.A. Moore (National Bureau of Standards, Galthersburg) p. 141.
HOUGARDY, H.P., 1981, in: Metallographie: Techniken der Gefügebeschreibung bei metallischen und keramischen Werkstoffen, Pract. Metallogr. Special Issue 12, eds. H.E. Bühler, K. Dieser and W.U. Kopp (Dr. Riederer Verlag, Stuttgart) p. 9.
HOWELLS, M.R., C. JACOBSEN, J. KIRZ, K. MCQUAID and S.S. ROTHMANN, 1990, in: Modern Microscopies, Techniques and Applications, eds. P.J. Duke and A.G. Michette (Plenum Press, New York) p. 119.
HUES, G.M., R.J. COLTON, E. MEYER and H.J. GÜNTHERODT, 1993, Mat. Res. Bullet. **18**,
HUMMEL, R.E., 1983, phys. stat. sol. (a) **76**, 11.
HUN-DE, Z., and W. JING-YUN, 1980, Pract. Metallogr. **17**, 608.
IBACH, H., ed., 1977, Electron Spectroscopy for Surface Analysis (Springer, Berlin).
IRANI, R.S., and R.W. CAHN, 1971, Metallography **4**, 91.
JÄHNE, B., 1993, Digital Image Processing: Concepts, Algorithms and Software Applications (Springer Publ., Berlin).
JEGLITSCH, F., and R. MITSCHE, 1967, Radex Rundschau, 587.
JEGLITSCH, F., 1968, in: Handbuch der Mikroskopie in der Technik, Band 3, Teil 1, ed. H. Freund (Umschau Verlag, Frankfurt), p. 187 and 247.
JEGLITSCH, F., 1989, Pract. Metallogr. **26**, 389.
JOST, N., and E. HORNBOGEN, 1988, Pract. Metallogr. **25**, 157.
JOY, D.C., 1992, in: Materials Science and Technology, Vol. 2A; ed. E. Lifshin (VCH, Weinheim) p. 221.
KAISER, W.J., L.D. BELL, M.H. HECHT and L.C. DAVIS, in: Scanning Tunneling Microscopy and Spectroscopy, Theory, Techniques and Applications, ed. D.A. Bonnell (VCH Publ., New York) p. 251.
KARAGÖZ, S., and H. FISCHMEISTER, 1987, Steel Research **58**, 46.
KELLER, H.E., 1977, in: Interpretive Techniques for Microstructural Analysis, eds. J.L. McCall and P.M. French (Plenum, New York) p. 105.
KELLOG, G.L., and T.T. TSONG, 1980, J. Appl. Phys. **51**, 1184.
KIESSLER, G., L. GESSNER and G. ELSSNER, 1978, in: Metallographie und Keramographie, Pract. Metallogr., Special Issue 9, eds. W.U. Kopp and H.E. Bühler (Dr. Riederer Verlag, Stuttgart) p. 113.

KIESSLER, G., and G. ELSSNER, 1980, in: Metallographie, Anschliff- und Dünnschlifftechnik an Metallen, Keramiken und Kunststoffen, Pract. Metallogr. Special Issue 11, eds. H.E. Bühler and W.U. Kopp (Dr. Riederer Verlag, Stuttgart) p. 21.

KIESSLER, G., H. RAPP and G. ELSSNER, 1982, in: Metallographie, Pract. Metallogr., Special Issue 13, eds. H.E. Bühler, K. Dieser and W.U. Kopp (Dr. Riederer Verlag, Stuttgart) p. 9.

KING, R.P., 1982, Powder Technol. **32**, 87.

KINGLAKE, R., ed., 1965, Applied Optics and Optical Engineering (Academic, New York).

KLEISER, T., and M. BOCEK, 1986, Z. Metallkde. **77**, 582.

KLOCKENKÄMPER, R., A. BEYER and M. MONES, 1979, Pract. Metallogr. **16**, 53.

KNEISSL, A., and S.S. BRENNER, 1990, Pract. Metallogr. **27**, 159.

KOCH, R., 1979, Pract. Metallogr. **16**, 11.

KOCH, E.E., (ed.), 1983: Handbook of Synchrotron Radiation (North Holland, Amsterdam).

KOPP, W.U., and G. MÜLLER, 1987, Pract. Metallogr. **24**, 336.

KORPEL, A., L.W. KESSLER, P.R. PALERMO, 1971, Nature **232**, 100.

KRONER, B., 1986, Pract. Metallogr. **23**, 400.

KÜHLMEYER, M., 1978, Einfluss der statistischen Korngrössenverteilung auf die Streckgrenze von Stahl (Verlag Stahleisen, Düsseldorf).

KUK, Y., 1993, in: Scanning Tunneling Microscopy, eds. J.A. Stroscio and W. J. Kaiser (Academic Press, Boston) p. 277.

KULMBURG, A., F. KORNTHEUER and P. SCHIMMEL, 1974, Pract. Metallogr. **11**, 183.

LAFOND, C., and F. MOLIERE, 1975, in: Quantitative Gefügeanalyse in Medizin, Biologie und Materialentwicklung, Pract. Metallogr., Special Issue 5, ed. H.E. Exner (Dr. Riederer Verlag, Stuttgart) p. 200.

LEAMY, H.J., 1982, J. Appl. Phys. **53**, R 51.

LEE, H.C., and J. GURLAND, 1978, J. Mater. Eng. **33**, 125.

LEHTINEN, B., and A. MELANDER, 1980, Metallography **13**, 283.

LEMONS, R.A., and C.F. QUATE, 1974, Appl. Phys. Lett. **24**, 162.

LIENKAMP, M., and H.E. EXNER, 1994, Mater. Sci. Technol. **10**, 526.

LIENKAMP, M. U. KUNAVER and H.E. EXNER, 1994, J. Comp. Assist. Microsc. **6**, 103.

LIFSHIN, E., ed., 1992, Characterization of Materials, Part I, Materials Science and Technology, Vol. 2A (VCH, Weinheim).

LIFSHIN, E., 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 286.

LIFSHIN, E., ed., 1994a, Characterization of Materials, Part II, Materials Science and Technology, Vol. 2B (VCH, Weinheim).

LIFSHIN, E., 1994b, in: Materials Science and Technology, Vol. 2B, ed. E. Lifshin (VCH, Weinheim) p. 351.

LIHL, F., and H. MEYER, 1960, Z. Metallk. **51**, 186.

LINDÉN, G., 1972, Pract. Metallogr. **9**, 3.

LIU, C.T., and J. GURLAND, 1968, Trans. Amer. Soc. Metals. **61**, 156.

LIU, Y., and B.R. PATTERSON, 1993, Metall. Trans. **24A**, 1497.

LÖBERT, P., 1982, Der Maschinenschaden **55**, 136.

LORETTO, M.H., 1984, Electron Beam Analysis of Materials (Chapman and Hall, London).

LOUTHAN, M.R., 1992, in: ASM Handbook, Vol. 10: Materials Characterization (ASM International, Metals Park, OH) p. 299.

LOVELAND, R.P., 1970, Photomicrography — A Comprehensive Treatise (Wiley, New York).

LOZINSKIJ, M.G., 1961, High-Temperature Microscopy (Pergamon Press, Oxford).

LÜCKE, K., 1984, Z. Metallkde. **75**, 948.

LYMAN, C.E., 1990, Scanning Electron Microscopy, X-Ray Microanalysis and Analytical Electron Microscopy (Plenum Press, New York).

LYMAN, C.E., J.I. GOLDSTEIN, A.D. ROMIG, P. ECHLIN, D.C. JOY, D.E. NEWBURY, D. WILLIAMS, J.T. ARMSTRONG, C.E. FIORI and E. LIFSHIN, 1990, Scanning Electron Microscopy, X-Ray Microanalysis and Analytical Electron Microscopy — A Laboratory Workbook (Plenum Press, New York).

MADESKI, A., 1980, Pract. Metallogr. **17**, 598.

MALIES, H.M., 1959, Applied Microscopy and Photo-Micrography (Fountain Press, London).

MALLARD, J.R., 1990, in: Modern Microscopies: Techniques and Applications, eds. P.J. Duke and A.G. Michette (Plenum Press, New York, London) p. 133.
MATHY, H., 1977, Microsc. Acta, Suppl. **3**, 3.
MATTHAEI, E., H. VETTERS and P. MAYR, 1988a, Pract. Metallogr. **25**, 361.
MATTHAEI, E., A. SCHULZ, H. VETTERS and V. SCHWARZMANN, 1988b, in: Microstructural Science, Vol. 16, Metallography of Advanced Materials, eds. H.J. Cialone, G.W.E. Johnson, M.E. Blum and G.F. Vander Voort (ASM International, Metals Park) p. 465.
MC CRONE, W.C., L.B. MC CRONE and J.G. DELLY, 1978, Polarized Light Microscopy (Ann Arbor Science, Ann Arbor).
MCCALL, J., 1973, in: Microstructural Analysis: Tools and Techniques, eds. J.L. McCall and W.M. Mueller (Plenum, New York) p. 93.
MCINTYRE, N.S., ed., 1978, Quantitative Surface Analysis of Materials, ASTM Spec. Tech. Publ. **634** (American Society for Testing and Materials, Philadelphia).
MEDALIA, A.I., 1980, in: Testing and Characterization of Fine Particles, eds. J.K. Beddow and T. Meloy (Heyden, London) p. 66.
MELOY, T.P., 1980, in: Testing and Characterization of Fine Particles, eds. J.K. Beddow and T. Meloy (Heyden, London) p. 1.
METALS HANDBOOK, 1985, 9th ed., Vol. 9: Metallography and Microstructures (American Society for Metals, Metals Park).
METALS HANDBOOK, 1986, 9th ed., Vol. 10: Materials Characterization (American Society for Metals, Metals Park).
MEYER, E., and H. HEINZELMANN, 1992, in: Scanning Tunneling Microscopy, eds. R. Wiesendanger and H.J. Güntherodt (Springer, Berlin) p. 99.
MICHETTE, A.G., 1990, in: Modern Microscopies, Techniques and Applications, eds. P.J. Duke and A.G. Michette (Plenum Press, New York) p. 41.
MICHOT, G., and A. GEORGE, 1982, Script. Metall. **16**, 519.
MIDDLETON, C.J., and D.V. EDMONDS, 1977, Metallography **10**, 55.
MILLER, M.K., and G.D.W. SMITH, 1989, Atom Probe Microanalysis: Principles and Applications (Materials Research Society, Pittsburgh).
MILLER, M.K., A. CEREZO, M.G. HETHERINGTON and G.D.W. SMITH, 1992, Atom Probe Field-Ion Microscopy (Oxford University Press, Oxford).
MINSKY, M., 1988, Scanning **10**, 128.
MITSCHE, R., F. JEGLITSCH and F. GABLER, 1964, Berg- und Hüttenmänn. Monatsh. **109**, 110.
MITSCHE, R., and H. SCHEIDL, 1964, Berg- und Hüttenmänn. Monatsh. **109**, 82.
MITSCHE, R., F. GABLER and F. JEGLITSCH, 1969, in: Handbuch der Mikroskopie in der Technik, Band 3, Teil 2, ed. H. Freund (Umschau Verlag, Frankfurt), p. 269.
MODIN, H., and S. MODIN, 1973, Metallurgical Microscopy (Butterworths, London).
MORNHEIM, A.F., 1977, in: Interpretive Techniques for Microstructural Analysis, eds. J.L. McCall and P.M. French (Plenum, New York) p. 117.
MOTT, B.W., 1956, Micro-Indentation Hardness Testing (Butterworths, London).
MÜCKLICH, F., J. OHSER, S. HARTMANN, M.J. HOFFMANN and G. PETZOW, 1994, in: Tayloring of High Temperature Properties of Silicon Nitride (Kluwer Academic Publ., Dordrecht) p. 73.
MUGGLI, R.Z., and W.C. MC CRONE, 1977, in: Interpretative Techniques for Microstructural Analysis, eds. J.L. McCall and P.M. French (Plenum, New York) p. 127.
MÜLLER, E.W., 1936, Z. Phys. **102**, 734.
MÜLLER, E.W., and T.T. TSONG, 1969, Field Ion Microscopy: Principles and Applications (American Elsevier, New York).
MÜLLER, G., and W.U. KOPP, 1989, Pract. Metallogr. **26**, 640.
MULLINS, W.W., 1961, Phil. Mag. **6**, 1313.
MURPHY, J.C., 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 515.
MÜRRLE, U., H.E. EXNER and D. STÖCKEL, 1980, Metall **34**, 617.
MYKURA, H., 1955, Acta Metall. **3**, 436.

NAKAGAWA, S., 1987. Pract. Metallogr. **24**, 431.
NANEV, C., 1981, Jenaer Rundschau **5**, 219.
NAWRATH, R., and J. SERRA, 1979, Microsc. Acta **82**, 101.
NELSON, J.A., 1989, Pract. Metallogr. **26**, 225.
NEWBURY, D.E., 1977, in: Practical Scanning Electron Microscopy, eds. J.I. Goldstein and H. Yakowitz (Plenum, New York) ch. 4, p. 95.
NEWBURY, D.E., and H. YAKOWITZ, 1977a, in: Practical Scanning Electron Microscopy, eds. J.I. Goldstein and H. Yakowitz (Plenum, New York) ch. 6, p. 211.
NEWBURY, D.E., and H. YAKOWITZ, 1977b, in: Practical Scanning Electron Microscopy, eds. J.I. Goldstein and H. Yakowitz (Plenum, New York) ch. 5, p. 149.
NEWBURY, D.E., 1986, D.C. JOY, P. ECHLIN, C.E. FIORI and J.I. GOLDSTEIN, 1987, Advanced Scanning Electron Microscopy and X-Ray Microanalysis (Plenum Press, New York).
NICHOLSON, W.L., 1978, J. Microsc. **113**, 223.
OATLEY, C.W., 1972, The Scanning Electron Microscope (Cambridge University Press).
ÖCHSNER, H., 1975, Appl. Phys. **8**, 185.
OETTEL, W.O., 1959, Grundlagen der Metallmikroskopie (Akademische Verlagsgesellschaft Geest und Portig, Leipzig).
OHSER, J., and U. LORZ, 1994, Quantitative Gefügeanalyse, Theoretische Grundlagen und Anwendung (Deutscher Verlag für Grundstoffindustrie, Leipzig, Stuttgart).
OHSER, J., and F. MÜCKLICH, 1995, Adv. Appl. Prob. (SGSA) **27**, 384.
OKAZAKI, K., and H. CONRAD, 1972, Trans. Japan. Inst. Metals **13**, 198.
OLIVER, W.C., 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 232.
ONDRACEK, G., and K. SPIELER, 1973, Pract. Metallogr. **10**, 324.
ONDRACEK, G., 1978, in: Quantitative Analysis of Microstructures in Materials Science, Biology and Medicine, Pract. Metallogr. Special Issue 8, ed. J.L. Chermant (Dr. Riederer Verlag, Stuttgart) p. 103.
ONDRACEK, G., 1982, Acta Stereol. **1**, 5.
OPIELKA, H., H. LAHODNY, S. BRAND and E. BISCHOFF, 1990, Pract. Metallogr. **27**, 171.
OPIELKA, H., P. MCKAY and B. RZEHAK–OPIELKA, 1993, Pract. Metallogr. **30**, 399.
ORTNER, H.M., and P. WILHARTITZ, 1991, Mikroch. Acta II, 177.
PAGE, T.F., 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 43.
PANGBORN, R.N., 1992, in: ASM Handbook, Vol. 10: Materials Characterization (ASM International, Metals Park) p. 365.
PAUL, J., and U. MÜRRLE, 1981, Pract. Metallogr. **18**, 418.
PAUL, J., H.E. EXNER and D. MÜLLER-SCHWELLING, 1982, Z. Metallk. **73**, 50.
PAUL, J., and B. BAUER, 1983, Pract. Metall. **20**, 213.
PAYNE, B.O., 1957, Microscope Design and Construction, 2nd Ed. (Cooke, Troughton and Simms, York, UK).
PAYNE, S.M., 1982, in: Electron Microscopy and Analysis 1981, Conf. Series 61, ed. M.J. Goringe (The Institute of Physics, Bristol, London) p. 287.
PECHENYAKOV, I., and R. KOVACHEVA, 1995, Pract. Metallogr. **32**, 17.
PELLOUX, R.M., 1970, in: Applications of Modern Metallography Techniques, ASTM Spec. Techn. Publ. **480** (American Society for Testing and Materials, Philadelphia) p. 127.
PEPPERHOFF, W., and H.H. ETTWIG, 1970, Interferenzschichten-Mikroskopie (Dr. Steinkopf Verlag, Darmstadt).
PETTEGREW, J. W., ed., 1990, NMR: Principles and Applications in Biomedical Research (Springer-Verlag, Berlin).
PETZOW, G., and H.E. EXNER, 1968, in: Handbuch der Mikroskopie in der Technik, Band 3, Teil 1, ed. H. Freund (Umschau Verlag, Frankfurt), p. 37.
PETZOW, G., and E. HORNBOGEN, 1970, Z. Metallk. **61**, 81.
PETZOW, G., and H. KNOSP, 1973, Metallography **6**, 249.
PETZOW, G., and H.E. EXNER, 1975, in: Microstructural Science, eds. P.M. French, R.J. Gray and J.L. McCall (American Elsevier, New York) vol. 3, p. 291.

PETZOW, G., 1978, Metallographic Etching (ASM, Metals Park, OH).
PETZOW, G., 1994, Metallographisches, Keramographisches, Plastographisches Ätzen, (Gebr. Borntraeger, Berlin, Stuttgart).
PFEFFERKORN, G., and R. BLASCHKE, 1982, Beiträge elektronenmikroskopische Direktabbildung von Oberflächen, vol. 15 (Verlag R.A. Remy, Münster). p. 1.
PHILLIPS, V.A., 1971, Modern Metallographic Techniques and their Applications (Wiley-Interscience, New York).
PIVIN, J.C., 1983, J. Mater. Sci. **18**, 1267.
POECH, M.H., and D. RUHR, 1994, Pract. Metallogr. **31**, 70.
POHL, M., and W.G. BURCHARD, 1980a, in: Metallographie, Anschliff- und Dünnschlifftechnik, Pract. Metallogr., Special Issue 11, eds. H.E. Bühler and W.U. Kopp (Dr. Riederer Verlag, Stuttgart) p. 42.
POHL, M., and W.G. BURCHARD, 1980b, Scanning **3**, 251.
POHL, D.W., 1992, in: Scanning Tunneling Microscopy II, eds. R. Wiesendanger and H.J. Güntherodt (Springer, Berlin) p. 233.
POHL, D.W., 1994, Z. Metallkde. **86**, 22.
POKORNY, A., 1980, Pract. Metallogr. **17**, 23.
POMPE, O., 1994, Pract. Metallogr. **31**, 274.
PORTER, D.A., and K.E. EASTERLING, 1981, Phase Transformations in Metals and Alloys (Van Nostrand-Reinhold, Wokingham, New York).
PUSCH, R., 1979, Pract. Metallogr. **16**, 26 and 79.
RANDLE, V., 1993, The Measurement of Grain Boundary Geometry (Institute of Physics Publ., Bristol, Philadelphia).
RAZA, S.M., 1982, Scripta Metall. **16**, 1325.
REIMER, L., and G. PFEFFERKORN, 1977, Rasterelektronenmikroskopie (Springer, Berlin).
REINACHER, G., 1965, Pract. Metallogr. **2**, 45.
RHINES, F.N., K.R. CRAIG and R.T. DE HOFF, 1974, Metallurg. Trans. **5**, 413.
RHINES, F.N., 1976, in: 4th Int. Congr. for Stereology, NBL Spec. Techn. Publ. **431**, eds. E.E. Underwood, R. De Wit and G.A. Moore (National Bureau of Standards, Gaithersburg) p. 233.
RHINES, F.N., and B.R. PATTERSON, 1982, Metall. Trans. **13A**, 985.
RHINES, F.N., 1986, Microstructurology: Behaviour and Microstructure of Materials (Dr. Riederer-Verlag, Stuttgart, now Carl Hanser Verlag, München), also published in Pract. Metallogr. **30**, 216, 287, 322 and 389.
RICHARDS, B.P., and A.D. TRIGG, 1982, in: Electron Microscopy and Analysis 1981, Conf. Series 61, ed. M.J. Goringe (The Institute of Physics, Bristol, London) p. 227.
RIEDLE, J., R. DENNOCHWEILER and H. OPIELKA, 1994, Pract. Metall. **31**, 580.
RIEGGER, H., J.A. PASK and H.E. EXNER, 1980, in: Sintering Processes, ed. G.C. Kuczynski (Plenum, New York) p. 219.
RIVIÈRE, J.C., 1990, Surface Analytical Techniques (Clarendon Press, Oxford).
ROOSZ, A., Z. GACZI and M.K. BAAN, 1980, Metallography **13**, 299.
RÖSCHENBLECK, B., and K. WOLTER, 1979, in: Fortschritte in der Metallographie, Pract. Metallogr., Special Issue 10, eds. F. Jeglitsch and G. Petzow (Dr. Riederer Verlag, Stuttgart) p. 95.
ROSENBERGER, H.E., 1977, in: Interpretative Techniques for Microstructural Analysis, eds. J.L. McCall and P.M. French (Plenum, New York) p. 79.
ROSENCWAIG, A., 1982, Science **218**, 223.
ROST, F.W.D., 1981, Proc. Roy. Microsc. Soc. **16**, 44.
RÜCKERT, M., and L. BJERREGAARD, 1993, Pract. Metallogr. **30**, 579.
RUDOLPH, D., G. SCHMAHL and B. NIEMANN, 1990, in: Modern Microscopies: Techniques and Applications, eds. P.J. Duke and A.G. Michette (Plenum Press, New York, London) p. 59.
RUSS, J.C., 1986, Practical Stereology (Plenum Press, New York).
RUSS, J.C., 1990, Computer-Assisted Microscopy: The Measurement and Analysis of Images (Plenum Press, New York).
RUSS, J.C., 1992, The Image Processing Handbook (CRC Press, Boca Raton, FL).
RUSS, J.C., 1995, J. Comp. Assist. Microsc. **7**, 35.

SAKURAI, T., S. SAKAI and H. W. PICKERING, 1989, Atom Probe Field-Ion Microscopy and its Applications (Academic Press, New York).
SALTYKOV, S. A., 1974, Stereometrische Metallographie (Deutscher Verlag für Grunstoffindustrie, Leipzig).
SAMUELS, L. E., 1971, Metallographic Polishing by Mechanical Methods (Pitman, Melbourne, London).
SARID, D., 1991, Scanning Force Microscopy with Applications to Electric, Magnetic and Atomic Forces (Oxford University Press, New York).
SCHAARWÄCHTER, W., 1968, in: Handbuch der Mikroskopie in der Technik, Band 3, Teil 1, ed. H. Freund (Umschau Verlag, Frankfurt), p. 291.
SCHADE, K. H., 1993, Lichtmikroskopie: Technologie und Anwendung (Verlag Moderne Industrie, Landsberg).
SCHAPER, M., and D. BÖSEL, 1985, Pract. Metallogr. **22**, 197.
SCHATT, W., 1991, Einführung in die Werkstoffwissenschaft (Deutscher Verlag für Grundstoffindustrie, Leipzig).
SCHRÖDER, B., and H. P. HOUGARDY, 1985, Pract. Metallogr. **22**, 587.
SCHRÖDER, J., H. OPIELKA, B. HAUKE and E. ARZT, 1990, Pract. Metallogr. **27**, 17.
SCHUMANN, H., 1990, Metallographie (Deutscher Verlag für Grundstoffindustrie, Leipzig, Stuttgart).
SCHWARZ, H., 1980, Mikroskopie **37**, Suppl., p. 64.
SCHWARZ, H., and H. E. EXNER, 1983, J. Microsc. **129**, 155.
SCHWARZER, R. A., 1981, Microsc. Acta **84**, 51.
SEAH, M. P., 1975, Surf. Sci. **52**, 186.
SEEGER, J., J. KLEIN and H. MECKING, 1980, Pract. Metallogr. **27**, 236.
SEIDMANN, D. N., 1993, Mater. Sci. Eng. **A137**, 57.
SEILER, H., 1968, Abbildung von Oberflächen mit Elektronen, Ionen und Röntgenstrahlen, Hochschultaschenbücher 428/428a (Bibliographisches Institut, Mannheim).
SEMLITSCH, M., and B. BERGMANN, 1969, in: Handbuch der Mikroskopie in der Technik, Band 3, Teil 2, ed. H. Freund (Umschau Verlag, Frankfurt), p. 347.
SERRA, J., 1982, Image Analysis and Mathematical Morphology (Academic, London).
SERRA, J., 1988, Image Analysis and Mathematical Morphology, Vol. 2: Theoretical Advances (Academic Press, London).
SHEPPARD, C. J., 1987, Optica Acta **25**, 269.
SHIGOLEV, P. V., 1974, Electrolytic and Chemical Polishing of Metals (Freund, Tel Aviv).
SIBILIA, J. P., 1988, A Guide to Material Characterization and Chemical Analysis (VCH, Weinheim).
SIEBER, B., 1982, in: Electron Microscopy and Analysis 1981, Conf. Series 61, ed. M.J. Goringe (The Institute of Physics, Bristol, London) p. 223.
SIEGBAHN, M., C. N. NORDLING, A. FAHLMANN, R. NORDBERG, K. HAMRIN, J. HEDMAN, G. JOHANSSON, T. BERHARK, S. E. KARLSSON, I. LINDGREN and B. LINDBERG, 1967, ESCA: Atomic, Molecular and Solid State Structure Studied by Means of Electron Spectroscopy (Almquist and Wicksell, Uppsala).
SIEGENTHALER H., 1992, in: Scanning Tunneling Microscopy, eds. R. Wiesendanger and H.J. Güntherodt (Springer, Berlin) p. 7.
SIGL, L., and H. E. EXNER, 1989, Mater. Sci. Eng. **A108**, 121.
SIGNORELLI, A. J., E. A. LEONE and R. L. CHIN, 1988, in: A Guide to Materials Characterization and Chemical Analysis, ed. J.P. Sibilia (VCH, Weinheim) p. 167.
SIMPSON, I. D., and N. STANDISH, 1977, Metallography **10**, 149 and 443.
SMITH, C. S., 1960, A History of Metallography (University of Chicago Press).
SMITH, G. D. W., 1986, in: Metals Handbook, 9th ed., Vol. 10, Material Characterization (American Society for Metals, Metals Park) p. 583.
SMITH, R., and J. M. WALLS, 1992, in: Methods of Surface Analysis: Techniques and Applications, ed. J. M. Walls (Cambridge University Press, Cambridge) p. 20.
SOMEKH, M. G., 1990, in: Modern Microscopies: Techniques and Applications, eds. P.J. Duke and A.G. Michette (Plenum Press, New York, London) p. 205.
SONNENFELD, R., J. SCHNEIR and P. K. HANSMA, 1990, in: Modern Aspects of Electrochemistry, Vol. 21, eds. R.E. White, O. Bockris and B.E. Conway (Plenum Press, New York) p. 1.
SPIESS, H. J., 1965, Freiberger Forschh. B111, 87.
STAPF, I., U. KOPACZ and P. SCHLÜTER, 1986, Pract. Metallogr. **23**, 297.

STOYAN, D., W. S. KENDALL and J. MECKE, 1987, Stochastic Geometry and Its Applications (Akademie-Verlag, Berlin).
STOYAN, D., 1990, Intern. Statist. Rev. **58**, 227.
STROSCIO, J. A., and R. M. FEENSTRA, 1993, in: Scanning Tunneling Microscopy, eds. J.A. Stroscio and W.J. Kaiser (Academic Press, Boston) p. 96.
STROSCIO, J. A., and W. J. KAISER, eds., 1993, Scanning Tunneling Microscopy (Academic Press, Boston).
SYNGE, E. H., 1928, Phil. Mag. **6**, 356.
TAKAJO, S., 1982, Teilchenwachstum durch Koaleszenz während des Flüssigphasensinterns von Fe-Cu und Cu-Ag, Ph. D. thesis (University of Stuttgart).
TALBOT, H., 1994, Rev. Métallurg. **91**, 211.
TAM, A. C., 1986, Rev. Mod. Phys. **58**, 384.
TANNER, B., 1976, X-ray Diffraction Topography (Pergamon Press, Oxford).
TATE, D. R., 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 183.
TAYLOR, D. S., and G. POLLARD, 1982, Metallography **15**, 225.
TEGART, W. J. McG., 1957, The Electrolytic and Chemical Polishing of Metals in Research and Industry (Pergamon Press, London).
TELLE, R., and G. PETZOW, 1992, in: Materials Science and Technology, Vol. 2A, ed. E. Lifshin (VCH, Weinheim) p. 357.
TENSI, H. M., 1968, in: Handbuch der Mikroskopie in der Technik, Band 3, Teil 2, ed. H. Freund (Umschau Verlag, Frankfurt), p. 1.
TERSOFF, J., and D. R. HAMANN, 1985, Phys. Rev. **B31**, 805.
TERSOFF, J., 1993, in: Scanning Tunneling Microscopy and Spectroscopy, Theory, Techniques and Applications, ed. D.A. Bonnell (VCH Publ., New York) p. 31.
TERSOFF, J., and N. D. LANG, 1993, in: Scanning Tunneling Microscopy, eds. J.A. Stroscio and W.J. Kaiser (Academic Press, Boston) p. 1.
THOMPSON, A. W., 1972, Metallography **5**, 366.
TOMANEK, D., G. OVERNEY, H. MIYAZAKI, S. D. MAHATINI and H. J. GÜNTHERODT, 1989, Phys. Rev. Lett. **63**, 876.
TÖRRÖNEN, K., 1980, Metallography **13**, 329.
TOUSEK, J., 1981, Pract. Metallogr. **18**, 471.
TSONG, T. T., 1990, Atom Probe Field-Ion Microscopy (Cambridge University Press, Cambridge).
TURLEY, D. M., and L. E. SAMUELS, 1981, Metallography **14**, 275.
TURNER, J. N., and D. H. SZAROWSKY, 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 68.
UNDERWOOD, E. E., 1986, in: Applied Metallography, ed. G.F. Vander Voort (Van Nostrand Reinhold, Amsterdam) p. 101.
UNDERWOOD, E. E. (1987), in: Metals Handbook, 9th ed., Vol. 9: Fractography (American Society for Metals, Metals Park) p. 193.
UNDERWOOD, E. E., 1970, Quantitative Stereology (Addison Wesley, Reading, MA).
UNDERWOOD, E. E., 1976, in: 4th Int. Conf. for Stereology, NBL Spec. Techn. Publ. 431, eds. E.E. Underwood, R. De Wit and G.A. Moore (National Bureau of Standards, Gaithersburg) p. 91.
UNDERWOOD, E. E., 1980, in: Testing and Characterization of Fine Particles, eds. J.K. Bedow and T. Meloy (Heyden, London) p. 77.
UNERTL, W., 1993, in: Scanning Tunneling Microscopy and Spectroscopy: Theory, Techniques and Applications, ed. D.A. Bonnell (VCH Publ., New York) 107.
VANDER VOORT, G. F., 1984a, Metallography: Principles and Applications (McGraw Hill, New York).
VANDER VOORT, G. F., 1984b, Metal Progress **127**, March, 31 and 36.
VANDER VOORT, G. F., 1985a, in: Metals Handbook, 9th ed., Vol. 9: Metallography and Microstructures (American Society for Metals, Metals Park) p. 139.
VANDER VOORT, G. F., 1985b, in: Metals Handbook, 9th ed., Vol. 9: Metallography and Microstructures (American Society for Metals, Metals Park) p. 71.

VANDER VOORT, G.F., 1985c, in: Metals Handbook, 9th Edition, Vol. 9, Metallography and Microstructures (American Society for Metals, Metals Park) p. 531.
VANDER VOORT, H.T.M., G.J. BRAKENHOFF and M.W. BAASLAG, 1989, J. Microsc. **153**, 123.
VENABLES, J.A., and G.G. HEMBREE, 1991, in: Inst. Phys. Conf. Ser. No. 119 (Institute of Physics, Bristol) p. 33.
VERHOEVEN, J.D., 1986, in: Metals Handbook, 9th ed., Vol. 10, Materials Characterization (American Society for Metals, Metals Park) p. 490.
VISSCHER, W., 1973, Z. Werkstofftechn. (J. Mater. Techn.) **4**, 320.
VOLBERT, B., 1981, Beiträge elektronenmikroskopische Direktabbildung von Oberflächen, vol. 14 (Verlag R.A. Remy, Münster).
WACHTMAN, J.B., 1993: Characterization of Materials (Butterworth-Heinemann, Stoneham, USA).
WAGNER, C., 1954, J. Electrochem. Soc. **101**, 225.
WAGNER, R., and S.S. BRENNER, 1978, Acta Metall. **26**, 197.
WAGNER, R., 1980, Phys. Blätter **36**, 65.
WAGNER, R., 1982, Field-Ion Microscopy in Materials Science (Springer, Berlin).
WAGNER, R., 1985, Field-Ion Microscopy (Springer, Berlin).
WALLS, J.M. (ed.), 1992, Methods of Surface Analysis (Cambridge University Press, Cambridge).
WANG, S., and L.E. MURR, 1980, Metallography **13**, 203.
WASCHULL, H., 1993, Präparative Metallographie (Deutscher Verlag für Grundstoffindustrie, Leipzig, Stuttgart).
WATTS, J.T., 1982, Proc. Roy. Microsc. Soc. (London) **17**, 192.
WAVER, G., 1973, Z. Werkstofftechn. (J. Mater. Techn.) **4**, 298.
WECHSUNG, R., 1977, Vacuum-Techn. **26**, 227.
WECK, E., and E. LEISTNER, 1986, Metallographic Instruction for Colour Etching by Immersion (In German) Fachbuchreihe Vol. 77 (Deutscher Verlag für Schweisstechnik, Düsseldorf).
WEGMANN, L., 1972, J. Microsc. **96**, 1.
WEIBEL, E.R., 1979, Stereological Methods, vol. 1, Practical Methods for Biological Morphometry (Academic, London).
WEIBEL, E.R., 1980, Stereological Methods, vol. 2, Theoretical Foundations (Academic, London).
WEISSMANN, S., F. BALIBAR and J.F. PETROFF, 1984, Applications of X-Ray Topographic Methods (Plenum Press, New York).
WELLNER, P., 1980, Pract. Metallogr. **17**, 525.
WELLS, O.C., 1993, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon Press, Oxford) p. 423.
WENDLER, B., and B. NEUBAUER, 1979, Pract. Metallogr. **16**, 3.
WENDROCK, H., R. HÜBEL and M. BERTRAM, in: Fortschritte in der Metallographie, Pract. Metallogr., Spec. Issue 26, ed. A. Kneissl (Carl Hanser Verlag, München) 365.
WERLEFORS, T., C. ESKILLSON and S. EKELUND, 1979, Scand. J. Metallurgy **8**, 221.
WETZIG, K., A. MASLOV and J. EDELMANN, 1984, Pract. Metallogr. **21**, 161.
WICKRAMASINGHE, H.K., R.C. BRAY, V. JIPSON, C.F. QUATE, J.R. SAKEDO, 1978, Appl. Phys. Lett. **33**, 923.
WICKRAMASINGHE, H.K., 1983, J. Microsc. **129**, 63.
WICKRAMASINGHE, H.K., 1989, Scanned-Probe Microscopes, Sci. Amer. **261**, 74.
WICKRAMASINGHE, H.K. ed., 1992a, Scanned Probe Microscopies, STM and Beyond (AIP, New York).
WICKRAMASINGHE, H.K., 1992b, in: Scanning Tunneling Microscopy, eds. R. Wiesendanger and H.J. Güntherodt (Springer, Berlin) p. 109.
WICKRAMASINGHE, H. K., 1993, in: Scanning Tunneling Microscopy, eds., J.A. Stroscio and W.J. Kaiser (Academic Press, Boston) p. 77.
WIELAND, H.J., 1993, Pract. Metallogr. **30**, 261.
WIESENDANGER, R., and H.J. GÜNTHERODT, eds., 1992, Scanning Tunneling Microscopy II (Springer, Berlin).
WILSON, T., and C.J.R. SHEPPARD, 1984, Theory and Practice of Scanning Optical Microscopy (Academic Press, London).
WINSCHUH, E., P. VETTERLEIN and M. KRÖNING, 1989, DGZP-Berichte, Vol. 18/1, Deutsche Gesellschaft für Zerstörungsfreie Prüfung, Berlin) p. 186.
WINSCHUH, E., D. SKUDLIK and J. HIMMELMANN, 1993, Pract. Metallogr. **30**, 595.

WOLF, E., 1986, Electron Tunneling Spectroscopy (Oxford University Press, Oxford).
WORCH, H., K. NOCKE, C. BLANK, W. OELSSNER and F. BERTHOLD, 1994, Pract. Metallogr. **31**, 245.
WRIGHT, K., and B. KARLSSON, 1982, J. Microsc. **129**, 185.
WRIGHT, S.I., 1993, J. Comp. Assist. Microsc. **5**, 207.
WU, W.T., I. AYDIN and H.E. BÜHLER, 1982, Pract. Metallogr. **19**, 322 and 347.
YAMAMOTO, K., and A. TAIRA, 1983, J. Microsc. **129**, 49.
YANG, H.S., 1990, Pract. Metallogr. **27**, 539.
YANKOVITH, H., 1970, in: Application of Modern Metallographic Techniques, ASTM Spec. Techn. Publ. **480** (Amer. Soc. Testing Mater., Philadelphia) p. 49.
YOUNG, J.Z., and F. ROBERTS, 1951, Nature **167**, 231.
YUZAWICH, P.M., and C.W. HUGHES, 1978, Pract. Metallogr. **17**, 607.
ZAHN, I., 1992, Pract. Metallogr. **29**, 158.
ZANGWILL, A., 1988, Physics at Surfaces (Cambridge University Press, Cambridge).
ZHAO, X., and Z. CHEN, 1993, J. Comp.-Assist. Microsc. **5**, 257.
ZHOU, J., F. ZHONG, W. SCHMITZ and S. ENGLER, 1993, Pract. Metallogr. **30**, 122.
ZOGG, H., S. WEBER and H. WARLIMONT, 1977, Pract. Metallogr. **14**, 553.

Further reading

Most of the standard journals in the field of physical metallurgy and materials science report frequently on metallographic techniques, surface microscopy, surface analysis, and stereology, with emphasis on the application of these techniques.

Journals

The following journals are specifically devoted to aspects of metallographic sample preparation and microscopic investigation of materials:
Practical Metallography, bilingual English/German (Carl Hanser Verlag, München).
Materials Characterization (Elsevier Science Publishing Co., New York).
Journal of Computer Assisted Microscopy (Plenum Publ. Corp., New York).
Journal of Microscopy (Royal Microscopical Society, London).
Ultramicroscopy (North-Holland Publishing Co., Amsterdam).
Acta Stereologica (International Society for Stereology, Ljubljana).

Books

Ash, E.A., ed., Scanned Image Microscopy (Academic Press, London, 1980).
ASM Handbook, Vol. 10, Materials Characterization (ASM International, Metals Park, 1992).
Beyer, H., Theorie und Praxis der Interferenzmikroskopie (Theory and Practice of Interference Microscopy) in German (Akademische Verlagsgesellschaft Geest und Portig, Leipzig, 1974).
Bonnell, D.A., (ed.), Scanning Tunneling Microscopy: Theory and Practice (VCH Publ., New York, 1990).
Bonnell, D.A., (ed.), Scanning Tunneling Microscopy and Spectroscopy (VCH Publ., New York, 1993).
Briggs, D., and M.P. Seah (eds.), Practical Surface Analysis, Vol. I: Auger and X-Ray Photoelectron Spectroscopy (Wiley, Chichester, 1990).
Briggs, D., and M.P. Seah (eds.), Practical Surface Analysis, Vol. II: Ion and Neutral Spectroscopy (Wiley, Chichester, 1992).
Brümmer, O., J. Heydenreich, K.H. Krebs and H.G. Schneider, Handbuch der Festkörperanalyse mit Röntgenstrahlen (Handbook of Solid State Analysis by Means of X-Rays) in German, (Vieweg, Braunschweig, 1980).
Bühler, H.E., and H.P. Hougardy, Atlas of Interference-Layer Metallography (Deutsche Gesellschaft für Metallkunde, Oberursel, and The Metals Society, London, 1980).
Cahn, R.W., and E. Lifshin, (eds.), Concise Encyclopedia of Materials Characterization (Pergamon Press, Oxford, 1993).

Coster, M., and J. L. Chermant, Précis d'Analyse d'Image (Introduction to the Analysis of Images) in French (Presses du Centre National de la Recherche Scientifique, Paris, 1989).

Czanderna, A. W., ed., Methods of Surface Analysis (Elsevier Scientific, Amsterdam, 1975).

Doyen, G., and D. Drakova, The Physical Principles of STM and AFM Operation (Akademie Verlag, Berlin, 1995).

Duke, P. J., and A. G. Michette (eds.), Modern Microscopies: Techniques and Applications (Plenum Press, New York, London, 1990).

Exner, H. E., and H. P. Hougardy, Quantitative Image Analysis of Microstructures (DGM Informationsgesellschaft-Verlag, Oberursel, 1988).

Feldmann, L. C., and J. W. Mayer, Fundamentals of Surface and Thin Film Analysis (PTR Prentice Hall, Englewood Cliffs, 1986).

Flewitt, P. E., and R. K. Wild, Microstructural Characterization of Metals and Alloys (The Institute of Metals, London, 1985).

Freund, H., ed., Handbuch der Mikroskopie in der Technik, (Handbook of Microscopy in Technology) in German, 8 Volumes (Umschau Verlag, Frankfurt 1960–1969).

Fuchs, E., H. Oppolzer and H. Rehme, Particle Beam Microanalysis: Fundamentals, Methods and Applications (VCH, Weinheim, 1990).

Gabriel, B. L., SEM: A User's Manual for Materials Science (American Society for Metals, Metals Park, 1992).

Galopin, R., and N. F. M. Henry, Microscopic Study of Opaque Materials (W. Heffer and Sons, Cambridge, 1972).

Goldstein, J. I., and H. Yakowitz, eds., Practical Scanning Electron Microscopy (Plenum, New York, 1977).

Goldstein, J. I., D. E. Newbury, P. Echlin, D. C. Joy, A. D. Romig, C. E. Lyman, C. Fiori and E. Lifshin, Scanning Electron Microscopy and X-Ray Microanalysis (Plenum Press, New York, 1994).

Goodhew, P. J., Specimen Preparation in Materials Science (North-Holland, Amsterdam, and American Elsevier, New York, 1973).

Göpel, W., and C. Ziegler, Struktur der Materie: Grundlagen, Mikroskopie und Spektroskopie (Structure of Matter: Fundamentals, Microscopy and Spectroscopy) in German (Teubner Verl. Ges., Stuttgart, Leipzig, 1994).

Güntherodt, H. J., and R. Wiesendanger, (eds.), Scanning Tunneling Microscopy I (Springer-Verlag, Berlin, 1992).

Hamann, C., and M. Hietschold, Raster-Tunnel-Mikroskopie (Akademie Verlag, Berlin, 1991).

Hearle, J. W. S., J. T. Sparrow and P. M. Cross, The Use of Scanning Electron Microscopy (Pergamon Press, Oxford, 1972).

Hornbogen, E., and B. Skrotzki, Werkstoffmikroskopie (Microscopy of Materials) in German (Springer-Verlag, Berlin, 1993).

Lifshin, E. (ed.), Characterization of Materials Part I, Materials Science Engineering, Vol. 2A (VCH Publ., Weinheim, 1992).

Lifshin, E. (ed.), Characterization of Materials, Part II, Materials Science and Engineering, Vol. 2B (VCH Weinheim, 1994).

Loretto, M. H., Electron Beam Analysis of Materials (Chapman and Hall, London, 1988).

Loveland, R. P., Photomicrography: A Comprehensive Treatise (Wiley, New York, 1970).

Lyman, C. E. (with 10 authors), Scanning Electron Microscopy, X-Ray Microanalysis, and Analytical Electron Microscopy: A Laboratory Workbook (Plenum Press, New York, London, 1990).

Metals Handbook, 9th ed., Vol. 9: Metallography and Microstructures (American Society for Metals, Metals Park, 1985).

Metals Handbook, 9th ed., Vol. 10: Materials Characterization (American Society for Metals, Metals Park, 1986).

Metals Handbook, 9th ed., Vol. 12: Fractography (ASM International, Metals Park, OH, 1987).

Modin, H., and S. Modin, Metallurgical Microscopy (Butterworths, London, 1973).

Müller, E. W. and T. T. Tsong, Field Ion Microscopy: Principles and Applications (American Elsevier, New York, 1969).

Mulvey, T., and R. K. Webster, Modern Physical Techniques in Materials Technology (Oxford University Press, 1974).

Newbury, D. E., D. C. Joy, P. Echlin, C. E. Fiori and J. I. Goldstein, Advanced Scanning Electron Microscopy and X-Ray Microanalysis (Plenum Press, New York, 1986).

Pepperhoff, W., and H. H. Ettwig, Interferenzschichtenmikroskopie (Interference Layer Microscopy) in German (Dr. Steinkopf Verlag, Darmstadt, 1970).

Petzow, G., Metallographic Etching (ASM, Metals Park, OH, 1978).

Petzow, G., Metallographisches Ätzen (Metallographic Etching) in German (Gebr. Borntträger Verlag, Stuttgart, 1994).

Phillips, V. A., Modern Metallographic Techniques and their Applications (Wiley, New York, 1971).

Reimer, L. and G. Pfefferkorn, Rasterelektronenmikroskopie (Scanning Electron Microscopy) in German (Springer, Berlin, 1977).

Rhines, F. N., Microstructology: Behaviour and Microstructure of Materials (Carl Hanser Verlag, München, 1986).

Rivière, J. C., Surface Analytical Techniques (Clarendon Press, Oxford, 1990).

Russ, J. C., Practical Stereology (Plenum Press, New York, 1986).

Russ, J. C., Computer-Aided Microscopy: The Measurement and Analysis of Images (Plenum Press, New York, 1990).

Russ, J. C., The Image Analysis Handbook (CRC Press, Boca Raton, 1992).

Saltykov, S. A., Stereometrische Metallographie (Stereometric Metallography) in German (Deutscher Verlag für Grundstoffindustrie, Leipzig, 1974).

Samuels, L. E., Metallographic Polishing by Mechanical Methods (Pitman, Melbourne, London, 1971).

Sarid, D., Scanning Force Microscopy with Applications to Electric, Magnetic and Atomic Forces (Oxford University Press, New York, 1991).

Schatt, W., Einführung in die Werkstoffwissenschaft (Introduction to Materials Science) in German (Deutscher Verlag für Grundstoffindustrie, Leipzig, 1991).

Schumann, H., Metallographie (Metallography) in German (Deutscher Verlag für Grundstoffindustrie, Leipzig, Stuttgart, 1990).

Serra, J., Image Analysis and Mathematical Morphology, Vol.2: Thereotical Advances (Academic Press, London, 1988).

Stroscio, J. A., and W. J. Kaiser, (eds.), Scanning Tunneling Microscopy (Academic Press, Boston, 1993).

Sybilia, J. P., A Guide to Materials Characterization and Chemical Analysis (VCH Publ, Weinheim 1988).

Underwood, E. E., Quantitative Stereology (Addison–Wesley, Reading, MA, 1970).

Vander Voort, G. F., Metallography: Principles and Applications (McGraw Hill, New York, 1984).

Walls, J. M. (ed.), Methods of Surface Analysis (Cambridge University Press, Cambridge, 1992).

Waschul, H., Präparative Metallographie (Preparative Metallography) in German (Deutscher Verlag für Grundstoffindustrie, Leipzig, 1993).

Weibel, E. R., Stereological Methods, vol. 2, Theoretical Foundations (Academic, London, 1980).

Wells, O. C., Scanning Electron Microscopy (McGraw-Hill, New York, 1974).

Wiesendanger, R., and H. J. Güntherodt, (eds.), Scanning Tunneling Microscopy II (Springer-Verlag, Berlin, 1992).

CHAPTER 11

TRANSMISSION ELECTRON MICROSCOPY

M. RÜHLE and M. WILKENS

Max-Planck-Institut für Metallforschung
7000 Stuttgart, FRG

R. W. Cahn and P. Haasen†, eds.
Physical Metallurgy; fourth, revised and enhanced edition

1. Introductory remarks

In the fifties, when transmission electron microscopes became commercially available, their potentialities for enabling applied and fundamental research in physical metallurgy and materials science were realized soon. Within a few years the resolution limit for a direct imaging of structural details of solids — up to that time limited by the light-wavelength in optical microscopy to some fraction of a μm at best — was reduced to about one nm; and in the course of this rapid development the horizon was opened for completely new aspects of research with a spatial resolution to nearly an atomic level. However, at that time it was quickly realized that transmission electron microscopy (TEM) differs from the classical optical microscopy significantly in the sense that TEM, in particular when applied to crystalline specimens, requires a much more profound knowledge of the interaction of the imaging (electron) waves with matter than in the case of optical microscopy. This holds for the imaging of crystal inhomogeneities (lattice defects, precipitates etc.) by TEM via *elastic* interaction of the imaging electrons with the specimen atoms. But it holds also for a local material analysis via various processes of *inelastic* interactions.

The present article attempts to present a brief, and necessarily incomplete, introduction (i) to a number of methods for imaging different kinds of lattice defects and precipitates by *diffraction contrast*, constituted mainly by elastic interaction, and (ii) to problems of the new field of *analytical electron microscopy*, based on inelastic interactions.

Because of the restricted space available, the references quoted are necessarily incomplete and, perhaps, not altogether free of arbitrariness. Other, more special aspects of TEM such as, e.g., Lorentz microscopy of ferromagnetic domains and special aspects of high-voltage electron microscopy (e.g., the critical voltage effect) are not dealt with at all.

In order to facilitate the step into the topics of this chapter the reader is recommended to study first the Appendix, *Elements of kinematical diffraction theory*, which may provide him with the basic knowledge necessary for an understanding of the content of the present chapter. This appendix may also be useful in connection with ch. 12.

2. The instrument

The technology of modern electron microscopes has been developed so far that nowadays commercially available instruments and the sophisticated attachments fulfill practically all requirements necessary for studying crystalline and non-crystalline thin films. Some essential properties of the instrument are sketched in this section.

The *resolution* of an electron microscope is governed by errors of the magnetic lenses, in particular by the *spherical aberration* of the objective lens. The ultimate resolution $r_{\min}$ can be reached for an *optimum objective aperture angle* α_{opt}, with

$$\alpha_{\text{opt}} = A\lambda^{1/4}C_s^{-1/4}, \qquad r_{\min} = B\lambda^{3/4}C_s^{1/4}, \tag{1}$$

where λ is the wavelength of the incident electrons and C_s the constant of spherical

aberration. A and B are constants, their actual values depend on the way in which different contributions to lens errors are combined (HIRSCH *et al.* [1977]).

High-resolution instruments (with accelerating voltages of 200 kV or 300 kV) possess a point-to-point resolution of ≤ 0.25 nm. If ultimate resolution has to be reached then stringent requirements are necessary for the stabilities of the high voltage and the lens currents. The pole piece of the objective lens possesses a narrow gap and small borings, so that in such instruments the specimens can be manipulated only in a very limited range.

Usually, ultimate resolution is not necessary for electron microscopy studies in materials science. For these studies it is more important that the specimen can be shifted and tilted over large ranges and that different signals of scattered electrons and X-rays can be detected in analytical microscopy studies. The lenses of the standard instruments possess pole pieces with larger boring which allow specimen manipulations and the mounting of detectors for analytical purpose. Thereby, the resolution of the instrument is reduced to ~ 0.4 nm. This resolution is, however, sufficient for most TEM studies in materials science.

An electron microscope can be split into three components according to its function: (i) the illumination system composed of the electron gun and the condensor lenses produces a fine electron beam to "illuminate" the specimen; (ii) the objective lens, which immediately follows the specimen, produces the diffraction pattern and a first magnified image of the specimen; (iii) the magnification system produces the final image. In addition to the optical components the microscope encloses also a specimen chamber (specimen handling system) and a recording system (HIRSCH *et al.* [1977] and HREN *et al.* [1979]).

The illumination system provides a beam of adjustable size, intensity and convergence angle at the specimen, resulting in a limited coherency. The term coherence refers to the range of phase differences in the illuminating beam as it approaches the specimen. If the electrons come from a single point source, then all the waves in the incident beam are in phase with each other and the illumination is coherent. On the other hand, if the source of electrons is so large that there is *no* phase relation between the incident waves, the illumination is completely incoherent. In reality, the filament in an electron microscope is somewhere between these two extremes and the incident illumination is defined as partially coherent (SPENCE [1981]).

In TEM the dimensions of the region on the specimen, a, over which the illumination appears coherent is related to the angular aperture of the illumination, α_i, by (GEISS [1979]):

$$a = \frac{\lambda}{2\alpha_i}. \tag{2}$$

Field emission guns possess a very small α_i which results in a large coherence length.

Charged electrons interact strongly with the transmitted specimen, the scattering cross-section σ_e for electrons is rather large compared to the cross-section of neutrons or X-rays (see appendix). Specimen thicknesses are required in the range of 10 nm to 1 μm

depending on the imaging mode and on the voltage of the TEM. Recipes exist for electrochemical thinning of conducting materials (THOMPSON–RUSSELL and EDINGTON [1977]) as well as for ion beam etching of insulators (TIGHE [1976]).

Different kinds of electrons and electromagnetic waves are emitted from a specimen which is irradiated with high-energy electrons. The different waves result from elastic or inelastic scattering processes. Different signals (fig. 1) are used for different imaging modes. Information on the crystal structure and on defects in the specimen can be obtained by studying the elastically scattered electrons (in micrographs inelastically scattered electrons are also present and contribute to the background intensity) whereas investigations of inelastically scattered electrons and of other waves leaving the specimen allow the determination of chemical composition and topology of the specimen surface.

The microscope can be operated in different modes as described briefly in the following. For more details the reader is referred to textbooks on electron microscopy given in the list of references.

In *standard or conventional transmission electron microscopy* (TEM) mode the microscope is operated to form images by bright field (BF), dark field (DF), or lattice image (phase) contrast, see fig. 2. A BF (DF) image is formed when only the direct (one diffracted) beam is used for image formation. The objective aperture prevents all other beams to pass to the recording system. Usually, the specimen is oriented so that the Bragg condition is nearly fulfilled for a set of lattice planes. Then one reflected beam is strongly excited besides the incident beam (see § 4). A weak-beam dark-field (WBDF) image is produced if a weakly excited DF beam is used for imaging (COCKAYNE [1978]).

A *lattice image* is formed by the interference of at least two beams in the image plane of the objective lens. Lattice fringes can be observed if a row of systematic beams (reflected at the lattice planes in question) is used for imaging, while a structure image is formed by using many beams present in a low-indexed Laue zone. Special adjustments of the microscope are required for the formation of the high resolution electron microscope images (SPENCE [1981] and THOMAS and GORINGE [1979]), see § 9.

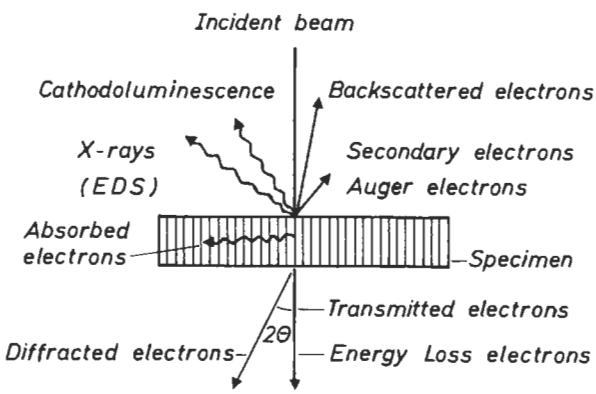


Fig. 1. Electrons and electromagnetic waves emitted from a transmitted specimen as a result of elastic and inelastic scattering or diffraction of the incident electron waves.

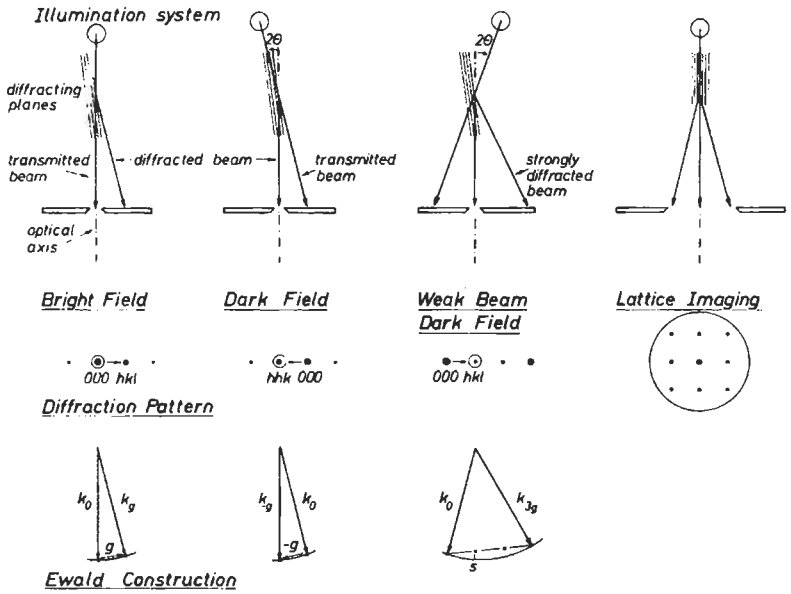


Fig. 2. Ray diagrams (including Ewald sphere construction) for: (a) a conventional two-beam bright-field (BF) image; (b) dark-field (DF) image; (c) weak-beam dark-field (WBDF) image; (d) lattice image. The gun of the electron microscope is tilted by the appropriate angle in going from (a) to (b) or from (a) to (c).

The objective lens produces a *diffraction pattern* of the specimen in its backfocal plane (see fig. 3). The first image of the object is rotated by 180° against the diffraction pattern. The diffraction pattern and the first image are magnified by the subsequent intermediate lenses and projector lenses. The information obtainable from the diffraction pattern is summarized in § 3.

In the *scanning transmission electron microscopy* (STEM) mode, the electron beam is focused as a fine probe on the specimen by the prefield of the objective lens. The beam probe is scanned over the specimen (by scanning coils) and the transmitted intensity is recorded.

The STEM mode is usually applied in materials science for analytical microscopy (see § 10), where the probe is fixed on a selected small area and either the energy losses of the transmitted electrons are studied (EELS) or the X-rays emitted from the specimens are investigated (usually by EDS) for the determination of the chemical composition of the specimen. The use of a very small probe size (< 10 nm) is often limited by a strongly enhanced contamination rate (HREN [1979]).

In the STEM mode backscattered electrons can be collected as well as secondary electrons. These scanning micrographs are similar to those obtained by regular scanning electron microscopy (SEM). Information on the topography of the specimen surface can be obtained.

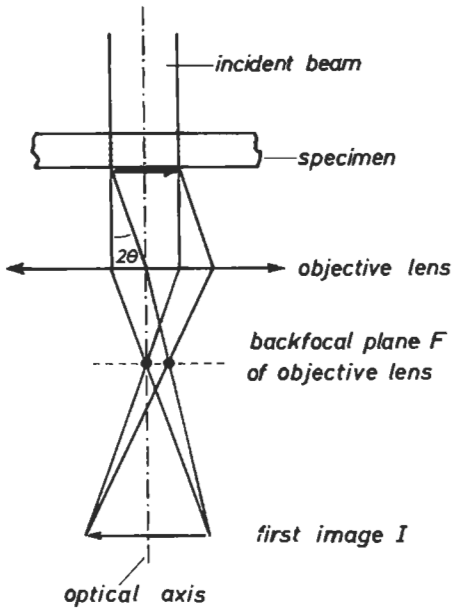


Fig. 3. Ray diagram in the area of the objective lens of an electron microscope. A diffraction pattern is formed in the backfocal plane F of the objective lens, whereas the (first) image I of the object O lies in the plane I. The first image of the object is rotated by 180° relative to the diffraction pattern.

3. *Information from the diffraction pattern*

3.1. Diffraction spot pattern

The diffraction pattern provides crystallographic information on a qualitative phase identification and on the orientation-relations between crystals and the direction of the incoming electron beam. Kikuchi patterns (see § 3.2) can be used for the exact determination of the orientation, whereas the convergent-beam technique allows statements on crystal symmetry and determination of the foil thickness.

The possibilities and the accuracy of analysis of diffraction patterns are discussed in many textbooks on electron microscopy and special papers (e.g., ANDREWS *et al.* [1971]). Important factors for the accuracy are: (i) the shape factor described in the appendix which determines the intensity distribution in the reciprocal space; (ii) instrumental alignment and beam divergence; (iii) specimen perfection; (iv) curvature of the Ewald sphere and its orientation with respect to the foil; (v) double diffraction.

3.1.1. Double diffraction

From the structure factor it follows that certain reciprocal lattice points are not present (zero intensity) for certain crystal symmetries. However, each diffracted beam (within the crystal) can act as an incident beam and can diffract electrons to a reciprocal lattice point forbidden by the structure-factor rules, especially in orientations where

several different reflections are excited simultaneously. For example, in the diamond cubic structure the (002) reflection is not allowed, but this reflection can be excited in a [110] foil via double diffraction: if a $(1\bar{1}1)$ reflection is excited then this reflected beam can act as a primary beam for $(\bar{1}11)$ planes; this gives a total reflection $\boldsymbol{g}_1+\boldsymbol{g}_2=(1\bar{1}1)+(\bar{1}11)=(002)$.

3.1.2. Patterns from ordered crystals

The symmetries of ordered crystal structures are often changed compared to the disordered crystal structure. This results in the appearance of superlattice reflections at positions that are forbidden for the disordered structure. The intensities of the superlattice reflections correspond to the difference between the atomic scattering factors of the different atoms, in contrast to the intensities of the fundamental reflections, which are related to the sum of the scattering factors. As an example, the B2 superlattice is selected. It is based on the bcc structure of the CsCl lattice with one kind (A) of atoms at 000 and other (B) at $\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2}$. The structure factors are (for complete ordering):

$$\begin{aligned} F &= f_A + f_B \quad \text{for } h+k+l = \text{even: fundamental reflections,} \\ F &= f_A - f_B \quad \text{for } h+k+l = \text{odd: superlattice reflections.} \end{aligned} \tag{3}$$

The corresponding diffraction pattern is shown in fig. 4a. In general, superlattices can be identified from their diffraction pattern, either by comparison with structure-factor calculations for different possible superstructures, or by an analytical method which can be viewed as the reversal of the structure-factor calculations (KHACHATURYAN [1974]). In this method the ordered lattice is represented by a superposition of concentration waves. Theoretically, it allows unequivocally the determination of the real space lattice from the complete set of experimentally determined superlattice diffraction vectors.

This method by Khachaturyan can be used to identify superlattices in substitutional (DAS *et al.* [1973]) and interstitial solutions. However, it must be carefully observed that superstructures cannot always be identified by diffraction patterns alone. If, as in many

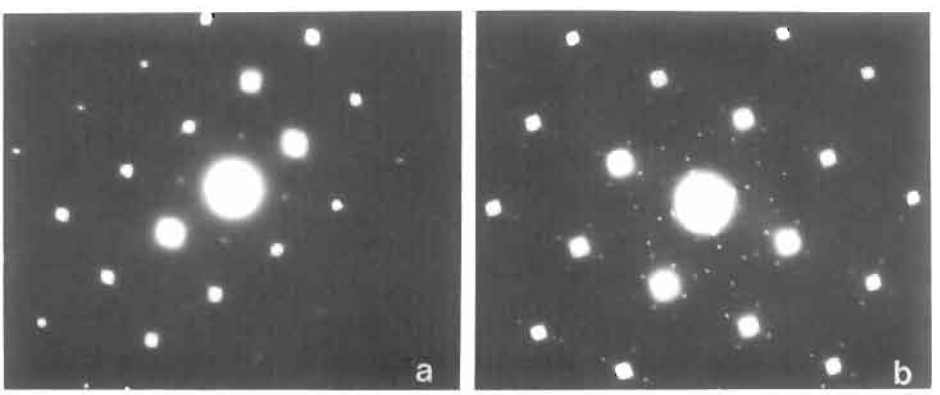


Fig. 4. Diffraction patterns from ordered crystals: (a) diffraction pattern of an ordered β-AlNi alloy (B2 superlattice), [100] zone axis, superlattice reflections are visible; (b) Diffraction pattern of an ordered Ta–O alloy, [110] zone axis, variants of different orientations contribute to the pattern.

disordered cubic alloys, an ordered structure has non-cubic symmetry a number of orientational variants may exist. Figure 4b shows, as an example, a (110) pattern of an ordered Ta–O interstitial phase. The superlattice spots arise from very small ordered domains. If the selected area contributing to the diffraction pattern is large compared to the domain size, then the different domain variants contribute to the diffraction pattern. It is then impossible to determine the superstructure from the diffraction pattern alone. The size, shape, and the number of the different variants of the ordered domains can be observed, however, from different DF images taken with superlattice reflections.

3.2. Kikuchi lines

Electrons can be scattered inelastically by interaction with the atoms of the specimen. Those electrons lose energies in the range of about a hundred eV. The inelastically scattered electrons can subsequently be diffracted coherently when Bragg's law is fulfilled at a suitable set of reflection planes. Since the (primary) inelastic scattering process occurs in different directions, the loci of the different subsequent coherent scattering are cones with semivertex angles of $(90° - \theta)$ to each side. The two cones are bisected by the reflection plane. The lines are, therefore, produced in pairs which are in contrast to the background. A *deficiency line* of less intensity than the background occurs nearer the origin of reciprocal space than its associated *excess line* with intensity above the background. Two conditions must be fulfilled so that these *Kikuchi lines* are observed. One is that the crystal is thick enough, so that enough inelastic scattering processes occur, the second condition is that the crystal must be nearly perfect, especially not bent. The width of the lines indicates the curvature range of the crystal planes over the thickness traversed. If this becomes excessive, the lines disappear into the background as their intensity is spread over a larger angle.

The Kikuchi line pattern can be used to determine the orientation of the crystal with respect to the incoming electron beam. There exist two ways for doing this. Either three pairs of Kikuchi lines have to be indexed and then the orientation can be calculated, or the observed Kikuchi pattern has to be compared to Kikuchi maps (THOMAS and GORINGE [1979]). The orientation of the specimen with respect to the electron beam can be determined with an accuracy of better than 0.3°. The excitation error s (cf. § 4) can also be determined from the relative position of Kikuchi lines compared to the diffraction spots as demonstrated in fig. 5.

3.3. Convergent-beam diffraction

The size of the area giving rise to the diffraction pattern can be substantially reduced if — similarly as in STEM — the electron beam is focused onto the specimen. Under this condition it is practically impossible to maintain the nearly parallel illumination condition. Some convergence of the beam is introduced. As a result, the diffraction "spots" become "discs". For a convergence angle $\alpha < 2\theta_B$ this does not introduce problems since the discs do not overlap (θ_B is the Bragg angle).

In the intensity distribution of those diffraction "discs" much information is available (STEEDS [1979]):

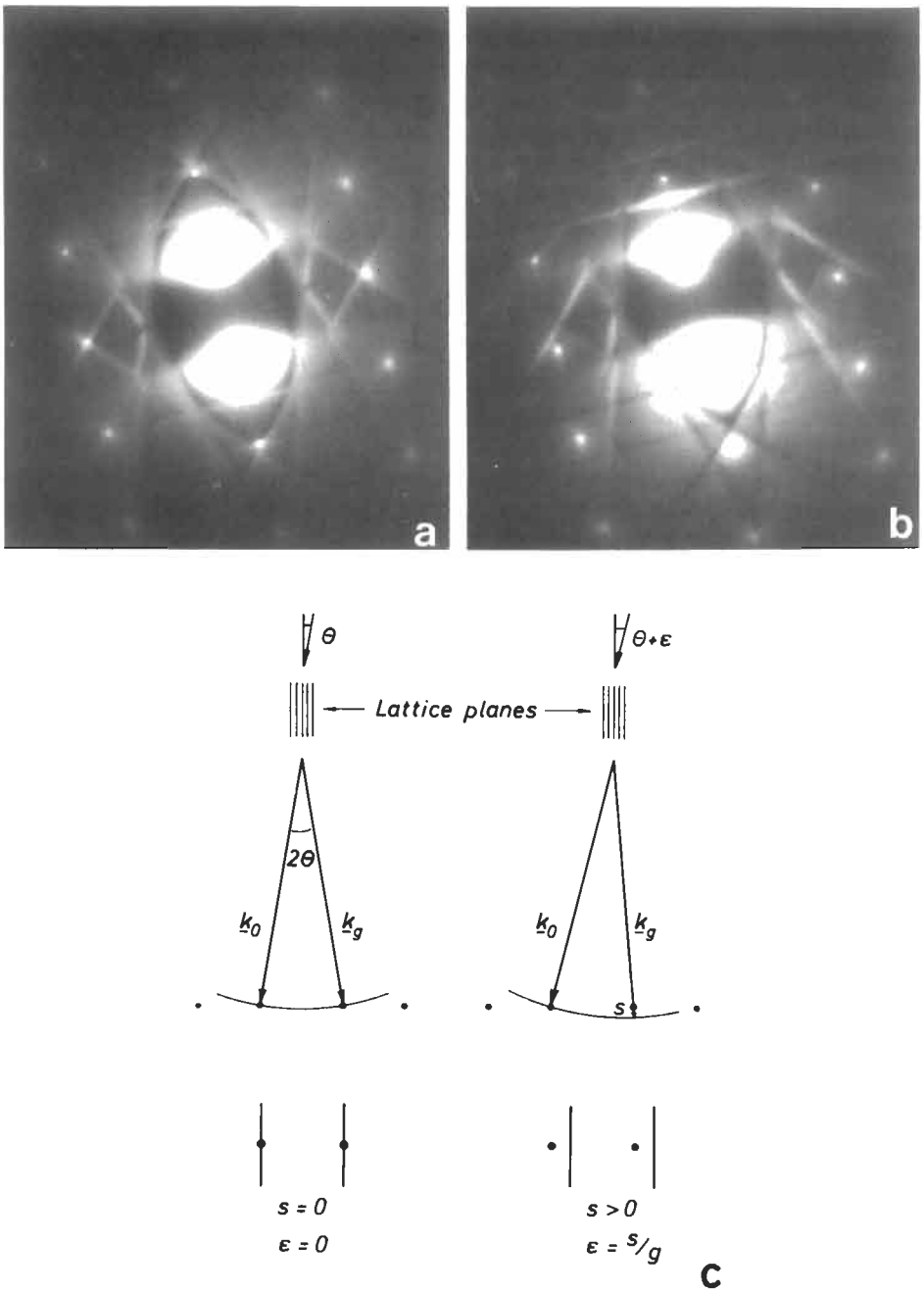


Fig. 5. Kikuchi lines on a diffraction pattern: (a) excitation error $s \approx 0$; (b) $s > 0$. (c) Sketch showing shift in Kikuchi lines produced by a tilt ε.

References: p. 1105.

(i) Each diffraction disc is produced from exactly the same small area (microdiffraction) without any shift introduced by lens errors (HIRSCH *et al.* [1977]). Therefore, these diffraction patterns can be used for the determination of the orientation and the crystal structure of the small illuminated area.

(ii) Each disc contains the intensity diffracted by this same area of crystal for a range of incident orientations. That means that diffraction discs (convergent-beam electron diffraction pattern — CBED) are two-dimensional maps of diffraction intensities as a function of inclination between the incident electrons and a particular crystal direction.

The lattice parameters of a crystal can be determined very accurately from the fine structure of CBED patterns. A change in Bragg angle $\Delta\theta$ results in a change of the positions of lines within a pattern. The method is very strong in measuring *relative* changes in lattice parameters. Absolute numbers do not have a high accuracy.

If the change of lattice parameters of an alloy or compound is directly and unambiguously related to its chemical composition, then the chemical composition may be deduced from lines observable in CBED patterns. As an example, for dual-phase steels the local concentration of carbon could be determined with an accuracy of <0.1 at% (STEEDS [1981]). The spatial resolution (including beam broadening) is governed just by the geometry of the electron probe.

Local strains (STEEDS [1981]) and also the foil thickness (KELLY *et al.* [1975]) can be measured from CBED patterns. Furthermore, CBED patterns allow the determination of the symmetry and of the space group of the crystals (STEEDS [1981]).

3.4. Moiré pattern

Moiré patterns (see fig. 8c for an example) occur from overlapping crystals as in composite films or in two- (or more) phase systems. Two general cases must be considered:

(i) Parallel moiré fringes are formed if two parallel planes of different spacings are reflecting. The lattice spacings ($d_1 = 1/g_1$ and $d_2 = 1/g_2$) differ only in magnitude. Fringes can be observed with distances $d = 1/\Delta g = 1/|g_1 - g_2|$.

(ii) A rotational moiré pattern is formed when planes with equal spacing d, but rotated through an angle α, diffract together. For this situation the moiré spacing d_r is given by $d_r = 1/g\sin\alpha$.

Sometimes moiré patterns have to be distinguished from other periodic defects in the crystals, e.g., a set of parallel dislocations. This can always be done best by imaging the same area with different diffraction vectors. The direction of the moiré lines is usually perpendicular to the diffraction vector.

4. *Theory of diffraction contrast*

4.1. Introduction

Diffraction contrast of defects in crystalline specimens is mainly a problem of high-energy electron diffraction in non-perfect crystals. It will turn out later that, by intrinsic

reasons, for the imaging of lattice defects (dislocations, stacking faults etc.) an electron–optical resolution of about 1 nm is in general sufficient. Accordingly, lens errors of the objective lens can be neglected. We assume that the objective lens is focused onto the lower specimen surface through which the electrons leave the specimen. Then it suffices for the purpose of this section to calculate the electron wave function at the lower specimen surface since the electron–optical imaging system is assumed to be perfect.

In a first part, we deal with the diffraction in a perfect crystal specimen. In a second part, it will be outlined how the diffraction theory must be extended for specimens containing lattice defects. Electron refraction effects, due to the mean inner potential of the crystal, are neglected throughout.

4.2. Specimen, reciprocal lattice and excitation error

We consider a specimen of constant thickness t of the order of 100 nm. A Cartesian coordinate system is introduced with its origin in the upper specimen surface and with the z-axis (unit vector $\boldsymbol{e}_z$) perpendicular to the specimen plane and pointing downwards. The lateral dimensions L_x, L_y of the specimen are orders of magnitude larger than t. According to the appendix, this means that the intensity distributions $|F(\boldsymbol{\kappa})|^2$ at the reciprocal lattice points $\boldsymbol{g}$ are rod-shaped or spike-shaped with the spike axis parallel to $\boldsymbol{e}_z$ and a spike length of the order of $1/t$. This is indicated in fig. 6, where also the wavevector $\boldsymbol{k}_0$ of the incident wave and the Ewald sphere are inserted. Note that this figure is not to scale: assuming a modulus of $(0.2\ \text{nm})^{-1}$ for the low-order diffraction vector $\boldsymbol{g}_1$, and $k_0 = 1/\lambda = (3.7\ \text{pm})^{-1}$ (100 keV electrons), we have $k_0/g_1 = 54$. Having this in mind, one realizes that near the low-order reflections the Ewald sphere is so flat that, if the sphere cuts through the spike of one of the reciprocal lattice points, this will happen also for the spikes of other adjacent reciprocal lattice points. Accordingly, in high-energy electron diffraction, in general several Bragg reflections are excited

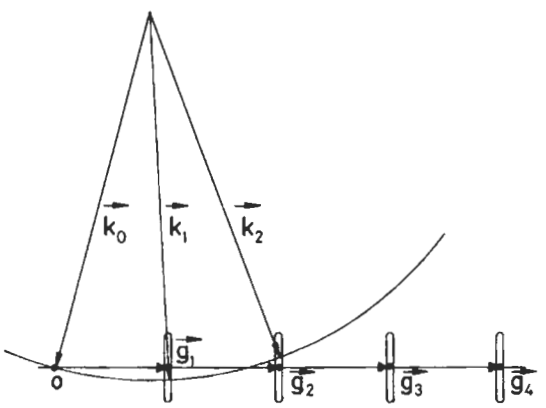


Fig. 6. The Ewald sphere and the reciprocal space. The intensity distributions $|F(\boldsymbol{\kappa})|^2$ around the reciprocal lattice points are spike-shaped with spike lengths inversely proportional to the specimen thickness. The excitation error s is positive for $\boldsymbol{g} = \boldsymbol{g}_1$ and negative for the other $\boldsymbol{g}$.

References: p. 1105.

simultaneously; besides the primary wave a number of diffracted plane waves leave the crystal. However, in general the specimen can be orientated so that only one diffracted wave is strong. Then, we speak about a *two-beam case* (primary plus diffracted beam). Images taken with either of these two beams are called *strong-beam images*. On the other hand, dark-field images taken with an extremely weakly excited beam are called *weak-beam images*.

The direction of incidence of the primary beam with respect to the specimen can be characterized by the *excitation error* s_g, or simply s, which is given by the distance between the reciprocal lattice point $\boldsymbol{g}$ considered, and that point on the intensity spike where the Ewald sphere cuts through it. s is positive (negative) if the point $\boldsymbol{g}$ lies inside (outside) the Ewald sphere. If θ is the glancing angle between the direction of $\boldsymbol{k}_0$ and the lattice planes belonging to $\boldsymbol{g}$ and if θ_g is the corresponding Bragg angle, then

$$s = g\Delta\theta, \quad \text{with} \quad \Delta\theta = \theta - \theta_g. \tag{4}$$

4.3. Outline of the dynamical diffraction theory

The essential points of the dynamical diffraction theory, in contrast to the kinematical theory as outlined in the appendix, are that diffractions between all plane waves involved have to be treated as being equivalent and that, in the absence of absorption effects, conservation of intensity must be fulfilled. There are several ways for formulating the dynamical theory for a perfect crystal. Here we refer very briefly to the classical way first developed by BETHE [1928]. Another way will be outlined later when crystals with lattice defects are considered.

Following Bethe, the time-independent Schrödinger equation with a periodic potential is solved in terms of independent "eigen solutions" or Bloch waves ψ_B. Generally, Bloch waves can be written in the form

$$\psi_B(\boldsymbol{r}) = b(\boldsymbol{r}) \exp(2\pi i \boldsymbol{K} \cdot \boldsymbol{r}), \tag{5}$$

where the wavevector $\boldsymbol{K}$ has to be found for a given electron energy and a given direction of $\boldsymbol{K}$. The function $b(\boldsymbol{r})$ is periodic with the crystal periodicity. In the two-beam case, to which we restrict ourselves, there are two independent Bloch waves belonging to a given tangential component of $\boldsymbol{K}$ parallel to the electron entrance surface of the crystal. We assume that the two-beam diffraction vector $\boldsymbol{g}$ is perpendicular to $\boldsymbol{e}_z$ and that the origin of the coordinate system lies on a reflecting lattice plane. Before the solutions are presented we introduce some terms which will be used repeatedly later. One important diffraction parameter is the *extinction length* ξ_g,

$$\xi_g = \frac{k_0 V_e \pi}{F_s(g)}, \tag{6}$$

where V_e is the volume of the elementary cell and F_s is the structure amplitude, see appendix. For low-order reflections ξ_g is of the order of some 10 nm. Besides ξ_g we use the "wavenumber" notations

$$1/\xi_g = \sigma_0, \quad \sigma \equiv \sigma(w) = \sigma_0 \cdot \left(1 + w^2\right)^{1/2}, \quad w = s\xi_g. \tag{7}$$

w is the "normalized" excitation error and σ is the reciprocal of the effective extinction length $\xi_{g,\mathrm{eff}}$ as a function of w. The wanted solutions, expressed in these terms, are (e.g., HIRSCH *et al.* [1977]):

$$\left.\begin{aligned} \psi_B^{(1)} &= c_0^{(1)} \exp\left[2\pi i\left(\boldsymbol{K}_0^{(1)} \cdot \boldsymbol{r}\right)\right], \\ \psi_B^{(2)} &= c_0^{(2)} \exp\left[2\pi i\left(\boldsymbol{K}_0^{(2)} \cdot \boldsymbol{r}\right)\right]. \end{aligned}\right\} \tag{8}$$

$$\left.\begin{aligned} c_0^{(1)} &= c_g^{(2)} = \frac{1}{\sqrt{2}}\left[1 - \frac{w}{\left(1+w^2\right)^{1/2}}\right]^{1/2}, \\ c_g^{(1)} &= -c_0^{(2)} = \frac{1}{\sqrt{2}}\left[1 + \frac{w}{\left(\sqrt{1+w^2}\right)^{1/2}}\right]^{1/2} \\ \boldsymbol{K}_0^{(l)} &= \boldsymbol{k}_0 + \Delta\boldsymbol{k}^{(l)}, \qquad \boldsymbol{K}_g^{(l)} = \boldsymbol{K}_0^{(l)} + \boldsymbol{g}, \quad l = 1, 2, \\ \Delta\boldsymbol{k}^{(l)} &= \tfrac{1}{2}(s+\sigma)\boldsymbol{e}_z, \qquad \Delta\boldsymbol{k}^{(2)} = \tfrac{1}{2}(s-\sigma)\boldsymbol{e}_z. \end{aligned}\right\} \tag{9}$$

The $\Delta\boldsymbol{k}^{(l)}$ are the "eigen values" and the $c_n^{(l)}$ ($n = 0$, g) are the components of the "eigen vectors" of the Schrödinger equation. The relation between eqs. (5) and (8) is easy to realize if the phase factor $\exp(2\pi\boldsymbol{K}_0^{(l)})$ r) is put in front of the two terms of $\psi_B^{(1)}$ in eq. (8).

If the ψ_B are known, a linear combination of the $\psi_B^{(l)}$,

$$\psi = \varphi^{(1)}\psi_B^{(1)} + \varphi^{(2)}\psi_B^{(2)}, \tag{10}$$

has to be found which satisfies the boundary conditions at the specimen surface at z = 0 where both partial waves with the wave vector $\boldsymbol{K}_0$ have to add up to unity so that at $z = 0$ they fit to the incident primary wave:

$$\psi_0 - \phi_0 \exp\left[2\pi i\left(\boldsymbol{k}_0 \cdot \boldsymbol{r}\right)\right], \qquad \phi_0 = 1 \text{ for } \; z < 0. \tag{11}$$

Further, both partial waves with wave vectors $\boldsymbol{K}_g$ have to cancel at $z = 0$. In our case, this leads to

$$\varphi^{(1)} = c_0^{(1)}, \qquad \varphi^{(2)} = c_0^{(2)}. \tag{12}$$

At the lower specimen surface at $z = t$, the Bloch waves decompose into their partial waves, and those having the same subscript, 0 or g, interfere and constitute a plane wave below the specimen. Thus, we obtain for $z > t$:

$$\psi_0 = \phi_0(t) \exp\left[2\pi i\left(\boldsymbol{k}_0 \cdot \boldsymbol{r}\right)\right], \quad \psi_g = \phi_g(t) \exp\left[2\pi i\left(\boldsymbol{k}_g \cdot \boldsymbol{r}\right)\right],$$

References: p. 1105.

with

$$\boldsymbol{k}_g = \boldsymbol{k}_0 + s\boldsymbol{e}_z + \boldsymbol{g}, \quad |\boldsymbol{k}_g| = |\boldsymbol{k}_0|, \tag{13}$$

and

$$\phi_0(t) = \cos \pi\sigma t - \mathrm{i}\frac{\mathrm{w}}{\left(1+\mathrm{w}^2\right)^{1/2}} \sin \pi\sigma t, \quad \phi(t) = \mathrm{i}\frac{1}{\left(1+w^2\right)^{1/2}} \sin \pi\sigma t. \tag{14}$$

4.4. Normal and anomalous absorption

High-energy electrons, when passing through a crystal, experience, besides the elastic scattering at the atom potential, also inelastic scattering, e.g., by interaction with the thermal vibration of the crystal atoms (phonons) or with the crystal electrons (plasmons, inner-shell excitation). By these events energy is transferred between high-energy electrons and the crystal, which leads to a loss of coherency of the wave fields of the elastically and inelastically scattered electrons. This effect can formally be described as an "absorption" (although, of course, the inelastically scattered electrons are not really absorbed). This "absorption" is different for different Bloch waves, depending on the high-energy electron density distribution $|\psi_B|^2$ with respect to the atom positions (HASHIMOTO *et al.* [1960]). From eqs. (8) and (9) it is easy to derive that $|\psi_B^{(1)}|^2$ reveals a maximum electron density at the lattice planes characterized by $\boldsymbol{g}$ and a minimum in between; for $|\psi_B^{(2)}|^2$ the reverse is true:

$$\begin{aligned} \left|\psi_B^{(1)}\right|^2 &= 1 + \frac{1}{\left(1+w^2\right)^{1/2}} \cos 2\pi(\boldsymbol{g}\cdot\boldsymbol{r}), \\ \left|\psi_B^{(2)}\right|^2 &= 1 - \frac{1}{\left(1+w^2\right)^{1/2}} \cos 2\pi(\boldsymbol{g}\cdot\boldsymbol{r}). \end{aligned} \tag{15}$$

Accordingly, $\psi_B^{(1)}$ interacts more strongly with the crystal atoms and is thus more strongly absorbed than $\psi_B^{(2)}$. We subdivide the absorption into *normal* absorption, which a high-energy electron would experience when travelling through the crystal far away from any Bragg reflection, and *anomalous* absorption, which takes the structure of $|\psi_B^{(j)}|^2$ into account. Normal absorption is accounted for by adding a common factor $\exp(-\mu_0 z/2)$, where μ_0 is usually expressed by the "normal" absorption length ξ_0', with $\mu_0 = 2\pi/\xi_0'$. The anomalous absorption is introduced by adding a positive imaginary part to $1/\xi_g$,

$$1/\xi_g \to \frac{1}{\xi_g} + \frac{1}{\xi_g'}, \tag{16}$$

where typically $\xi_0' \approx (10\text{–}20)\xi_g$ (RADI [1970]). Accurate values of ξ_0' are not well known since they depend, e.g., on the size of the objective aperture. Normal absorption acts only as a scaling factor, independently of the actual diffraction conditions. Therefore, accurate values of ξ_0' are not required, and often $\xi_0' = \xi_g'$ is used for intensity calculations.

Working through the abbreviations introduced in eq. (7) leads in first order of ξ_g/ξ'_g to the substitution

$$\sigma \to \sigma + i\sigma', \quad \sigma' = \frac{1}{\xi'_g\left(1+w^2\right)^{1/2}}. \tag{17}$$

This finally gives:

$$\psi_B^{(l)} \propto \exp\left[-\tfrac{1}{2}\mu^{(l)}z\right], \quad l = 1, 2,$$

with

$$\mu^{(1)} = \mu_0 + 2\pi\sigma', \quad \mu^2 = \mu_0 - 2\pi\sigma', \tag{18}$$

i.e., the absorption of $\psi_B^{(1)}(\psi_B^{(2)})$is enhanced (reduced). ($\xi_g$ is usually left uncorrected where entering into the $c_n^{(l)}$.)

4.5. Dynamical bright-field and dark-field intensities

The same substitution as eq. (17) must be applied to the argument $\pi\sigma\tau$ in eq. (14). After some algebraic operations, we find for the bright-field (I_0) and dark-field (I_g) intensities:

$$\begin{aligned} I_0 &= \exp\left(-\mu_0 t\right)\left[\cosh 2\pi\sigma' t + \frac{w\sinh 2\pi\sigma' t}{\left(1+w^2\right)^{1/2}} - \frac{\sinh^2\pi\sigma' t + \sinh^2\pi\sigma t}{1+w^2}\right], \\ I_g &= \exp\left(-\mu_0 t\right)\frac{\sinh^2\pi\sigma' t + \sin^2\pi\sigma t}{1+w^2}. \end{aligned} \tag{19}$$

In fig. 7a, b we have plotted $I_0(t)$ and $I_g(t)$ as a function of t for $w=0$ and $w=1$ (with $\xi'_0=\xi'_g$, $\xi'_g=10\xi_g$). We see that the total intensity oscillates between I_0 and I_g, a maximum in I_0 corresponds to a minimum in I_g and vice-versa ("Pendellösung"). These oscillations are due to the fact that the z-components of the wavevectors of the Bloch waves $\psi_B^{(1)}$ and $\psi_B^{(2)}$ differ by σ, which leads to a beating of the partial waves constituting ψ_0 and ψ_g, respectively. The full oscillation period $\Delta t=\xi_{g,\mathrm{eff}}=\sigma^{-1}$ and the oscillation amplitude decreases with increasing $|w|$. Further, the oscillations are damped with increasing t which is a consequence of the anomalous strong absorption of $\psi_B^{(1)}$; if $\psi_B^{(1)}$ has decayed, a beating between partial waves is no longer possible. On wedge-shaped specimens the Pendellösung-oscillations give rise to "thickness fringes" or "thickness contours" along lines of constant specimen thickness. An example is shown in fig. 7c, d.

In fig. 8a, b we show I_0 and I_g as a function of w for some values of t. Both terms show oscillations with varying w ("bend contours"), which decrease in amplitude with increasing t. Whereas I_g is symmetric in w, this is not true for I_0: the maximum in I_0 (i.e., best transmittivity!) occurs at $w>0$. A complicated system of bend contours is visible in fig. 8c.

References: p. 1105.

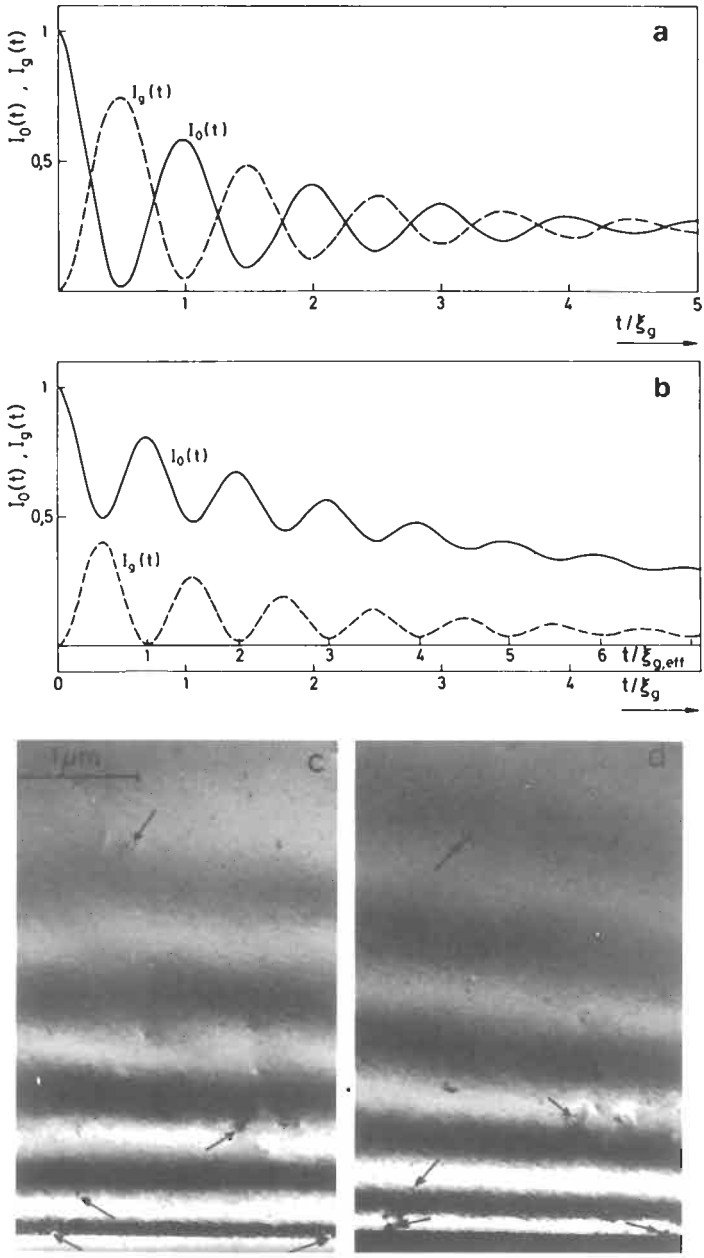


Fig. 7. (a, b) Two-beam thickness contours calculated for $\xi_g' = \xi_0' = 10\xi_g$ as a function of specimen thickness t: solid line = bright-field intensity I_0, dashed line = dark field intensity I_g, with (a) excitation error $w = 0$, and (b) $w = 1$. (c, d) Thickness contours in copper, $g = (111)$, $w \approx 0$, in (c) bright fields and (d) dark field. Arrows indicate equivalent points on the images. The lower arrows point to the specimen edge.

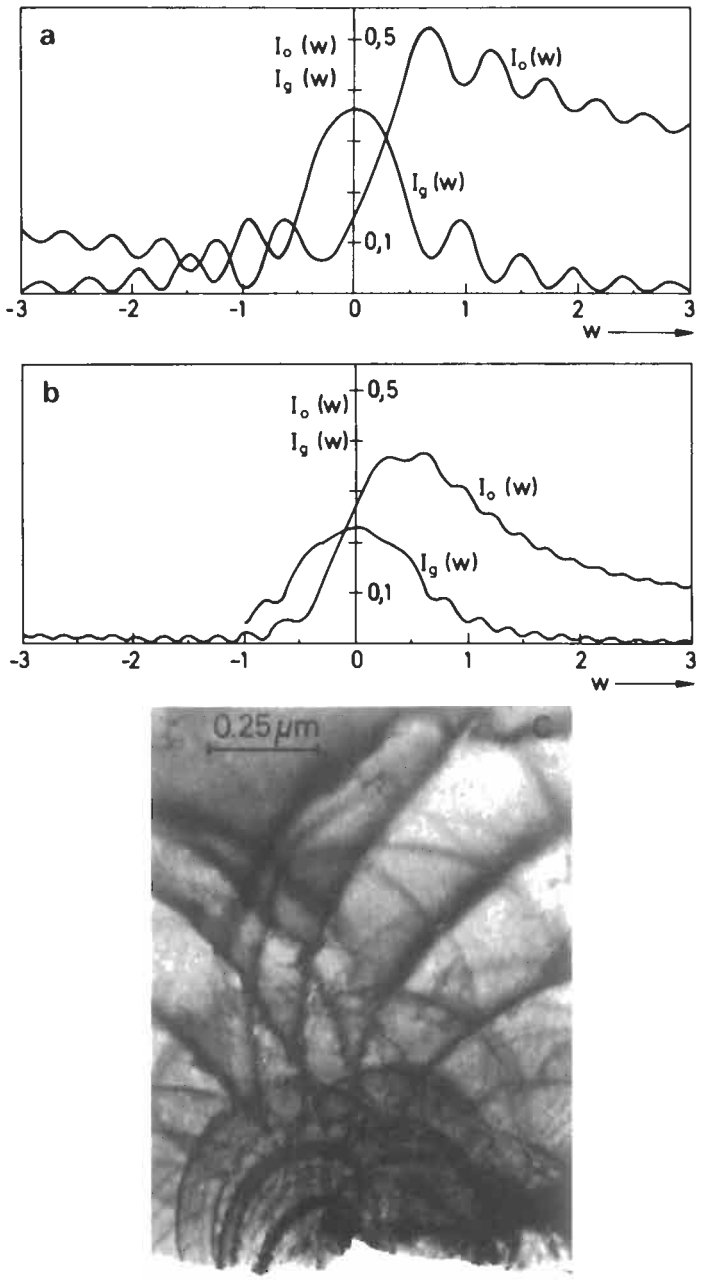


Fig. 8. (a, b) Two-beam bend contours calculated for $\xi_g' = \xi_0' = 10\xi_g$ as a function of w, with (a) $t = 2.5\xi_g$, and (b) $t = 5\xi_g$. (c) Bright field bend contours in copper. The dark lines correspond to specimen orientations where a certain set of lattice planes is in Bragg orientation. The moire patterns in the centre of the figure are due to a thin oxide layer on the specimen surface.

References: p. 1105.

4.6. The column approximation

We consider in fig. 9 the point C at the lower specimen surface. The lines through A and C and through B and C are parallel to the wavevectors k_0 and kg of the primary and diffracted wave, respectively. Triangle ABC is known as the *Takagi-triangle*. It has been shown (TAKAGI [1962], HOWIE and BASINSKI [1968]) that the wave functions ψ_0 and ψ_g at point C are determined in very good approximation only by the specimen parameters within this triangle. With specimen parameters we mean local specimen thickness, local orientation, i.e., local excitation error, and local displacements of atoms from their positions in a perfect lattice. For high-energy electrons the Takagi triangle is very narrow: The angle at C is $2\theta_g$ with the Bragg angle θ_g being about 1° or less. Taking $t = 100$ nm, it turns out that the distance A–B is some nm at most. Accordingly, if variations of the specimen parameters over such a distance are sufficiently small, then ψ_0 and ψ_g at point C can be calculated under the assumption that the specimen parameters along the "column" above C, i.e., along the z-axis in our case, are the same for the entire specimen. This means that for calculation of ψ_0 and ψ_g the "column" coordinate is the only intrinsic variable, whereas the lateral coordinates (x, y in our case) act only as parameters. This is the *column approximation* introduced by HIRSCH *et al.* [1960].

4.7. Diffraction at imperfect crystals

4.7.1. The displacement field

Lattice defects cause displacements $\boldsymbol{R}(\boldsymbol{r}_n) \equiv \boldsymbol{R}n$ of the atoms from their positions rn in the defect-free reference lattice*. In general, $|\boldsymbol{R}_n|$ is of the order of, or smaller than, the interatomic distances. Here we assume for simplicity that the displacements do not vary appreciably over the atom positions within the elementary cell so that the structure-

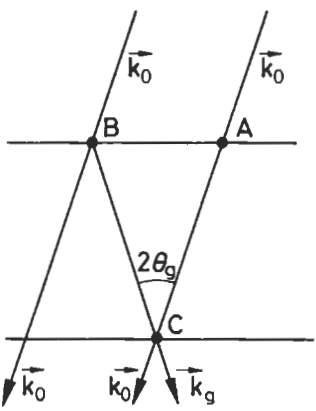


Fig. 9. The Takagi triangle. The wave functions ψ_0 and ψ_g at point C are determined by the specimen parameters within triangle ABC.

* Regarding the notation $\boldsymbol{R}$, cf. the footnote in § 5.2.

amplitude F_s is unaffected by the lattice defects. Further we confine ourselves to cases where $\boldsymbol{R}$ is a continuous function in space, giving thus rise to "strain contrast". Contrast due to stacking faults and other planar defects will be considered briefly in § 8.

4.7.2. The kinematical approach

A first insight in the intensities $I_0(x, y)$ and $I_g(x, y)$ for perfect and imperfect crystals may be obtained by the kinematical diffraction theory outlined in the appendix. We start from eq. (A.46), where we neglect, as usual, the term $\boldsymbol{s}\bullet\boldsymbol{R}$ in the exponent. Thus, the kinematical diffraction amplitude $F(\boldsymbol{s})$ follows as:

$$F(\mathbf{s}) = \frac{F_s(\boldsymbol{g})}{V_c} \int_{V_c} \exp\{-2\pi i[\boldsymbol{g} \cdot \boldsymbol{R}(\boldsymbol{r}) + \boldsymbol{s} \cdot \boldsymbol{r}]\} d^3\boldsymbol{r}, \tag{20}$$

where the integration runs over the specimen volume $V_c = L_x L_y t$ with $t \ll L_x, L_y$. The diffracted intensity follows by inserting eq. (20) into the integral of eq. (A.35). We assume that $\boldsymbol{g}$ is perpendicular to the z-axis and make use of the fact that — because of the smallness of the Bragg angle — the Ewald sphere near $\boldsymbol{g}$ is (nearly) parallel to the x–y plane. Thus, the differential df_E in eq. (A.35) can be replaced by $ds_1 ds_2$ [the coordinate s_3 in the appendix is identical with the excitation error s in eq. (4)]. Now having in mind that, within the column approximation, the terms x, y in $\boldsymbol{R}(x, y; z)$ are not coordinates but only parameters, it turns out that all integrations in eq. (A.35) with $F(\boldsymbol{s})$ according to eq. (20) can be carried out, except those over the z-axis. After dividing the result by the specimen area we get for the diffracted intensity per unit area:

$$I_g(x, y) = \left(\frac{F_s}{kV_c}\right)^2 \left| \int_0^t \exp\{-2\pi i[\boldsymbol{g} \cdot \boldsymbol{R}(z) + sz]\} dz \right|^2. \tag{21}$$

Thus, with eq. (6), the amplitude ϕ_g follows as:

$$\phi_g(x, y; t) = \frac{i\pi}{\xi_g} \phi_0 \int_0^t \exp\{-2\pi i[\boldsymbol{g} \cdot \boldsymbol{R}(x, y; z) + sz]\} dz. \tag{22}$$

Here we have added the amplitude ϕ_0, of the primary wave, which is constant ($|\phi_0| = 1$) within the kinematical approach. Equation (22) is the *kinematical contrast integral* derived by HIRSCH *et al.* [1960] who have shown that the imaginary unit i should be added.

The argument in the exponent of eq. (22) can be re-interpreted as

$$\boldsymbol{g} \cdot \boldsymbol{R}(z) + sz = \int_0^z s_\ell(\zeta) d\zeta, \tag{23a}$$

with

$$s_\ell(z) = s + \frac{d(\boldsymbol{g} \cdot \boldsymbol{R}(z))}{dz} \tag{23b}$$

being the *local* excitation error experienced by the primary wave as due to the *local* orientation of the reflecting lattice planes. Diffraction from ϕ_0 into ϕ_g is especially strong

where $s_i(z)$ is small (COCKAYNE *et al.* [1969]). For $\boldsymbol{R}=0$ (perfect crystal) eq. (22) can be integrated and yields, with $|\phi_0|=1$,

$$I_g = \frac{1}{\xi_g^2}\frac{\sin^2 \pi t s}{s^2}, \tag{24}$$

to which the "dynamical" solution eq. (14) degenerates if absorption is neglected and $|s\xi_g| = |w| \gg 1$. Thus, the range of applicability of eqs. (21)–(24) is restricted to large values of $|w|$, for which $I_g \ll 1$, for instance to the weak-beam imaging mode (COCKAYNE *et al.* [1969], and COCKAYNE [1978]), where the diffracted beam taken for the image formation is excited with $|w| \gg 1$.

Equation (22) can be rearranged in order to separate the defect-induced part of the contrast integral. By partial integration we get

$$\phi_g = \phi_g' + \phi_g'', \tag{25}$$

with

$$\phi_g' = \frac{i\pi}{\xi_g}\phi_0 \left.\frac{\exp\{-2\pi i[\boldsymbol{g}\cdot\boldsymbol{R}(z) + sz]\}}{-2\pi i s}\right|_{z=0}^{t}, \tag{26a}$$

$$\phi_g'' = -\frac{i\pi}{\xi_g}\phi_0 \frac{1}{s}\int_0^t \frac{d(\boldsymbol{g}\cdot\boldsymbol{R}(z))}{dz}\exp\{-2\pi i[\boldsymbol{g}\cdot\boldsymbol{R}(z) + sz]\}dz. \tag{26b}$$

If $\boldsymbol{g}\bullet\boldsymbol{R}(t)$ and $\boldsymbol{g}\bullet\boldsymbol{R}(0)$ are neglected in eq. (26a), then ϕ_g' represents the "background amplitude" with $|\phi_g'|^2$ identical to (24). ϕ_g'' is the defect-induced part of ϕ_g. Since the integrand of eq. (26b) is proportional to the displacement derivative $d(\boldsymbol{g}\bullet\boldsymbol{R}(z))/dz$, the integration limits can in general be extended to $\pm\infty$. In the dark-field intensity $I_g = |\phi_g|^2$ the interference term between ϕ_g' and ϕ_g'' is often neglected (cf. e.g., HIRSCH *et al.* [1960] and WILKENS and HORNBOGEN [1964]):

$$I_g(x,y) = |\phi_g'|^2 + |\phi_g''(x,y)|^2. \tag{27}$$

Then adjusting to conservation of intensity gives the bright-field intensity:

$$I_0(x,y) = \left(1 - |\phi_g'|^2\right) - |\phi_g''(x,y)|^2, \tag{28}$$

i.e., the kinematical contrast in dark field (bright field) is always bright (dark). Later we will see that this is not true if $|w|$ is not $\gg 1$.

4.7.3. Dynamical diffraction theory in terms of plane waves

In § 4.3 we have outlined that one of the prerequisites of a dynamical diffraction theory is that all plane waves involved must be treated equivalently, i.e., diffraction between all plane waves must be taken into account. Equation (22) suggests intuitively one way by which this condition can be met: Both sides of eq. (22) are differentiated

with respect to t, then we set $t=z$ and concede that also ϕ_0 may be z-dependent. In a next step we construct an equivalent equation, describing transition from the diffracted into the primary wave, i.e., $\boldsymbol{k}_g$ and $\boldsymbol{k}_0$ change their roles. This requires a change of the signs of g and s. Then we end up with a set of coupled differential equations:

$$\begin{aligned} \mathrm{d}\phi_g/\mathrm{d}z &= \mathrm{i}\left(\pi/\xi_g\right)\phi_0(z)\exp\left\{-2\pi\mathrm{i}\left[\boldsymbol{g}\cdot\boldsymbol{R}(z)+sz\right]\right\}, \\ \mathrm{d}\phi_0/\mathrm{d}z &= \mathrm{i}\left(\pi/\xi_g\right)\phi_g(z)\exp\left\{+2\pi\mathrm{i}\left[\boldsymbol{g}\cdot\boldsymbol{R}(z)+sz\right]\right\}. \end{aligned} \tag{29}$$

These equations are indeed one form of the differential equations established by HOWIE and WHELAN [1961] in a more detailed way for describing strain contrast by dynamical diffraction in imperfect crystals. These equations have to be integrated down to $z=t$ with the boundary condition $\phi_0=1$, $\phi_g=0$ at $z=0$. For $\boldsymbol{R}=0$ (perfect crystal) they can be integrated analytically; the result is identical to that given in eq. (14).

4.7.4. Dynamical diffraction theory in terms of Bloch waves

A different approach for deriving dynamical diffraction in imperfect crystals starts from the Bloch-wave solution of the perfect reference lattice as given in § 4.3 (WILKENS [1964], HÄUSSERMANN *et al.* [1973], WILKENS *et al.* [1973]). A trial solution, in the two-beam case given by

$$\psi = \varphi^{(1)}(z)\psi_B^{(1)} + \varphi^{(2)}(z)\psi_B^{(2)}, \tag{30}$$

with z-dependent Bloch-wave amplitudes $\varphi^{(j)}$ is inserted into the Schrödinger equation which now contains a potential, the periodicity of which is perturbed by the atomic displacement field $\boldsymbol{R}(\boldsymbol{r})$. This leads in the column approximation to a set of ordinary differential equations for the $\varphi^{(j)}$. In the two-beam case one possible representation is given by:

$$\begin{aligned} \mathrm{d}\varphi^{(1)}/\mathrm{d}z &= \mathrm{i}\pi\frac{\mathrm{d}(\boldsymbol{g}\cdot\boldsymbol{R})}{\mathrm{d}z}\frac{1}{\left(1+w^2\right)^{1/2}}\varphi^{(2)}(z)\cdot\exp\left\{-2\pi\mathrm{i}\left[\frac{s}{\sigma}\frac{\mathrm{d}(\boldsymbol{g}\cdot\boldsymbol{R})}{\mathrm{d}z}+\sigma z\right]\right\}, \\ \mathrm{d}\varphi^{(2)}/\mathrm{d}z &= \mathrm{i}\pi\frac{\mathrm{d}(\boldsymbol{g}\cdot\boldsymbol{R})}{\mathrm{d}z}\frac{1}{\left(1+w^2\right)^{1/2}}\varphi^{(1)}(z)\cdot\exp\left\{+2\pi\mathrm{i}\left[\frac{s}{\sigma}\frac{\mathrm{d}(\boldsymbol{g}\cdot\boldsymbol{R})}{\mathrm{d}z}+\sigma z\right]\right\}. \end{aligned} \tag{31}$$

These differential equations describe the defect-induced transition from Bloch-state (2) into Bloch-state (1) (first equation) and in the opposite direction (second equation). This transition is proportional to the displacement derivative $\mathrm{d}(\boldsymbol{g}\bullet\boldsymbol{R})/\mathrm{d}z$ and is essentially controlled by the phase factors $\exp(\pm 2\pi\mathrm{i}\sigma z)$. Sufficiently far away from the defect, where $\mathrm{d}(\boldsymbol{g}\bullet\boldsymbol{R})/\mathrm{d}z$ vanishes, the differential equations obey trivial solutions, $\varphi^{(1)}$ and $\varphi^{(2)}$ are constant. Once eq. (31) is integrated down to $z=t$ with $\varphi_n(0)=c_0^{(j)}$ the wanted amplitudes $\varphi_n(t)$ $(n=0,\mathrm{g})$ are obtained as outlined in § 4.3 with $\varphi^{(j)}(t)$ inserted into eq. (10). Anomalous absorption is taken into account by substituting σ in the exponents of eqs. (8) and (31) according to eq. (17). Normal absorption can be added subsequently.

Both sets of differential equations, eqs. (29) and (31), are in principle equivalent

(WILKENS *et al.* [1973]). However, one consequence of the special form of the equations (31) is that they are especially suitable for deriving semiquantitative analytical solutions (e.g., WILKENS [1964], CHIK *et al.* [1967], WILKENS and RÜHLE [1972] and KATERBAU [1981]).

4.7.5. Properties of strain contrast in strong-beam images

Figure 10 shows strong-beam images in bright and dark field of dislocations traversing the specimen from the top to the bottom surface. Near the surface the dislocation contrast reveals characteristic bright-dark oscillations, which are "in phase" in bright and dark field near the top surface and which are "in anti-phase" near the bottom surface. Further, the bright-dark oscillations are reversed if the sign of $\boldsymbol{g}$ is changed. In the middle of the specimen the contrast is mainly dark in both bright and dark field. These phenomena, which are of quite a general nature, will be explained by a qualitative discussion of eq. (31). For this purpose we re-define the Bloch-wave amplitudes $\varphi^{(l)}$ into $\hat{\varphi}^{(l)}$ by incorporating the absorption terms of eq. (18), which are incorporated in the usual procedure when going from the $\varphi^{(l)}(t)$ to $\varphi_n(t)$,

$$\hat{\varphi}^{(l)}(z) = \varphi^{(l)}(z) \exp - \frac{\mu^{(l)}}{2} z, \qquad l = 1, 2, \tag{32}$$

and restrict ourselves to dynamical diffraction, i.e., $w = 0$. The results, however, are valid also for $|w| < 1$. In a perfect crystal, where $\varphi^{(1)}$ and $\varphi^{(2)}$ are constant, $\hat{\varphi}^{(1)}(z)$ decreases rapidly with z, whereas $\hat{\varphi}^{(2)}$ is only weakly damped ($\xi'_0 < \xi'_g$). We assume a fairly thick specimen, so that $\hat{\varphi}^{(1)}$ has already decayed significantly in the middle of the specimen, cf. fig. 11. We consider the three regions (1), (2) and (3), indicated in fig. 11, separately. Now concerning contrast formation by a defect (e.g., a dislocation segment in fig. 10), we consider one particular column close to the defect, and the depth position z_0 of the latter is varied over the regions (1), (2), and (3).

If z_0 lies in region (1), where $\hat{\varphi}^{(1)}$ and $\hat{\varphi}^{(2)}$ are still of about equal strength, the

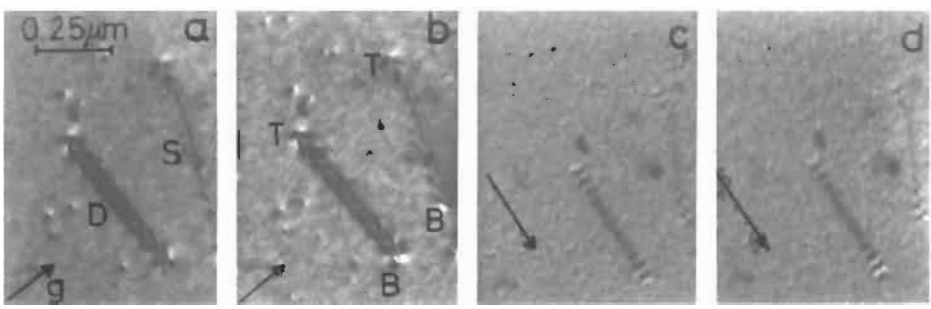


Fig. 10. Diffraction contrast of a single dislocation (S) and a narrow dislocation dipole (D) in copper traversing the specimen from the top (T) to the bottom (B) surface $\boldsymbol{g} = \{220\}$. (The black–white dots in the background are due to ion damage (cf. § 6.2.3) inside the microscope, as caused by an insufficient vacuum in the microscope column.) (a) Bright field $(\boldsymbol{g} \cdot \boldsymbol{b}) = 2$; (b) dark field $(\boldsymbol{g} \cdot \boldsymbol{b}) = 2$; (c) bright field $(\boldsymbol{g} \cdot \boldsymbol{b}) = 0$; (d) dark field $(\boldsymbol{g} \cdot \boldsymbol{b}) = 0$.

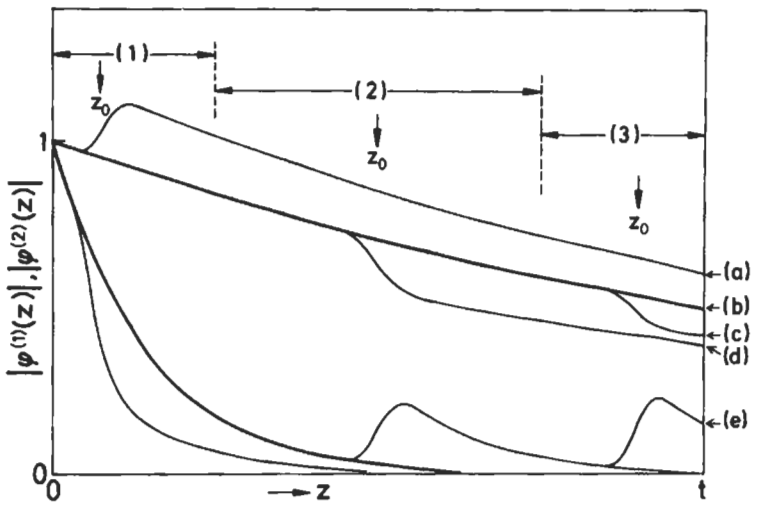


Fig. 11. Properties of strain contrast in strong-beam images taken with $w=0$, schematically. Shown are the moduli of the Bloch-wave amplitudes $\hat{\varphi}^{l}(z)$, $l=1, 2$, for defect depth-positions z_0 in either of the regions (1), (2) or (3), for details see text. (a): $|\hat{\varphi}^{(2)}(t)|$, z_0 in (1); (b): $|\hat{\varphi}^{(2)}(t)|_{bg}$; (c): $|\hat{\varphi}^{(2)}(t)|$, z_0 in (3); (d): $|\hat{\varphi}^{(2)}(t)|$, z_0 in (2); (e): $|\hat{\varphi}^{(1)}(t)|$, z_0 in (3).

scattering within a narrow interval near z_0 will increase (or decrease) $|\hat{\varphi}^{(1)}|$ and at the same time decrease (or increase) $|\hat{\varphi}^{(2)}|$, depending on the sign of $\mathrm{d}(\boldsymbol{g}\bullet\boldsymbol{R})/\mathrm{d}z$ *and* on the phase factors $\exp(\pm 2\pi i\sigma z_0)$ on the right-hand sides of eq. (31). The phase factor is periodic in z_0 with period $\xi_{g,\mathrm{eff}}$. Below the defect, the $\varphi^{(1)}$ in eq. (31) are constant again and the $\hat{\varphi}^{(l)}$ follow the decay as described by eq. (32). Then, going down to $z=t$, $\hat{\varphi}^{(1)}$ has decayed irrespective of what happened to $\hat{\varphi}^{(1)}$ at, or near to, z_0. Thus, the amplitudes $\phi_0(t)$ and $\phi_g(t)$ at $z=t$ are both determined solely by $\hat{\varphi}^{(2)}(t)$. And since the modulus of the latter oscillates with the period $\Delta z_0=\xi_{g,\mathrm{eff}}$ we get bright-dark depth oscillations in both bright field and dark field, which are "in phase": bright (dark) contrast in bright field corresponds to bright (dark) contrast in dark field.

In region (2) we have already $|\hat{\varphi}^{(1)}| \ll |\hat{\varphi}^{(2)}|$. Accordingly, for defects with z_0 within this region the scattering goes essentially only in one way from $\hat{\varphi}^{(2)}$ to $\hat{\varphi}^{(1)}$, since there is (nearly) nothing to scatter from $\hat{\varphi}^{(1)}$ to $\hat{\varphi}^{(2)}$. So $|\hat{\varphi}^{(2)}|$ is decreased irrespective of the actual value of z_0. On the way down to $z=t$, $\hat{\varphi}^{(1)}$ has again decayed, and the amplitudes $\phi_0(t)$ and $\phi_g(t)$ are, as in region (1), solely determined by $\hat{\varphi}^{(2)}(t)$. Since now $|\hat{\varphi}^{(2)}(t)| < |\hat{\varphi}^{(2)}(t)|_{bg}$, where the subscript bg refers to the corresponding value in the background far away from the defect, we obtain a dark contrast in both bright field and dark field, in striking contrast to the kinematical approach, cf. eqs. (27) and (28).

If z_0 lies in region (3), the scattering occurs even more than in region (2) only from $\hat{\varphi}^{(2)}$ to $\hat{\varphi}^{(1)}$. However, now $\hat{\varphi}^{(1)}$, on its enhanced level below z_0, can "survive" during the short distance from z_0 to t, along which absorption may be negligible. Consequently, the amplitudes ϕ_0 and ϕ_g at $z=t$ are determined by the interference of partial waves as described in § 4.3. The mutual phase shifts of these partial waves are given by exp

$[\pm 2\pi i\sigma(t-z_0)]$, and since absorption can be disregarded over the distance $t-z_0$, the depth oscillations in bright field and dark field must be complementary or in "antiphase". Of course, in practice, these three regions are not sharply separated. Depending on the actual specimen thickness, we expect depth oscillations, which are "in phase" in bright field and dark field for defects at positions z_0 close to the top surface of the specimen. These depth oscillations are damped out with increasing z_0. For thick, or moderately thick, specimens they disappear, or nearly disappear, for z_0 in the interior of the specimen, resulting in only a dark contrast in bright and dark field. Finally, they reappear in "antiphase" for z_0 close to the lower specimen surface, see fig. 10.

4.7.6. Structure-factor contrast

Coherent precipitates may reveal a mismatch between the lattice parameters in the precipitate and the surrounding matrix. This mismatch gives rise to a space-dependent displacement field and, accordingly, to strain contrast as described by eqs. (22), (29) or (31). If, in addition, the structure-amplitude F_s inside the precipitate differs from that of the matrix, e.g., by a change of the chemical composition, an additional contrast, named *structure-factor contrast* may be caused. The reason is that now the extinction length ξ_g (proportional to $1/F_s$) becomes space-dependent, $\xi_g = \xi_g(\boldsymbol{r})$. This can be incorporated into eqs. (22) and (29) by setting $1/\xi_g = 1/\xi_g(x, y; z)$ and taking $1/\xi_g$ in eq. (22) under the integral (cf. ASHBY and BROWN [1963] and WILKENS [1981]). For eq. (31) an appropriate extension for including structure-factor contrast was given by WILKENS *et al.* [1977].

4.8. Practical applications of the differential equations

Standard programs are available for the numerical integration of the differential equations (29) or (31) by electronic computers. The user has only to take care for subroutines for computing the displacement field $\boldsymbol{R}(x, y; z)$ or its derivative $\partial\boldsymbol{R}(x, y; z)/\partial z$. If the diffraction contrast of a particular defect is to be calculated as a function of the specimen thickness t and the depth position z_0 of the defect centre, considerable computer time can be saved by making use of the method of "generalized cross-section" (HEAD *et al.* [1973]), referring to eq. (29), or of the scattering matrix method (HÄUSSERMANN *et al.* [1973], and KATERBAU [1980]), referring to eq. (31). In many cases computer time can additionally be saved by application of symmetry relations inherent in the particular diffraction contrast problem or in the displacement field of the defect (HOWIE and WHELAN [1961], BALL [1964], POGANY and TURNER [1968] and KATERBAU [1980]).

5. *Dislocations*

5.1. Introduction

Dislocations and dislocation distributions, e.g., in deformed crystal, are generally best investigated in the two-beam bright-field mode operated at a sufficiently positive excitation error, which ensures best transmittivity, cf. § 4.5. This means a maximum specimen volume accessible by TEM and avoidance of pronounced black–white depth

oscillations for dislocations close to the specimen surfaces (§ 4.7.5), which may perturb the image. On the other hand, if details of the dislocation structure, e.g., splitting into partials etc., are of interest, the weak-beam technique $|w| \gg 1$ is superior (COCKAYNE *et al.* [1969]), at the cost of transmittivity.

5.2. The displacement field

We assume elastic isotropy and introduce the displacement field $\boldsymbol{R}$ of a straight dislocation in its own (right-handed) coordinate system x', y', z'. The dislocation line runs along the y'-axis with its line unit vector $\boldsymbol{u}$ pointing from $y' < 0$ to $y' > 0$. The Burgers vector $\boldsymbol{b}$ lying in the x'–y' plane can be split according to $\boldsymbol{b} = \boldsymbol{b}_s + \boldsymbol{b}_e$, where the screw component $\boldsymbol{b}_s$ is parallel, and the edge component $\boldsymbol{b}_e$ is perpendicular to $\boldsymbol{u}$. Then $\boldsymbol{R}$ is given by (e.g., HIRSCH *et al.* [1977]) *:

$$\boldsymbol{R} = \boldsymbol{R}_a + \boldsymbol{R}_s, \tag{33a}$$

$$\boldsymbol{R}_a = \frac{1}{2\pi}\left[\boldsymbol{b} \operatorname{artan} \frac{z'}{x'} + \boldsymbol{b}_e \frac{1}{2(1-\nu)} \frac{x'z'}{\hat{r}^2}\right], \tag{33b}$$

$$\boldsymbol{R}_s = -\frac{1}{2\pi}(\boldsymbol{b} \times \boldsymbol{u})\left[\frac{1-2}{2(1-\nu)} \ln \hat{r}/r_0 + \frac{1}{4(1-\nu)} \cdot \frac{x'^2 - z'^2}{\hat{r}^2}\right], \tag{33c}$$

with ν = Poisson's ratio, $\hat{r}^2 = x'^2 + z'^2$, r_0 = inner cut-off radius. Note that $\boldsymbol{R}_a$ is antisymmetric and $\boldsymbol{R}_s$ is symmetric in x',z'. Further, $\boldsymbol{R}$ changes its sign if the sign of either $\boldsymbol{b}$ *or* $\boldsymbol{u}$ is changed. This is obvious for $\boldsymbol{R}_s$ but must be noticed also for $\boldsymbol{R}_a$.

5.3. Contrast profiles of single perfect dislocations

We assume that $\boldsymbol{b}$ is a translation vector of the crystal structure ($\boldsymbol{b}$ is a "perfect" Burgers vector) so that $\boldsymbol{g} \cdot \boldsymbol{b}$ is integer. Regarding the contrast profiles of such dislocations we have to distinguish between the cases $\boldsymbol{g} \cdot \boldsymbol{b} \neq 0$ and $\boldsymbol{g} \cdot \boldsymbol{b} = 0$.

5.3.1. The g•b ≠ 0 contrast

In case of $\boldsymbol{g} \cdot \boldsymbol{b} \neq 0$ the contrast is mainly governed by the component $\boldsymbol{R}_a$ of eq. (33b). This component causes an S-shaped bending of the reflecting lattice planes as schematically indicated in fig. 12. We assume $s \neq 0$ and recall the fact that diffraction contrast is especially strong (weak) where the local excitation error $s_g(z)$ is decreased (increased) as compared to the background value s, cf. § 4.7.2, in other words where the reflecting lattice planes are bent by the dislocation displacement field towards (away from) the exact Bragg orientation. Consequently, for $s \neq 0$ the centre of gravity of a dislocation-

* In the TEM literature it is customary world-wide to denote the displacement field vector by $\boldsymbol{R}$ and the dislocation line unit vector by $\boldsymbol{u}$. Here we adopt the same notation, although in ch. 12 the symbol $\boldsymbol{u}$ is used for the displacement vector and $\boldsymbol{R}$ is used for denoting particle radii in various connections.

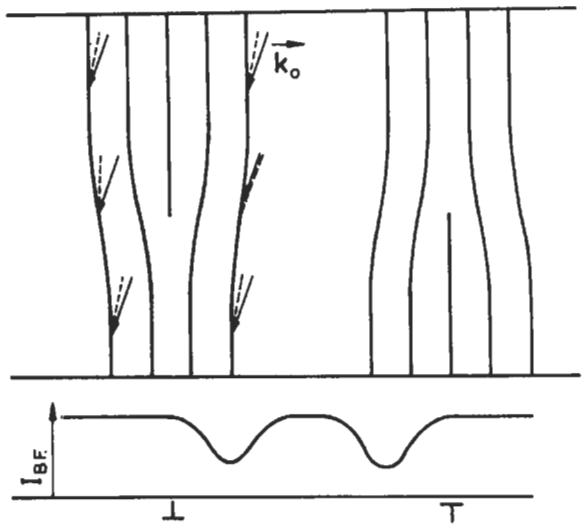


Fig. 12. Diffraction contrast of an edge dislocation for $(\boldsymbol{g}\bullet\boldsymbol{b})s \neq 0$, $s > 0$, schematically. Solid arrows represent the direction of incidence of the primary beam. Dashed lines represent local directions of incidence which would fulfill the Bragg condition. The angle between the two types of arrows is a measure of $s_i(z)$. Bright field intensity is denoted by I_{BF}.

contrast profile does not coincide with, but is shifted with respect to the image position of the dislocation line. Assuming a given direction of the line unit vector $\boldsymbol{u}$ the direction of this lateral contrast shift depends on the sign of $(\boldsymbol{g}\bullet\boldsymbol{b})s$, as will be demonstrated below in connection with the images of dislocation dipoles.

The kinematical contrast integral ϕ_g'', cf. eqs. (26)–(28), was evaluated for different types of dislocations by HIRSCH *et al.* [1960] and GEVERS [1962]. It turned out that the width of the kinematical dislocation contrast is about $(1/3–1/2)\xi_K$, where $\xi_K = s^{-1}$ denotes the kinematical extinction length. If dynamical diffraction has to be taken into account ($|w| < 1$) this result remains essentially valid, with ξ_K substituted by $\xi_{g,\mathrm{eff}} = \sigma^{-1}$ [eq. (17)] (HOWIE and WHELAN [1962]). It should be added that the contrast of a dislocation line is always single-lined for $|\boldsymbol{g}\bullet\boldsymbol{b}| = 1$ but may be (asymmetrically) double-lined for $|\boldsymbol{g}\bullet\boldsymbol{b}| = 2$ and $0 < |w| < 1$.

Under normal strong-beam (bright-field) conditions, $\xi_{g,\mathrm{eff}}$ can be as large as several 10 nm, whereas under weak-beam conditions $\xi_K = s^{-1}$ can be kept as small as 5 nm (COCKAYNE *et al.* [1969] and COCKAYNE [1978]). Consequently, dislocation images are much sharper in the latter case, cf. fig. 13 where the same specimen area is imaged under strong-beam and weak-beam imaging conditions, respectively. Of course, the gain in resolution is paid for by a reduction of the useful specimen thickness.

5.3.2. The g•b = 0 contrast

In case of $\boldsymbol{g}\bullet\boldsymbol{b} = 0$ the dislocation contrast is completely extinguished ($\boldsymbol{g}\bullet\boldsymbol{R} = 0$) for screw dislocations and also for edge dislocations if in the latter $\boldsymbol{g}$ is parallel to the

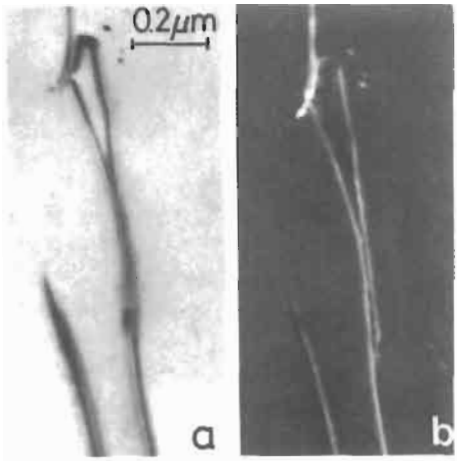


Fig. 13. Dislocations in silicon: (a) strong-beam bright field image, $w > 0$; (b) the same area imaged under dark-field weak beam conditions, $w > 1$.

dislocation line. For $\boldsymbol{g}\cdot\boldsymbol{b} = 0$ but $|\boldsymbol{g}\cdot\boldsymbol{b} \times \boldsymbol{u}| \neq 0$ a "residual" contrast due to the displacement component $\boldsymbol{R}_s$ in eq. (33c) is produced, which is symmetric with respect to the image position of the dislocation. This residual contrast is controlled in strength by the modulus $|\boldsymbol{g}\cdot\boldsymbol{b} \times \boldsymbol{u}|$ and is in general significantly weaker than the $\boldsymbol{g}\cdot\boldsymbol{b} \neq 0$ contrast and can thus be distinguished from the latter. Depending on the strength of $|\boldsymbol{g}\cdot\boldsymbol{b} \times \boldsymbol{u}|$ the contrast may consist of a single, a double or a triple contrast line, cf. e.g., HOWIE and WHELAN [1962] and HIRSCH *et al.* [1977].

Typical dislocation images for $\boldsymbol{g}\cdot\boldsymbol{b} \neq 0$ and $\boldsymbol{g}\cdot\boldsymbol{b} = 0$ are shown in fig. 10. Note that the depth-oscillations where the dislocations approach the specimen surfaces (§ 4.7.5) show different symmetries for $\boldsymbol{g}\cdot\boldsymbol{b} \neq 0$ and $\boldsymbol{g}\cdot\boldsymbol{b} = 0$. These differences are occasionally helpful for discriminating between the two cases.

For dislocation of mixed type, complete contrast extinction is not possible at all, since in such a case the terms $\boldsymbol{g}\cdot\boldsymbol{b}$, $\boldsymbol{g}\cdot\boldsymbol{b} \times \boldsymbol{u}$ and $\boldsymbol{g}\cdot\boldsymbol{b}_e$ cannot be zero simultaneously. If the first two terms are zero but $\boldsymbol{g}\cdot\boldsymbol{b}_e \neq 0$ [cf. eq. (33b)] an asymmetrical residual contrast is produced. This is demonstrated in fig. 14 where a curved dislocation segment, which changes in type from screw to edge, is imaged under $\boldsymbol{g}\cdot\boldsymbol{b} = 2$ conditions (fig. 14a, strong symmetrical double-line contrast) and $\boldsymbol{g}\cdot\boldsymbol{b} = 0$ conditions (fig. 14b). In the latter case the dislocation is out of contrast in the pure screw and the pure edge orientation but reveals an easily detectable asymmetric contrast around the "45°-orientation", where $\boldsymbol{g}\cdot\boldsymbol{b}$ and $(\boldsymbol{g}\cdot\boldsymbol{b} \times \boldsymbol{u}) = 0$, but $\boldsymbol{g}\cdot\boldsymbol{b}_e \neq 0$. Here $\boldsymbol{b}_e$ refers to the component of $\boldsymbol{b}$ perpendicular to the *local* line direction $\boldsymbol{u}$.

5.4. Contrast of dislocation pairs

Two parallel dislocations having opposite Burgers vectors $\boldsymbol{b}$ constitute a dislocation dipole. Another description is that $\boldsymbol{b}$ is the same for both dislocations but their line

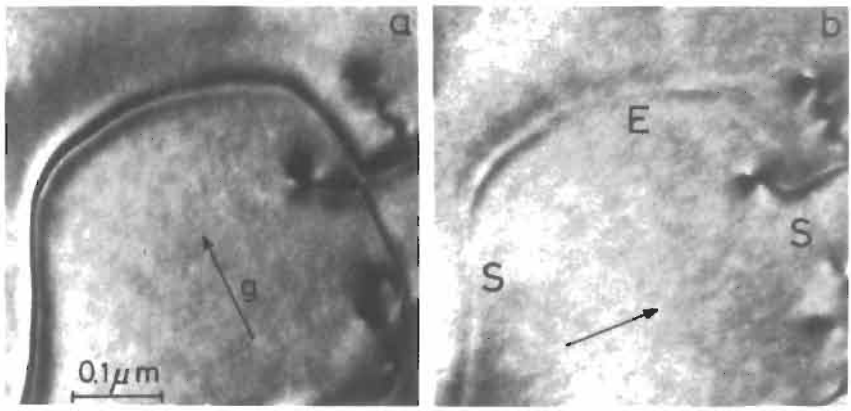


Fig. 14. Strong-beam images of a curved dislocation segment of Burgers vector $\boldsymbol{b}=\frac{1}{2}[110]$: (a) $\boldsymbol{g}=(220)$, $(\boldsymbol{g}\bullet\boldsymbol{b})=2$, dislocation is in strong contrast; (b) $\boldsymbol{g}=(2\bar{2}0)$, $(\boldsymbol{g}\bullet\boldsymbol{b})=0$, the dislocation contrast is extinguished only at the edge- (E) and the screw (S) orientations. (From KORNER and KARNTHALER [1980].)

vectors $\boldsymbol{u}$ are opposite. The most common representative is a dipole constituted by two pure edge dislocations. If such a dipole is imaged with $s\neq 0$ and $\boldsymbol{g}\bullet\boldsymbol{b}\neq 0$, then $(\boldsymbol{g}\bullet\boldsymbol{b})s>0$ for one and $(\boldsymbol{g}\bullet\boldsymbol{b})s<0$ for the other dislocation. Consequently, the contrasts of the two dislocations are displaced into opposite directions: The distance between the two contrast lines is either smaller or larger than the true projected distance between the dislocations. If $\boldsymbol{g}$ is changed into $-\boldsymbol{g}$, while leaving the sign of s unchanged, the two cases are interchanged. This phenomenon is called inside-outside contrast. An example is shown in fig. 15a, b.

In other cases, the two parallel dislocations may show up the same sign of $\boldsymbol{g}\bullet\boldsymbol{b}$. This may occur in an ordered alloy where two dislocations with the same perfect Burgers vector $\boldsymbol{b}$ of the disordered lattice constitute a split superdislocation ($2\boldsymbol{b}$ = perfect Burgers vector of the ordered alloy) with an antiphase boundary between them. Another example refers to a perfect dislocation which is split into partial dislocations (Burgers vectors $\boldsymbol{b}_p$) and where the signs of $\boldsymbol{g}\bullet\boldsymbol{b}_p$ are the same for both partials. If in such a case the distance between the two dislocations becomes of the order of, or smaller than, $\xi_{g,\mathrm{eff}}$ (which may be $\xi_K=1/s$ in the weak-beam case) the contrast of one of the two dislocations increases in strength and the contrast of the other decreases, depending on the signs of the $(\boldsymbol{g}\bullet\boldsymbol{b}_p)$ (WILKENS and HORNBOGEN [1964], STOBBS and SWORNE [1971] and COCKAYNE [1978]). This is demonstrated in fig. 15c, d, showing the weak-beam images of an edge-dislocation dipole in Ag where the individual dislocations D_1 and D_2 are split into partials. For both dislocations the two partials give rise to the same sign of $\boldsymbol{g}\bullet\boldsymbol{b}_p$, which, however, is opposite for D_1 and D_2. Accordingly, the sequence "strong-weak" is opposite for D_1 and D_2. Further, the sequence is changed when changing the sign of $\boldsymbol{g}$ (indicated by the arrows).

Below a critical distance the contrast of the weakly imaged dislocation disappears in the contrast tail of the strongly imaged one. This effect limits the resolution of dislocation pairs having the same sign of $(\boldsymbol{g}\bullet\boldsymbol{b})$ even on weak-beam images to about 2 nm. This

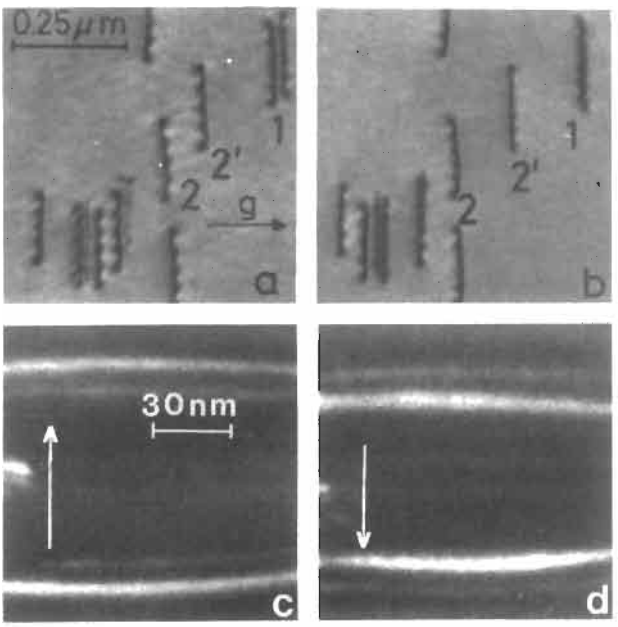


Fig. 15. Diffraction contrast of dislocation pairs taken with $g = \pm(20\bar{2})$: (a, b) Strong-beam bright-field images of dislocation dipoles in copper, taken with $w > 0$. Dipole 1 is clearly resolved in (a) but unresolved in (b). The distance between the dipole dislocations 2 and 2′ is narrow in (a) and wide in (b). (cs d) weak-beam images of dislocations in silver which are split into partials. The dislocation at the top (bottom) is named D_1 (D_2) in the text. (From KORNER and KARNTHALER [1981].)

lower limit is important for the evaluation of stacking-fault energies from the measurement of equilibrium distances of split dislocations, cf., e.g., STOBBS and SWORNE [1971].

5.5. Determination of the dislocation Burgers vectors and the dislocation densities

The determination of the Burgers vector $\boldsymbol{b}$ of individual dislocations requires a number of images of the corresponding specimen area taken under different twobeam conditions. If a given dislocation shows contrast extinction or only residual ($\boldsymbol{g}\cdot\boldsymbol{b} = 0$) contrast for two (non-parallel) diffraction vectors, say $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$, then its Burgers vector $\boldsymbol{b}$ is perpendicular to $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ (regarding the sign of $\boldsymbol{b}$ see below). The Burgers vector analysis is considerably facilitated if the general crystallographic direction of $\boldsymbol{b}$ can be anticipated, for instance $\boldsymbol{b} = \frac{1}{2}\langle 110\rangle$ in a fcc lattice. In this case three images taken with three diffraction vectors $\boldsymbol{g} = \{111\}$ are sufficient for indexing all dislocations with $\boldsymbol{b} = \frac{1}{2}\langle 110\rangle$, since the six significantly different vectors $\boldsymbol{b}$ of this kind give rise to a different contrast-extinction behaviour for the three $g = \{111\}$ (HIRSCH *et al.* [1977]). The series of micrographs shown in fig. 16 stems from a deformed and subsequently annealed Ni crystal and was taken from a specimen parallel to the primary slip plane (111) with $\boldsymbol{g}_a = (11\bar{1})$, $\boldsymbol{g}_b = (\bar{1}11)$, $\boldsymbol{g}_c = (1\bar{1}1)$, where the subscripts a, b, c refer to figs 16a, b, c,

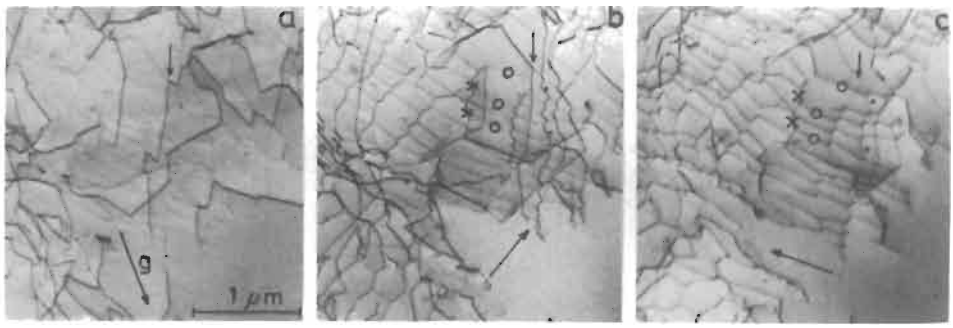


Fig. 16. Dislocation network in nickel. Two-beam bright-field images of the same specimen area, specimen normal (111). Three types of segment directions (see text) of the network are indicated by open circles (1), crosses (2), and dots (3). (a) $\boldsymbol{g}=(11\bar{1})$, (b) $\boldsymbol{g}=(\bar{1}11)$, (c) $\boldsymbol{g}=(1\bar{1}1)$. The small arrows point to the same position on the three images.

respectively. In the central part of figs. 16b, c a network is visible constituted by three types of dislocation segments, as indicated in the figures. The segments (1) give rise to a $\boldsymbol{g}\cdot\boldsymbol{b}\neq 0$ contrast on fig. 16c and to a residual ($\boldsymbol{g}\cdot\boldsymbol{b}=0$) contrast on fig. 16b and (scarcely to see) on fig. 16a. Accordingly, we have $\boldsymbol{b}_1=\pm\frac{1}{2}[101]$. The segments (2) show $\boldsymbol{g}\cdot\boldsymbol{b}\neq 0$ contrast on figs. 16b and c, but not on fig. 16a. This leads to $\boldsymbol{b}_2=\pm\frac{1}{2}[1\bar{1}0]$. The segments (3) are in $\boldsymbol{g}\cdot\boldsymbol{b}\neq 0$ contrast only on fig. 16b, thus $\boldsymbol{b}_3=\pm\frac{1}{2}[011]$. Taking always the upper sign, we have $\boldsymbol{b}_2+\boldsymbol{b}_3=\boldsymbol{b}_1$ as it is required for dislocation networks. Other dislocation segments in figs. 16a–c can be indexed in the same way.

For a full indexing of a dislocation a discrimination between $+\boldsymbol{b}$ and $-\boldsymbol{b}$ (for a given choice of the line direction $\boldsymbol{u}$) is required. This can be achieved by making use of the fact that the lateral shift of $(\boldsymbol{g}\cdot\boldsymbol{b})s\neq 0$ contrast lines is indicative for the sign of the S-shaped bending of the reflecting lattice planes from which the signs of $\boldsymbol{u}$ and $\boldsymbol{b}$ can be determined by physical arguments. In this context, certain contrast peculiarities due to surface-stress relaxation effects may also be helpful, cf., e.g., SIEMS *et al.* [1962], TUNSTALL *et al.* [1964], WILKENS *et al.* [1967] and HAZZLEDINE *et al.* [1975].

The dislocation density ρ is defined as the (average) dislocation length per unit volume or, in a somewhat misleading way, as the number of dislocations per unit area. A full and reliable evaluation of ρ requires rather extensive experimental work. Here we stress the attention to one aspect only: On a given set of micrographs taken with a given two-beam diffraction vector $\boldsymbol{g}$ dislocations of some slip systems may be fully in contrast ($\boldsymbol{g}\cdot\boldsymbol{b}\neq 0$). Dislocations of other slip systems may be visible in residual contrast ($\boldsymbol{g}\cdot\boldsymbol{b}=0$) only or may be completely invisible. This fact suggests immediately that any reliable determination of ρ requires a careful evaluation of different sets of micrographs of the same specimen area taken with different $\boldsymbol{g}$-vector, cf., e.g., STEEDS [1966], ESSMANN [1966].

5.6. Elastic anisotropy

Experience has shown that regarding the fundamental properties of the dislocation contrast of dislocations, elastic isotropy is a useful approximation even for moderately

anisotropic crystals like copper. However, for strongly anisotropic materials such as β-brass the contrast rules, described above on the base of elastic isotropy, break down. This holds in particular for the discrimination between $\boldsymbol{g}\cdot\boldsymbol{b} \neq 0$ and $\boldsymbol{g}\cdot\boldsymbol{b} = 0$ contrast. In such cases a Burgers vector analysis requires the comparison of observed contrast figures, in particular of those parts showing depth oscillations, with computer-simulated images for modelled dislocation types, cf. HEAD *et al.* [1973] and HUMBLE [1978].

6. *Point-defect agglomerates, radiation damage*

6.1. Introduction

Irradiation of crystals with energetic particles (e.g., electrons, neutrons, or ions) leads to the formation of interstitials and vacancies. Point defects may also be created by plastic deformation or by quenching-in of thermal vacancies. If these point defects are mobile, they may cluster together, forming point-defect agglomerates such as dislocation loops. In fcc metals they may also form stacking-fault tetrahedra. This holds especially for vacancies, but in principle also for interstitials. Vacancies may also agglomerate into cavities (bubbles, voids). One of the problems to be solved by TEM is the determination of the crystallographic nature of such agglomerates and the determination of their type, vacancy (V) or interstitial (I), i.e., whether they are produced by an agglomeration of vacancies or of interstitials. In this section some of such methods are briefly outlined. For a review see WILKENS [1978].

6.2. Dislocation loops

6.2.1. Formation of loops

Dislocation loops are formed by agglomeration of point defects into plates on densely packed lattice planes. For instance, a monolayer (diameter D) of agglomerated vacancies on a {111}-plane in a fcc metal gives rise to an intrinsic stacking fault surrounded by a dislocation loop (diameter D) with a partial Burgers vector $\boldsymbol{b}_F = \frac{1}{3}\langle 111\rangle$ of Frank type, perpendicular to the loop plane. Agglomeration of interstitials on {111} gives rise to an extrinsic stacking fault over the loop area. If the loop size exceeds a critical value, it becomes energetically more favourable to eliminate the stacking fault area by sweeping a Shockley partial $\boldsymbol{b}_S = \frac{1}{6}\langle 112\rangle$ over the loop area to convert the Frank partial into a perfect Burgers vector $\boldsymbol{b}$, e.g.,

$$\boldsymbol{b}_F + \boldsymbol{b}_S = \boldsymbol{b}, \tag{34a}$$

$$\tfrac{1}{3}[111] + \tfrac{1}{6}[11\bar{2}] = \tfrac{1}{2}[110]. \tag{34b}$$

There is evidence that similar two-step mechanisms for the formation of loops with perfect Burgers vectors are existing also for bcc (EYRE and BULLOUGH [1965], JÄGER and WILKENS [1975]) and for hcp metals (FÖLL and WILKENS [1977]).

6.2.2. Analysis of large dislocation loops

If the loop diameter D is sufficiently large compared to $\xi_{g.eff}$ then the loop nature is clearly visible on the micrograph and the Burgers vector $\boldsymbol{b}$ of the loop can be determined by the contrast-extinction rules outlined in § 5.5. Approximate information about the habit plane of the loop can be obtained by following the change of the projected shape and width of the loop when tilting the specimen around an axis parallel to the operating $\boldsymbol{g}$-vector (for details see, e.g., MAHER and EYRE [1971]).

Segments of a dislocation loop which are opposite to each other have the same direction of $\boldsymbol{b}$, but opposite line direction $\boldsymbol{u}$. Their contrast behaviour is therefore similar to that of a dislocation dipole, cf. § 5.4. Accordingly, if the loop is imaged with $(\boldsymbol{g}\bullet\boldsymbol{b})s \neq 0$, the loop is imaged either as inside contrast (loop contrast inside the true projected loop position) or as outside contrast, depending on the sign of $(\boldsymbol{g}\bullet\boldsymbol{b})s$. The question, inside- or outside contrast, can best be clarified by comparing images taken with $\boldsymbol{g}$ and $-\boldsymbol{g}$ and leaving the sign of s unchanged, cf. fig. 17 where dislocation loops of interstitial type in electron-irradiated molybdenum are shown; the apparent sizes of the loop contrasts change remarkably when going from $\boldsymbol{g}$ to $-\boldsymbol{g}$. Then, the loop type, V or I, can be determined if the inclination of the loop plane with respect to the primary beam (direction of $\boldsymbol{k}_0$) and the operating $\boldsymbol{g}$-vector is known (HIRSCH *et al.* [1977]).

In application to practical cases the inside outside method suffers from a number of pitfalls which may lead to the wrong answer regarding the loop type. Recipes have been worked out in order to avoid these pitfalls (MAHER and EYRE [1971], and FÖLL and WILKENS [1975]).

6.2.3. Small dislocation loops

If the loop diameter D becomes smaller than $\xi_{g.eff}$, dislocation loops are no longer imaged "as loops". On kinematical strong-beam images the loop contrast degenerates to

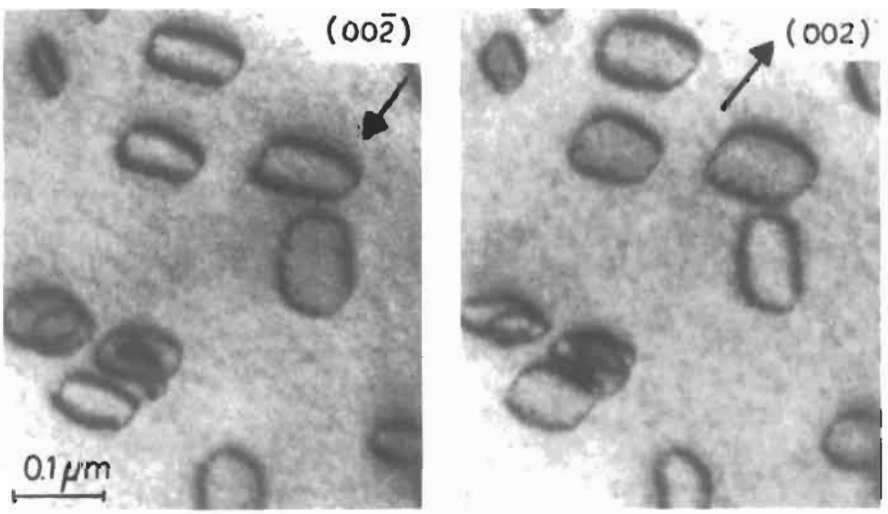


Fig. 17. Dislocation loops of interstitial type in electron-irradiated molybdenum. Two-beam bright-field images taken with $w > 0$. Note the apparent change in the loop size when changing the sign of $\boldsymbol{g}$. For details see text.

a black-dot contrast containing very little information on details of the loop (RÜHLE [1967]). Even under weak-beam conditions the inside-outside contrast method breaks down for $D < 7$ nm. (HÄUSSERMANN *et al.* [1973]). For dislocation loops below this limit the black–white contrast method becomes applicable when the specimen is imaged under two-beam dynamical conditions with $w = 0$, in bright or dark field (RÜHLE [1967]). If under such an imaging condition a small dislocation loop is located within the surface-near regions of depth-oscillation (§ 4.7.5) it gives rise to a characteristic black–white contrast figure. An example is shown in fig. 18 referring to a Cu specimen containing mainly small Frank dislocation loops of V-type (produced by irradiation with 30 keV Cu-ions) in a very surface-near layer close to the bottom surface of the specimen: Such black–white contrast figures may be characterized by a black–white vector $\boldsymbol{l}$, pointing from the centre of the black to the centre of the white lobe. Because of the very nature of the depth oscillation contrast, the black–white contrast of a loop is inverted, i.e., bright is changed into dark and vice versa, if either the loop type, V or I, is changed or if the loop depth-position z_0 is shifted towards $z_0 \pm \frac{1}{2}\xi_{g,\mathrm{eff}}$. Accordingly, in order to discriminate between loops of V or I type we need information about the sign of $(\boldsymbol{g} \cdot \boldsymbol{l})$ *and* of the depth position z_0 of the loop. The latter information can be obtained by careful stereo measurements (e.g., RÜHLE [1967] and RÜHLE and WILKENS [1967]).

Additional information about the direction of the Burgers vector and the loop plane normal $\boldsymbol{n}$ can be obtained from the fine structure in the centre of a black–white contrast figure (KATERBAU [1976] and EYRE *et al.* [1977]) and from the outer shape of the black–white contrast figures by comparison of experimentally obtained and calculated contrast figures. For elastically isotropic, or nearly isotropic, materials the shape of the contrast figure and in particular the angle between $\boldsymbol{l}$ and $\boldsymbol{g}$ depend in a characteristic manner on the direction of $\boldsymbol{b}$ and $\boldsymbol{n}$ with respect to $\boldsymbol{k}_0$ and $\boldsymbol{g}$, respectively. These dependencies have been worked out by numerical integration of the differential equations

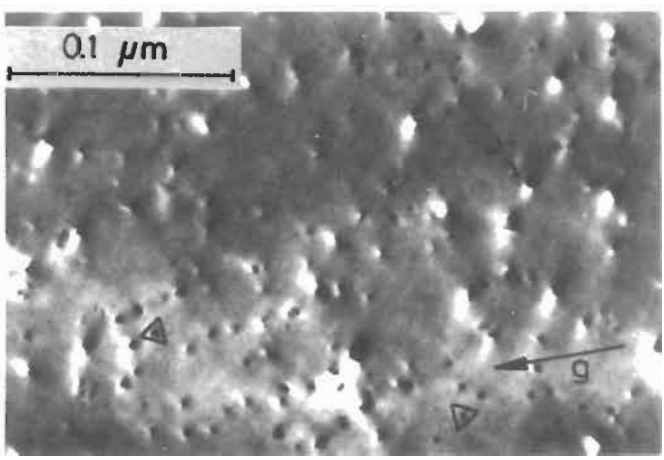


Fig. 18. Black–white contrast of small dislocation loops (Frank loops) in a copper specimen irradiated with 30 keV copper ions. Dynamical dark field image taken with $w = 0$, specimen normal (110). $\boldsymbol{g} = (002)$. The inserted triangles indicate small stacking-fault tetrahedra.

References: p. 1105.

(29) or (31) (e.g., RÜHLE [1967], KATERBAU [1976], EYRE *et al.* [1977] and ENGLISH *et al.* [1980]) or by an analytical first-order perturbation integration of eq. (31) (e.g., WILKENS and RÜHLE [1972]).

The latter was successfully applied to the analysis of small dislocation loops in ion-irradiated tungsten (HÄUSSERMANN *et al.* [1972] and JÄGER and WILKENS [1975]) and cobalt (FÖLL and WILKENS [1977]). More recent analytical calculations by WILKENS and KIRCHNER [1981] have shown that the shapes of the black–white contrasts of small loops are sensitive to elastic anisotropy.

6.3. Stacking-fault tetrahedra

Frank dislocation loops in fcc metals can dissociate into *stacking-fault tetrahedra* (SFT) constituted by stacking faults on the four {111}-type tetrahedra faces interconnected by stair rod dislocations with $\boldsymbol{b}=\frac{1}{6}\langle 110\rangle$ at the edges of the tetrahedra. This transformation is favoured in particular in metals of low stacking-fault energy such as gold, silver or copper.

One consequence of the small strength of the stair–rod dislocations bounding the SFT is that their strain contrast is rather weak. Accordingly, the contrast of a SFT is predominantly determined by the stacking-fault areas (which may give rise to a fringe pattern if the SFT is larger than the extinction length (CHIK [1966]), i.e., the contrast shape of a SFT is in general close to the shape of the SFT projected onto the image plane. Some SFT, as produced by ion-damage in Cu, are indicated in fig. 18 by a symbol reflecting their outer shapes, as expected for a specimen orientation close to (110).

In spite of the weakness of the strain field, SFT may reveal black–white contrast. For large SFT with edge lengths of the order of, or larger than, $\xi_{g,\mathrm{eff}}$ the black–white contrast degenerates to the so-called Ashby–Brown contrast (ASHBY and BROWN [1963]) where the depth oscillations are suppressed due to surface-stress relaxation (MCINTYRE and BROWN [1966] and CHIK *et al.* [1967]). Under suitable conditions even small SFT may reveal black–white contrast similar to that of small dislocation loops (SALDIN *et al.* [1978]). However, in contrast to small loops, black–white vector $\boldsymbol{l}$ points always parallel or anti-parallel to $\boldsymbol{g}$ (because the strain field does not show up a pronounced preferential direction). The latter property is sometimes helpful in order to discriminate between small loops and small SFT.

6.4. Cavities

Cavities (voids, bubbles) are best imaged under two-beam conditions with $s=0$. Then cavities act like a local reduction of the specimen thickness t. Thus, in a specimen with foil thickness t in that range where thickness fringes are observable the contrast of a cavity is dark on the front flank of a bright thickness contour and bright on the rear side (VAN LANDUYT *et al.* [1965]). If t is beyond the region of thickness fringes, the contrast is bright but decreases rapidly with decreasing diameter d of the cavity. Finally, if d is below a critical value of some nm, cavities are no longer visible on in-focus images. Then the contrast is considerably improved by an appropriate defocusing of the objective lens (RÜHLE and WILKENS [1975]).

6.5. Displacement cascades and disordered zones

Energetic ions (in the case of fast neutron irradiation the primary knocked-on atoms) deposit their kinetic energy in a crystal in so-called cascades, within which a large number of atomic replacements takes place. In an ordered alloy these replacement events lead to a zone of reduced long-range order which reflects size and structure of the cascade. Such disordered zones can be made visible on dark-field images taken with a superlattice diffraction vector $\boldsymbol{g}$ (JENKINS and WILKENS [1976]). An example of disordered zones in Cu_3Au irradiated with Cu^+-ions is shown in fig. 19.

7. *Precipitates*

The precipitation of a second phase can be pursued by TEM, which was done successfully since the early days of TEM of materials (HIRSCH *et al.* [1977]). Depending on the structure of the interface, we talk about coherent, partially coherent or incoherent precipitates. A *partially coherent* particle may possess one coherent interface and one where coherency is lost (i.e., is semi-coherent or incoherent). This can be caused either by different structures in the precipitate and the matrix or by a large displacement along the interface (e.g., OLSON and COHEN [1979]). An incoherent particle has a crystal structure different from the matrix. Often there exist no orientation relationships between the two phases. A precipitate can give rise to TEM contrast for two reasons: (i) due to the alteration of the electron waves passing through the particle (*precipitate contrast*) and (ii) due to alteration of the electron waves passing through columns near the particle where the crystal has been distorted due to the presence of the precipitate (matrix contrast or strain contrast).

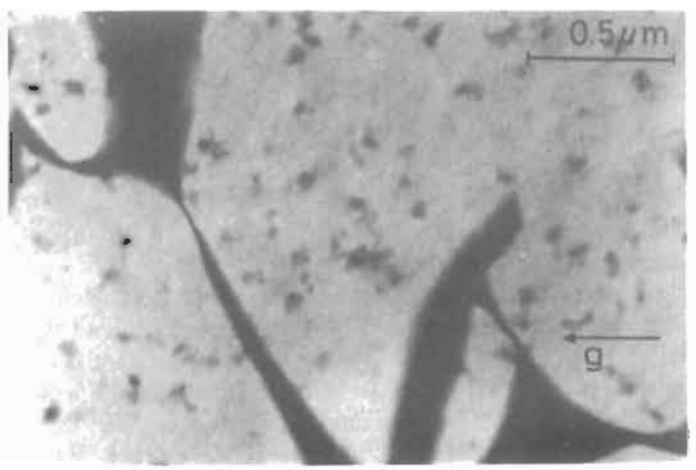


Fig. 19. Dark-field image taken with a superlattice reflection $\boldsymbol{g} = (110)$ of a fully ordered Cu_3Au specimen irradiated with 30 keV copper ions. The dark dots represent zones Of reduced long-range order as produced by the displacement cascades of the incident ions. The dark bands are due to antiphase boundaries.

References: p. 1105.

The interpretation of the *matrix contrast* is straightforward and can be done in a similar way as for radiation-induced defects or dislocations. The strain contrast has to be observed for different foil orientations (and different diffraction vectors) under well-defined dynamical two-beam or well-defined kinematical conditions (see § 4). If a strain contrast can be observed then contrast calculations have to be performed for different strain distributions surrounding the precipitate in a thin foil. Matching the results of the calculations with the observations gives a qualitative and sometimes even quantitative model of the precipitate (ASHBY and BROWN [1963]).

MCINTYRE and BROWN [1966] and CHIK *et al.* [1967] showed that black–white contrasts (BW contrasts) can be observed if certain conditions for the size and magnitude of the strain are fulfilled. The strain is thereby expressed in terms of a dimensionless quantity,

$$Q = \varepsilon r_0^3 \cdot g \cdot \frac{1}{\xi_g^2}, \tag{35}$$

where r_0 is the radius of the particle, g the modulus of the diffraction vector, ξ_g the extinction length, and ε a parameter describing the constrained strain of the particle (ASHBY and BROWN [1963]). BW contrasts are expected for certain combinations of r_0 and Q (CHIK *et al.* [1967]). Small values of r_0 result in BW contrasts with depth oscillation similar to those of radiation-induced defects (§ 6), whereas for large values of r_0 and Q the depth oscillations of the BW contrasts are suppressed by the stress relaxations at the foil surfaces (Ashby–Brown contrast). In this case the sign of the displacement field can be determined by DF images unambiguously (ASHBY and BROWN [1963]). Very large defects produce no BW contrast at all (CHIK *et al.* [1967]). The calculation of TEM contrast was performed for inclusions of different shapes and included also elastic anisotropy (see e.g., LEPSKI [1974]). The main problem in doing such computer simulations is the determination of the displacement field of the precipitate placed in a thin foil.

Spherical precipitates form in Cu–Co alloys (fig. 20) and the sizes depend on the annealing treatment (MCINTYRE and BROWN [1967]). The typical *coffee-bean contrast* can be observed. A "line-of-no-contrast" is perpendicular to the diffraction vector $\boldsymbol{g}$. The width of the lobes measured perpendicular to $\boldsymbol{g}$ is a measure of the size of the precipitate.

The small plate-like precipitates in Al–Cu also produce a strain contrast. The visibility depends on the operating diffraction vector since the displacement field is strongly anisotropic. Coherent Ag precipitates in Al–Ag alloys do not form a strain contrast, since there exists (nearly) no constrained strain around the particle (see fig. 21a).

Semicoherent precipitates may produce a strain contrast which can be evaluated similarly as for coherent precipitates. This was demonstrated for small Mo_2C precipitates in Mo by LEPSKI and BURCK [1981, 1982]. HIRSCH *et al.* [1977] showed that coherent, semicoherent, or incoherent precipitates can be studied by utilizing the strain contrast (DF and BF techniques). If the symmetry of the precipitate is different from the symmetry of the matrix, then additional reflections in the diffraction pattern can be

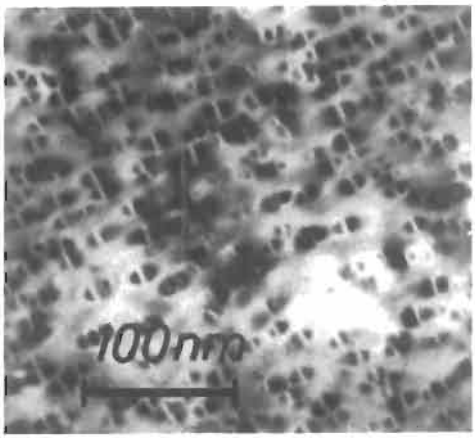


Fig. 20. Contrast from spherically symmetrical strain fields due to precipitations in a Cu–Co alloy. The coffee-bean type contrast is visible.

observed (fig. 21e). The shape of the precipitate is revealed by DF images taken with such a reflection belonging to the precipitate (fig. 21c, d).

The *precipitates* themselves can be made visible (i) if difference in the structure factor in the two phases ("structure-factor contrast"), of which an example is shown in fig. 21, (ii) if the orientation of the foil is such that certain planes in the precipitate are near the exact Bragg condition (strong reflection) while no planes in the matrix are in such a position ("orientation contrast"), (iii) if the matrix displacements induced by the precipitate cause an abrupt phase-change at the precipitate ("stacking-fault contrast"), or (iv) if a moire pattern is formed due to different lattice parameters in the matrix and precipitate. The contrast caused for the different cases is already extensively described by HIRSCH *et al.* [1977].

Particles of a second phase change the *diffraction pattern* of the pure matrix because: (i) there are extra reflections due to the diffraction pattern from the precipitate itself, (ii) the second-phase particles exert an influence on the matrix diffraction spots, (iii) there is the possibility of double diffraction (if the particle is embedded in the matrix), and (iv) since the precipitates are usually small, the shape of the diffraction spots is reciprocal to the shape of the (small) particles.

8. *Structure of grain boundaries and interfaces*

Internal surfaces determine many properties of materials. For an understanding of the properties it is essential to know the structure of the defects, which are actually the regions in space at which two crystals meet. Interfaces between dissimilar materials are called *heterophase boundaries* (or just interfaces). Interfaces between crystals differing only in relative orientation and/or translation are called *homophase boundaries*. Homophase boundaries include grain boundaries, stacking faults, twins, and antiphase

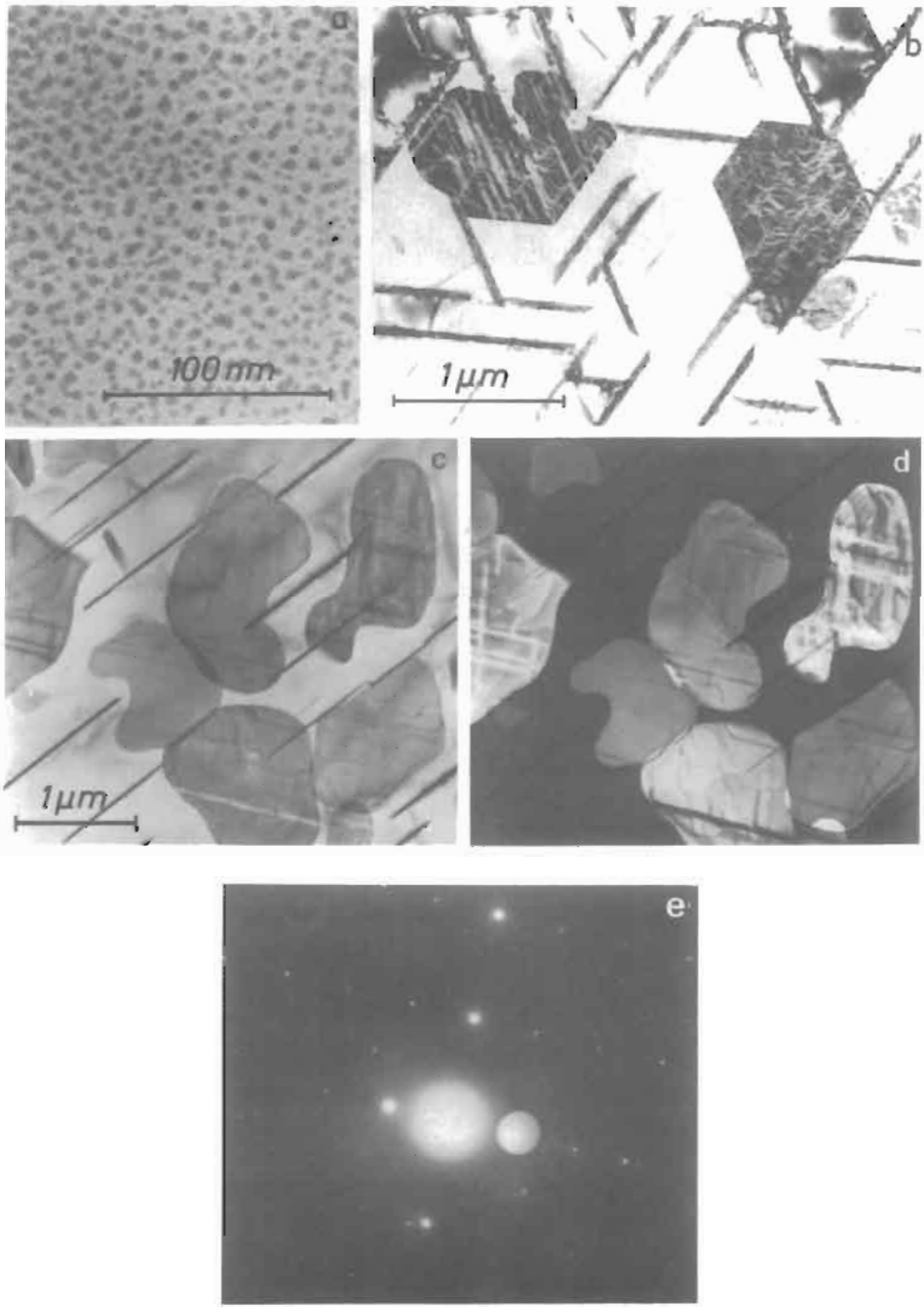


Fig. 21. Contrast of precipitates. (a) Structure-factor contrast of small coherent precipitates in Al–6% Ag alloys; no strain contrast is visible. (b) Structure-factor contrast of incoherent (large) precipitates in Al–6% Ag alloys annealed for 1000 min at 400°C (θ'' precipitates); foil orientation ~ (110); interface dislocations are visible between the Al matrix and the Ag precipitates. (c) Contrast of incoherent Al–4% Cu alloys; structure-factor contrast; foil orientation (106), BF image. (d) DF image with a reflexion of the Cu precipitate. Only Cu precipitates of one variant are in contrast. (e) Diffraction pattern of figs. (c, d).

boundaries in ordered alloys. In the first part of this section the possibilities for revealing the structure of homophase boundaries will be described. The second part deals with some observations on heterophase boundaries.

KALONJI and CAHN [1982] manifested that ten parameters are required for the geometrical description of an interface, see fig. 22. In this concept one assumes two interpenetrating crystals which are then misoriented (four parameters) and shifted against each other (three parameters). Three additional parameters are required to specify the location of a plane interface in the interpenetrating crystals. Finally, atoms of one crystal are removed from one side of the surface and those of the second crystal from the other. In this way, the *symmetry* of the interface can be described very elegantly by means of group theory (GRATIAS *et al.* [1979] and KALONJI and CAHN [1982]). Stacking faults and antiphase boundaries are pure translation interfaces, whereas grain boundaries represent general homophase boundaries. The geometric parameters are potentially capable of undergoing relaxation.

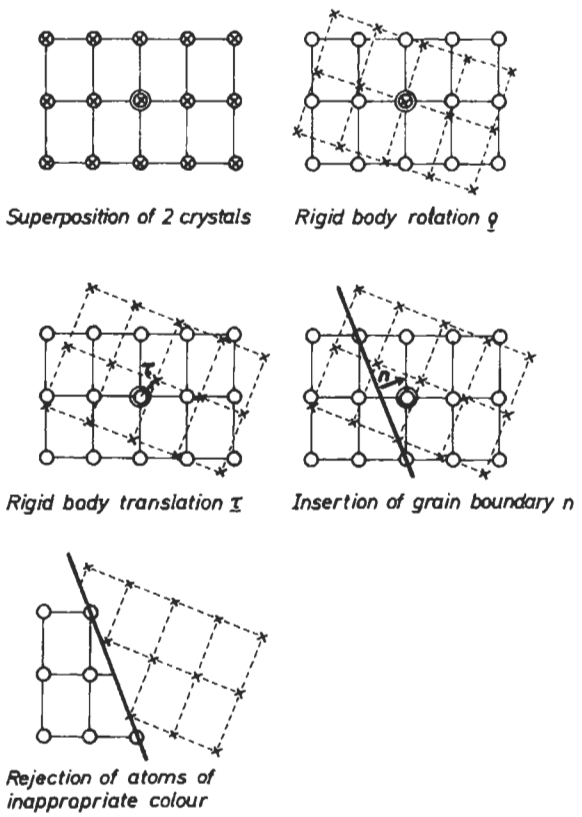


Fig. 22. The idealized construction of a crystalline interface: (a) superposition of the two crystals; (b) imposition of rigid body rotation (four parameters); (c) imposition of rigid body translation (three parameters); (d) insertion of surface S of the interface (three parameters); (e) rejection of inappropriate atoms at the interface.

References: p. 1105.

8.1. Transmission electron microscopy of pure translation interfaces

At a pure translation interface (stacking faults, antiphase boundaries) two perfect crystal parts are shifted against each other. Usually, the plane of the interfaces coincides with a close-packed plane. For example, the translation vector is $\tau = \frac{1}{6}\langle 112\rangle$ for a stacking fault (STF) in fcc materials, whereas τ is a lattice vector of the disordered crystal structure for antiphase boundaries in an ordered structure. TEM contrast is expected in the projected area of the (usually) inclined translation interface. The depth position of the interface may be t_1. Then the crystal part lying below the interface is shifted against the upper reference crystal by the translation vector τ which defines also the sign of τ.

The diffraction contrast can easily be calculated by utilizing the so-called "matrix-method" introduced by AMELINCKX and VAN LANDUYT [1978]. With this method the contrast can be calculated for a general n-beam situation. We restrict ourselves to the two-beam case which takes into account only one scattered beam ϕ_g besides the transmitted beam ϕ_0. The amplitudes of the scattered beam ϕ_g and of the transmitted beam ϕ_0 are represented by a column vector

$$\psi = \begin{pmatrix} \phi_0 \\ \phi_g \end{pmatrix}. \tag{36}$$

AMELINCKX and VAN LANDUYT [1978] described the changes of the amplitudes after passing through a perfect crystal of thickness t [normalized excitation error w, cf. eq. (7)] by:

$$\begin{pmatrix} \phi_0(t,w) \\ \phi_g(t,w) \end{pmatrix}_{\text{out}} \boldsymbol{M}(t,w) \begin{pmatrix} \phi_0 \\ \phi_g \end{pmatrix}_{\text{in}}, \tag{37}$$

where the subscripts "in" and "out" refer to the incoming and outgoing waves for a perfect crystal, respectively. The initial values are of course $\phi_0 = 1$, $\phi_g = 0$. $\boldsymbol{M}(t,w)$ represents a "response matrix" given by:

$$\boldsymbol{M}(t,w) = \begin{pmatrix} \phi_0^0 & \phi_g^{0-} \\ \phi_g^0 & \phi_0^{0-} \end{pmatrix}, \tag{38}$$

with [cf. eq. (14)]:

$$\phi_0^0(t,w) = \cos \pi\sigma t - \mathrm{i}\,\frac{w \sin \pi\sigma t}{\left(1+w^2\right)^{1/2}}, \quad \phi_g^0(t,w) = \frac{\mathrm{i} \sin \pi\sigma t}{\left(1+w^2\right)^{1/2}}, \tag{39}$$

anomalous absorption is included, which means σ is complex, whereas normal absorption in neglected. The normalized excitation error w has to be replaced by $-w$ in eq. (39) to form ϕ_0^{0-} and ϕ_g^{0-} of eq. (38).

The influence of a planar defect included in a transmitted specimen on the amplitude distribution of the different waves can be described by a response matrix of the defect $\boldsymbol{M}^{\mathrm{d}}$. This matrix is called $\boldsymbol{M}^{\alpha}$ for a pure translation interface, with

$$\boldsymbol{M}^{\alpha} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix}, \tag{40}$$

and $\alpha = 2\pi \boldsymbol{g} \cdot \boldsymbol{\tau}$ ($\boldsymbol{g}$ = diffraction vector, $\boldsymbol{\tau}$ = translation of the perfect crystal below the defect with respect to the perfect crystal in which the electrons enter).

The amplitudes of the waves transmitted through a specimen (total foil thickness t) containing such a translation interface (depth position $t_1(x)$, see § 4) are determined by three contributions: (i) the contribution of the perfect crystal [foil thickness $t_1(x)$] lying above the planar defect, described by a response matrix $\boldsymbol{M}(t_1, w)$, [eq. (38)]; (ii) the contribution of the planar defect itself [response matrix $\boldsymbol{M}^{\alpha}$, eq. (40)]; and (iii) the contribution of the perfect crystal (thickness $t_2 = t - t_1$) lying below the defect [response matrix $\boldsymbol{M}(t_2, w)$].

The amplitudes of the waves at the lower foil surface are just expressed by

$$\begin{pmatrix} \phi_0(t) \\ \phi_g(t) \end{pmatrix} = \boldsymbol{M}(t_2, w) \cdot \boldsymbol{M}^{\alpha} \cdot \boldsymbol{M}(t_1, w) \begin{pmatrix} \phi_0(0) \\ \phi_g(0) \end{pmatrix}. \tag{41}$$

The contrast caused by an inclined defect can be calculated by an evaluation of eq. (41) for different depth positions t_1.

The following results can be obtained for an inclined STF (for details see AMELINCKX and VAN LANDUYT [1978]):

(i) Bright and dark fringes are expected in the projected area of the STF.

(ii) With increasing thickness new fringes are created in the centre of the foil.

(iii) The fringes are parallel to the closest surface.

(iv) The BF fringe pattern is symmetrical with respect to the foil centre, whereas the DF image is similar to the BF image close to the top surface but complementary close to the bottom surface.

Figure 23 shows an example of the contrast of a stacking fault in a Cu–10 at% Al alloy. The image of a stacking fault (STF) is characterized by the value of $\alpha = 2\pi \boldsymbol{g} \cdot \boldsymbol{\tau}$. In fcc materials stacking faults are predominantly produced on {111} planes by a shear of $\boldsymbol{\tau} = \frac{1}{6}\langle 112 \rangle$ or by removal or insertion of a plane of atoms. This may happen when vacancies or interstitial atoms condense on close-packed planes. In fcc materials the STF are classified as intrinsic or extrinsic.

The determination of the type of STF is identical with the determination of the sign of α. We know from the symmetry rules that the sign of the contrast fringe in BF and DF at the top of the transmitted foil is positive (bright fringe) for $\alpha > 0$ (thick foil). From the BF and Df images of a STF (see fig. 23) we can determine the top surface and the bottom surface. Since we know the sign of α from the edge fringe of the BF image, we have the information necessary to determine the sense of $\boldsymbol{\tau}$. This method was first given by HASHIMOTO *et al.* [1962]. GEVERS *et al.* [1963] pointed out that the type of the fault can be determined from a DF image alone; the sense of inclination need not be known. GEVERS *et al.* showed that $\boldsymbol{\tau}$ can be determined by the following simple method. The diffraction vector $\boldsymbol{g}$ is drawn as an arrow with its origin at the centre of contrast fringes on the DF image. The nature of the stacking fault is intrinsic (extrinsic) if the limiting

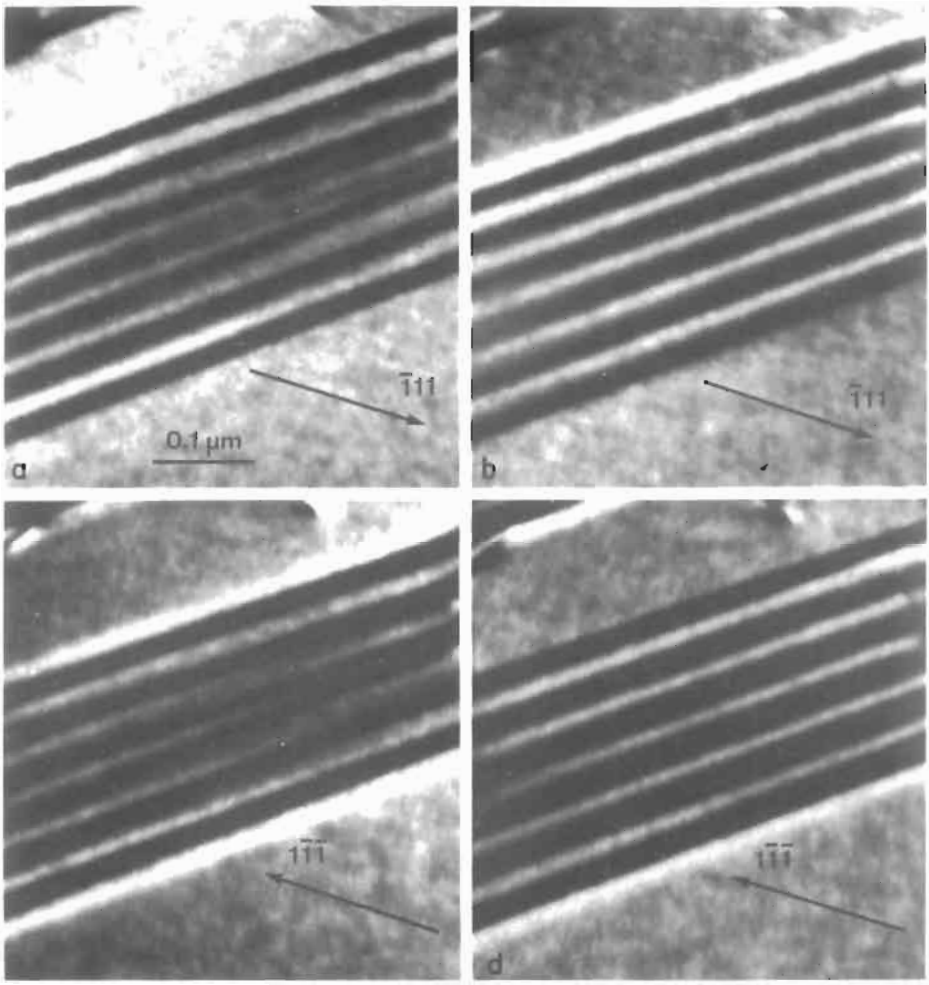


Fig. 23. Images of a stacking fault in Cu–10 at% Al, intrinsic stacking fault: (a) BF, $\boldsymbol{g}=(\bar{1}11)$; (b) DF, $\boldsymbol{g}=(\bar{1}11)$; (c) BF, $\boldsymbol{g}=(1\bar{1}\bar{1})$; (d) DF, $\boldsymbol{g}=(1\bar{1}\bar{1})$. The rule of GEVERS *et al.* [1963] is fulfilled. (Courtesy A. KORNER and H. P. KARNTHALER.)

fringe on the side of the arrow of $\boldsymbol{g}$ is dark (bright). This rule is valid in fcc materials for {111}, {220} and {400} reflexions while it is reversed for {200}, {222} and {440} reflexions. The contrast of the STF vanishes for $\boldsymbol{g}\cdot\boldsymbol{\tau}=$ integer. The direction of τ can be determined from two images taken with different diffraction vectors on which the STF is out of contrast.

The contrast of domain boundaries, twins, antiphase boundaries can also be calculated with the matrix method, see AMELINCKX and VAN LANDUYT [1978].

8.2. Transmission electron microscopy of grain boundaries

Grain boundaries (GBs) have been the object of intensive TEM studies for many years, and in that time many papers and review lectures have appeared (eg., CHADWICK and SMITH [1976], CLAREBROUGH and FORWOOD [1978] and BALLUFFI [1980]. It is well recognized that the structure of GBs is important for the understanding of the problems of segregation, recrystallization texture and intergranular embrittlement, and of course the bonding across the boundary also governs those properties. In the papers quoted above it is shown that much information can be obtained by TEM diffraction contrast studies, diffraction studies and recently also by direct lattice imaging.

By TEM diffraction contrast studies, information on the intrinsic structure of GBs can be obtained; mainly dislocations can be analyzed. For special situations the geometrical parameters can be determined by TEM work. Diffraction studies allow the determination of the "width" and symmetry of GBs, whereas direct imaging by means of high-resolution TEM allows an insight into the atomic structure of the boundary.

In ch. 9 it is explained that GBs may be described in different ways: the dislocation model, the plane-matching model, and the coincidence-site model. All models will summarize the relaxation phenomena at the GB in different configuration of defects. The models assume that there exist certain misorientations which possess a low GB energy, and these misorientations are characterized by a low Σ value of the coincidence-site lattice *.

If the symmetry of a boundary and the included dislocations are determined, then the GB is completely specified. The parameters included in the description can be determined by TEM in principle by comparing experimental micrographs taken under well defined diffraction conditions with results of computer simulations carried out using the dynamical theory of electron diffraction. For the computer simulations the displacement of the atoms (lattice planes) due to relaxation as well as many experimental parameters such as the thickness, orientation of the interface, and extinction- and absorption lengths must be known. In addition, experimental diffraction parameters, such as the number of beams excited and their s value, must be established accurately.

The contrast calculations can be very complicated. Each beam excited in the upper crystal is incident on the lower crystal and can give rise to further beams in that crystal. Thus, the total number of beams propagating in the lower crystal can be substantial, and the coupling of the beams is strongly influenced by the crystallographic relationship of the two crystals. There are three experimental conditions which allow a more straightforward interpretation:

(i) Two-beam diffraction in one crystal and negligible diffraction in the other. Either the upper or lower crystal can be oriented for two-beam diffraction, and the non-diffracting crystal is regarded as a block in which normal but not anomalous absorption

* The CSL can be obtained by allowing the two misonented crystal lattices adjoining the boundary to interpenetrate and translate so that lattice points of each crystal coincide. The space lattice made up of the coincident lattice points is called the CSL. The fraction of lattice points (in one crystal) in good coincidence is defined as $1/\Sigma$. Coincidence models of GBs are discussed in ch. 10B, § 2.2.1.3.

References: p. 1105.

occurs. The contrast behaviour is similar to thickness fringes. This diffraction condition can only be obtained approximately, since weak beams are inevitably excited in the "non-diffracting" crystal, and these may be coupled to some extent with the beam selected for image formation. This was demonstrated by HUMBLE and FORWOOD [1975] who compared microdensitometer traces taken across experimentally obtained GB thickness fringes to computer profiles. Humble and Forwood observed that the "non-diffracting" crystal does not behave simply as an absorbing wedge, and that no simple function could be found which would adequately describe the attenuation in this grain.

(ii) Simultaneous two-beam diffraction in the two crystals. In this mode the specimen is exactly oriented so that only one diffracted beam is strongly excited in the upper crystal — and in such a way that this (excited) beam does not excite further strong beams in the lower crystal. At the same time, only one beam is strongly excited in the lower crystal by the incoming beam. Usually, only BF images are taken (HUMBLE and FORWOOD [1975]).

(iii) In cases where the crystallography permits, two-beam or systematic diffraction by sets of planes with identical spacing and orientation in the adjacent crystals may occur. Such sets of planes are referred to as common, and this case is a special instance of condition (ii). This orientation has been used for the measurement of the relative shift of the two adjacent crystals.

Experimentally, the orientation relation between the two adjacent grains of an interface and the normal on the interface can be obtained by three micrographs and accompanying diffraction patterns taken under different, well established orientations. The lateral shift τ can be measured for special configurations (MATTHEWS and STOBBS [1977] and POND [1979]).

The TEM contrast of GB dislocations can be treated similarly as for lattice dislocations. However, for the GB dislocations further parameters, including the geometry of the interface and the diffraction conditions pertaining to both crystals must be considered. The direction of the Burgers vector $\boldsymbol{b}$ of the GB dislocation can be determined by using the $\boldsymbol{g}\cdot\boldsymbol{b}=0$ criterion. However, the accuracy of this semi-quantitative method is small, since the magnitude of the Burgers vector of secondary GB dislocations is usually small.

Boundaries between two crystals are assumed to have a low-energy structure for certain misorientations (e.g., BROKMAN and BALLUFFI [1981]). This low-energy structure can be thought to exist of finely spaced arrays of dislocations, the so-called "primary" dislocations. Any deviation from such low-energy orientation relations is accommodated by a network of dislocations, the "secondary" dislocations. Experimental observations on those dislocations were carried out by Balluffi and coworkers (e.g., SCHOBER and BALLUFFI [1969]) mainly on "artificial" (001) twist boundaries in Au. The TEM studies reveal that a grid of undissociated secondary dislocations is present, see fig. 24. The distance of dislocations with the same Burgers vectors depends on the deviation $\Delta\theta$ from the exact coincidence position. The observations are in agreement with the predictions of the theory (BROKMAN and BALLUFFI [1981]). Secondary dislocations can split into partial secondary dislocations and stacking-fault-like structures. This is observed not only for metals but also for semiconductors (BACMAN *et al.* [1981]).

Lattice dislocations can interact with grain boundaries during yielding, creep and

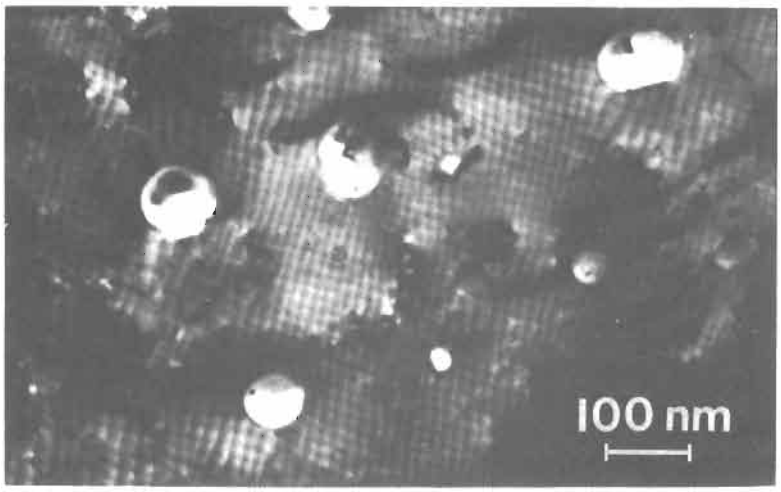


Fig. 24. "Artificial" twist boundary in a gold bicrystal, twist angle near $\Sigma = 5$ (36.9°). Square grids of secondary screw grain-boundary dislocations can be observed, Burgers vector $b = (a/10)\langle 310\rangle$. The dislocations accommodate the deviations from the exact $\Sigma = 5$ misorientation. (Courtesy T. SCHOBER and R. W. BALLUFFI.)

recrystallization (SMITH [1976]). Many TEM observations on reactions of lattice dislocations with GB dislocations can be explained on the basis of the dislocation model for GB structures. For the reactions of the lattice dislocations with grain boundaries one has to observe that Burgers vectors are conserved and dislocation lines can end only on other dislocations or at a free surface.

8.3. Diffraction studies on the structure of grain boundaries

SASS and BRISTOWE [1980] demonstrated that diffraction techniques also in the TEM can answer questions concerning the atomic structure of GBs. Each GB represents a periodic arrangement of strains (misfit) and good coincidences. Electrons and X-rays are scattered by this periodic grid which possesses a unit cell which corresponds to the CSL. If the (periodic) displacements at the GBs are known, then the scattering factor of such boundaries (mostly twist boundaries) can be calculated for the different reflections caused by the GBs. The reciprocal lattice, e.g., of a twist boundary, exists of thin rods, the integral intensity in the rod depends on the scattering factor (see appendix). The length of the rod is proportional to the inverse of the " thickness" of the disturbed region near the GB. This grain-boundary " thickness" approximately equals the distance of the dislocations present in the GB.

8.4. Direct imaging of grain boundaries

High-resolution electron microscopy (HREM) allows the direct imaging of certain GBs. In HREM the structure of the specimen must be periodic in the direction of the transmitted electron beam. Therefore, only tilt boundaries can be studied where the tilt

axis is parallel to the beam direction. Since the strong periodicity is disturbed in the core region of the boundaries, the point-to-point resolution of the instrument must be good enough so that information on the relaxation of atoms is transferred through the objective lenses despite the still rather high spherical aberration of the best lenses available to date.

Tilt boundaries in silicon (BOURRET and DESSEAUX [1979]), germanium (KRIVANEK *et al.* [1977]) and molybdenum (PENISSON *et al.* [1982]) were studied by HREM. For these studies the specimen thickness has to be below 15 nm. Figure 25 shows a HREM micrograph of a tilt boundary in molybdenum. For the determination of the positions of the atoms close to the GB, contrast simulations have to be performed. In the simulation work one assumes a certain atomic configuration near the tilt boundary. The simulated micrographs have to be compared with the experimentally observed pictures for different defocusing values. The atomic arrangements have to be modified until a complete fit between the observed and calculated images is reached.

8.5. TEM contrast of heterophase boundaries

Heterophase boundaries are formed when two materials of different structure and/or different chemical composition meet. If, for example, any type of phase transformation occurs then a heterophase boundary is present where the different phases meet. This is true for the interface of a martensitic phase transformation where two materials with the same chemical composition but different structures meet, as well as for an interface of thermally grown precipitate.

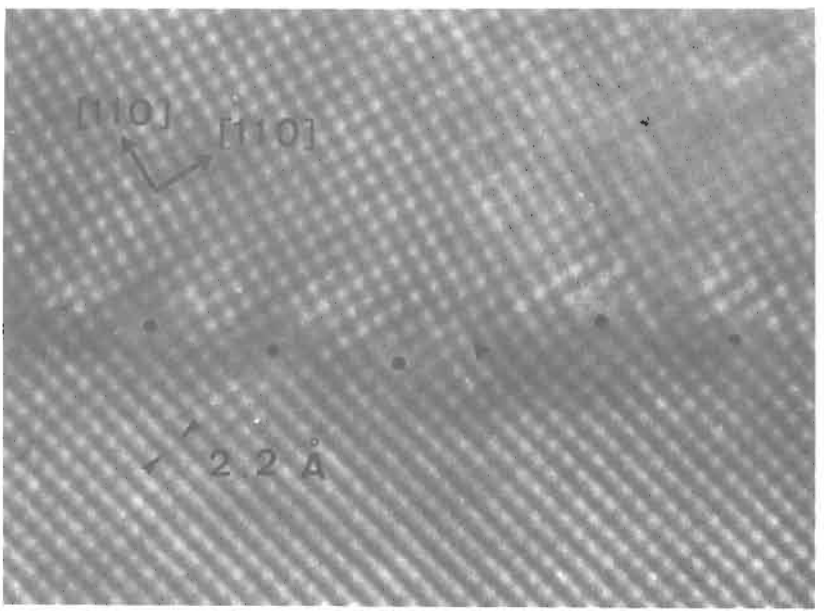


Fig. 25. High-resolution image of a $\Sigma = 1$ pure tilt boundary in Mo. Foil orientation [001] (for both grains). (Courtesy A. BOURRET and J. M. PENISSON.)

The importance of the structure of these heterophase boundaries for the nucleation and growth process of the precipitate is recognized. However, only few systematic TEM studies were performed investigating the structure of such interfaces. The studies would require experimental imaging of the inclined interface under different diffraction conditions under which reflections from both crystal structures are excited simultaneously or one after the other. The observations have to be compared with contrast simulations. General expressions were derived even in the framework of anisotropic elasticity theory for the elastic displacement and stress fields by BONNET [1981]. The results, however, have as yet not been applied to contrast simulations of an interface.

The geometry of the heterophase interfaces can be described by a generalization of the geometry of grain boundaries (BALLUFFI *et al.* [1982]). These models are essentially fit–misfit models in which the regions of a good "fit" are patched where partial lattice matching across the boundary is achieved, and the regions of 'misfit' are boundary line defects which possess the character either of a dislocation or of a boundary step (facet). (See also ch. 9, § 2.3 for a discussion of heterophase ["interphase"] boundaries).

The model has been applied to interfaces between two cubic crystals — copper and silver (LAFFONT and BONNET [1982]) — and for hexagonal materials (BONNET *et al.* [1981]). A metal–metal-carbide interface was studied by different authors (DAHMEN *et al.* [1981] and FLORJANCIC *et al.* [1982]). FLORJANCIC *et al.* studied in detail the interface between Mo and Mo2C precipitates which forms after cooling from a supersaturated solution. In TEM micrographs (fig. 26) a dislocation-type contrast can be observed. The determination of the Burgers vector by means of the simple $\boldsymbol{g}\cdot\boldsymbol{b}=0$ rule is not unambiguously possible. The contrast behaviour suggests that the Burgers vector is of type $\frac{1}{2}\langle 111\rangle$, diffraction patterns yield the orientation relation between Mo and Mo_2C. Dense packed planes of the Mo lattices are about parallel to dense-packed planes of Mo_2C (Burgers relation). Careful diffraction studies were performed for an edge-on configuration of the interface, fig. 26. It was found that besides the reflections of Mo and Mo_2C, streaks can also be observed lying between them. The formation of such additional streaks can simply be explained with the assumption that an interface crystal is present between the (pure) Mo and Mo_2C. From the maximum streak length the minimum thickness of the interface crystal can be determined to within nine lattice planes of the Mo lattice.

9. *High-resolution TEM*

9.1. Introduction

By high-resolution TEM we understand all attempts of utilizing modern electron microscopes up to the ultimate resolution limit. The latter is determined by the wave aberrations of the objective lens — spherical aberration, chromatic aberration, astigmatism etc. — under optimal adjustments of size and position of the objective aperture, of the direction of the incident electron beam with respect to the optical axis and of the actual focusing of the objective lens, cf. also § 2. Under such optimum conditions a point-to-point resolution of about 0.2 nm and a lattice-fringe resolution of < 0.1 nm are

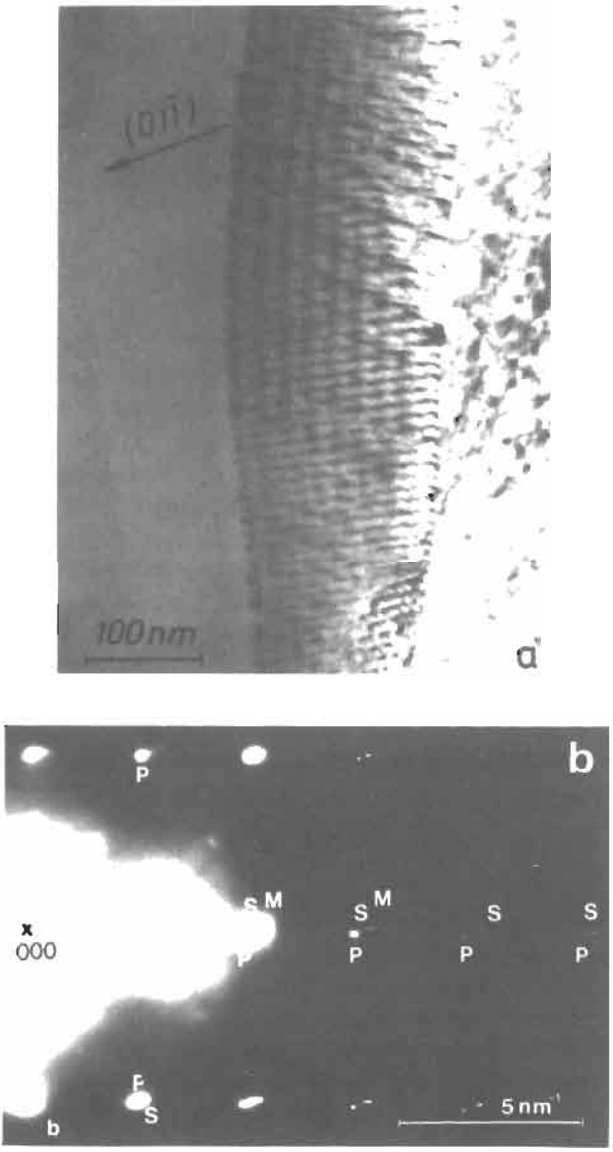


Fig. 26. Interface between Mo and Mo_2C: (a) dislocation contrasts are visible, the Burgers vector cannot be analyzed unambiguously; (b) diffraction pattern of the same interface in an edge-on configuration, additional "streaks" are observable due to the "interface" crystal.

achievable by modern microscopes specialized for high-resolution work. At the present time, high-resolution TEM plays only a marginal role in physical metallurgy, but this may change in the near future when the modern generation of electron microscopes becomes widely available in laboratories concerned with physical metallurgy. Therefore,

we present here only a brief outline of the basic principles of high-resolution microscopy. In the following we restrict ourselves to spherical aberration as the most important lens error in the high-resolution work. For more extended introductions to high-resolution TEM we refer to the references in the bibliography.

9.2. The optical transfer function

We assume that the specimen is orientated perpendicular to the optical axis of the microscope and introduce a Cartesian coordinate system in the lower specimen surface with the z-axis (unit vector $\boldsymbol{e}_z$) pointing downwards, i.e., towards the objective lens. Axial illumination is assumed, i.e., the incident beam is parallel to $\boldsymbol{e}_z$. A space vector in the plane perpendicular to $\boldsymbol{e}_z$ will be denoted by $\hat{\boldsymbol{r}}$. The objective lens may be focused onto the plane $z=\delta$ where $\delta>0$ ($\delta<0$) means overfocusing (underfocusing) with respect to the lower specimen surface ($z=0$). We denote the electron wave function in the plane $z=0$ by $\psi_0(\hat{\boldsymbol{r}})$ *. By propagating from $z=0$ to the plane $z=\delta$, $\psi_0(\hat{\boldsymbol{r}})$ is modified by Fresnel diffraction to $\psi_\delta(\hat{\boldsymbol{r}})$. Note that for $\delta>0$, the wave function $\psi_\delta(\hat{\boldsymbol{r}})$ is really existing at the plane $z=\delta$, whereas for $\delta<0$, $\psi_\delta(\hat{\boldsymbol{r}})$ means a virtual wave function, which would merge into $\psi_0(\hat{\boldsymbol{r}})$ at $z=0$ when propagating from $z=\delta<0$ to $z=0$ in an empty space, i.e, in the absence of the specimen.

The propagation of the electron wave through the imaging system of the microscope is considered in terms of partial waves or Fourier waves of amplitudes $\tilde{\psi}(\boldsymbol{p})$, where the Fourier waves $\tilde{\psi}_0$ and $\tilde{\psi}_\delta$ are assumed to emerge either from the plane $z=0$ or from the plane $z=\delta$, respectively,

$$\tilde{\psi}_j(\boldsymbol{p}) = \int \psi_j(\hat{\boldsymbol{r}}) \exp(-2\pi \mathrm{i}\boldsymbol{p}\cdot\hat{\boldsymbol{r}})\, \mathrm{d}^2\hat{\boldsymbol{r}}, \qquad j = 0, \delta. \tag{42}$$

$\boldsymbol{p}$ is a "space frequency" vector perpendicular to $\boldsymbol{e}_z$, starting from the optical axis. Since $\boldsymbol{p}$ has the dimension of a reciprocal length, it is a vector in reciprocal space. $\tilde{\psi}_0$ and $\tilde{\psi}_\delta$ are interlinked by the Fresnel diffraction theorem,

$$\tilde{\psi}_j(\boldsymbol{p}) = \tilde{\psi}_0(\boldsymbol{p}) \cdot \exp - \mathrm{i}\chi_\delta(p), \tag{43}$$

with

$$\chi_\delta(p) = \pi\lambda\delta p^2, \tag{44}$$

with λ = electron wavelength.

The partial wave, belonging to the space frequency vector $\boldsymbol{p}$ and making an angle $\alpha \approx p/k$ ($k = 1/\lambda$) with the optical axis, suffers when passing through the objective lens an additional phase shift $\chi_s(p)$ caused by the spherical aberration of this lens,

$$\chi_s(p) = \frac{\pi}{2} C_s \lambda^3 p^4, \tag{45}$$

* In this section the subscript of ψ refers to the z-coordinate.

References: p. 1105.

where C_s is the spherical aberration constant, typically of the order of some mm. Then, in the backfocal plane of the objective lens (subscript "b"), we get the amplitude distribution

$$\tilde{\psi}_b(\boldsymbol{p}) = \tilde{\psi}_0(\boldsymbol{p}) \cdot \exp - i\chi(p), \tag{46}$$

with the wave aberration function

$$\chi(p) = \chi_f + \chi_s = 2\pi\left[\tfrac{1}{2}\lambda\delta p^2 + \tfrac{1}{4}C_s\lambda^3 p^4\right]. \tag{47}$$

The objective aperture in the backfocal plane (assumed to be circular in shape) cuts off all partial waves making an angle $\alpha > \alpha_{max}$ with the optical axis. With $\alpha = p/k$, see above, we introduce an aperture function $A(p)$ by

$$\begin{aligned} &A(p) = 1 \quad \text{for } p \lesssim p_{max} = k\alpha_{max}, \\ &A(p) = 0 \quad \text{for } p > p_{max}. \end{aligned} \tag{48}$$

Thus, immediately behind the objective aperture we are left with

$$\tilde{\psi}_b(\boldsymbol{p}) = \tilde{\psi}(\boldsymbol{p}) \cdot T(p), \qquad T(p) = A(p) \cdot \exp - i\chi(p). \tag{49}$$

$T(p)$ is called the optical transfer function. The subsequent imaging system (intermediate lenses, projector lens) can be assumed to be aberration-free. Thus the final image amplitude on the viewing screen (or on the photographic plate) is found from $\tilde{\psi}_b(\boldsymbol{p})$ according to eq. (49) by an inverse Fourier transformation,

$$\psi_m(\hat{\boldsymbol{r}}) = \int \tilde{\psi}_0(\boldsymbol{p}) \cdot T(p) \exp\left(-2\pi i \frac{\boldsymbol{p} \cdot \hat{\boldsymbol{r}}}{M}\right) d^2\boldsymbol{p}, \tag{50}$$

where the index "m" stands for "magnified" image, M = electron–optical magnification. [The same sign of the exponents of the first, eq. (42), and of the inverse, eq. (50), Fourier transformations accounts for the 180° rotation between the diffraction pattern and the image, cf. § 2.]

9.3. Consequences of the wave aberration

By the objective aperture all partial waves $\tilde{\psi}_0(\boldsymbol{p})$ with $p > p_{max}$ are cut off. These partial waves are produced by structural details of $\psi_0(\hat{\boldsymbol{r}})$ of "wavelengths" $\lesssim 1/p_{max}$. Accordingly, these structural details are completely suppressed in $\psi_m(\hat{\boldsymbol{r}})$. Moreover, the phases of the partial waves $\tilde{\psi}_0(\boldsymbol{p})$ with $p \le p_{max}$ are shifted by the wave-aberration term $\chi(p)$ so that, after recombination of these partial waves in the image plane, the "structure" of $\psi_m(\hat{\boldsymbol{r}})$ may be significantly different from $\psi_0(\hat{\boldsymbol{r}})$ even for structures of wavelength $> 1/p_{max}$. For illustration, we calculate $\chi(p)$ assuming $C_s = 3$ mm, $\lambda = 3.7$ pm, $\delta = 0$ (exactly focused objective lens), and $p = p_1 = (1\text{ nm})^{-1}$ and $p_2 = (0.7\text{ nm})^{-1}$. From eq. (47) we find $\chi(p_1) = 13.7°$ and $\chi(p_2) = 57°$. Accordingly, in this case the wave aberration is probably negligible for the imaging of structures of dimension of 1 nm and above, but it becomes important near 0.7 nm or below. Obviously, the deteriorating influence of $\chi(p)$ on the image formation can be reduced for a certain p-interval if both terms of $\chi(p)$

act against one another: Since C_s is always >0, it is expedient to operate the microscope with $\delta < 0$, i.e., with an appropriately underfocused objective lens. For a special case this will be quantified in the next section.

9.4. The weak-phase object approximation

At present, high-resolution images with an intensity distribution $I_m(\hat{r}) = |\psi_m(\hat{r})|^2$, which are directly interpretable in terms of structural details of the specimen, are only obtainable from extremely thin specimens of thickness $t \lesssim 10$ nm depending on the composition of the specimen. In such cases the weak-phase object approximation may be applicable. It can be shown that, assuming axial illumination, the wave function $\psi_0(\hat{r})$ of sufficiently thin specimens can be represented by

$$\psi_0(\hat{r}) \approx \exp\left[+i\omega_0(\hat{r})\right] \approx 1 + i\omega_0(\hat{r}), \tag{51}$$

where $\omega_0(\hat{r})$, being real and of modulus $\gg 1$, is proportional to the electrostatic potential $V(\hat{r}, z)$ in the specimen projected along the electron beam direction ($=z$-axis, origin in the lower specimen surface!):

$$\omega_0(\hat{r}) \propto \int_{-t}^{0} V(\hat{r}, z)\mathrm{d}z. \tag{52}$$

Then, the formalism outlined in the preceding section leads to:

$$\psi_m(\hat{r}) = 1 + \omega_m^{(r)}(\hat{r}) + i\omega_m^{(i)}(\hat{r}) \tag{53}$$

with

$$\omega_m^{(r)}(\hat{r}) = \int \tilde{\omega}_0(\boldsymbol{p}) \cdot A(p) \cdot \sin\chi(p) \cdot \exp\left(-2\pi i \frac{\boldsymbol{p} \cdot \hat{r}}{M}\right) \mathrm{d}^2\boldsymbol{p}. \tag{54}$$

For $\omega_m^{(i)}$ the same equation is applied with $\sin\chi$ substituted by $\cos\chi$. From $\chi(p) = \chi(-p)$, $A(p) = A(-p)$, and $\omega_0(\hat{r})$ real it follows by symmetry arguments that $\omega_m^{(r)}$ and $\omega_m^{(i)}$ are real as well. Accordingly, in linear approximation of the "perturbation" term $\omega_0(\hat{r})$ we get the image intensity

$$I_m(\hat{r}) \approx 1 + 2\omega_m^{(r)}(\hat{r}), \tag{55}$$

i.e., the only phase-modulated wave function $\psi_0(\hat{r})$ at the specimen [eq. (51)], which would give rise to no contrast in a perfect microscope ($|\psi_0(\hat{r})|^2 = 1$), is converted by the wave aberration term $\sin\chi(p)$ into an amplitude-modulated wave function $\psi_m(\hat{r})$ at the image plane, eq. (55).

In order to keep an utmost similarity between the observable structure of $I_m(\hat{r})$ and the unknown structure of $\psi_0(\hat{r})$, we have to adjust the focusing parameter and the size of the objective aperture in such a way that the partial waves of a maximum range of space frequencies p are transferred through the imaging system by the $\sin\chi(p)$ term with a reasonable value close to 1 or -1 and with the same sign of $\sin\chi(p)$. This is achieved by the socalled *Scherzer focus*, δ_0, which correlates δ to C_s and λ by:

References: p. 1105.

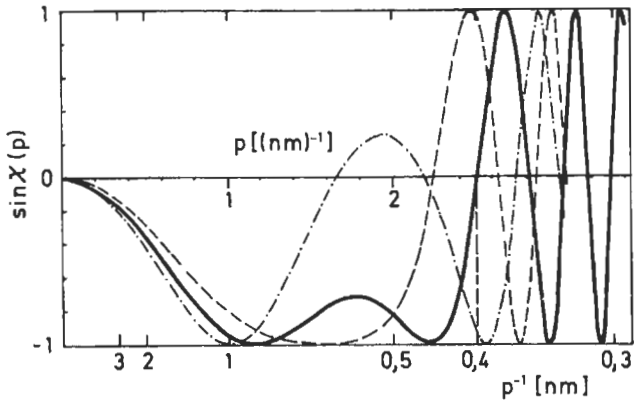


Fig. 27. The transfer function $\sin\chi(p)$ for $C_s = 3$ mm. Solid curve: $\delta = \delta_0 =$ Scherzer focus; dashed curve: $\delta = 0.8\delta_0$, dashed-dotted curve: $\delta = 1.2\delta_0$.

$$\delta_0 = -\left(\tfrac{3}{2} C_s \lambda\right)^{1/2}. \tag{56}$$

Figure 27 displays the term $\sin\chi(p)$ for three values of δ assuming $C_s = 3$ mm and $\lambda = 3.7$ pm. In the case of $\delta = \delta_0$ (solid curve) we find a broad band-pass with $\sin\chi(p)$ close to -1 (the shallow dent near $p = 1.8$ (nm)$^{-1}$ is tolerable) and with the first zero at $p = p_1 = 2.5$ (nm)$^{-1}$. This means that all partial amplitudes of space frequencies with $p < p_1$ are transferred with the same sign, i.e., they may contribute to a correct reconstruction of the image. The partial amplitudes associated to space frequencies $p > p_1$ change in sign due to the $\sin\chi(p)$ term more and more rapidly with increasing p. So their contribution to the final image may be confusing rather than constructive (e.g., they may be washed out by a finite beam divergence). Therefore, the size of the objective aperture should be chosen so that p_{max} (eq. 48) coincides approximately with p_1, thus cutting off the oscillating part of $\sin\chi(p)$. In this way a resolution limit of $1/p_1$ (≈ 0.4 nm in fig. 27) is achieved.

If δ is reduced by 20% to $0.8\delta_0$ (dashed curve) we find a reduction of the first zero to $p_1 = 2.2$ (nm)$^{-1}$ which corresponds to a reduction of the resolution limit to 0.45 nm. If, on the other hand, δ is increased by 20% to $1.2\,\delta_0$ (dashed-dotted curve) the shallow dent near $p = 1.8$ (nm)$^{-1}$ becomes so deep that it crosses the abscissa. As a consequence, the broad band-pass of the solid curve is completely disturbed and the resolution becomes drastically worse. We learn from these examples that fully utilizing the resolving power of an electron microscope requires a precise knowledge of the spherical aberration constant C_σ and a precise control over the actual defocusing value δ.

In the meantime several methods have been developed in order to make the partial waves of space frequencies beyond p_1 also useful for the image interpretation. Here the key-word "image reconstruction" must be sufficient, cf., e.g., HAWKES [1980].

9.5. Some remarks to the high-resolution images of crystalline specimens

The weak-phase object approximation, originally developed for the study of non-crystalline specimens, is valid also for crystalline specimens, provided they are sufficiently thin. The space frequency spectrum $\tilde{\psi}(\boldsymbol{p})$ of a perfect crystalline weak-phase object consists of a discrete set of sharp Bragg diffraction spots. If such a specimen is carefully orientated with the incident beam along a low-order zone axis and if the lattice constant is large compared to that of pure metals or alloys, a large number of diffraction spots may pass through the objective aperture. In the frame of the weak-phase object approximation they form an image which reveals many details on an atomic scale of the (projected) crystal structure. An example is shown in fig. 28. In particular, complicated stacking or shear variants of otherwise comparatively simple crystal structures, which are hardly, or not at all, determinable by X-ray or neutron diffraction techniques, may become easy to analyze on high-resolution TEM. Regarding stacking variants of ordered alloys (antiphase structures) we refer as an example to VAN TENDELOO and AMELINCKX [1977, 1978]. Various applications of high-resolution TEM to crystal-structure research in mineralogy and inorganic chemistry may be found in a book edited by WENK [1976].

Silicon is a substance especially suitable for high-resolution TEM because of its simple structure and because of its comparatively large lattice constant. Detailed information on an atomic scale of split dislocation configurations and of the structure of dislocation cores have been obtained, cf., e.g., SPENCE and KOLAR [1979] and BOURRET *et al.* [1982].

One point of special interest in physical metallurgy concerns the imaging of a local variation of lattice-plane distances as a consequence of a local variation of the chemical composition, e.g., in the course of a spinodal decomposition. SINCLAIR *et al.* [1976],

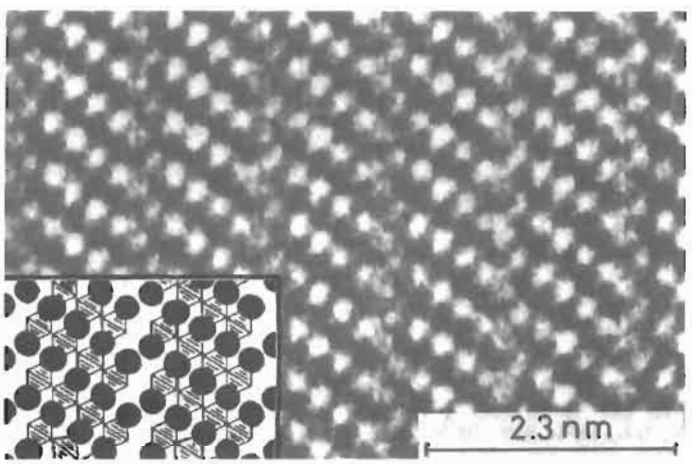


Fig. 28. High-resolution image of $Ba_5Nb_4O_{15}$. (hexagonal unit cell, $a = 0.572$ nm, $c = 1.172$ nm). Projection along $<10\bar{1}0>$. The insert shows the projected structure: solid circles indicate positions of Ba atoms, hatched diamonds indicate NbO_6 octahedra. (Courtesy J. L. HUTCHISON; inset from SHANNON and KATZ [1970].)

References: p. 1105.

studying a spinodally decomposed Ni_3Au alloy, found indeed a periodic modulation of the lattice plane fringes, in accordance with the expectation.

However, subsequent model calculations of COCKAYNE and GRONSKY [1981] have clearly demonstrated that, due to the wave aberration $\chi(p)$, the spatial modulation of the lattice fringes observed on the image may be significantly different from the spatial modulation of the lattice plane spacings in the specimen. Even the number of lattice fringes per modulation period may be different from the corresponding number of lattice planes in the specimen.

In conclusion it is worthwhile to emphasize that at the present state of the art all kinds of high-resolution work, in particular that concerning lattice defects, must be accompanied by extended theoretical work, i.e., by corresponding model calculations. This holds even more if one crosses the border of the restricted regime of applicability of the weak-phase object approximation.

10. Analytical electron microscopy

10.1. Basic considerations

Electron scattering can be divided into two categories: elastic and inelastic scattering. When elastic scattering occurs, the direction of the electron velocity v is changed, but the magnitude $|v|$ remains constant, so that the kinetic energy, $E=\frac{1}{2}m_e v^2$ (m_e=electron mass) is unchanged.

During an inelastic scattering event, energy is transferred to the target atoms and electrons, and the kinetic energy of the beam electron decreases.

Inelastic scattering can happen by: (i) plasma excitation; (ii) excitation of conducting electrons leading to secondary electron (low-energy) emissions; (iii) deceleration of the electron beam in the Coulomb field of an atom (*Bremsstrahlung*); (iv) excitation of phonons; (v) ionization of the inner shells of atoms. Signals arising by the latter process can be used for analytical studies (fig. 29).

In the ionization process of inner shells, a sufficiently energetic electron can interact with an atom and cause the ejection of a tightly bound inner-shell electron, leaving the atom in an ionized and highly energetic state. The energy transferred to the bound electrons must be at least the energy of the bound state of the electron in a shell of the

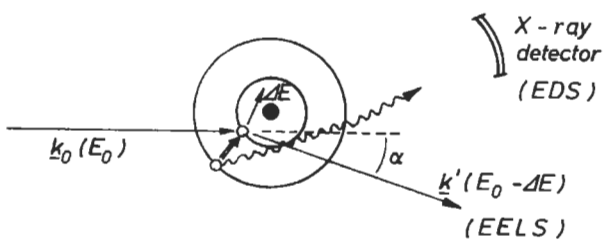


Fig. 29. Inelastic scattenng of incoming electrons at electrons of inner shells (schematic drawing). The characteristic X-rays and the energy losses are used for analytical purposes.

atoms. Subsequent decay of this excitation state results in the emission of characteristic X-rays and Auger electrons. The X-ray yield is plotted in fig. 30 for different elements. Qualitative and quantitative information on the presence of the different elements can be obtained from the wavelength of a characteristic X-ray and from its intensity. A detailed treatment of the properties of characteristic X-rays is found in the literature (e.g., BERTIN [1975] and GOLDSTEIN *et al.* [1981]).

In *electron energy loss spectroscopy* (EELS) the inelastic scattering of the incident beam is studied directly. The intensity of the inelastically scattered beam is measured as a function of energy loss by spectral analysis. At characteristic energy losses the intensities of the inelastically scattered electrons increase due to the ionization of inner shells. From the position of the edge and the increase in intensity, information is again available on the chemical composition. EELS is most powerful for light elements (atomic number $Z < 12$), whereas the analysis of the characteristic X-rays is possible for $Z > 9$.

The cross-section for the inelastic processes (i) to (v) has been calculated (see GOLDSTEIN *et al.* [1981]) for different metal targets as a function of energy. All of the cross-sections are observed to decrease with increasing energy. Considering a range of atomic numbers, inelastic scattering is favoured at low atomic numbers and elastic scattering at high atomic numbers.

Inelastic scattering occurs by a variety of discrete processes, with a variable amount of energy transferred to the solid, depending on the strength of each interaction. Cross-sections for the individual processes are difficult to obtain for all targets of interest. It is useful in many calculations to consider all inelastic processes grouped together to give a "continuous energy loss" (BETHE [1933]).

It is found experimentally that an appreciable fraction of the beam electrons which strike a target subsequently escape where the electrons enter the solid (backscattered

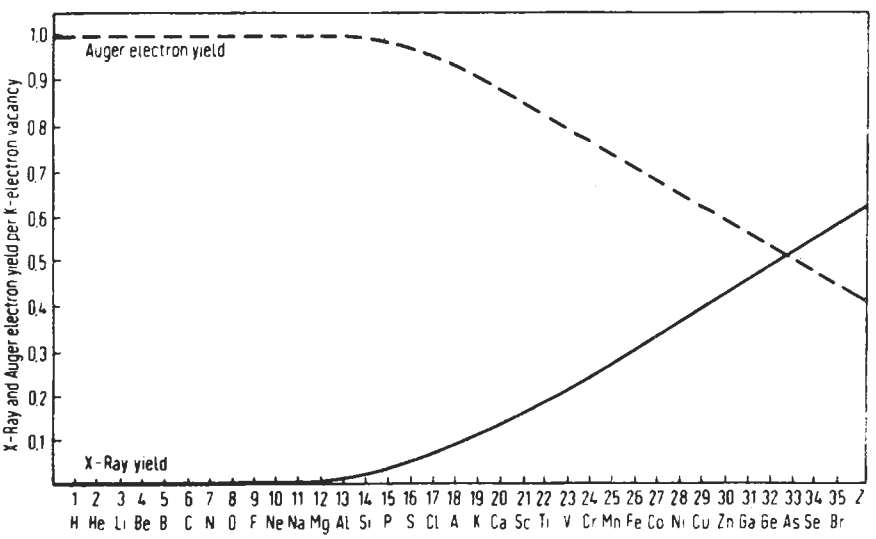


Fig. 30. X-ray yield and Auger-electron yield for K-shell excitation for different elements.

References: p. 1105.

electrons). Backscattering electrons provide an extremely useful signal for imaging in scanning electron microscopy and scanning transmission electron microscopy. For proper image interpretation, it is necessary to understand the properties of the backscattered electrons as a function of the parameters of the beam and the characteristics of the specimen (for a detailed review see NIEDRIG [1978]).

It is well established by experiments and by Monte Carlo calculations that the volume of interaction between a high-energy electron beam and a solid possesses the shape of a pear (fig. 31). The width of the neck of the pear is mainly determined by the shape of the electron beam, while the length of the neck depends on the energy of the electrons as well as on the material investigated. The diameter of the pear can exceed several microns (μm). Therefore, studies with high spatial resolution can only be performed in thin foils.

Those studies are preferentially performed in a conventional TEM fitted with a STEM unit, appropriate detectors for X-rays and energy loss spectra.

10.2. Quantitative analytical electron microscopy of thin foils: analysis of X-rays

In X-ray microanalysis the detector is as close as possible to the specimen. Usually, the detectors are energy-dispersive devices, which offer the advantages of speed of operation and ease of attachment to the system. Crystal spectrometers (wavelength-dispersive devices) are usually preferred for quantitative studies. However, it is very cumbersome to attach those to a transmission electron microscope. An example of an EDS X-ray microanalysis spectrum in the STEM mode of a TEM is shown in fig. 32.

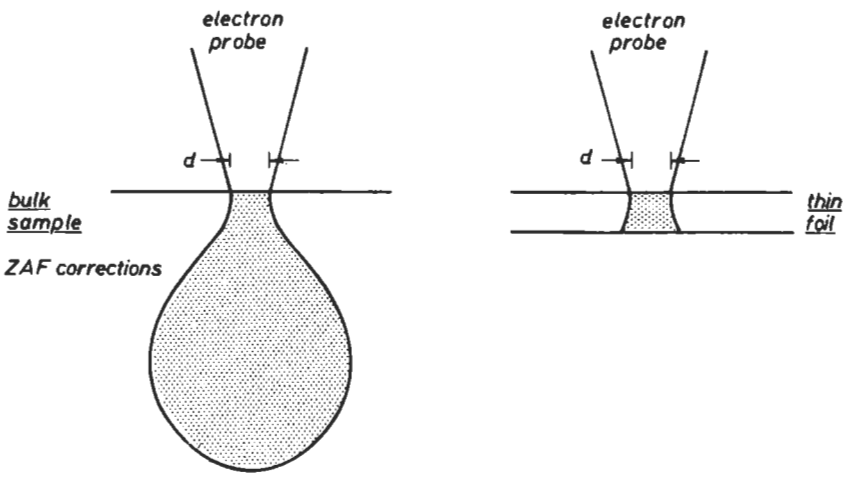


Fig. 31. Schematic representation of the interaction of a high-energy electron beam with a bulk sample and a thin foil, respectively To the first approximation absorption and fluorescence effects can be ignored in the thin specimen.

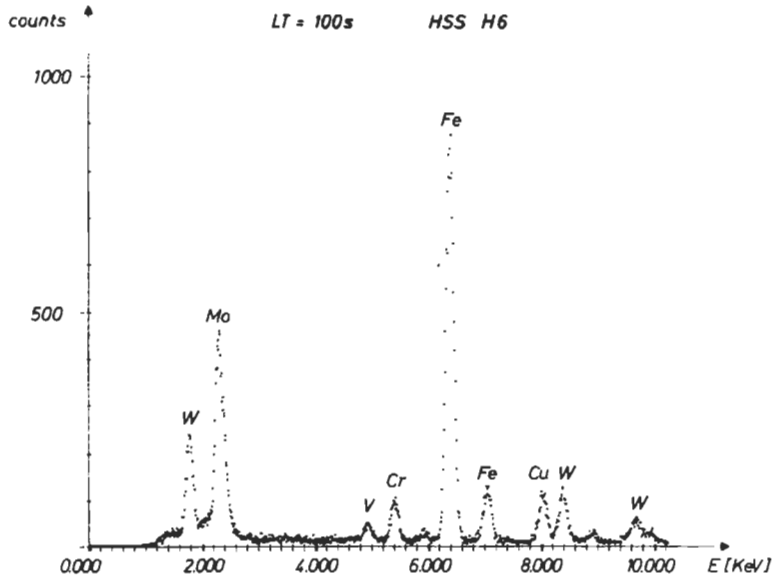


Fig. 32. STEM X-ray spectrum from a low-carbon steel. The Cu peak is caused by the Cu grid of the TEM.

10.2.1. Cross-section for inner-shell ionization

Numerous cross-sections for inner-shell ionization can be found in the literature; this has been reviewed by POWELL [1976]. The number of X-ray photons produced per incident electron in a thin foil of thickness Δx can be predicted from

$$n_x = \sigma w N_0 \frac{1}{A} \rho \Delta x, \tag{57}$$

where σ is the ionization cross-section, w the fluorescence yield (see fig. 30), N_0 Avogadro's number, A the atomic weight, and ρ the density of the material. Experimental measurements of the X-ray yield (absolute efficiency) give a range of $\sim 0.11 \times 10^{-4}$ for Mg to $\sim 0.933 \times 10^{-4}$ for Ni in units of X-ray photons per electron and per angular unit.

10.2.2. Thin-film approximation

If the sample is sufficiently thin to carry out quantitative transmission electron microscopy at ≥ 100 kV, the activated volume is approximately a cylinder equal to the beam diameter (see fig. 31); that means a high lateral resolution can be achieved. This high lateral resolution, however, results in a disadvantage in the use of thin specimens in that activated volume for X-ray production is small compared to a bulk-sample; x-ray counting rates are low and hence the potentioal accuracy is inferior to that generally accepted for bulk specimens in scanning electron miscroscopy.

CLIFF and LORIMER [1975] showed that for a thin foil, to a first approximation X-ray absorption and fluorescence in the specimen can be neglected. Therefore, the ratio of two

observed X-ray intensities I_A/I_B can be related to the corresponding weight fraction ratio c_A/c_B by the equation

$$c_A/c_B = k_{AB} I_A/I_B, \tag{58}$$

where k_{AB} is a constant at a given accelerating voltage and is independent of specimen thickness and composition. A normalization procedure, e.g., $\Sigma c_n = 1$, must be used to convert the ratios of the weight fractions into weight percentages. In mineral specimens, which are frequently used for calibration, assumptions must also be made concerning oxidation states, e.g., it is impossible to differentiate between Fe_3O_4 and Fe_2O_3 if ratios are measured, and oxygen cannot be detected (LORIMER *et al.* [1981]).

For quantitative analysis the parameters k_{AB} in eq. (58) have to be determined experimentally and also compared to theoretical results. There are discussions in the literature on the reliability of different kAB values available (e.g., GEISS [1981]). Furthermore the effects of absorption of the X-rays and fluorescence in the specimen have to be considered for a reliable quantitative analysis (see § 10.2.4).

10.2.3. Beam-spreading in the specimen

Beam-spreading in thin foils is currently an active field of research in both experimental investigations and theoretical analysis (LORIMER *et al.* [1981] and GEISS [1981]).

A simple approach to the problem of beam-spreading has been proposed by GOLDSTEIN *et al.* [1977] who assumed that a single elastic scattering event occurs at the centre of the foil, and defined the X-ray source as that volume in which 90% of the electron trajectories lie. Experimental observations surprisingly show good agreement with the simple theory by GOLDSTEIN *et al.* [1977]. The beam-spreading limits the spatial resolution of analytical TEM, as shown in table 1.

10.2.4. Errors limiting the data of X-ray analysis

X-ray counting statistics are usually assumed to obey Gaussian behaviour and at the 2σ confidence level the relative error in the number of counts I is $2\sqrt{I}$. Using eq. (58)

Table 1
Spatial resolution for X-ray microanalysis at 120 kV for incident electron probe sizes defined by $d = 0$ nm and $d = 5$ nm (after LORIMER *et al.* [1981]).

Foil thickness (nm)	d (nm)	Spatial resolution			
		C	Al	Cu	Au
40	0	–	1	4	10
	5	22	22	22	23
100	0	4	6	17	38
	5	22	22	26	42
200	0	11	18	47	107
	5	24	27	50	108
400	0	30	50	132	302
	5	35	53	134	302

a ratio of two counts is converted into a weight-fraction ratio via the constant k_{AB}. Thus it is necessary to add the total relative errors in I_A, I_B and k_{AB} to obtain the relative error in the weight fraction c_A/c_B would be 6%. Unfortunately, X-ray counts often a premium in thin specimens, and it may be necessary to accept only a few hundred counts for one element. If, for example, I_A contained only 900 counts and I_B stilled equalled 10000, then the relative error in the c_A/c_B ratio would be ± 10%. The statistical error was often not taken into account in quantitative analytical electron microscopy (AEM). Therefore, materials-science interpretation of AEM work is bases on weak or incorrect foundations.

In order to make corrections for absorption, fluorescence and beam spreading, it is necessary to know the thickness of the sample. Various parallax techniques can be used, including the contamination spots formed on the top and the bottom of the foil during the analysis. However, it seems that the most accurate technique is that of convergent-beam diffraction as described by AMELINCKX [1964] and KELLY *et al.* [1975]. This enables specimen thickness to be determined with an accuracy of ± 2%.

Many investigators have been aware that the measured X-ray spectrum may contain X-rays generated from other than the sample region excited by the focused electron beam. These additional X-rays are called "spurious" or "extraneous" radiation. This problem has received much attention (cf. GOLDSTEIN and WILLIAMS [1981]). The spurious radiation can be avoided by a correct experimental arrangement of the specimen in the analytical electron microscope. The absence of spurious radiation can be checked by a hole count.

10.2.5. Examples

Since about 1979 reliable results obtained by AEM appeared. CHAMPNESS *et al.* [1982] determined the partitioning of various alloying additions during the austenite–pearlite transformation in eutectoid steels. The composition of different precipitates in a variety of steels was obtained and first results on segregation of impurities at grain boundaries and interfaces were reported. More results are published in the proceedings of recent workshops and conferences on AEM (HREN *et al.* [1979], LORIMER *et al.* [1981] and GEISS [1981]).

10.3. Quantitative analytical electron microscopy of thin foils — electron energy loss spectroscopy

A transmission microscope fitted with an electron spectrometer can be used to produce energy-filtered images (KRAHL *et al.* [1981]) or for electron energy-loss spectroscopy (EELS) (JOY [1981]). The latter can be used for measurements of inner-shell loss for quantitative microanalysis of light elements.

There are two instrumental principles used nowadays. They are based on (i) the use of a magnetic analyzer below the camera chamber of a conventional TEM or a TEM with STEM attachments (KRIVANEK and SWANN [1981]) and (ii) the use of an analyzer between the objective and the intermediate lens in a conventional TEM (ZANCHI *et al.* [1977]). Figure 33 shows schematically the energy spectrum of 100 kV electrons after transmission through a specimen of about 25 nm thickness. The elastically or (by

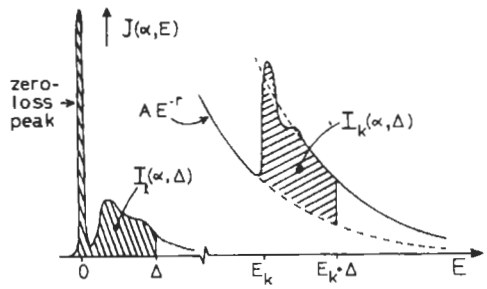


Fig. 33. Schematic electron energy-loss spectrum of the signal intensity $I(E)$ versus energy loss E. The hatched area close to the energy E_K is proportional to the number of atoms present in the specimen.

phonons) quasi-elastically scattered electrons give rise to a zero-loss peak whose width provides an indication of the instrumental energy resolution. For small energy losses (< 100 eV) surface and bulk plasmon losses are observed (JOUFFREY *et al.* [1978]). For higher energy losses E, inelastic scattering is caused at electrons of the outer atomic shells (e.g., valence electrons). The intensity of the inelastically scattered electrons decays rapidly, approximately as AE^{-r}, where A is a constant and r is between 2 and 6. At an energy loss equal to the ionization energy E_k of an inner atomic shell k (where k = K, L, M) the energy loss intensity increases rather abruptly, followed by a slower decay. By determining A and r from a region just preceding the ionization edge, the outer-shell background can be subtracted (by a computer, or manually using a log–log plot). The area $I_k(\alpha, \Delta)$ above the background is measured up to an energy loss $E_k + \Delta$ (see fig. 33). The number of atoms per unit area of the corresponding element is then given by (JOY [1981])

$$N = \left[I_k(\alpha, \Delta)/I_1(\alpha, \Delta)\right]/\sigma_k(\alpha, \Delta), \tag{59}$$

where $\sigma_k(\alpha, \Delta)$ is a partial cross-section for the ionization of shell k within an energy range Δ and for scattering angles up to α. The angle α is determined by the collecting angle of the spectrometer. For a quantitative analysis $\sigma_k(\alpha, \Delta)$ can be determined experimentally or by calculations (EGERTON [1979] and JOY [1981]).

Unlike X-ray or Auger microanalysis, energy-loss spectroscopy therefore provides a convenient standardless method for determining the absolute number N of atoms of light elements. In addition, the ratio of two (or more) elements may be determined from the appropriate ionization edge, provided an equal integration range Δ is chosen for both elements. For that case $I_1(\alpha, \Delta)$ will cancel in eq. (59) so that the low-loss region needs not to be measured.

As the specimen thickness increases, multiple-scattering contributions to the background beneath an inner-shell edge cause the background to increase faster than the inner-shell "signal", so that the signal/background ratio falls, which reduces the visibility of an edge and may make the background-fitting and subtraction process insufficiently accurate (JOY [1981]). Therefore, judging by the criterion of signal/background ratio, the specimen should be as thin as possible within the analyzed region.

However, by making the sample very thin we reduce the amount of valence-electron- and inner-shell scattering, so that the relative noise level in both signal and background is increased. Consequently, the signal/noise ratio (due to the finite number of electrons detected within the energy window chosen for analysis) initially increases with thickness and then goes through a maximum at a sample thickness of the order of Λ_i, the total mean free path for inelastic scattering. This last statement however ignores the practical problem of separating the signal from the background (i.e., background-fitting and subtraction) which generally becomes more difficult as the thickness increases and the signal/background ratio deteriorates. In practice it is generally found that the optimum specimen thickness is closer to $\Lambda_i/2$, i.e., 60 nm for 100 keV electrons (EGERTON [1979, 1981]), 50–100 nm for 200 keV, and 100–200 nm for 1000 keV incident energy (JOUFFREY *et al.* [1978]).

The sensitivity of a microanalytical technique is an important factor in determining its usefulness. For EELS, as for EDS, the limiting condition will be the statistical requirement for a minimum signal-to-noise ratio between the edge and the background. Because of the dominance of the background in the spectrum the sensitivity will depend very much on the experimental parameters such as beam current, recording time and acceptance angle. Consequently, detailed calculations are necessary in order to obtain realistic results. JOY and MAHER [1980] calculated the sensitivity for different atoms distributed in a carbon matrix, under a variety of experimental conditions typical for STEM operation.

Although the absolute detection limits are very good, typically 10^{-19} g, this represents, however, a mass- or atomic fraction of only between 0.5 and 5% in most cases. This indicates that EELS is best suited for studies of precipitates, segregation effects and similar situations where the local mass fraction is high or can be made so by choosing the probe diameter. It can also be seen that much of the potential gain in microanalytical sensitivity derived from EELS operation is sacrificed when the normal sequential mode of spectrum recording is used, since the analysis time is only of the order of a few hundred milliseconds per channel compared with 100 s or more for EDS. However, recent developments in parallel detector arrays which acquire data simultaneously in many channels will ultimately remove this disability and yield a substantial gain in sensitivity.

10.3.1. Examples

EELS is applied mainly to ceramic specimens since it is the only technique for the determination of the distribution of light elements; an example is shown in fig. 34, below. Grain-boundary phases could be detected and analyzed chemically (GEISS [1981], RÜHLE and PETZOW [1981]). Ion implantation is sometimes employed to introduce light-element dopants into semiconductors. EELS could be used to measure local dopant concentrations, if the latter exceed ~0.1% (JOY and MAHER [1980]). Light elements occur in certain metal specimens in the form of nitride or carbide precipitates; these materials have also been analyzed by the EELS technique (ROSSOUW *et al.* [1976], GEISS [1981]).

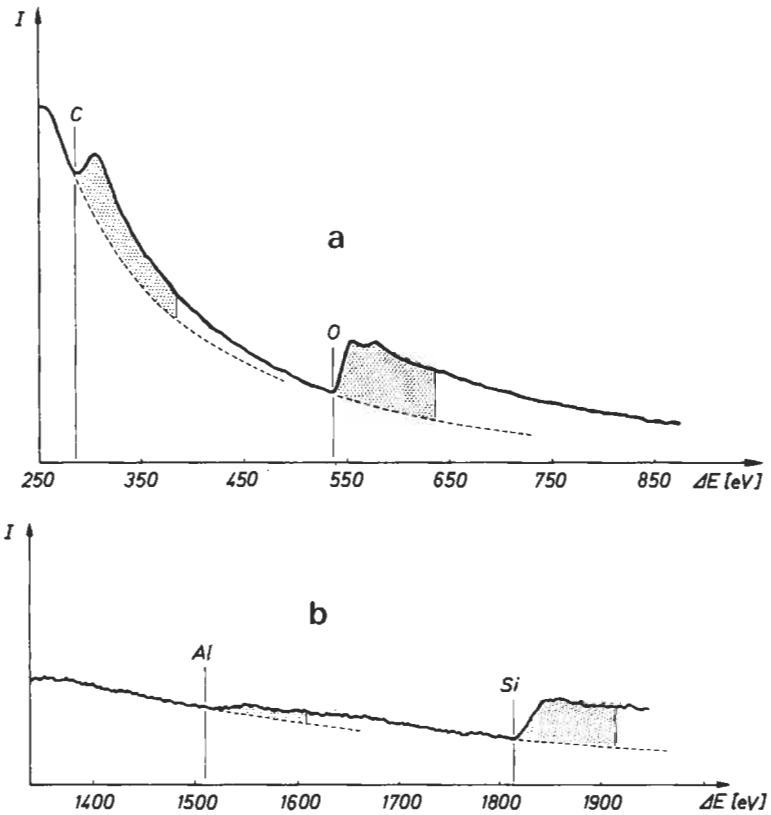


Fig. 34. Electron energy-loss spectrum of a 25 nm thick amorphous area in a silicon nitride ceramic. Edges in the spectrum are found at characteristic energy losses. The concentration of the different elements can be determined from the hatched areas: a) Energy losses 250–850 eV. b) Energy losses 1350–1950 eV. (From BISCHOFF [1983].)

Acknowledgement

The authors would like to express their sincere thanks to all their colleagues in and outside Stuttgart who have helped them in preparing the present article by critical comments and/or by providing them with suitable micrographs. The latter holds in particular for their colleagues Dr. Bourret, Dr. Hutchison and Dr. Karnthaler.

Appendix. Elements of the kinematical diffraction theory

A.1. Introduction

In materials science, scattering or diffraction experiments using some kind of waves of wavelengths X comparable to, or smaller than, interatomic distances play a key-role

for the study of the atomistic structure of solids *. This holds in particular for solids of more or less perfect crystalline structure. Such experiments are carried out either by *X-rays*, as provided by conventional X-ray generators or, more recently, by synchrotron radiation of electron- or positron storage rings, or by the de Broglie waves of thermalized *neutrons*, as provided by nuclear reactors orn more recently, by spallation sources or by the de Broglie waves of high-energy *electrons* as used in a transmission electron microscope (TEM).

One of the basic theories which describe the interaction of the above-mentioned radiation with matter is the so-called *kinematical diffraction theory* which can be derived as a first Born approximation, for X-rays from the Maxwell equations, and for neutron and electron waves from the Schrödinger equation. The kinematical diffraction theory is applicable under the following conditions:

(i) The scattering at individual atoms is *elastic*, which means that the wavelength of the scattered wave is the same as the wavelength of the incoming wave.

(ii) The scattering at atoms occurs *coherently*, i.e., the phase relationship between the scattered wave and the incoming wave is always the same for a given kind of atoms.

(iii) The scattering at individual atoms is so weak that a second scattering of the already scattered wave at other atoms can be neglected.

(iv) The *total diffraction* at an ensemble of atoms remains so *weak* that the intensity of the incoming wave is practically not affected, i.e., all the atoms of the ensemble (specimen) are irradiated with the same intensity.

Under these assumptions a fairly sophisticated kinematical diffraction theory has been developed and applied — using X-ray and neutron waves - to various problems in crystallography and materials science, as set out in ch. 12. However, it turns out that this kind of theory is only of restricted use for TEM investigations of crystalline material: Conditions (iii) and (iv) are in general strongly violated. Therefore, in TEM the theory must be extended to the so-called *dynamical diffraction theory*. Nevertheless, a number of elements of the kinematical diffraction theory will be needed also for understanding TEM of crystalline materials. Therefore, we give here a brief introduction to the general concept of the kinematical diffraction theory, which should be understood as an introduction to the more detailed reports in this chapter and ch. 12.

The amplitudes of X-rays are vectors, because X-rays are electromagnetic waves, whereas the amplitudes of neutron and electrons are scalar. This gives some marginal differences in the formulation of the kinematical diffraction theory for these two kinds of radiation. For simplicity we consider here waves of scalar amplitudes only.

A.2. Fundamental equations

The *incident plane wave* $A(\boldsymbol{r}, t)$ is given by

* In the following the term *scattering* is used for interaction of *individual* scattering centres, as e.g., individual atoms, with radiation. On the other hand, we speak about *diffraction* if the coherent superposition of scattered waves of *more* than one scattering centre is considered.

$$A(\boldsymbol{r},t) = A_0 \exp\left[2\pi\mathrm{i}(\boldsymbol{k}_0\boldsymbol{r} - \nu t)\right], \tag{A.1}$$

where $\boldsymbol{r}$ is the space vector, t the time coordinate, $\boldsymbol{k}_0$ the wave vector with $|\boldsymbol{k}_0| = k = 1/\lambda$, λ the wavelength and ν the wave frequency. Further, A and A_0 are scalar terms with $|A_0|^2 =$ intensity per unit area perpendicular to $\boldsymbol{k}_0$. From

$$\boldsymbol{k}_0 \cdot \boldsymbol{r} - \nu t = \text{const.} \tag{A.2}$$

we obtain the *phase velocity* v in direction of $\boldsymbol{k}_0$,

$$v = \nu/k = \nu \cdot \lambda. \tag{A.3}$$

Since we consider only elastic scattering, where the wavenumber k and the frequency ν remain unchanged by scattering events, we can omit the time-dependent part of A. Thus, in the following we start with

$$A(\boldsymbol{r}) = A_0 \exp(2\pi\mathrm{i}\boldsymbol{k}_0 \cdot \boldsymbol{r}) \tag{A.4}$$

as the primary wave incident on a number of N atoms. For the moment these atoms are assumed to be point-like. They are characterized by atomic scattering amplitudes f_n and are located at positions $\boldsymbol{r}_n$ near the origin $\boldsymbol{r} = 0$; $n = 1 \cdots N$. At all these atoms, the primary wave excites the emission of spherical waves in coherency with the primary wave. Thus the sum of the scattered waves can be represented by $A_s(\boldsymbol{r})$, with

$$A(\boldsymbol{r}) = -\sum A(\boldsymbol{r}_n) \cdot f_n \cdot \frac{\exp\left[2\pi\mathrm{i}k|\boldsymbol{r} - \boldsymbol{r}_n|\right]}{|\boldsymbol{r} - \boldsymbol{r}_n|}, \tag{A.5}$$

On the righthand side the term $A(\boldsymbol{r}_n)$ represents amplitude and phase of the incident wave at the positions of the atoms. The term behind f_n represents a spherical wave emerging with phase π from the atom positions. It has the dimension of a reciprocal length. Thus, since A_s and A must be of the same dimension, the atomic scattering amplitude f_n must have the dimension of a length. Accordingly, in neutron diffraction f_n is called the "scattering length", b_n.

A detector for receiving the resultant intensity of the scattered, or diffracted, waves may be placed at a position $\boldsymbol{r} = \boldsymbol{R}$ so far away from the N atoms that in the denominator of eq. (A.5) the approximation

$$|\boldsymbol{R} - \boldsymbol{r}| \approx |\boldsymbol{R}| = R \tag{A.6}$$

is applicable. This is the condition for *Frauenhofer diffraction*. Further, we assume that near the detector all spherical wave fronts can be approximated by plane waves with a common wave vector $\boldsymbol{k}_s$ parallel to $\boldsymbol{R}$. Thus, near the detector, we have

$$\boldsymbol{k}_s \cdot (\boldsymbol{R} - \boldsymbol{r}_n) = k|\boldsymbol{R} - \boldsymbol{r}_n| \cos\gamma, \qquad |\boldsymbol{k}_s| = k, \tag{A.7}$$

where γ is the angle between $\boldsymbol{k}_s$ and $\boldsymbol{R} - \boldsymbol{r}n$. Since $\gamma \to 0$ for $\boldsymbol{R} \to \infty$, we set $\cos\gamma = 1$. Then we can rearrange eq. (A.5) with eqs. (A.6) and (A.7) to the form

$$A_s(R) = A_0 \frac{1}{R} F(\boldsymbol{\kappa}) \exp(2\pi i \boldsymbol{k}_s \cdot \boldsymbol{R}), \tag{A.8}$$

with

$$\boldsymbol{\kappa} = \boldsymbol{k}_s - \boldsymbol{k}_0, \qquad |\boldsymbol{\kappa}| = 2k_0 \sin\theta \tag{A.9}$$

and

$$F(\boldsymbol{\kappa}) = \sum f_n \exp(-2\pi i \boldsymbol{\kappa} \cdot \boldsymbol{r}_n). \tag{A.10a}$$

$\boldsymbol{\kappa}$ is called the *diffraction vector* (in neutron diffraction, $\boldsymbol{\kappa}$ is often used for $4\pi k \sin\theta$). θ is half the angle between $\boldsymbol{k}_0$ and $\boldsymbol{k}_s$, and $F(\boldsymbol{\kappa})$ may be called the *diffraction amplitude* of the ensemble of scattering atoms. Note that information about this ensemble ($\boldsymbol{r}_n$,f_n) is only contained in $F(\boldsymbol{\kappa})$ and that F depends only on $\boldsymbol{\kappa}$ but not explicitly on the moduli and directions of $\boldsymbol{\kappa}_0$ and $\boldsymbol{\kappa}_s$. Equation (A.10a) may easily be extended to a continuous distribution of the atomic scattering amplitude. If $\rho_f(\boldsymbol{r})$ denotes the scattering amplitude per unit volume at the point $\boldsymbol{r}$ then eq. (A.10a) is modified to

$$F(\boldsymbol{\kappa}) = \int \rho_f(\boldsymbol{r}) \exp(-2\pi i \boldsymbol{\kappa} \cdot \boldsymbol{r}) \mathrm{d}^3\boldsymbol{r}. \tag{A.10b}$$

The detector may receive a diffracted intensity I_s via a window of area Q, with $Q/R^2 \ll 1$. Then I_s is given by:

$$I_s(\boldsymbol{\kappa}) = |A_s(\boldsymbol{R})|^2 Q = |A_0|^2 |F(\boldsymbol{\kappa})|^2 \frac{Q}{R^2}. \tag{A.11}$$

$Q/R^2 = \Delta\Omega$ is the solid-angle interval under which the detector receives radiation from the scattering atoms. Thus, $|F(\boldsymbol{\kappa})|^2$, which has the dimension of an area, will be called the *differential diffraction cross-section* of the ensemble, i.e., the cross-section per unit solid angle, $\mathrm{d}\sigma/\mathrm{d}\Omega$, which is a function of $\boldsymbol{\kappa}$ only. Accordingly, the significance of $|F(\boldsymbol{\kappa})|^2$ is rather general, and it is in particular free of "instrumental" parameters of the experiment such as the actual values of A_0, R and Q. Therefore, in the following we concentrate mainly on a more detailed evaluation of $F(\boldsymbol{\kappa})$ or $|F(\boldsymbol{\kappa})|^2$.

A.3. Real space and reciprocal space, description of perfect crystal structures

We assume now that the atoms are periodically distributed, i.e., they form a perfect crystal lattice. The elementary cell of the lattice is given by three non-coplanar basis vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$, or briefly $\boldsymbol{a}_j$, $j = 1, 2, 3,$. For simplicity we assume here that the crystal is shaped like a parallel epiped, i.e., in the direction of $\boldsymbol{a}_j$ the elementary cells are repeated N_j times so that the lattice consists of $N_1\, N_2\, N_3$ unit cells. The elementary cell contains M atoms, associated with atomic scattering amplitudes f_m, $m = 1 \cdots M$. Then the positions of the lattice atoms are given by:

$$\boldsymbol{r}_{nm} = \boldsymbol{r}_n + \boldsymbol{r}_m, \tag{A.12}$$

where r_n denotes the position of the elementary cell,

$$\boldsymbol{r}_n = \sum_j n_j \boldsymbol{a}_j, \qquad 1 \geqslant n_j \geqslant N_j. \tag{A.13}$$

and $\boldsymbol{r}_m$ denotes the position of the mth atom within the elementary cell. We express $\boldsymbol{r}_m$ in dimensionless coordinates ξ_j related to the $\boldsymbol{a}_j$ as:

$$\boldsymbol{r}_m = \sum_j \xi_j^m \boldsymbol{a}_j, \qquad 1 \geqslant m \geqslant M \tag{A.14}$$

Besides the basis vectors $\boldsymbol{a}_j$ we introduce so-called *reciprocal basis vectors* which constitute the *reciprocal lattice*:

$$\boldsymbol{b}_1 = \frac{\boldsymbol{a}_1 \times \boldsymbol{a}_3}{V_e}, \qquad \boldsymbol{b}_2 = \frac{\boldsymbol{a}_3 \times \boldsymbol{a}_1}{V_e}, \qquad \boldsymbol{b}_3 = \frac{\boldsymbol{a}_1 \times \boldsymbol{a}_2}{V_e}, \tag{A.15}$$

where $\times$ means the vector product and V_e the volume of the elementary cell. The $\boldsymbol{b}_j$ are related to the $\boldsymbol{a}_j$ by:

$$\boldsymbol{a}_j \cdot \boldsymbol{b}_{j'} = \delta_{jj'}, \tag{A.16}$$

where

$$\delta_{jj'} = 1 \text{ for } j = j',$$
$$= 0 \text{ for } j \neq j'.$$

Note that the $\boldsymbol{b}_j$ have the dimension of a reciprocal length, as is also true for the wave vectors $\boldsymbol{k}_0$, $\boldsymbol{k}_s$ and the scattering vector $\boldsymbol{\kappa}$. So $\boldsymbol{\kappa}$ can be expressed by a linear combination of the $\boldsymbol{b}_j$,

$$\boldsymbol{\kappa} = \sum_j h_j \boldsymbol{b}_j, \tag{A.17}$$

where the h_j are dimensionless numbers. Within the $\boldsymbol{\kappa}$-space, or reciprocal space, the reciprocal lattice points given by h_j = integers play a special role. We denote them by $\boldsymbol{g}$ or, if necessary, by $\boldsymbol{g}_{hkl}$, with

$$\boldsymbol{g} = \boldsymbol{g}_{hkl} = h\boldsymbol{b}_1 + k\boldsymbol{b}_2 + l\boldsymbol{b}_3, \qquad h, k, l = \text{integers}. \tag{A.18}$$

The integers h, k, l are the well-known *Miller indices* of a given set of lattice planes. In standard text books of crystallography it is shown that $\boldsymbol{g}_{hkl}$ is perpendicular to the set (hkl) of lattice planes and that

$$|\boldsymbol{g}_{hkl}| = 1/d_{hkl}, \tag{A.19}$$

where d_{hkl} is the repetition distance of the lattice planes (hkl).

In the following we will denote reciprocal lattice points, depending on the context, by $\boldsymbol{g}_{hkl}$, or $\boldsymbol{g}$, or simply by (hkl). If the h, k, l refer to the ensemble of crystallographically equivalent reciprocal lattice vectors we use the symbol $\{hkl\}$. If a particular one is meant we write (hkl). For instance, concerning a crystal of cubic symmetry, $\{100\}$ comprises six $\boldsymbol{g}$-vectors, namely $\pm(100)$, $\pm(010)$, and $\pm(001)$. In quite a similar way we will denote vectors in real space in an abbreviated form by $\boldsymbol{r} = [uvw]$ with the meaning

$$\boldsymbol{r} = u\boldsymbol{a}_1 + v\boldsymbol{a}_2 + w\boldsymbol{a}_3. \tag{A.20}$$

If the ensemble of all crystallographically equivalent vectors $\boldsymbol{r}$ is meant, we write $\boldsymbol{r} = \langle uvw \rangle$. For instance, in a cubic crystal the symbol <100> means all vectors of the kind ±[100], ±[010], and ±[001]. Sets of different lattice planes (hkl) with a common direction in real space, e.g., $\boldsymbol{r}_z$, are called a zone, and $\boldsymbol{r}_z$ is the zone axis. The corresponding vectors $\boldsymbol{g}_{hkl}$ constitute a plane in the reciprocal lattice with $\boldsymbol{g}_{hkl} \cdot \boldsymbol{r}_z = 0$. In a cubic crystal, $\boldsymbol{r}_z = [100]$, for instance, is the zone axis for all lattice planes given by $(0kl)$. Note that for crystals of cubic symmetry a real space vector $\boldsymbol{r} = [uvw]$ is parallel to a reciprocal lattice vector $\boldsymbol{g} = (hkl)$ with $u = h$, $v = k$, $w = l$. For crystals of lower symmetry this is, in general, not the case. Considering, for instance, a crystal of tetragonal symmetry ($|\boldsymbol{a}_3| = c \neq |\boldsymbol{a}_1| = |\boldsymbol{a}_2| = a$), it turns out that the real space vector $\boldsymbol{r} = [101]$ is not parallel to the reciprocal lattice vector $\boldsymbol{g} = (101)$. Accordingly, if crystallographic elements of crystals of lower than cubic symmetry have to be denoted, a careful distinction between vectors $\boldsymbol{r}$ in real space and lattice plane normals $\boldsymbol{g}$ in reciprocal space is necessary.

A.4. The kinematical diffraction amplitude F(K) of a perfect crystal

Applying the coordinates of the crystal atoms as given by eqs. (A.12)–(A.14), eq. (A.10a) can be written as

$$F(\boldsymbol{\kappa}) = F_{\mathrm{L}}(\boldsymbol{\kappa}) \cdot F_{\mathrm{s}}(\boldsymbol{\kappa}) \tag{A.21}$$

with

$$F_{\mathrm{L}}(\boldsymbol{\kappa}) = \sum_n \exp\left(-\pi \mathrm{i} \boldsymbol{\kappa} \cdot \boldsymbol{r}_n\right), \tag{A.22}$$

where the summation extends over all n_1, n_2, n_3, i.e., over all unit cells, and with

$$F_{\mathrm{s}}(\boldsymbol{\kappa}) = \sum f_m \exp\left(-2\pi \mathrm{i} \boldsymbol{\kappa} \cdot \boldsymbol{r}_m\right). \tag{A.23}$$

The subscript "L" stands for "lattice amplitudes". Accordingly, F_{L} reflects the outer shape of the crystal and is independent of details of the structure of the elementary cell. The subscript "s" stands for "structure amplitude" by which the structure of the elementary cell is taken into account. Both terms have to be considered separately. First we derive F_{L}. For simplicity we assume an orthogonal elementary cell where the $\boldsymbol{a}_j$ are parallel to $\boldsymbol{b}_j$ and $|\boldsymbol{b}_j| = 1/|\boldsymbol{a}_j|$. (This restriction is not necessary, but facilitates the interpretation of the final result.) We consider the environment of a particular diffraction vector $\boldsymbol{g}$ and write:

$$\boldsymbol{\kappa} = \boldsymbol{g} + \boldsymbol{s}, \quad \boldsymbol{s} = \sum_j s_j \boldsymbol{b}_j / |\boldsymbol{b}_j| = \sum_j s_j a_j \boldsymbol{b}_j. \tag{A.24}$$

Then, with (A.13)–(A.18),

$$\boldsymbol{\kappa} \cdot \boldsymbol{r}_n = \left[n_1 h + n_2 k + n_3 l\right] + \sum_j n_j s_j a_j. \tag{A.25}$$

References: p. 1105.

The term in square brackets is integer and can be omitted in the exponent of eq. (A.22). Thus:

$$F_L(\boldsymbol{\kappa}) = F_L(s) = \sum_{\substack{n_j=1 \\ j=1,2,3}}^{N_j} \exp\left(-2\pi i \sum_j n_j s_j a_j\right). \tag{A.26}$$

For $s_j a_j \ll 1$, the argument in the exponent varies smoothly with varying the integers n_j. Thus, the summations can be replaced by an integration

$$F_L(s) = \frac{1}{V_e} \int_{x_j=0}^{L_j} \exp\left(-2\pi i \sum s_j x_j\right) dx_1\, dx_2\, dx_3 \tag{A.27}$$

$$= \frac{\sin \pi s_1 L_1}{\pi s_1 a_1} \frac{\sin \pi s_2 L_2}{\pi s_2 a_2} \frac{\sin \pi s_3 L_3}{\pi s_3 a_3}, \tag{A.28}$$

with $L_j = N_j a_j$ denoting the length of the crystal in the direction of $\boldsymbol{a}_j$ and $V_e = a_1 a_2 a_3$. In eq. (A.28) a phase factor was omitted, since only $|F_L|^2$ is accessible by experiment.

An indication of the "structure" of the lattice factor $|F_L|^2$ in the reciprocal space is obtained by considering the square of one of the three factors of F_L, e.g.,

$$I_1(s_1) = \frac{\sin^2 \pi s_1 L_1}{\pi^2 s_1^2 a_1^2}. \tag{A.29}$$

The following properties of I_1 can be recognized: $I_1(s_1)$ is sharply peaked at $s_1 = 0$ where I_1 reaches the value N_1^2. In the tails I_1 oscillates and decreases proportional to s_1^{-2}. The first zero on both sides occurs at $|s_1| = 1/L_1 = 1/N_1 a_1$. Accordingly, the half-width of I_1 is about $1/N_1 a_1$. Thus, one may suppose that

$$\int I_1(s_1) ds_1 = \frac{N_1}{a_1}, \tag{A.30}$$

which is indeed the case. (If eq. (A.26) is summed exactly, it turns out that the terms $\pi s_j a_j$ in the denominator of eq. (A.28) have to be replaced by $\sin \pi s_j a_j$. This small "error" in (A.28) is, however, without any practical relevance in general.)

Having understood the structure of $I_1(s_1)$, it is easy to understand the three-dimensional structures of $|F_L|^2$,

$$|F_L|^2 = I_1(s_1) \cdot I_2(s_2) \cdot I_3(s_3). \tag{A.31}$$

Around every reciprocal lattice point $\boldsymbol{g}$ the intensity of $|F_L|^2$ is sharply peaked, and the shape of the peak is reciprocal to the shape of the crystal: parallel to a direction where the crystal is thin (L_j small), $|F_L|^2$ is broad, and parallel to a direction where the crystal is large (L_j large) $|F_L|^2$ is narrow. For instance, a plate-like crystal gives a rod-like structure of the peaks of $|F_L|^2$ and vice versa.

Now we proceed to the term $F_s(\boldsymbol{\kappa})$, cf. eq. (A.23). We recall that for $N_1, N_2, N_3 \gg 1$, the lattice factor $|F_L|^2$ is essentially $\neq 0$ only for $\boldsymbol{\kappa}$ close to $\boldsymbol{g}$, or for $(h_1 h_2 h_3)$ close to (hkl).

Further, we realize from eqs. (A.14) and (A.23) that F_s is, comparative to F_L, a smooth function of $\boldsymbol{\kappa}$ [or $(h_1h_2h_3)$]. Therefore, considering $\boldsymbol{\kappa}$ close to $\boldsymbol{g}$ it is allowed to approximate F_s by

$$F_s(h_1h_2h_3) \approx F_s(hkl) = \sum_{m=1}^{M} f_m \exp\left[-2\pi i\left(h\xi_1^m + k\xi_2^m + l\xi_3^m\right)\right]. \tag{A.32}$$

We consider $F_s(hkl)$ for some simple cubic lattices.

(i) face-centred cubic (fcc) lattice; four equivalent atoms in the unit cell are located at $r=[000]$, $\frac{1}{2}[011]$, $\frac{1}{2}[101]$, $\frac{1}{2}[110]$. Thus:

$$F_s(hkl) = f\{1+\exp[-i\pi(k+l)]+\exp[-i\pi[h+l)]+\exp[-i\pi(h+k)]\}$$
$$= 4f \text{ for } h, k, l \text{ unmixed, i.e. all even or all uneven}$$
$$= 0 \text{ for } h, k, l \text{ mixed.}$$

(ii) body-centred cubic (bcc) lattice; two equivalent atoms at $r=[000]$, $\frac{1}{2}[111]$. Then:

$$F_s(hkl) = 2f \text{ for } h+k+l = \text{even},$$
$$= 0 \text{ for } h+k-l = \text{uneven}.$$

For these two types of Bravais lattices, F_s takes either full value (f multiplied with the number of atoms per unit cell) or vanishes.

(iii) Ordered lattice of type Ll_2 (e.g. Cu_3Au); one atom of kind A in $r=[000]$, three atoms of kind B in $r=\frac{1}{2}[011]$, $\frac{1}{2}[101]$, $\frac{1}{2}[110]$. This gives:

$$F_s(hkl) = f_A + 3f_B \text{ for } hkl \text{ unmixed},$$
$$= f_A - f_B \text{ for } hkl \text{ mixed}.$$

Lattice plane of the first category are called fundamental lattice planes and those of the second category are *superlattice* planes. The F_s-values of the latter are in general much smaller than those of the former.

A.5. The Ewald sphere and Bragg's law

In order to decide whether a crystal gives rise to a diffracted beam for a given direction and a given wavelength of the incident radiation we introduce the so-called *Ewald Sphere* in the reciprocal lattice. Figure 35 gives a two-dimensional representation. A reciprocal lattice is schematically drawn and the wave vector $\boldsymbol{k}_0$ of the primary beam is inserted in such a way that $\boldsymbol{k}_0$ ends at the origin of the reciprocal lattice. A sphere, the Ewald sphere, is drawn around the starting poine of $\boldsymbol{k}_0$ (point C) with $|\boldsymbol{k}_0| = k$ as radius. All vectors starting from C and ending on the Ewald sphere represent possible wave vectors $\boldsymbol{k}_s$ of diffracted waves of intensity $|F(\boldsymbol{\kappa})|^2$, with $\boldsymbol{\kappa} = \boldsymbol{k}_s - \boldsymbol{k}_0$. If the Ewald sphere cuts a reciprocal lattice point $\boldsymbol{g}$ where $|F(\boldsymbol{\kappa})|^2$ takes high values a strong intensity will be diffracted in the direction of the particular wave vector $\boldsymbol{k}_s$, named $\boldsymbol{k}_g$, pointing from C to $\boldsymbol{g}$. We denote the angle between $\boldsymbol{k}_0$ and $\boldsymbol{k}_g$ by $2\theta_{hkl}$. Then we have:

$$|g_{hkl}| = 2k \sin\theta_{hkl}, \tag{A.33}$$

or with eq. (A.19):

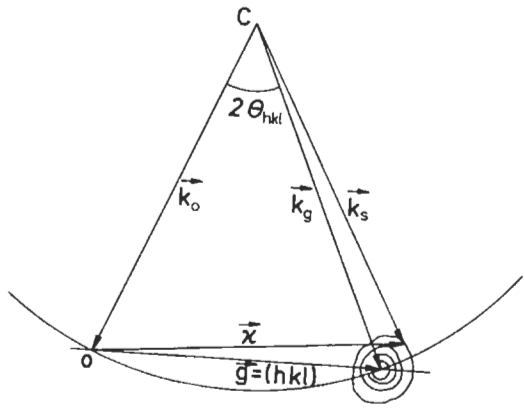


Fig. 35. The Ewald sphere and the intensity distribution $|F(\kappa)|^2$ in the reciprocal space.

$$\lambda = 2d_{hkl} \sin\theta_{hkl}. \tag{A.34}$$

This is *Bragg's law* with θ_{hkl} being the Bragg angle of the set of lattice planes (hkl).

These *Bragg peaks* (or "Bragg reflections") are not infinitely sharp but rather reveal a certain finite width (reciprocal to the crystal dimensions). Therefore, appreciable intensity will be diffracted not only for *one* discrete direction of $\boldsymbol{k}_0$ but for a (more or less) small interval of directions. In the same way the wave vectors $\boldsymbol{k}_s$ of the diffracted beam will cover a similarly small interval of directions. Consequently, for a particular direction of $\boldsymbol{k}_0$ the diffracted intensity I_{hkl} is given by a two-dimensional integration,

$$I_{hkl} = \int |F(\boldsymbol{\kappa})|^2 \frac{\mathrm{d}f_E}{k_2} \tag{A.35}$$

over the Ewald sphere in the neighbourhood of $\boldsymbol{g}_{hkl}$. $\mathrm{d}f_E/k^2$ is a differential solid angle $\mathrm{d}\Omega$ with $\mathrm{d}f_E$ = differential area on the Ewald sphere.

In a practical diffraction experiment the crystal will be rocked with respect to $\boldsymbol{k}_0$ over an angular interval so that the Ewald sphere is swept over the peak at the reciprocal lattice point under consideration. If effective source size and detector aperture are sufficiently large to capture all the intensity belonging to $\boldsymbol{g}$, the diffracted intensity I_{hkl} is proportional to the integral over the diffraction spot,

$$I_{hkl} \propto \int |F(\boldsymbol{\kappa})|^2 \mathrm{d}^3\boldsymbol{\kappa} = \frac{N}{V_c} |F_s(hkl)|^2, \tag{A.36}$$

cf. § A.4, where the integration runs over the neighbourhood of $\boldsymbol{g}_{hkl}$, N is the number of unit cells and $|F_s(hkl)|^2$ is the structure factor of the Bragg reflection (hkl).

A.6. The atomic scattering amplitudes and the Debye–Waller factor

In § A.1 we had assumed that the atoms are point-like scattering centres and that, consequently, the atomic scattering amplitudes, or atomic scattering lengths, are constants

which are characteristic for the kind of atoms considered.

This assumption is a very good approximation for nuclear scattering of thermal neutrons having wavelengths of the order of 0.1 nm. Nuclear scattering occurs at the atomic nuclei the diameters of which are of the order of some 10^{-14} m, i.e., smaller by a factor of $\sim 10^4$ than the neutron wavelengths. In neutron diffraction physics it is customary to denote the atomic scattering length by the symbol b. In general b is positive and of the order of 10^{-14} m; however, for some isotopes b may be negative. Considered as a function of the atomic number Z or the atomic mass A, no smooth dependence of b on Z or A exists. The values of b may even vary considerably for different isotopes of a given element. However, apart from these variations the general trend is that b increases roughly proportional to $A^{1/3}$, i.e., to the size of the nuclei.

X-rays are (nearly) exclusively scattered by the electrons constituting the electron cloud of an atom. From Maxwell's equations it follows that a single electron scatters X-rays of wavelength of the order of 0.1 nm like a point-like scattering centre with a scattering amplitude f_0 with

$$f_0 - e^2/(m_0 c^2) \approx 2.8 \times 10^{-15}\,\mathrm{m}, \tag{A.37}$$

where e is the elementary charge, m_0 is the rest mass of the electron and c is the velocity of light. Around an atomic nucleus, located at $r=0$, the electrons are distributed according to a density function $\rho_e(r)$. If the atom of atomic number Z is neutral,

$$\int \rho_e(\boldsymbol{r})\mathrm{d}^3\boldsymbol{r} = Z. \tag{A.38}$$

Applying eq. (A.10b) we get for the atomic scattering amplitude

$$f_X = f_0 f_a(\boldsymbol{\kappa}),$$

with

$$f_a(\boldsymbol{\kappa}) = \int \rho_e(\boldsymbol{\kappa}) \exp(-\boldsymbol{\kappa}\cdot\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r}. \tag{A.39}$$

The index "X" stands for X-rays. The term f_a is named the *atomic form factor* and is tabulated in general as a function of $|\boldsymbol{\kappa}|/2 = \sin\theta/\lambda$. Obviously, we have $f_a = Z$ for $\sin\theta/\lambda = 0$. With increasing $\sin\theta/\lambda$, f decreases monotonously. In a very rough approximation it decreases to about half its maximum value for $\sin\theta/\lambda \approx 1/d_A$ where d_A is the atomic diameter.

Transmission electron microscopes operate at accelerating voltages U of 100 kV or higher. Taking $U = 100$ kV, we get the corresponding de Broglie wavelength from

$$\lambda = \frac{h}{\left[2m_0 eU\left(1 + \frac{eU}{2m_0 c^2}\right)\right]^{1/2}}, \tag{A.40}$$

giving $A = 3.7$ pm (h = Planck's constant). The wavelengths of such electrons are thus smaller by a factor of about fifty than the wavelengths normally used in X-ray and neutron diffraction work. Because of Bragg's law [eq. (A.34)] the Bragg angles θ_{hkl} are

References: p. 1105.

also smaller by about the same factor. Electrons are scattered at an atom by the atomic electrostatic potential $V(\boldsymbol{r})$ built up by the charge $+Ze$ of the atomic nucleus and by the charge distribution $-e\rho_e(\boldsymbol{r})$ of the surrounding electron cloud. Since charge distribution and potential are linked by Poisson's equation,

$$\nabla^2 V(\boldsymbol{r}) = 4\pi e \cdot \rho_e(\boldsymbol{r}), \tag{A.41}$$

($\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$), the atomic scattering amplitudes for electrons are closely related to f_a. The result is

$$f_e(\sin\theta/\lambda) = \frac{m_0 e^2}{2h^2} \frac{Z - f_a(\sin\theta/\lambda)}{(\sin\theta/\lambda)^2}, \tag{A.42}$$

with $m_0e^2/2h^2 = 2.39 \times 10^8/\text{m}$. f_e decreases rapidly (more rapidly than f_a) for small values of $\sin\theta/\lambda$ but only slowly, $\propto Z/(\sin\theta/\lambda)^2$, for large $\sin\theta/\lambda$ where f_a becomes small compared to Z. The value $f_e(0)$ will not be discussed here.

Diffracted intensities are proportional to the square of the atomic scattering amplitudes. In order to compare the different scattering strengths for neutron-, X-ray- and electron waves, we choose as a representative example the 200 Bragg reflection of a copper crystal. With $\sin\theta_{200}/\lambda = 1/2d_{200}$ we get for copper $\sin\theta_{200}/\lambda = 2.76\ \text{nm}^{-1}$. Taking tabulated values for b, f_a and f_e we obtain:

neutron waves: $b^2 = 5.8 \cdot 10^{-29}\ \text{m}^2$,
X-rays: $f_X^2 = 3.10^{-27}\ \text{m}^2$,
electron waves: $f_e^2 = 8.7 \cdot 10^{-20}\ \text{m}^2$.

It is obvious that neutrons are least effectively scattered and that electron scattering is orders of magnitude stronger than X-ray scattering. For Bragg reflections of high order ($|\boldsymbol{g}_{hkl}|$ large or d_{hkl} small), the difference between neutron and X-ray scattering becomes less spectacular, because f_x decreases, but not b, with increasing $\sin\theta/\lambda$.

There is another effect which reduces the intensities of Bragg reflections with increasing $|\boldsymbol{g}|$ (or decreasing d) even for "perfect" crystals: The atoms are not at rest because of thermal vibrations. This leads to an additional "smearing-out" in space of the scattering power of the individual atoms. If $\langle u_t^2 \rangle$ is the mean square displacement of an atom from its ideal lattice position by thermal vibration the atomic scattering amplitudes f have to be multiplied by a factor e^{-M}, with

$$M = 8\pi^2 \langle u_t^2 \rangle \frac{\sin^2\theta}{\lambda^2}. \tag{A.43}$$

Of course, $\langle u_t^2 \rangle$ increases with temperature T. For $T > \Theta_D$, where Θ_D is the *Debye temperature* (for copper, $\Theta_D = 320$ K), $<u_t^2>$, and thus M, increases about linearly with T. The diffracted intensities are damped by the factor e^{-2M}. Because $M \propto g^2$, this damping is particularly effective for high-order reflections. The factor e^{-M} is named Debye–Waller factor.

A.7. Imperfect crystals

The Debye–Waller factor as an unavoidable consequence of thermal vibrations is but

one example for the effect of displacements of atoms from ideal lattice positions. Apart from thermal displacements, "static" displacements are usually encountered with lattice defects, e.g., stacking faults, point defects, dislocations (including solute atoms), clusters and coherent precipitates. Displacements will generally vary with the position in the crystal. Let $\boldsymbol{r}_n$ denote the position for the nth unit cell of the ideal reference lattice and $\boldsymbol{R}_n = \boldsymbol{R}(\boldsymbol{r}_n)$ the displacement of the first atom in this cell, $\boldsymbol{R}_{nm}$ are the corresponding displacements from the other ideal positions $\boldsymbol{r}_{nm}$ in the nth cell. A factorization of $|F(\boldsymbol{\kappa})|^2$ as in eq. (A.21) is now no longer possible. Instead, we can write:

$$F(\boldsymbol{\kappa}) = \sum_n F_{sn}(\boldsymbol{\kappa}) \exp(-2\pi \mathrm{i} \boldsymbol{\kappa} \cdot \boldsymbol{r}_n) \exp(-2\pi \mathrm{i} \boldsymbol{\kappa} \cdot \boldsymbol{R}_n), \tag{A.44}$$

with

$$F_{sn}(\boldsymbol{\kappa}) = \sum_m f_{nm} \exp(-2\pi \mathrm{i} \boldsymbol{\kappa} \cdot \boldsymbol{r}_{nm}) \exp\left[-2\pi \mathrm{i} \boldsymbol{\kappa} \cdot (\boldsymbol{R}_{nm} - \boldsymbol{R}_n)\right] \tag{A.45}$$

for the structure amplitude of the nth cell.

Equation (A.45) shows that *positional disorder* alone leads to different structure amplitudes for different cells. Furthermore, eq. (A.45) also shows the general effect of *compositional* disorder (chemical disorder), as f_{nm} reflects exactly the occupation of the various sites by different atoms. In the special case, where compositional disorder is not considered and where the displacement vector $\boldsymbol{R}_{nm}$ can be regarded as constant for all atoms of the elementary cell, the structure amplitude $F_s(\boldsymbol{\kappa})$ remains unchanged, and using eqs. (A.24) and (A.25) the lattice amplitude F_L can be written in the integral form:

$$F_L(\boldsymbol{\kappa}) = \frac{1}{V_e} \int_{V_s} \exp\left\{-2\pi \mathrm{i}\left[\boldsymbol{g} \cdot \boldsymbol{R}(\boldsymbol{r}) + \boldsymbol{s} \cdot \boldsymbol{r} + \boldsymbol{s} \cdot \boldsymbol{R}(\boldsymbol{r})\right]\right\} \mathrm{d}^3 r, \tag{A.46}$$

where the integration runs over the sample volume V_s. In many cases the term $\boldsymbol{s} \bullet \boldsymbol{R}$ can be neglected. These general expressions are the basis for more detailed evaluations in this chapter and ch. 12, where several special cases are discussed.

References

AMELINCKX, S., 1964, The Direct Observations of Dislocations (Academic, New York).

AMELINCKX, S., and J. VAN LANDUYT, 1978, in: Diffraction and Imaging Techniques in Material Science, eds. S. Amelinckx, R. Gevers and J. Van Landuyt (North-Holland, Amsterdam) vol. I, p. 107.

AMELINCKX, S., R. GEVERS and J. VAN LANDUYT, eds., 1978, Diffraction and Imaging Techniques in Material Science, vols. I and II (North-Holland, Amsterdam).

ANDREWS, K.W., D.J. DYSON and S.R. KEOWN, 1971, Interpretation of Electron Diffraction Patterns (Adam Hilger, London).

ASHBY, M.F., and L.M. BROWN, 1963, Phil. Mag. **8**, 1093 and 1649.

BACMAN, J.-J., G. SILVESTRE and M. PETIT, 1981, Phil. Mag. **A43**, 189.

BALL, C.J., 1964, Phil. Mag. **9**, 541.

BALLUFFI, R.W., ed., 1980, Grain-Boundary Structure and Kinetics (ASM, Metals Park, OH).

BALLUFFI, R.W., A. BROKMAN and A.H. KING, 1982, Acta Metall. **30**, 1453.

BERTIN, E.P., 1975, Principles and Practice of X-Ray Spectrometric Analysis (Plenum, New York).

BETHE, H.A., 1928, Ann. Phys. (Leipzig) **87**, 55.

BETHE, H. A., 1933, Handbuch der Physik, vol. 24 (Springer, Berlin) p. 273.
BISCHOFF, E., 1983, unpublished.
BONNET, R., 1981, Phil. Mag. **A43**, 1165.
BONNET, R., E. COUSINEAU and D. H. WARRINGTON, 1981, Acta Cryst. **A37**, 184.
BOURRET, A., and J. DESSEAUX, 1979, Phil. Mag. **A39**, 405.
BOURRET, A., J. DESSEAUX and A. RENAULT, 1982, Phil. Mag. **A45**, 1.
BROKMAN, A., and R. W. BALLUFFI, 1981, Acta Metall. **29**, 1703.
CHADWICK, G. A., and D. A. SMITH, eds., 1976, Grain-Boundary Structure and Properties (Academic, London).
CHAMPNESS, P. E., G. CLIFF and G. W. LORIMER, 1982, Ultramicroscopy **8**, 121.
CHIK, K. P., 1966, Phys. Stat. Sol. **16**, 685.
CHIK, K. P., M. WILKENS and M. RÜHLE, 1967, Phys. Stat. Sol. **23**, 113.
CLAREBROUGH, L. M., and C. T. FORWOOD, 1976, in: Proc. 9th Int. Congress on Electron Microscopy, ed. J.M. Sturgess (Microscopical Society of Canada, Toronto) vol. II, p. 38.
CLIFF, G., and G. W. LORIMER, 1975, J. Microsc. **103**, 203.
COCKAYNE, D. J. H., 1978, in: Diffraction and Imaging Techniques in Material Science, eds. S. Amelinckx, R. Gevers and J. Van Landuyt (North-Holland, Amsterdam) vol. I, p. 153.
COCKAYNE, D. J. H., and R. GRONSKY, 1981, Phil. Mag. **A44**, 159.
COCKAYNE, D. J. H., I. L. F. RAY and M. J. WHELAN, 1969, Phil. Mag. **20**, 1265.
DAHMEN, U., K. H. WESTMACOTT and G. THOMAS, 1981, Acta Metall. **29**, 627.
DAS, S. K., P. R. OKAMOTO, P. M. J. FISHER and G. THOMAS, 1973, Acta Metall. **21**, 913.
EGERTON, R. F., 1979, Ultramicroscopy **4**, 169.
EGERTON, R. F., 1981, in: Analytical Electron Microscopy 1981, ed. R.H. Geiss (San Francisco Press, San Francisco) vol. I, p. 154.
ENGLISH, C., B. L. EYRE and S. M. HOLMES, 1980, J. Phys. **F10**, 1065.
ESSMANN, U., 1966, Phys. Stat. Sol. **17**, 725.
EYRE, B. L., and R. BULLOUGH, 1965, Phil. Mag. **12**, 31.
EYRE, B. L., D. M. MAHER and R. C. PERRIN, 1977, J. Phys. **F7**, 1359 and 1371.
FLORJANCIC, M., M. RÜHLE and S. L. SASS, 1982, in: Proc. 10th Int. Congress on Electron Microscopy, (Deutsche Gesellschaft für Elektronenmikroskopie, Frankfurt) vol. 2, p. 359.
FÖLL, H., and M. WILKENS, 1975, Phys. Stat. Sol. (a) **31**, 519.
FÖLL, H., and M. WILKENS, 1977, Phys. Stat. Sol. (a) **39**, 561.
GEISS, R. H., 1979, in: Introduction to Analytical Electron Microscopy, eds. J.J. Hren, J.I. Goldstein and D.C. Joy (Plenum, New York) p. 43.
GEISS, R. H., ed., 1981, Analytical Electron Microscopy 1981, vols. I and II (San Francisco Press, San Francisco).
GEVERS, R., 1962, Phil. Mag. **7**, 59 and 651.
GEVERS, R., A. ART and S. AMELINCKX, 1963, Phys. Stat. Sol. **3**, 1563.
GEVERS, R., J. VAN LANDUYT and S. AMELINCKX, 1965, Phys. Stat. Sol. **11**, 689.
GOLDSTEIN, J. I., and D. B. WILLIAMS, 1981, in: Quantitative Microanalysis with High Spatial Resolution, eds G.W. Lonmer, M.H. Jacobs and P. Doig (The Metals Society, London) p. 5.
GOLDSTEIN, J. I., J. L. COSTLEY, G. W. LORIMER, and S. J. B. REED, 1977, in: Scanning Electron Microscopy 1977, ed. O. Johari (IIT Research Inst., Chicago, IL) vol. I, p. 315.
GOLDSTEIN, J. I., D. E. NEWBURY, P. ECHLIN, D. C. JOY, C. FIORI and E. LIFSHIN, 1981, Scanning Electron Microscopy and X-Ray Microanalysis (Plenum, New York).
GRATIAS, D., R. PORTIER, M. FAYARD and M. GUYMOT, 1979, Acta Cryst. **A35**, 885.
HÄUSSERMANN, F., M. RÜHLE and M. WILKENS, 1972, Phys. Stat. Sol. (b) **50**, 445.
HÄUSSERMANN, F., K. H. KATERBAU, M. RÜHLE and M. WILKENS, 1973, J. Microsc. **98**, 135.
HASHIMOTO, H., A. HOWIE and M. J. WHELAN, 1960, Phil. Mag. **5**, 967.
HASHIMOTO, H., A. HOWIE and M. J. WHELAN, 1962, Proc. Roy. Soc. **A269**, 80.
HAWKES, P. W., ed., 1980, Computer Processing of Electron Microscope Images, Topics in Current Physics, vol. 13 (Springer, Berlin).
HAZZLEDINE, P. M., H. P. KARNTHALER and E. WINTNER, 1975, Phil. Mag. **32**, 81.
HEAD, A. K., P. HUMBLE, L. M. CLAREBROUGH, A. J. MORTON and C. T. FORWOOD, 1973, Computer Electron Micrographs and Defect Identification (North-Holland, Amsterdam).

HIRSCH, P. B., A. HOWIE and M. J. WHELAN, 1960, Phil. Trans. Roy. Soc. **A252**, 499.
HIRSCH, P. B., A. HOWIE, R. B. NICHOLSON, D. W. PASHLEY and M. J. WHELAN, 1977, Electron Microscopy of Thin Films (Kneger, Huntington, NY).
HOWIE, A., and Z. S. BASINSKI, 1968, Phil. Mag. **17**, 1039.
HOWIE, A., and M. J. WHELAN, 1961, Proc. Roy. Soc. **A263**, 217.
HOWIE, A., and M. J. WHELAN, 1962, Proc. Roy. Soc. **A267**, 206.
HREN, J. J., 1979, in: Introduction to Analytical Electron Microscopy, eds. J.J. Hren, J.I. Golstein and D.C. Joy (Plenum, New York) p. 481.
HREN, J. J., J. I. GOLDSTEIN and D. C. JOY, eds., 1979, Introduction to Analytical Electron Microscopy (Plenum, New York).
HUMBLE, P., 1978, in: Diffraction and Imaging Techniques in Material Science, eds. S. Amelinckx, R Gevers and J. Van Landuyt (North-Holland, Amsterdam) vol. I, p. 315.
HUMBLE, P., and C. T. FORWOOD, 1975, Phil. Mag. **31**, 1011 and 1025.
JÄGER, W., and M. WILKENS, 1975, Phys. Stat. Sol. **32**, 89.
JENKINS, M. L., and M. WILKENS, 1976, Phil. Mag. **34**, 1155.
JOUFFREY, B., Y. KIHN, J. P. PEREZ, J. SEVELY and G. ZANCHI, 1978, in: Proc. 9th Int. Congress on Electron Microscopy, ed. J. M. Sturgess (Microscopical Society of Canada, Toronto) vol. III, p. 292.
JOY, D. C., 1981, in: Quantitative Microanalysis with High Spatial Resolution, eds. G.W. Lorimer, M.H. Jacobs and P. Doig (The Metals Society, London) p. 127.
JOY, D., and D. M. MAHER, 1980, Ultramicroscopy **5**, 333.
KALONJI, G., and J. W. CAHN, 1982, J. Physique **C6**, 25.
KATERBAU, K. H., 1976, Phys. Stat. Sol. (a) **38**, 463.
KATERBAU, K. H., 1980, Phys. Stat. Sol. (a) **59**, 211.
KATERBAU, K. H., 1981, Phil. Mag. **A43**, 409.
KELLY, P. M., A. JOSTSONS, R. G. BLAKE and J. G. NAPIER, 1975, Phys. Stat. Sol. (a) **31**, 77.
KHACHATURYAN, A. G., 1974, in: Order–Disorder Transformations in Alloys, ed. H. Warlimont (Springer, Berlin) p. 114.
KORNER, A., and H. P. KARNTHALER, 1980, Phil. Mag. **A42**, 753.
KORNER, A., and H. P. KARNTHALER, 1981, Phys. Stat. Sol. (a) **68**, 19.
KRAHL, D., K.-H. HERMANN and W. KUNATH, 1978, in: Proc. 9th Int. Congress on Electron Microscopy, ed. J.M. Sturgess (Microscopical Society of Canada, Toronto) vol. I, p. 42.
KRIVANEK, O. L., and P. R. SWANN, 1981, in: Quantitative Microanalysis with High Spatial Resolution (The Metals Society, London) p. 136.
KRIVANEK, O. L., S. ISODA and K. KOSAYASHI, 1977, Phil. Mag. **36**, 331.
LAFFONT, A., and R. BONNET, 1982, Acta Metall. **30**, 763.
LEAPMAN, R. D., S. J. SANDERSON and M. J. WHELAN, 1978, Met. Sci. **23**, 215.
LEPSKI, D., 1974, Phys. Stat. Sol. (a) **24**, 99.
LEPSKI, D., and P. BURCK, 1981, Phys. Stat. Sol. (a) **64**, 625.
LEPSKI, D., and P. BURCK, 1982, Phys. Stat. Sol. (a) **70**, 571.
LORIMER, G. W., M. H. JACOBS and P. DOIG, eds., 1981, Quantitative Microanalysis with High Spatial Resolution (The Metals Society, London).
MAHER, D. M., and B. L. EYRE, 1971, Phil. Mag. **23**, 409.
MATHEWS, J. W., and W. M. STOBBS, 1977, Phil. Mag. **36**, 373.
MCINTYRE, K. G., and L.M. BROWN, 1966, J. Physique **27**, C3-178.
NIEDRIG, H., 1978, Scanning **1**, 17.
OLSON, G. B., and M. COHEN, 1979, Acta Metall. **27**, 1907.
PENISSON, J. M., R. GRONSKY and J. B. BROSSE, 1982, Scripta Metall. **16**, 1239.
POGANY, A. P., and P. S. TURNER, 1968, Acta Cryst. **A24**, 103.
POND, R. C., 1979, Phil. Mag. **37**, 699.
POWELL, C. J., 1976, in: National Bureau of Standards, Spec. Publ. 460, eds. K.K.J. Heinrich, D.E. Newbury and H. Yakowitz (National Bureau of Standards, Washington) p. 97.
RADI, G., 1970, Acta Cryst. **A26**, 41.
ROSSOUW, C. J., R. F. EGERTON and M. J. WHELAN, 1976, Vacuum **26**, 427.

RÜHLE, M., 1967, Phys. Stat. Sol. **19**, 263 and 279.
RÜHLE, M., and G. PETZOW, 1981, in: Surface and Interfaces in Ceramic and Ceramic–Metal Systems, eds. J. Pask and A. Evans (Plenum, New York) p. 167.
RÜHLE, M., and M. WILKENS, 1967, Phil. Mag. **15**, 1075.
RÜHLE, M., and M. WILKENS, 1975, Cryst. Lattice Defects **6**, 129.
SALDIN, D.K., A.Y. STATHOPOULOS and M.J. WHELAN, 1978, Inst. Phys. Conf. Ser. No. **41**, 350.
SASS, S.L., and P.D. BRISTOWE, 1980, in: Grain-Boundary Structure and Kinetics, ed. R.W. Balluffi (ASM, Metals Park, OH) p. 71.
SCHOBER, T., and R.W. BALLUFFI, 1969, Phil. Mag. **20**, 511.
SHANNON, J., and L. KATZ, 1970, Acta Cryst. **B26**, 102.
SIEMS, R., P. DEI AVIGNETTE and S. AMI LINCKX, 1962, Phys. Stat. Sol. **2**, 421.
SINCLAIR, R., R. GRONSKY and G. THOMAS, 1976, Acta Metall. **24**, 789.
SMITH, D.A., 1976, in: Grain-Boundary Structure and Properties, eds. G.A. Chadwick and D.A. Smith (Academic, London) p. 321.
SPENCE, J.C.H., 1981, Experimental High-Resolution Electron Microscopy (Clarendon Press, Oxford).
SPENCE, J.C.H., and H. KOLAR, 1979, Phil. Mag. **A39**, 59.
STEEDS, J.W., 1966, Proc. Roy. Soc. **A292**, 343.
STEEDS, J.W., 1979, in: Introduction to Analytical Electron Microscopy, eds. J.J. Hren, J.I. Goldstein and D.C. Joy (Plenum, New York) p. 481.
STEEDS, J.W., 1981, in: Quantitative Microanalysis with High Spatial Resolution, eds. G.W. Lorimer, M.H. Jacobs and P. Doig (The Metals Society, London) p. 210.
STOBBS, W.M., and C.H. SWORNE, 1971, Phil. Mag. **24**, 1365.
TAKAGI, S., 1962, Acta Cryst. **15**, 1310.
THOMAS, G., and M.J. GORINGE, 1979, Transmission Electron Microscopy of Materials (Wiley, New York).
THOMPSON–RUSSELL, K.C., and J.W. EDINGTON, 1977, in: Practical Electron Microscopy in Matenals Science, vol. 5 (Mac Millan, London).
TIGHE, N.J., 1976, in: Electron Microscopy in Mineralogy, ed. H.R. Wenk (Springer, Berlin).
TUNSTALL, W.J., P.B. HIRSCH, and J.W. STEEDS, 1964, Phil. Mag. **9**, 99.
VAN LANDUYT, J., R. GEVERS and S. AMELINCKX, 1965, Phys. Stat. Sol. **10**, 319.
VAN TENDELOO, G., and S. AMELINCKX, 1977, Phys. Stat. Sol. (a) **43**, 553.
VAN TENDELOO, G., and S. AMELINCKX, 1978, Phys. Stat. Sol. (a) **50**, 53.
WENK, H.R., 1976, Electroii Microscopy in Mineralogy (Springer, Berlin).
WILKENS, M., 1964, Phys. Stat. Sol. **6**, 939.
WILKENS, M., 1978, in: Diffraction and Imaging Techniques in Material Science, eds. S. Amelinckx, R. Gevers and J. Van Landuyt (North-Holland, Amsterdam) vol. I, p. 185.
WILKENS, M., 1981, Phys. Stat. Sol. (a) **69**, 123.
WILKENS, M., and E. HORNBOGEN, 1964, Phys. Stat. Sol. **4**, 557.
WILKENS, M., and H.O.K. KIRCHNER, 1981, Phil. Mag. **A43**, 139.
WILKENS, M., and M. RÜHLE, 1982, Phys. Stat. Sol. (b) **49**, 749.
WILKENS, M., M. RÜHLE and F. HÄUSSERMANN, 1967, Phys. Stat. Sol. **22**, 689.
WILKENS, M., K.H. KATERBAU and M. RÜHLE, 1973, Z. Naturf. **28a**, 681.
WILKENS, M., M.L. JENKINS and K.H. KATERBAU, 1977, Phys. Stat. Sol. (a) **39**, 103.
ZANCHI, G., J. SEVELY and B. JOUFFREY, 1977, in: Proc. 5th Int. Conf. on High Voltage Electron Microscopy, Kyoto, eds. T. Imura and H. Hashimoto (Japanese Society for Electron Microscopy, Tokyo) p. 117.

General bibliography for transmission electron microscopy

Textbooks

Amelinckx, S., R. Gevers and J. Van Landuyt, eds., 1978, Diffraction and Imaging Techniques in Material Science, vols. I and II (North-Holland, Amsterdam).
Bethge, H., and J. Heydenreich, eds., 1982, Elektronenmikroskopie in der Festkörperphysik (Springer, Berlin).

Cowley, J. M., 1975, Diffraction Physics (North-Holland, Amsterdam).
Head, A. K., P. Humble, L. M. Clarebrough, A. J. Morton and C. T. Forwood, 1973, Computer Electron Micrographs and Defect Identification (North-Holland, Amsterdam).
Hirsch, P. B., A. Howie, R. B. Nicholson, D. W. Pashley and M. J. Whelan, 1977, Electron Microscopy of Thin Crystals (Krieger, Huntington, NY).
Hornbogen, E., 1971, Durchstrahlungs-Elektronenmikroskopie fester Stoffe (Verlag Chemie, Weinheim).
Reimer, L., 1967, Elektronenmikroskopische Untersuchungs- und Präparationsmethoden (Springer, Berlin).
Thomas, G., and M. J. Goringe, 1979, Transmission Electron Microscopy of Materials (Wiley, New York).

Evaluation of diffraction patterns

Andrews, K. W., D. J. Dyson and S. R. Keown, 1971, Interpretation of Electron Diffraction Patterns (Adam Hilger, London).

Observation of dislocations

Amelinckx, S., 1964, The Direct Observation of Dislocations (Academic, Oxford).
Amelinckx, S., 1979, Dislocations in Particular Structures, in: Dislocations in Solids, ed. F. R. N. Nabarro (North-Holland, Amsterdam) vol. II, ch. 6.

TEM of grain boundaries

Balluffi, R. W., 1980, Grain-Boundary Structure and Kinetics (ASM, Metals Park, OH).
Chadwick, G. A., and D. A. Smith, eds., 1976, Grain-Boundary Structure and Properties (Academic, London).

Analytical electron microscopy

Geiss, R. H., ed., 1981, Microbeam Analysis 1981, vols. I and II (San Francisco Press, San Francisco).
Goldstein, J. I., D. E. Newbury, P. Echlin, D. C. Joy, C. Fiori and E. Lifshin, 1981, Scanning Electron Microscopy and X-Ray Microanalysis (Plenum, New York).
Hren, J. J., J. I. Goldstein and D. C. Joy, eds., 1979, Introduction to Analytical Electron Microscopy (Plenum, New York).
Lorimer, G. W., M. H. Jacobs and P. Doig, eds., 1981, Quantitative Microanalysis with High Spatial Resolution (The Metals Society, London).

High-resolution electron microscopy

Hawkes, P. W., ed., 1980, Computer Processing of Electron Microscope Images, Topics in Current Physics, vol. 13 (Springer, Berlin).
Spence, J. C. H., 1981, Experimental High-Resolution Electron Microscopy (Clarendon Press, Oxford).

Microscopy of Minerals

Wenk, H.-R., 1976, Electron Microscopy in Mineralogy (Springer, Berlin).

Addendum

Since the release of the third edition of Physical Metallurgy (R. W. Cahn and P. Haasen, eds.) enormous progress has been made regarding instrumention as well as further development in transmission electron microscopy of materials. The description of the basic physics of the scattering processes of electrons within a specimen still holds as described in the third edition. The main applications of transmission electron microscopy in materials science are still concentrated mainly on conventional transmission electron microscopy; however, high-resolution transmission electron mciroscopy and analytical electron microscopy are becoming powerful tools, too. This addendum will focus on

progress made in different areas of transmission electron microscopy.

Transmission electron microscopy (TEM) is becoming more specialized and can be divided into conventional transmission electron microscopy (CTEM), analytical electron microscopy (AEM) and high-resolution transmission electron microscopy (HRTEM). In CTEM the morphology of materials as well as the distribution of different phases can be determined as described in sect. 4. Information on structure can be obtained with great success. In addition, the analysis of defects such as interfaces (phase boundaries, etc.), dislocations is possible (sect. 5 to 8). Recently, major progress has been made in determining the dislocation structure in intermetallics with emphasis on detailed analysis of the splitting of superdislocations into partial dislocations using the weak beam technique (YU and KARNTHALER [1993], KARNTHALER *et al.* [1993]).

In addition, *in situ* experiments in the microscope allow heating, cooling and deformation of the specimen. The latter technique revealed dynamic processes (LOUCHET and SAKA [1993] and FUJITA *et al.* [1991]). Interesting results could be obtained for an understanding of deformation mechanisms in metals and ceramics, radiation damage, crystal growth and corrosion.

In AEM the main emphasis of present studies is on extracting information on chemical composition with high spatial resolution (EGERTON [1986], JOY *et al.* [1986] and REIMER [1993]). The technique is applied especially for the determination of the composition at or close to the interface and gradients in materials (KRIVANEK [1991] and DISKO *et al.* [1992]). Recently, special techniques of electron energy loss spectroscopy (EELS) have come to allow the determination of bonding and electronic structure of materials.

HRTEM allows the determination of the structure of materials, especially the analysis of specific defects in materials (SPENCE [1988], BUSECK *et al.* [1988], COWLEY *et al.* [1992] and HORIUCHI [1994]). These defects are dislocations in specific orientations in the thin film and interfaces. With great success the structure of complex high-temperature superconductors could be revealed by TEM (HORIUCHI [1994], MITCHELL and DAVIES [1988], MATSUI and KRAKOW [1995]).

This progress was possible because of major advances in instrumentation as described in the first section below.

B.1. Instrumentation

Over the last ten years enormous progress has been made towards better instrumentation and auxiliary parts. In conventional TEM it is now quite common that near ultra-high vacuum conditions can be reached close to the specimen in a microscope. The conditions of the specimen environment are extremely well-defined and the resolution power of conventional TEMs (operating at 200 kV to 400 kV) approaches 0.2 nm or better depending on the high-voltage of the instrument. This enormous progress was made through middle-voltage instruments where the high voltage is in the range of 300 kV to 400 kV resulting in a resolution of better than 0.2 nm.

The specimen stages of the different instruments are mechanically extremely stable. New developments in fine mechanics allow a stability of the specimen which often results in a drift rate smaller than 0.001 nm/s. Large progress has therefore been made in the analysis of

defects as described in sect. 5 since extremely long exposure times are now possible.

The acquisition of data is now also possible with slow-scan CCD (charge-coupled device) cameras (REIMER [1993]) which possess a large linear detection range.

Field emission guns are available that operate at high voltages between 100 kV and 300 kV. These guns possess an extreme brightness and coherence (REIMER [1993] and Spence [1988]).

In AEM, the high-voltage stability and sensitive X-ray detectors (NEWBURY *et al.* [1986]) together with very sensitive energy parallel detection systems for energy loss spectroscopy, result in shorter exposure times (by an order of magnitude compared to ten years ago). Of course, the same physical principles apply; however, the detection limit in AEM was drastically reduced by more sensitive EDS detectors and by parallel EELS detection systems.

Since 1990, energy-filtering electron microscopes or imaging filters are available which allow imaging with elastically scattered electrons only (REIMER [1995]). All inelastically scattered electrons are filtered out of the image, resulting in micrographs of much higher contrast. The filtering can be done either by a system of magnetic sector field segments built into the microscope column or by an imaging filter mounted at the bottom of the instrument (REIMER [1995]).

In HRTEM the resolution power of the instruments has been pushed towards 0.1 nm. This resolution is reached by using high-voltage high-resolution electron microscopy instruments operating at 1 MeV or beyond. A resolution of 0.1 nm can be reached (RÜHLE *et al.* [1994]).

B.2. Conventional transmission electron microscopy including weak beam

In conventional TEM the resolution of weak beam techniques has reached its theoretical limits. The limitation is caused by the width of the contrast lines which in turn is correlated to the extinction length of the operating reflection. The observation of splitting of dislocations down to the 1 nm level could be achieved and valuable information on the splitting, especially, of superdislocations in intermetallics could be obtained (LIU *et al.* [1992]).

B.3. Analytical electron microsopy

In AEM the energy-filtered convergent beam electron diffraction (CBED) (SPENCE and ZUO [1992], REIMER [1995]) has become an important tool for the analysis of crystal structures. A convergent beam with a converging angle in the order of 10^{-2} is falling on a specimen and scattering occurs in different directions. From the structure within the diffraction disc of a CBED information on the structure of the material can be obtained with high spatial resolution. Scattering factors for different reflections can be calculated by comparing the calculated intensity distributions within a disc of the CBED with corresponding experimental energy-filtered micrographs. Using this method, the scattering factors of reflections in silicon could be reproduced with high accuracy. Recently, the scattering factors of intermetallic alloys could also be determined and the Debye–Waller factor could be determined for the different elements existing in the ordered alloy.

The electronic structure of the material with high spatial resolution can also be determined by spatially resolved valence electron energy loss spectroscopy (SREELS) (MÜLLEJANS *et al.* [1994], BRULEY [1993]). Recently, it was shown that the spatial variation of the electronic structure of interfaces can be determined from valence spectroscopy. This value is critical to both interatomic bonding at atomically abrupt interfaces such as grain boundaries and also for the development of van der Waals attraction forces at wetted interfaces (HOFFMANN and PETZOW [1994]). The physics for the quantitative analysis of SREELS requires a series of corrections of transformation of data including single scattering deconvolution, convergence correction and Kronig–Kramer analysis (MÜLLEJANS *et al.* [1993] and FRENCH *et al.* [1995]). With this technique the bonding across clean interfaces and wetted interfaces could be determined. The near-edge structure of EELS gives some information on bonding across the interfaces and the atomic environment at defects. The experimental technique is conducted as follows. First a spectrum is taken at the bulk material (next to the interface) and then another spectrum at the interface itself. The difference of both spectras results in a interface specific contribution. For this analysis a linear addition of the different components is assumend. This technique was applied for grain boundaries in alumina (MÜLLEJANS *et al.* [1993]) as well as metal/ceramic interfaces (BRULEY *et al.* [1994]). The type of bonding and the environment of the bonding could be determined with high precision.

B.4. High-resolution transmission electron microscopy

Similar advances have also been made in the analysis of micrographs obtained by HRTEM. Methods and programs have been developed that allow the simulation of HRTEM images of any given atom arrangement. These simulated images are used to interpret experimentally obtained micrographs. The recent developments enable us to use HRTEM as a method for solving problems in materials science. Atomic structures of different lattice defects, such as phase boundaries, grain boundaries and dislocations, can be determined by HRTEM (FINNIS and RÜHLE [1993], SUTTON and BALLUFFI [1995]). The difficulties in the interpretation of HRTEM are similar to those in conventional transmission electron microscopy of lattice defects in crystalline materials: the experimentally obtained micrographs do not usually present a direct image of the object.

The extraction of information on the structure of lattice defects from HRTEM micrographs is complicated (SPENCE [1988] and BUSECK et al. [1988]). Despite this, very useful information has been obtained on the structure of lattice defects using HRTEM, particularly for semiconductors. The new generation of instruments also allows the characterization of defects in ceramics and metals by HRTEM.

Addendum References

BRULEY, J., 1993, Microsc. Microanal. Microstruct. **4**, 23.

BRULEY, J., R. BRYDSON, H. MÜLLEJANS, J. MAYER, G. GUTEKUNST, W. MADER, D. KNAUSS and M. RÜHLE, 1994, J. Mater. Res. **9**, 2574.

BUSECK, P., J. COWLEY and L. EYRING, eds., 1988, High-Resolution Transmission Electron Microscopy (Oxford University Press, Oxford).

COWLEY, J.M., ed., 1992, Electron Diffraction Techniques, vol. 1 (Oxford University Press, Oxford).

COWLEY, J.M., ed., 1992, Electron Diffraction Techniques, vol. 2 (Oxford University Press, Oxford).

DISKO, M.M., C.C. AHN and B. FULTZ, eds., 1992, Transmission Electron Energy Loss Spectrometry in Materials Science (TMS, Warrendale, PA).

EGERTON, R.F., 1986, Electron Energy-Loss Spectroscopy, (Plenum Press, New York).

FINNIS, M.W., and M. RÜHLE, 1993, in: Materials Science and Technology, eds. R.W. Cahn, P. Haasen and E. Krämer, vol. 1, ed. V. Gerold (VCH, Weinheim), p. 533.

FRENCH, R.H., C. SCHEU, G. DUSCHER, H. MÜLLEJANS, M.J. HOFFMANN and R.M. CANNON, 1995, in: Interfaces in Ceramics, eds. D. Bonnell, U. Chowdhry and M. Rühle, vol. 357 (MRS, Pittsburgh, PA).

FUJITA, H., K. URA and H. MORI, eds., 1991, Proc. 2nd Int. Symp. on High-Voltage Electron Microscopy, Osaka, Ultramicroscopy **39**, 1–402.

HOFFMANN, M.J., and G. PETZOW, eds., 1994, Tailoring of Mechanical Properties of Si_3N_4 Ceramics (Kluwer Academic Publ., Dordrecht).

HORIUCHI, S., 1994, Fundamentals of High-Resolution Transmission Electron Microscopy (North-Holland, Amsterdam).

JOY, D.C., A.D. ROMIG Jr. and J.I. GOLDSTEIN, eds., 1986, Principles of Analytical Electron Microscopy (Plenum Press, New York).

KARNTHALER, H.P., C. RENTENBERGER and E. MÜHLBACHER, 1993, in: High-Temperature Ordered Intermetallic Alloys V, eds. J.D. Whittenberger, R. Darolia and M.H. Yoo, vol. 288 (MRS, Pittsburgh, PA) 293.

KRIVANEK, O.J., ed., 1991, Proc. Electron Energy Loss Workshop, Tahoe City, CA, Microsc. Microanal. Microstruct. **2**, 143–411.

LIFSHIN, E. (ed.) Characterization of Materials, Part I, Vol. 2A of Materials Science and Technology, ed. R.W. Cahn, P. Haasen and E.J. Kraamer (VCH, Weinheim). Includes chapters by S. Amelinckx ("Electron Diffraction and Transmission Electron Microscopy", p. 1) and by E.L. Hall ("Analytical Electron Microscopy", p. 147).

LIU, C.T., R.W. CAHN, S. SAUTHOFF, eds., 1992, Ordered Intermetallics — Physical Metallurgy and Mechanical Behaviour (Kluwer Publ., Dordrecht).

LOUCHET, F., and H. SAKA, eds., 1993, Proc. French-Japanese Seminar on In Situ Electron Microscopy, Nagoya, vol. 4, Microsc. Microanal. Microstruct. **4**, 101–346.

MATSUI, Y., and W. KRAKOW, eds., 1995, Microscopy of High Temperature Superconductors, Part I, Microscopy Research and Technique **30**, 101–192.

MITCHELL, T.E., and P.K. DAVIES, eds., 1988, Electron Microscopy of Ceramic Superconductors, J. of Electron Microscopy Technique **8**, 245–341.

MÜLLEJANS, H., J. BRULEY, R.H. FRENCH and P.A. MORRIS, 1993, in: Proc. Electron Microscopy and Analysis Conference, Liverpool, vol. 138 (IOP Publ., Bristol).

MÜLLEJANS, H., J. BRULEY, R.H. FRENCH and P.A. MORRIS, 1994, in: Determining Nanoscale Physical Properties of Materials by Microscopy and Spectroscopy, eds. M. Sarikaya, H.K. Wickramasinghe, M. Isaacson, vol. 332 (MRS, Pittsburgh, PA).

NEWBURY, D.E., D.C. JOY, P. ECHLIN, C.E. FIORI and J.I. GOLDSTEIN, eds., 1986, Advanced Scanning Electron Microscopy and X-Ray Microanalysis (Plenum Press, New York).

REIMER, L., 1993, Transmission Electron Microscopy (Springer Verlag, Berlin).

REIMER, L., ed. 1995, Energy-Filtering Transmission Electron Microscopy (Springer Verlag, Berlin).

RÜHLE, M., F. PHILLIPP, A. SEEGER and J. HEYDENREICH, eds., 1994, Proc. Int. Workshop on High-Voltage and High-Resolution Electron Microscopy, Stuttgart, Ultramicroscopy **56**, 1–232.

SPENCE, J.C.H., 1988, Experimental High-Resolution Electron Microscopy (Oxford University Press, Oxford).

SPENCE, J.C.H., and J.M. ZUO, 1992, Electron Microdiffraction (Plenum Press, New York).

SUTTON, A.P., and R.W. BALLUFFI, 1995, Interfaces in Crystalline Solids (Oxford University Press, Oxford).

YU, H., and H.P. KARNTHALER, 1993, Materials and Design **14**, 69.

CHAPTER 12

X-RAY AND NEUTRON SCATTERING

G. KOSTORZ

Institut für Angewandte Physik
Eidgenössische Technische Hochschule
Zürich, Switzerland

R. W. Cahn and P. Haasen†, eds.
Physical Metallurgy; fourth, revised and enhanced edition

1. Introduction

There are many textbooks and monographs on X-ray and neutron scattering (see the selection at the end of this chapter), and the number of publications is immense. This short chapter cannot give a complete account of the history or the current state of the field. The basic theory, standard methods and many of the more classical applications are well described in various textbooks (a brief compendium of fundamentals of scattering and crystallography is given in the Appendix to ch. 11). The main purpose of the present chapter will be to demonstrate that there is much more we can "learn from scattering experiments besides the average structure" (SCHWARTZ and COHEN [1987], ch. 7).

One immediate question concerns the range of stability of a given structure, and scattering experiments are helpful in locating phase transformations. There are classical methods (e.g., powder diffraction) and very sophisticated recent techniques (e.g., quasi-elastic neutron scattering) to study transformations and phase separation in metallic systems. Some examples will be discussed. Much space, however, will be devoted to studies of inhomogeneities, i.e., deviations from the average structure (point defects, clusters, short-range order, precipitates, etc.).

In § 2, some remarks on scattering from crystalline material that contains defects are followed by a discussion of X-rays and neutrons as the two types of radiation now commonly used for scattering studies of essentially bulk materials. Scattering of low-energy electrons, important in the investigation of surfaces and surface phenomena, will not be covered. High-energy electron diffraction is discussed in ch. 11. Throughout this chapter, scattering intensities are expressed in terms of scattering cross-sections (differential cross-sections without energy analysis, double-differential cross-sections for scattering experiments with energy analysis), mostly applicable to both X-rays and neutrons. The reader should not underestimate the difficulties in converting measured intensities to absolute cross-sections. Calibration, absorption, background, polarization and other corrections may introduce considerable errors. These problems and questions of instrument design and optimization can only be mentioned occasionally.

§§ 3–6 cover recent applications, and the subject matter has been divided according to simple criteria. Diffraction and elastic scattering at and near Bragg peaks, between Bragg peaks, and near the incident beam are treated in §§ 3–5, and § 6 gives a few examples for inelastic and quasi-elastic scattering. Theoretical expressions are kept to a minimum but some are required to interpret the main effects. There are several systematically elegant and quite general schemes to describe the same scattering phenomena (starting, e.g., from lattice sums, correlation functions or convolutions). No formal elaboration of this type will be attempted. Results of the kinematical theory and of first-order Born approximation adapted to simple but manageable cases will mostly be relied upon. Dynamic diffraction theory as the basis of transmission electron microscopy is discussed in ch. 11, and related X-ray and neutron applications can be found there.

2. *Scattering from real crystals*

2.1. General predictions of the kinematical theory

As can be seen in ch. 11, eq. (A.29), the kinematical scattering theory for an otherwise perfect, but finite crystal yields Bragg peaks whose width $\Delta\kappa_g$ equals $1/L_\kappa$ where L_κ is the linear dimension of the crystal parallel to κ. As $\Delta\kappa_g$ is independent of the magnitude of g, this "finite size effect" broadening can be recognized by comparing the width of different Bragg peaks (at Bragg angles θ_{hkl}). For powders, we thus have a method of determining crystallite sizes, since a constant $\Delta\kappa$ means:

$$\Delta(2\theta_{hkl}) \approx \frac{\lambda}{L\cos\theta_{hkl}}, \tag{1}$$

with L as the average crystallite diameter. However, experimental conditions will severely limit the range over which peak widths are controlled by the crystallite size according to eq. (1). If L is too large ($\geq 1\ \mu$m), we may reach the resolution limit of the experiment, and if L is too small (≤ 100 Å = 10 nm), it becomes difficult to separate the diffraction lines (see, e.g., BARRETT and MASSALSKI [1980] for a detailed discussion). For large bulk samples, the size broadening may be ignored, and the differential cross-section for Bragg scattering may be written as [cf. eq. (A.36), ch. 11]:

$$\left(\frac{d\sigma}{d\Omega}\right)_B = N|F_s|^2 \frac{\delta(\kappa - g)}{V_c} \tag{2}$$

where $\delta(\kappa - g)$ is the (three-dimensional) delta function and V_c is the unit-cell volume. Nevertheless, a line or peak broadening is frequently found in real crystals — caused by imperfections.

We discuss the consequences of imperfections on the scattering pattern starting from the general expressions given in § A.7 of ch. 11. If defects are present in a crystal, the displacements u_n and the structure amplitudes F_{sn} will be different in different cells. For specific defects it is useful to express these quantities in terms of the properties of individual defects and their mutual arrangement. If t denotes all possible positions for a defect, and u_{tn} is the displacement vector at position n related to the defect at t, the total displacement u_n may be expressed as

$$u_n = \sum_t c_t \mathbf{u}_{tn}, \tag{3}$$

where

$$c_t = \begin{cases} 1 & \text{if a defect is present at site } t, \\ 0 & \text{if no defect is present at site } t. \end{cases} \tag{4}$$

This simple superposition is certainly justified for low defect concentrations but may be questionable in the case of concentrated alloys.

The corresponding expression for the structure amplitude of the nth cell reads

$$F_{sn} = F_{s0} + \sum_t c_t \Delta F_{stn}, \tag{5}$$

where F_{s0} is the structure amplitude of the defect-free crystal. Replacing $\boldsymbol{u}_n$ and F_{sn} in eq. (A.44), ch. 11 by eqs. (3) and (5), the mean value for $|F(\boldsymbol{\kappa})|^2$, i.e., the scattering cross section $\mathrm{d}\sigma/\mathrm{d}\Omega$ of the scattering ensemble, calculated for a random distribution of defects of concentration c, is (see KRIVOGLAZ [1969]):

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \sum_{n,n'} \left\{\exp\left[-2\pi\,\mathrm{i}\boldsymbol{\kappa}\cdot\left(\boldsymbol{r}_n - \boldsymbol{r}_{n'}\right)\right]\right\} \mathrm{e}^{-\mathcal{T}} \times \left\{\left|\bar{F}_s\right|^2 + c\sum_t \Delta F_{stn}\Delta F^*_{stn'} \exp\left[-2\pi\,\mathrm{i}\boldsymbol{\kappa}\cdot\left(\boldsymbol{u}_{tn} - \boldsymbol{u}_{tn'}\right)\right]\right\}, \tag{6}$$

where

$$\mathcal{T} = c\sum_t \left\{1 - \exp\left[-2\pi\,\mathrm{i}\,\boldsymbol{\kappa}\cdot\left(\boldsymbol{u}_{tn} - \boldsymbol{u}_{tn'}\right)\right]\right\}\left[1 + \frac{1}{F_{s0}}\left(\Delta F_{stn} + \Delta F_{stn'}\right)\right], \tag{7}$$

and $\bar{F}_s$ is the average structure amplitude of the crystal with defects:

$$\bar{F}_s = F_{s0} + c\sum_t \Delta F_{stn}\,. \tag{8}$$

It is now possible to split $\mathrm{d}\sigma/\mathrm{d}\Omega$ into two terms, a Bragg-like term with sharp maxima at the Bragg positions, and a second term describing scattering for all other scattering vectors. Taking into account that for the Bragg peaks, $\mathrm{d}\sigma/\mathrm{d}\Omega$ from the double sum of eq. (6) is controlled by terms corresponding to large distances $\rho = |\boldsymbol{r}_n - \boldsymbol{r}_{n'}|$, we may write:

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_B = N\left|\bar{F}_s\right|^2 \mathrm{e}^{-2M}\,\frac{\delta(\boldsymbol{\kappa} - \boldsymbol{g})}{V_c}, \tag{9}$$

where $2M$ is the real part of $\mathcal{T}$ according to eq. (7) for $\rho \to \infty$:

$$2M = c\lim_{\rho\to\infty}\sum_t \left\{1 - \cos\left[2\pi\boldsymbol{\kappa}\cdot\left(\boldsymbol{u}_{tn} - \boldsymbol{u}_{tn'}\right)\right]\right\}\left[1 + \frac{\Delta F_{stn} + \Delta F_{stn'}}{F_{s0}}\right]. \tag{10}$$

Equation (9) defines a generalized Debye–Waller factor e^{-M}. The reciprocal lattice vectors $\boldsymbol{g}$ occurring in eq. (9) are defined in the average lattice that *contains* defects. Defects may cause a shift of Bragg peaks because the average lattice constant changes, but also a reduction in intensity due to the factor e^{-2M} [eqs. (9) and (10)]. As long as $2M$ remains finite, Bragg peaks will still be present, but there are cases where $2M$ tends to infinity.

If we subtract the Bragg intensity according to eq. (9) from the total scattering cross-section, eq. (6), we obtain the *diffuse scattering* cross-section

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_d = \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} - \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_B. \tag{11}$$

The cross-section $(d\sigma/d\Omega)_d$ contains no δ-function but varies smoothly with κ, even for an infinitely large crystal. KRIVOGLAZ [1969] distinguishes two types of defects depending on whether or not the Bragg intensities are reduced to zero [eq. (9) with $2M \to \infty$]. The limiting behaviour of $2M$ [eq. (10)] may be discussed by considering a displacement $\boldsymbol{u}_{tn}$ that decreases rapidly towards zero with increasing distance between t and n. Then, only one of the displacements $\boldsymbol{u}$ and one of the ΔF_s in eq. (10) are markedly different from zero, and:

$$M = c\sum_t \left[1 - \cos(2\pi\boldsymbol{\kappa}\cdot\boldsymbol{u}_{tn})\right]\left(1 + \frac{\Delta F_{stn}}{F_{s0}}\right). \tag{12}$$

The convergence of the sum depends on the contributions from large distances between n and t. There, $\boldsymbol{u}_{tn}$ is small and the cosine can be expanded, so that $1-\cos(2\pi\boldsymbol{\kappa}\cdot\boldsymbol{u}_{tn}) = (2\pi\boldsymbol{\kappa}\cdot\boldsymbol{u}_{tn})^2/2$, and if one writes $\boldsymbol{u}_{tn}$ as a continuous function of $\boldsymbol{r}' = \boldsymbol{r}_t - \boldsymbol{r}_n$, the convergence of the integral $\int(\boldsymbol{\kappa}\cdot\boldsymbol{u}_{tn})^2 d^3\boldsymbol{r}'$, $\infty \geq r' > r'_0$ (≫ lattice constant), will assure a finite value of M. We see that for a large class of defects where $u \propto (r')^{-2}$ (see ch. 17), M remains finite. Others, e.g., straight dislocations, small-angle grain boundaries and stacking faults, cause e^{-2M} to decrease to very small values for large crystals, and the total scattering is diffuse although it will be concentrated, but not in a δ-like fashion, near the original Bragg positions (line broadening).

The distinction between Bragg intensity [eq. (9)] and diffuse intensity (line broadening and scattering far away from Bragg peaks), discussed here for systems with a random distribution of equivalent defects, will not always be possible experimentally, as the width of the Bragg peak is also affected by the resolution function of the instrument. Correlations in the arrangement of defects may reduce long-range displacements and modify the expected effect, as will the presence of different types of defects. Figure 1 shows schematically, how the scattering of an ideal crystal is modified by defects.

2.2. X-rays and neutrons

Apart from electrons (see ch. 11), X-rays (and γ-rays) and thermal neutrons are frequently used for structure determination and the study of defects as they provide the appropriate range of wavelengths for such investigations. Table 1 gives some of the important properties of both types of radiation. For the wavelengths of interest in normal diffraction work (0.5 to 20 Å), corresponding photon energies are in the range of about 1–40 keV whereas neutron energies are between 0.85 and 400 meV. Excitations in condensed matter (phonons, magnons, etc.) are in the range of a few meV and above. The relative energy change of X-rays scattered inelastically (with energy loss or gain) by any sample is then very small (say $< 10^{-6}$) and cannot easily be resolved (except if Mössbauer sources or synchrotron radiation combined with backscattering techniques are used). In contrast, neutrons can experience an appreciable relative change in energy, so that elastic (no energy change) and inelastic scattering can be distinguished (see below, § 2.4). Another important difference arises from the magnetic moment of the neutron which interacts with the local magnetization density. This leads *to magnetic scattering* which has very important applications as a sensitive probe in the study of magnetic

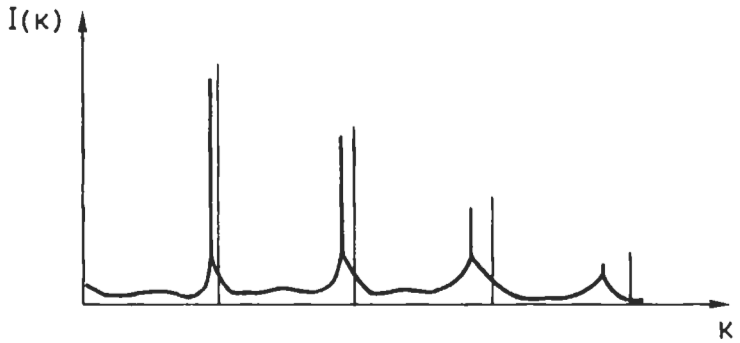


Fig. 1. Scattering from a crystal containing defects. Bragg positions of a perfect crystal are indicated by thin vertical lines.

substances (see § 2.3). Finally, absorption differs appreciably for the two types of radiation. (Measurements of absorption can also yield considerable insight in the structure of matter, e.g., in EXAFS = extended X-ray absorption fine structure, see § 5.4.).

The linear absorption coefficient Σ_t is defined by

$$\ln(I/I_0) = -\Sigma_t D_s, \tag{13}$$

where I_0 and I are the intensities of incident and transmitted beam, respectively, and D_s is the sample thickness. The absorption coefficient has the dimension of 1/length or area/volume, and can be understood as a total macroscopic removal cross-section, as is

Table 1
Some properties of X-rays and neutrons.

Property	Value[a]	
	For X-rays (photons)	For neutrons
mass m [kg]	0	1.675×10^{-27}
momentum	h/λ	$h/\lambda = mv$
energy	hc/λ	$h^2/2m\lambda^2$
scattering length	Zf_0 (for $\kappa \rightarrow 0$)	b(nuclear)
absorption	$\propto Z^4\lambda^3$ (strong)	$\propto \lambda$ (mostly weak)
speed v[m/s][b]	c	$437\ E^{1/2}$
wavelength λ[Å][b]	$12.4/E$	$3956/v$
magnetic moment	0	$-1.913\ \mu_n$
frequency ν[s^{-1}][b]	$2.42 \times 10^{17}\ E$	$2.42 \times 10^{11}\ E$

[a] h = Planck's constant, c = velocity of light, f_0 = classical electron radius, μ_n = nuclear magneton.
[b] For photons, E in keV; for neutrons, E in meV.

common in neutron scattering. In the X-ray literature, a mass absorption coefficient is defined by Σ_t/ρ with ρ = density. Its dimensions are area/mass, and values are independent of the thermodynamic state of the sample material.

All beam-attenuating processes (including coherent and incoherent scattering) are included in Σ_t, but for X-rays the excitation of fluorescence radiation can be singled out as the most important true absorption mechanism. When the energy of the incident X-rays approaches a resonance energy in the electronic states, an absorption edge is observed, corresponding to the excitation of electrons in the K, L, etc., levels (see fig. 2 where the mass absorption coefficient is shown as a function of wavelength). The electrons associated with the absorption move out of phase from the others and a destructive interference results that reduces the atomic scattering factor by the equivalent of twice the number of resonance electrons (e.g., at least a total of four at the K edge). The atomic scattering factor has to be corrected according to

$$f_a = f_{a0} + f' + if'', \tag{14}$$

where f' and f'' are the real and the imaginary part of the *Hönl corrections*. The wavelength dependence of f' and f'' is shown schematically in fig. 3 (see International Tables [1952–74], vol. III. pp. 214–216 and vol. IV, p. 148, and SASAKI [1989, 1990] for tabulated values). As synchrotron radiation sources now provide strong X-ray beams with continuously tunable wavelength, diffraction work in the neighbourhood of absorption edges can take advantage of the variation of specific scattering factors.

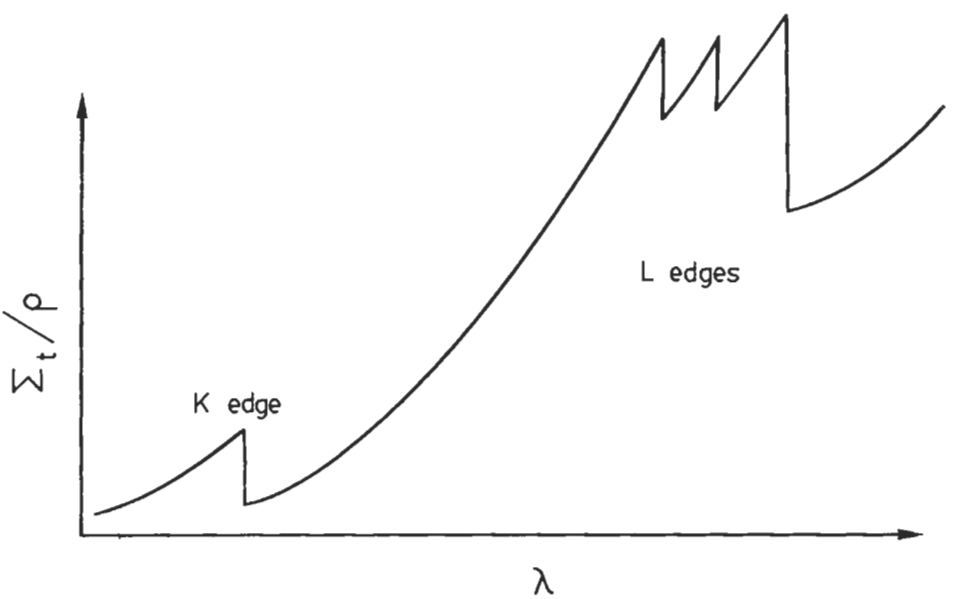


Fig. 2. Mass absorption coefficient Σ_t/ρ for X-rays as a function of wavelength near the K and L edges of an element (schematic).

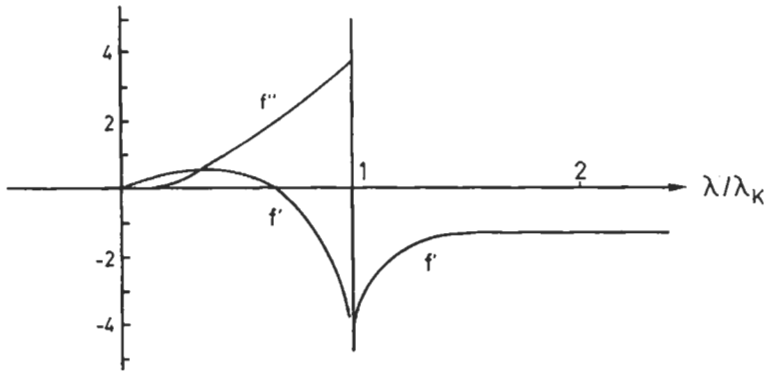


Fig. 3. Variation of real and imaginary components f' and f'' of the Hönl corrections as a function of X-ray wavelength near the K absorption edge (at λ_K) of an element. (After COWLEY [1981].)

Table 2 gives a comparison of linear (and mass) absorption coefficients for a few metals. According to eq. (13), $\Sigma_t D_s = 1$ corresponds to a reduction in intensity by a factor $1/e$ which gives an estimate for the typical thickness of a sample. Table 2 shows that even Cd, a material frequently used for shielding purposes in thermal neutron work, has a smaller absorption for neutrons than most common metals for X-rays. On the other hand, lead, a good shielding material for X-rays, is almost transparent to neutrons. (See BACON [1975] and SCHWARTZ and COHEN [1987] for more details on the different properties of X-rays and neutrons.)

For high-energy X-rays (about 100 keV or more), the absorption by matter is very

Table 2
Linear absorption coefficients of some metals, for neutrons (true absorption, Σ_a) and X-rays (Σ_t, total removal) and corresponding mass absorption coefficients (values from BACON [1975], and KOSTORZ and LOVESEY [1979]).

Metal	Neutrons Σ_a (cm^{-1}) 1.8 Å	X-rays Σ_t (cm^{-1}) 1.54 Å	Neutrons Σ_a/ρ (m^2/kg) 1.8 Å	X-rays Σ_t/ρ (m^2/kg) 1.54 Å
Be	0.001	2.7	0.00006	0.150
Al	0.0014	131	0.0005	4.86
Cu	0.326	474	0.0036	5.29
Cd	119	2000	1.29	23.1
W	1.22	3320	0.0061	17.2
Au	5.72	4170	0.029	20.8
Pb	0.005	2630	0.00005	23.2

weak, and the incident beam is mainly attenuated by Compton scattering. Diffraction experiments with the 412 keV γ-radiation from radioactive gold were initiated in Grenoble in the sixties (see SCHNEIDER [1981] for a detailed account) as a means to study the mosaic structure of large single crystals, of sizes comparable to those used for monochromators of thermal neutrons or in neutron scattering experiments. As the γ-wavelength is small (0.03 Å in this case), the Bragg angles are also correspondingly small (~ 1°). Thus, lattice tilts may be monitored directly for large crystals, also, owing to the easy penetration of windows, etc., during crystal growth (see BREUER and BLEICHERT [1993] and references therein for further applications). Diffractometers in the 100–150 keV range may now also be found at suitable synchrotron radiation facilities (see, e.g., BOUCHARD *et al.* [1993]).

2.3. Magnetic scattering

Although there is a measurable interaction between photons and electronic magnetic moments, it is too small (cross-sections about six to eight orders of magnitude smaller than for Bragg scattering, see DE BERGEVIN and BRUNEL [1972]) to be used conveniently in the study of magnetic structures.

The magnetic moment of the neutron has a value of $\gamma\mu_n$ (see table 1) with $\gamma = -1.913$. Examination of the dipole interaction between neutron and local magnetic induction (see GUREVICH and TARASOV [1968], LOVESEY [1984]) shows that of the Fourier transform of the local magnetization density, $\boldsymbol{M}(\boldsymbol{\kappa})$, only the component $\boldsymbol{M}^{\perp}(\boldsymbol{\kappa})$, perpendicular to the scattering vector $\boldsymbol{\kappa}$, contributes to the scattering. Therefore, a *magnetic interaction vector* may be defined by

$$\boldsymbol{q}_{\mathrm{M}} = \boldsymbol{M}^{\perp}(\boldsymbol{\kappa})/|\boldsymbol{M}(\boldsymbol{\kappa})|. \tag{15}$$

The *magnetic scattering length*, commonly denoted by $\boldsymbol{p}$, is *a vector*:

$$\boldsymbol{p} = p\boldsymbol{q}_{\mathrm{M}}, \tag{16}$$

with p proportional to $|\boldsymbol{M}(\boldsymbol{\kappa})|$. If we consider magnetic scattering due to unpaired spins (without orbital moments), p is given by

$$p(\boldsymbol{\kappa}) = \frac{e^2}{m_0 c^2}|\gamma| S f_{\mathrm{mag}}(\boldsymbol{\kappa}). \tag{17}$$

Here, e^2/m_0c^2 is the classical electron radius f_0 known from X-ray scattering [$f_0 = 2.8 \times 10^{-15}$ m, see eq. (A.37), ch. 11], S is the total number of unpaired spins, and $f_{\mathrm{mag}}(\boldsymbol{\kappa})$ is the Fourier transform of the spin density, normalized to $f_{\mathrm{mag}}(0) = 1$. As the number μ of Bohr magnetons ($\mu_{\mathrm{B}} = eh/m_0c$) equals $\mu = 2S$, we can write

$$p(\boldsymbol{\kappa})\left[10^{-14}\mathrm{m}\right] = 0.27\mu f_{\mathrm{mag}}(\boldsymbol{\kappa}). \tag{18}$$

For not too large values of $\boldsymbol{\kappa}$, p is comparable to the values of the nuclear scattering amplitude b (for the pure elements Fe, Co, Ni one has $b = 0.96$, 0.28, 1.03 and $p(0) = 0.6$, 0.47 and 0.16, respectively, all in units of 10^{-14} m). The magnetic form factor $f_{\mathrm{mag}}(\boldsymbol{\kappa})$

falls off more rapidly with $\sin\theta/\lambda$ than the atomic form factor for X-rays (ch. 11), as only a few electrons in an outer shell contribute to f_{mag}.

The total scattering of a magnetic material will show a combination of nuclear and magnetic scattering and will also depend on the polarization of the incident beam. Figure 4 shows the scattering geometry, assuming a fully polarized beam with the polarization vector parallel to the z-axis (up: +, down: −). If the scattered beam is analyzed along the same direction, four different scattering cross-sections can be defined (+ +, + −, − + and − −). In an otherwise perfect crystal, the *nuclear* unit-cell structure factor, $F_s(\boldsymbol{\kappa})$ [see eq. (A.32), ch. 11]), can be combined with the *magnetic structure factor*,

$$\boldsymbol{F}_{\mathrm{M}}^{\perp}(\boldsymbol{\kappa}) = \sum_m \boldsymbol{q}_{\mathrm{M}m} p_m \exp\left(-2\pi\, \mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{r}_m\right), \tag{19}$$

where $\boldsymbol{q}_{\mathrm{M}m}$ [eq. (15)] is defined according to the direction of the moment at site m, and p_m at each site is given by eq. (18). This yields $|F_s + F_{\mathrm{M}z}^{\perp}|^2$ for (+ +), $|F_{\mathrm{M}x}^{\perp}|^2 + |F_{\mathrm{M}y}^{\perp}|^2$ for (+ −) and (− +), and $|F_s - F_{\mathrm{M}z}^{\perp}|^2$ for (− −), where the $F_{\mathrm{M}i}^{\perp}$ are the Cartesian components of $\boldsymbol{F}_{\mathrm{M}}^{\perp}$.

An unpolarized beam can be assumed to be composed of 50% positive and 50% negative polarization, and the scattering cross-section (per atom) without polarization analysis is $\mathrm{d}\sigma/\mathrm{d}\Omega = |F_s|^2 + |\boldsymbol{F}_{\mathrm{M}}|^2$. For a simple ferromagnet, $\boldsymbol{q}_{\mathrm{M}m}$ is the same for all sites, and with $\boldsymbol{F}_{\mathrm{M}}^{\perp} = \boldsymbol{q}_{\mathrm{M}} F_{\mathrm{mag}}$ [see eq. (19)]:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |F_s|^2 + |F_{\mathrm{mag}}|^2 \sin^2\alpha, \tag{20}$$

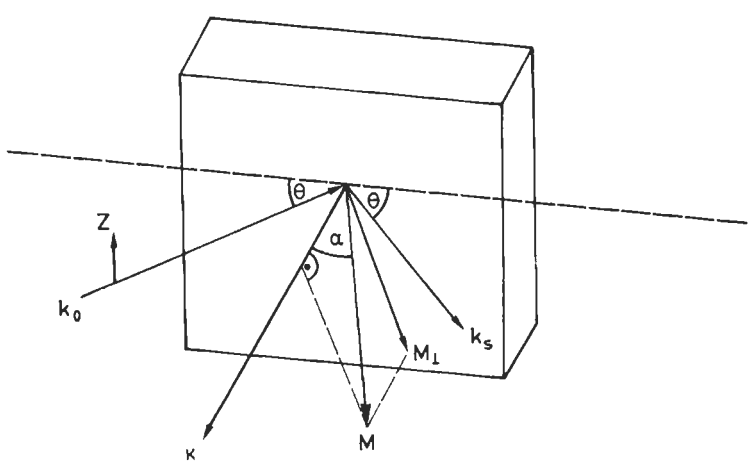


Fig. 4. Scattering geometry for magnetic scattering. The incident neutrons are polarized along the z-axis.

where α is the angle between κ and M (see fig. 4). By varying α (via the external magnetic field or the scattering geometry), nuclear and magnetic contributions can be separated. Equation (20) also applies to antiferromagnets if we define α relative to the direction of the moment at site one and write $F_{mag} = \Sigma_m \hat{\sigma}_m p_m \exp(-2\pi i \boldsymbol{\kappa} \cdot \boldsymbol{r}_m)$ where $\hat{\sigma}_m$ is either +1 or −1 depending on whether or not the moment at site m is parallel or antiparallel to $\boldsymbol{q}_1$. Figure 5 shows the consequences of ferro- and antiferromagnetic order in the bcc lattice ($\alpha \neq 0$). Additional peaks ("magnetic peaks") occur for antiferromagnetic order whereas in ferromagnets the nuclear peaks are enhanced.

If a polarized beam is used, the cross-section for scattering without polarization analysis is the sum of, e.g., (+ +) and (+ −) scattering and contains an interference term between nuclear and magnetic scattering with F_s and $\boldsymbol{F}_M^{\perp}$ appearing unsquared;

$$\left(\frac{d\sigma}{d\Omega}\right)_{P_0} = |F_s|^2 + 2F_s \boldsymbol{F}_M^{\perp} \cdot \boldsymbol{P}_0 + \left|\boldsymbol{F}_M^{\perp}\right|^2 . \tag{21}$$

where $\boldsymbol{P}_0$ is a unit vector indicating the direction of polarization. "Flipping" $\boldsymbol{P}_0$ from +1 to −1, we have a very sensitive method to measure $\boldsymbol{F}_M^{\perp}$ and consequently $F_{mag}(\boldsymbol{\kappa})$, $p_m(\boldsymbol{\kappa})$ or $f_{mag}(\boldsymbol{\kappa})$. Equation (21), if generalized as indicated by eq. (8) for structural disorder, is also the basis for a separation of structural and magnetic disorder in alloys.

Finally we mention the original result of HALPERN and JOHNSON [1939] for paramag-

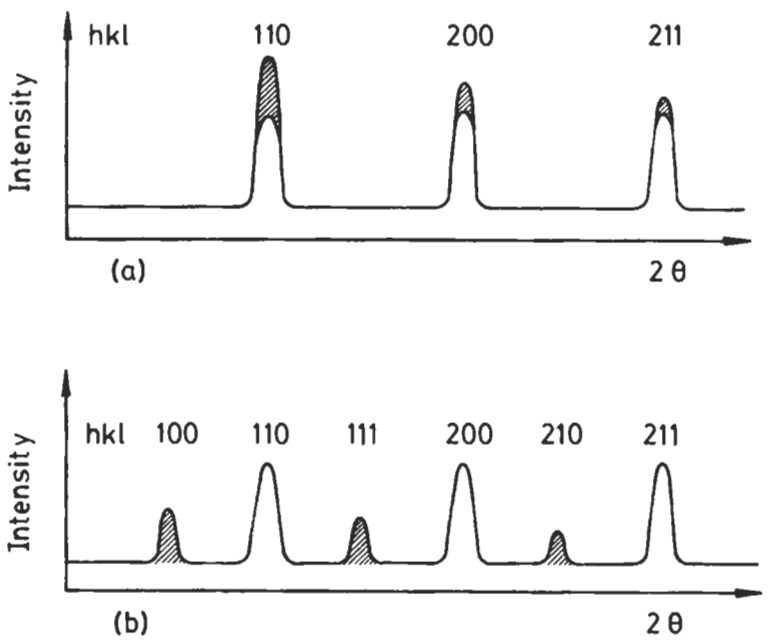


Fig. 5. Schematic neutron diffraction patterns from a bcc poylcrystalline (a) ferromagnet and (b) antiferromagnet. The shaded areas represent the magnetic scattering contributions which decrease with θ because f_{mag} decreases. (After SCHWARTZ and COHEN [1987].)

netic scattering, i.e., for a randomly oriented ensemble of identical magnetic moments (spins S). The paramagnetic scattering cross-section (per atom) reads

$$\left(\frac{d\sigma}{d\Omega}\right)_p = \tfrac{2}{3} S(S+1) f_0^2 \gamma^2 f_{mag}^2 . \tag{22}$$

This scattering is similar to the Laue scattering term in diffuse scattering. Deviations from the simple monotonic κ-dependence are of interest in both cases as they relate to correlations in the atomic or the spin arrangement.

There are many more special features of magnetic scattering than this brief discussion can indicate (see BROWN [1979], HICKS [1979] for more details) and many complex magnetic structures have been studied in recent years.

2.4. Inelastic and quasi-elastic scattering

As mentioned in § 2.2., the relative energy gain or loss of X-rays scattered from a sample with lattice vibrations (phonons) is very small. Near an absorption edge resonant Raman scattering (see SPARKS [1974]) occurs just below the edge, and incoherent fluorescence radiation is emitted above the edge. Compton scattering is another inelastic scattering process of X-rays. From the conservation of energy and momentum for the scattering of a photon from an individual electron, the wavelength shift of Compton-modified radiation is (for electrons assumed at rest):

$$\Delta\lambda\left[\text{Å}\right] = 0.0243(1 - \cos 2\theta), \tag{23}$$

independent of the incident wavelength. The scattering is incoherent as there is no fixed phase-relationship between the different inelastic scattering events. The relative contribution of Compton scattering to the total scattering is given by $1 - \Sigma_j f_{aj}^2 / Z]$, where f_{ai} is the form factor (normalized to one) of each of the Z electrons of an atom. Compton scattering increases the background in X-ray diffraction experiments, but it can be eliminated experimentally (e.g., with a monochromator in the diffracted beam) or by calculation (see SCHWARTZ and COHEN [1987]). Detailed study of the energy distribution of Compton intensities provides information on the momentum distribution of electrons in condensed matter (see, e.g., COOPER [1977]), one of the few techniques also applicable to alloys.

Much smaller (absolute) energy transfers can be detected by neutron scattering, either by analyzing the change of wavelength with a single crystal or the change of neutron momentum by time-of-flight methods. Here, we consider one-phonon scattering only (for a complete account see, e.g., LOVESEY [1984] and BACON [1975]).

Figure 6 shows two possible scattering configurations in reciprocal space. The scattering vector $\boldsymbol{\kappa}$ can be written as

$$\boldsymbol{\kappa} = \boldsymbol{k}_s - \boldsymbol{k}_0 = \boldsymbol{g} + \boldsymbol{q} \tag{24}$$

where $\boldsymbol{q}$ is the phonon wavevector, counted from the nearest reciprocal lattice point. If the wavevector after scattering is $\boldsymbol{k}_{s1}$, the neutron has lost energy ($k_{s1} < k_0$), and a phonon

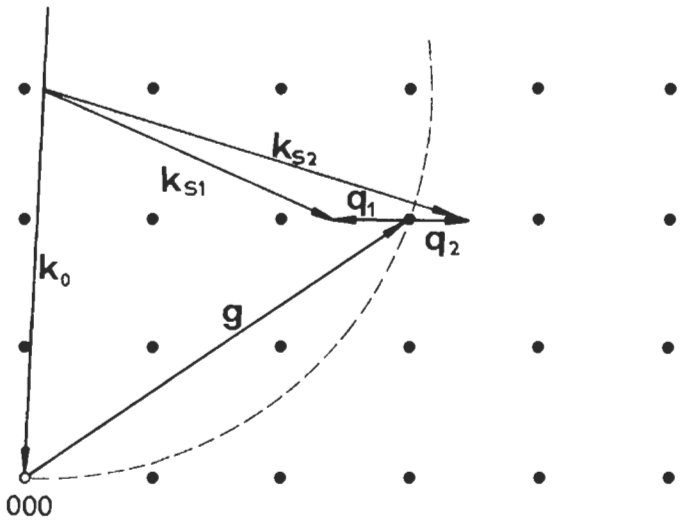


Fig. 6. Two possible inelastic scattering events in a plane of reciprocal space involving the creation ($\kappa = k_{s1} - k_0$) and the annihilation ($\kappa = k_{s2} - k_0$) of a phonon.

has been created. For k_{s2}, a phonon has been annihilated. As the neutron momentum is h/λ, eq. (24) states the conservation of momentum in the scattering process ($\lambda_q = 1/q$ is the phonon wavelength). Simultaneously, the conservation of energy,

$$h^2/2m_n\lambda_0^2 - h^2/2m_n\lambda_s^2 = h\nu_p\,, \tag{25}$$

defines the energy of the phonon participating in the scattering (λ_0 and λ_s denote the wavelength of the incident and scattered neutrons, respectively). There are only discrete values of phonon frequency ν_p for a given q, and appropriate scans can be designed to obtain directly the phonon dispersion curves of a crystal.

Apart from coherent inelastic scattering processes, analysis of the incoherent inelastic scattering of neutrons may often be interesting. True incoherent scattering processes are due to the interaction of the neutron spin with nuclear spins $I \neq 0$. The scattering length of the compound nucleus depends on its total spin which is $I + 1/2$ or $I - 1/2$. The tabulated values of coherent scattering lengths and incoherent cross-sections for individual isotopes represent properly weighted averages (for unpolarized nuclei). Natural elements are frequently a mixture of different isotopes, each with its own nuclear spin. Coherent scattering lengths and spin-incoherent cross-sections are simply arithmetic averages, but owing to the random distribution of nuclei with different coherent scattering lengths over the sites of the sample, another κ-independent term, $\sigma_i = 4\pi\left\{\overline{b^2} - \left(\overline{b}\right)^2\right\}$, occurs in complete analogy to

References: p. 1188.

the monotonic Laue scattering term in diffuse X-ray scattering (see § 4). For neutrons, this part is included in the total value of incoherent scattering of an element. Incoherent neutron scattering cross-sections are quite large in several cases, and for coherent scattering experiments (see § 2.5), they may impose severe limitations. However, similar to the case of Compton scattering and electrons, energy analysis of the incoherently scattered neutrons will reveal details of the motion of nuclei. The incoherent scattering function can be calculated for different processes (see, e.g., BACON [1975], LOVESEY [1984]). If the energy transfer is centred around $h\nu = 0$, i.e., is small, the scattering is called quasi-elastic. In condensed matter, the motion of atoms is restricted by the environment, and quasi-elastic neutron scattering has become a widely used technique to study atomic and molecular motion, especially if hydrogen with its high incoherent cross-section is involved. Especially in crystalline solids, a study of the quasi-elastic line width (coherent or incoherent) may help to identify diffusion mechanisms (see § 6.2).

2.5. Some experimental considerations

Incoherent scattering as a source of background and absorption was already mentioned in § 2.4. Extensive discussions on general experimental problems in X-ray and neutron scattering can be found in the book by SCHWARTZ and COHEN [1987]), and, more specifically oriented towards defects in crystals, in the articles by SCHMATZ [1973, 1978]. Some remarks will be found in subsequent chapters, relating to particular applications. Here, we only state a few quite general points.

Laboratory X-ray sources with rotating anodes have been built, the maximum power of 100 kW (see HAUBOLD [1975]) yielding 4×10^{16} quanta/s cm^2 sr (CuK_α). Higher luminosities are obtained (originally as a by-product, now from dedicated sources) from electron or positron synchrotrons and storage rings. Depending on the energy of the electrons or positrons, synchrotron radiation emerges in a wavelength range from infrared to ~ 0.2 Å. Synchrotron radiation is continuous and very intense (10^{12}–10^{14} quanta/s mrad per 0.1% bandwidth, integrated over the vertical direction). The radiation is highly collimated perpendicular to the orbital plane. In the GeV region, the divergence is $\sim 10^{-4}$ rad. In the plane of the ideal orbit, synchrotron radiation is 100% polarized with the electrical vector parallel to this plane. There is a well-defined time structure of the beam with pulse durations as short as 10^{-10} s with repetition rates of 1 MHz or more. The use of synchrotron radiation for materials studies is currently experiencing a very rapid growth (see e.g. BARUCHEL *et al.* [1993]).

Compared even with a classical sealed X-ray tube, neutron sources are not very powerful. Figure 7 shows a comparison of several X-ray and neutron sources according to SCHMATZ [1973]. As suggested by MAIER–LEIBNITZ [1966], the momentum space density $p(\boldsymbol{k}_0)$ ($h\boldsymbol{k}_0$, where h is Planck's constant, is the linear momentum of a particle) is an adequate quantity for comparison, as the count rate $\dot{Z}$ at a detector is proportional to $p(\boldsymbol{k}_0)$ multiplied by the momentum space elements $\mathrm{d}^3\boldsymbol{k}_0$ and $\mathrm{d}^3\boldsymbol{k}_s$ that can be optimized instrument parameters (within the bounds of k_x, k_y, k_z fixed by the properties of the source), $\dot{Z} \propto p(\boldsymbol{k}_0) S(\boldsymbol{\kappa}, \nu) \mathrm{d}^3\boldsymbol{k}_0 \mathrm{d}^3\boldsymbol{k}_s$, where $S(\boldsymbol{\kappa}, \nu)$ is the scattering law to be studied. We see that $p(\boldsymbol{k}_0)$ for neutrons is several orders of magnitude lower than the values of all

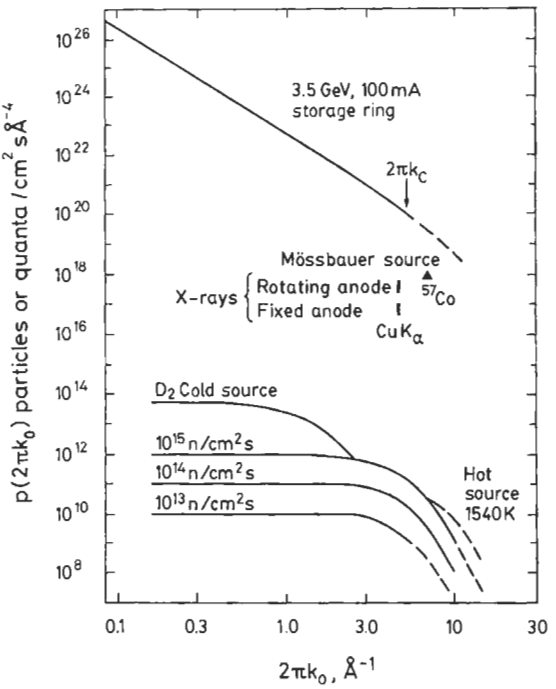


Fig. 7. Momentum space density $p(2\pi k_0)$ for different X-ray and neutron sources (after SCHMATZ [1973]).

X-ray sources. On the other hand, restrictions on the choice of parameters of the incident beam ($\Delta k_0/k_0$ or vertical collimation), which are common for X-rays, are less severe for neutrons where $d^3\boldsymbol{k}_0$ can be chosen more flexibly to match the resolution requirements for a given $S(\boldsymbol{\kappa}, \nu)$. This may compensate in part for the lower values of $p(\boldsymbol{k}_0)$.

Traditionally, thermal neutrons are produced by fission and subsequent moderations i.e., in a nuclear reactor. A steady-state thermal flux of about 10^{15} n/cm^2 is virtually impossible to surpass as heat removal is a limiting factor. With hot and cold sources, the Maxwellian spectrum of neutrons may be shifted to smaller or larger wavelengths (see fig. 7). Pulsed neutron sources may provide a higher peak flux. If the pulsed structure of the neutron beam is maintained after moderation, time-of-flight experiments are advantageous at pulsed sources because the relevant flux at the sample will be higher than the average flux. One concept for the production of pulsed neutrons is based on the use of charged particles from accelerators (spallation). A variety of spallation sources are now operational.

References: p. 1188.

Apart from film techniques (with a converter foil for neutrons), the *detection* of the scattered radiation is based on electronic counting circuits attached to gas or solid state detectors. For X-rays, gas-filled proportional "counters" and solid-state scintillation "counters" detect the incident quanta via the ionization of a gas or the production of photoelectrons by scintillation photons. The energy resolution is poor for both detector types, typically ~ 20% in the 10 keV range for gas detectors and ~50% for scintillation detectors. Solid-state detectors based on electron–hole pair production in doped (Ge or Si with Li) or intrinsic (Ge) semiconductors have a theoretical resolution of about 1% at 10 keV, about 200 eV in practice.

Another important aspect for many experiments is the spatial resolution of large detectors. Linear position-sensitive detectors with a resolution of about 30 μm employ a resistive wire, and peak heights or pulse shapes are analyzed as a function of position of the detected event. Two-dimensional gas-filled detectors for X-rays employ (sets of) mutually perpendicular wires and different electronic techniques to locate the detected events. The simultaneous measurement of scattering over a large solid angle is of particular importance for weak scattering signals but also for in-situ kinetic studies.

Neutrons can only be detected after they have participated in a nuclear reaction with the emission of charged particles or γ-rays. As beam dimensions are usually much larger in neutron scattering than in X-ray scattering, the use of large arrays of individual detectors is feasible without loss of resolution (and with time-of-flight resolution, i.e., energy resolution for each desired detector). Position-sensitive detectors have been developed for diffraction work, planar detectors for small-angle scattering and curved ones for powder patterns.

3. *Bragg peaks and vicinity*

Scattering from real crystals may occur anywhere in reciprocal space but Bragg peaks, though modified, will remain a predominant feature as long as an average lattice can be defined. A diffraction pattern, after all instrumental corrections, reflects the distribution of scattering matter within the (average) unit cell of a substance, and the atomic coordinates (location of the centres of scattering objects) as well as the scattering-length density distribution (electrons with X-rays, nuclei and magnetic moments with neutrons) can be determined from an analysis of Bragg peaks (see, e.g., LIPSON and COCHRAN [1953], WARREN [1969], SCHWARTZ and COHEN [1987]). Particular Bragg peaks occur for different phases in a sample, and diffraction methods are thus essential in the study of phase diagrams and phase transitions. The orientation distribution of Bragg peaks for a polycrystalline sample reveals its texture (see, e.g., BARRETT and MASSALSKI [1980]).

3.1. Peak shifts

Changes in peak position may be caused by residual stresses, by faulting on certain crystallographic planes (WARREN [1969], WAGNER [1966], COWLEY [1981]), or by lattice parameter changes as a function of alloying or of defect concentration. Whereas a lattice

parameter change affects all crystallographically equivalent Bragg peaks in the same way, the other two causes may lead to different shifts depending on the orientation of the reflecting planes relative to a preferred axis of the disturbance. Figure 8 illustrates this for a polycrystal with a surface under compression. As residual stresses and applied stresses combined determine load carrying capabilities, X-ray methods in this field have received considerable attention (e.g., BARRETT and MASSALSKI [1980], CULLITY [1977]). The book by NOYAN and COHEN [1987] gives a detailed account of procedures and applications.

As the penetration depth of X-rays is limited and strongly depends on wavelength, only near-surface regions are accessible. This requires some caution in the interpretation, but may also be used to obtain some information on stress gradients perpendicular to the surface (see, e.g., OETTEL [1989]). The preferred Bragg diffraction from near-surface regions leads to important applications in the study of epitaxial layers and hetero-structures (see, e.g., CHU *et al.* [1985], ZOGG *et al.* [1994]), as accommodation stresses, their inhomogeneity and their relaxation may be determined.

With the availability of powerful X-ray beams at synchrotron radiation sources, the surface sensitivity for Bragg scattering may be considerably improved (several nm instead of μm) by grazing incidence. For an incident beam which impinges on the sample at a glancing angle at or below the limiting angle of total reflection, diffraction peaks may arise from the evanescent wave travelling in the crystal parallel to the surface and decaying exponentially perpendicular to the surface (see MARRA *et al.* [1979]).

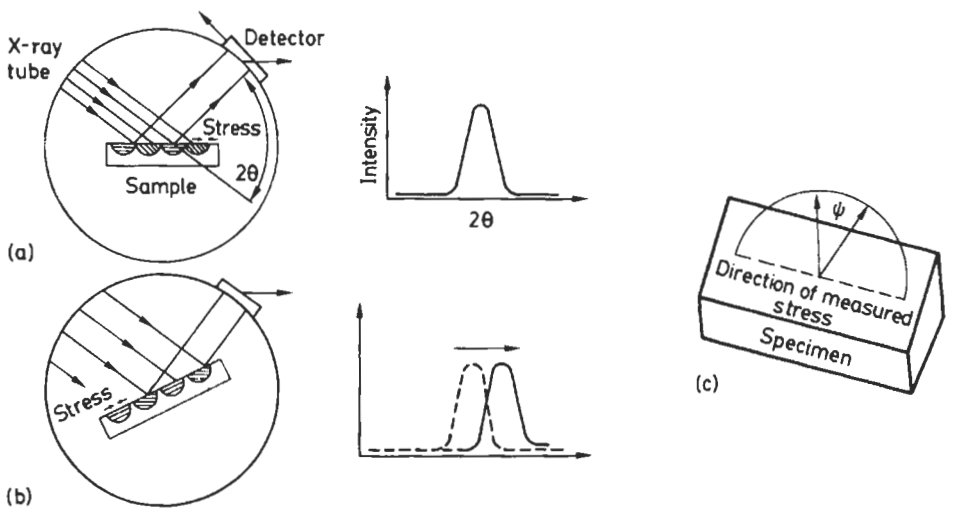


Fig. 8. Residual stress measurements with an X-ray diffractometer for a surface of a polycrystal under compression (schematic, not to scale). (a) Bragg's law is satisfied for lattice planes parallel to the surface. These planes are further apart than in the stress-free state (Poisson's effect). Their spacing is obtained from the position of the Bragg peak. (b) The specimen has been tilted, and other grains now present suitable lattice planes for Bragg scattering. As these planes are more nearly perpendicular to the compressive stress, they are less separated than in (a), and the Bragg peak moves to higher angles. (c) The direction of the measured stress is given by the intersection of the circle of tilt and the surface. (After M.R. JAMES and COHEN [1980].)

References: p. 1188.

Evanescent scattering of X-rays and neutrons offers many new experimental possibilities in surface science (see FEIDENHANS'L [1989], ZABEL and ROBINSON [1992], and DOSCH [1992, 1993]). As examples for evanescent X-ray Bragg scattering we quote the studies on depth-resolved near-surface ordering of Cu_3Au (ZHU *et al.* [1988], DOSCH *et al.* [1991a] and on surface-melting of an Al single crystal (DOSCH *et al.* [1991b]).

Gradients of lattice constants may also occur in diffusion couples, e.g., when a plated sample is annealed. The (nondestructive) analysis of such diffusion profiles using X-ray diffraction starts from similar ideas (see HOUSKA [1980]).

Neutron diffraction averages over larger volumes, and several nondestructive measurements of internal stresses in "real" components have been reported (see e.g., ALLEN *et al.* [1985], PRIESMEYER *et al.* [1994]). The method has a special value for "hidden" stress concentrations.

3.2. Peak broadening and intensity changes

In § 2.1 we have already mentioned a broadening due to sample size that could be relevant for powders [eq. (1)]. However, for bulk samples we have tacitly ignored size broadening and written the Bragg cross-section as a delta-function. Thus, for a well collimated X-ray beam of characteristic radiation, the width of a Bragg peak should be controlled by the natural linewidth, since Bragg's law yields (θ being a Bragg angle):

$$\Delta(2\theta) = 2\frac{\Delta\lambda}{\lambda}\tan\theta. \tag{26}$$

With $\Delta\lambda/\lambda \approx 5\times 10^{-4}$, $\Delta(2\theta)$ is about one minute of arc at $2\theta \approx 30°$ but increases dramatically for $\theta \to 90°$, and eq. (26) can be confirmed for large perfect crystals. However, quite frequently, the Bragg peaks are much broader (many minutes at moderate Bragg angles), and only crystals with such *mosaicity* will actually approach the predictions of the kinematical theory as may be judged from the integrated reflecting power (which should be proportional to $|F_s|^2$). According to DARWIN ([1922], see ZACHARIASEN [1945] and R.W. JAMES [1963]), the kinematical theory is valid for an ideal mosaic crystal consisting of small perfect domains which are tilted against each other by small angles (a few minutes). According to this concept, the effect of coherent multiple reflection of the same beam can be reduced to just the planes within one *mosaic block* and finally neglected (i.e., primary extinction plays no role). The dynamical theory of diffraction shows (see ch. 11) that the tolerable size of the blocks depends on wavelength. As the extinction length (for Bragg reflections) is proportional to $(\lambda|F_s|)^{-1}$, a sample of given thickness or mosaicity will approach the case of the ideal mosaic crystal with decreasing wavelength.

Mosaicity in a real crystal is a merely formal concept for peak broadening which may be caused by various defects, e.g., dislocations. Referring to crystals containing dislocations in connection with Bragg peaks appears contradictory since eq. (12) and the subsequent discussion imply that dislocations cause Bragg peaks to disappear as their strain fields decrease proportionally to $(r')^{-1}$. However, this would hold only for a random arrangement of dislocations in a large crystal. As WILKENS [1969] has shown, a random distribution is not a very suitable model for the calculation of X-ray line

broadening or of the elastic energy. A so-called restrictedly random distribution of dislocations was used to calculate the line broadening and the peak profile more realistically (WILKENS [1970, 1975]) from the kinematical theory. The linewidth $\Delta\kappa$ is proportional to $g_{hkl}\rho_d^{1/2}\ln M_e$, where ρ_d is the dislocation density and M_e is a parameter influencing the lineshape. It is given by $M_e = \rho_d^{1/2} R_e$ where R_e is an effective outer cut-off radius, indicating the range over which a random dislocation arrangement can be admitted. Beyond this distance, long-range stresses should compensate each other (see UNGÁR *et al.* [1984] and MUGHRABI *et al.* [1986] for an application to deformed Cu single crystals). For peak profiles with neutrons, see SWALLOWE and OSBORN [1992], who studied *in situ* the effects of elastic and plastic deformation in Cu.

Rocking-curve broading, on the other hand, is only very indirectly related to dislocation densities. It rather reflects the spread of lattice rotations around the rocking axis. An analysis of rocking curves taken at different angles with respect to the tensile axis of a deformed crystal may provide useful details on the slip geometry (for [001]-Cu, see WILKENS *et al.* [1987]).

An analysis of the "static Debye–Waller" factor may be useful to determine displacements around defects (see KRIVOGLAZ [1969], DEDERICHS [1973], TRINKAUS [1975], DIETRICH and FENZL [1989a]). In a dilute alloy, ($c \ll 1$), eq. (12) simplifies to

$$M = c\sum_n \left[1 - \cos\left(2\pi\boldsymbol{\kappa}\cdot\boldsymbol{u}_n\right)\right], \tag{27}$$

or even $2M = c\sum_n \left(2\pi\boldsymbol{\kappa}\cdot\boldsymbol{u}_n\right)^2$ if $|\boldsymbol{\kappa}\cdot\boldsymbol{u}_n| \ll 1$. The $\boldsymbol{u}_n$ are displacements caused by a defect taken at the n origin, and the summation extends over all cells.

In martensite, where interstitially dissolved carbon causes a tetragonal distortion, M_{002} is expected to be different from M_{200} or M_{020}, and for polycrystals $\ln(2I_{002}/I_{200,020}) = 2M_{002} - 2M_{200,020}$. CHEN *et al.* [1980] have evaluated such (X-ray) intensity ratios for martensite formed in quenched fine-grained Fe-18 wt% Ni-(0.9/1.15)wt% C. A similar study of cubic NbH_x was performed by METZGER *et al.* [1982] who obtained $u_1 = (0.100 \pm 0.007)$ Å for H occupying tetrahedral sites, a value very close to theoretical calculations using lattice theory. In binary ordered crystals (e.g., Ni_3Al, see MORINAGA *et al.* [1988]) containing dilute solutes, the static atomic displacements in the two sublattices may be determined individually from the analysis of a large number of fundamental and superstructure peaks. For alloying elements substituting mainly for Al in Ni_3Al, it was found that there is a reverse correlation between Al and Ni displacements and that the average lattice constant is primarily controlled by the Al displacements.

In these applications, a separation of static and dynamic displacements relies on the thermal displacements known from the defect-free reference state. Neutron and X-ray results for pure Cu_3Au just above or below T_c, the critical temperature for appearance of long-range order, showed a very strong increase of the Debye–Waller exponent when T_c was approached (BARDHAN *et al.* [1977]). From the difference in the intensity ratios of fundamental and superstructure/diffuse peaks, individual mean-square displacements for Au and Cu (based on an expansion of eq. (27) for M) were calculated, and large displacements of Au near T_c were found. From a wealth of other information, including phonon dispersion curves (CHEN *et al.* [1977]) near T_c, the authors concluded that these

displacements are static and related to a high density of antiphase boundaries within large long-range ordered regions. Elastic neutron scattering (LANDER and BROWN [1985]), kinetic measurements of domain growth (LUDWIG *et al.* [1988]) and new inelastic neutron scattering studies of phonons (GAULIN *et al.* [1990]) around T_c have provided further details on this first-order phase transformation.

3.3. Diffuse scattering near Bragg peaks

We now turn to the scattering cross-section given formally by eq. (11), i.e., continuous scattering for $\boldsymbol{\kappa} \neq \boldsymbol{g}$. There is no general analytical solution, and approximations have to be introduced to obtain manageable expressions.

As pointed out in § 2.4, thermal vibrations (phonons) are always present in a crystal, and their effect on the Bragg intensity is usually contained in the Debye–Waller factor. The diffuse intensity is called thermal-diffuse scattering (TDS) and has been treated extensively (e.g., R. W. JAMES [1963], WARREN [1969], BACON [1975], WILLIS and PRYOR [1975]). TDS occurs for all values of $\boldsymbol{\kappa}$ but is particularly pronounced near reciprocal lattice points. This is also true for static displacements but there are some important differences regarding the symmetry of diffuse scattering in reciprocal space. The theory of diffuse scattering from point defects and clusters has been considered by KRIVOGLAZ [1969], DEDERICHS [1969, 1973], TRINKAUS [1972], and DIETRICH and FENZL [1989a, b]. For small defect concentrations c one obtains from eqs. (6) and (11) the diffuse scattering cross-section:

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{d}} = cN\left|f(\boldsymbol{\kappa})\right|^2, \tag{28}$$

$$\text{with} \quad F(\boldsymbol{\kappa}) = \Delta f_{\mathrm{B}} + f_{\mathrm{A}} \sum_n \exp\left(-2\pi\,\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{r}_n^{\mathrm{a}}\right)\left[\exp\left(-2\pi\,\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{u}_{0n}\right) - 1\right], \tag{29}$$

where $\boldsymbol{u}_{0n}$ is the displacement at site n due to *one* defect at the origin, $\boldsymbol{r}_n^{\mathrm{a}}$ denotes the position vector in the *average* lattice, and $\Delta f_{\mathrm{B}}, f_{\mathrm{A}}$ are the scattering lengths of defect and host atoms (one atom per unit cell assumed for simplicity), including thermal and static Debye–Waller factors. The value of Δf_{B} depends on the type of defect (e.g., DEDERICHS [1973]):

$$\Delta f_{\mathrm{B}} = \begin{cases} f_{\mathrm{I}} \exp\left(-2\pi\,\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{r}_{\mathrm{I}}\right) & \text{for interstitial impurities,} \\ f_{\mathrm{B}} - f_{\mathrm{A}} & \text{for substitutional impurities,} \\ -f_{\mathrm{A}} & \text{for vacancies} \\ f_{\mathrm{A}}\left[2\cos\left(2\pi\boldsymbol{\kappa}\cdot\boldsymbol{r}_{\mathrm{I}}\right) - 1\right] & \text{for split (self-)interstitials,} \end{cases} \tag{30}$$

where $\boldsymbol{r}_{\mathrm{I}}$ is the position of the interstitial (or of one of the split interstitials). For $\boldsymbol{\kappa}$ close to a Bragg peak ($\boldsymbol{\kappa} = \boldsymbol{g} + \boldsymbol{s}$, and $s \ll g$), but also, with a different reasoning, for any $\boldsymbol{\kappa}$ if $|2\pi\boldsymbol{\kappa}\cdot\boldsymbol{u}_{0n}| \ll 1$, eq. (29) yields:

$$F(\boldsymbol{\kappa}) = \Delta f_{\mathrm{B}} - 2\pi\,\mathrm{i} f_{\mathrm{A}} \boldsymbol{\kappa}\cdot \sum_n \boldsymbol{u}_{0n} \exp\left(-2\pi\,\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{r}_n^{\mathrm{a}}\right), \tag{31}$$

and finally, for the *Huang scattering* region ($s \ll g$, Δf_B neglected):

$$\left(\frac{d\sigma}{d\Omega}\right)_{Hd} = cNf_A^2\left|2\pi \boldsymbol{g}\cdot\boldsymbol{u}(\boldsymbol{\kappa})\right|^2, \tag{32}$$

$$\text{with } \boldsymbol{u}(\boldsymbol{\kappa}) = \sum_n u_{0n} \exp\left(-2\pi\, i\boldsymbol{\kappa}\cdot\boldsymbol{r}_n^a\right) \approx \frac{1}{V_c}\int u_o(\boldsymbol{r}) \exp\left(-2\pi\, i\boldsymbol{\kappa}\cdot\boldsymbol{r}\right) d^3\boldsymbol{r}\,. \tag{33}$$

The displacements $\boldsymbol{u}$ are generally determined by lattice statics (see, e.g., DEDERICHS *et al.* [1978] and § 4.1) relating displacements to forces exerted on atoms around the defect. For large distances from the defect, the asymptotic behaviour of the displacement is given by continuum theory,

$$\boldsymbol{u}_0(\boldsymbol{r}) \propto \frac{1}{r^2}\frac{\boldsymbol{r}}{r}. \tag{34}$$

(For isotropic materials, the proportionality factor is $\Delta V/4\pi$, where ΔV = volume change per defect in an infinite medium.) Thus, the long-range displacement field is always centrosymmetric, and the Fourier transform can be written as [$\boldsymbol{u}(\boldsymbol{\kappa})$ is periodic in reciprocal space, and only $\boldsymbol{s} = \boldsymbol{\kappa} - \boldsymbol{g}$ is relevant]:

$$\boldsymbol{u}(\boldsymbol{\kappa}) = \boldsymbol{u}(\boldsymbol{s}) = \frac{i}{sV_c}\boldsymbol{S}(\hat{\boldsymbol{s}})\cdot[\boldsymbol{P}\cdot\hat{\boldsymbol{s}}]. \tag{35}$$

Here, $\boldsymbol{S}$ is a tensor containing combinations of elastic compliances depending on the orientation of $\hat{\boldsymbol{s}}$ where $\hat{\boldsymbol{s}}$ is a unit vector in the direction of $\boldsymbol{s}$. The tensor $\boldsymbol{P}$ is the dipole force tensor of the defect (see, e.g., DEDERICHS [1973], SEEGER [1961]) with components

$$P_{jk} = P_{kj} = \sum_n p_{jn}x_{kn}. \tag{36}$$

where p_{jn} are the components of the forces $\boldsymbol{p}_n$ and x_{kn} the components of the position vectors $\boldsymbol{r}_n$. We note that the volume change caused by defects is proportional to the trace of $\boldsymbol{P}$, and the lattice parameter change $\Delta a/a$ can be obtained from

$$\frac{\Delta V}{V} = 3\frac{\Delta a}{a} = \frac{c}{V_c}\frac{Tr\boldsymbol{P}}{C_{11}+2C_{12}} \tag{37}$$

for a cubic crystal with the elastic constants C_{11}, C_{12}. Equation (37) is valid for isotropic *and* anisotropic defects, provided the latter show no preferred orientational arrangement relative to one of the cubic axes (see PEISL *et al.* [1974]).

Substituting eq. (35) into eq. (32), the Huang intensity will be:

$$\left(\frac{d\sigma}{d\Omega}\right)_{Hd} = \frac{cN}{V_c^2}f_A^2\left(\frac{g}{s}\right)^2\left|2\pi\hat{\boldsymbol{g}}\cdot\left\{\boldsymbol{S}(\hat{\boldsymbol{s}})\cdot[\boldsymbol{P}\cdot\hat{\boldsymbol{s}}]\right\}\right|^2. \tag{38}$$

We see that Huang scattering has the same general κ dependence as first-order TDS for $\boldsymbol{\kappa} \to \boldsymbol{g}$, especially inversion symmetry around $\boldsymbol{g}$, but the intensity distribution around a reciprocal lattice point is different, as the intensity depends on $\boldsymbol{g}\cdot\boldsymbol{u}(\boldsymbol{s})$. For example, an

isotropic defect with three equal, mutually perpendicular double forces P_0 in an isotropic cubic crystal has

$$u(s) = \frac{\mathrm{i}P_0 s}{s^2 V_c C_{11}}, \tag{39}$$

and it follows from eq. (32) that there is a zero-intensity plane for s perpendicular to $\kappa \approx g$. Figure 9a (from PEISL [1976]) illustrates this case. Double-force tensor, defect configuration and lines of equal intensity in reciprocal space are indicated. In fig. 9b, c, d some other cases are shown where the defect has a preferred axis ("anisotropic" defect, in an isotropic matrix). Because the defect axis can be oriented along several crystallographically equivalent directions, the scattering cross-section [eq. (32) or (39)] has to be averaged over all these orientations. As the orientation of a zero-intensity plane (nodal surface) is determined by the axis of the defect, nodal planes will then be reduced to a nodal line or disappear entirely, as shown in fig. 10. The isointensity surfaces in anisotropic crystals look deformed and rotated but additional nodal planes can also appear (see DEDERICHS [1973]).

As the Huang-scattering cross-section is proportional to the square of the dipole tensor [eq. (38)], interstitials will scatter considerably more than vacancies (e.g., for Cu, the ratio exceeds 20). A Huang-scattering study of a sample containing Frenkel pairs will therefore yield information predominantly related to interstitials. In a classical series of experiments on electron-irradiated Al and Cu single crystals, EHRHART *et al.* [1974] (see also EHRHART and SCHILLING [1973], EHRHART and SCHLAGHECK [1974]) have analyzed the Huang scattering from as-irradiated and step-annealed samples in terms of single defects and clusters. Clustering of defects leads to an increase of Huang scattering [proportional to the number of clustering defects, z, if the dipole strengths superpose linearly, see eq. (38)]. Interstitial clusters ($z \approx 3$) in Al irradiated with fast neutrons at low temperatures (4.6 K) have been reported by V. GUERARD *et al.* [1980] who compared Huang scattering

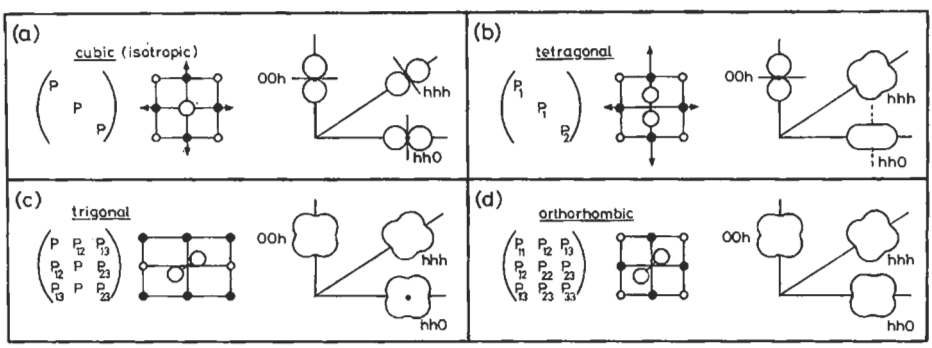


Fig. 9. Schematic contours of equal scattering intensity (right column) for various defect symmetries in a cubic lattice (central column), for anisotropic defects averaged over all crystallographically equivalent defect orientations. The column on the left indicates the form of the dipole force tensor. (After PEISL [1976].)

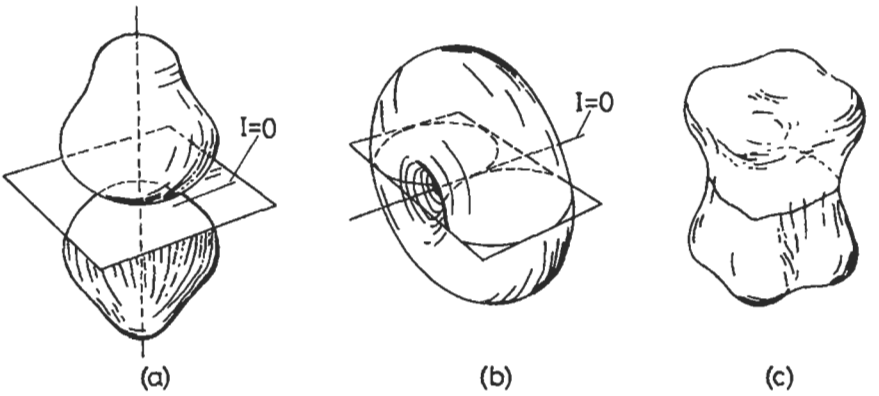


Fig. 10. Types of surfaces of equal scattering intensity in reciprocal space: (a) double-drop type with one zero-intensity surface; (b) apple-shaped type with one zero-intensity line; (c) single-bubble type. (After TRINKAUS [1972].)

from these samples with single-defect scattering after electron irradiation. The analysis yields further details on the size (radius ~ 50 Å) of displacement cascades, the number of defects per cascade (~ 200), overlap at larger irradiation doses and annealing behaviour.

Strong distortions as they may be expected for clustered point defects reduce the range of validity of the Huang approximation. This can be seen in neutron-irradiated Al and is also found when clusters form upon recovery of electron-irradiated Al and Cu (EHRHART and SCHILLING [1973], EHRHART and SCHLAGHECK [1974]). Figure 11 shows results for Al. The s^{-4} dependence at the larger s is expected from the theory for strongly distorting defects (Stokes–Wilson approximation, see, e.g., DEDERICHS [1973]). This type of scattering occurs if closely spaced scattering centres introduce large phase shifts [i.e., $2M$ becomes very large, see eqs. (10) and (12)]. The so-called asymptotic distortion scattering is given by ($\boldsymbol{g} \approx \boldsymbol{\kappa}$):

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{Ad}} = \frac{cN}{V_c^2} f_A^2 \frac{|\boldsymbol{P}\|\boldsymbol{g}|}{s^4} \phi(\hat{s}), \tag{40}$$

and depends linearly on the modulus of the double–force tensor $\boldsymbol{P}$ and on g, in contrast to the square dependence for Huang scattering [eq. (38)].

The comparable magnitudes of TDS and Huang scattering by point defects suggest that for X-rays one should work at low temperatures. This is not always possible, as the defects may not be present in the state of interest. METZGER *et al.* [1976] and METZGER and PEISL [1978] have shown that Huang scattering from H in Nb and Ta can be studied at room temperature (where the solubility in the bcc α-phase is sufficiently high), and that after appropriate corrections, especially for the static Debye–Waller factor, the results can be interpreted in terms of single-defect scattering up to surprisingly high concentrations of H in Ta (11%, METZGER and PEISL [1978]). The long-range stress field

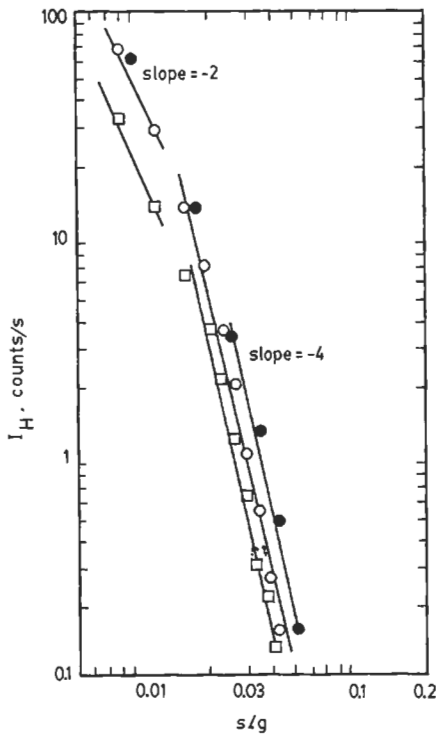


Fig. 11. Symmetrical part (average for $+s$ and $-s$) of the scattering intensity of electron-irradiated Al after annealing at 130 K, measured near the 400 reflection. Open circles: molar fraction of Frenkel pairs 5×10^{-4}, κ parallel to ⟨100⟩; squares: molar fraction 3×10^{-4}, κ parallel to ⟨100⟩; solid circles: 5×10^{-4}, κ parallel to ⟨110⟩. (After EHRHART and SCHILLING [1973].)

shows cubic symmetry in both systems (with a possible anisotropy of less than 7%), i.e., the distortion caused by H atoms occupying tetrahedral sites (as known from channeling experiments, see CARSTANJEN and SIZMANN [1972]) is isotropic for large distances. Although a Kanzaki-force model involving two forces can be adjusted to account for the cubic symmetry of the long-range distortions (see BAUER [1979]), more recent results of diffuse elastic neutron scattering between Bragg peaks have led to a modified model (DOSCH *et al.* [1992], see § 4).

For otherwise perfect crystals, one remaining "defect" is the surface. Even if this is very smooth, it will cause diffuse intensity at each Bragg peak along a line perpendicular

to the surface (ANDREWS and COWLEY [1985], ROBINSON [1986]). This so-called truncation-rod scattering is now more easily accessible with the use of synchrotron radiation. Surface roughness, misorientations and other surface properties will be reflected in changes of this scattering. For example, atomic displacements at the surface of (111)-Si wafers have been revealed using this technique (KASHIHARA *et al.* [1987]). Truncation-rod scattering for neutrons has also been seen from the same type of surface (AL USTA *et al.* [1990]).

4. *Between Bragg peaks*

If we now extend the range of $\boldsymbol{\kappa}$ and look for intensity at some distance from the Bragg peaks of a given structure, we will find diffuse intensity from TDS and diffuse scattering from static displacements if defects are present. This scattering (occasionally called structural diffuse scattering or also "Zwischenreflex" scattering) is particularly sensitive to the symmetry of displacements in the immediate vicinity of defects and can help to confirm (or eliminate) certain defect models. In addition to this "positional disorder", alloys will also exhibit "compositional disorder". Whereas the Bragg peaks are controlled simply by the average structure factor of the unit cell, diffuse scattering will reflect deviations from a random arrangement (short-range order or clustering). Basic interactions in alloys can thus be studied. Finally, new Bragg peaks may appear when (chemical or magnetic) order or (reconstructive or displacive) phase transformations occur.

The experimental requirements for diffuse scattering between Bragg peaks are somewhat different from those in § 3. There, very good resolution near Bragg peaks was desirable. Now, as small intensities are to be monitored over a wide $\boldsymbol{\kappa}$ range, resolution demands may often be relaxed in favour of higher intensity, and simultaneous observations at various values of $\boldsymbol{\kappa}$ are advantageous. Multidetector arrangements, position-sensitive detectors or films (these last for a broad survey of diffuse scattering, and still as an important tool in the study of phase transformations) are therefore used wherever possible.

4.1. Displacement scattering

The diffuse scattering for $\boldsymbol{\kappa}$ not too close to a Bragg peak depends on the displacements in the immediate vicinity of the defect. If we remain in the dilute limit $c \ll 1$, eq. (28) is still valid with $F(\boldsymbol{\kappa})$ according to eq. (29). But eq. (31) may not be valid for all displacements, and we write

$$F(\boldsymbol{\kappa}) = \Delta f_B + f_A \sum_n \exp\left(-2\pi\, i\boldsymbol{\kappa}\cdot\boldsymbol{r}_n^a\right)\left[\exp\left(-2\pi\, i\boldsymbol{\kappa}\cdot\boldsymbol{u}_{0n}\right) - 1 + 2\pi\, i\boldsymbol{\kappa}\cdot\boldsymbol{u}_{0n}\right] - 2\pi\, i f_A \boldsymbol{\kappa}\cdot\boldsymbol{u}(\boldsymbol{\kappa}) \tag{41}$$

Thus, if there are any $\boldsymbol{u}_{0n}$, too large to permit an expansion of the exponential, they are taken into account by the second term of eq. (41). If $|\boldsymbol{\kappa}\cdot\boldsymbol{u}_{0n}| \ll 1$, eq. (41) reduces to eq. (31). The $\boldsymbol{u}_{0n}$ may be obtained from lattice statics (see, e.g., DEDERICHS *et al.* [1978]). A set of (virtual) forces (called *Matsubara–Kanzaki forces*) is assumed to be applied to the

neighbouring atoms of a defect in order to produce the displacements u_{0n} within the harmonic approximation, i.e., with the dynamical matrix of the ideal, defect-free crystal. In reciprocal space, we have:

$$f(\kappa) = \phi(\kappa) \cdot u(\kappa), \tag{42}$$

where $f(\kappa)$ is the Fourier transform of the virtual forces f_{0n}, and $\phi(\kappa)$, the dynamical matrix (inverse of the Fourier transform of the lattice Green's function), contains combinations of squares of phonon frequencies. In symmetry directions, $\phi(\kappa)$ can be expressed in terms of the three phonon branch frequencies ν_L^2 and ν_{T1}^2, ν_{T2}^2. The transition to continuum theory ($\nu^2 \propto q^2$) has already been mentioned in § 3.3. Although the actual displacements may sometimes be more interesting, the concept of virtual forces has the advantage of giving a more rapidly converging method to calculate the scattering of (model) defects as the relevant forces may be restricted to the immediate surroundings of the defect (see SCHMATZ [1973] and BAUER [1979] for further discussions). Figure 12 illustrates the sensitivity of diffuse scattering between Bragg peaks to the details of the displacement field near the defect. Experiments with X-rays have first been reported for electron-irradiated Al (HAUBOLD [1975]) and subsequently for Cu (HAUBOLD and MARTINSEN [1978]). These experiments clearly confirmed the dumbbell configuration in ⟨100⟩ direction (see ch. 18, § 3.3 for a further discussion of the results).

Early displacement scattering studies with neutrons have been reviewed by BAUER [1979]. Although the incident beam intensities are much smaller than with X-rays, experiments are feasible, and the energy analysis usually achieved with a chopper allows one to extend diffuse scattering measurements to higher temperatures. Several dilute alloys with low incoherent scattering cross-sections have been studied (Al, Nb, Pb are favourable solvents). Figure 13 shows diffuse elastic scattering contours for an Al-0.8 at% Cu single crystal (see BAUER [1979]). The scattering at 300 K shows signs of precipitation, the results at 800 K are indicative of a random distribution of defects. Figure 14 shows the scattering cross-sections along the three main symmetry directions, compared with calculated values from a model with radial virtual forces acting on the nearest and next-nearest neighbours of the substitutional Cu solute. As the strength of the dipole force tensor is given by the lattice parameter change and c is known, no adjustable parameters remain in a nearest-neighbour model, and the forces f_0 must be -4.22×10^{-10} N, directed towards the Cu atom (lattice contraction). For this and any other model with centrosymmetric forces (inversion symmetry), the distortion scattering must be zero halfway between the Bragg peaks. For a centrosymmetric defect, one thus obtains just a diffuse scattering contribution from Δf_B [see eq. (41)] at these particular positions. Figure 14 shows that for Cu in Al, inversion symmetry is fully compatible with the results. The nearest-neighbour forces f_0 alone yield quite a satisfactory fit. Detailed analysis of the slope near $\kappa = g/2$ led to the best fit indicated in the figure. Different signs for the radial forces f_1 and f_2 are not in conflict with the oscillatory behaviour of the effective interaction potential in Al (see ch. 2).

Pseudopotential theory has been applied to Mg in Al to calculate diffuse elastic scattering cross-sections, and neutron scattering results obtained at 720 K (Al-3.2 at% Mg) along [110] compare favourably (WERNER *et al.* [1978]). Simple virtual-force

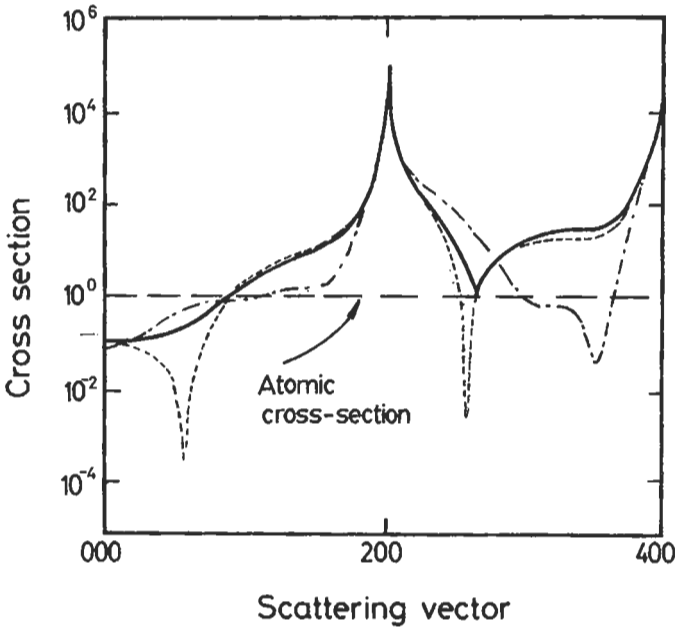


Fig. 12. Calculated cross-sections for diffuse scattering by self-interstitials in Al. Solid line: ⟨100⟩ dumbbell, dashed-dotted line: tetrahedral, dotted line: octahedral configuration. (After HAUBOLD [1975].)

models fail in this case, as they do in Al–Zn and, even more so, in Al–Mn (see BAUER [1979]). Another complicated case is Pb–Bi. Alloy crystals with 2 and 4% Bi (and pure Pb) were studied with neutrons (see BAUER [1979]). An unexpected peak at $\kappa = g_{220}/2$ was found, which cannot be attributed to short-range order because a similar peak would then be required at a distance g_{111} from $g_{220}/2$, i.e., at $g_{200}/2$. If displacements are responsible, the inversion symmetry one intuitively expects for a substitutional solute has to be perturbed. The suggested model assumes equal outward forces f_1 on all nearest neighbours and three extra forces f_e pulling three neighbours in one plane closer to the Bi atom (an extension of the $8-N$ rule).

In the high-temperature bcc phase of Zr–Co and Zr–Nb, strong diffuse scattering was found in the $(1\bar{1}0)$ and (001) planes of reciprocal space (HEIMING *et al.* [1991b]). A centrosymmetric model with Kanzaki forces on nearest neighbours (inward) and next nearest neighbours (outward) can be fitted to the results. Phonon modes (see § 6) indicate a tendency toward the formation of the ω-phase which appears at lower temperatures by a periodic collapse of {111} planes along ⟨111⟩ directions (see, e.g., CURRAT and PYNN [1979]). HEIMING *et al.* [1991b] conclude that the observed diffuse scattering is due to a superposition of all displacements for which low-energy phonons exist.

In Nb single crystals containing small quantities of nitrogen (0.15 and 1.4 at%, see ROWE and MAGERL [1980], DOSCH and PEISL [1985]), elastic neutron scattering showed diffuse peaks reminiscent of the ω-phase. Assuming octahedral sites for N in Nb, DOSCH

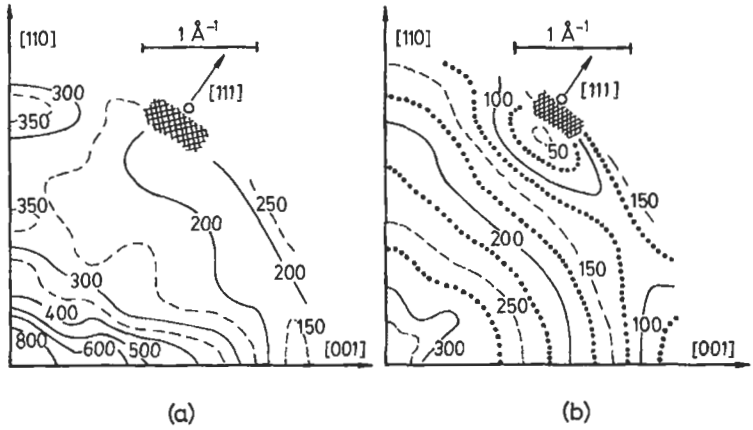


Fig. 13. Lines of equal elastic diffuse scattering cross-section (in mbarns/sr per Cu atom) for Al-0.8 at% Cu in the $(1\bar{1}0)$ plane of the reciprocal lattice for (a) $T = 300$ K and (b) $T = 800$ K. Data in the cross-hatched regions have been omitted. Scale: 1 $Å^{-1} = 2\pi\kappa$. (After BAUER [1979].)

and PEISL [1985] showed that their measurements could not be rationalized with a Kanzaki force model including only two radial force constants (nearest and next-nearest neighbours). As the individual elements of the dipole force tensor are known from independent diffuse X-ray measurements in the Huang regime (SCHUBERT *et al.* [1984]), the values of the forces are again fixed. Using a set of three forces (for the three neighbouring shells) compatible with the dipole force tensor, the continuous curves shown in fig. 15 were obtained. The displacements of the nearest Nb neighbours surrounding N remain large (~0.5 Å) in the model involving three forces and are compatible with other experimental evidence, the next-nearest neighbours are displaced toward the interstitial (~ −0.2 Å). These displacements lead to a local approach of neighbouring {111} planes towards a small ω-like region. Similar results have been obtained for Nb–O and interpreted quite similarly (DOSCH *et al.* [1986]). Although a Kanzaki force model involving two forces roughly reproduces the measured diffuse scattering of Nb–D (deuterium is used to reduce the incoherent scattering which is much higher for protons) at room temperature and the forces can be adjusted to account for the (almost) cubic symmetry of the long-range stress field (see BAUER [1979]), more recent work by DOSCH and PEISL [1986] and DOSCH *et al.* [1987, 1992] shows that the fit is not perfect (see fig. 16) and can be considerably improved by a new model which adds a finite probability for D to be in a mobile state, appearing delocalized over three tetrahedral sites. That the jump rate of the diffusing defect (D) is much larger than previously assumed, and increasingly so at higher temperatures, is revealed by quasielastic scattering (DOSCH *et al.* [1992]). The mobile state is modelled by combining scattering contributions from a

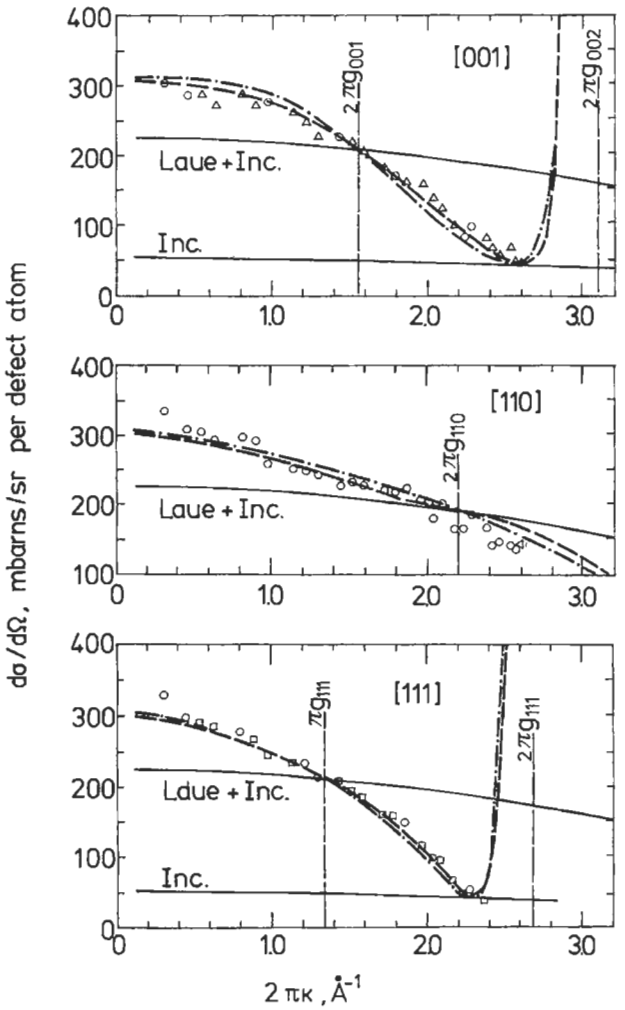


Fig. 14. Comparison of measured cross-sections for elastic diffuse scattering from Al-0.8 at% Cu at 800 K with the results of model calculations. The solid lines marked "Laue + Inc." indicate the expected behaviour without static displacements (but including the Debye–Waller factor). Dashed lines: model with forces on nearest neighbours only; dashed dotted lines: "best fit" with forces f_1 on nearest and f_2 on next-nearest neighbours, $f_2/f_1 = -0{,}2^{+0.2}_{-0.1}$. The symbols represent experimental data. (After BAUER [1979].)

modified environment around the three neighbouring tetrahedral sites involved, a nonvanishing occupation probability of the triangular site within this same environment, and an adjustment for the high mobility of the defect (see fig. 16).

References: p. 1188.

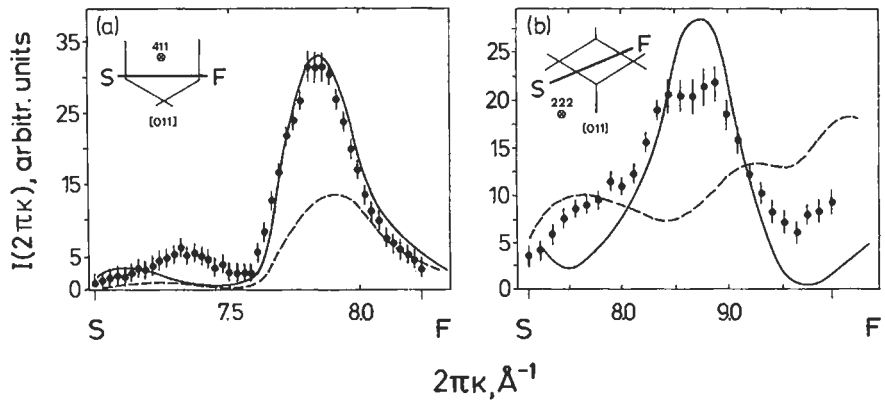


Fig. 15. Diffuse elastic scattering for a Nb-1.4 at% N single crystal at room temperature for two different κ-scans (along the lines in reciprocal space indicated by SF in the insets, crystal axis parallel [011]). Experimental results are compared with a model using two forces (dashed lines) and three forces (solid lines). (After DOSCH and PEISL [1985].)

4.2. Short-range order

So far, we have considered dilute solutions of defects and assumed a random distribution. This approach yields simple scattering laws and is sufficient for many cases. However, in concentrated solid solutions, the notion of single defects is difficult to maintain, and additional scattering effects will occur if there is any deviation from a random occupation of lattice sites by the different scattering centres, i.e., short-range order (SRO) occurs.

Much work has been devoted to binary substitutional alloys. If we neglect positional disorder, eq. (28) for diffuse scattering (i.e., $\boldsymbol{\kappa} \neq \boldsymbol{g}$) can be replaced by (KRIVOGLAZ [1969])

$$\left(\frac{d\sigma}{d\Omega}\right)_{SRO} = N|c(\boldsymbol{\kappa})|^2 (f_B - f_A)^2 \tag{43}$$

for an A–B alloy of concentration $c = c_B$. (We consider alloys with one sublattice only. More general expressions for several sublattices, important for SRO in intermetallic compounds, hydrogen in metals, non-stoichiometric compounds etc., are given, e.g., by HAYAKAWA and COHEN [1975] and BAUER [1979]. For some further discussions including ternary alloys see COHEN [1970] and DE FONTAINE [1971, 1979].) In eq. (43), Debye–Waller factors are again included in the scattering lengths f_A, f_B. The term $|c(\boldsymbol{\kappa})|^2$ contains the Fourier transform of the compositional fluctuations (of component B),

$$c(\boldsymbol{\kappa}) = \frac{1}{\sqrt{N}} \sum_n (c_{Bn} - c_B) \exp\left(-2\pi i \boldsymbol{\kappa} \cdot \boldsymbol{r}_n^a\right), \tag{44}$$

and c_{Bn} is defined as in eq. (4) to be one if a B atom is present at site n and zero otherwise. In the limit of small concentrations, $N|c(\boldsymbol{\kappa})|^2 \to Nc$ as in eq. (28). For a

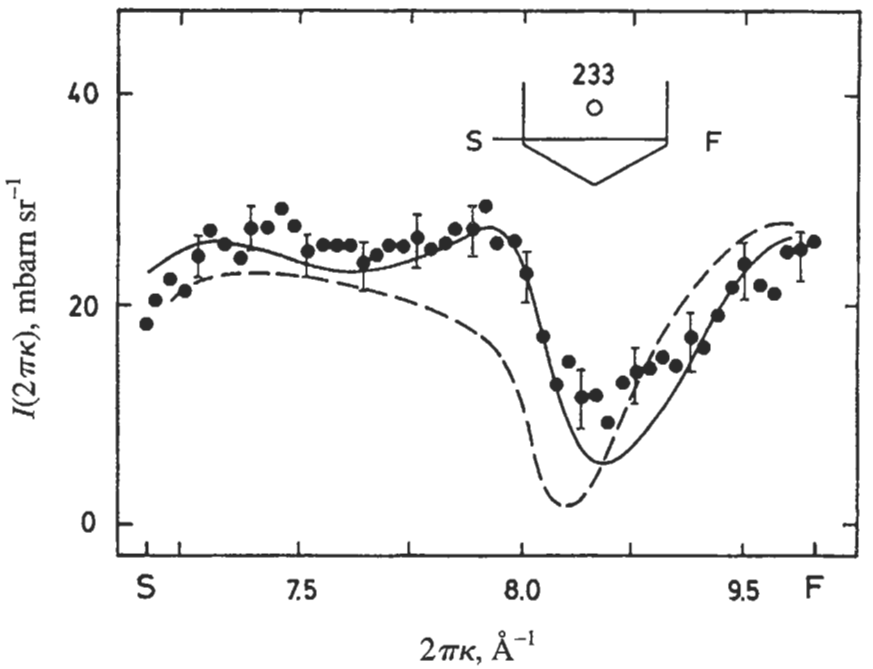


Fig. 16. Coherent diffuse (quasi-)elastic neutron scattering $I(2\pi\kappa)$ (energy resolution 3.8 meV) from Nb-1.7 at% D at room temperature (DOSCH and PEISL [1986]) along a line S–F in reciprocal space as shown in the inset (crystal axis along [011]). The dashed line is the expected scattering from a static cubic defect with the correct long-range stress field. The solid line is the best fit for a delocalized defect (DOSCH *et al.* [1992]).

random distribution of B atoms of arbitrary concentration, $N|c(\boldsymbol{\kappa})|^2 \to Nc(1-c)$, and eq. (43) then yields the well-known *monotonic Laue scattering* $\propto c(1-c)(f_B - f_A)^2$. For the theory of alloys, deviations from a random distribution are of course more interesting. If displacements were negligible, $|c(\boldsymbol{\kappa})|^2$ could be obtained directly from diffuse scattering measurements, but quite frequently, displacements have to be taken into account.

Figure 17 shows a classical experimental result (MOZER *et al.* [1968]) on a polycrystalline Cu-47.5 at% Ni sample studied with neutrons. A unique possibility offered by neutrons has been exploited here. Some isotopes (see § 2.2) have a negative scattering length, and for one (or more, if one changes the isotopic composition of the alloying partners) chemical composition of a binary alloy, it is thus possible to obtain a "null-matrix alloy" with an average scattering length $\bar{f}=0$. This eliminates Bragg peaks that normally, in the case of polycrystals, severely limit the analysis of diffuse scattering. The alloy used by MOZER *et al.* [1968] contained natural Cu and Ni enriched to 99% in ^{62}Ni. In fig. 17, only very small Bragg intensities are visible, but a very pronounced modulation of the diffuse scattering indicates a non-random distribution of the alloying elements on the fcc lattice sites.

It is common to describe SRO in terms of the Warren–Cowley parameters α_{0n}

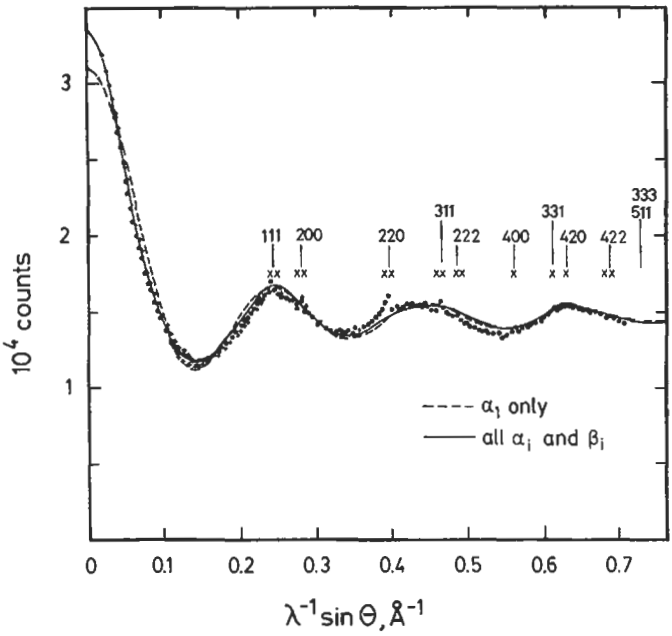


Fig. 17. Diffuse neutron scattering from polycrystalline Cu-47.5 at% ^{62}Ni, furnace-cooled from 1021°C. The data indicated by crosses were taken with a neutron wavelength $\lambda = 1.951$ Å, those indicated by solid circles with $\lambda = 1.024$ Å. The Bragg positions are marked. The dashed curve is a fit with one SRO parameter only, the solid curve is the "best fit" described in the text. (After MOZER *et al.* [1968].)

(COWLEY [1950], see e.g., SCHWARTZ and COHEN [1987], and WELBERRY and BUTLER [1994] for a survey of other descriptions). They are related to $|c(\boldsymbol{\kappa})|^2$ by:

$$|c(\boldsymbol{\kappa})|^2 = c(1-c)\alpha(\boldsymbol{\kappa}), \tag{45}$$

$$\text{where} \quad \alpha(\boldsymbol{\kappa}) = \sum_n \alpha_{0n} \exp(-2\pi\, i\boldsymbol{\kappa}\cdot\mathbf{r}_{0n}), \tag{46}$$

and the α_{0n} are defined using "conditional probabilities" P_{0n}^{BB} (indicating the probability that a B atom is at position n if an A atom is at the origin), P_{0n}^{BB} etc.

$$\alpha_{0n} = \frac{P_{0n}^{BB} - c}{1-c} = \frac{c - P_{0n}^{AB}}{c}. \tag{47}$$

Obviously, $\alpha_{00} = 1$, independent of the state of SRO, and for a random distribution $\alpha_{0n} = 0$ ($n \neq 0$). In cubic structures, the α_{0n} are usually labeled per coordination shell i or with indices *lmn* referring to all combinations of coordinates for positions within one shell.

For a polycrystalline sample, the diffuse SRO scattering from eqs. (43)–(46), averaged over all orientations (e.g., WARREN [1969]), is

$$\left(\frac{d\sigma}{d\Omega}\right)_{SRO} = Nc(1-c)\left(f_B - f_A\right)^2 \sum_i Z_i \alpha_i \frac{\sin 2\pi \boldsymbol{\kappa} \boldsymbol{r}_i}{2\pi \boldsymbol{\kappa} \boldsymbol{r}_i}, \tag{48}$$

where Z_i is the number of neighbours in the ith shell with radius r_i (the sum includes $i = 0$). The dashed curve in fig. 17 is a fit according to eq. (48) with only one α_i adjusted ($\alpha_1 = 0.131$). This fit already looks quite reasonable and indicates that short-range clustering is present in this alloy. According to eq. (47), a positive SRO parameter means more BB pairs than expected from the average concentration. This "negative SRO" leads to a reduction of diffuse scattering (relative to the Laue scattering) between Bragg positions, and an enhancement near Bragg positions (including 000, i.e., $\kappa = 0$). The opposite is true for normal ("positive") SRO where $\alpha_1 < 0$.

The deviations from the simple α_i fit visible in fig. 17 cannot be explained by contributions of more α_i terms alone (except for κ near zero). As discussed in § 3, following eq. (38), displacements from ideal lattice sites introduce Huang scattering (symmetric around Bragg peaks and $\propto s^{-2}$) and "size effect" scattering (antisymmetric around reciprocal lattice points). Both contributions have to be considered in concentrated alloys, and the treatment, even for the approximation of small displacements, becomes rather involved (see, e.g., WARREN [1969] and SCHWARTZ and COHEN [1987]). In the approximation $\exp[-2\pi\, i\boldsymbol{\kappa}\cdot(\boldsymbol{u}_n - \boldsymbol{u}_{n'})] \approx 1 - 2\pi\, i\boldsymbol{\kappa}\cdot(\boldsymbol{u}_n - \boldsymbol{u}_{n'})$, neglecting the Huang term [(32), $\boldsymbol{\kappa}$ is no longer restricted to the vicinity of a Bragg peak], the diffuse scattering can be written (see WARREN [1969]):

$$\left(\frac{d\sigma}{d\Omega}\right)_d = Nc(1-c)\left(f_B - f_A\right)^2 \sum_{i,j} \left(\alpha_i - 2\pi\, i\beta_i \boldsymbol{\kappa}\cdot\boldsymbol{r}_{ij}\right) \exp\left(-2\pi\, i\boldsymbol{\kappa}\cdot\boldsymbol{r}_{ij}\right), \tag{49}$$

where i indicates the shell and j the position in the ith shell, and the "size effect coefficients" are given by

$$\beta_i = \frac{f_A}{f_A - f_B}\left(\frac{1-c}{c} + \alpha_i\right)\varepsilon^i_{AA} - \frac{f_B}{f_A - f_B}\left(\frac{c}{1-c} + \alpha_i\right)\varepsilon^i_{BB}, \tag{50}$$

where the $\varepsilon^i_{\mu\nu}$ are defined as the average relative deviations of the distance between a μ atom at the origin and a ν atom in the ith shell,

$$r_{\mu\nu,i} = r_i^a\left(1 + \varepsilon^i_{\mu\nu}\right). \tag{51}$$

Equation (49), averaged for polycrystals, was used to fit the data in fig. 17 (MOZER *et al.* [1968]). The "best fit", shown by the continuous curve, was obtained with nine adjustable α_i and size effect coefficients for five shells. In this case, $\alpha_1 = 0.121$ and $\alpha_2 = -0.008$, indicating a preference for unlike second-nearest neighbours.

References: p. 1188.

In a more extensive study, VRIJEN and RADELAAR [1978] investigated Cu–Ni alloys containing 20–80% Ni and were able to quench from thermal equilibrium at various temperatures prior to the neutron scattering measurements at room temperature.

Linear size effect corrections, frequently limited to nearest neighbours, were also used to separate SRO and size effect scattering from single crystals, exploiting the different symmetry of the α_i- and β_i-terms in eq. (49) (see SPARKS and BORIE [1966], MOSS and CLAPP [1968] and SCHMATZ [1973] for surveys of early experimental results), until BORIE and SPARKS [1971] and GRAGG and COHEN [1971] extended the evaluation to include size-effect contributions up to second order, i.e., starting from

$$\exp\left[-2\pi\,\mathrm{i}\boldsymbol{\kappa}\bullet\left(\boldsymbol{u}_n-\boldsymbol{u}_{n'}\right)\right]\approx 1-2\pi\,\mathrm{i}\boldsymbol{\kappa}\bullet\left(\boldsymbol{u}_n-\boldsymbol{u}_{n'}\right)-2\pi^2\left[\boldsymbol{\kappa}\bullet\left(\boldsymbol{u}_n-\boldsymbol{u}_{n'}\right)\right]^2. \tag{52}$$

The "Borie–Sparks approach" is fully described in textbooks (e.g., SCHWARTZ and COHEN [1987] and COWLEY [1981]). For cubic crystals, the total diffuse intensity, divided by $Nc(1-c)(f_B-f_A)^2$ (the result is then given in "Laue units"), can be written in this approximation as (see SCHWARTZ and COHEN [1987], a corresponding treatment for hexagonal crystals is given by KHANNA [1989])

$$\begin{aligned} I_d = {} & I_\alpha + h_1Q_x + h_2Q_y + h_3Q_z + h_1^2R_x + h_2^2R_y + h_3^2R_z + h_1h_2S_{xy} + h_2h_3S_{yz} \\ & + h_3h_1S_{zx}\,, \end{aligned} \tag{53}$$

where I_α is the short-range order term,

$$I_\alpha = \sum_{l,m,n}\alpha_{lmn}\exp\left[-2\pi\,\mathrm{i}\left(h_1l+h_2m+h_3n\right)\right] \tag{54}$$

(with l,m,n from $-\infty$ to $+\infty$), h_1,h_2,h_3 are the components of the scattering vector in units of the reciprocal lattice vectors $\boldsymbol{b}_i$ [see eq. (A.17), ch. 11] and the l,m,n coordinates are measured in units of the lattice vectors $\boldsymbol{a}_i$. The functions Q are the size effect terms containing corresponding components of the displacements β_{lmn} [see eqs. (50) and (51)]. Similarly, the R and S functions contain pair probabilities and squares and products of displacements. All functions Q, R and S also contain ratios of atomic form factors, $f_A/(f_B-f_A)$, $f_B/(f_B-f_A)$, and only if these are constant, Q, R and S are periodic functions in reciprocal space. Figure 18 shows schematically the three contributions along $[h_100]$. Using the symmetry properties of Q, R and S, it is now possible to construct a minimum volume in reciprocal space where the diffuse intensity must be measured in order to form appropriate sums and differences for certain sets of reciprocal lattice points that will yield the various terms of eq. (53) separately (see SCHWARTZ and COHEN [1987]). Figure 19 shows the minimum volume in reciprocal space for fcc crystals. It corresponds to one eighth of a unit cell in reciprocal space and has been placed in a region accessible in reflection geometry. The method was first applied to diffuse X-ray scattering from a clustering system (Al-5 at% Ag) by GRAGG and COHEN [1971].

Apart from the fact that the linear combination of the different $\varepsilon^i_{\mu\nu}$ cannot be separated, the assumption of constant ratios of the atomic form factors is the main limitation of the Borie–Sparks approach if X-rays are used (see TIBALLS [1975] and GEORGOPOULOS and COHEN [1977a, 1979]), whereas in this respect, it is ideally suited

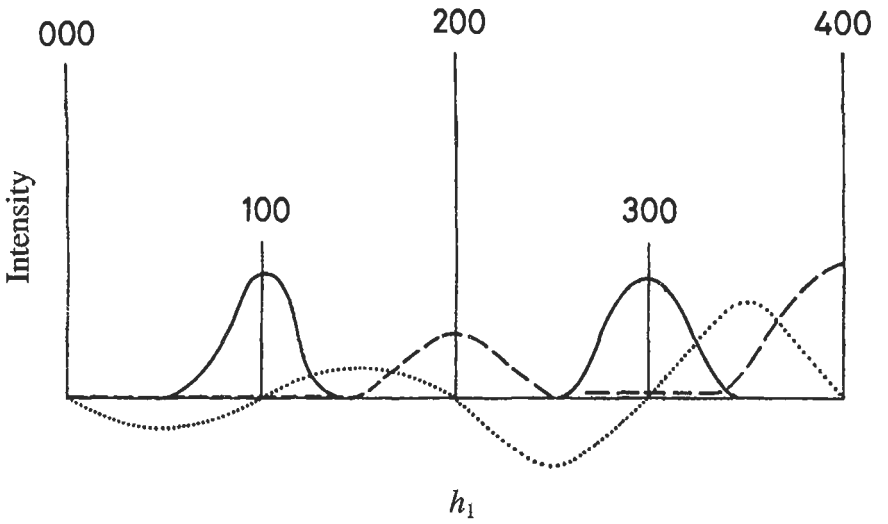


Fig. 18. Schematic of the diffuse intensity components according to eq. (53) (after SCHWARTZ and COHEN [1987]). The intensity scale is not the same for the three components. Continuous line: I_α, dashed line: $h_1 Q_x$, dotted line: $h_1^2 R_x$.

for neutron scattering. To exploit the substantial variation of the atomic form factors across the experimental κ range for the X-ray case, TIBALLS [1975] has proposed to write the functions Q, R and S of the Borie–Sparks approach as products of the κ-dependent ratios of form factors and new functions Q', R' and S' that are strictly periodic in reciprocal space. There are now 25 Fourier series instead of 10 in the Borie–Sparks method. GEORGOPOULOS and COHEN [1977a, 1979] have developed a separation method based on this approach. The procedure is considerably more laborious than for the Borie–Sparks method. From a detailed error analysis the authors conclude that the extra effort may be worthwile, especially for clustering systems. Apart from more reliable values for α_i, the method also yields individual (AA, BB, AB) values for the displacement terms. With the availability of continuously tunable X-ray wavelengths at synchrotron radiation sources, the anomalous scattering variation of the atomic form factors offers additional advantages since with a proper choice of f_A, f_B, the thermal diffuse background may be removed and a reliable separation of short-range order and individual displacement terms is achieved (RAMESH and RAMESESHAN [1971], ICE *et al.* [1992], REINHARD *et al.* [1992], SCHÖNFELD *et al.* [1994]).

From a given set of SRO parameters, the local atomic arrangement in SRO structures can be simulated on a computer by rearranging many thousands of atoms on a given lattice until sufficient agreement is obtained with the SRO parameters (GEHLEN and COHEN [1965], GRAGG *et al.* [1971]). The simulated structure can then be analyzed in terms of specific atomic arrangements. It is not yet possible to completely simulate local atomic arrangements and displacements, though first attempts have been reported (KYOBU *et al.* [1994]). If the set of SRO parameters belongs to an equilibrium state, they may be

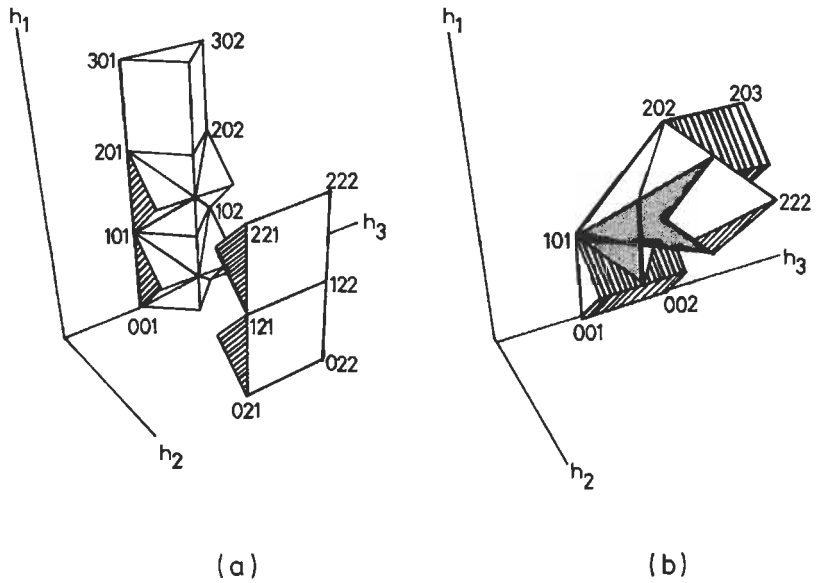


Fig. 19. (a) Typical volumes in reciprocal space required to separate the diffuse intensity from a fcc crystal into components due to local order and due to atomic displacements (to second order in the displacements). (b) As (a), after symmetry of intensity in reciprocal space across the planes $h_i = h_j$ has been employed. (After HAYAKAWA *et al.* [1975].)

used to calculate basic alloy parameters (i.e., effective pair interaction potentials V_{lmn}). Apart from approximate analytical methods (COWLEY [1950], KRIVOGLAZ [1969] and CLAPP and MOSS [1966, 1968], MASANSKII *et al.* [1991], see REINHARD and MOSS [1994] for an assessment), exact numerical calculations are now feasible. The inverse Monte Carlo method proposed by GEROLD and KERN [1987] starts from an equilibrium model crystal (obtained using the Gehlen–Cohen method) and makes use of the fact that for any equilibrium quantity of the system, the average over many fluctuations (weighted according to conditions of detailed balance) of such quantity must be zero. From many (virtual) pair exchanges in the model crystal, one thus obtains a set of effective pair interaction potentials V_{lmn} (defined by the difference of $V_{lmn}^{AA} + V_{lmn}^{BB}$ and $2V_{lmn}^{AB}$). These may in turn be used in Monte Carlo simulations for other situations (see fig. 20 for a schematic diagram). An inverse cluster variation algorithm has also been introduced (SOLAL *et al.* [1988]).

A brief survey of experimental results must start with the Cu–Au system, as this is certainly the most extensively studied binary system, providing several SRO (and long-range order) effects that have been used to test theoretical ideas at various stages (see e.g., MOSS and CLAPP [1968], KRIVOGLAZ [1969], MOSS [1969], COWLEY and WILKINS [1972], MOSS and WALKER [1975], DE FONTAINE [1979]). BARDHAN and COHEN [1976] measured the diffuse scattering of Cu_3Au in a volume of reciprocal space at six temperatures ranging from 2°C above T_c for long-range order (~ 394°C) to 930°C, and used the

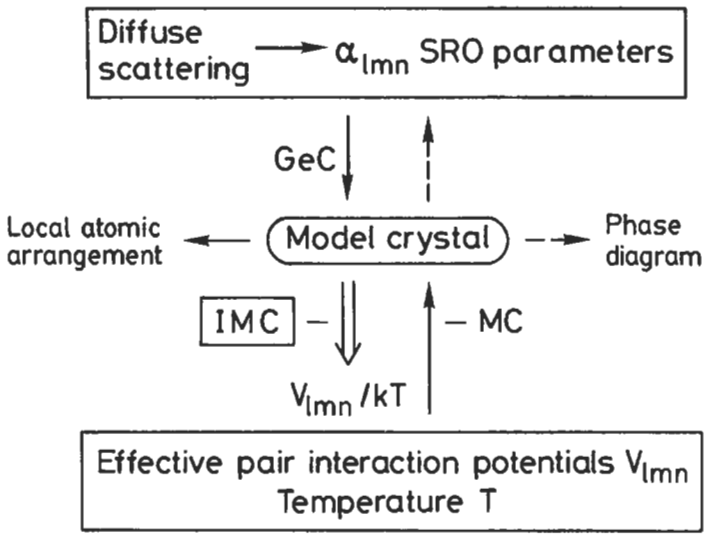


Fig. 20. Schematic of the steps involved in the inverse Monte Carlo method (IMC) and the Monte Carlo method (MC). GeC = simulation of a model crystal according to GEHLEN and COHEN [1965]. (After GEROLD and KERN [1987].)

Borie–Sparks approach to separate the corrected, absolute intensities (including detailed error analysis, also for higher-order displacement terms, see GRAGG *et al.* [1973]). More limited measurements with smaller temperature intervals indicated discontinuous changes in slope of the diffuse intensity versus temperature near specific-heat anomalies reported in the literature. The SRO scattering separated from complete measurements at selected temperatures revealed several new features, especially a change of symmetry of the separated short-range order scattering from the expected diffuse peaks at 100 and 110 ($L1_2$ is the long-range ordered structure below T_c) to a concentration of SRO scattering around $\frac{111}{222}$ near 930°C, see fig. 21a. This type of scattering would result e.g. from alternating Cu- and Au-rich {111} planes. More recently, BUTLER and COHEN [1989] have compared all previous diffuse X-ray results for Cu-25 at% Au for temperatures just above T_c and evaluated new measurements at 430°C using the Georgopoulos–Cohen approach. The SRO scattering component is shown in fig. 21b. The SRO parameters agree best with the early results of COWLEY [1950]. The simulated short-range ordered structure contains small clusters (1–10 unit cells in diameter) resembling $L1_2$ ordered domains. The individual first-order displacement terms are all negative (up to $lmn = 530$) for Au–Au and positive for Au–Cu (except for $lmn = 200$), although the lattice parameter of ordered Cu_3Au is slightly smaller than in the disordered state.

Apart from the Cu–Au system (see also BESSIÈRE *et al.* [1982] for other compositions) and the Al–Ag crystal mentioned above, the diffuse X-ray scattering of several other alloys has been analysed with the Borie–Sparks approach; Ni-20 at% Mo (CHAKRAVARTI *et al.* [1974]), Cu-(9.1-14.8) at% Al (EPPERSON *et al.* [1978a]), Cu-29.8 at% Pd (OHSHIMA *et al.* [1976], Au-18 at% Mn (FÜRNROHR *et al.* [1980]) and Au-20.3 at% Mn

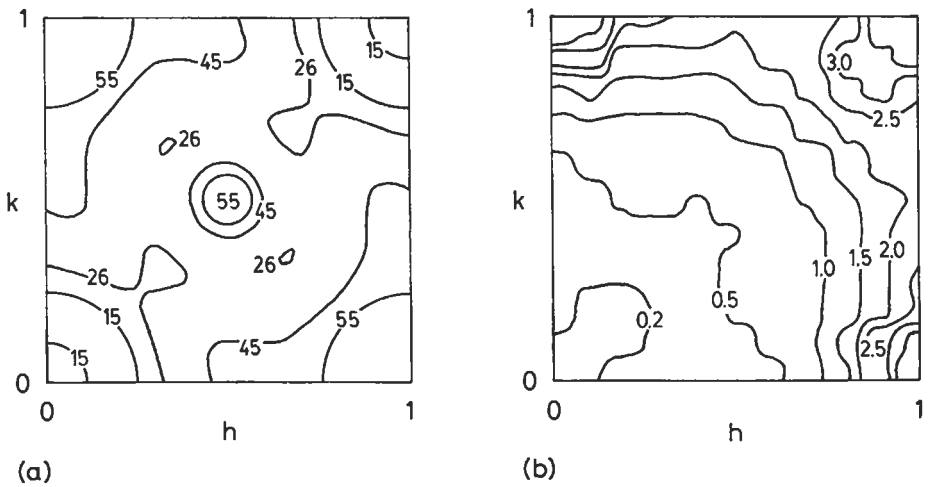


Fig. 21. Short-range order diffuse X-ray scattering in Laue units from Cu-25 at% Au at two different temperatures: (a) 930°C, (*hk* 0.5) section of reciprocal space, scale: $1 = 1.9 \times 10^{-2}$ (Borie–Sparks analysis, after BARDHAN and COHEN [1976]), (b) 430°C, (*hk*0) section (Georgopoulos–Cohen evaluation, from BUTLER and COHEN [1989]).

(SUZUKI *et al.* [1982]), Ni-12.7 at% Al (EPPERSON and FÜRNROHR [1983]) and Ni-7.3/10.5 at% Al (CHASSAGNE *et al.* [1989]), and Ag-15 at% Mg (OHSHIMA and HARADA [1986]). The X-ray study on Cu-Al (EPPERSON *et al.* [1978a], EPPERSON [1979]) was performed at room temperature, with samples quenched from different temperatures ranging from 150–900°C. The apparent degree of SRO was found to pass through a minimum near 350°C for Cu-14.8 at% Al. The values obtained for samples quenched from higher temperatures do not reflect thermal equilibrium. For some of the equilibrium states (1580 h annealing at 150°C), a computer simulation was performed using the experimental α_i and requiring agreement for the first six of them. The resulting structures were searched for specific nearest-neighbour configurations (CLAPP [1971]), and fig. 22 shows those found in all conditions investigated. In these configurations, none of the Al atoms are nearest neighbours. The configurations C9, C16 and C17 are particularly frequent in the equilibrium states. That the Borie–Sparks evaluation might suffice in some cases (CHASSAGNE *et al.* [1989]), has been shown explicitely by KLAIBER *et al.* [1987] for a non-equilibrium quenched state of Ni-9.5/9.8 at% Al by a detailed comparison of the two methods mentioned and a third one based on a least-squares fitting of the various Fourier coefficients as originally proposed by R. WILLIAMS [1972]. A similar comparison for Au-25, 48, 75 at% Ag in thermodynamic equilibrium (SCHÖNFELD *et al.* [1992a]), another system where, as in Ni–Al, linear displacement terms are sufficient, shows that differences between the Borie–Sparks and the Georgopoulos–Cohen approaches are noticeable but quite subtle as the degree of SRO is very low in the equilibrium states accessible for investigation. The diffuse SRO peaks are at 100 positions for all three alloys. Figure 23 shows the typical localized $L1_2$ features expected,

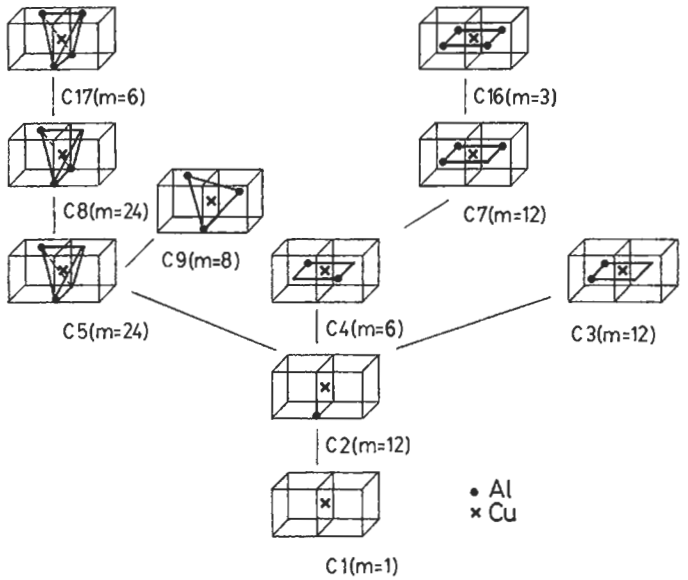


Fig. 22. The nearest-neighbour configurations of Al atoms (solid circles) around Cu atoms (crosses) found in a Cu-14.76 at% Al single crystal annealed for 1 h at 650°C and quenched to room temperature. Two adjacent unit cells are outlined. Not specially marked lattice sites are occupied by Cu atoms. The number of crystallographically equivalent variants of a configuration is indicated by *m*. (After EPPERSON *et al.* [1978a].)

but there are no indications of larger domains at this low degree of SRO ($\alpha_1 \approx -0.08$). The Clapp configurations C16 and C17 are four to seven times more frequent in all three crystals than in a random solid solution. Table 3 shows the first three effective pair interaction potentials V_{lmn} obtained for all three crystals and used for a comparison of internal energies for the $L1_2$, $D0_{22}$, $D0_{23}$ structures for 3:1 stoichiometry and the $L1_0$ and $1\frac{1}{2}0$ (see CLAPP and MOSS [1968]) structures for AuAg. The most stable states (with one exception where the differences are within 0.1 meV for the structures considered) are $L1_2$ and $L1_0$. The V_{lmn} were also used in Monte-Carlo simulations to determine critical temperatures for disordering, starting from completely ordered stoichiometric crystals. The results (T_c between 100 and 200 K) compare well with ab-initio calculations (see SCHÖNFELD *et al.* [1992a] for more details). In these alloys, the effective interactions are weak enough to employ the approximation of CLAPP and MOSS [1966]. The resulting V_{lmn} are only slightly smaller than the "exact" values. A similar X-ray study of Ni-11.2 at% Mo (SCHWANDER *et al.* [1992]) equilibrated at 500°C, using the Georgopoulos–Cohen method, shows that only the inverse Monte-Carlo method gives reliable values for V_{lmn} (V_{110}

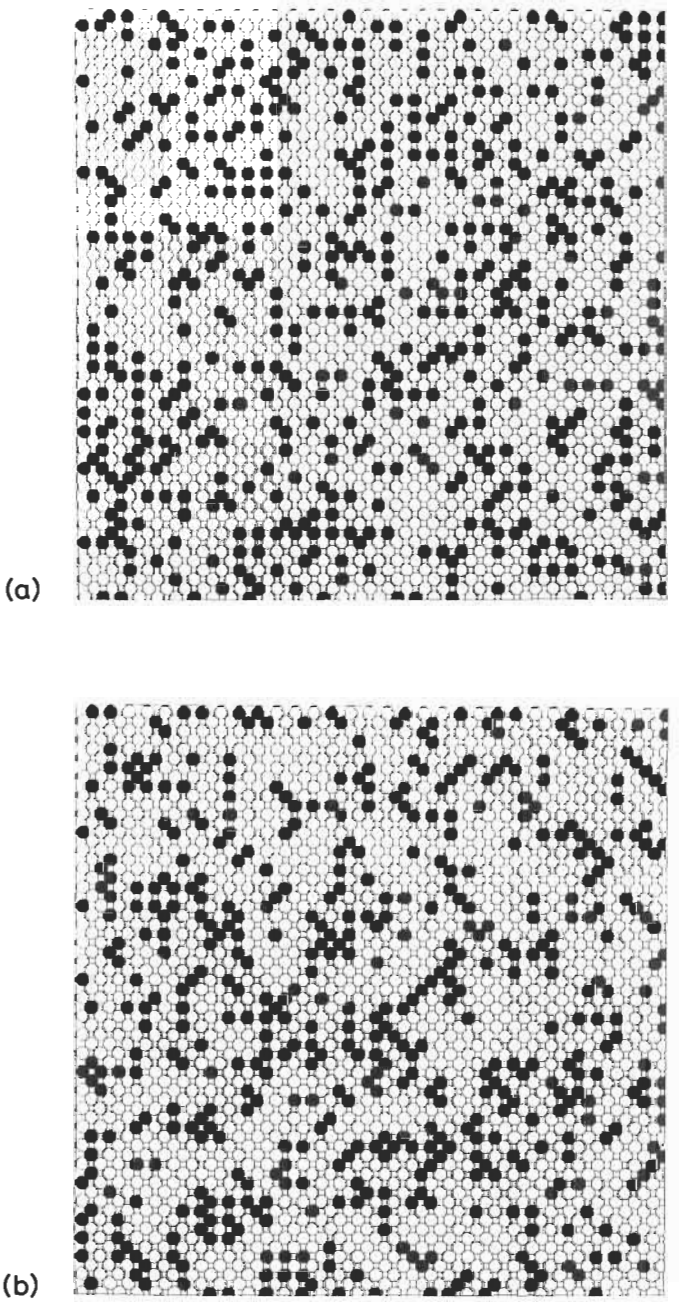


Fig. 23. A (001) plane of Au-25% Ag modelled (a) from the measured SRO parameters corresponding to equilibrium at 215°C and (b) with a random distribution of Au (○) and Ag (●) atoms. (From SCHÖNFELD *et al.* [1992a].)

Table 3

Effective pair interaction potentials V_{lmn} (in meV) for the first three shells as determined by the inverse Monte Carlo method with the short-range-order parameters α_{lmn} from a Georgopoulos–Cohen analysis (GC) or a Borie–Sparks analysis (BS) according to SCHÖNFELD *et al.* [1992]. The errors are the standard deviations resulting from the use of five independently modelled crystals in each case.

Shell index *lmn*	Au-25.0 at% Ag		Au-47.7 at% Ag		Au-74.7 at% Ag	
	GC	BS	GC	BS	GC	BS
110	15.6(2)	10.4(1)	10.7(1)	10.4(1)	16.1(2)	12.2(2)
200	2.3(2)	−1.0(1)	−0.1(1)	0.7(1)	2.2(2)	1.0(1)
211	2.0(1)	0.2(1)	0.7(1)	0.4(1)	1.0(1)	0.2(1)
Ground states	$L1_2$ $L1_0$	$L1_2$ $L1_0$	$L1_2$ $L1_0$	$L1_2$ $L1_0$	$L1_2$ $L1_0$	$D0_{22}$ $1\frac{1}{2}0$

is about 100 meV, i.e., larger than kT). The work of SCHWANDER *et al.* [1992]) also shows how the change of configurational energy per unit area, γ_{SRO}, upon successive passages of unit dislocations may be calculated for arbitrary numbers of SRO parameters and effective pair potentials. This quantity intervenes in the transition from wavy glide to planar glide in concentrated alloys (see SCHWANDER *et al.* [1992] and WOLF *et al.* [1994] for details).

The first applications of the Georgopoulos–Cohen method were in fact to systems with strong displacements, namely non-stoichiometric β'-NiAl (GEORGOPOULOS and COHEN [1977b, 1981]) and aged Al-1.7 at% Cu containing Guinier–Preston zones (AUVRAY *et al.* [1981]). The resulting α_{lmn} for Al–Cu were used to synthesize a model crystal which was subsequently searched for Cu-rich zones. A mixture of single- and multi layer zones with {100} faces was found, consisting essentially of pure Cu whereas practically no Cu was found in the matrix. The original model for Guinier–Preston-I zones (GEROLD [1954, 1958]) is a Cu monolayer disc about 5 nm in diameter with the adjacent Al planes collapsing toward the zone. The displacements found by AUVRAY *et al.* [1981] were also different (substantially larger but decreasing more rapidly).

The results of AUVRAY *et al.* [1981] were essentially confirmed by MATSUBARA and COHEN [1985] in a study employing X-rays from a synchrotron radiation source (see also COHEN [1986]), but questioned by BUBECK and GEROLD [1986] (see the viewpoint set edited by GEROLD [1988], and COHEN [1992] for further discussions) who demonstrated with a simple one-dimensional model that diffuse intensities from within the first Brillouin zone, i.e., where the displacement terms are minimal, are required to obtain a reliable separation. Smaller scattering vectors than for X-rays (reflection geometry) are easily accessible with neutrons as sizeable samples may also be studied in transmission on the same instrument. For Al-1.75 at% Cu (two Cu isotopes were used to enhance and vary the different scattering contributions), aged for 60 min at 80°C to obtain Guinier–Preston-I zones, a linear scan of diffuse intensities (inelastic scattering was eliminated experimentally by measuring the elastic scattering on a three-axes spectrometer with an energy resolution ± 0.7 or ± 4 meV) covering scattering vectors [*h*00] from 0.175 to 6.9 reciprocal lattice units was analyzed (MÜLLER *et al.* [1988, 1989]), and it became

References: p. 1188.

apparent that results for I_α (see eq. (53) depend on the range of h values used and that I_α retains no structure (i.e., becomes independent of h) if the lowest h ranges are used for the separation (see fig. 24). Thus, there remains only a small (< 10%) probability for multi-layer zones. These investigations have been extended to obtain complete three-dimensional scattering data (MALIK *et al.* [1996]).

The diffuse X-ray scattering of another system forming platelet-shaped zones, Cu–Be, has been studied in as-quenched (KOO *et al.* [1988]) and aged (KOO and COHEN [1989]) Cu-10.9 at% Be crystals. Clusters after quenching and single and multilayer {100} zones after aging are identified and modelled, and ⟨110⟩ diffuse streaks in X-ray scattering are found to be due to thermal diffuse scattering (as confirmed by elastic neutron scattering where no streaks are found, KOO *et al.* [1988]).

Neutron scattering studies of SRO have become more numerous in recent years, taking advantage of the low absorption cross-sections (i.e., lower scattering angles are accessible), the ease of energy resolution to eliminate thermal diffuse scattering, and the possibility of good scattering contrast also for neighbouring elements (e.g., Ni–Fe, Cu–Zn, etc.). Neutrons are, of course, particularly useful if magnetic SRO is to be investigated.

The SRO study of Ni-25 at% Fe (quenched from ~535°C) (LEFÈBVRE *et al.* [1980,

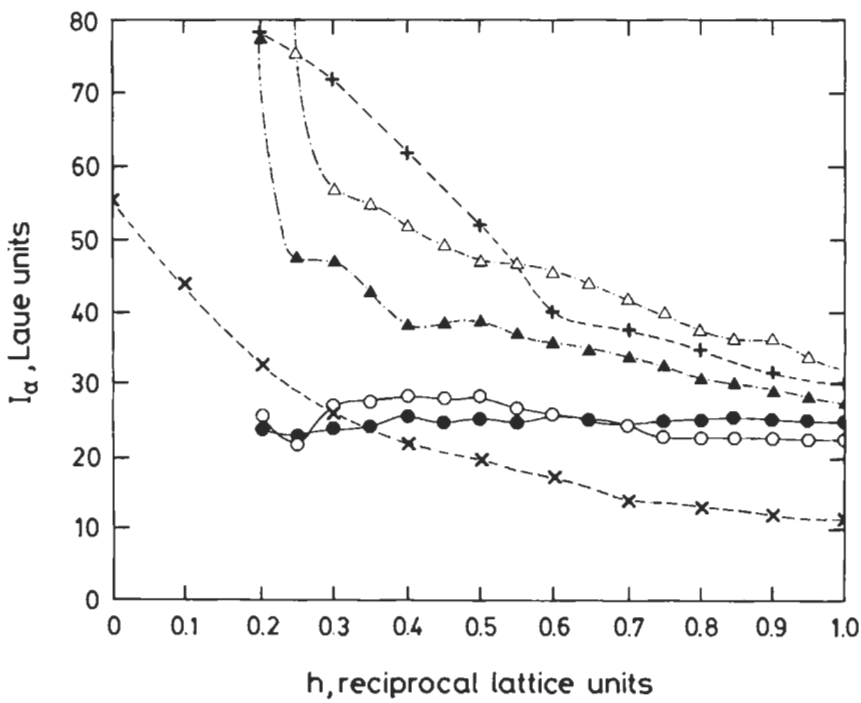


Fig. 24. Short-range order intensity I_α for Al–Cu. ×: calculated from the α_{lhm} of AUVRAY *et al.* [1981], +: taken from fig. 2 of MATSUBARA *et al.* [1985], incomparison with values obtained by MÜLLER *et al.* [1989] within different separation ranges, ○, ●: h = 3 and △, ▲: h = 2 to 5, using the quadratic approximation for Al–^{63}Cu (open symbols) and Al–^{65}Cu (solid symbols).

1981]) demonstrated for the first time the ideal suitability of the Borie–Sparks method for the analysis of diffuse neutron scattering. Furthermore, the least-squares fitting method developed by R. WILLIAMS [1972, 1976] which also includes second-order size-effect corrections, was compared with the Borie–Sparks approach. All displacements were found to be very small, in agreement with the small size effect in Ni–Fe. The α_0 values were always close to one (within 7%, with an uncertainty of 6% attributed to the calibration alone). Effective pair potentials were calculated in the CLAPP and MOSS [1966, 1968] approximation. They show an oscillatory behaviour and a rapid decrease, similar to CuAu (METCALFE and LEAKE [1975]). The ratios V_{200}/V_{110} and V_{211}/V_{110} place Ni_3Fe well within the $L1_2$ stability region, in contrast to Cu_3Au. The use of three different wavelengths at a synchrotron radiation source offers another method to remove the thermal diffuse background, as recently demonstrated for Ni-22.5 at% Fe by ICE *et al.* [1992]. The results for α_{lmn} of the neutron study are essentially reproduced, but in addition, individual displacement parameters are obtained which may be useful in the theoretical modelling of size-effects in these (magnetic) alloys.

In Ni–Cr, displacements are also small, and it is sufficient to consider only linear displacement scattering terms. Three diffuse neutron scattering studies, one on Ni-20 at% Cr (SCHÖNFELD *et al.* [1986, 1988]), the other on Ni-10 at% Cr (SCHWEIKA and HAUBOLD [1988]) – where the ^{58}Ni isotope was used to enhance the diffuse elastic cross-section and simultaneously reduce the incoherent scattering – and on Ni-25/33 at% Cr (CAUDRON *et al.* [1992]) have been published. Figure 25 shows that diffuse peaks are found at $1\frac{1}{2}0$ positions and that the displacement scattering introduces only a small asymmetry. Using IMC procedures, very similar results for the effective pair potentials V_{lmn} were obtained for all these Ni–Cr alloys, indicating that any concentration dependence of V_{lmn} must be very small. There is also no systematic variation of V_{lmn} with equilibration temperature (SCHÖNFELD *et al.* [1988]). For example V_{110} is 26 meV for Ni-11 at% Cr quenched from 833 K (SCHWEIKA and HAUBOLD [1988]), and for Ni-20 at% Cr, values of 26.1(4), 29.3(2) and 28.0(2) are found for crystals quenched from 741, 828 and 973 K, respectively. (The subsequent effective pair potentials decrease rapidly in magnitude, with varying signs.) The Clapp–Moss approximation also yields results comparable to the IMC values, with some deviation for V_{110}, which is smaller than in Ni–Mo, but larger than in Ni–Al. Figure 26 summarizes (SCHÖNFELD [1994]) V_{lmn} results for Ni–Al (8.9 at% Al, 23.5 h at 500°C, neutron scattering), Ni–Cr and Ni–Mo. The γ/γ' phase boundary in Ni–Al can only be reproduced if a concentration dependence of the V_{lmn} is admitted (SCHÖNFELD [1994, 1995]).

In similar neutron scattering studies for Cu-31 at% Zn (REINHARD *et al.* [1990, 1993]) and Cu-22 at% Zn (REINHARD *et al.* [1993]) (using ^{65}Cu to enhance the contrast) equilibrium was established at 473 K or 493 K. Although several other properties had been related to SRO in the α-phase of the Cu–Zn system, no direct evidence was available. Diffuse elastic neutron scattering showed four distinct maxima in the (001) plane (see fig. 27), distorted by displacement scattering. The peak positions may be related to flat portions of the Fermi surface across the ⟨110⟩ directions, as indicated in fig. 27b. A minimum in the Fourier transform of the V_{lmn} will lead to a diffuse scattering maximum. The static dielectric function $\varepsilon(\kappa)$ has a singularity at the Fermi surface,

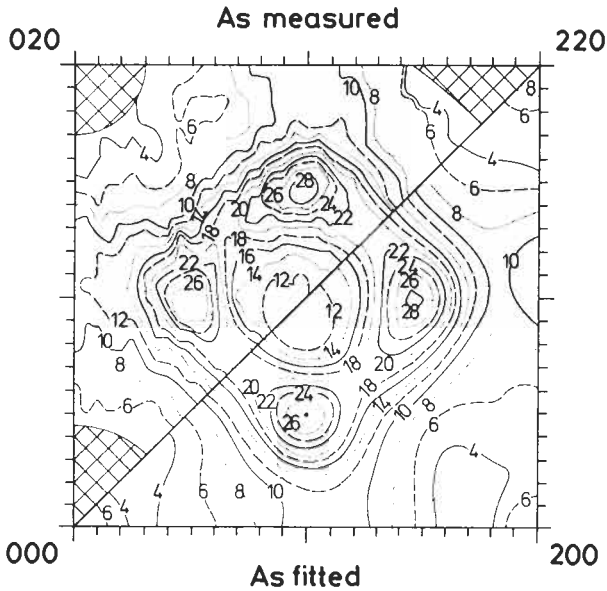


Fig. 25. Diffuse elastic scattering in 0.1 Laue units for a Ni-20 at% Cr single crystal aged for 480 h at 741 K, in a (100) reciprocal plane. Experimental (as-measured) results except for shaded areas (near Bragg peaks or incident beam) are compared with recalculated lines of equal intensity using 21 α_{lmn} and 26 γ_{lmn}. (After SCHÖNFELD *et al.* [1988].)

leading to a reduction of the pure ion-ion interaction. After separation of I_{SRO} and linear displacement scattering, effective pair potentials for Cu–Zn were again obtained from the IMC method (they show no marked temperature dependence and, again, compare favourably with values from the Clapp–Moss approximation). Using these effective pair potentials, the ground-state energies for different (hypothetical) ordered structures of Cu_3Zn were calculated. The $D0_{23}$ structure was found to be more stable than other candidates if more than nine V_{lmn} were used. (The $D0_{23}$ superstructure would introduce order peaks at 100 and $1\frac{1}{4}0$ positions. Diffuse maxima near these positions were also found by TURCHI *et al.* [1991a] based on first-principles concentration-functional theory.) Monte Carlo simulations were performed (REINHARD *et al.* [1990]) to estimate the critical temperature T_c for the long-range order transition. Values of T_c near 330 K were obtained, too low to be easily established experimentally.

An extensive combined X-ray and neutron scattering study to elucidate the interplay of atomic (chemical) and magnetic short-range order has been performed by KOGA *et al.* [1993a, b] on a series of three Ag–Mn single crystals. Ag–Mn shows spin-glass behaviour

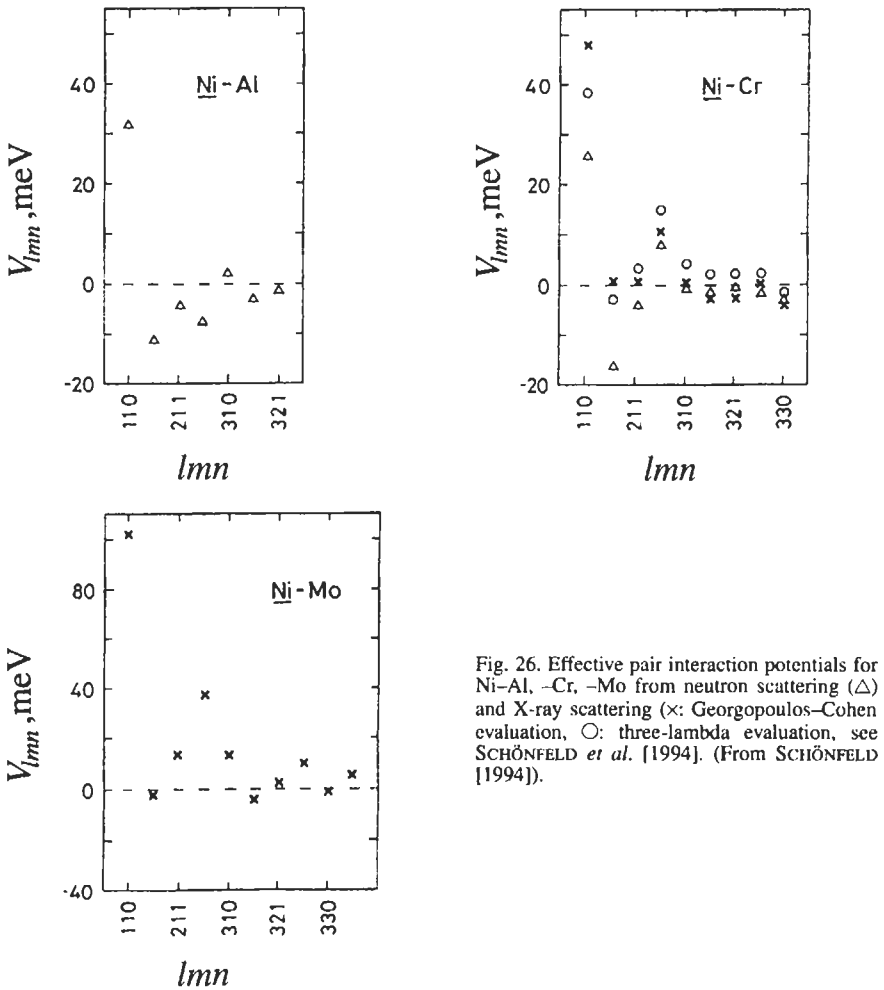


Fig. 26. Effective pair interaction potentials for Ni–Al, –Cr, –Mo from neutron scattering (△) and X-ray scattering (×: Georgopoulos–Cohen evaluation, ○: three-lambda evaluation, see SCHÖNFELD *et al.* [1994]. (From SCHÖNFELD [1994]).

for 10 to 25 at% Mn (see KOUVEL [1961]), and BOUCHIAT *et al.* [1981] had reported on diffuse X-ray peaks near $1\frac{1}{2}0$ and proposed several complex short-range order models, as $D0_{22}$ can be ruled out as the basic unit. The X-ray measurements of KOGA *et al.* [1993a, b] show diffuse peaks with maxima at $1\ \frac{1}{2}\pm\delta\ 0$. After a Borie–Sparks analysis, there is a preference of next-nearest neighbour Mn–Mn pairs for quenched crystals containing 13.3 and 20.8 at% Mn, whereas the crystal containing 28.1 at% Mn shows a dramatic decrease of such correlations. The magnetic diffuse (neutron) scattering confirms that there is antiferromagnetic coupling for nearest-neighbour Mn pairs and ferromagnetic coupling for next-nearest neighbours. A second, longer-ranged ferromag-

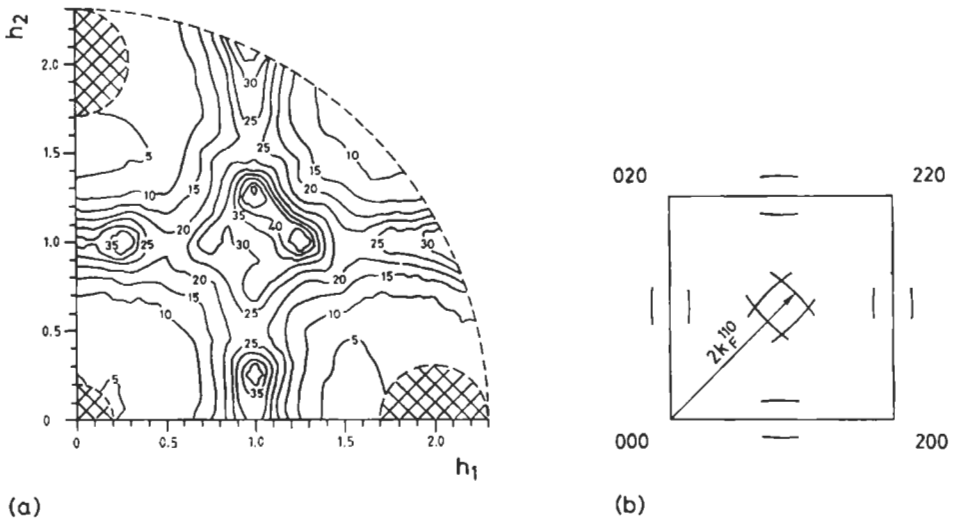


Fig. 27. Diffuse elastic neutron scattering in 0.1 Laue units for a Cu-31 at% Zn single crystal aged for 160 h at 473 K. (a) Lines of equal intensity as measured in a (001) reciprocal plane (the cross-hatched areas near the Bragg peaks are excluded as they may be affected by the weak wings of the mosaic spread), (b) schematic indication of $2k_F^{110}$-positions (k_F^{110} = Fermi vector in ⟨110⟩ direction). (From REINHARD *et al.* [1990].)

netic coupling component has been identified, too. It is believed that this component, attributed to Rudermann–Kittel–Kasuga–Yoshida (RKKY) interactions (see HARDER and WELLS [1983]), is connected with the spin-glass freezing temperature which varies linearly with Mn concentration, whereas the locally coupled spin structure has a strong, "confusing" influence on the details of spin-freezing (KOGA *et al.* [1993a, b]). The homologous system Cu–Mn has been the subject of numerous magnetic scattering studies (see HICKS [1979] and TSUNODA and CABLE [1992] for a complete literature list). Here, the atomic SRO peak seems to be centred at $1\frac{1}{2}0$ whereas the magnetic SRO peaks develop as satellites at $1\ \frac{1}{2}\pm\delta\ 0$ (Cu-25 at% Mn, polarized neutrons, CABLE *et al.* [1984]). These satellites are strongly inelastic, but if atomic SRO is well-developed, an elastic part remains at room temperature in Cu-34.8 at% Mn (TSUNODA and CABLE [1992]). It is suggested that with the development of atomic SRO, there is a tetragonal distortion and spin-density-wave clusters couple locally to ⟨210⟩ directions in the basal plane (as in Ag–Mn, $D0_{22}$ is not a compatible long-range ordered structure, and an A_2B_2 structure seems to be approached). A more extensive diffuse neutron scattering study of Cu-17 at% Mn, separating SRO and linear static displacement scattering, has recently provided first results on the effective pair interaction potentials (equilibrium at 483 K where all magnetic scattering is inelastic) (ROELOFS *et al.* [1995]). The results resemble those obtained for Ni–Cr (see above), with both components just one atomic number below Cu–Mn, but V_{222} and all V_{ll0} up to $l=5$ are positive.

Almost all the diffuse scattering work quoted so far concerned fcc alloys. Transition metal carbides and nitrides of NaCl structure can also be treated as fcc. There are frequently vacancies in the metalloid sublattice which distribute non-randomly. The long-range

superstructures and SRO as well as displacement effects have been reviewed by DE NOVION and LANDESMANN [1987]. The diffuse scattering is evaluated following the same principles as described above. A number of elastic diffuse neutron scattering studies have been reported more recently, e.g., on $TiC_{0.64}$, $TiC_{0.76}$, $TiN_{0.82}$, $NbC_{0.73}$ and $NbC_{0.83}$ single crystals (PRIEM *et al.* [1989a]), $TiN_{0.82}$ (PRIEM *et al.* [1989b]) and on $NbC_{0.73}$ (BEUNEU *et al.* [1990]).

No complete SRO separation is known for any other structures, except for the recent X-ray work on bcc Fe-47 at% Cr exploiting the anomalous scattering when tuning the wavelength close to the Fe and Cr K edges (REINHARD *et al.* [1992]) and a third wavelength minimizing SRO and linear displacement scattering. The alloy shows clustering (e.g., $\alpha_{111} = 0.16$), and the effective pair interaction potentials V_{lmn} determined from IMC calculations and from the Clapp–Moss approximation agree well with the calculations of TURCHI *et al.* [1991b] based on the generalized perturbation method of DUCASTELLE and GAUTIER [1976] combined with the Korringa–Kohn–Rostoker coherent-potential (KKR-CPA) approximation, see, e.g., TURCHI *et al.* [1988].

Diffuse neutron scattering in (magnetic) bcc Fe-13.5 at% V was measured along a few reciprocal lattice lines and fitted with SRO parameters and effective pair potentials in the Clapp–Moss approximation (CABLE *et al.* [1989]). This system shows a tendency towards ordering (β'-CuZn type), and the fitted V_{lmn} agree well with results of a self-consistent field KKR-CPA treatment of the electronic structure including magnetic order within the local spin-density approximation (STAUNTON *et al.* [1990]).

Diffuse neutron scattering of a quenched β-Cu-40 at% Zn crystal was measured in a volume of reciprocal space and evaluated according to Borie and Sparks including second-order displacement scattering (ZOLLIKER *et al.* [1992]). The β' superstructure peaks were already well developed and were excluded from the analysis. This non-equilibrium state shows an almost uncorrelated distribution of surplus Cu atoms on the Zn sublattice of the B2 structure.

Some high-temperature elastic neutron scattering work on Fe-30 at% Al showed broad diffuse peaks that were evaluated as long-range ordered domains (HILFRICH *et al.* [1991]) but might deserve a more elaborate investigation in terms of SRO. Finally, liquid-quenched polycrystalline dilute Fe–Mn and Fe–Ti were studied by neutron diffuse elastic scattering (polarized and unpolarized neutrons) to characterize atomic SRO and magnetic ordering (PARETTE [1991]), indicating a clustering tendency for Mn and displacement scattering only around Ti solutes.

An extension of the Borie–Sparks and Georgopoulos–Cohen methods to multiple sublattices has been presented by HAYAKAWA and COHEN [1975]. A recent diffuse X-ray scattering study (synchrotron radiation) on $Hg_{0.8}Cd_{0.2}Te$ (zincblende structure) with two wavelengths, one near the HgL_{III} edge, revealed small ordered regions with local 3:1 Hg:Cd stoichiometry and second-order displacement effects from the mixed-cation sublattice (QUINTANA and COHEN [1993]).

5. *Near the incident beam*

This section will be mostly concerned with the region around the origin of the reciprocal lattice, i.e., $\kappa \ll g_m$ where g_m is the modulus of the smallest reciprocal lattice

vector. This field is commonly called low- or small-angle scattering which means diffuse scattering for small κ, not just Bragg diffraction at small angles as in the electron microscope or with γ-rays. After a summary of the principles of small-angle scattering, some applications to the study of inhomogeneities in alloys and of defect clusters will be presented. Finally, some other effects and methods for $\kappa \approx 0$ will be discussed.

5.1. Small-angle scattering

Small-angle scattering (SAS) is caused by the variation of scattering length density over distances exceeding the normal interatomic distances in condensed matter. If κ_{max} is the maximum value of the scattering vector accessible in a SAS experiment, details on a scale smaller than about $1/\kappa_{max}$ will not be resolved, and the discrete arrangement of scattering centres (atoms or nuclei) can be replaced by a continuous distribution of scattering length over volumes of about $1/\kappa_{max}^3$. Sums may then be replaced by integrals (cf. eq. (A10.b), ch. 11), and the theory of SAS specializes on the evaluation of these integrals for a variety of cases. For a full account of the theoretical principles, developed for X-rays but equally valid for neutrons, see GUINIER and FOURNET [1955], BEEMAN *et al.* [1957], GUINIER [1963], GEROLD [1967] and POROD [1982], for neutrons see SCHMATZ [1978] and KOSTORZ [1979]. The basic expression for the SAS cross-section is

$$\left(\frac{d\sigma}{d\Omega}\right) = \left| \int_{V_s} \rho_f(\boldsymbol{r}) \exp(-2\pi\, i\boldsymbol{\kappa}\cdot\boldsymbol{r}) d^3\boldsymbol{r} \right|^2 , \tag{55}$$

where the integration extends over the sample volume, V_s, and $\rho_f(\boldsymbol{r})$ is the scattering length density, locally averaged as indicated above. It is also useful to write

$$\rho_f(\boldsymbol{r}) = \Delta\rho_f(\boldsymbol{r}) + \bar{\rho}_f , \tag{56}$$

where $\bar{\rho}_f$ is averaged over distances much larger than $1/\kappa_{min}$, and κ_{min} is the smallest κ value accessible in the experiment. For the κ range considered, only $\Delta\rho_f(\boldsymbol{r})$ will contribute to the scattering, i.e.:

$$\left(\frac{d\sigma}{d\Omega}\right) = \left| \int_{V_s} [\rho_f(\boldsymbol{r})] \exp(-2\pi\, i\boldsymbol{\kappa}\cdot\boldsymbol{r}) d^3\boldsymbol{r} \right|^2 . \tag{57}$$

Equations (55) or (57) may be used as a starting point for analytical or numerical calculations of scattering for model distributions. With the help of the equations of § 2.3, magnetic SAS may also be included for the case of neutrons.

The simplest and very widely used approach to SAS is based on the two-phase model, assuming small particles with a homogeneous scattering-length density ρ_{fp} embedded in a homogeneous matrix of scattering-length density ρ_{fm} ($\rho_{fm} = 0$ covers the case of small particles in vacuo). If there are N_p identical particles without spatial correlations (dilute limit), but – if they are anisometric – all aligned in the same direction:

$$\left(\frac{d\sigma}{d\Omega}\right) = N_p\left(\rho_{fp} - \rho_{fm}\right)^2 V_p^2 S(\kappa), \tag{58}$$

where V_p is the volume of one particle and $S(\kappa)$, the *single-particle scattering function*, is given by:

$$S(\boldsymbol{\kappa}) = \left| \frac{1}{V_p} \int_{V_p} \exp(-2\pi i \boldsymbol{\kappa} \cdot \boldsymbol{r}) d^3 r \right|^2 . \tag{59}$$

$S(\boldsymbol{\kappa})$ can be calculated for a variety of particle shapes (see POROD [1982]) and depends on the direction of $\boldsymbol{\kappa}$ relative to any given axis, except for spheres. For all other particle shapes, $S(\boldsymbol{\kappa})$ has to be averaged for a given orientational distribution. Only if the orientational distribution is isotropic, $\boldsymbol{\kappa}$ can be replaced by κ for anisometric particles.

In many real scattering systems, the single-particle approach is not valid as particle sizes and possibly particle shape as a function of size may vary and spatial as well as orientational correlations must be considered. Generalized cases can only be treated by finite-element methods, but several approximations have been suggested, e.g., evaluation of size distributions of spheres or other uncorrelated particles of known shape, or linear distance distributions of polydisperse spheres (GUINIER [1963]). These problems are extensively discussed e.g., by POROD [1982], GLATTER [1982] and PEDERSEN [1993]. If eq. (58) can be used (possibly after correction for interparticle interference), the evaluation of SAS yields important parameters of the scattering systems. We summarize some general properties.

a. Extrapolation to $\boldsymbol{\kappa = 0}$. If the extrapolation can be performed reliably (experimental problems at very small angles may interfere), eq. (58) yields [because $S(0) = 1$]:

$$\left(\frac{d\sigma}{d\Omega}\right)(\kappa \to 0) = N_p\left(\rho_{fp} - \rho_{fm}\right)^2 V_p^2 . \tag{60}$$

Three parameters determine $d\sigma/d\Omega(\kappa \to 0)$: number of particles (i.e., particle density $n_p = N_p/V_s$), scattering contrast and particle size. Combined with information from other parts of the scattering curve, they may be obtained separately.

b. Guinier approximation. For any particle shape, the scattering function at small values of κa (where a is the relevant size of the particle) can be approximated by an exponential function (GUINIER [1939, 1963]),

$$S(\boldsymbol{\kappa}) = \exp\left(-4\pi^2 \kappa^2 R_\kappa^2\right), \tag{61}$$

where R_κ is the average inertial radius (GUINIER and FOURNET [1955]) of the particle defined by

$$R_\kappa^2 = \frac{1}{V_p} \int_{V_p} r_\kappa^2 q(r_\kappa) dr_\kappa , \tag{62}$$

where $q(r_\kappa)$ is the geometrical cross section of the particle perpendicular to a distance vector r_κ parallel to the direction of κ, with its origin inside the particle (defined by $\int r_\kappa q(r_\kappa) \mathrm{d}r_\kappa = 0$). The scattering function reflects the anisometry of a particle as its decrease with increasing κ is the steeper the larger R_κ. For randomly oriented particles:

$$S(\kappa) = \overline{S(\boldsymbol{\kappa})} = \exp\left(-4\pi^2\kappa^2 R_G^2/3\right), \tag{63}$$

with the radius of gyration R_G defined by

$$R_G^2 = \frac{1}{V_p}\int_{V_p} r^2 \mathrm{d}^3\boldsymbol{r}. \tag{64}$$

Confusion often arises if the term "Guinier radius" is used. It may signify the radius of a sphere, R_s, calculated from R_G, but sometimes R_G itself. For a sphere, $R_G^2 = 3R_s^2/5$, for an ellipsoid with axes $2a$, $2b$, $2c$, $R_G^2 = (a^2+b^2+c^2)/5$. For other particle shapes, see e.g., GLATTER [1982].

The Guinier approximation results from an expansion of the exponential function in eq. (59) and a reinterpretation of the resulting series for $S(\kappa)$ as an exponential [eq. (61)]. The term proportional to κ^2 in the Guinier approximation is exact for any particle shape. For spheres, the term with κ^4 from eq. (63) is correct to better than 10%, and higher-order terms are in fair agreement. The Guinier approximation is therefore acceptable for $2\pi R_G\kappa \leq 1.2$. For not too strongly anisometric particles, a wider range may even be possible, whereas for extremely anisometric particles (e.g., rods or platelets) the higher-order terms will contribute at smaller values of κ (see, e.g., POROD [1982]).

c. Integrated intensity. From eq. (57):

$$Q = \frac{1}{V_s}\int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}(\boldsymbol{\kappa})\,\mathrm{d}^3\boldsymbol{\kappa} = \overline{\left\{\Delta\rho_f(\boldsymbol{r})\right\}^2}, \tag{65}$$

where the integration extends over the entire reciprocal space ($0 \leq \kappa \leq \infty$) and the bar denotes averaging over the sample. As Q represents the mean-square fluctuation of the scattering length density of the system which is insensitive to detailed structural features, it is sometimes (see POROD [1982]) called 'invariant'. In the two-phase model:

$$Q = \left(\rho_{fp} - \bar{\rho}_f\right)\left(\bar{\rho}_f - \rho_{fm}\right) \tag{66a}$$

$$\text{or} \quad Q = C_p\left(1 - C_p\right)\left(\rho_{fp} - \rho_{fm}\right)^2, \tag{66b}$$

where $C_p = N_p V_p / V_s$ is the volume fraction of particles. The "invariant" is, as eq. (65) indicates, the Fourier transform of the scattering cross-section for $r = 0$, i.e., a special value $\gamma(0)$ of a correlation function $\gamma(r)$ (see POROD [1982]) which, however, has a simple meaning only for dilute isotropic systems.

d. Characteristic length. An average correlation length L_c can be defined by (the bar denotes an average of all directions of κ if the system is anisotropic):

$$L_c = \frac{2\pi}{V_s Q} \int \kappa \overline{\frac{d\sigma}{d\Omega}(\kappa)}\, d\kappa, \tag{67}$$

with Q as defined in eq. (65), for any scattering system. In the two-phase model, L_c can be interpreted as the mean length of all lines passing through all points in all directions. The average of all the chords is L_p, the *characteristic length*, given by

$$L_p = 2\pi V_p \int \kappa \overline{S(\kappa)}\, d\kappa. \tag{68}$$

e. Porod approximation. For scattering systems with well-defined interfaces between the two phases (see POROD [1951, 1982]), the final slope of the scattering function is proportional to κ^{-4}. For particles of any shape (κ must be larger than the inverse of the shortest dimension of the particle):

$$\overline{S(\kappa)} \approx (2\pi)^{-3} A_p V_p^{-2} \kappa^{-4}, \tag{69}$$

where A_p is the surface area of the particle. Inserting eq. (69) into eq. (58), we get:

$$\frac{d\sigma}{d\Omega} = (2\pi)^{-3} N_p A_p \left(\rho_{fp} - \rho_{fm}\right)^2 \kappa^{-4}; \tag{70}$$

the cross-section becomes proportional to the total surface area of the scattering particles. The Porod approximation is not valid e.g., for particles with sharp edges or corners (POROD [1982]), diffuse boundaries (RULAND [1971]) or rough interfaces (see WONG and BRAY [1988], SINHA *et al.* [1988]). If the scattering objects or the interfaces are self-similar fractals (see WONG and BRAY [1988], BALE and SCHMIDT [1984], MARTIN and HURD [1987], TEIXEIRA [1988], SCHMIDT [1991]), power laws reflecting the fractal dimensions are found, but power law scattering deviating from a κ^{-4} dependence may also be caused by specific size distributions (see SCHMIDT [1982]). Interparticle interference effects diminish at large κ.

The advantage of the Guinier approximation is that the radius of gyration, i.e., a size parameter, can be determined from uncalibrated SAS data [cf. eqs. (61) and (63), inserted into eq. (58)]. The combination of several of the above relationships [(60) and (61) or (63), (66), (68) and (70)] allows one to evaluate shape, size, number and composition of uniform particles from precise measurements over a sufficiently large range of κ. For example, eqs. (66b) and (70) yield the surface-to-volume ratio of particles, again without calibration of the measured intensities.

Although X-rays are more widely available than neutrons, the use of neutrons in SAS studies of metallic systems is basically more generally applicable. As SAS is measured in transmission geometry, the most severe restriction for X-rays stems from the absorption properties of materials (see § 2.2, table 2). For Cu K_α radiation, the optimum thickness [see eq. (13), $\Sigma_t D_s = 1$] of a pure Al sample is $D_s = 76\ \mu$m, but less than 10 μm for many other metals (atomic numbers 23–27 and above 41, except 55). With Mo K_α, a thickness of 713 μm is ideal for Al, but for most heavier metals, 20–100 μm should not be exceeded. Although with more powerful sources or yet shorter wavelengths somewhat thicker samples may be used, the sample volume may still not be representative of

the bulk material. Undesired scattering from surface irregularities (PARKER [1972], ROTH [1977], KOSTORZ [1976], SIMON *et al.* [1992a]) and sample environment (windows, heat shields etc.) will be more perturbing with X-rays than with neutrons. Another difficulty arises from double Bragg scattering which can obscure the SAS effects and sometimes exceed them. The obvious remedy, increasing the incident wavelength above the Bragg cut-off, works well for neutrons but not for X-rays as absorption becomes prohibitively large. Double Bragg scattering may then only be avoided using properly oriented single crystals. Finally, inelastic scattering may be appreciable at high sample temperatures, and if an energy-resolved separation is required, neutrons must be used.

The applications of SAS of X-rays and neutrons to metallurgical problems are numerous and have frequently been reviewed (GEROLD and KOSTORZ [1978] and KOSTORZ [1979, 1982, 1983, 1988a, b, c, 1991, 1992, 1993], WAGNER and KAMPMANN [1992], WILLIAMS *et al.* [1993]). From a simple verification of sample homogeneity to the determination of sizes, size distributions and interparticle interference effects, the degree of complexity of evaluation procedures varies from simple analytical methods to very involved computer routines. Here, only a few examples of SAS research related to alloys and defects can be given.

5.2. Alloys

GUINIER [1938, 1939] was the first to report SAS results from various Al-rich alloys, e.g. Al–Cu and Al–Ag, and found evidence for the existence of very small coherent precipitates, now known as *Guinier–Preston zones* (GP zones). An important problem which still has not been solved completely is the question how phase separation is initiated and progresses during the early stages when large parts of the sample are still in a supersaturated state (see ch. 15, § 3.1). The SAS technique is sensitive to small (in scale and in amplitude) compositional variations, bridging and overlapping with the domains accessible by field-ion microscopy and transmission-electron microscopy. Many SAS experiments have therefore been performed to study phase separation, but also coarsening reactions, precipitate parameters in relation to other properties, and dissolution of precipitates (see the reviews quoted above). For X-rays, Al-rich alloys are most suitable, and very few other systems have been investigated until very recently when synchrotron radiation sources became available. Initial experiments with neutrons concentrated on Al alloys too, but many other systems have been studied during the last ten years.

As an example, fig. 28 (MESSOLORAS [1974]) shows a set of SAS curves obtained for a polycrystalline Al-6.8 at% Zn alloy (neutron measurements in a cryostat at 4.2 K). As the scattering did not depend on the direction of κ, the SAS cross-sections (calibration with the incoherent scattering of vanadium) are given as a function of the modulus of κ. A peak of the SAS intensity at a value $\kappa_m \neq 0$ is quite common for decomposing or decomposed alloys containing – in their fully decomposed metastable or stable state – a few percent of precipitates ($C_p \leq 7\%$ in the present example). An early appearance of a peak has sometimes been taken as evidence for the mechanism of spinodal decomposition (see ch. 15) but this feature of SAS alone is insufficient to distinguish between concentration fluctuations and well-defined homogeneous particles. Quantitative kinetic

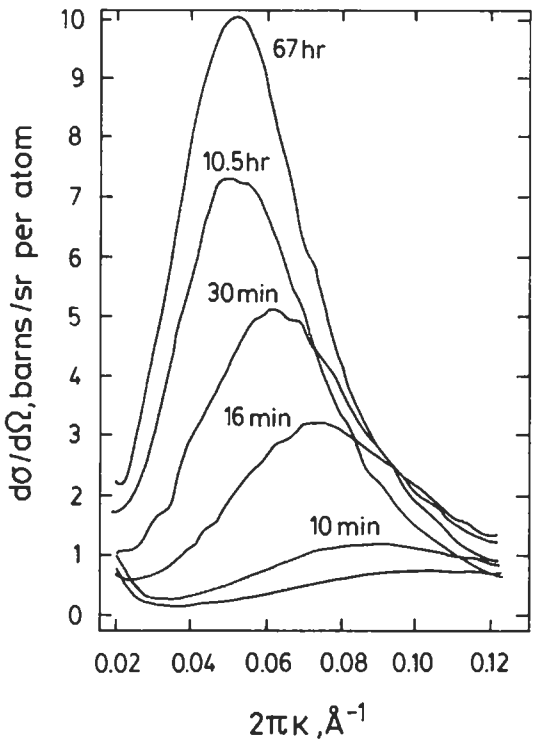


Fig. 28. Neutron SAS cross-sections as a function of $2\pi\kappa$ for an Al-6.8 at% Zn alloy quenched from 310°C into ice water and aged at room temperature as indicated (from MESSOLORAS [1974]).

measurements with a variation of important parameters (initial concentration, quenching conditions, aging conditions) are necessary and must be compared with specific predictions of the different theoretical models. A SAS peak will of course also occur in a two-phase system as the first maximum of the interparticle interference function (see GUINIER and FOURNET [1955], LASLAZ *et al.* [1977] and GLATTER [1982]), and many other cases can be constructed, e.g., a "three-phase" system representing particle, surroundings depleted of the alloying element, and supersaturated matrix (WALKER and GUINIER [1953]). Figure 29 shows measurements obtained at a synchrotron (LURE-DCI, Orsay) with $\lambda = 1.54$ Å, revealing a SAS peak at very small κ for Al-6.8 at% Zn aged at 135°C, a temperature which is definitely above any spinodal at this composition.

Considerable progress has been made in recent years in the theoretical description of the evolution of thermodynamically unstable (spinodal decomposition) or metastable (nucleation-and-growth) systems, and a large variety of analytical and numerical methods have been employed (see ch. 15 and GUNTON *et al.* [1983], FURUKAWA [1985], BINDER [1992], and WAGNER and KAMPMANN [1992] for reviews), providing scattering functions for different situations and stages of phase separation, primarily for statistically isotropic

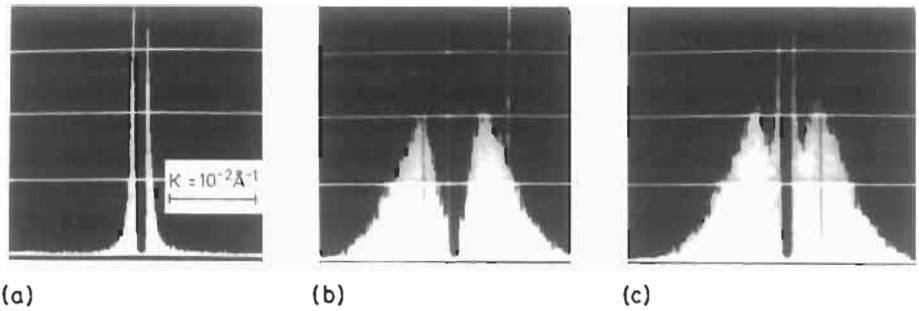


Fig. 29. Multichannel analyzer display of SAS patterns obtained with a linear position-sensitive detector. The centre was covered by a beam stop. (a) Pure Al; (b) Al-6.8 at% Zn, aged for 7 h at 135°C; (c) difference between alloy and pure Al (long white lines in the centre are artefacts). Measuring time 100 s, intensity range (full scale) 51200 counts. (From NAUDON *et al.* [1979].)

systems. Anisotropy may be due to an anisotropic interfacial energy, leading to a variation of the single-particle scattering function $S(\kappa)$, see eq. (59), or it may be related to a correlated arrangement of the scattering entities (in the two-phase model, this could be taken into account by an interparticle interference function in eq. 58) owing to elastic interactions (see, e.g., JOHNSON *et al.* [1990] and WANG *et al.* [1993]).

Anisotropies of a decomposing alloy will, of course, not be visible for a sufficiently untextured polycrystal. Polycrystalline Al–Zn alloys (on the Al-rich side) are ideal candidates for an experimental verification of many theoretical predictions as the matrix is almost elastically isotropic and the Zn-rich GP zones are fully coherent with the matrix and initially spherical because of a small size effect. Under these circumstances, changes in scattering length density can be exclusively attributed to compositional changes (displacement effects, small according to eq. (32), can be neglected). Some years ago, an oversimplified interpretation of the original paper on spinodal decomposition (CAHN [1961]) led to a search for special features at the spinodal line and for the "linear spinodal regime" where the SAS curve should show a time invariant peak at a specific value of κ, κ_m, and an exponential growth/decay of the SAS intensity for all $\kappa \leq \sqrt{2}\,\kappa_m$. The overwhelming evidence of SAS experiments (in accordance with other experimental and more elaborate, more realistic theoretical results, see BINDER [1992]) indicates that for metallic systems, a linear regime is found only exceptionally (and may then be explained by partial phase separation during quenching), and that the search for singularities at the spinodal line may be pointless in a system with short-range interactions. In these cases, the spinodal singularity disappears and there is a gradual transition from the nucleation regime to spinodal decomposition in a region of the phase diagram where the free-energy barrier of a nucleus becomes comparable to the thermal energy, kT. A sequence of scattering curves as shown in fig. 28 is thus more typical. Nonlinear extensions of the spinodal theories, but many other models as well, may yield a similar sequence of scattering curves (see SIMON *et al.* [1984]).

A general analysis of the scattering curves, which is capable of distinguishing

different regimes of the decomposition process is a test of self-similarity (see the literature cited above, and FRATZL and LEBOWITZ [1989]). If a system evolves controlled by a single characteristic length scale $L(t)$, a scaling test of the SAS function $I(\kappa, t)$ (we use I for simplicity for the cross-section) according to

$$I(\kappa,t)/Q(t) = L^3(t)\, s(x,t) \tag{71}$$

where t is time, Q is the integrated intensity as defined in eq. (65), $x = 2\pi L(t)\kappa$, yields a scaled scattering (or "structure") function $s(x)$ if $s(x,t)$ does not change with time. The characteristic length $L(t)$ may be obtained from the radius of gyration of a Guinier plot (see eq. 63), or from the peak position, $L_m = (2\pi\kappa_m)^{-1}$, or as the inverse of the first moment of the scattering curve. If scaling holds, these different parameters should not give different results – but this test is rarely made. Figure 30 shows two tests for scaling where κ_m^{-1} was used to obtain $L(t)$. Power laws for κ_m and the peak intensity, I_m, are also frequently tested, as they hold approximately in several model calculations (though this may be accidental, see e.g., MILCHEV *et al.* [1988]),

$$\kappa_m(t) = A\, t^{-a}, \quad I_m(t) = B\, t^b \tag{72}$$

with constants A, B, and constant exponents a, b. If Q is constant, $b = 3a$, cf. eq. (66b). Scaling is usually observed for advanced stages of decomposition. The scaled structure function depends on the volume fraction and on the morphology of the decomposition product (see KOSTORZ [1991] and LANGMAYR *et al.* [1992] for a comparison of various results). There is agreement that in the dilute "droplet" region, i.e., for isolated precipitates, the exponents a, b coincide with those already predicted by LIFSHITZ and SLYOSOV [1961] and WAGNER [1961], i.e., $a = 1/3$, $b = 1$ for coarsening under the sole influence of an isotropic interfacial energy. In real systems, however, coherent precipitates introduce strain fields, and the resulting elastic interactions may introduce different growth laws. Early stages are expected to follow different time laws as shown, for example, in fig. 31. This figure also illustrates one of the assets of synchrotron radiation (see SIMON *et al.* [1985] and KOSTORZ [1986] for a comparison of X-rays and neutrons). Owing to the high intensity of incident radiation decomposition processes may be followed with a time resolution of about 10 s. With neutrons, a few minutes are usually necessary to accumulate statistically relevant data (see, e.g., WILLIAMS *et al.* [1993] for an assessment of experimental features). Further examples for time-resolved SAS using synchrotron radiation are the studies on decomposition (OSAMURA *et al.* [1993]) and reversion (OKUDA *et al.* [1993]) in Al–Li alloys.

The earliest stages of decomposition, in favourable cases well accessible with the SAS technique, are still not well understood. There is a difficulty in determining a (small) critical radius as even the most elaborate methods to calculate particle size distributions from scattering curves (see, e.g., PEDERSEN [1993]) will not yield reliable values for the lower end as the scattering curve is insensitive to the smallest particles in the distribution. The role of quenched-in vacancies deserves further experimental attention and theoretical consideration, too. For example, the more concentrated Al–Zn alloys studied by ACUÑA [1980] with X-ray SAS show a reduced rate of decomposition

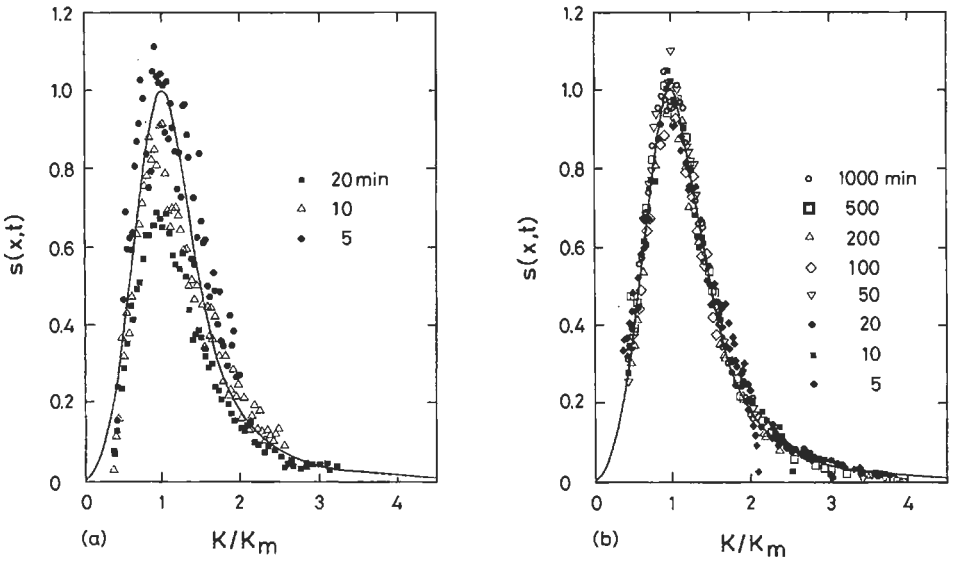


Fig. 30. Scaled structure function $s(\kappa/\kappa_m, t)$ according to eq. (71) for (a) Al-6.8 at% Zn aged at 40°C and (b) Al-10 at% Zn aged at 18°C for the times (total) indicated. The continuous line is a scaling function proposed by FURUKAWA [1979]. (Neutron SAS results of KOMURA *et al.* [1984, 1985].)

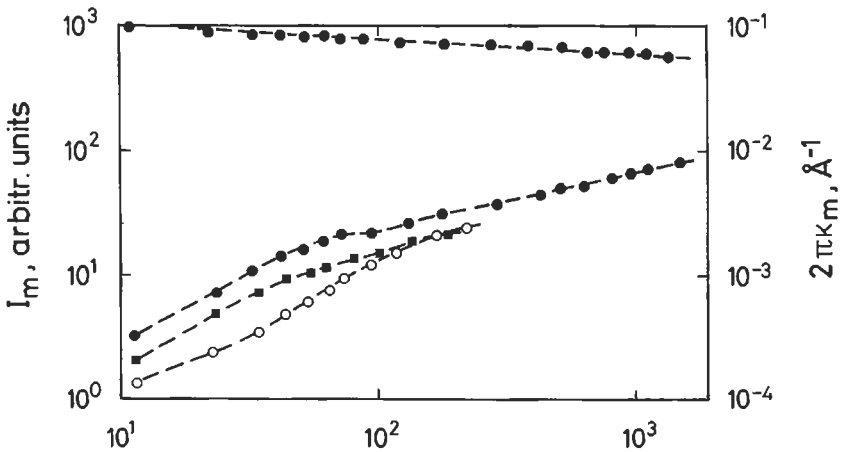


Fig. 31. Maximum intensity, I_m (lower curves), and position of the SAS peak, κ_m, as a function of aging time t, for an Al-12 at% Zn alloy aged *in situ* in a SAS experiment using synchrotron radiation at ■: 139°C, ●: 163°C and ○: 195°C (after SIMON and LYON [1989]).

probably because of a smaller amount of quenched-in vacancies. Computer simulations with vacancies have recently been initiated (see YALDRAM and BINDER [1991], FRATZL and PENROSE [1994]). Besides quenching, unconventional methods of reaching non-equilibrium states in alloys may be helpful, e.g., neutron transmutation or ion implantation (see DELAFOND *et al.* [1981] for $Zn^{+(+)}$ implantation in Al and GP zones studied by X-ray SAS).

The second asset of synchrotron radiation is the possibility of tuning the incident wavelength and thereby exploiting the anomalous scattering terms (see § 2.2). This may be useful in binary alloys to distinguish between surface and volume effects (as the scattering varies proportional to the square of the average scattering length in the former case and proportional to the square of the difference in scattering lengths in the latter, see e.g., SIMON *et al.* [1992b], who showed that a neutron SAS signal from a thin-foil isotope-enriched Fe–Ni Invar alloy (WIEDENMANN *et al.* [1989]) was most likely not related to phase separation). FRATZL *et al.* [1992] used contrast enhancement in the bulk for a study of dilute Cu–Fe alloys where a nucleation regime was identified. The anomalous dispersion also allows decomposition in ternary alloys to be studied (see LYON *et al.* [1985] and SIMON and LYON [1994] for details). In a disordered alloy with three components, there are six independent partial scattering functions, for substitutional solid solutions, this number is reduced to three. The SAS cross-section may then be written proportional to

$$I(\kappa) = |f_A - f_C|^2 S_{AA} + \mathrm{Re}\{(f_A - f_B)(f_B - f_C)^*\} S_{AB} + |f_B - f_C|^2 S_{BB} \tag{73}$$

(A, B, C are the three components, density variations are neglected). A variation of the scattering lengths may also be realized by using different (stable) isotopes in neutron scattering, but this requires that different samples be prepared under exactly the same conditions (see SALVA–GHILARDUCCI *et al.* [1983] for an example). With X-rays, the f_A, f_B, f_C variations according to eq. (14) may be used within the same sample. Experiments on Al–Zn–Ag, Cu–Ni–Fe, Cu–Ni–Sn, Fe–Co–Cr, Fe–Ni–W and quaternary Fe–Ni–Co–Mo have been reviewed by SIMON and LYON [1993] and LYON and SIMON [1994]. As an example, fig. 32 shows results for Cu-42.5 at% Ni-15 at% Fe, obtained from data near the K edges of Fe and Ni. The three independent functions S_{NiNi}, S_{FeFe} and S_{FeNi} were obtained from an overdetermined set of equations of the type given in eq. (73), while S_{CuCu} was calculated. If a two-phase description is adopted, $S^2_{FeNi} = S_{NiNi} S_{FeFe}$ should hold, which is obviously not fulfilled. The best interpretation (SIMON and LYON [1994]) is based on interfacial segregation of Fe between coherent Cu- and (Ni, Fe)-rich domains. The decomposition kinetics are in good agreement with earlier neutron scattering results (WAGNER *et al.* [1984]).

These decomposing alloys in fact show a strongly anisotropic microstructure which is immediately obvious if a two-dimensional SAS pattern is taken from a single-crystalline sample (see LYON and SIMON [1992]). An example from a neutron SAS study of Al-6.8 at% Zn (BUBECK *et al.* [1985], KOSTORZ [1988a]) is shown in fig. 33. Two types of "anisotropy" may be observed. In fig. 33a, there is radial symmetry for the larger values of κ, i.e., the precipitate zones are spherical (single-particle scattering

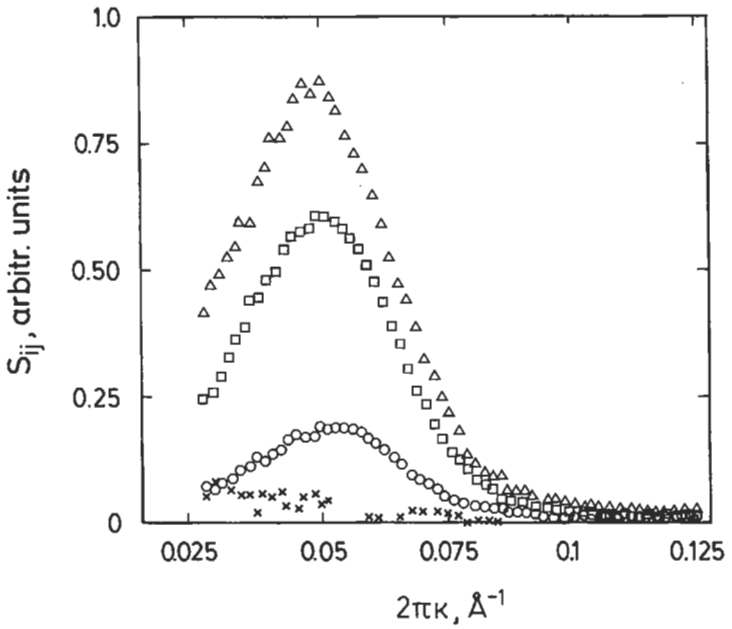


Fig. 32. Partial scattering functions of polycrystalline Cu-42.5 at% Ni-15 at% Fe aged at 500°C for 56 h. Δ: S_{CuCu}, □: S_{NiNi}, ○: S_{FeFe}, ×: S_{FeNi}. (From SIMON and LYON [1994].)

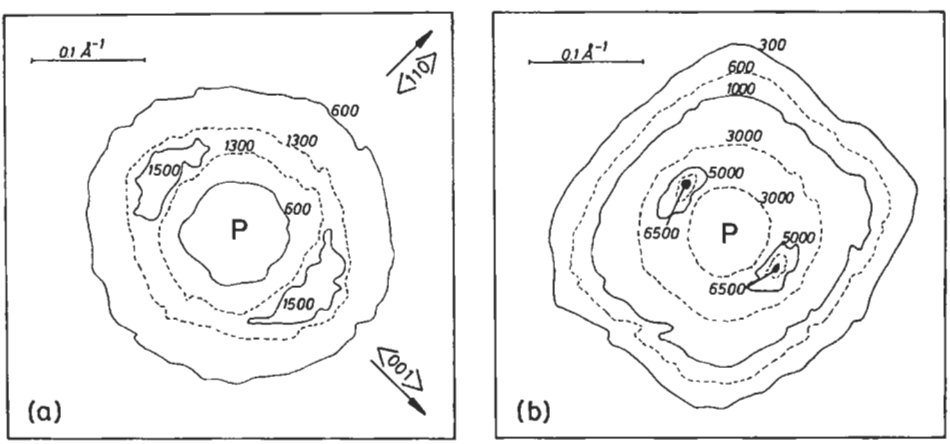


Fig. 33. Small-angle scattering pattern (lines of equal intensity) for an Al-6.8 at% Zn single crystal with (110) faces perpendicular to the incident beam. Two other crystallographic orientations are indicated. The crystal was aged *in situ* at room temperature for (a) 2 h and (b) 7.5 h. The azimuthal dependence of the peak intensity is visible. P indicates the area of the primary beam which has been shaded off by a beamstop. Scale: 0.1 Å^{-1} = $2\pi\kappa$. (From KOSTORZ [1988a]).

dominates at larger κ). The peak intensity at a value $\kappa_m \approx 0.085$ Å, however, varies with the azimuthal angle. This is due to a preferred alignment of zones along the elastically soft ⟨100⟩ directions. The alignment may be approximately described by a one-dimensional interparticle correlation function (BUBECK *et al.* [1985], GUINIER [1956]). Analysis of the kinetics shows (BUBECK *et al.* [1985]) that the zones remain spherical until they reach a radius of about 25 Å, transforming to oblate ellipsoids with the short axes along ⟨111⟩. A large-angle diffuse X-ray scattering study of Al-4 at% Zn (using synchrotron radiation) aged at 80°C (HAEFFNER and COHEN [1989]) shows a shape transition of zones modelled from the experimental SRO parameters at a radius of about 10–15 Å. These are individually identified particles in a population, and a smaller value than obtained from a scattering pattern involving a large number of particles seems acceptable. The work of HAEFFNER and COHEN [1989] also shows that the zones are not faceted but assume rather irregular interfacial structures and high zinc concentrations in the interior.

SAS from single crystals has been observed for many other decomposing alloy crystals, e.g. Al–Cu (GUINIER [1938], SATO *et al.* [1988] – time resolved, strong anisotropy), Al–Ag (DUBEY *et al.* [1991], combined with diffuse scattering, see SCHÖNFELD *et al.* [1992b], KOSTORZ [1994]), Cu–Co (STEINER *et al.* [1983]), Au–Pt (GLAS *et al.* [1992]), Ti–Mo (FRATZL *et al.* [1993], LANGMAYR *et al.* [1994]), Ni–Al and Ni–Al–Mo (BEDDOE *et al.* [1984], CALDERON and KOSTORZ [1990], SEQUEIRA *et al.* [1995]), Ni–Si (POLAT *et al.* [1986]) and Ni–Ti (CERRI *et al.* [1990], VYSKOCIL [1994], see KOSTORZ [1994]).

Most of these results have recently been reviewed (KOSTORZ [1991, 1993]). With the exception of Cu–Co and Ni–Si, all SAS patterns show some anisotropy caused by misfit strains or anisotropic interfacial energy and often preferred correlations along the elastically soft directions. An example where the scattering anisotropy is entirely related to the platelike character of the scattering objects is shown in fig. 34. In this sample, Guinier–Preston I zones are formed which are very flat Cu-rich platelets in the {100} planes. Despite the dramatic difference between the two directions, these results alone do not allow mono- and multilayers to be distinguished (HAEFFNER *et al.* [1988]). Thin platelets have also been identified in internally oxidized Cu–Fe single crystals (FRATZL *et al.* [1993], PARIS *et al.* [1994]). If very sharp streaks or peaks are to be measured, the curvature of the Ewald sphere must be taken into account, as misleading results might otherwise be obtained for a fixed scattering arrangement.

An example for the strong influence of elastic interactions between precipitates is shown in fig. 35. The mismatch between γ matrix and γ' (Ni_3Al) precipitates in the binary Ni–Al alloy is about 0.31% whereas it is close to zero for the ternary alloy. While there is still an interference ring indicating a sharp distance distribution, no azimuthal intensity variation is seen in this case. The precipitates are spherical as confirmed by electron microscopy, and there is no preferred alignment. In binary Ni–Al, the cuboidal particles show pronounced correlations along ⟨100⟩. The growth rate decreases dramatically in the alloy containing Mo.

Some more applied SAS studies with neutrons on technical alloys have been summarized (KOSTORZ [1979, 1988b]). The advantage of low absorption of neutrons may be used to study large components non-destructively or to apply mechanical stresses on sufficiently large samples, prior to or during the scattering experiment. For example,

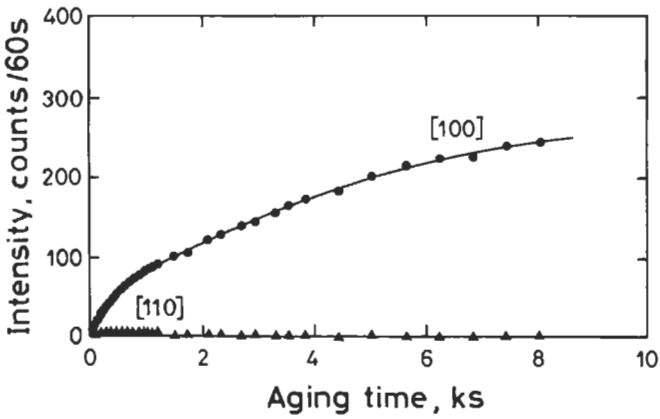


Fig. 34. Small-angle scattering intensity of an Al-1.7 at% Cu single crystal in the region of larger κ values, taken *in situ* at 25°C with a linear position-sensitive detector along [100] or [110]. The incident beam is parallel to [001], and the intensity is monitored at $\kappa \approx 0.07$ Å^{-1} as function of aging time. (From SATO *et al.* [1988].)

neutron SAS of fatigued Cu–Co single crystals (STEINER *et al.* [1983]) and a Ni-base superalloy (Nimonic PE16, SUNDARARAMAN *et al.* [1992]) has confirmed that precipitates dissolve in the zones of extensive dislocation activity. In polycrystalline Cu-2 at% Co, uniaxial compressive deformation was found to introduce SAS anisotropy due to sheared coherent or semicoherent precipitates (ANCRENAZ and SERVANT [1992]). A special nondestructive neutron SAS study of a very complex material has been reported by GOTODA *et al.* [1989]). From multifilamentary composite wires less than 1 mm in

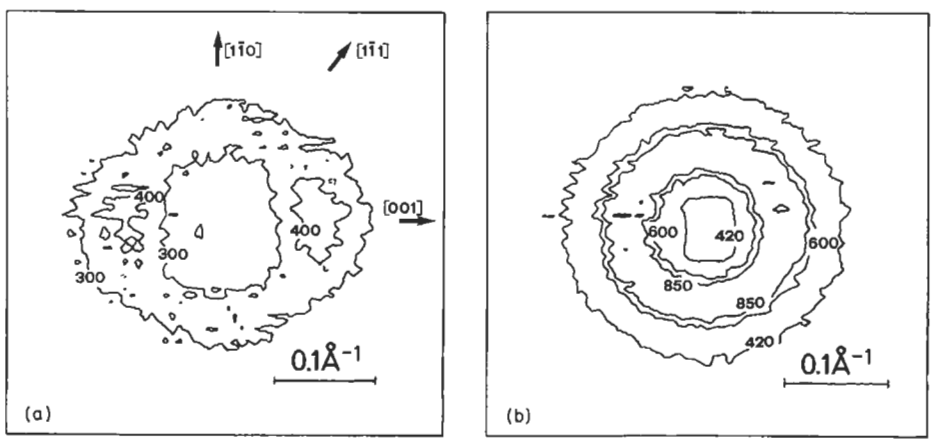


Fig. 35. Lines of equal intensity (counts in 20 min) of neutron SAS data for single crystals of (a) Ni-12 at% Al aged for 35 min at 560°C and (b) Ni-6.1 at% Al-9.4 at% Mo aged for 3.9 h at 610°C. The incident beam is parallel to [110]. Some other crystallographic directions are indicated. Scale: 0.1 Å^{-1} = $2\pi\kappa$. (From CALDERON and KOSTORZ [1990].)

diameter, containing bundles of very thin (16 to 39 μm) filaments of superconducting Nb–Ti, embedded in a normal copper matrix, short pieces were arranged with the axes parallel to each other across the incident neutron beam. The global pinning force density controlling the critical current in the superconducting state depends mainly on the volume fraction of α precipitates in the β matrix of Nb–Ti. It was possible by neutron SAS to determine changes of the volume fraction of α-phase particles from 1.3 to 13.4%, depending on the thermomechanical treatment, and to interpret the differences in critical current on this basis, using supporting results from transmission electron microscopy.

Decomposition processes in amorphous alloys (with and without crystallization) have received particular attention. SAS (mostly of X-rays) has been reported for a variety of substances, see e.g., OSAMURA *et al.* [1982, 1984] for Fe–Zr, YAVARI *et al.* [1988] for Pd–Ni–P and Cu–Zr, LAMPARTER *et al.* [1988] for Fe–B and Fe–B–Si–C, also using neutrons, RODMACQ *et al.* [1984] for Pd–Si using neutrons, JANOT and GEORGE [1985] for Fe–Cr–P–C (these authors showed that surface effects are predominant for thin ribbons), GOUDEAU *et al.* [1987] (X-rays, anomalous scattering) and MANGIN *et al.* [1985] (neutrons) for hydrogenated Cu–Ti, SCHAAL *et al.* [1986] for Mg–Zn, MARET *et al.* [1992a] for Fe–Mn–Y, MARET *et al.* [1992b] for Tb–Cu and Gd–Cu, and LORENTZ *et al.* [1994] for Fe–Ge. Figure 36 shows an example of a rolled and subsequently heat-treated sample of $Fe_{90}Zr_{10}$. The drastic change of the SAS intensity is attributed to fine-scale separation (~60 Å) into α-Fe and $Fe_3Zr/Fe_{23}Zr_6$ (OSAMURA *et al.* [1984]). No Bragg peaks are found with X-rays or electrons.

Compositional fluctuations and clustering above the critical temperature for phase separation may also be studied by SAS techniques. For large-scale fluctuations, i.e., those relevant for the SAS signal, the continuum theory of fluctuation waves (see KRIVOGLAZ [1969]) is appropriate, and one can write [starting from eq. (43)] for a binary alloy (v_a is the atomic volume):

$$\frac{d\sigma}{d\Omega} = \frac{NkT}{v_a}\left(f_B - f_A\right)^2\left(g'' + 4\pi^2 B\kappa^2\right)^{-1}, \tag{74}$$

as

$$\overline{|c(\kappa)|^2} = \frac{kT}{v_a}\left(g'' + 4\pi^2 B\kappa^2\right)^{-1} \tag{75}$$

with $g'' = d^2g/dc^2$ and g = Gibbs free energy density, and B = gradient energy term. The free energy density g may contain, apart from the chemical term g_0, a strain energy term g_δ taking into account the elastic strain energy around a cluster. Defining a correlation length ξ by

$$\xi^2 = \frac{B}{g_0'' + g_\delta''} \tag{76}$$

and using

$$\frac{d\sigma}{d\Omega}(0) = \frac{NkT\left(f_B - f_A\right)^2}{v_a\left(g_0'' + g_\delta''\right)} \tag{77}$$

one can write

References: p. 1188.

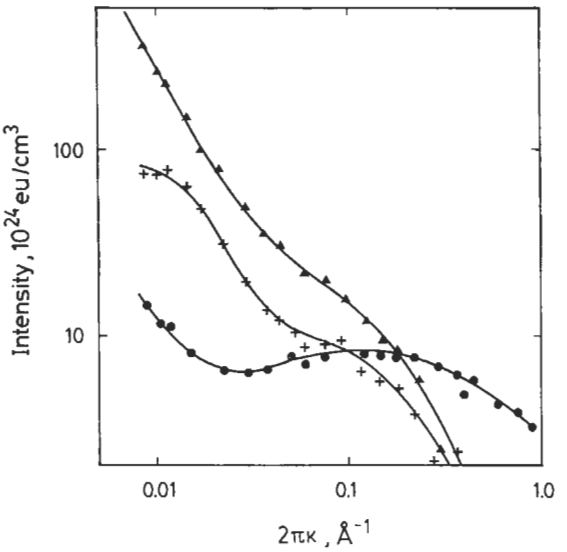


Fig. 36. X-ray SAS for amorphous $Fe_{90}Zr_{10}$, ●: after rolling, and after subsequent annealing at 350°C for +: 10 min, and ▲: 50 min (after OSAMURA *et al.* [1982]).

$$\frac{d\sigma}{d\Omega}(\kappa) = \frac{d\sigma}{d\Omega}(0)\frac{1}{1+4\pi^2\xi^2\kappa^2} \tag{78}$$

According to eqs. (76) and (77), $d\sigma/d\Omega(0)$ and the correlation length diverge if $g_0'' \to g_\delta'' \to 0$, i.e. when a coherent critical point is approached. This has been confirmed by SCHWAHN and SCHMATZ [1978] for Al–Zn of the critical composition $c = 39.5$ at% Zn. The stable phase diagram shows a critical temperature of 351.5°C, but neutron SAS measurements (at temperature) indicated no special behaviour at this temperature. Upon further cooling, large changes occurred between 324 and 322°C, corresponding to the coherent critical point at 323°C. A depression of the coherent spinodal by 28 K is well understood on the basis of CAHN's [1961] formulation of the elastic energy term.

Neutron SAS on frozen-in homophase fluctuations was measured by POERSCHKE *et al.* [1986] in Cu–Ni–Fe alloys where the decomposition kinetics are very slow. Samples were quenched from several temperatures above the coherent miscibility gap. Scattering from large precipitates at grain boundaries and some decomposition products picked up during the quench have to be separated, but the remaining SAS intensity yields equilibrium parameters $\xi^{-2}(T)$ from an Ornstein–Zernike–Debye plot [reciprocal intensity vs. κ^2, see eq. (78)]. A plot of $\xi^{-2}(T)$ vs. T can be used to determine T_{coh}, the temperature of the coherent miscibility gap (see fig. 37).

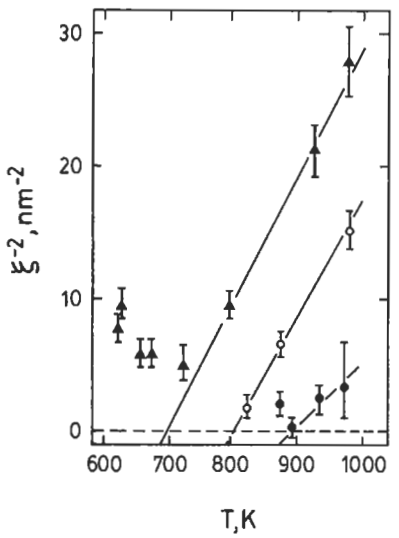


Fig. 37. Temperature dependence of the square of the reciprocal correlation length for homophase fluctuations, ξ^{-2}, for ▲: Cu-48 at% Ni-2 at% Fe, ○: Cu-46 at% Ni-4 at% Fe and ●: Cu-48 at% Ni-8 at% Fe derived from an Ornstein–Zernike analysis of neutron SAS data. Extrapolation to $\xi^{-2} = 0$ yields T_{coh}. (After POERSCHKE *et al.* [1986].)

Similar measurements have been performed by FURUSAKA *et al.* [1986] for Fe–Cr, another system with very slow decomposition kinetics. Ornstein–Zernike behaviour is also found for short aging times within the miscibility gap and large κ. From a recent neutron SAS study, BLEY [1992] concluded that two characteristic lengths are necessary to describe the aging behaviour in this system.

Magnetic clusters and precipitates can be studied with neutron SAS. The magnetic structure factor, eq. (19), can, in analogy to nuclear scattering, be reinterpreted by introducing an integration over fluctuations of the local magnetization density rather than discrete magnetic moments:

$$\boldsymbol{M}(\boldsymbol{\kappa}) = \frac{1}{V_s} \int_{V_s} \boldsymbol{M}(\boldsymbol{r}) \exp(-2\pi \mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{r}) \mathrm{d}^3\boldsymbol{r} \,. \tag{79}$$

A homogeneously magnetized sample will therefore not show any magnetic SAS, but fluctuations in the orientation and/or the magnitude of $\boldsymbol{M}(\boldsymbol{r})$ will be revealed.

A number of useful applications of eq. (20) for unpolarized neutrons are possible. In the simple case of a magnetic phase embedded in a non-magnetic matrix or *vice versa*, the ratio $I^{\perp}/I^{\parallel}$ is independent of the magnitude of $\boldsymbol{\kappa}$ for complete saturation. Any

deviation is an indication of incomplete saturation (then spin flip scattering will occur and an analysis is worthwhile) or of some other scattering contribution. Using measurements of $I^{\perp}/I^{\parallel}$, BEAVEN *et al.* [1986] were able to distinguish scattering contributions from small voids in the presence of somewhat larger Cu precipitates in Fe (the ratio $I^{\perp}/I^{\parallel}$ is 1.4 for voids in ferromagnetic Fe, but 11 for Cu precipitates).

In inhomogenous alloys, $\boldsymbol{M}(\boldsymbol{\kappa})$ may vary considerably over a wide range of temperatures, as the Curie temperature depends on the chemical composition. Without polarization analysis, magnetic scattering as a function of temperature will give a first indication of the extent and magnitude of magnetized regions. fig. 38 shows the SANS curves $I^{\perp}$ and $I^{\parallel}$ for a partially decomposed Ni-14 at% Al alloy containing small (~5 nm diameter) nonmagnetic Ni_3Al particles in an inhomogeneous matrix, measured at a temperature where the matrix is completely ferromagnetic (194 K). The Al concentration of the matrix is low near the precipitates (~11 at%) and still high at some distance. The variation of $I^{\perp}/I^{\parallel}$ is related to the gradient of saturation magnetization density around the precipitates. As the temperature is increased, those parts containing higher Al concentrations will become paramagnetic first and some additional scattering in the $I^{\perp}$ part is expected which will initially increase with increasing temperature until, at a higher temperature, the magnetized regions are reduced to shells around the precipitates that finally disappear once T exceeds the highest T_c (corresponding to regions with the lowest Al concentration). Preliminary tests using unpolarized neutrons have shown the expected results. As for those parts of the sample where the concentration implies a T_c near the temperature of measurement, the magnetization will fluctuate, and measurements with polarized neutrons will be even more instructive.

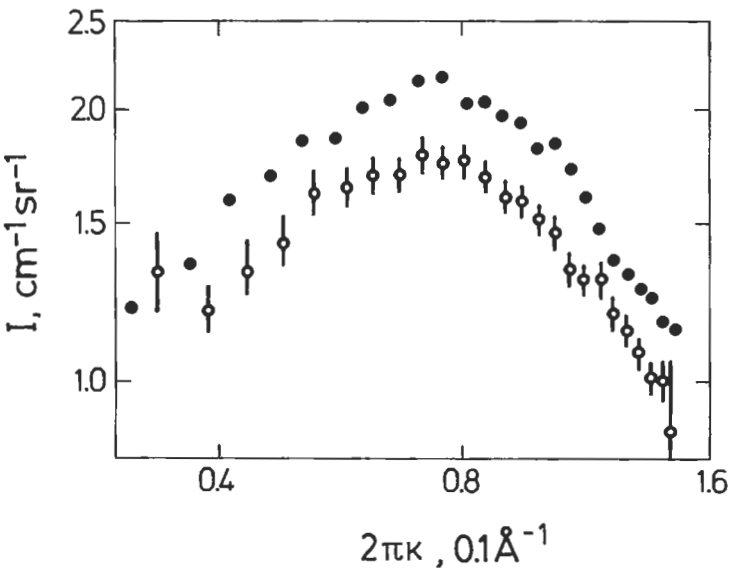


Fig. 38. Macroscopic neutron SAS cross sections at 194 K, ○: parallel and ●: perpendicular to the magnetic field for inhomogeneous ferromagnetic Ni–Al containing nonmagnetic Ni_3Al precipitates of an average radius of about 25 Å (KOSTORZ [1991]).

Measurements of magnetic neutron SAS in zero field for bulk ferromagnets are not advisable if domain walls are present as these will introduce multiple refraction. However, the scattering due to fluctuations of magnetic moments over length scales accessible by neutron SAS may be easily followed for dynamical fluctuations (giant clusters, superparamagnetism, spinglass behaviour, etc.) as a function of temperature. Thus, the formation of finite spin clusters has been revealed by SAS of neutrons, e.g., in certain Au–Fe alloys in the early experiments of MURANI (1976). No sizeable Fe clusters for 15 or 25 at% Fe were observed in recent X-ray SAS experiments (YOSHIDA *et al.* [1989], FRATZL *et al.* [1991]) but the same authors found a short-range-order tendency by Mössbauer spectroscopy and attributed the strong dependence of the magnetic properties on the thermal history in these alloys to temperature-dependent (higher-order) correlations, especially third-neighbour Fe–Fe pairs. Other neutron SAS experiments have shown magnetic clustering of finite range in Ni–Mn (MOZE *et al.* [1985]), Pt-10 at% Fe, (RADHAKRISHNA *et al.* [1989]) and in amorphous $Fe_{91}Zr_9$ (RHYNE and FISH [1985]). For a discussion of SANS from disordered magnets in a magnetic field of variable strength, see CULLEN and CALLEN [1986].

A rare example for the use of polarized neutrons in SAS is the study of compositional and magnetic inhomogeneities in the Fe-65 at% Ni Invar alloy (see TAKEDA *et al.* [1987]). These authors used two different samples, one with natural Ni (scattering length 1.03×10^{-14} m), the other with the isotope ^{62}Ni (-0.87×10^{-14} m), to introduce a change of sign between chemical (nuclear) and magnetic scattering-length densities, $\rho(\boldsymbol{r})$ and $\rho_0 M(\boldsymbol{r})$ (compared in saturation).

By taking the difference of $I^{\perp++}$ and $I^{\perp--}$, one obtains the interference scattering function between nuclear and magnetic scattering which is controlled by correlations given by

$$\int \{\rho(\boldsymbol{r}) - \bar{\rho}\}\{M(\boldsymbol{r}+\boldsymbol{r}') - \overline{M}\}\mathrm{d}^3\boldsymbol{r}'. \tag{80}$$

A change of sign of $\{\rho(\boldsymbol{r})-\bar{\rho}\}$ will also change the sign of this term. The experiments (after correction for depolarization, calibration, etc.) showed a small positive result for $\{I^{\perp++}-I^{\perp--}\}$ for the sample prepared with natural Ni, whereas for Fe–^{62}Ni the difference was negative. A model with paramagnetic Fe-rich regions embedded in a ferromagnetic Ni-rich matrix is compatible with the results. However, dynamic spin fluctuations also contribute significantly to SAS in this temperature range.

Neutron depolarization of the transmitted beam is due to magnetic inhomogeneities larger than about 100 Å (up to a few mm) and is complementary to SAS (see, e.g., ROSMAN and REKVELDT [1991] for antiferromagnetic particles and MITSUDA *et al.* [1991, 1992] for spin glass behaviour in Fe–Al and Au–Fe).

5.3. Defects

Here, we discuss the SAS from dislocations and clusters of point defects as further examples of inhomogeneities revealed by this technique, and some other defect scattering results.

In an otherwise homogeneous material, dislocations yield a SAS signal according to

eq. (57) because of local density changes which can be described by elastic continuum theory. This scattering is very small, and as the presence of dislocations simultaneously increases the probability of double Bragg scattering, early attempts to reveal SAS from dislocations with X-rays failed. ATKINSON [1959] and ATKINSON and LOWDE [1957] used neutrons of long wavelength and finally found true SAS of the correct intensity from deformed metals. Theoretical work by ATKINSON and HIRSCH [1958a, b] and SEEGER and coworkers (SEEGER and KRÖNER [1959], SEEGER [1959], KRONMÜLLER *et al.* [1963], SEEGER and RÜHLE [1963]) for isotropic cubic crystals and by SEEGER and KRÖNER [1959] and SCHMATZ *et al.* [1974], who considered elastic anisotropy, has been reviewed by SCHMATZ [1975] along with the (scarce) experimental results. The SAS cross-sections are so small that even for high dislocation densities, not many specific dislocation arrangements have been studied, and surface irregularities or a few large inclusions or pores may easily obscure the scattering pattern.

In magnetic crystals, magnetoelastic coupling leads to an additional SAS term (for unpolarized neutrons), and deformed single crystals of Ni (ANDERS *et al.* [1984]) and Fe (GÖLTZ *et al.* [1986]) have been studied making use of the additional intensity and its symmetry properties near magnetic saturation. Despite the magnetic scattering, experiments of this type are very tedious and require a high incident neutron flux.

The agglomeration of point defects should lead to more easily measurable SAS effects as the scattering contrast is much higher. The structure and annealing behaviour of voids in Al single crystals, produced by irradiation with fast neutrons, have been extensively studied (see HENDRICKS *et al.* [1977]), LINDBERG *et al.* [1977] and references therein). X-rays, neutrons, transmission electron microscopy and positron annihilation were combined to obtain a rather complete picture of the microstructure of these samples. From the SAS point-of-view, the two-phase model can be employed, and no interparticle interference effects are visible at volume fractions of about 1% and void sizes of several hundred Å. Whereas the voids in irradiated Al showed faceting attributed to truncated octahedra with {111} faces, voids formed in quenched and annealed β'-NiAl single crystals of strictly stoichiometric composition showed faceting on {110} planes, forming a rhombic dodecahedron (EPPERSON *et al.* [1978b]). Figure 39 shows an example. The crystallographic symmetry of the faceted voids is reflected in the scattering pattern. SAS at smaller angles than shown in the figure is "isotropic", and a radius of gyration $R_G \approx 240$ Å can be determined for this particular case.

The void concentrations in the studies mentioned above, about 0.1–1%, were easily measurable in otherwise perfect single crystals. Among the factors influencing mechanical properties of metals and alloys, void formation during fatigue or high-temperature creep presents a major problem, and it is desirable to recognize void formation as early as possible. SAS, especially with neutrons because bulk samples can be examined without further preparation, is a useful method if other scattering contributions are not prohibitively large (see WALTHER and PIZZI [1980]). Grain boundary cavitation in fatigued Cu was studied by PAGE *et al.* [1982]) who found cavities in the range of 100–900 Å. YOO *et al.* [1982] measured neutron SAS from fatigued polycrystalline Ni of commercial purity. A strong influence of temperature on void formation was found. Small voids, about 20 Å in diameter, have been seen by neutron SAS after fatigue

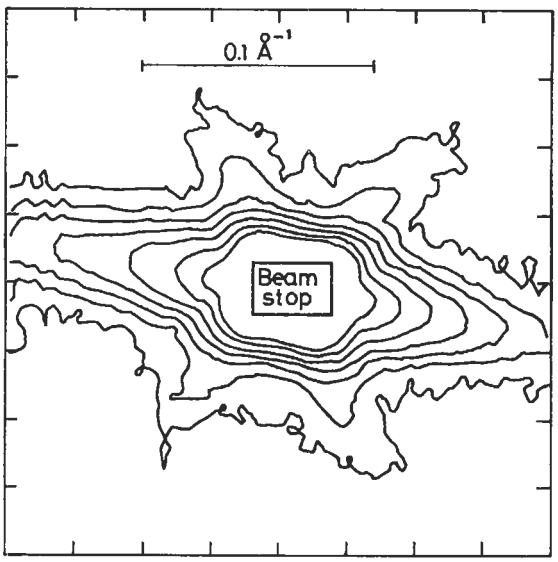


Fig. 39. Lines of equal intensity of neutron SAS measured for a β'-NiAl single crystal containing faceted voids. The crystal was annealed at 1600°C, quenched and aged for 30 min at 650°C. The [110] direction was parallel to the incident beam. The two-dimensional position-sensitive detector was placed 2.5 m away from the sample, the neutron wavelength was 6.64 Å. Scale: 0.1 Å^{-1} = $2\pi\kappa$. (After EPPERSON *et al.* [1978b].)

deformation of ⟨111⟩-oriented Cu single crystals (KETTUNEN *et al.* [1981], LEPISTÖ *et al.* [1991]). Figure 40 shows the SAS increment between a late and early state of fatigue life. A major scattering contribution at low κ must be attributed to dislocations, but at higher κ, voids (or small dislocation loops) are present (volume fraction ~2×10^{-5}). The existence of small vacancy clusters in the persistent slip bands of fatigued Cu has been confirmed by positron annihilation (LEPISTÖ *et al.* [1981]).

Finally, helium bubbles forming in implanted Fe–Ni–Cr have been studied by neutron SAS (CARSUGHI [1994]). Their size-dependent gas density requires a size-dependent contrast factor.

5.4. Special topics

A special case of small-κ scattering is Bragg diffraction from structures with large lattice constants or large-scale periodicity, including magnetic structures (see, e.g., LEBECH *et al.* [1989] for a long-range spiral magnetic structure of a period of ~700 Å in a cubic FeGe single crystal). The flux line lattice of type II superconductors with a lattice constant typically around 1000 Å and first revealed by neutron scattering by CRIBIER *et al.* [1964], has subsequently been studied in great detail for many different materials and conditions (see GEROLD and KOSTORZ [1978], HICKS [1979]). Recently, flux line lattices have also been found in Y–Ba–Cu–O superconductors (FORGAN *et al.* [1990], YETHIRAJ *et al.* [1993]) by neutron diffraction.

References: p. 1188.

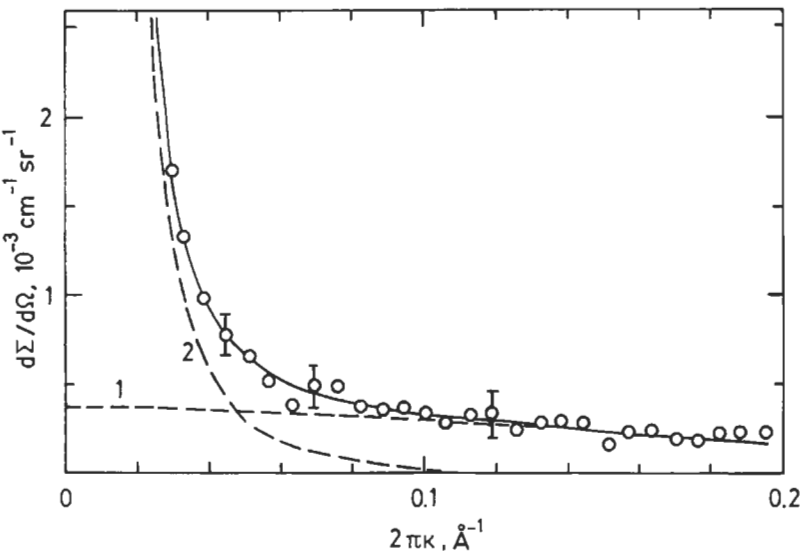


Fig. 40. Neutron SAS increment between a ⟨111⟩ Cu single crystal fatigued to about 80% of its expected fatigue life and another ⟨111⟩ crystal fatigued to only 20% (peak hardened). The circles represent experimental values, the solid curve is the sum of curve 1 (void scattering) and curve 2 (dislocation scattering). (After KETTUNEN *et al.* [1981].)

Periodic and near-periodic multilayers can also be studied in specular reflection near the critical angle of total reflection (see e.g., PENFOLD [1992]). Such investigations are essential in the development of X-ray and neutron mirrors. With the improved collimation at a synchrotron radiation source, SAS measurements under grazing incidence are providing a new surface-sensitive tool (LEVINE *et al.* [1989]) which has been used to study agglomerates and pores at surfaces (see also NAUDON *et al.* [1994]).

At small angles, multiple refraction in two-phase systems may obscure scattering, especially for neutrons of large wavelength and for thick samples (see BACON [1975], SCHMATZ [1978], KOSTORZ [1979]). A separation of refraction and scattering is in principle possible by varying the incident wavelength. Strong multiple neutron refraction from magnetic domains has been evaluated by GALOTTO *et al.* [1976]. BERK and HARDMAN–RHYNE [1985, 1988] have considered multiple scattering and refraction and developed a multiple-scattering treatment of neutron SAS data from dense powders and porous systems. Multiple scattering was found to be the appropriate description of the (wavelength dependent) scattering from alumina powders (~30% dense, average particle diameter 340 nm, samples 2 to 10 mm thick, HARDMAN–RHYNE and BERK [1985]). The theory was also successfully applied to SAS measurements on microporous silica and alumina for which the densification was studied (LONG *et al.* [1991]).

A very simple method of monitoring changes in the microstructure of a sample is to measure its transmission. As indicated by eq. (13), the transmission is a function of the total removal cross-section, i.e., scattering into any direction outside the range covered by the transmission monitor appears like additional absorption. The method has been

proposed for nondestructive testing of materials containing large inhomogeneities, e.g., cavities (PALACIOS *et al.* [1981]). With a very highly collimated beam, and if SAS is the only source of additional/variable scattering, a change in the transmission corresponds to the integrated SAS intensity change as given by eq. (65), or, for isotropic scattering

$$Q = \frac{4\pi}{V_s} \int_0^\infty \kappa^2 \frac{d\sigma}{d\Omega}(\kappa) d\kappa. \tag{81}$$

In reality, the lower limit of the integral will be finite, κ_{min}. It is also possible to work with a fixed acceptance angle $4\theta_a$ for transmission measurements and vary the wavelength of the incident radiation, thus testing the value of the integral of eq. (81) as a function of its lower limit $\kappa_{min} \approx 2\theta_a/\lambda$ (see ENGELMANN *et al.* [1979] for an experiment with ultracold neutrons). As many details of the scattering function will thus be lost, angular-resolved measurements are preferred if the intensity is high enough.

The extended X-ray absorption fine structure (EXAFS), i.e., the energy dependence of absorption up to about 1000 eV above the absorption edge (e.g., the K edge, see fig. 2) can now be measured with sufficient accuracy. The basic physical process leading to oscillations of the absorption coefficient is the modification of the final state of the photoelectron of the excited atom by neigbouring atoms (see, e.g., STERN [1974]). The effect of different chemical species on phase and amplitude of the oscillations is rather complex, but in principle specific for each type of atom. Nearest neighbours of the excited atom have the most pronounced influence, and the EXAFS technique is thus a useful tool (see, e.g., FONTAINE [1993]), complementary to scattering studies in all areas discussed and mentioned in this chapter.

6. *Energy transfers*

As mentioned in § 2.4, inelastic neutron scattering (coherent and incoherent) is a versatile method to study phonons [eqs. (24) and (25)] whereas incoherent (as well coherent) energy-resolved quasi-elastic scattering of neutrons reveals diffusive motion. A few examples will now be given.

6.1. Phonons in real crystals

The dynamic properties of crystals are very sensitive to lattice defects and change dramatically near phase transitions. Phonons and defects have been discussed by NICKLOW [1979] and phonons and phase transformations e.g., by CURRAT and PYNN [1979], and by SCHOBER and PETRY [1993]. As an example, fig. 41 (NICKLOW *et al.* [1979]) shows the frequency distribution for several phonon wavevectors in a Cu single crystal irradiated with thermal and fast neutrons to produce Frenkel pairs (concentration $\sim 1.3 \times 10^{-4}$). The measured phonon peaks are shifted relative to those measured for defect free Cu, and an additional component on the high-frequency side occurs. Subsequent annealing results in an elimination of the peak shifts (72 K), but the additional high-frequency structure is only removed at 800 K. Part of the results can be explained by a

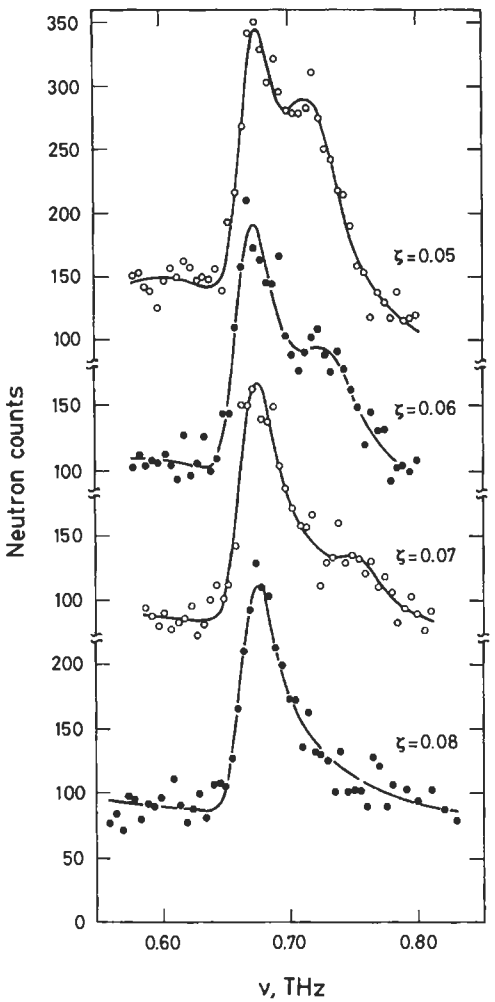


Fig. 41. The frequency distributions of neutrons scattered from irradiated Cu at 4 K for $q = (0, 0, \zeta)/a$. Open and solid circles are experimental data, the lines are fits including resonant coupling of librational modes of the split interstitial. (After NICKLOW *et al.* [1979].)

resonant coupling of phonons to the librational modes of the split interstitial, but other defects, possibly small vacancy clusters, also contribute (see also DEDERICHS and ZELLER [1980] for details).

Larger interstitial defect concentrations are present in metal-hydrogen systems (for

coherent neutron scattering, D is used instead of H), and local modes are more easily observed (see BLASCHKO *et al.* [1988] for D in rare-earth metals).

The phonon spectra of many crystals show phonon anomalies at temperatures well above the transformation temperature. A very clear example is shown in fig. 42 where the phonon dispersion surface for acoustic phonons travelling along the chain direction of a one-dimensional conductor is indicated. The anomaly for the longitudinal modes is related to the special shape of the Fermi surface which is restricted, in the one-dimensional case, to a pair of points. Because of the electron–phonon coupling, phonon anomalies are expected (KOHN [1959]) for scattering processes involving opposite sides of the Fermi surface connected by the wavevector $2k_F$ (k_F is the Fermi wavevector), i.e., when the energy of the electrons is not changed. For a one-dimensional metallic system the electronic susceptibility diverges for $2k_F$. The resulting "giant" Kohn anomaly is responsible for the displacive transformation at lower temperature which may be interpreted as a "condensation" of these special phonons into a pattern of static displacements, leading to a modulation of period $1/2k_F$ (see COMÈS and SHIRANE [1979] for details).

In two and three dimensions, Kohn anomalies, i.e., singularities in the phonon dispersion curves due to the abrupt change of the screening response of conduction electrons when wave vectors at the Fermi surface are involved, depend on the topology of the Fermi surface. Sections of low curvature are particularly effective. A structural phase transformation may in many cases be induced by the condensation ("softening") of phonons of the appropriate wavevectors to create the new structure. An instructive case is β'-NiAl which shows a broad shallow dip in the transverse acoustic phonon branch with [110] propagation and [1$\bar{1}$0] polarization (see MOSTOLLER *et al.* [1989]) for

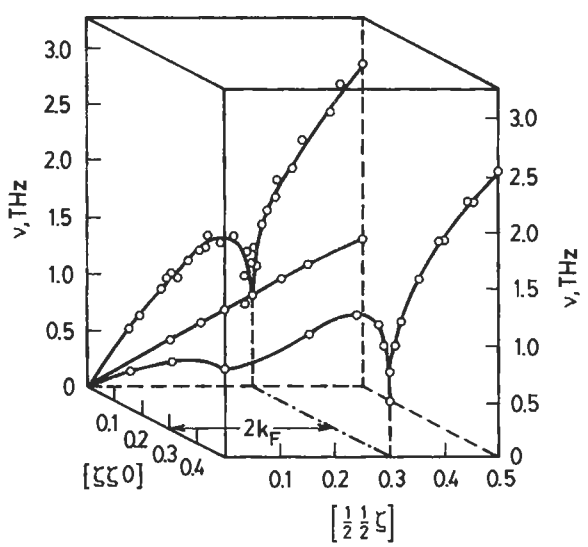


Fig. 42. Room-temperature dispersion surface of KCP $[K_2Pt(CN)_4Br_{0.3}\cdot xH_2O]$ for acoustic phonons propagating along the chain direction. A giant $2k_F$ anomaly is clearly visible for the longitudinal mode. (After COMÈS and SHIRANE [1979].)

the stoichiometric composition. Upon cooling, these phonon modes do not condense, and there is no structural transition. For the off-stoichiometric composition Ni-37.5 at% Al, a martensitic phase transformation occurs at about 80 K (the martensitic start temperature M_S varies from 0 K to 600 K, approximately linearly with off-stoichiometry between about 60 at% and 68 at% Ni, see OCHIAI and UENO [1988]). The transverse phonon branch shows a deeper dip when the transformation temperature is approached (TANNER *et al.* [1990]), see fig. 43. The transformation sets in before the condensation is complete (the corresponding elastic constant C′ is also reduced but not to zero), indicating that the transformation must be triggered at strain defects, e.g., dislocations (TANNER *et al.* [1990]). Application of an external stress leads to a shift of the dip in the dispersion curve (SHAPIRO *et al.* [1993]). ZHAO and HARMON [1992]) have calculated the band structure of Ni_xAl_{1-x} and related the martensitic transformation to Fermi nesting and strong electron-phonon interactions and suggest that local variations in concentrations and thus in order parameter could trigger the transformation. In contrast to NiAl, isostructural NiTi shows a static instability when M_S is approached, and the same phonon branch collapses (see TIETZE *et al.* [1984]).

Dynamic and/or static instabilities are also expected for pure bcc metals undergoing a structural transformation. For the group-IV metals, the transition temperatures lie far above room temperature, and if a single crystal is to be studied, it must always be kept above the transition. In a specially designed furnace (see FLOTTMANN *et al.* [1987]),

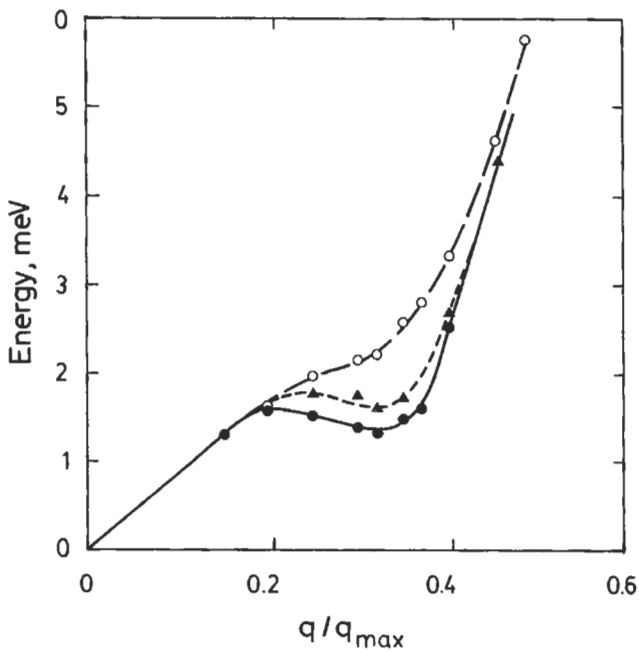


Fig. 43. Neutron scattering results (energy tranfer $E = h\nu$) for the transverse acoustic phonon branch in [110] direction ([1$\bar{1}$0] polarization) of Ni-37.5 at% Al (B2 structure) at ○: 290, ▲: 150, ●: 85 K (after TANNER *et al.* [1990].)

crystals of bcc Ti (PETRY *et al.* [1991a]), Zr (HEIMING *et al.* [1991a]), and Hf (TRAMPENAU *et al.* [1991]) have been grown and investigated at high temperatures by neutron inelastic scattering. The phonon branches show dynamical (but no static) precursor effects for the displacements necessary to reach the hdp α-phase and the ω-phase (found under pressure or by alloying these metals).

In the vicinity of the order–disorder transition in Cu_3Au, selected phonon scans show marked changes at temperatures about 25 K below T_c (GAULIN *et al.* [1990]) which require further experimental and theoretical study.

6.2. Diffusive motion

The use of incoherent quasi-elastic neutron scattering (see § 2.4) to study the motion of hydrogen in metals has been reviewed, e.g., by SPRINGER [1977] and by SKÖLD *et al.* [1979]. The incoherent scattering is somewhat simpler to interpret, but coherent quasi-elastic scattering may also be evaluated (see SPRINGER [1977], LOVESEY [1984], and PETRY *et al.* [1991b] for the glass transition). The availability of backscattering spectrometers (BIRR *et al.* [1971] and spin-echo instruments (MEZEI [1980]) allows changes of neutron energies to be detected between 10^{-5} and 10^{-9} eV, corresponding to atomic motion on a time scale of 10^{-11} and 10^{-7} s, or diffusion coefficients in the range of 10^{-7} to 10^{-11} cm^2/s. Self-diffusion in solid metals can thus also be studied at temperatures not too far below the melting point if the absorption is low. AIT–SALEM *et al.* [1979] and GÖLTZ *et al.* [1980] studied self-diffusion in Na(bcc), using incoherent quasi-elastic neutron scattering, at temperatures between 323 and 370.2 K (melting point 370.9 K). The κ dependence of the energy width gives very detailed information on the jump mechanism. The monovacancy mechanism with nearest-neighbour jumps is predominant, with a divancy contribution of about 20%.

A more recent example concerns the unusually fast self diffusion in bcc β-Ti (PETRY *et al.* [1988], VOGL *et al.* [1989], measured on a backscattering spectrometer directly after growth *in situ* (to avoid the transition to the α-phase at 882°C). Figure 44 shows the results obtained at 1460 and 1530°C along with a visualization of the two jump vectors used for the fitted curves. Such fits can be attempted for all reasonable combinations of jump mechanisms, and the best fits are shown; in fig. 44(a) the admixture of next-nearest neighbours tends to improve the fit for some orientations, but in fig. 44(b), the nearest-neighbour jumps are better. The authors conclude that, if any, there might be 10–15% of direct next-nearest-neighbour jumps. The diffusion coefficients are in the range of 10^{-7} to 10^{-8} cm^2/s. Spin-echo measurements have been performed down to 1100°C (polycrystals) where the diffusion coefficient is only $\sim 10^{-9}$ cm^2/s. The phonon softening observed in the inelastic scattering measurements (see above) provides a means to explain the low migration enthalpy for nearest-neighbour jumps, as explained in detail by PETRY *et al.* [1989]. Diffusion of Fe, Co, Ni and other metallic solutes in group-IV metals is at least two orders of magnitude faster than self-diffusion. In a similar study of dilute Zr–Co, Co diffusion via a vacancy mechanism was identified as the predominant mechanism (PETRY *et al.* [1987]), but only an enhancement of diffusivity by a factor of ten was found via this path, calling for another mechanism, e.g., via interstitial sites. While

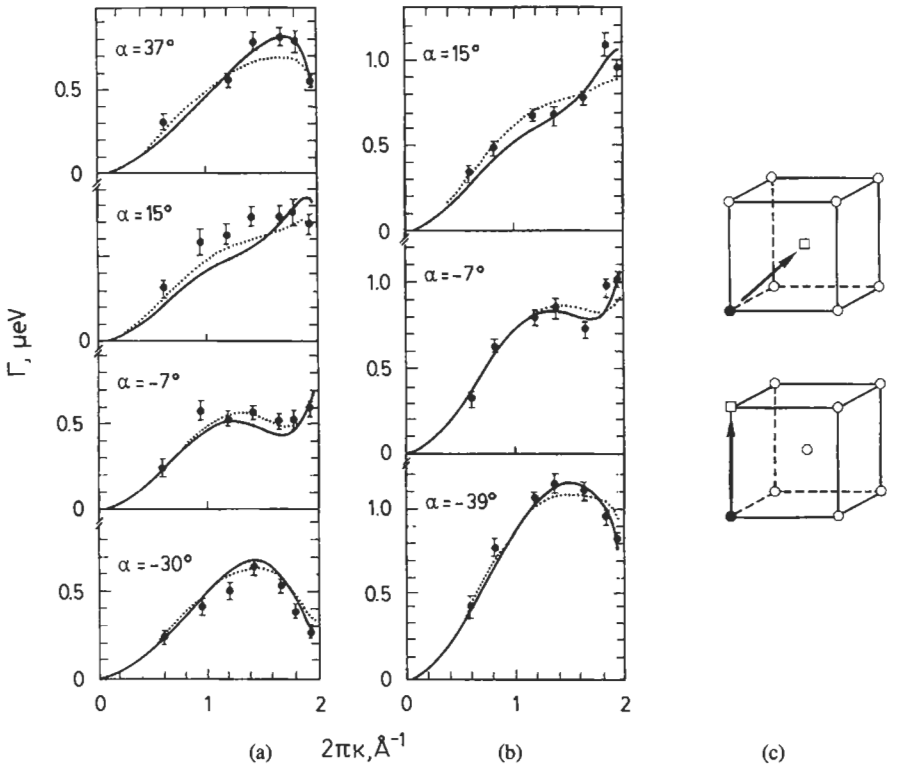


Fig. 44. Quasi-elastic line broadening Γ (full width at half maximum) for two single crystals of Ti, at (a) 1460°C and (b) 1530°C, as a function of scattering vector κ for different angles α with respect to the lattice axes (two different scattering planes). The solid curve is calculated for nearest-neighbour (NN) jumps, the dotted curve for 75% NN jumps and 25% direct next-nearest-neighbour jumps. (c) Schematic of the two types of jump used in the fits of (a) and (b). (From VOGL *et al.* [1989].)

a fraction of 28% of interstitials was still compatible with these measurements, lattice parameter changes (HEIMING *et al.* [1992]) and diffuse scattering (HEIMING *et al.* [1991b]) show that only a small fraction, i.e., a few percent of the solutes present, may be on interstitial sites. This means that the interstitial diffusion path is considerably faster than anticipated.

Acknowledgements

The author is grateful to many colleagues, especially to J. B. Cohen, V. Gerold, B. Schönfeld, C. J. Sparks, and G. Vogl for helpful comments and suggestions.

References

ACUÑA, J., 1980, J. Mater. Sci. **15**, 20.
AIT-SALEM, M., T. SPRINGER, A. HEIDEMANN and B. ALEFELD, 1979, Phil. Mag. **39**, 797.

AL USTA, K., H. DOSCH and J. PEISL, 1990, Z. Phys. **B79**, 409.
ALLEN, A.J., M.T. HUTCHINGS, C.G. WINDSOR and C. ANDREANI, 1985, Adv. Phys. **34**, 445.
ANCRENAZ, P., and C. SERVANT, 1992, J. Phys. I (France) **2**, 1113.
ANDERS, R., M. GIEHRL, E. RÖBER, K. STIERSTADT and D. SCHWAHN, 1984, Solid State Commun. **51**, 111.
ANDREWS, S.R., and R.A. COWLEY, 1985, J. Phys. C **18**, 6427.
ATKINSON, H.H., 1959, J. Appl. Phys. **30**, 637.
ATKINSON, H.H., and P.B. HIRSCH, 1958a, Phil. Mag. **3**, 213.
ATKINSON, H.H., and P.B. HIRSCH, 1958b, Phil. Mag. **3**, 862.
ATKINSON, H.H., and R.D. LOWDE, 1957, Phil. Mag. **2**, 589.
AUVRAY, X., P. GEORGOPOULOS, and J.B. COHEN, 1881, Acta Metall. **29**, 1061.
BACON, G.E., 1975, Neutron Diffraction (Clarendon Press, Oxford).
BALE, H.D., and P.W. SCHMIDT, 1984, Phys. Rev. Lett. **53**, 596.
BARDHAN, P., and J.B. COHEN, 1976, Acta Cryst. **A32**, 597.
BARDHAN, P., H. CHEN and J.B. COHEN, 1977, Phil. Mag. **35**, 1653.
BARRETT, C.S., and T.B. MASSALSKI, 1980, Structure of Metals (Pergamon Press, Oxford).
BARUCHEL, J., J.L. HODEAU, M.S. LEHMANN, J.R. REGNARD and C. SCHLENKER, eds., 1993, Neutron and Synchrotron Radiation for Condensed Matter Studies, vol. III (Applications to Solid State Physics and Chemistry) (Les Editions de Physique, Les Ulis, France).
BAUER, G.S., 1979, Diffuse Elastic Neutron Scattering, in: Neutron Scattering, ed. G. Kostorz (Academic, New York) p. 291.
BEAVEN, P.A., F. FRISIUS, R. KAMPMANN and R. WAGNER, 1986, Analysis of Defect Microstructures in Irradiated Ferritic Alloys, in: Atomic Transport and Defects in Metals by Neutron Scattering, eds. C. Janot, W. Petry, D. Richter and T. Springer (Springer, Berlin) p. 228.
BEDDOE, R.E., P. HAASEN and G. KOSTORZ, 1984, Early Stages of Decomposition in Ni–Al Single Crystals Studied by Small-Angle Neutron Scattering, in: Decomposition of Alloys: the early stages, eds. P. Haasen, V. Gerold, R. Wagner, M.F. Ashby (Pergamon, Oxford) p. 233.
BEEMANN, W.W., P. KAESBERG, J.W. ANDEREGG and M.B. WEBB, 1957, Size of Particles and Latttice Defects, in: Handbuch der Physik, ed. S. Flügge, vol. 32 (Springer, Berlin) p. 321.
BENEU, B., T. PRIEM, C.H. DE NOVION, S. LEFEBVRE, J. CHEVRER and A.N. CHRISTENSEN, 1990, J. Appl. Cryst. **23**, 497.
BERK, N.F., and K.A. HARDMAN-RHYNE, 1985, J. Appl. Cryst. **18**, 467.
BERK, N.F., and K.A. HARDMAN-RHYNE, 1988, J. Appl. Cryst. **21**, 645.
BESSIÈRE, M., S. LEFÈBVRE, Y. CALVAYRAC, F. BLEY and M. FAYARD, 1982, J. Appl. Cryst. **15**, 94.
BINDER, K., 1992, Spinodal Decomposition, in: Materials Science and Technology, vol. 5, Phase Transformations in Materials, ed. P. Haasen (VCH Verlagsgesellschaft, Weinheim), p. 405.
BIRR, M., A. HEIDEMANN and B. ALEFELD, 1971, Nucl. Instr. and Meth. **95**, 435.
BLASCHKO, O., G. KREXNER, L. PINTSCHOVIUS, P. VAJDA, and J.N. DAOU, 1988, Phys. Rev. **B38**, 9612.
BLEY, F., 1992, Acta Metall. Mater. **40**, 1505.
BORIE, B., and C.J. SPARKS, 1971, Acta Cryst. **A27**, 198.
BOUCHARD, R., S. KOUPTSIDIS, H.B. NEUMANN, T. SCHMIDT and J.R. SCHNEIDER, 1993, J. Appl. Phys. **73**, 3680.
BOUCHIAT, H., E. DARTYGE, P. MONOD and M. LAMBERT, 1981, Phys. Rev. **B23**, 1375.
BREUER, F.J., and H. BLEICHERT, 1993, Gamma-Ray Diffractometry, in: Concise Encyclopedia of Materials Characterization, eds. R.W. Cahn and E. Lifshin (Pergamon, Oxford).
BROWN, P.J., 1979, Neutron Crystallography, in: Neutron Scattering, ed. G. Kostorz (Academic, New York) p. 69.
BUBECK, E., and V. GEROLD, 1986, J. Appl. Cryst. **19**, 164.
BUBECK, E., V. GEROLD and G. KOSTORZ, 1985, Crystal Res. & Technol. **20**, 97.
BUTLER, B.D., and J.B. COHEN, 1989, J. Appl. Phys. **65**, 2214.
CABLE, J.W., S.A. WERNER, G.P. FELCHER and N. WAKABAYASHI, 1984, Phys. Rev. **B29**, 1268.
CABLE, J.W., H.R. CHILD and Y. NAKAI, 1989, Physica B **156/157**, 50.
CAHN, J.W., 1961, Acta Metall. **9**, 795.
CAHN, J.W., 1968, Trans. AIME **242**, 166.

CALDERON, H., and G. KOSTORZ, 1990, Lattice Misfit and Decomposition in Ni–Al–Mo Alloys, in: Neutron Scattering for Materials Science, eds. S.M. Shapiro, S.C. Moss, and J.D. Jorgensen (MRS Symp. Proc. No. 166, Mat. Res. Soc., Pittsburgh) p. 255.
CARSTANJEN, H.D., and R. SIZMANN, 1972, Phys. Lett. **40**, 93.
CARSUGHI, F., 1994, J. Appl. Cryst. **27**, 326.
CAUDRON, R., M. SARFATI, M. BARRACHIN, A. FINEL, F. DUCASTELLE and F. SOLAL, 1992, J. Physique (I) **2**, 1145.
CERRI, A., SCHÖNFELD, B. and G. KOSTORZ, 1990, Phys. Rev. **B42**, 958.
CHAKRAVARTI, B., E.A. STARKE, JR., C.J. SPARKS and R.O. WILLIAMS, 1974, J. Phys. Chem. Solids **35**, 1317.
CHASSAGNE, F., M. BESSIÈRE, Y. CALVAYRAC, P. CENEDÈSE, and S. LEFÈBVRE, 1989, Acta Metall. **37**, 2329.
CHEN, H., J.B. COHEN and R. GHOSH, 1977, J. Phys. Chem. Solids **38**, 855.
CHEN, P.C., B.O. HALL and P.G. WINCHELL, 1980, Metallurg. Trans. **11A**, 1323.
CHU, S.N.G., A.T. MACRANDER, K.E. STREGE and W.D. JOHNSTON, JR., 1985, J. Appl. Phys. **57**, 249.
CLAPP, P.C., 1971, Phys. Rev. **B4**, 255.
CLAPP, P.C., and S.C. MOSS, 1966, Phys. Rev. **142**, 418.
CLAPP, P.C., and S.C. MOSS, 1968, Phys. Rev. **171**, 754.
COHEN, J.B., 1970, The Order–Disorder Transformation, in: Phase Transformations (ASM, Metals Park, Ohio) p. 561.
COHEN, J.B., 1986, Solid State Phys. **39**, 131.
COHEN, J.B., 1992, Met. Trans. **23A**, 2685.
COMÈS, R., and G. SHIRANE, 1979, X-ray and Neutron Scattering from One-Dimensional Conductors, in: Highly Conducting One-Dimensional Solids, eds. J.T. Devreese, R.P. Evrard and V.E. van Doren (Plenum, New York) p. 17.
COOPER, M., 1977, Contemp. Phys. **18**, 489.
COWLEY, J.M., 1950, J. Appl. Phys. **21**, 24.
COWLEY, J.M., 1981, Diffraction Physics, 2nd Ed. (North-Holland, Amsterdam).
COWLEY, J.M., and S. WILKINS, 1972, Derivation of Long-Range Interaction Energies from Diffuse Scattering in Diffraction Patterns, in: Interatomic Potentials and Simulation of Lattice Defects, eds. P.C. Gehlen, J.R. Beeler and R.I. Jaffee (Plenum, New York) p. 265.
CRIBIER, D., B. JACROT, L. MADHAV RAO and B. FARNOUX, 1964, Phys. Lett. **9**, 106.
CULLEN, J., and E. CALLEN, 1986, J. Magn. Mater. **54–57**, 311.
CULLITY, B.D., 1977, Elements of X-ray Diffraction (Addison-Wesley, Reading, MA).
CURRAT, R., and R. PYNN, 1979, Phonons and Phase Transitions, in: Neutron Scattering, ed. G. Kostorz (Academic, New York) p. 131.
DARWIN, C.G., 1922, Phil. Mag. **43**, 800.
DE BERGEVIN, F., and M. BRUNEL, 1972, Phys. Lett. **39A**, 141.
DE FONTAINE, D., 1971, J. Appl. Crystallogr. **4**, 15.
DE FONTAINE, D., 1979, Solid State Phys. **34**, 74.
DE NOVION, C.H., and J.P. LANDESMAN, 1987, Nonstoichiometric Compounds, in: Advances in Ceramics, vol. 23 (The American Ceramic Society, Westville, Ohio) p. 499.
DEDERICHS, P.H., 1969, Phys. Rev. **188**, 1175.
DEDERICHS, P.H., 1973, J. Phys. **F3**, 471.
DEDERICHS, P.H., and R. ZELLER, 1980, Dynamical Properties of Point Defects in Metals, in: Point Defects in Metals II, eds. G. Hohler, E.A. Niekisch, Springer Tracts in Modern Physics (Springer, Berlin).
DEDERICHS, P.H., C. LEHMANN, H.R. SCHOBER, A. SCHOLZ and R. ZELLER, 1978, J. Nucl. Mater. **69/70**, 176.
DELAFOND, A., A. NAUDON and C. TEMPLER, 1981, Nucl. Instr. and Meth. **182/183**, 379.
DIETRICH, S., and W. FENZL, 1989a, Phys. Rev. **B39**, 8873.
DIETRICH, S., and W. FENZL, 1989b, Phys. Rev. **B39**, 8900.
DOSCH, H., 1992, Int. J. Modern Phys. **B6**, 2773.
DOSCH, H., 1993, Critical Phenomena at Surfaces and Interfaces — Evanescent X-ray and Neutron Scattering, Springer Tracts in Modern Physics, vol. 126 (Springer Verlag, Berlin, Heidelberg).
DOSCH, H., and J. PEISL, 1985, Phys. Rev. **B32**, 623.
DOSCH, H., and J. PEISL, 1986, Phys. Rev. Lett. **56**, 1385.

DOSCH, H., A.v. SCHWERIN and J. PEISL, 1986, Phys. Rev. **B34**, 1654.
DOSCH, H., J. PEISL and B. DORNER, 1987, Phys. Rev. **B35**, 3069.
DOSCH, H., L. MAILÄNDER, H. REICHERT, J. PEISL and R.L. JOHNSON, 1991a, Phys. Rev. **B43**, 13172.
DOSCH, H., T. HÖFER, J. PEISL and R.L. JOHNSON, 1991b, Europhys. Lett. **15**, 527.
DOSCH, H., F. SCHMID, P. WIETHOFF and J. PEISL, 1992, Phys. Rev. **B46**, 55.
DUBEY, PH. A., B. SCHÖNFELD and G. KOSTORZ, 1991, Acta Metall. Mater. **39**, 1161.
DUCASTELLE, F., and F. GAUTIER, 1976, J. Phys. **F6**, 2039.
EHRHART, P., and W. SCHILLING, 1973, Phys. Rev. **B8**, 2604.
EHRHART, P., and U. SCHLAGHECK, 1974, J. Phys. **F4**, 1589.
EHRHART, P., W. SCHILLING and H.-G. HAUBOLD, 1974, Adv. Solid State Phys. **14**, 87.
ENGELMANN, G., A. STEYERL, A. HEIDEMANN, G.KOSTORZ and H. MUGHRABI, 1979, Z. Phys. **B35**, 345.
EPPERSON, J.E., 1979, J. Appl. Crystallogr. **12**, 351.
EPPERSON, J.E., and P. FÜRNROHR, 1983, Acta Cryst. **A39**, 740.
EPPERSON, J.E., P. FÜRNROHR and C. ORTIZ, 1978a, Acta Cryst. **A34**, 667.
EPPERSON, J.E., K.W. GERSTENBERG, D. BERNER, G. KOSTORZ and C. ORTIZ, 1978b, Phil. Mag. **A38**, 529.
FEIDENHANS'L, R., 1989, Surf. Sci. Rep. **10**, 3.
FLOTTMANN, Th., W. PETRY, R. SERVE and G. VOGL, 1987, Nucl. Instrum. Methods **A260**, 165.
FONTAINE, A., 1993, Interaction of X-rays with Matter: X-ray Absorption Spectroscopy, in: Neutron and Synchrotron Radiation for Condensed Matter Studies, vol. I Theory, Instruments and Methods), eds. J. Baruchel, J.L. Hodeau, M.S. Lehmann, J.R. Regnard and C. Schlenker (Les Editions de Physique, Les Ulis, France) p. 323.
FORGAN, E.M., D.McK. PAUL, H.A. MOOK, P.A. TIMMINS, H. KELLER, S. SUTTON and J.S. ABELL, 1990, Nature (London) **343**, 735.
FRATZL, P., and J.L. LEBOWITZ, 1989, Acta Metall. **37**, 3245.
FRATZL, P., and O. PENROSE, 1994, Phys. Rev. **B50**, 3477.
FRATZL, P., F. LANGMAYR and Y. YOSHIDA, 1991, Phys. Rev. **B44**, 4192.
FRATZL, P., Y. YOSHIDA, G. VOGL and H.G. HAUBOLD, 1992, Phys. Rev. **B46**, 11323.
FRATZL, P., F. LANGMAYR and O. PARIS, 1993, J. Appl. Cryst. **26**, 820.
FÜRNROHR, P., J.E. EPPERSON, and V. GEROLD, 1980, Z. Metallkd. **71**, 403.
FURUKAWA, H., 1979, Phys. Rev. Lett. **43**, 136.
FURUKAWA, H., 1985, Adv. Phys. **34**, 703.
FURUSAKA, M., Y. ISHIKAWA, S. YAMAGUCHI and Y. FUJINO, 1986, J. Phys. Soc. Jap. **55**, 2253.
GALOTTO, C.P., P. PIZZI, H. WALTHER, V. ANGELASTRO, N. CERULLO, and G. CHERUBINI, 1976, Nucl. Instr. and Meth. **134**, 369.
GAULIN, B.D., E.D. HALLMAN and E.C. SVENSSON, 1990, Phys. Rev. Lett. **64**, 289.
GEHLEN, P.C., and J.B. COHEN, 1965, Phys. Rev. **139**, A844.
GEORGOPOULOS, P., and J.B. COHEN, 1977a, J. Physique **12**, C7–191.
GEORGOPOULOS, P., and J.B. COHEN, 1977b, Scripta Metall. **11**, 147.
GEORGOPOULOS, P., and J.B. COHEN, 1979, Direct Methods of Analyzing Diffuse Scattering, in: Modulated Structures — 1979, Proc. Conf. Kailua Kona, Hawaii 1979, eds. J.M. Cowley, J.B. Cohen, M.B. Salamon and B.J. Wuensch (AIP, New York) p. 21.
GEORGOPOULOS, P., and J.B. COHEN, 1981, Acta Metall. **29**, 1535.
GEROLD, V., 1954, Z. Metalik. **45**, 593, 599.
GEROLD, V., 1958, Acta Cryst. **11**, 230.
GEROLD, V., 1967, Application of Small-Angle X-ray Scattering to Problems in Physical Metallurgy and Metal Physics, in: Small-Angle X-ray Scattering, Proc. Conf. Syracuse University, 1965, ed. H. Brumberger (Gordon and Breach, New York) p. 277.
GEROLD, V., 1988, Scripta Metall. **22**, 927 and subsequent papers.
GEROLD, V., and J. KERN, 1987, Acta Metall., **35**, 393.
GEROLD, V., and G. KOSTORZ, 1978, J. Appl. Crystallogr. **11**, 376.
GLAS, R., O. BLASCHKO and L. ROSTA, 1992, Phys. Rev. **B46**, 5972.
GLATTER, O., 1982, Data Treatment, in: Small-Angle X-ray Scattering, eds. O. Glatter and O. Kratky, (Academic, London) p. 119; Interpretation, ibid. p. 167.

GÖLTZ, G., A. HEIDEMANN, H. MEHRER, A. SEEGER and D.WOLF, 1980, Phil. Mag. **A41**, 723.
GÖLTZ, G., H. KRONMÜLLER, A. SEEGER, H. SCHEUER and W. SCHMATZ, 1986, Phil. Mag. **A54**, 213.
GOTODA, H., K. OSAMURA, M. FURUSAKA, M. ARAI, J. SUZUKI, P.J. LEE, D.C. LARBALESTIER and Y. MONJU, 1989, Phil. Mag. **B60**, 819.
GOUDEAU, P., A. NAUDON, B. RODMACQ, P. MANGIN and A. CHAMBEROD, 1987, Europhys. Lett. **3**, 269.
GRAGG, J.E., JR., and J.B. COHEN, 1971, Acta Metall. **19**, 507.
GRAGG, J.E., JR., P. BARDHAN and J.B. COHEN, 1971, The "Gestalt" of Local Order, in: Critical Phenomena in Alloys, Magnets and Superconductors, eds. R.E. Mills, E. Ascher and R.I. Jaffee (McGraw–Hill, New York) p. 309.
GRAGG, J.E., M. HAYAKAWA and J.B.COHEN, 1973, J. Appl. Crystallogr. **6**, 59.
GUINIER, A., 1938, Nature (London) **142**, 569.
GUINIER, A., 1939, Ann. Phys. (Paris) [11] **12**, 161.
GUINIER, A., 1956, Theorie et Technique de la Radiocrystallographie (Dunod, Paris).
GUINIER, A., 1963, X-ray Diffraction, transl. P. Lorrain and D.Lorrain (Freeman, San Francisco).
GUINIER, A., and G. FOURNET, 1955, Small-Angle Scattering of X-rays, transl. C.B. Walker (Wiley, New York).
GUNTON, J.D., M. SAN MIGUEL and P.S. SAHNI, 1983, The Dynamics of First-order Phase Transitions, in: Phase Transitions and Critical Phenomena, vol. 8, eds. C. Domb and J.L. Lebowitz (Academic Press, London), p. 267.
GUREVICH, I.I., and L.V. TARASOV, 1968, Low-Energy Neutron Physics, transl. eds. R.I. Sharp and S. Chomet (North-Holland, Amsterdam).
HAEFFNER, D.R., and J.B. COHEN, 1989, Acta Metall. **37**, 2185.
HAEFFNER, D.R., A. WINHOLTZ, JR. and J.B. COHEN, 1988, Scripta Metall. **22**, 1821.
HALPERN, O., and M.H. JOHNSON, 1939, Phys. Rev. **55**, 898.
HARDERS, T.M., and P. WELLS, 1983, J. Phys. **F13**, 1017.
HARDMAN–RHYNE, K.A., and N.F. BERK, 1985, J. Appl. Cryst. **18**, 473.
HAUBOLD, H.-G., 1975, J. Appl. Crystallogr. **8**, 175.
HAUBOLD, H.-G., and D. MARTINSEN, 1978, J. Nucl. Mater. **69/70**, 644.
HAYAKAWA, M., and J.B. COHEN, 1975, Acta Cryst. **A31**, 635.
HAYAKAWA, M., P. BARDHAN and J.B. COHEN, 1975, J. Appl. Crystallogr. **8**, 87.
HEIMING, A., W. PETRY, J. TRAMPENAU, M. ALBA, C. HERZIG, H.R. SCHOBER and G. VOGL, 1991a, Phys. Rev. **B43**, 10948.
HEIMING, A., W. PETRY, G. VOGL, J. TRAMPENAU, H.R. SCHOBER, J. CHEVRIER and O. SCHÄRPF, 1991b, Z. Phys. B (Condensed Matter) **85**, 239.
HEIMING, A., W. PETRY, J. TRAMPENAU, W. MIEKELEY, and J. COCKCROFT, 1992, J. Phys.: Condens. Matter **4**, 727.
HENDRICKS, R.W., J. SCHELTEN and G. LIPPMANN, 1977, Phil. Mag. **36**, 907.
HICKS, T.J., 1979, Magnetic Inhomogeneities, in: Neutron Scattering, ed. G. Kostorz (Academic, New York) p. 337.
HILFRICH, K., TH. EBEL, W. PETRY, O. SCHÄRPF and E. NEMBACH, 1991, Scripta Metall. Mater. **25**, 1857.
HOUSKA, C.R., 1980, The Investigation of Composition Variations by Diffraction, in: Treatise on Materials Science and Technology, vol. 19 (Experimental Methods, Part A) ed. H. Herman (Academic, New York) p. 63.
ICE, G.E., C.J. SPARKS, A. HABENSCHUSS and L.B. SHAFFER, 1992, Phys. Rev. Lett. **68**, 863.
International Tables for Crystallography, vol. I (1952), vol. II (1959), vol. III (1962), ed. K. LONSDALE, vol. IV (1974), eds. J.A. IBERS and W.C. HAMILTON (Kynoch Press, Birmingham).
JAMES, M.R., and J.B. COHEN, 1980, The Measurement of Residual Stresses by X-ray Diffraction Techniques, in: Treatise on Materials Science and Technology, vol. 19 (Experimental Methods, Part A), ed. H. Herman (Academic New York) p. 2.
JAMES, R.W., 1963, The Optical Principles of the Diffraction of X-rays, 3rd Ed. (Bell, London).
JANOT, CHR., and B. GEORGE, 1985, J. Phys. Lett. **46**, L85.
JOHNSON, W.C., T.A. ABINANDANAN and P.W. VOORHEES, 1990, Acta Metall. **38**, 1349.
KASHIHARA, Y., K. KAWAMURA, N. KASHIWAGURA and J. HARADA, 1987, Jpn. J. Appl. Phys. **26**, L1029.

KETTUNEN, P.O., T. LEPISTÖ, G. KOSTORZ, and G. GÖLTZ, 1981, Acta Metall. **29**, 969.
KHANNA, R., 1989, Acta Cryst. **A45**, 353.
KLAIBER, F., B. SCHÖNFELD, G. KOSTORZ, 1987, Acta Cryst. **A43**, 525.
KOGA, K., K.-I. OHSHIMA and N. NIIMURA, 1993a, Phys. Rev. **B47**, 5783.
KOGA, K., and K.-I. OHSHIMA and N. NIMURA, 1993b, J. Appl. Phys. **73**, 5457.
KOHN, W., 1959, Phys. Rev. Lett. **2**, 393.
KOMURA, S., K. OHSHIMA, H. FUJI, and T. TAKEDA, 1984, Phys. Rev. **B30**, 2944.
KOMURA, S., K. OHSHIMA, H. FUJI, and T. TAKEDA, 1985, Phys. Rev. **B31**, 1278.
KOO, Y.N., and J.B. COHEN, 1989, Acta Metall. Mater. **37**, 1295.
KOO, Y.M., J.B. COHEN, S.M. SHAPIRO, and L.E. TANNER, 1988, Acta Metall. **36**, 591.
KOSTORZ, G., 1976, Z. MetalBlc. **67**, 704.
KOSTORZ, G., 1979, Small-Angle Scattering, in: Neutron Scattering, ed. G. Kostorz (Academic, New York) p. 227.
KOSTORZ, G., 1982, Inorganic Substances, in: Small-Angle X-ray Scattering, eds. O. Glatter and O. Kratky (Academic, London) p. 467.
KOSTORZ, G., 1983, Physica **120B**, 387.
KOSTORZ, G., 1986, A Comparison of Neutrons and Synchrotron Radiation for Defect Studies, in: Atomic Transport and Defects in Metals by Neutron Scattering, eds. C. Janot, W. Petry, D. Richter and T. Springer (Springer, Berlin) p. 1.
KOSTORZ, G., 1988a, Makromol. Chem., Makromol. Symp. **15**, 131.
KOSTORZ, G., 1988b, Mater. Sci. Forum **27/28**, 325.
KOSTORZ, G., 1988c, Experimental Studies of Ordering and Decomposition Processes in Alloys, in: Dynamics of Ordering Processes in Condensed Matter, eds. S. Komura and H. Furukawa (Plenum, New York, London) p. 199.
KOSTORZ, G., 1991, J. Appl. Cryst. **24**, 444.
KOSTORZ, G., 1992, Scattering from Structural Defects, in: Industrial and Technological Applications of Neutrons, eds. M. Fontana, F. Rustichelli, R. Coppola (Soc. Italiana di Fisica, Bologna) p. 85.
KOSTORZ, G., 1993, Physica Scripta **T49**, 636.
KOSTORZ, G., 1994, Decomposition in Alloys — Some Recent Experiments, in: Solid–Solid Phase Transformations, eds. W.C. Johnson and J.M. Howe (The Minerals, Metals and Materials Society, Warrendale, Penn.) p. 133.
KOSTORZ, G., and S.W. LOVESEY, 1979, Neutron Scattering — General Introduction, in: Neutron Scattering, ed. G. Kostorz (Academic, New York) p. 1.
KOUVEL, J.S., 1961, J. Phys. Chem. Solids **21**, 57.
KRIVOGLAZ, M.A., 1969, Theory of X-ray and Thermal Neutron Scattering by Real Crystals, translated from Russian, ed. S.C. Moss (Plenum, New York).
KRONMÜLLER, H., A. SEEGER and M. WILKENS, 1963, Z. Phys. **171**, 291.
KYOBU, J., Y. MURATA and M. MORINAGA, J. Appl. Crystallogr. **27**, 772.
LAMPARTER, P., S. STEEB, D.M. KROEGER and S. SPOONER, 1988, Mater. Sci. Eng. **97**, 227.
LANDER, G.H., and P.J. BROWN, 1985, J. Phys. **C18**, 2017.
LANGMAYR, F., P. FRATZL and G. VOGL, 1992, Acta Metall. Mater. **40**, 3381.
LANGMAYR, F., P. FRATZL, G. VOGL and W. MIEKELEY, 1994, Phys. Rev. **B49**, 11759.
LASLAZ, G., G. KOSTORZ, M. ROTH, P. GUYOT and R.J. STEWART, 1977, Phys. Stat. Sol. (a) **41**, 577.
LEBECH, B., J. BERNHARD and T. FRELTOFT, 1989, J. Phys. Condens. Matter **1**, 6105.
LEFÈBVRE, S., F. BLEY, M. BESSIÈRE, M. FAYARD, M. ROTH and J.B. COHEN, 1980, Acta Cryst. **A36**, 1.
LEFÈBVRE, S., F. BLEY, M. FAYARD and M. ROTH, 1981, Acta Metall. **29**, 749.
LEPISTÖ, T., J. YLI-KAUPPILA, P. KETTUNEN and P. HAUTOJÄRVI, 1981, Phys. Stat. Sol. (a) **67**, K93.
LEPISTÖ, T.K., G. KOSTORZ, V.-T. KUOKKALA and P. KETTUNEN, 1991, Mater. Sci. Eng. **A131**, 171.
LEVINE, J.R., J.B. COHEN, Y.W. CHUNG and P. GEORGOPOULOS, 1989, J. Appl. Cryst. **22**, 528.
LIFSHITZ, I.M., and V.V. SLYOSOV, 1961, J. Phys. Chem. Solids **19**, 35.
LINDBERG, V.W., J.D. MCGERVEY, R.W. HENDRICKS and W. TRIFTSHÄUSER, 1977, Phil. Mag. **36**, 117.
LIPSON, H. and W. COCHRAN, 1953, The Determination of Crystal Structures (Bell, London).
LONG, G.G., S. KRUEGER, R.A. GERHARDT and R.A. PAGE, 1991, J. Mater. Res. **6**, 2706.
LORENTZ, R.D., A. BIENENSTOCK and T.I. MORRISON, 1994, Phys. Rev. **B49**, 3172.

LOVESEY, S.W., 1984, Theory of Neutron Scattering from Condensed Matter, 2 vols. (Clarendon, Oxford).
LUDWIG, JR., K.F., G.B. STEPHENSON, J.L. JORDAN-SWEET, J. MAINVILLE, Y.S. YANG and M. SUTTON, 1988, Phys. Rev. Lett. **61**, 1859.
LYON, O., and J.P. SIMON, 1992, J. Condens. Matter **4**, 6073.
LYON, O., and J.P. SIMON, 1994, J. Phys. IV (France) **4**, C3-25.
LYON, O., J.J. HOYT, R. PRO, B.E.C. DAVIS, B. CLARK, D. DE FONTAINE and J.P. SIMON, 1985, J. Appl. Cryst. **18**, 480.
MAIER-LEIBNITZ, H., 1966, Nukleonik **8**, 61.
MALIK, A., B. SCHÖNFELD, G. KOSTORZ, and W. BÜHRER, 1996, to be published.
MANGIN, P., B. RODMACQ and A. CHAMBEROD, 1985, Phys. Rev. Lett. **55**, 2899.
MARET, M., J.P. SIMON and O. LYON, 1992a, J. Phys. (Condensed Matter) **4**, 1241.
MARET, M., J.P. SIMON, B. BOUCHER, R. TOURBOT and O. LYON, 1992b, J. Phys. (Condensed Matter) **4**, 9709.
MARRA, W.C., P. EISENBERGER and A.Y. CHO, 1979, J. Appl. Phys. **50**, 6927.
MARTIN, J.E., and A.J. HURD, 1987, J. Appl. Cryst. **20**, 61.
MASANSKII, I.V., V.I. TOKAR and T.A. GRISCHENKO, 1991, Phys. Rev. **B44**, 4647.
MATSUBARA, E., and J.B. COHEN, 1985, Acta Metall. **33**, 1945.
MESSOLORAS, S., 1974, Ph.D. thesis, Univ. of Reading.
METCALFE, E., and J.A. LEAKE, 1975, Acta Metall. **23**, 1135.
METZGER, H., and H. PEISL, 1978, J. Phys. **F8**, 391.
METZGER, H., J. PEISL, and J. WANAGEL, 1976, J. Phys. **F6**, 2195.
METZGER, H., H. BEHR and J. PEISL, 1982, Z. Phys. **B46**, 295.
MEZEI, F., ed., 1980, Neutron Spin Echo, Lecture Notes in Physics, vol. 128 (Springer, Berlin).
MILCHEV, A., D.W. HEERMANN and K. BINDER, 1988, Acta Metall. **36**, 377.
MITSUDA, S., H. YOSHIZAWA, T. WATANABE, S. ITOH, Y. ENDOH and I. MIREBEAU, 1991, J. Phys. Soc. Jpn **60**, 1721.
MITSUDA, S., H. YOSHIZAWA, and Y. ENDOH, 1992, Phys. Rev. **B45**, 9788.
MORINAGA, M., K. SONE, T. KAMIMURA, K. OTHAKA and N. YUKAWA, 1988, J. Appl. Cryst. **21**, 41.
MOSS, S.C., 1969, Phys. Rev. Lett. **22**, 1108.
MOSS, S.C., and P.C. CLAPP, 1968, Phys. Rev. **171**, 764.
MOSS, S.C., and R.H. WALKER, 1975, J. Appl. Crystallogr. **8**, 96.
MOSTOLLER, M., R.M. NICKLOW, D.M. ZWEHNER, S.C. LUI, J.M. MUNDENAR and E.W. PLUMMER, 1989, Phys. Rev. **B40**, 2856.
MOZE, O., E.J. LINDLEY, B.D. RAINFORD and D. McK.PAUL, 1985, J. Magn. Mater. **53**, 167.
MOZER, B., D.T. KEATING and S.C. MOSS, 1968, Phys. Rev. **175**, 868.
MUGHRABI, H., T. UNGÁR, W. KIENLE and M. WILKENS, 1986, Phil. Mag. **A53**, 793.
MÜLLER, P.P., B. SCHÖNFELD, G. KOSTORZ and W. BÜHRER, 1988, Scripta Metall. **22**, 959.
MÜLLER, P.P., B. SCHÖNFELD, G. KOSTORZ and W. BÜHRER, 1989, Acta Metall. **37**, 2125.
MURANI, A., 1976, Phys. Rev. Lett. **37**, 450.
NAUDON, A., M. LEMONNIER and F. ROUSSEAUX, 1979, C.R. Acad. Sci. (Paris) **288**, 21.
NAUDON, A., P. GOUDEAU, A. HALIMAOUI, B. LAMBERT and B. BOMCHIL, 1994, J. Appl. Phys. **75**, 780.
NICKLOW, R.M., 1979, Phonons and Defects, in: Neutron Scattering, ed. G. Kostorz (Academic, New York) p. 191.
NICKLOW, R.M., W.P. CRUMMET and J.M. WILLIAMS, 1979, Phys. Rev. **B20**, 5034.
NOYAN, I.C., and J.B. COHEN, 1987, Residual Stress — Measurement by Diffraction and Interpretation (Springer, Berlin).
OCHIAI, S., and M. UENO, 1988, J. Jpn. Inst. Met. **52**, 157.
OETTEL, H., 1989, Methodical Aspects of X-ray Diffraction Analysis of Surface Treated Materials, in: X-ray and Neutron Structure Analysis in Materials Science, ed. J. Hašek (Plenum, New York) p. 171.
OHSHIMA, K.-I., and J. HARADA, 1986, Acta Cryst. **B42**, 436.
OHSHIMA, K.-I., D. WATANABE and J. HARADA, 1976, Acta Cryst. **A32**, 883.
OKUDA, H., M. TANAKA, K. OSAMURA, and Y. AMEMYA, 1993, Acta Metall. Mater. **41**, 1733.

OSAMURA, K., R. SUZUKI, and Y. MURAKAS, 1982, SAXS Study on the Structure of Amorphous Metallic Alloys, in: Proc. 4th Int. Conf. on Rapidly Quenched Metals, Sendai 1981 (Japan. Inst. Met., Sendai) p. 431.

OSAMURA, K., S. OCHIAI and S. TAKAYAMA, 1984, J. Mater. Sci. **19**, 1917.

OSAMURA, K., H. OKUDA, M. TANAKA, M. NAGAO and Y. AMEMYA, 1993, Experimental Study of Phase Decomposition Process in Al–Li Alloy, in: Kinetics of Ordering Transformations in Metals, eds. H. Chen and V. K. Vasudevan (The Minerals, Metals & Materials Society, Warrendale, Ohio) p. 291.

PAGE, R., J. R. WEERTMANN and M. ROTH, 1982, Acta Metall. **30**, 1357.

PALACIOS, J., D. SCHWAHN and H. RAUCH, 1981, NDT International, August 1981, 189.

PARETTE, G. 1991, J. Appl. Phys. **69**, 6135.

PARIS, O., P. FRATZL, F. LANGMAYR, G. VOGL and H. G. HAUBOLD, 1994, Acta Metall. Mater. **42**, 2019.

PARKER, B. A., 1972, J. Appl. Crystallogr. **5**, 372.

PEDERSEN, J. S., 1993, Phys. Rev. **B47**, 657.

PEISL, H., 1976, J. Physique **37**, C7–47.

PEISL, H., R. BALSER and H. PETERS, 1974, Phys. Lett. **46A**, 263.

PENFOLD, J., 1992, The Application of Specular Neutron Reflection to the Study of Surfaces and Interfaces, in: Industrial and Technological Applications of Neutrons, eds. M. Fontana, F. Rustichelli and R. Coppola (Soc. Italiana di Fisica, Bologna) p. 303.

PETRY, W. G. VOGL, H. HEIDEMANN and K. H. STEINMETZ, 1987, Phil. Mag. **A55**, 183.

PETRY, W., T. FLOTTMANN, A. HEIMING, J. TRAMPENAU, M. ALBA and G. VOGL, 1988, Phys. Rev. Lett. **61**, 722.

PETRY, W., A. HEIMING, J. TRAMPENAU and G. VOGL, 1989, On the Diffusion Mechanism in the bcc Phase of the Group 4 Metals, in: Proc. of the International Conference on Diffusion in Metals and Alloys, eds. F. J. Kedves and D. L. Beke (TransTech, Aedermannsdorf, Switzerland).

PETRY, W., A. HEIMING, J. TRAMPENAU, M. ALBA, C. HERZIG, H. R. SCHOBER and G. VOGL, 1991a, Phys. Rev. **B43**, 10933.

PETRY, W., E. BARTSCH, F. FUJARA, M. KIEBEL, H. SILLESCU and B. FARAGO, 1991b, Z. Phys. B (Condensed Matter) **83**, 175.

POERSCHKE, R., W. WAGNER and H. WOLLENBERGER, 1986, J. Phys. **F16**, 421.

POLAT, S., C. MARSH, T. LITTLE, C. P. JU, J. E. EPPERSON and H. CHEN, 1986, Scripta Metall. **20**, 1759.

POROD, G., 1951, Kolloid-Z. **124**, 83.

POROD, G., 1982, General Theory, in: Small-Angle X-ray Scattering, eds. O. Glatter and O. Kratky (Academic, London) p. 17.

PRIEM, T., B. BENEU, C. H. DE NOVION, J. CHEVRER, F. LIVET, A. FINEL and S. LEFÈBVRE, 1989a, Physica B **156/157**, 47.

PRIEM, T., B. BENEU, C. H. DE NOVION, A. FINEL and F. LIVET, 1989b, J. Phys. **F50**, 2217.

PRIESMEYER, H. G., J. LARSEN and K. MEGGERS, 1994, J. Neutron Res. **2**, 31.

QUINTANA, J. P., and J. B. COHEN, 1993, J. Mater. Res. **8**, 855.

RADHAKRISHNA, P., H. M. GILDER, G. PARETTE and A. MENELLE, 1989, Phys. Rev. **B40**, 2435.

RAMESH, T. G., and S. RAMESESHAN, 1971, Acta Cryst. **A27**, 569.

REINHARD, L., and S. C. MOSS, 1994, Ultramicroscopy **52**, 223.

REINHARD, L., B. SCHÖNFELD, G. KOSTORZ and W. BÜHRER, 1990, Phys. Rev. **B41**, 1727.

REINHARD, L., J. L. ROBERTSON, S. C. MOSS, G. E. ICE, P. ZSCHACK and C. J. SPARKS, 1992, Phys. Rev. **B45**, 2662.

REINHARD, L., B. SCHÖNFELD, G. KOSTORZ and W. BÜHRER, 1993, Z. Metallkd. **84**, 251.

RHYNE, J. J., and G. E. FISH, 1985, J. Appl. Phys. **57**, 3407.

ROBINSON, I. K., 1986, Phys. Rev. **B33**, 3830.

RODMACQ, B., Ph. MANGIN and A. CHAMBEROD, 1984, Phys. Rev. **B30**, 6188.

ROELOFS, H., B. SCHÖNFELD, G. KOSTORZ and W. BÜHRER, 1995, Phys. Stat. Sol. (b) **187**, 31.

ROSMAN, R., and M.Th. REKVELDT, 1991, Phys. Rev. **B44**, 4437.

ROTH, M., 1977, J. Appl. Crystallogr. **10**, 172.

ROWE, J. R. and A. MAGERL, 1980, Phys. Rev. **B21**, 1706.

RULAND, W., 1971, J. Appl. Cryst. **4**, 70.

SALVA–GHILARDUCCI, A., J.P. SIMON, P. GUYOT and I. ANSARA, 1983, Acta Metall. **31**, 1705.
SASAKI, S., 1989, Numerical Tables of Anomalous Scattering Factors Calculated by the Cromer and Liberman Method, KEK Report 88–14 (KEK, Tsukuba, Japan).
SASAKI, S., 1990, X-ray Absorption Coefficients of the Elements (Li to Bi, U), KEK Report 90–16 (KEK, Tsukuba, Japan).
SATO, T., T. MURAKAMI, Y. AMEMYA, H. HASHIZUME and T. TAKAHASHI, 1988, Acta Metall. **36**, 1335.
SCHAAL, M., P. LAMPARTER and S. STEEB, 1986, Z. Naturforsch. **41a**, 1123.
SCHMATZ, W., 1973, X-ray and Neutron Scattering Studies on Disordered Crystals, in: Treatise on Materials Science and Technology, vol. 2, ed. H. Herman (Academic, New York) p. 105.
SCHMATZ, W., 1975, Riv. Nuovo Cim. **5**, 398.
SCHMATZ, W., 1978, Disordered Structures, in: Neutron Diffraction, ed. H. Dachs (Springer, Berlin) ch. 5.
SCHMATZ, W., P.H. DEDERICHS and H. SCHEUER, 1974, Z. Phys. **270**, 337.
SCHMIDT, P.W., 1982, J. Appl. Cryst. **15**, 567.
SCHMIDT, P.W., 1991, J. Appl. Cryst. **24**, 414.
SCHNEIDER, J.R., 1981, Nucl. Sci. Appl. **1**, 227.
SCHOBER, H.R., and W. PETRY, 1993, Lattice Vibrations, in: Materials Science and Technology, vol. 1, Structure of Solids, ed. V. Gerold (VCH, Weinheim) p. 289.
SCHÖNFELD, B., 1994, Short-Range Order and Pair Interactions in Binary Nickel Alloys, in: Statics and Dynamics of Alloy Phase Transformations, eds. P.E.A. Turchi and A. Gonis (Plenum, New York) p. 175.
SCHÖNFELD, B., 1995, to be published.
SCHÖNFELD, B., F. KLAIBER, G. KOSTORZ, U. ZAUNE and G. MCINTYRE, 1986, Scripta Metall. **20**, 385.
SCHÖNFELD, B., L. REINHARD, G. KOSTORZ and W. BÜHRER, 1988, Phys. Stat. Sol (b) **148**, 457.
SCHÖNFELD, B., J. TRAUBE, and G. KOSTORZ, 1992a, Phys. Rev. **B45**, 613.
SCHÖNFELD, B., A. GÖCMEN and G. KOSTORZ, 1992b, J. Phys. I (France) **2**, 1075.
SCHÖNFELD, B., G.E. ICE, C.J. SPARKS, H.-G. HAUBOLD, W. SCHWEIKA and L.B. SHAFFER, 1994, Phys. Stat. Sol. (b) **183**, 79.
SCHUBERT, U., H. METZGER and J. PEISL, 1984, J. Phys. **F14**, 2457.
SCHWAHN, D., and W. SCHMATZ, 1978, Acta Metall. **26**, 1571.
SCHWANDER, P., B. SCHÖNFELD and G. KOSTORZ, 1992, Phys. Stat. Sol. (b) **172**, 73.
SCHWARTZ, L.H., and J.B. COHEN, 1987, Diffraction from Materials, 2nd ed. (Springer, Berlin).
SCHWEIKA, W., and H.-G. HAUBOLD, 1988, Phys. Rev. **B37**, 9240.
SEEGER, A.K., 1959, J. Appl. Phys. **30**, 629.
SEEGER, A., 1961, Phys. Stat. Sol. **1**, 670.
SEEGER, A., and E. KRÖNER, 1959, Z. Naturf. **A14**, 74.
SEEGER, A., and M. RÜHLE, 1963, Ann. Phys. (Leipzig) **11**, 216.
SEQUEIRA, A., H. CALDERON, J.S. PEDERSEN and G. KOSTORZ, 1995, Acta Metall. Mater. **43**, 3427; ibid., 3441.
SHAPIRO, S.M., E.C. SVENSSON, C. VETTIER and B. HENNION, 1993, Phys. Rev. **B48**, 13223.
SIMON, J.P., and O. LYON, 1989, J. de Chim. Physique **86**, 1523.
SIMON, J.P., and O. LYON, 1994, Anomalous Small Angle X-ray Scattering in Materials Science, in: Resonant and Anomalous X-ray Scattering, eds. G. Materlik, C.J. Sparks and K. Fischer (Elsevier, Amsterdam).
SIMON, J.P., P. GUYOT and A. GHILARDUCCI DE SALVA, 1984, Phil. Mag. **49**, 151.
SIMON, J.P., O. LYON, and D. DE FONTAINE, 1985, J. Appl. Cryst. **18**, 230.
SIMON, J.-P., O. LYON, F. FAUDOT, L. BOULANGER and O. DIMITROV, 1992a, Acta Metall. Mater. **40**, 2693.
SIMON, J.P., J. MAINVILLE, E. ZIEGLER and O. LYON, 1992b, J. Appl. Cryst. **25**, 785.
SINHA, S.K., E.B. SIROTA, S. GAROFF and H.B. STANLEY, 1988, Phys. Rev. **B38**, 2297.
SKÖLD, M.H. MUELLER and T.O. BRUN, 1979, Hydrogen in Metals, in: Neutron Scattering, ed. G. Kostorz (Academic, New York) p. 423.
SOLAL, F., R. CAUDRON and A. FINEL, 1988, Physica **B156/157**, 75.
SPARKS, C.J., 1974, Phys. Rev. Lett. **33**, 262.
SPARKS, C.J. and B. BORIE, 1966, Methods of Analysis for Diffuse X-ray Scattering Modulated by Local Order and Atomic Displacement, in: Local Atomic Arrangements Studied by X-ray Diffraction, eds. J.B. Cohen and J.E. Hilliard (Gordon and Breach, New York) p. 5.

SPRINGER, T., 1977, Molecular Rotations and Diffusion in Solids, in Particular Hydrogen in Metals, in: Dynamics of Solids and Liquids by Neutron Scattering, eds. S. W. Lovesey and T. Springer (Springer, Berlin, Heidelberg, New York) p. 255.

STAUNTON, J. B., D. D. JOHNSON and F. PINSKI, 1990, Phys. Rev. Lett. **65**, 1259.

STEINER, D., R. BEDDOE, V. GEROLD, G. KOSTORZ and R. SCHMELCZER, 1983, Scripta Metall. **17**, 733.

STERN, E. A., 1974, Phys. Rev. **B10**, 3027.

SUNDARARAMAN, M., W. CHEN, R. P. WAHI, A. WIEDENMANN, W. WAGNER and W. PETRY, 1992, Acta Metall. Mater. **40**, 1023.

SUZUKI, H., J. HARADA, T. NAKASHIMA and K. ADACHI, 1982, Acta Cryst. **A38**, 522.

SWALLOWE, G. M., and J. C. OSBORN, 1992, Physica B **180/181**, 1039.

TAKEDA, T., S. KOMURA, T. MIYAZAKI, Y. ENDOH and S. ITOH, 1987, J. Magn. Mater. **70**, 431.

TANNER, L. E., D. SCHRYVERS and S. M. SHAPIRO, 1990, Mater. Sci. Eng. **A127**, 205.

TEIXEIRA, J., 1988, J. Appl. Cryst. **21**, 781.

TIBALLS, J. E., 1975, J. Appl. Crystallogr. **8**, 111.

TIETZE, H., M. MÜLLNER and B. RENKER, 1984, J. Phys. **C17**, L529.

TRAMPENAU, J., A. HEIMING, W. PETRY, M. ALBA, C. HERZIG, W. MIEKELEY and H. R. SCHOBER, 1991, Phys. Rev. **B43**, 10963.

TRINKAUS, H., 1972, Phys. Stat. Sol. (b) **51**, 307.

TRINKAUS, H., 1975, Z. Naturf. **28a**, 980.

TSUNODA, Y., and J. W. CABLE, 1992, Phys. Rev. **B46**, 930.

TURCHI, P. E. A., G. M. STOCKS, W. H. BUTLER, D. M. NICHOLSON and A. GONIS, 1988, Phys. Rev. **B37**, 5982.

TURCHI, P. E. A., M. SLUITER, F. J. PINSKI, D. D. JOHNSON, D. M. NICHOLSON, G. M. STOCKS and J. B. STAUNTON, 1991a, Phys. Rev. Lett. **67**, 1779; 1992, errata, ibid. **68**, 418.

TURCHI, P. E. A., M. SLUITER, and G. M. STOCKS, 1991b, A Comparative Study of Short Range Order in Fe–Cr and Fe–V Alloys around Equiatomic Composition, in: High-Temperature Ordered Intermetallic Alloys IV, eds. L. A. Johnson, D. P. Pope and J. O. Stiegler (Materials Research Soc. Sym. Proc. 213, MRS, Pittsburgh, Ohio) p. 75.

UNGÁR, T., H. MUGHRABI, D. RÖNNPAGEL and M. WILKENS, 1984, Acta Metall. **32**, 333.

V. GUERARD, B., D. GRASSE and J. PEISL, 1980, Phys. Rev. Lett. **44**, 262.

VOGL, G., W. PETRY, TH. FLOTTMANN and A. HEIMING, 1989, Phys. Rev. **B39**, 5025.

VRIJEN, J., and S. RADELAAR, 1978, Phys. Rev. **B17**, 409.

VYSKOCIL, P., 1994, Decomposition and Coarsening in Ni–Ti (Dr. sc. nat. Dissertation, ETH Zürich).

WAGNER, C., 1961, Z. Elektrochem. **65**, 581.

WAGNER, C. N. J., 1966, Analysis of the Broadening and Changes in Position of Peaks in an X-ray Powder Pattern, in: Local Atomic Arrangement Studied by X-ray Diffraction, eds. J. B. Cohen and J. E. Hilliard (Gordon and Breach, New York) p. 219.

WAGNER, R., and R. KAMPMANN, 1992, Homogeneous Second Phase Precipitation, in: Materials Science and Technology, vol. 5, Phase Transformations, ed. P. Haasen (VCH, Weinheim) p. 213.

WAGNER, W., R. POERSCHKE and H. WOLLENBERGER, 1984, Thermal Decomposition in Cu–Ni–Fe Alloys — III. Neutron Diffraction Study, in: Decomposition of Alloys: the early stages, eds. P. Haasen, V. Gerold, R. Wagner, M. F. Ashby (Pergamon, Oxford) p. 170.

WALKER, C. B., and A. GUINIER, 1953, Acta Metall. **1**, 568.

WALTHER, H., and PIZZI, 1980, Small Angle Neutron Scattering for Nondestructive Testing, in: Research Techniques in Nondestructive Testing, vol. IV, ed. R. S. Sharpe (Academic, London) p. 341.

WANG, Y., L.-Q. CHEN and A. G. KHACHATURYAN, 1993, Acta Metall. Mater. **41**, 279.

WARREN, B. E., 1969, X-ray Diffraction (Addison–Wesley, Reading, MA).

WELBERRY, T. R., and B. D. BUTLER, 1994, J. Appl. Cryst. **27**, 205.

WERNER, K., W. SCHMATZ, G. S. BAUER, E. SEITZ, A. BARATOFF and H. J. FENZL, 1978, J. Phys. **F8**, L207.

WIEDENMANN, A., W. WAGNER and H. WOLLENBERGER, 1989, Scripta Met. **23**, 603.

WILKENS, M., 1969, Acta Metall. **17**, 1155.

WILKENS, M., 1970, Phys. Stat. Sol. (a) **2**, 359.

WILKENS, M., 1975, J. Appl. Crystallogr. **8**, 191.

WILKENS, M., T. UNGÁR and H. MUGHRABI, 1987, Phys. Stat. Sol. (a) 104, 157.

WILKINS, S., 1970, Phys. Rev. **B2**, 3935.
WILLIAMS, C., R.P. MAY and A. GUINIER, 1993, Small-Angle Scattering of X-rays and Neutrons, in: Materials Science and Technology, vol. 2B, Materials Characterization, ed. E. Lifshin (VCH Verlagsgesellschaft, Weinheim) p. 611.
WILLIAMS, R.O., 1972, ORNL Report No. 4828 (Oak Ridge National Laboratory, Oak Ridge, TN).
WILLIAMS, R.O., 1976, ORNL Report No. 5140 (Oak Ridge National Laboratory, Oak Ridge, TN).
WILLIS, B.T.M., and A.W. PRYOR, 1975, Thermal Vibrations in Crystallography (Cambridge Univ. Press, London).
WOLF, K., H.-J. GUDLADT, H.A. CALDERON and G. KOSTORZ, 1994, Acta Metall. Mater. **42**, 3759.
WONG, P.-Z., and A.J. BRAY, 1988, Phys. Rev. **B37**, 7751.
YALDRAM, K., and K. BINDER, 1991, Z. Phys. B — Condensed Matter **82**, 405.
YAVARI, A.R., K. OSAMURA, H. OKUDA and Y. AMEMYA, 1988, Phys. Rev. **B37**, 7759.
YETHIRAJ, M., H.A. MOOK, G.D. WIGNALL, R. CUBITT, E.M. FORGAN, D.McK. PAUL and T. ARMSTRONG, 1993, Phys. Rev. Lett. **70**, 857.
YOO, M.H., J.C. OGLE, B.S. BORIE, E.H. LEE and R.W. HENDRICKS, 1982, Acta Metall. **30**, 1733.
YOSHIDA, Y., F. LANGMAYR, P. FRATZL and G. VOGL, 1989, Phys. Rev. **B39**, 6395.
ZABEL, H., and I.K. ROBINSON, eds., 1992, Surface X-ray and Neutron Scattering, Springer Proceedings in Physics, vol. 61 (Springer, Berlin).
ZACHARIASEN, W.H., 1945, Theory of X-ray Diffraction in Crystals (Wiley, New York).
ZHAO, G.L., and B.N. HARMON, 1992, Phys. Rev. **B45**, 2818.
ZHU, X.-M., R. FEIDENHANS'L, H. ZABEL, J. ALS-NIELSEN, R. DU, C.P. FLYNN and F. GREY, 1988, Phys. Rev. **B37**, 7157.
ZOGG, H., S. BLUNIER, A. FACH, C. MAISSEN, P. MÜLLER, S. TEODOROPOL, V. MEYER, G. KOSTORZ, A. DOMMANN and T. RICHMOND, 1994, Phys. Rev. **B50**, 10801.
ZOLLIKER, M., W. BÜHRER and B. SCHÖNFELD, 1992, Physica B **180/181**, 303.

Further reading

General

The books referred to in the text and contained in the list of references are recommended for general reading, in particular the books by SCHWARTZ and COHEN and by COWLEY and, more extensively covering real crystals, the book by KRIVOGLAZ. An up-to-date account on theory, methods and applications of neutrons and X-rays in condensed matter research is given in the three volumes Neutron and Synchrotron Radiation for Condensed Matter Studies, edited by J. BARUCHEL, J.L. HODEAU, M.L. LEHMANN, J.R. REYNARD and C. SCHLENKER (Les Editions de Physique, Les Ulis, France, 1993). Volume I is entitled Theory, Instruments and Methods, Volume II Applications to Solid State Physics and Chemistry, Volume III Applications to Soft Condensed Matter and Biology.

Neutron Scattering

In addition to the books referred to, G.L. SQUIRES, Introduction to the Theory of Thermal Neutron Scattering (Cambridge Univ. Press, 1978) may be consulted. Several chapters in G. KOSTORZ, ed., Neutron Scattering, vol. 15, Treatise in Materials Science and Technology (H. HERMAN, series ed., Academic, New York, 1979) contain further details about metallurgical applications. More recent materials studies with neutrons are described in the Enrico Fermi Course proceedings edited by M. FONTANA, F. RUSTICHELLI, R. COPPOLA, Industrial and Technological Applications of Neutrons (Soc. Italiana di Fisica, Bologna, 1992).

Diffuse scattering

Beyond the general coverage of the subject in the book by SCHWARTZ and COHEN [1987], BAUER's [1979] chapter (see references) gives more details on the use of neutrons.

Small-angle scattering

The classical monograph by GUINIER and FOURNET [1955] is complemented by "Small-Angle X-ray Scattering", eds. O. GLATTER and O. KRATKY (Academic, London, 1982) which gives a solid theoretical background and describes many methods and applications primarily for statistically isotopic systems. C. WILLIAMS *et al.* [1993] (see references) report on a broad range of recent X-ray and neutron studies, and neutrons are more extensively covered by KOSTORZ [1979] (see references). Reviews and research papers of the two most recent International Small-Angle Scattering Meetings, 1990 in Leuven and 1993 in Saclay, can be found in J. Appl. Crystallogr. 24 (1991) and in Journal de Physique D (1994/1995), respectively.

CHAPTER 13

INTERFACIAL AND SURFACE MICROCHEMISTRY

E. D. HONDROS

Department of Materials,
Imperial College
London, UK

M. P. SEAH

Division of Materials Metrology
National Physical Laboratory
Teddington, Middlesex, UK

S. HOFMANN

Max-Planck-Institut für Metallforschung
Institut für Werkstoffwissenschaft
Stuttgart, Germany

P. LEJČEK

Institute of Physics
Academy of Sciences of the Czech Republic
Prague, Czech Republic

R. W. Cahn and P. Haasen†, eds.
Physical Metallurgy; fourth, revised and enhanced edition

1. Introduction — The chemistry of interfaces and physical metallurgy

Writing more than a hundred years ago, one of the progenitors of modern physical metallurgy, ROBERTS-AUSTEN [1888] stated: "One thousandth part of antimony converts first rate best selected copper into the worst conceivable". Although it was well known at that time that small amounts of certain elements could provoke remarkable changes in the physical and mechanical properties of metals — for example, the improvement in the strength of iron by carbon and the reduction in the electrical conductivity of copper by impurities — the above refers to the dramatic embrittlement of an otherwise ductile metal, associated with intergranular fractures. This is a cogent illustration of the subject matter in the present chapter, in which we shall be concerned with how the chemical constitution of the various interfaces in polycrystals influences the physical metallurgical properties. Observations on the transformation of ductile metals to intergranularly brittle matter clearly implicate the chemical constitution of the grain boundaries and point to the existence of basic, underlying mechanisms which lead to interfacial regions having highly localized and altered chemistry. In this chapter, the generic term *interfacial microchemistry* denotes the science and practice of those phenomena in polycrystalline materials which give rise to zones of chemical heterogeneity. This may occur from either the equilibrium or the non-equilibrium partitioning of those solutes from solid solutions which are present as alloying elements or as residual impurities.

In metallurgical practice, the term *segregation* is used in its original sense to denote the build-up of concentrations of elements during crystallization from casting. In more recent times, metallurgists studied, by indirect procedures, various effects which implied very local chemical-concentration changes at grain boundaries. The latter concentrations were thought to be the origin of metallurgical problems such as *hot cracking* or *temper-embrittlement* in low alloy steels, as demonstrated by increases in the ductile-to-brittle transition temperature during impact testing. Such indirect observations associated with the presence of certain residual impurities at grain boundaries were commonly referred to as "segregation" in metallurgical circles (MCLEAN [1957]) and although common usage dictates that we shall continue to use this term in the context of interfacial microchemistry, we clearly distinguish between the phenomena which give rise to these types of equilibrium segregation and the phenomena which give rise to the original solidification segregation. Indeed, as will be demonstrated, this redistribution of solutes by *equilibrium segregation* is truly reversible as analyzed by thermodynamic procedures, and it is quite analogous to the chemisorption or adsorption phenomena long studied by surface chemists for gas atoms interacting with free solid surfaces. Although the terms "segregation" and "adsorption" have acquired historically these specific usages in their respective disciplines, in order to emphasize the fact that the basic phenomena involved are identical, we shall apply the terms "segregation" and "adsorption" interchangeably to both the *external interface*, such as the free surface and the *internal interface*, such as the grain boundary.

In addition to grain boundaries which are the principal interfaces with metallurgical consequences, interfacial microchemistry embraces a variety of interfaces in polycrystalline materials, among them free surfaces, the interfaces between various constituent

phases in the solid, stacking faults, and precipitate/matrix interfaces. From a macroscopic point of view of this theme, we underline that all these interfaces behave essentially in a similar way in spite of widely differing, often very complex structures at the atomic level — the same thermodynamic analyses are applicable to all these types of interface, which may be described by the classical concepts of surface adsorption chemistry; indeed similar adsorption isotherms hold, the conditions for the existence of any particular form and the magnitude of the adsorption parameters depending on the nature of the interface alone.

Interfacial microchemical phenomena in polycrystals can be visualized by aid of the schema in fig. 1. The multicomponent solid mass is held at equilibrium in an isothermal enclosure and at a temperature which permits rapid thermal transport. The chemical potentials of all species, assumed here to be in solid solution, are constant throughout the system. Instead of a random distribution of solute species, as in an ideal single crystal, it is found that at equilibrium there is a heterogeneous partitioning which results in the enrichment of interfaces by certain surface-active species, as shown diagrammatically in fig. 1. The levels of the enrichment are defined only by the system parameters at equilibrium and not by the history of the material. The chemical enrichment at any interface can be reproduced simply by re-establishing the identical physicochemical conditions. It is in this sense that the segregation is truly reversible.

An essential condition of equilibrium segregation, whether at a free surface, a grain boundary or precipitate/matrix interface, is that the width of matter over which it builds up is influenced by the structural width of the interface. The field of structural perturbation of a grain boundary should not extend beyond a few atom distances, and it has now been experimentally confirmed that the space occupied by the segregation is constrained to within the structurally perturbed region of the interface (§ 3.2). Hence a distinguishing feature of equilibrium segregation is that it is localized to within one or two atom distances of the plane interface. Thus we speak of a fractional monatomic layer level of segregation at grain boundaries and other interfaces, by which we mean, by analogy with the concepts of monomolecular adsorbed layers in surface chemistry, that a fraction of a single equivalent atomic layer consisting of a sheet of close-packed atoms lies along the interface. Of course this is an idealization, and in practice monatomic layers of this nature probably do not form. However, it is a convenient way of visualizing the extent of segregation and of quantifying it, whether this is expressed as a fraction of a monatomic layer or as several such layers.

In contrast to this equilibrium segregation, various phenomena exist in which the apparent levels of segregation build-up may extend to distances of as much as several micrometers across grain boundaries. This effect is sensitive to the rate of cooling from a high temperature. The earliest effect of this type was announced by Westbrook and his collaborators (WESTBROOK [1964]) in a series of experiments in which they claimed to detect such an effect through changes in indentation microhardness profiles across grain boundaries in certain dilute alloys. Although falling within the scope of a general treatment of interfacial microchemistry, this phenomenon must be distinguished from that involving equilibrium segregation, which is probably the most ubiquitous form of interfacial enrichment in solid systems. The origin of the former segregation is also believed to be entirely different from that involving reversible segregation — namely,

References: p. 1284.

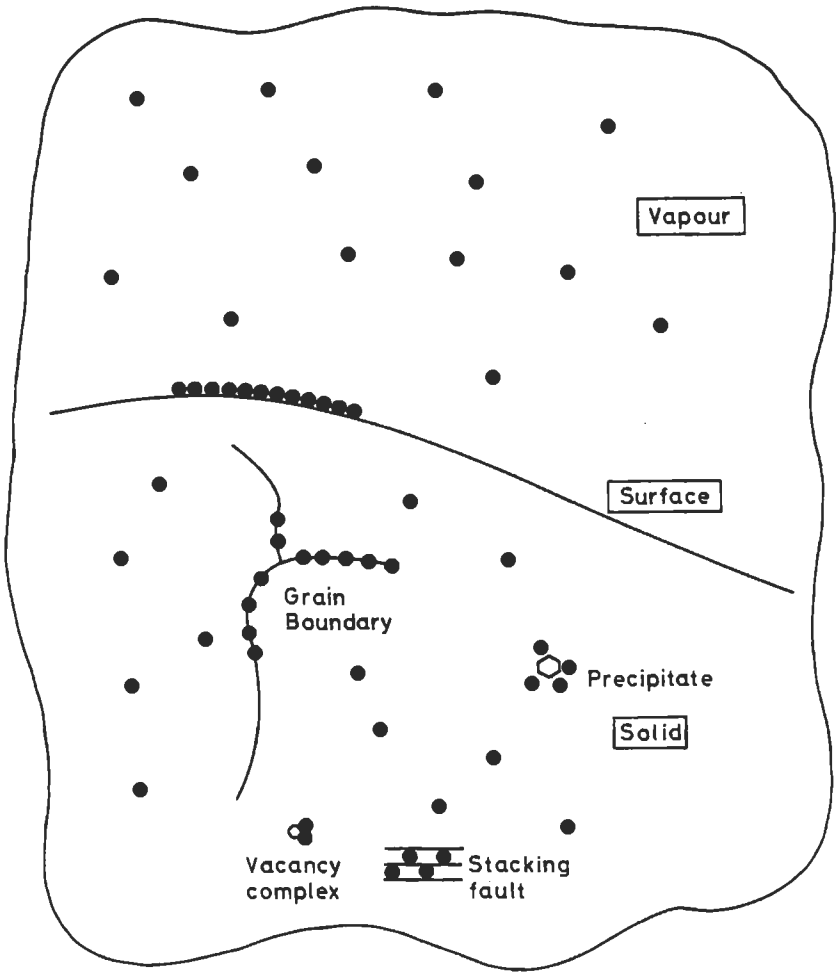


Fig. 1. Schematic representation of crystalline interfaces showing adsorption or microsegregation.

during a temperature fall, vacancies flow to the grain-boundary sinks in order to preserve their thermal equilibrium value. A solute concentration is thus built up about the boundary because of the dragging of vacancy — solute pairs. Although highly system-specific, the phenomenon may be of consequence in certain metallurgical operations and this, together with other related manifestations, will be considered later in this chapter. To distinguish this phenomenon from reversible segregation, we refer to it as *non-equilibrium segregation*.

In order to understand the impact of interfacial microchemistry on physical metallurgical processes, our aim in this chapter is to provide descriptions of the equilibrium and non-equilibrium segregation phenomena in terms of basic physical, chemical and kinetic

parameters, and then to relate such segregation to changes in mechanical properties, embrittlement in particular. Indeed, the manner by which a narrow sub-monatomic layer of a foreign species along a grain-boundary path may so affect bulk mechanical properties as to convert a tough metal into a fragile one, has been the point of departure of many intriguing basic studies. However, it is beyond the scope of this chapter to enlarge in detail on all the mechanical effects that have been scrutinized in recent years in terms of the role of interfacial microchemistry.

2. *Thermodynamic features of interfacial adsorption*

The thermodynamic property possessed by all interfaces in common is a free energy, which is a measure of the unsatisfied atomic bonding in the relatively disordered structure of the interface. In liquids, where it is referred to as *surface tension*, this free energy is easily demonstrated to be able to do work, as testified by many conventional capillarity phenomena, such as droplet formation or the shapes of soap bubbles. Foreign species may accumulate at these interfaces in order to satisfy better the atomic bonding, and hence to reduce the free energy: this is in simple terms the basic motivation for the segregation of foreign atoms at interfaces. It is in this way that soap added to water reduces its surface tension. This free energy is the central thermodynamic function which characterizes all interfaces whether in solids or liquids, and because its magnitude depends on temperature, crystalline orientation and the presence of foreign species, its determining presence is felt in many physical phenomena in nature and in turn in the mechanical and chemical properties of materials.

The fundamental connection between surface free energy and the chemical state of the surface was first demonstrated in the last century by GIBBS [reprinted 1948] in an elegant and generalized treatment. This treatment and the mathematical derivations are shown in detail in a number of textbooks on surface chemistry to which the reader is referred (e.g., DEFAY *et al.* [1966]). In his classical derivation of the fundamental theorem relating the "surface of tension" of a fluid to bulk composition and the excess of species at the surface, Gibbs introduced for convenience the device of a mathematical "dividing surface" which was placed arbitrarily in order to make the surface concentration of the solvent species vanish; however, for internal interfaces such as grain boundaries which were not considered by Gibbs, there are conceptual difficulties in locating this mathematical dividing surface in a manner that corresponds to a physical surface of separation, because the physicochemical properties vary continuously from one homogeneous phase to the adjoining phase. Therefore, in the treatment of adsorption processes at internal interfaces it is now customary to adopt the convenient approach of GUGGENHEIM [1950], in which the interface is considered as a separate phase of finite thickness, and where the thermodynamic relations so derived are independent of the thickness of the phase separating the two adjoining phases which may be two crystals or a crystal in contact with a vapour. For a multicomponent solid, following the approach of § 9 of ch. 5, leading to eq. (69) there, but replacing the surface tension, σ, by the scalar interfacial free energy γ, the Gibbs adsorption isotherm may be derived. In the dilute approx-

imation, which is that of general interest, for a bulk solute molar concentration $X_c \ll 1$, in which Henry's law holds, this becomes:

$$\left(\frac{d\gamma}{d \ln X_c}\right)_T = -RT\Gamma_2, \tag{1}$$

where Γ_2 is the interfacial excess of the solute species expressed in mol/m^2. This is the most useful form of the *Gibbs Adsorption Theorem* for a dilute binary system and in this basic form it has been used experimentally to relate changes in composition of the bulk to the changes in both the free energy and the surface excess composition for all interfaces in solids and in liquids. (See also ch. 5, § 9.1.).

To illustrate the use of the Gibbs Adsorption Theorem in the dilute form in order to measure interfacial excesses, fig. 2 shows for three interfaces in pure iron — the liquid

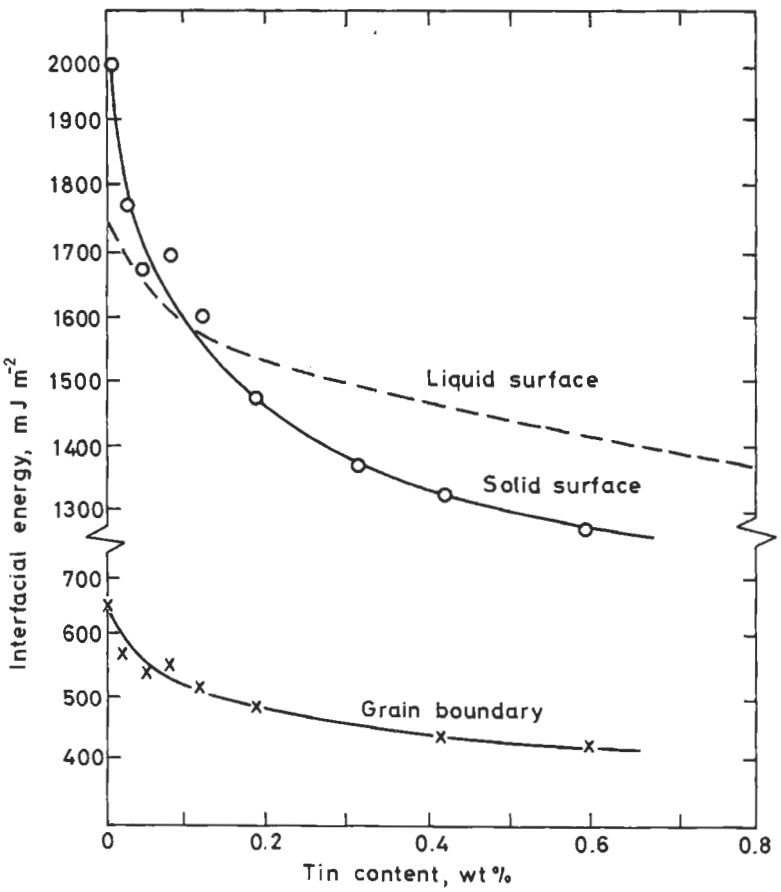


Fig. 2. Isotherms for interfacial free energies in the dilute binary system Fe–Sn: data for the liquid surface from KOZAKEVITCH and URBAIN [1961].

surface, the solid surface and the average grain boundary — the variation in the interfacial free energy as a function of the bulk tin content (SEAH and HONDROS [1973]). The three interfaces behave similarly in the general form of the dependence of interfacial energy on bulk composition, in that there is an early rapid fall in the isotherm, followed by a gradual weakening of the effect as the composition increases. This is typical of the behaviour that has been encountered in many such binary systems studied experimentally. The interfacial tensions for the pure metal are also typical in that the free energy of the solid surface is significantly greater than that of the liquid surface, while the grain-boundary free energy is approximately one third of that of the surface free energy. The data shown in this figure are analyzed by application of eq. (1) to the results pertaining to the specific temperature of the experiment. The amount of interfacial excess expressed by Γ_2 is in terms of moles of the solute species per unit area. In fig. 3 these quantities have been transformed into equivalent monatomic layers, and the surface composition is shown as a function of the bulk tin content for the three interfaces in question. The behaviour of the liquid surface is more ideal, in that it shows the approach towards a

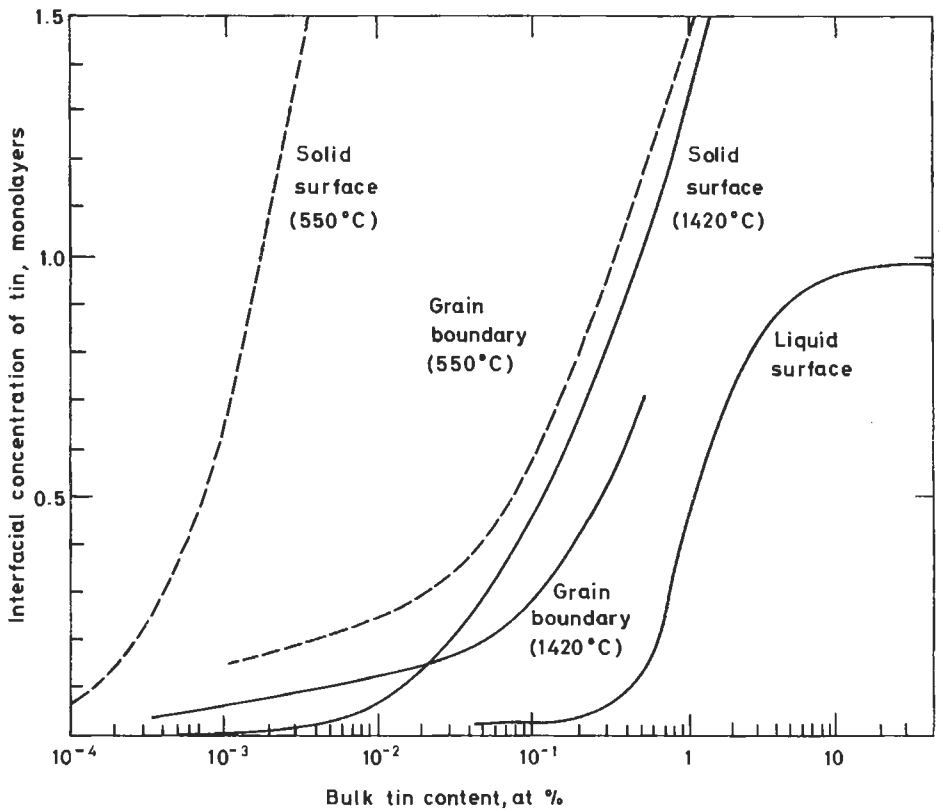


Fig. 3. Adsorption isotherms for tin in iron interfaces: the low-temperature isotherms (dashed) are measured directly by AES: the other are derived from the interfacial energy data shown in fig. 2.

References: p. 1284.

saturation adsorption of tin at the liquid iron surface, at a level of about a single monatomic layer. With increasing bulk content, the interfacial excesses shown for the grain boundary and for the solid surface increase, but in this particular system and in the conditions of this experiment, there is no apparent sign of saturation adsorption. In fact, for the system iron–tin the behaviour in the solid state is not typical in that there occurs multiple layer segregation (SEAH and HONDROS [1973]).

In the case of iron containing dilute amounts of phosphorus, which typifies the behaviour of iron in the presence of a number of species, including sulphur and oxygen, the dependence of the grain-boundary free energy on the bulk phosphorus content is shown in fig. 4 (HONDROS [1965]). From the slopes of the surface tension isotherm at 1723 K, the grain-boundary excesses are calculated and the resulting adsorption isotherm is superimposed on the same figure. At the saturation part of the isotherm, the calculated grain-boundary excess of phosphorus is 0.65×10^{19} atoms/m^2 which corresponds to approximately a third of a monatomic layer of phosphorus at grain boundaries, assuming the phosphorus atoms to be hard spheres spread out evenly and compactly across the interface. This compares with 1.4×10^{19} atoms of phosphorus per m^2 for the saturation coverage of the free surface which, in turn, corresponds to about 0.75 atoms of phosphorus per outermost atom of iron.

As will be shown later, there has been an enormous stimulus to the study of interfacial microchemistry in both binary systems as well as multicomponent engineering alloys as a result of the advent of the modern surface spectroscopy techniques which

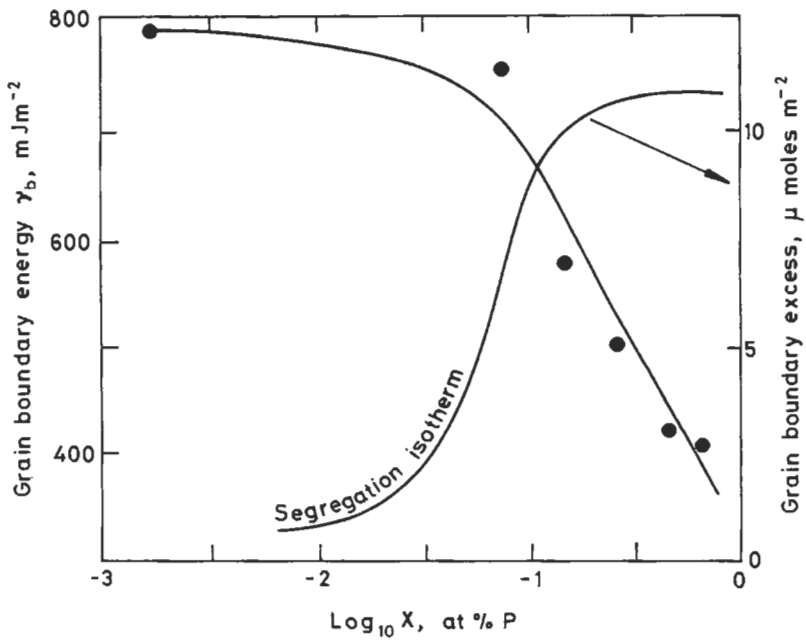


Fig. 4. Grain-boundary energy and segregation isotherms for phosphorus in iron (after HONDROS [1965]).

permit the direct measurement of segregation. In order to build up a body of reliable data, it is useful that the quantities determined by the different approaches be made compatible. Thus we have extended to all interfaces the notion of an *Interfacial Enrichment Factor*, β_s, defined as the ratio between the *interfacial* concentration, X_s, (in mole fraction of a monatomic layer) and the *bulk solute* concentration, X_c, (expressed as mole fraction). Thus:

$$\beta_s = \frac{X_s}{X_c} = \frac{\Gamma_s}{\Gamma_s^0 X_c}, \tag{2}$$

where Γ_s is the Gibbs excess defined earlier as Γ_2 and expressed as mol/m^2 and Γ_s^0 is the quantity of solute in moles constituting a close packed monatomic layer of unit area. This can be shown to be related to the slope of the surface tension isotherm (eq. (1)) as:

$$\beta_s = -\frac{1}{RT\Gamma_s^0}\left(\frac{d\gamma_s}{dX_c}\right)_T. \tag{3}$$

In this way comparisons can be made between enrichment levels obtained by the Gibbs technique and those obtained directly using surface and interface analysis techniques.

3. *Methods of measuring the microchemistry of interfaces*

There now exists a copious technical literature on the subject of the detection and measurement on interfacial segregation, and one will remark on the wide range of imaginative approaches which were deployed during this century in attempts, mostly qualitative, to detect the suspected contaminants at grain boundaries which gave rise to various metallurgical effects. These mostly indirect techniques include: electrode potential measurements on grain-boundary fracture surfaces; a number of subtle metallographic features observed at grain boundaries; variations in X-ray lattice parameter measurements with different grain sizes; internal friction measurements; micro-indentation hardness measurements; the spectrographic analysis of material carefully extracted from the grain-boundary region; auto-radiography; and deductions made from anomalous grain growth behaviour. From the point of view of furthering this subject matter and offering a deeper insight, these past methods are not recommended to the reader as vital to the pursuit of the theme. The interested reader will find them fully covered in the review by HONDROS [1976], but for most purposes these early techniques can be relegated to studies of the historical evolution of the subject: in one respect it is interesting to read these accounts as a testament to the highly ingenious approaches that the tenacious research worker is forced to adopt in those frequently encountered situations where the available tools are inadequate to satisfy his curiosity and sense of scientific exploration. These methods have been effectively superseded with the development of a range of sophisticated direct surface analysis techniques, based upon a variety of electron and ion emission processes; these and other significant techniques are considered below.

References: p. 1284.

3.1. The interfacial energy or Gibbsian approach

Not only is there a requirement for data on interfacial energies in alloys as quantities in the modelling of various metallurgical processes, but as we saw above, a knowledge of the dependence of the interfacial free energy on bulk composition yields a quantitative measure of the amount of the species concentrated at that interface. This is the classical Gibbsian approach to the problem, and its most serious disadvantage is that since it is an indirect procedure, one requires an *a priori* knowledge of the relevant species procuring the change in the interfacial energy. Thus it finds its greatest applicability in binary systems where one knows in advance the elements involved. This inability to detect chemical species is a serious disadvantage in the study of multicomponent alloys where one must presuppose a knowledge of the surface active species from prior general metallurgical observations.

The demand for relevant interfacial free energy data for the solid state arising, for example, from studies of grain growth or nucleation phenomena, has stimulated a variety of measurement approaches such as heat of solution calorimetry or controlled cleavage of the crystal. Most of these have been found inadequate or intrinsically not convenient for studying adsorption effects. For the measurement of the surface free energy of a metal, the most successful technique, and indeed the only one to have supplied data on Gibbsian segregation, is that described as the *Zero Creep Technique* in which the surface energy is deduced from a knowledge of the critical externally applied force necessary to counter-balance the capillarity forces which tend to shrink, at high temperature, samples with a high surface-to-volume ratio. The idea that at high temperature a solid will measurably shrink in order to reduce its surface area has been used in various forms by investigators, but we note that its main limitation is that because the shape change depends on solid-state diffusion, the technique is limited in practice to temperatures generally above $0.7\ T_m$. The zero creep technique, as well as other measurement techniques, are described in detail by HONDROS [1970].

As in the case of free surfaces, a wide range of techniques has been explored in order to measure grain-boundary free energies. These include isothermal calorimetry but in practice, most techniques are restricted to the determination of the grain-boundary energies for the pure metal. The only systematic data on the compositional dependence of grain-boundary energies have come through the application of polyphase equilibria techniques, which yield a *relative* grain-boundary/surface free energy value. The simple procedure is based on the observation that at high temperatures, the kinetics of mass transport allow intersecting interfaces to assume equilibrium configurations consistent with a minimum of energy, in which the local surface tensions can be analyzed as if in vectorial balance (SMITH [1948]). This idea has been used for a variety of situations and, in the case of grain boundaries, the microtopography of a grain boundary intersecting a free surface (fig. 5) is a distinct groove along the intersection of the boundary with the surface. This configuration indicates a balance in the tensions between the two crystalline surfaces and that of the grain boundary and in these conditions, assuming the surfaces to be isotropic, the equilibrium is given by:

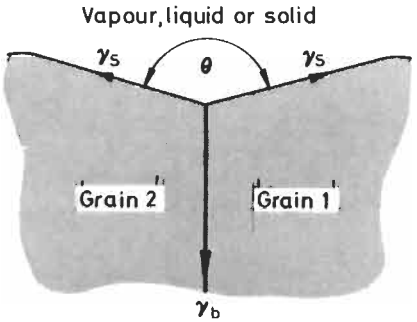


Fig. 5. Equilibrium configuration for a grain boundary intersecting the free surface, showing the dihedral angle.

$$\gamma_b = 2\gamma_s \cos(\theta/2), \tag{4}$$

where θ is the dihedral angle formed by the groove. In practice, the polycrystalline sample is equilibrated in physicochemical conditions which duplicate those in which γ_s, the reference value, is measured by the Zero Creep Technique. After a prolonged period of heating at temperature, which allows the equilibrium geometry to be attained, dihedral angles are measured on about 50 intersecting boundaries using optical interferometry or some other suitable technique. In this way, errors arising from the anisotropy of interfacial energies can be neglected, since it has been shown that eq. (4) is adequate if the grain-boundary/surface energy ratios are averaged statistically over a large number of observations. The method of thermal grooving applied to crystallographically fully characterized bicrystalline samples can be used for determination of the energy of *individual* grain boundaries. This can be documented, for example, by measurements recently performed by WOLF *et al.* [1992] who constructed the orientation dependence of grain-boundary energy for asymmetrical twin boundaries in copper. These measurements revealed an energetical minimum corresponding not to the symmetrical {112} interface as would be expected, but at an asymmetrical boundary deviated 8° from the {112} one.

3.2. Modern surface analysis techniques

The stimulus for much of the present-day work on surface and interface microchemistry arises from the ease of use and interpretation of measurements made with the surface analysis techniques developed in the last decades. To give some idea of the problems involved in this analysis, suppose a typical metal is fractured in air to provide a clean surface. This will be covered by a monolayer of oxygen atoms in a time of the order of one nanosecond. Thus, to study the surface or interface chemistry and reactions over a reasonable timescale, vacua of the order of 10^{-8} Pa (10^{-10} torr) are required to keep the environmental contamination under control. This technical capability became generally available only with the development of ultra-high vacuum (UHV) technology, in the mid-1960s. Thus, it was not until the 1970s that the surface analysis techniques became firmly established as tool for materials studies.

References: p. 1284.

There are four main techniques, Auger electron spectroscopy (AES), X-ray photoelectron spectroscopy (XPS), ion scattering spectroscopy (ISS) and secondary ion mass spectroscopy (SIMS), which have been used to study segregation and each of which provides slightly different information. In general, the techniques are operated in UHV and are used to study the free surface of solid. The solid may be cleaned by argon ion sputtering, and the free surface segregation studied as a function of time, temperature, ambient environment or a combination of these. For the study of grain boundaries, these are exposed by fracture in the UHV of the instrument, using a suitable fracture stage. If the segregant embrittles the grain boundary, this fracture may be achieved fairly readily either at room temperature or by cooling the sample to liquid nitrogen temperatures. In the absence of intrinsic embrittlement, the grain boundary may be parted in some materials by impact after cathodically charging the solid with hydrogen to promote hydrogen embrittlement. Of course, for grain-boundary studies, unlike those on free surfaces, a fresh sample must be prepared for each heat-treatment time and temperature. The relative ease of studying free surface explains, to some extent, the high level of effort devoted to free surfaces compared with that for grain boundaries.

The analysis of the prepared surface in UHV is usually undertaken with the most powerful and popular of the techniques, AES (BRIGGS and SEAH [1990]). In this, a focused electron beam of 1–30 keV energy is used to excite atoms in the surface layer of an appropriate area of the target, creating holes in their inner core energy levels. The atoms subsequently decay by emission of either a characteristic X-ray (which is the basis of bulk chemical analysis) or by the emission of a characteristic Auger electron, which is the basis for surface analysis. We show in fig. 6(i) the inner core level diagram for an element heavier than neon as an example. The initial core hole may be created in level K which is then filled by a higher energy electron from, say, L_3 with the energy balance taken by, say, a second L_3 electron which is ejected from the atom with an energy E_A and is known as an Auger electron. E_A is characteristic of the particular atom and in this case is given approximately by:

$$E_A = E_K - E_{L_3} - E_{L_3}. \tag{5}$$

Fortunately, of the very many transitions that may occur, only a few are strong for each element and these enable the surface atoms to be clearly identified from the emitted electron energy spectrum. Atoms from depths greater than a few monolayers (1 nm) also eject Auger electrons but these do not escape easily and so do not contribute to the emitted line spectrum. Thus AES is characteristic of the outermost atomic layers of a solid and, by focusing the electron beam, we may now localize the composition of heterogeneous surfaces to a resolution across the surface that currently approaches 20 nm. Elements, except for H and He, are generally detected with sensitivities in the range one part in 100 to 1000.

A classical example which illustrates the use of AES is the analysis of the segregation giving rise to the temper brittleness in the 3Cr$\frac{1}{2}$Mo steel rotor which initiated a well known failure at a power station (KALDERON [1972]). The fracture of this material is fully intergranular as shown by the scanning electron micrograph of fig. 7. AES analysis of a sample taken from the vicinity of the original crack, and fractured in the UHV AES

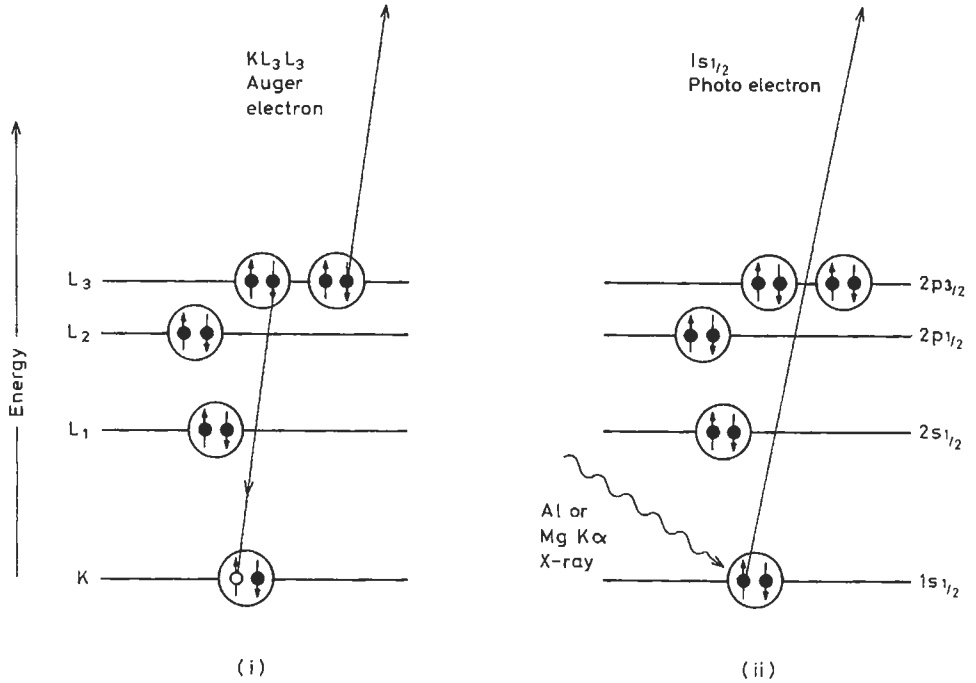


Fig. 6. Schematic transition level diagram for electrons in (i) AES and (ii) XPS.

system, shows 45% of a monatomic layer of phosphorus at the grain boundary. The spectrum shown in fig. 8 also identifies the alloying elements present and the carbides at the grain boundary.

That the segregant is located in the grain-boundary zone may be demonstrated by slowly peeling away the surface atoms, using *in situ* argon ion sputtering and monitoring the surface composition by AES. In this way a composition–depth profile may be obtained for all elements present. Measurements for many segregants at grain boundaries are shown in fig. 9 where it is seen that the results agree with the solid curve predicted on the hypothesis that the segregant atoms are on the outermost atom plane of the fracture surface and hence associated with the precise atom plane of fracture.

A particular attribute of AES is the ability to achieve a fine localization on the free or fracture surface with the focused electron beam. This allows anisotropy effects between different crystals and the segregation at inclusion–matrix interfaces to be analyzed with comparative ease. The present state of the method also enables to detect fine variations in segregants contents at different parts of both matching fracture surfaces of a bicrystal, connected with an asymmetry of the fracture process on the atomic scale (MENYHARD *et al.* [1991]).

The second of the techniques, XPS (BRIGGS and SEAH [1990]) has some advantages

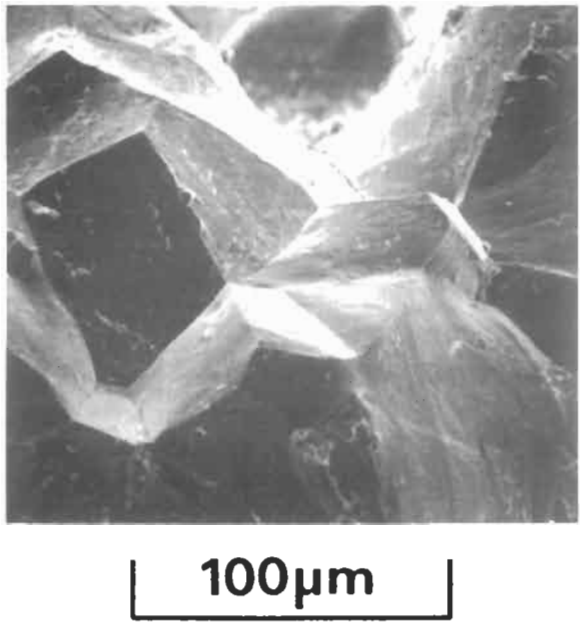


Fig. 7. Scanning electron micrograph of the service failure of a temper-brittle $3Cr\frac{1}{2}Mo$ rotor steel, showing the fully intergranular fracture surface (field of view approximately 200 μm) (from SEAH [1975]).

over AES although it lacks spatial resolution. In XPS, as shown in fig. 6(ii), electrons are ejected from the surface atom core levels by the characteristic X-rays of energy $h\nu$ from a metal target, usually Al or Mg. The *emitted electron energy*, E_x is given in the example of fig. 6(ii) by:

$$E_x = h\nu - E_{1s}. \tag{6}$$

Thus, each of the lines in the emitted photo-electron spectrum directly reflects the core level positions, identifying the emitting atoms which, for the same reasons as given for AES, are those in the outermost 1 nm of the surface. The precise position and shape of the photo-electron line depends on the chemical state of the surface atoms, and so XPS can be used to probe the chemical state of the segregant atoms and to separate out those elements which are in micro-precipitates on the surface. For instance, in fig. 10 details are shown of the 2p (or L_{23}) line for phosphorus on the free (100) surface of an iron single crystal. The various curves show: (1) a true segregant; (2) phosphorus segregated plus Fe_3P precipitates where the phosphorus level is above the solubility limit during the heat treatment; and (3) the dotted curve from solid Fe_3P.

During the past years, a veritable plethora of alternative surface analysis techniques has emerged. Of these, the only one that has been used for segregation studies is ion-scattering spectroscopy (ISS). In ISS (NIEHUS *et al.* [1993]) the mass of atoms in the surface atom layer is determined by recording the energy loss of ions bounced off those atoms as if in a billiard-ball collision. Using a detector at an angle θ to the incident beam

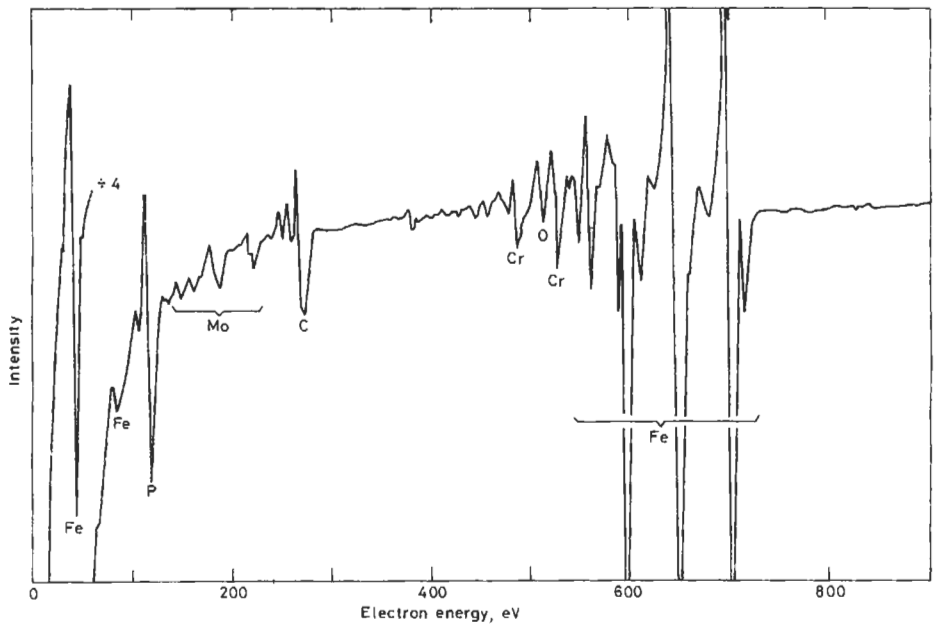


Fig. 8. Auger electron spectrum from the fracture surface of fig. 7, exhibiting 0.45 monolayers of phosphorus segregated at the grain boundary (after SEAH [1975]).

of ions of mass M_i and energy E_i, it is easy to show that surface atoms of mass M will reflect those ions into the detector with an energy E_M where:

$$M = M_i\left[\frac{E_i + 2\sqrt{E_i E_m}\cos\theta + E_m}{E_i - E_M}\right]. \tag{7}$$

This allows the analysis of the surface atoms. However, it will be seen from eq. (7) that high masses are all found bunched in a small energy range near E_i and thus, because of energy blurring terms which give rise to peak broadening, it is difficult to distinguish adjacent masses heavier than copper. ISS has, however, two main advantages over AES and XPS which are very pertinent to studies of surface segregation, especially in applications related to adsorption and catalysis. These are that by using neon ions of 1 keV energy, the analysis is limited to the outermost atom layer only and by suitably directing the beam at a single crystal surface, atom site positions can be determined very accurately by shadowing effects.

SIMS, as a surface analysis technique can well supply useful information. As with ISS, the surface is interrogated by an ion beam with an energy in the range 1–10 keV, with a flux of some nA/cm^2 (BENNINGHOVEN *et al.* [1982], BRIGGS and SEAH [1992]). Secondary ions that are then emitted may be mass analyzed to determine the atom groupings on the surface since these groupings are reflected in the total mass of the clusters detected. Complex atom groupings of twenty or more atoms can be observed which represent bonding arrangements at surfaces (BRIGGS [1982]) and so, in principle,

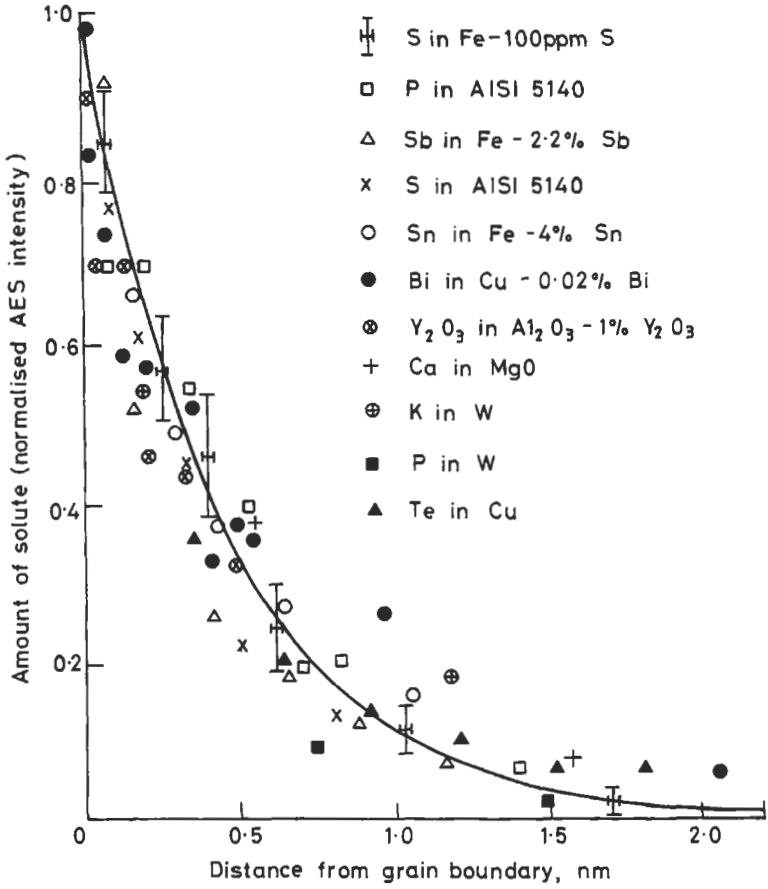


Fig. 9. Localization of segregant atoms at the grain boundary in many systems by AES with argon ion etching (after SEAH [1980a]).

the interactions of co-segregants can be studied. SIMS has other advantages of detecting hydrogen (FUKUSHIMA and BIRNBAUM [1984]) and of being much more sensitive than AES or XPS; however, it is very difficult to quantify the spectra and hence it has not yet become popular as a technique to study segregation. This may change with the development of the detection of "sputtered neutrals mass spectroscopy" (SNMS), using post-ionization of the sputtered matter to effectively reduce matrix effects. Time-of-flight SIMS (TOF–SIMS) is particularly useful for submonolayer detection and is often combined with SNMS (BRIGGS and SEAH [1992]).

3.3. Micrographic techniques

Three micrographic techniques have been used in the study of grain-boundary segregation, and although each has its own unique advantages, as noted below, none can

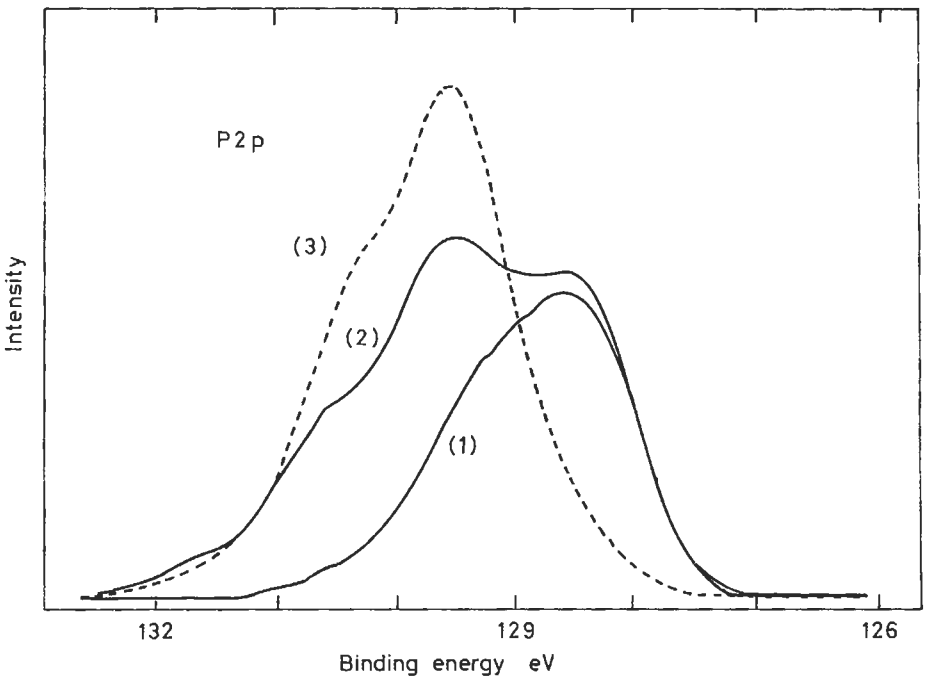


Fig. 10. X-ray photoelectron 2p core line of phosphorus: (1) segregated on Fe (100); (2) segregated with Fe_3P precipitates; (3) from solid Fe_3P (after EGERT and PANZNER [1982]).

be used for the parallel study of surface segregation. These three techniques have the crucial advantage over the techniques discussed above that, in addition to being analytical, they can provide measurements for boundaries that will not undergo fracture. They thus enable a new range of systems to be studied.

The most popular of the micrographic techniques is *ion microprobe analysis* or, as it is sometimes known, *secondary ion microscopy*. In this technique, a mass spectrometer is used to construct emitted-ion images of the surface bombarded by a primary ion beam. The grain boundary is arranged normal to the surface being imaged, so it intersects the surface in a line that does not move as the surface is eroded by the primary ion beam. In this way a trace of the boundary can be imaged using the segregant atoms, with a resolution as low as 20 nm (RÜDENAUER [1994]). Isotopes can be separated, and the elements hydrogen and helium, not observed in AES or XPS, may be clearly imaged.

Another micrographic technique which has unique advantages is *scanning transmission electron microscopy* (STEM) with X-ray analysis (DOIG *et al.* [1981]) or with electron-energy loss spectroscopy (EELS; BRULEY *et al.* [1994]). A thin section of material is prepared with the grain boundary again perpendicular to the surface and the sample is then imaged in the electron microscope. The segregant is measured using a nondispersive X-ray detector with the electron beam on the boundary or by determination

of the characteristic energy loss features in the transmitted primary beam by an electrostatic analyzer. These techniques have the advantage that small precipitates and inclusions on the boundary can also be structurally analyzed by electron diffraction, but have the disadvantage of requiring laborious and sophisticated specimen preparation (see also ch. 11, § 2.10). In the last decade, the X-ray method has been improved to identify quantitatively the crystallographic sites, distribution and types of impurities in many crystals. This method, which takes advantages of the dependence of localized atomic characteristic X-ray emission on the orientation of the incident electron beam, is known as ALCHEMI (Atom Location CHannelling Enhanced MIcroanalysis (SPENCE and TAFTØ [1983])). When the dimensions of segregant and matrix atoms substantially differ from each other (e.g. Bi and Cu), the atomic positions of a large segregant at the grain boundary can be identified by high-resolution electron microscopy combined with a theoretical study of a segregated boundary. In this synergistic way, LUZZI *et al.* [1991] were able to decipher with great accuracy the structure of segregated grain boundaries containing multiple atomic sites.

The third micrographic technique is *atom-probe field ion microscopy* (AP FIM; MILLER [1987]; TSONG [1990]). Here the atoms on the end of a sharp tip of the material may be imaged with atomic spatial resolution. Selected atoms may then be removed from the tip, using (electric) field evaporation, and mass-analyzed. Thus, in principle, individual atoms around a grain boundary or a precipitate/matrix interface may be analyzed and counted. In a alternative variant of the instrument, atoms of selected masses may be field-desorbed and imaged. In this way images of the trace of a grain boundary may be constructed with near atomic resolution. For example, a boron-segregated grain boundary in Ni_3Al is displayed by means of AP FIM in fig. 11.

4. *Theory of segregation processes*

4.1. Introduction: equilibrium and non-equilibrium segregation

In this section we are concerned with the theory of microsegregation in two forms: equilibrium and non-equilibrium. Equilibrium segregation occurs, as noted earlier, as a result of inhomogeneities in the solid giving rise to sites for which solute atoms have a lower free energy. These sites occur at interfaces such as the free surface, grain boundaries and phase interfaces as well as at defect sites, dislocations and stacking faults. All of these regions have concentrations of solute atoms which differ from each other and from that of the bulk materials as shown schematically in fig. 1. At any given temperature, there is a unique solute concentration for each of these sites that is asymptotically approached as time goes to infinity and at a rate governed by diffusion.

On the other hand, non-equilibrium segregation depends on rate processes and kinetic events and, in general, disappears as time approaches infinity if diffusion processes are allowed to reach full equilibrium. There are a number of discrete routes for producing this form of segregation, which include moderate rate quenching of samples from a high temperature, the growth of precipitates, the effect of stress at temperature, etc., which are dealt with in more detail in § 4.7. Here we concentrate on equilibrium segregation which

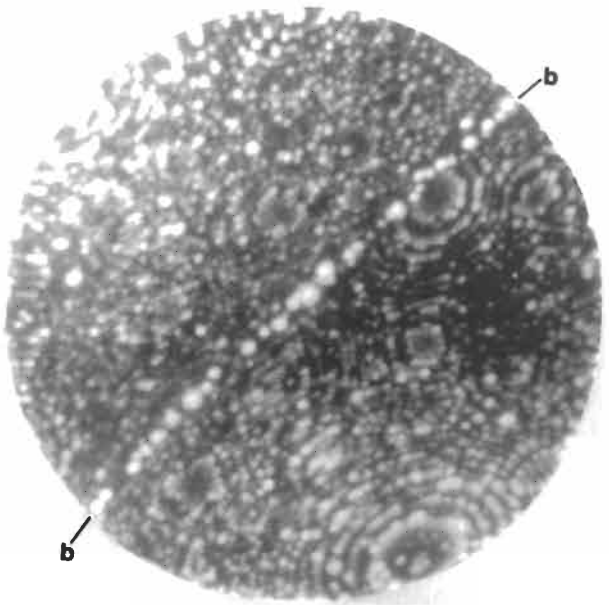


Fig. 11. Field-ion micrograph of boron-decorated boundary in Ni_3Al. Bright spots on boundary were identified as individual boron atoms using single atom identification technique of the atom probe (from MILLER [1987]).

not only is the more common form, but also has been studied more comprehensively.

The free energies of the sites for equilibrium segregation may be represented as shown in fig. 12, where each site has both enthalpy and entropy contributions. In the dilute limit of element A segregating in a matrix of element B, one may consider the relative concentrations $C_1(A)$ and $C_2(A)$, at different sites, 1 and 2, to obey the Maxwell–Boltzmann relation:

$$\frac{C_1(A)}{C_2(A)} = \exp\left(-\frac{\Delta G_{12}}{RT}\right), \tag{8}$$

where ΔG_{12} is the free energy difference between sites 1 and 2. Although eq. (8) provides a feel for what may generally occur in dilute systems, it greatly oversimplifies the problem. In the sections that follow we consider the various adsorption theories for non-dilute systems, which have been established as direct analogues or developments of theories well known in the field of gas adsorption on the free surfaces of solids. The theories will be presented as a series with successively more complex segregation conditions. The free energy term giving rise to segregation will then be analyzed and finally the kinetics of the segregation process will be presented.

4.2. The Langmuir–McLean theory

The earliest derivation was made specifically for grain boundaries, (MCLEAN [1957]), but is also valid for free surfaces. In this classical derivation, McLean proposed a model

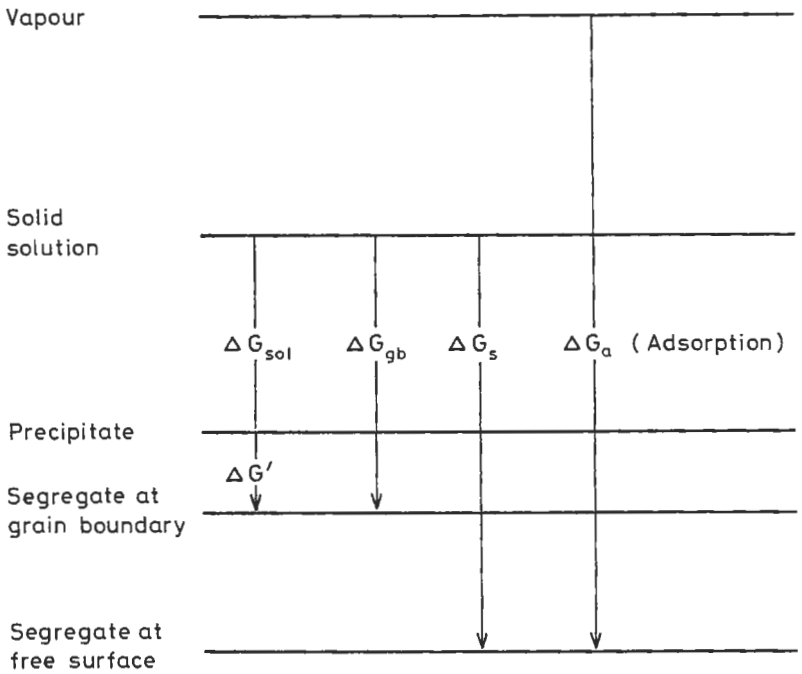


Fig. 12. Free energy diagram for solutes segregated at surfaces and grain boundaries.

of P solute atoms distributed at random amongst N lattice sites and p solute atoms distributed at random amongst n independent grain-boundary sites. The total free energy due to the solute atoms is then

$$G = pe + PE - kT\left[\ln n!\, N! - \ln(n-p)!\, p!(N-P)!\, P!\right], \tag{9}$$

where E and e are the energies of the solute atom in the lattice and in the grain boundary, respectively, and the final term is the configurational entropy for the arrangement of the solute atoms in the bulk and grain boundary. The equilibrium state of the system occurs at the minimum value of G. This may be determined by differentiating G with respect to p, noting that the sum of p and P remains constant. Corresponding to this minimum, the grain-boundary analogue of Langmuir adsorption at free surfaces is obtained:

$$\frac{X_b}{X_b^0 - X_b} = \frac{X_c}{1 - X_c} \exp\left(-\frac{\Delta G}{RT}\right), \quad \text{Langmuir–McLean,} \tag{10}$$

where the notation of HONDROS and SEAH [1977a] is now used: here X_b^0 is the fraction of the grain-boundary monolayer available for segregated atoms at saturation, X_b is the actual fraction covered with segregant, X_c is the bulk solute molar fraction and ΔG is the free energy of segregation per mole of solute. ΔG contains all the entropy terms such as vibrational and anharmonic, but not the configurational term which is written separately

in eq. (9). Segregation measurements obeying the form of eq. (10) have been reported for several systems at both free surfaces and grain boundaries. Figure 13, for example, reproduces the measurements by KUMAR and EYRE [1980] which show that the Langmuir–McLean relation is obeyed for the segregation of 8–12 at. ppm of oxygen to grain boundaries in molybdenum over a wide temperature range. For convenience eq. (10) has been rearranged as follows to obtain a linear plot: X_b^0 is assumed to be unity and $X_c \ll 1$ so that:

$$-\frac{\Delta G}{RT} = \ln \frac{X_b}{(1 - X_b)X_c}. \tag{11}$$

We recognize the basic features of segregation behaviour from this form: the element segregation increases as the bulk solute content, X_c, increases or as the temperature, T, falls. In this simple model, solute atoms are not interacting at the grain boundary and never exceed the saturation segregation level of one monolayer. At this point, we require some way of estimating the value of ΔG for segregation to both grain boundaries and free surfaces. This is discussed in the next sections.

4.2.1. Prediction of the free energy of segregation to grain boundaries

The value of free energy of grain-boundary segregation, ΔG_b, was estimated by MCLEAN [1957] from the elastic strain energy, E_{el}, of the solute in the lattice, all of which is assumed to be released on segregation. For a solute atom of radius r_1 in a site of radius r_0, the elastic energy associated with one mole of solute is given by (WYNBLATT and KU [1979]):

$$E_{el} = \frac{24\pi NKGr_0r_1(r_1 - r_0)^2}{3Kr_1 + 4Gr_0}, \tag{12}$$

where K is the solute bulk modulus, G is the solvent or matrix shear modulus and N is Avogadro's number. Two examples of the values for segregants in copper computed by MCLEAN [1957] are shown in table 1. The final column shows the experimental values.

Values estimated by this approach are generally correct to within a factor of two. However, much higher accuracy is required for the quantitative prediction of segregation behaviour. This was provided in the treatment by SEAH and HONDROS [1973]. Using the (BET) gas adsorption theory of BRUNAUER *et al.* [1940], they write the solid-state analogue as

$$\frac{X_b}{X_b^0 - X_b} = \frac{X_c}{X_c^0} \exp\left(\frac{\Delta G'}{RT}\right), \qquad \text{(Truncated BET).} \tag{13}$$

where X_c^0, the solid solubility, is the important parameter. For slightly soluble substances, $X_c^0 = \exp(\Delta G_{sol}/RT)$ so that the Langmuir–McLean and Truncated BET theories are in fact identical in the dilute limit with

$$\Delta G_b = \Delta G' + \Delta G_{sol}, \tag{14}$$

References: p. 1284.

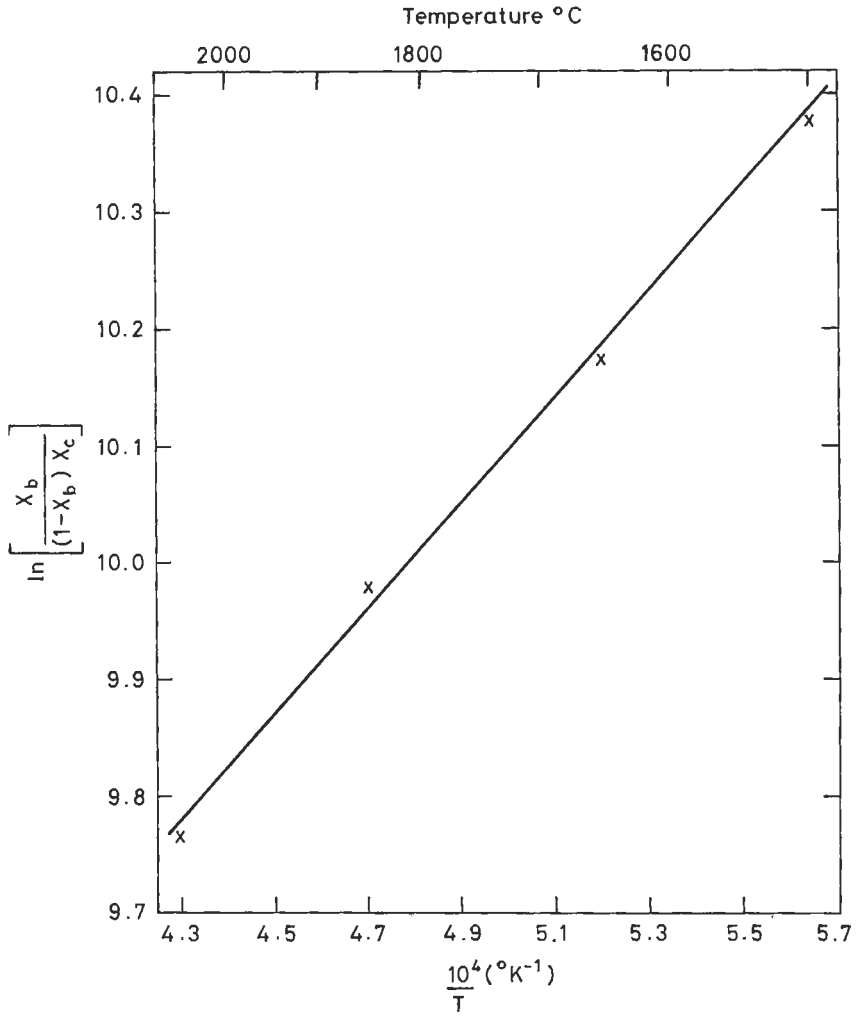


Fig. 13. Langmuir–McLean plot for oxygen grain-boundary segregation measured by AES in molybdenum (after KUMAR and EYRE [1980]).

as shown in the energy level diagram of fig. 12. The importance of the Truncated BET form is that one may define as a useful parameter the enrichment ratio, β_b which represents the ratio X_b/X_c in the dilute limit. Since X_b^0 generally turns out to be unity, we find:

$$\beta_b = \frac{K}{X_c^0} \quad \text{where} \quad K = \exp\left(-\frac{\Delta G'}{RT}\right). \tag{15}$$

Figure 14 shows a compilation of measurements of grain-boundary enrichment ratios, β_b, (estimated by the grain-boundary energy and surface-analysis approaches), in relation to

Table 1
Values of the elastic distortion energy, E_{el}, for Sb and Bi in Cu.

Element	K (GN/m^2)	r_1 (nm)	$(r_1 - r_0)/r_1$	E_{el} (kJ/mol)	ΔG_b (kJ/mol)
Sb	41	0.1528	+ 0.194	37	65
Bi	31.4	0.157	+ 0.266	66	100

the solubilities of the segregant at the measurement temperature. Thus the grain-boundary enrichment for systems in which no measurements are available may be accurately estimated from handbooks of phase diagrams (MASSALSKI *et al.* [1986]). It is clear that, in the dilute limit, eq. (15) provides the theoretical description of the experimental data of

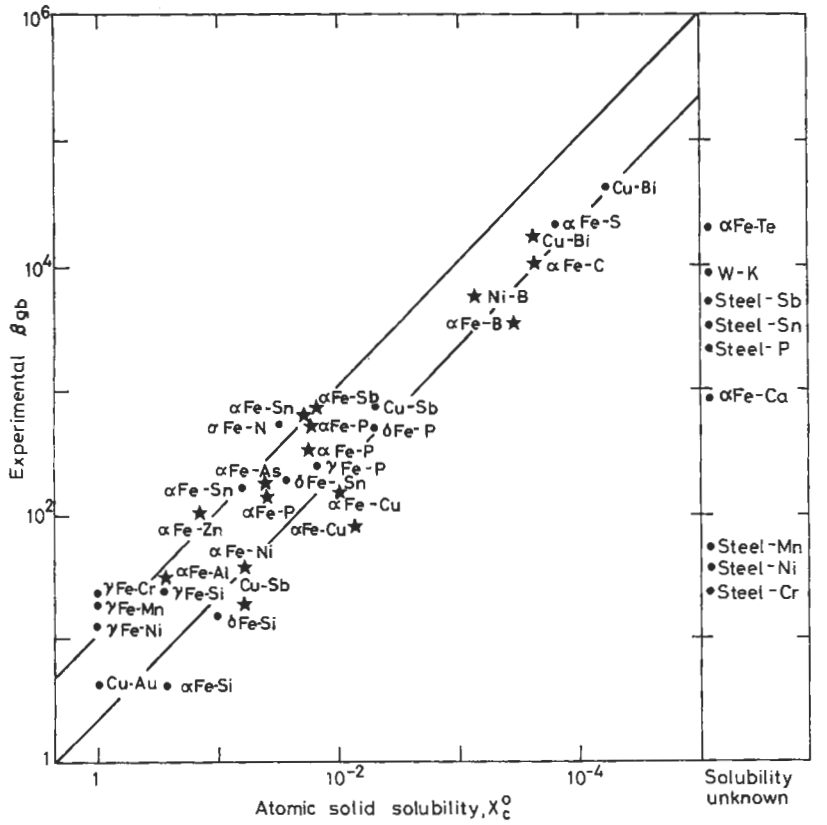


Fig. 14. Correlation of measured grain-boundary enrichment ratios with the atomic solid solubility (after SEAH [1980d]).

fig. 14 quite accurately. Indeed it is seen that whilst ΔG_{sol} ranges from zero to –80 kJ/mol, $\Delta G'$ has a mean value of –10 kJ/mol with a standard deviation of only 6 kJ/mol. Thus,

$$\Delta G_b = \Delta G_{sol} - 10 \pm 6 \text{ kJ/mol}. \tag{16}$$

Equation (13) is valid only for $X_c < X_c^0$. If excess solute is present, so that a second phase appears, the solute content is limited to X_c^0 and eq. (13) becomes

$$\frac{X_b}{X_b^0 - X_b} = \exp\left(-\frac{\Delta G'}{RT}\right). \tag{17}$$

The translation from eq. (13) to eq. (17) as the solubility limit is reduced through the solute level, is illustrated by the measurements for sulphur in iron by SEAH and HONDROS [1973], shown in fig. 15. Equation (13) is shown for the higher temperatures, where the solubility is high, and eq. (17) at the lower temperatures.

WATANABE *et al.* [1980] proposed to refine the dependence of grain-boundary enrichment ratio on solubility limit by considering the effect of grain-boundary orientation (see also § 4.5) and suggested the idea of constructing a *grain-boundary segregation diagram* as a three-dimensional function of β_b in dependence on X_c^0 and on the misorientation angle of adjacent grains. The first experimental grain-boundary segregation diagram was constructed by HOFMANN and LEJČEK [1991] for Si, P and C segregation at [100] symmetrical tilt grain boundaries in α-Fe. In this diagram, however, the

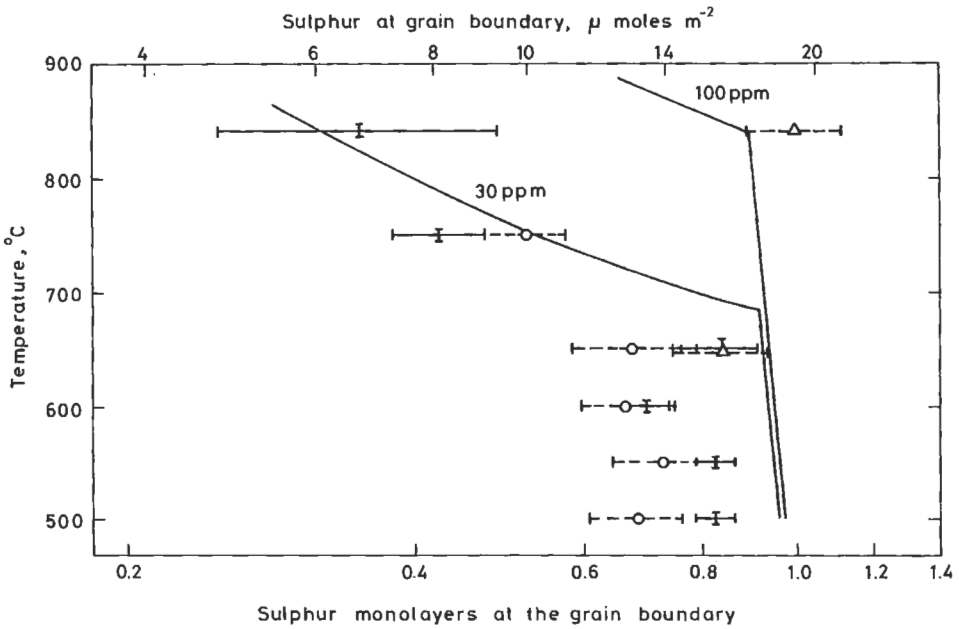


Fig. 15. AES measurements of sulphur grain-boundary segregation in iron as a function of temperature for two bulk compositions (after HONDROS and SEAH [1977b]).

grain-boundary enrichment ratio was replaced by a more general and independent thermodynamic quantity, ΔH^0. Recently, LEJČEK and HOFMANN [1993] theoretically derived the dependence of the segregation enthalpy on both the grain-boundary orientation and the bulk solid solubility. It was shown that this dependence can be expressed as a sum of two mutually independent terms

$$\Delta H^0\left(\Phi, X_c^0\right) = \Delta H^*\left(\Phi\right) + vR\left[T \ln X_c^0\right] \tag{18}$$

where $\Delta H^*(\Phi)$ is the enthalpy of segregation of a solute with unlimited solid solubility (i.e., $X_c^0 = 1$) in a given matrix at a given interface of orientation Φ, and v is the parameter characteristic for the matrix element. Expression (18) was derived under two basic assumptions, first,

$$T \ln a_l^* = \text{const} \tag{19}$$

where a_c^* is the activity of a solute at the solid solubility limit, and second,

$$a_c^* = \left(X_c^0\right)^v \tag{20}$$

which relates the activity a_c^* and concentration X_c^0 of any solute at the solid solubility limit in a given matrix. In fact, eq. (18) represents an extension of the model of SEAH and HONDROS [1973] by taking additionally into account (i) the structural dependence of interfacial segregation (i.e., $\Delta H^*(\Phi) \neq \text{const}$), and (ii) the non-ideal behaviour of the system (i.e., $v \neq 1$). The grain-boundary segregation diagram for the [100] symmetrical tilt grain boundaries in α-Fe is shown in fig. 16.

4.2.2. Prediction of the free energy of segregation to surfaces

The theory for the free energy of surface segregation, ΔG_s, has developed along less phenomenological lines than that for grain boundaries although both are merely special cases of one general effect. Each of the two approaches leads to a reasonably accurate prediction for the case to which it pertains. The reader may blend the two approaches to gain a greater insight into the nature of the contributions to the free energy, however, at the present time it does not lead to a greater accuracy in calculation.

The treatment for free surfaces is based on the derivation of WILLIAMS and NASON [1974] who use a layer-by-layer model of the surface region, with solute A atoms replacing the solvent B atoms substitutionally at the surface. They find that this model gives the Langmuir–McLean adsorption eq. (10) for the fractional coverage of A on the surface, which allows the calculation of ΔG_s. The basic model adopts the quasi-chemical approach, shown in fig. 17, where the nearest neighbour bond energies are given by ε_{AA}, ε_{BB} and ε_{AB} between AA, BB and AB neighbours respectively. ΔG_s is calculated as the energy released on removing an A atom from a lattice site and exchanging it for a B atom on the free surface. The energy required to remove A from the bulk is:

$$-\left(Z_l + 2Z_v\right)\left[X_c\varepsilon_{AA} + \left(1 - X_c\right)\varepsilon_{AB}\right], \tag{21}$$

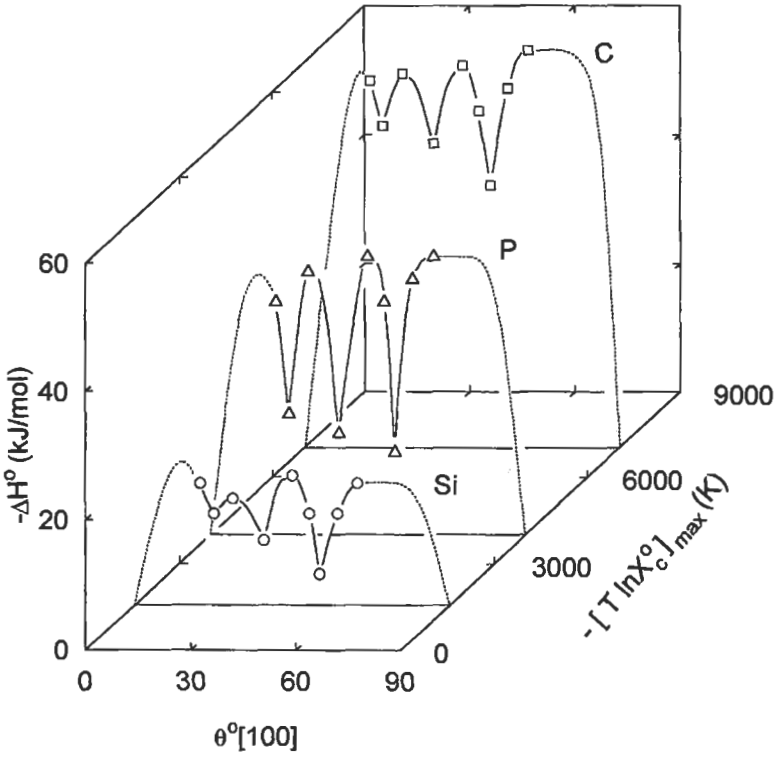


Fig. 16. Grain-boundary segregation diagram for [100] symmetrical tilt grain boundaries in α-Fe showing the dependence of enthalpies of Si, P and C segregation on misorientation angle θ of both adjacent crystals and on the bulk solid solubility data (after LEJČEK and HOFMANN [1993]).

where Z_l is the atomic coordination number in the layer and Z_v is the atomic coordination to one of the adjacent layers. The energy to replace A at the surface is similarly:

$$Z_l\left[X_s\varepsilon_{AA} + (1 - X_s)\varepsilon_{AB}\right] + Z_v\left[X_c\varepsilon_{AA} + (1 - X_c)\varepsilon_{AB}\right]. \tag{22}$$

Equivalent terms occur for the B atom moving in the reverse direction so that by using the *regular solution approximation* in which

$$\varepsilon_{AB} = \tfrac{1}{2}\left(\varepsilon_{AA} + \varepsilon_{BB}\right) + \omega, \tag{23}$$

we find that the free energy of surface segregation per atom, $\Delta G_s/N$, is given by

$$\Delta G_s/N = \tfrac{1}{2} Z_v\left(\varepsilon_{BB} - \varepsilon_{AA}\right) + 2\omega\left\{Z_l\left(X_c - X_s\right) + Z_v\left(X_c - \tfrac{1}{2}\right)\right\}. \tag{24}$$

In Williams and Nason's theory, successive layers away from the surface are considered and for nonzero values of ω the composition of the second and further layers can also differ from that of the bulk. For instance, if ω is negative, there is a perceptible

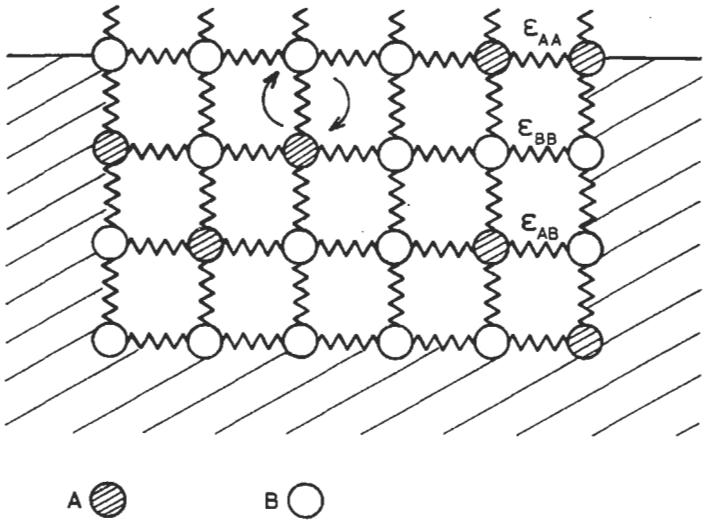


Fig. 17. Schematic substitutional model for the quasi-chemical segregation theory.

oscillation of composition decaying away from the surface in addition to the surface segregation. Here, however, to simplify the analysis and because it has not been possible to observe such small oscillations in practice, it will be assumed that all of the compositional variation occurs in the topmost atom layer.

Numerical evaluation of eq. (24) may proceed in several ways. Following WYNBLATT and KU [1979] the bond strengths ε_{AA} and ε_{BB} may be derived from the molar surface energies of the two constituents. At absolute zero these energies may be expressed as follows:

$$\gamma_A^s = -\tfrac{1}{2} Z_v \varepsilon_{AA} N; \qquad \gamma_B^s = -\tfrac{1}{2} Z_v \varepsilon_{BB} N. \tag{25}$$

The regular solution parameter, ω [eq. (23)], may then be calculated from tabulated values (HULTGREN *et al.* [1973]) of the enthalpy of mixing, H^m:

$$\Omega = N\omega = H^m / ZX_c(1 - X_c), \tag{26}$$

where Z is the total coordination number. Williams and Nason in their original paper evaluated ε_{AA} and ε_{BB} from the molar sublimation enthalpies, H^{sub}, where

$$H_i^{sub} = -\tfrac{1}{2} Z \varepsilon_{ii} N, \tag{27}$$

where i stands for A or B. The above terms now give some insight into the kinds of elements that will segregate to free surfaces. If the regular solution parameter can be ignored, it is seen from eqs. (24) and (25) that, other things being equal, the element with the lower surface energy enriches the surface of a binary system. Alternatively, from eqs. (24) and (27), this will also be the element of lower sublimation enthalpy.

The above is particularly true for liquid binary metallic systems as shown by

HONDROS [1980a]. However, for solids McLean's strain term given in eq. (12) is also important. The strain term is only significant for excess-sized solute atoms. TSAI *et al.* [1977] point out that, as there is little relaxation around vacancies, the strain energy associated with $r_0 < r_1$, must be much less than the converse.

Using the above approach, SEAH [1979] gained an overall view of surface segregation by combining eqs. (24) and (25) and adding the strain term of eq. (12). From what we have seen in § 4.2.1, this term will be of the order of $2.6 \times 10^{35} \times r_0(r_1 - r_0)^2$. In fact, using regression analysis of published data, Seah found for dilute binary systems:

$$-\Delta G_s = \left(\gamma_B^S - \gamma_A^S\right) + 1.8\Omega + M2.66 \times 10^{35} r_0 \left(r_1 - r_0\right)^2, \tag{28}$$

where M is unity for $r_1 > r_0$ and zero for $r_1 < r_0$. The r values are the atom sizes given by $\rho N(2r)^3 = A$, where ρ is the bulk density and A the atomic weight of the element. Figure 18 shows the resulting correlation for an improved version of eq. (28), using the relation $\beta_s = X_s/X_c = \exp(-\Delta G_s/RT)$, and the experimental data for the surface enrichment ratio. This analysis confirmed the correctness of simply adding the bond and strain energy terms directly and that the latter is significant only for $r_1 > r_0$. Thus we may write:

$$\begin{aligned}\Delta G_s = {} & \left[\text{bond term, eq. (24)}\right] + M\left[\text{strain term, eq. (12)}\right] \\ & - T\left(\text{entropy terms discussed in § 4.2.4}\right).\end{aligned} \tag{29}$$

Several other models have been developed to estimate thermodynamic parameters of surface segregation. On the basis of a simple tight-binding theory of surface segregation in alloys of transition metals, MUKHERJEE and MORÁN-LÓPEZ [1987] proposed a method to predict surface segregation assuming a rectangular shape of the d-band density of states. This method enables the equilibrium surface concentration to be calculated as a function of temperature. Another technique for calculating the heat of segregation is based on the BFS model (BOZZOLO *et al.* [1992]). The energy of segregation which is composed of two basic contributions, the strain term and the chemical term, can be obtained using a zero-temperature Monte Carlo calculation. As the input data, pure elemental parameters and only two alloy properties have to be used (BOZZOLO *et al.* [1993]). In contrast to these two models, a phenomenological approach was used in *modern thermodynamic calculation of interface properties* (MEZEY *et al.* [1990]). In this approach, the free energy of segregation of a solute I at a surface in a binary M–I alloy can be expressed as a sum of two terms, the dangling bond contribution:

$$\Delta G_I^d = \left(\gamma_I - \gamma_M\right) A_M, \tag{30}$$

and the real mixture contribution:

$$\Delta G_I^m = q\Delta\mu_I^E - \Delta\mu_M^E. \tag{31}$$

Here γ_i are the surface energies of pure elements i, $q = A_M/A_I$, A_i are the partial molar areas of species i, and $\Delta\mu_i^E$ are the excess chemical potentials of mixing related to the elements i. The total free energy of segregation, $\Delta G_I = \Delta G_I^d + \Delta G_I^m$, can be determined by means of individual factors evaluated from elemental data, and from experimental values of

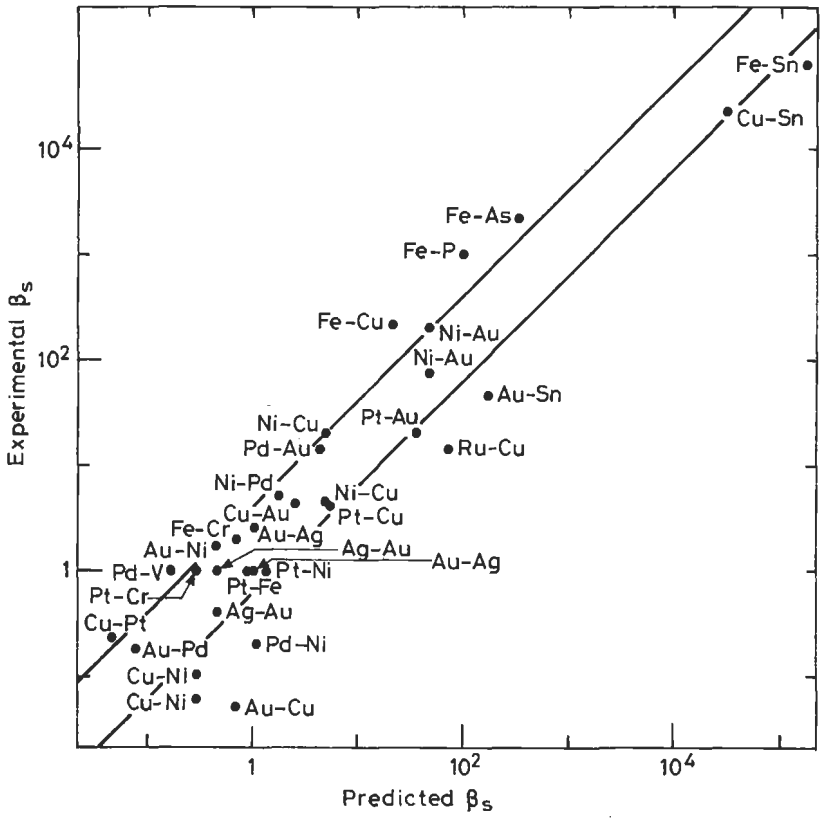


Fig. 18. Correlation of measured surface enrichment ratios with the predicted values from eq. (28) (after SEAH [1979]).

interfacial free energies and molar enthalpies of mixing. It was also shown that this approach can describe anisotropy of surface segregation of Pt in Ni (MEZEY and HOFER [1992]).

4.2.3. Segregation with adsorbate–adsorbate interactions

In the above we have concentrated on evaluating ΔG in the limit of dilute X_b, X_s and X_c to compare the overall behaviour with experiment. To study the effect of higher levels of segregation we may rewrite eq. (24) simply:

$$\Delta G_s = \Delta G_s^0 - 2Z_l N\omega X_s, \tag{32}$$

where $\Delta G_s^0 = NZ_v(\varepsilon_{BB} - \varepsilon_{AB})$ does not depend on the segregation level. This equation is based on the substitutional segregation model shown in fig. 17, which is a good description for many free surfaces and grain boundaries. A completely different description, valid only for free surfaces, is the *Fowler theory* in which adsorbed atoms on top of the free surface have, simply, the interaction energy ε_{AA}. In this case:

$$\Delta G_s = \Delta G_s^0 + Z_l N \varepsilon_{AA} X_s. \tag{33}$$

This is seen to be the same as the result for eq. (32) if the top atom layer B atoms are replaced by vacancies.

Using either of the above equations, if the presence of a segregated A atom enhances the probability of the next site being occupied by an A atom, ΔG becomes more negative as the segregation proceeds. As ε_{AA} becomes more and more negative, the segregation shows progressively sharper rises as the temperature falls until eventually the rise in segregation becomes discontinuous at a temperature given by the solution of

$$C = -4RT, \tag{34}$$

where C is the coefficient of X_s in eqs. (32) and (33). Measurements illustrating this form of segregation are shown in fig. 19 for Se and Te segregation in pure iron. There is also shown the simple Langmuir–McLean case with $C=0$ and a case in which a discontinuous rise occurs from a low to a high segregation level. As seen from the labelling of the ordinate, the lowest segregations occur at low bulk solute contents, X_c, and high temperatures, T. The Fowler type of discontinuous grain-boundary segregation with $C=-5RT$ was observed by MENYHARD *et al.* [1989] for Bi in Cu.

4.2.4. Temperature dependence of the free energies of segregation

The temperature dependence of segregation occurs mainly through the explicit temperature term in eq. (13). As the temperature rises the segregant "boils off", leading to a lower value. However, the free energy itself is temperature dependent through the customary relation:

$$\Delta G = \Delta H - T\Delta S \tag{35}$$

where ΔS is the entropy change on segregation to a grain boundary or free surface. SEAH and LEA [1975] write ΔS as the sum of three terms, ΔS_v, ΔS_a and ΔS_m. ΔS_v is associated with the changes in the vibrational behaviour on segregation, ΔS_a with the anharmonicity of the potential of the segregant atom site and ΔS_m with site multiplicity. The vibrational entropy term is given, in the Einstein model, as

$$S_v = 3R\left[1 + \ln\left(kT/h\nu\right)\right], \tag{36}$$

where $kT \gg h\nu$, ν being the Einstein frequency. $h\nu$ may be rewritten $k\theta_E$ or $0.775k\theta_D$ where θ_E and θ_D are the Einstein and Debye temperatures (KITTEL [1986]) so that

$$\Delta S_v = 3R \ln\left(\theta_D / \theta_D^*\right), \tag{37}$$

where θ_D and θ^* are the Debye temperatures for the solute atom in the matrix and at the distorted site of the grain boundary or free surface. Owing to the greater vibrational amplitude of atoms at surfaces, θ_D^*/θ_D has been calculated for tin surface segregation in iron to be approximately 2/3. In this way, taking into account the matrix atom exchanging places with the segregant atom, ΔS_v is about $3.5R$ for the free surface and $3.3R$ at grain boundaries.

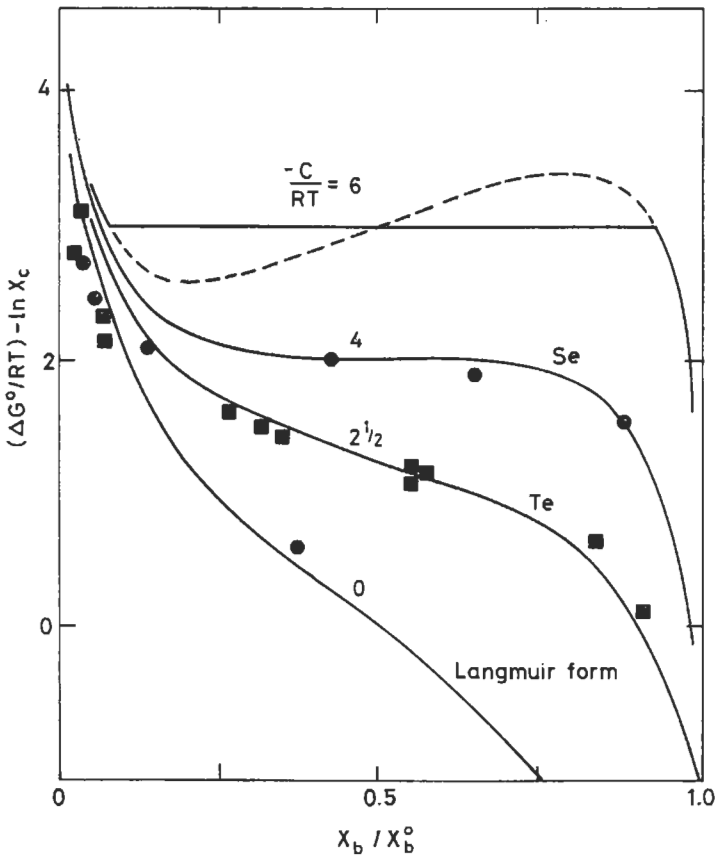


Fig. 19. Grain-boundary Fowler isotherms with the AES measurements of PICHARD *et al.* (after HONDROS and SEAH [1977a]).

In a similar manner, the two further entropy contributions, ΔS_a, due to the anharmonicity of the atomic potential at the termination of a lattice, either by the surface or by a grain boundary (EWING [1971]), and ΔS_m, due to the possible site multiplicity for the segregating atoms, may be calculated. Both terms are much smaller than ΔS_v, and so will not be treated here.

The correlation of the above full prediction of the temperature dependence of ΔG with the experimental data for tin in iron is shown in fig. 20 for both surface and grain-boundary segregations. This figure also serves to emphasize the generally observed fact that surface segregation at moderate temperatures is generally much greater than segregation at grain boundaries.

References: p. 1284.

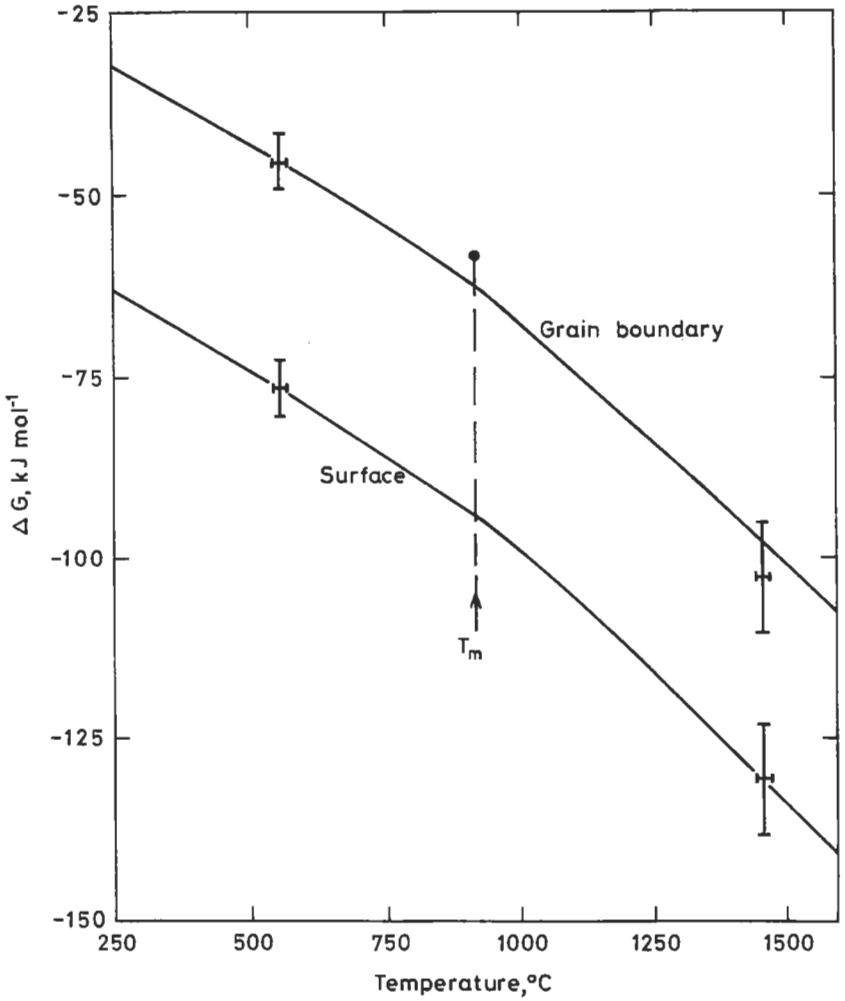


Fig. 20. The predicted temperature dependences of the free energies of surface- and grain-boundary segregations of tin in iron, together with the experimental data at 823 K (550°C) and 1693 K (1420°C) (after SEAH and LEA [1975]).

4.3. Segregation in simple ternary systems: site competition

Site competition in a ternary system occurs where there are two segregants which can fill the same sites at the free surface or grain boundary. Each segregant obeys the appropriate adsorption theory with the available empty sites reduced by the segregation of the other species. At saturation segregation, this leads to the situation where the sum of the segregants is constant at one monolayer. This has been observed on the free surface of iron for Sn, Sb and O in competition with S. Generally the Sn, Sb or O

segregate rapidly to an equilibrium level but, with time at temperature, S (which may be present at only a few ppm in solution), will replace much of the Sn, Sb and O. Such a competition is not generally expected at grain boundaries since the bond energy terms can continue to be reduced until two total monolayers are segregated, and this latter level of segregation is rarely reached. Grain-boundary site competition between N and S has been reported (TAUBER and GRABKE [1978]) in iron but not under all conditions (JONES *et al.* [1981]). More importantly, P is displaced by C (ERHART and GRABKE [1981]). This is of great significance in steels, as discussed in § 6.3.

4.4. Segregation in complex metallurgical systems

In 1975 GUTTMANN extended the Fowler theory to allow for coupling interactions between two co-segregating species in multicomponent systems. This development, vital to explain the segregation behaviour resulting in intergranular failures in engineering materials, gives the analogue to the Langmuir–McLean equation.

$$\frac{X_{bi}}{X_{bi}^{0}} = \frac{X_{ci}\exp\left(-\Delta G_i/RT\right)}{1+\sum_{j=1}^{2} X_{cj}\left[\exp\left(-\Delta G_j/RT\right)-1\right]}, \tag{38}$$

where X_{b1} and X_{b2} are the molar fractional monolayers segregated by a noxious impurity and by an alloying element of bulk contents X_{c1} and X_{c2} respectively. The ΔG's are given by equations of the form of eqs. (32) or (33) but depend not on the level of self-segregation but on each other through the equations:

$$\Delta G_1 = \Delta G_1^0 - 2C_{13}X_{b1} + \alpha'_{12}X_{b2}, \quad \Delta G_2 = \Delta G_2^0 - 2C_{23}X_{b2} + \alpha'_{12}X_{b1}, \tag{39}$$

where ΔG_1^0 and ΔG_2^0 are the free energies of segregation of the impurity and alloying elements separately in the matrix, and C_{13} and C_{23} are the respective Fowler-type interaction coefficients (cf. eq. (32)). The interaction coefficient α'_{12} refers to the changes in nearest neighbour bond energies ε_{ij} in forming the alloy–impurity bonds,

$$\alpha'_{12} = Z_l N\left(\varepsilon_{12} - \varepsilon_{11} - \varepsilon_{22}\right), \tag{40}$$

and can be obtained from measurements of the effects of the alloying elements on the solubilities of the impurities (GUTTMANN [1976]). The stronger the reduction in solubility, the stronger the coupling of the alloy element to the impurity and the more the alloying addition enhances the impurity segregation.

In a similar manner to that of the Fowler isotherm, as α'_{12} becomes more and more negative the curves exhibit a more pronounced S shape and eventually attain a region in which the segregation shows a discontinuous increase from a low to a high level. The general effect of Guttmann's coupling can be seen by reference to fig. 14. Alloy element additions with a negative value of α'_{12} cause the impurity element solubility to fall. Simultaneously, through eqs. (38) and (39), the segregation level rises. Thus, the addition of the alloying element causes the point for the original binary system in fig. 14 to move diagonally upwards to the right but to stay within the general correlation (GUTTMANN and MCLEAN [1979]).

The general behaviour of this model is now well accepted and allows an accurate

description of some of the behaviour observed in embrittlement problems. The increase in the free energy of segregation of one species by a second, described in eq. (39), has been validated in careful experiments on grain-boundary segregation of Sb and Ni in steel by GAS *et al.* [1982]. Although high $-\alpha'_{12}$ values lead to a high enrichment ratio, the result in practice often will be a low segregation. This is because the reduction in solubility and hence the number of solute atoms available to segregate make the rate of build-up of the segregant at the surface or grain boundary very slow indeed so that, with practical heat treatments, a negligible segregation occurs. This is precisely the case, for example, with the addition of manganese to remove the sulphur embrittlement in steels. The problem of the kinetics of segregation is very important in practice and this is considered in the next section.

Guttmann's original theory was modified by SEAH [1977] to remove the site competition between alloy elements and impurities which had not been observed in practice. In eq. (38) the sum over the j terms is simply replaced by the i^{th} term. This simplifies the calculations and gives a better description of the experimental results as discussed in § 6.3. Several more complex theories are detailed by GUTTMANN and MCLEAN [1979] and by GUTTMANN [1980].

Solute segregation in multicomponent systems can alternatively be described in terms of chemical equilibrium (BRIANT [1990]). In this approach, the reversible chemical reaction r between two segregating elements I and J at the boundary (b) and in bulk (c),

$$I^{(c)} + J^{(b)} = I^{(b)} + J^{(c)}, \tag{41}$$

can be characterized by an equilibrium reaction constant K_r,

$$K_r = \frac{a_{bI}a_{cJ}}{a_{cI}a_{bJ}} = \exp\left(-\frac{\Delta G_r}{RT}\right) \tag{42}$$

where a_{bi} and a_{ci} are the activities of individual solutes at the boundary and in the bulk, respectively. In contrast to the Guttmann's approach in which the solute interactions modify the value of segregation free energy (eqs. (38) and (39); GUTTMANN [1975]), the value of ΔG_r is constant for the reaction r at a given constant temperature in the Briant's approach (eq. (42)) and only the changes in bulk activities and solid solubilities are responsible for segregation behaviour of any element in the system, i.e., no solute atom interactions are considered. By means of the latter model, BRIANT [1988a] interpreted the grain-boundary segregation of Sb in α-Fe enhanced by additions of Ni as observed by GAS *et al.* [1982], by increasing the bulk solid activity of Sb in α-Fe by Ni. Strictly speaking, both Guttmann's and Briant's approaches lead to rather similar expression (compare eqs. (38) and (39) with eq. (42)), supposing the activity a_i to be expressed as the product of concentration X_i and activity coefficient f_i, and the product of activity coefficients to be represented by combinations of interaction coefficients in eq. (39).

4.5. Anisotropy of segregation

In the above we considered that all sites are equivalent prior to segregation. This is generally not the case either at free surfaces or at grain boundaries. The simplest case to

understand is that of free surfaces. The driving force for segregation is the lowering of the overall system free energy. If the surface energy is already low as in the case of the low-index surfaces (111), (110) and (100), we should expect that the segregation here would be lower than that for a high index surface, and lowest for (111) surfaces in an fcc lattice (ERLEWEIN and HOFMANN [1977]). Little direct work has been reported, but measurements by JOHNSON *et al.* [1978] for Au segregation on Ni, by ZHOU *et al.* [1981] for C and Sn segregation on Fe, and by SUZUKI and OKU [1990] for Cr segregation on Fe generally confirm this although Si segregation at Fe (111) was detected to be rather high (ZHOU *et al.* [1981], SUZUKI and OKU [1990]). Considerably more work has, however, been completed on the influences of the anisotropy of surface segregation on the anisotropy of surface energy, because of the technological significance, and this is discussed in § 5.1.

4.5.1. Segregation at symmetrical grain boundaries

In contrast to free surfaces which can fully be described by the orientation of surface normal (i.e., two degrees of freedom in a chosen reference system), five variables are necessary for a complete characterization of a grain boundary: Three of them determine misorientation of lattices of the two adjacent grains (rotation axis and misorientation angle) and two of them describe the orientation of the boundary plane (WOLF [1992]). Therefore, the orientation dependence of grain-boundary segregation cannot be displayed in the same simple manner as that for surface segregation. In principle, the complete anisotropy of grain-boundary segregation for a binary system should be constructed in an eight-dimensional space as an interdependence of interfacial composition (or grain-boundary enrichment ratio) on five grain-boundary parameters, and on temperature and pressure (SEIDMAN [1992]). To simplify the problem, orientation dependences of grain-boundary segregation are displayed as plots of grain-boundary composition vs. one grain-boundary variable keeping the other four ones constant, in addition to constant temperature and pressure. In this way, solute segregation was found to increase with increasing misorientation angle for both *tilt* (e.g., BISCONDI [1982]) and *twist* (e.g., SEIDMAN [1991]) *low-angle* grain boundaries. Among the *high-angle* grain boundaries, two types of interfaces can be distinguished: the majority of *general* grain boundaries with relatively high values of β_b, and a few *special* grain boundaries characterized by a low enrichment factor in comparison to the general ones (cf. fig. 21). From the viewpoint of atomic structure, the special grain boundaries are equivalent to so-called *favoured* or *delimiting* ones, the structures of which are created by single structural units whereas the structures of the general (or *non-favoured*) ones consist of a combination of structural units of the two neighbouring favoured grain boundaries (FISCHMEISTER [1985], see also ch. 9, § 2.2). To characterize the special grain boundaries, various bulk geometric parameters are used. One of the most widespread geometric characterizations of grain boundaries is the reciprocal density of coincidence sites Σ which can be determined for cubic crystals by

$$\Sigma = \delta\left(h^2 + k^2 + l^2\right) \tag{43}$$

where h, k and l are the Miller indices of the symmetrical grain boundary corresponding to the given misorientation θ, and $\delta = 1$ if $h^2+k^2+l^2$ is odd and $\delta = 0.5$ if $h^2+k^2+l^2$ is

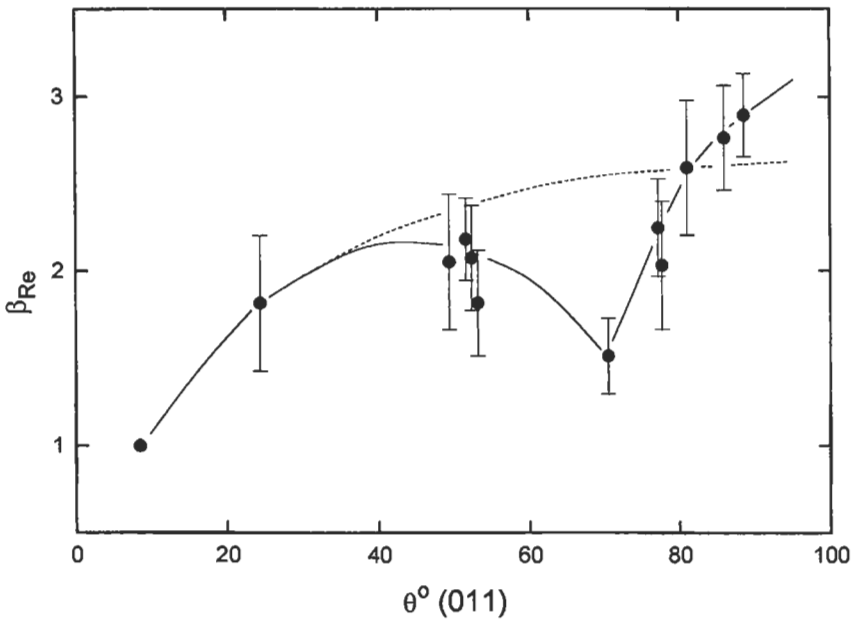


Fig. 21. Orientation dependence of the Re enrichment factor for (011) twist grain boundaries of a W–25at%Re alloy (after SEIDMAN [1991]).

even. Other geometric parameters are the inverse planar coincidence site density Γ defined as the ratio of the unit-cell area A_n and the smallest possible area per atom A_1

$$\Gamma = \frac{A_n}{A_1}, \tag{44}$$

and the interplanar spacing $d(hkl)$ defined for cubic lattices by

$$d(hkl) = \frac{\varepsilon a}{\sqrt{h^2 + k^2 + l^2}} \tag{45}$$

where a is the lattice parameter and ε has the value of either 0.5 or 1 depending on the particular combination of odd and even Miller indices and on the type of the cubic lattice (WOLF [1992]). Usually, it is accepted that special grain boundaries possess low values of Σ and Γ, or high values of $d(hkl)$; however, because these parameters are bulk and geometric characteristics, there is no one-to-one correlation between them and the structure of grain boundaries. Therefore, these parameters can serve as rough estimates only (SUTTON and BALLUFFI [1987]).

Owing to the great difference in the behaviour of special and general grain boundaries and because of the great importance of grain-boundary properties in the complex behaviour of a whole polycrystal, a large effort has been spent in the last two decades to learn the character of all possible special grain boundaries, their atomic structures and

properties. As regard grain-boundary segregation, low values of β_b were found at $\Sigma = 11$, 50.5° [110], {113} and $\Sigma = 3$, 70.5° [110], {112} symmetrical tilt grain boundaries in a Mo–O system (BISCONDI [1982]) or at the $\Sigma = 3$, 70.5° (011) twist grain boundary in a W–Re alloy (fig. 21; SEIDMAN [1991]). A low tendency to P, C and Si segregation was detected for $\Sigma = 13$, 22.5° [100], {015}, $\Sigma = 5$, 36.9° [100], {013} and $\Sigma = 5$, 53.1° [100], {012} symmetrical tilt grain boundaries in an Fe–Si alloy (HOFMANN *et al.* [1992]). In this case, grain-boundary segregation was represented by segregation enthalpies ΔH_I^0 of individual elements as shown in fig. 16. This parameter is more general than the enrichment ratio β_b: for example, it is independent of temperature and bulk composition of the alloy. This is very important in the case of complex multicomponent systems in which site competition and solute interaction can modify the grain-boundary composition to such an extent that β_b loses the ability to characterize grain-boundary segregation in a binary system (see also §§ 4.3. and 4.4). The above mentioned interfaces represent special grain boundaries selected on the basis of geometric classification of symmetrical tilt grain boundaries (PAIDAR [1987]).

4.5.2. Segregation at asymmetrical grain boundaries

For asymmetrical grain boundaries, the relative orientations of both crystals to the grain-boundary plane have to be considered and therefore their description is more complex. In a neat study by SUZUKI *et al.* [1981] of phosphorus grain-boundary segregation in iron, they show that after fracture, the level of phosphorus observed may be associated with the Miller indices of the fracture surface, as shown in fig. 22a. The measurements of the segregation in relation to the stereographic triangle are shown in fig. 22b. It is clear that the lower segregations are associated with the lower index poles since, as one may expect, the boundaries involving these surfaces are more ordered than the general high-index boundaries. The preponderant role of the orientation of the grain-boundary plane in solute segregation was also experimentally observed for sulphur in nickel. No segregation effects were detected at symmetrical {111} and {112} facets of a $\Sigma = 3$ twin boundary but some sulphur enrichment was observed at asymmetrical ones such as, for instance, (111)/(113) (BOUCHET and PRIESTER [1986, 1987]). Similarly, a symmetrical {113} grain boundary did not exhibit any S segregation in a nearly $\Sigma = 11$ Ni bicrystal whereas asymmetrical boundary facets with orientations close to e.g. $(1\bar{1}2)/(\bar{1}16)$ and $(1\bar{1}4)/(\bar{2}25)$ (i.e. with a deviation of the order of ~2° from ideal $\Sigma = 11$) were found to be enriched with sulphur (SWIATNICKI *et al.* [1990]). In contrast to the {113}, the other $\Sigma = 11$ symmetrical grain boundary {332} was found to be very unstable and to split into numerous asymmetrical parts considerably enriched with sulphur (JUHAS *et al.* [1994]). The conclusions of SUZUKI *et al.* [1981] were directly supported by systematic measurements of grain-boundary segregation in carefully prepared $\Sigma = 5$, 36.9° [100] tilt bicrystals of an α-Fe–Si alloy. Low tendency to Si, P and C segregation (low absolute values of segregation enthalpy) was observed at the boundaries created by at least one low-index plane, i.e., symmetrical {013} and {012}, and asymmetrical $(001)/(03\bar{4})$ and $(0\bar{1}7)/(01\bar{1})$ ones. On the other hand, high solute enrichment was detected at the high-index asymmetrical boundaries $(018)/(04\bar{7})$ and $(0\bar{3}11)/(09\bar{7})$ (LEJČEK *et al.* [1992]). In addition, different amounts of segregants were found at both

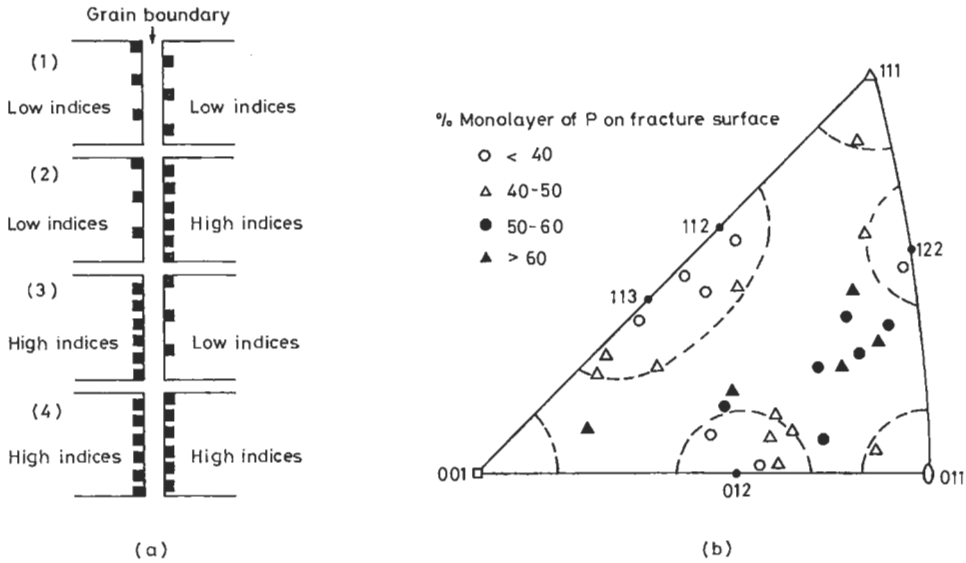


Fig. 22. The anisotropy of phosphorus grain-boundary segregation in Fe–1%P showing (a) the segregant atoms (dark squares) in relation to the fracture process and (b) the level of segregation in relation to the relative orientation of each grain to the grain boundary (after SUZUKI *et al.* [1981]).

halves of asymmetrical grain boundaries, suggesting an asymmetry of solute segregation in the boundary core (LEJČEK [1994]). A good correlation exists between the segregation enthalpy and the *effective interplanar spacing* d_{eff}: This form of spacing was introduced by D. Wolf for boundaries where two well defined planes join, e.g., $(001)_A/(011)_B$. d_{eff} is then given by $d_{eff}/a = (d_A + d_B)/2$. The boundaries characterized by high values of d_{eff} possess low absolute values of segregation enthalpy for a particular solute as is demonstrated in fig. 23 (HOFMANN and LEJČEK [1991], LEJČEK and HOFMANN [1995]).

4.5.3. Computer simulation of grain-boundary segregation

Measurements of chemical composition of individual grain boundaries in well-characterized bicrystals by AES represent a substantial step forward in understanding the anisotropy of grain-boundary segregation. However, the information provided by such measurements is still averaged over a relatively large interfacial area (at present, the lateral resolution limit of AES is about 20 nm) as compared to individual boundary sites which may possess different ability to solute adsorption. Due to the present limitations of experimental techniques (although methods such as high-resolution transmission electron microscopy, HREM, combined with electron-energy loss spectroscopy, or AP FIM, approach atomic resolution), the only method of evaluating solute segregation at individual boundary sites remains computer modeling.

Computer modeling has been successfully applied to the determination of atomic structure of grain boundaries in pure solids and to the development of the structural unit model (cf. ch. 9, § 2.2.7). Later, techniques such as molecular statics and dynamics

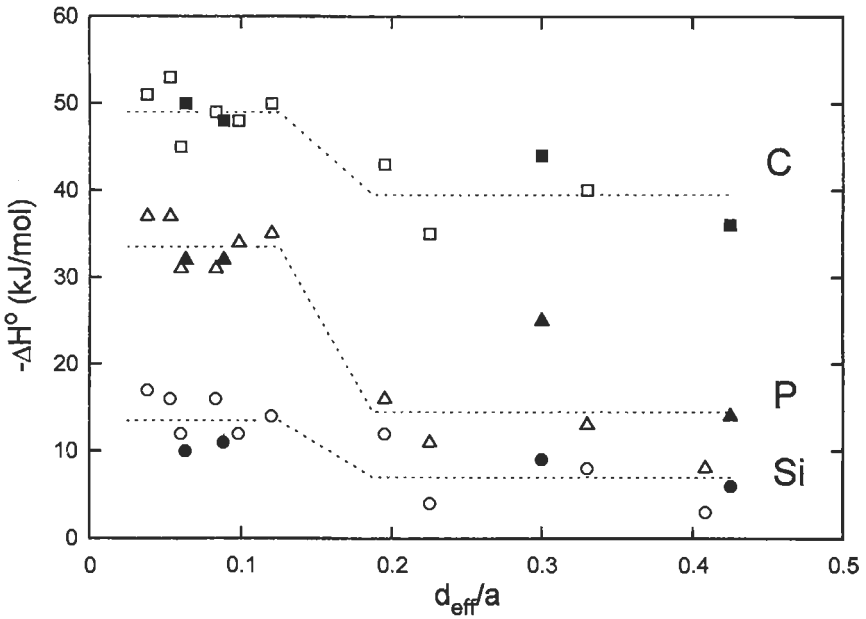


Fig. 23. Plot of segregation enthalpy ΔH_i^0 of carbon (□), phosphorus (△), and silicon (○) in α-iron vs. effective interplanar spacing d_{eff} normalized by the lattice parameter a, for some symmetrical (open symbols) and asymmetrical (solid symbols) tilt grain boundaries (after LEJČEK and HOFMANN [1995]).

or Monte Carlo calculations using many-body potentials of either the embedded-atom or Finnis–Sinclair types, and molecular orbital models, were extended to theoretically simulate equilibrium atomic structures of segregated grain boundaries.

Computer modeling of the structure of the $\Sigma = 5$, 36.9° [100] {012} grain boundary in Cu–Bi system revealed large differences in the segregation energy for particular sites. Because Bi tends to occupy the sites of the highest segregation energy, a two-dimensional (2-D) ordered structure occurs in the boundary layer where each Bi atom is surrounded by Cu atoms and vice versa (VITEK and WANG [1982]). The values of segregation energy of phosphorus and boron at {013} and {114} grain boundaries in α-Fe also show a high site sensitivity; however, their values calculated by HASHIMOTO *et al.* [1984] are much higher than those expected from experiments. More realistic values of the segregation energy of Mg, Al, Si, P, S and Cl at the sites at the exact plane of a $\Sigma = 5$, {013} symmetrical tilt grain boundary in α-Fe were calculated using a tight-binding type electronic theory of s-, p-, and d-basis orbitals by MASUDA-JINDO [1986]. Segregation energies of P and Si obtained in this way (–14.6 kJ/mol and –9.8 kJ/mol, respectively) are in excellent agreement with the values of segregation enthalpy of these elements experimentally determined at the same grain boundary by LEJČEK and HOFMANN [1991] (–13.3 kJ/mol and –8.5 kJ/mol, respectively). This agreement suggests that both these solutes segregate in a dilute Fe–Si(P) system at a specific type of sites at the boundary at all temperatures and therefore are characterized by specific segregation enthalpies.

References: p. 1284.

A very efficient and straightforward approach is based upon a point approximation for the configurational entropy, an Einstein model for vibrational contributions to the free energy with respect to atomic coordinates and composition of each site (NAJAFABADI *et al.* [1991]). The approach was applied to the study of Cu segregation at $\Sigma = 5$, (001) twist boundary in Ni–Cu alloys and revealed an oscillatory depth distribution of Cu in the equiatomic alloy (WANG *et al.* [1992]).

Computer modeling can also reveal an orientation dependence of grain-boundary segregation as is demonstrated in fig. 24 for the example of Au segregation at (001) twist grain boundaries of Pt. Absolute values of segregation enthalpy were found to increase with increasing rotation angle θ from 1.1 kJ/mol for $\theta = 5°$ to 11.7 kJ/mol for $\theta = 36.9°$ ($\Sigma = 5$). The average grain-boundary enrichment ratio β_b can be expressed by

$$\beta_b(T,\theta) = 1 + m(T)\sin(\theta/2) \tag{46}$$

where the slope $m(T)$ depends only on temperature (SEKI *et al.* [1991]).

There are numbers of experimental studies on grain-boundary segregation on the one hand and computer simulations of segregated boundaries on the other hand. Nevertheless, there is still a lack of direct correlations between these two approaches. The first attempt to make a direct comparison of HREM observation of a segregated grain boundary and its computer-simulated structure was performed very recently to interpret the atomic structure of the $\Sigma = 3$, $(111)/(11\bar{1})$ grain boundary in the Cu–Bi system (LUZZI [1991]). In a non-segregated state, this boundary is the symmetrical coherent twin boundary with the stacking

$$ABCAB\underline{\overline{C}}BACBA \tag{47}$$

where A, B and C represent individual {111} atomic planes in an fcc lattice and $\underline{\overline{C}}$ denotes the boundary layer. Usually, this boundary stays intact during solute segregation. However, Bi segregates at other grain boundaries and its segregation at non-coherent twin and $\Sigma = 9$ boundaries induces their transformation to $(111)/(11\bar{1})$ facets in which Bi atoms remain involved. The Bi-containing grain boundaries can be represented in the same way as the pure boundary but in the boundary layer $\underline{\overline{C}}$ one third of the Cu atoms is replaced by hexagonally arranged Bi atoms. However, the Bi atoms are centered outside this atomic plane so that the structure of this boundary can be represented more appropriately by

$$ABCAB\underline{\overline{C' - C(\mathrm{Bi})}}BACBA \tag{48}$$

where C' denotes a plane of Cu atoms in which one third of atoms is replaced by hexagonally arranged vacancies, and C(Bi) the plane of Bi atoms positioned above these vacancies. Due to the splitting of the {111} boundary plane C into two planes C' and C(Bi) containing Cu and Bi, respectively, this boundary is *not symmetrical* in the sense of mirror symmetry across the geometrical boundary plane (YAN and *et al.* [1993]).

4.5.4. Correlation between grain-boundary and free-surface segregation

Owing to the similar nature of solute segregation to grain boundaries and to free surfaces, measurements of free-surface segregation have often been used to estimate

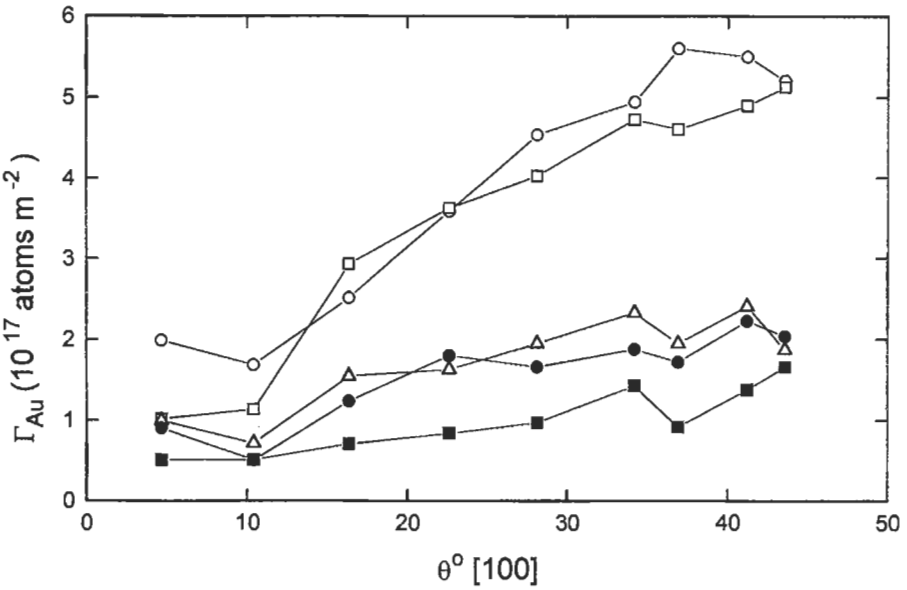


Fig. 24. Orientation dependence of the Gibbsian interfacial excess Γ_{Au} at (001) twist grain boundaries in an Pt–1at%Au alloy at 1900 K (○), 1500 K (□), 1300 K (△), 1000 K (●) and 900 K (■) (after SEKI *et al.* [1991]).

solute segregation at buried grain boundaries qualitatively or semiquantitatively. Because of the similar nature of solute segregation to both types of interfaces, the same formalism is generally used for a thermodynamic description of solute segregation. Since a high-angle grain boundary can roughly be considered as consisting of a complicated mixture of free surfaces and bulk binding states, the average free energy of general grain boundaries is about 1/3–1/2 of that for the free surface. Because solute segregation is associated with a lowering of the interfacial free energy, it is clear that grain-boundary segregation enthalpies are generally lower than those of surface segregation for the same system (HONDROS and SEAH [1977b]). On the basis of this similarity, SEAH and LEA [1975] proposed a method of quantification of grain-boundary segregation in a binary system. This method consists of the construction of a conversion diagram from the temperature dependence of surface segregation and a single measurement of grain-boundary segregation at suitable temperatures (see fig. 20, § 4.2.4.).

However, significant differences in segregation behaviour of the two types of interfaces can occur in three- or multicomponent systems resulting from site competition and solute interaction. For example, boron segregates significantly at grain boundaries in Ni_3Al whereas no sulphur is observed there. On the other hand, extensive segregation of sulphur but low segregation of boron occur at its free surfaces (WHITE *et al.* [1984]).

Qualitatively more pronounced differences between surface and grain-boundary compositions can be observed in systems with strong attractive interaction of two segregants resulting in the formation of two-dimensional surface compounds. In an

Fe–6at%Si alloy containing P, S, N, and C as trace impurities, cosegregation of Si and N results in a stable Si_xN_y surface compound at the (013) free surface thus preventing significant segregations of other elements. In contrast, enrichment of P and C, and depletion of Si characterize the segregation behaviour at the {013} grain boundary while N was practically not detected there. In the denitridized and decarburized alloy, phosphorus enrichment and silicon depletion were again observed at the grain boundary but pronounced sulphur segregation, accompanied by weaker phosphorus and silicon enrichments, dominates at the free surface (LEJČEK *et al.* [1993]). From a phenomenological point of view these qualitative differences may be considered as a *generalized anisotropy of interfacial segregation.*

4.6. The kinetics of segregation

Most models of the kinetics of segregation follow MCLEAN's [1957] approach. Solute atoms are assumed to segregate to a grain boundary from two infinite half crystals of uniform solute content or to a surface from one infinite half crystal. Diffusion in the crystals is described by Fick's laws and the ratio of the solute in the grain boundary to that in the adjacent atom layer of the bulk is given by the enrichment ratio, β, assumed to be constant and independent of the actual values of X_b and X_c. The kinetics of the segregation are thus described by:

$$\frac{X_b(t) - X_b(0)}{X_b(\infty) - X_b(0)} = 1 - \exp\frac{FDt}{\beta^2 f^2}\,\mathrm{erfc}\left[\frac{FDt}{\beta^2 f^2}\right]^{1/2}, \tag{49}$$

where $F = 4$ for grain boundaries and 1 for the free surface, $X_b(t)$ is the boundary content at time t, D is the solute bulk diffusivity, and f is related to the atom sizes b and a of the solute and matrix, respectively, by $f = a^3 b^{-2}$. For short times, eq. (49) approximates to

$$\frac{X_b(t) - X_b(0)}{X_b(\infty) - X_b(0)} = \frac{2b^2}{\beta a^3}\left[\frac{FDt}{\pi}\right]^{1/2}. \tag{50}$$

Equations (49) and (50) are, in fact, limiting extremes of a general problem. In practice β is only constant for dilute systems with low segregation levels. As segregation proceeds β generally falls as a result of saturation in X_b. If β starts high and falls rapidly as the segregation saturates, eq. (50) is valid up to saturation (LEA and SEAH [1977]). A detailed analysis for the saturation occurring in the Langmuir–McLean adsorption theory has been presented by ROWLANDS and WOODRUFF [1979]. Their analysis, re-interpreted in fig. 25, shows how the time dependence of the segregation changes from eq. (49) to eq. (50) as the final equilibrium segregation level $X_b(\infty)$ approaches the saturation level X_b^0. Calculations for particular systems in relation to temper brittleness are given in § 6.3. Measurements of $X_b(t)$ at surfaces by AES at different temperatures allow the determination of the bulk diffusion constant through eq. (50), as shown by HOFMANN and ERLEWEIN [1976] for Sn in Cu.

The more complex problem of the kinetics of segregation in ternary systems has also been studied by TYSON [1978]. Tyson uses McLean's approach but incorporates Gutt-

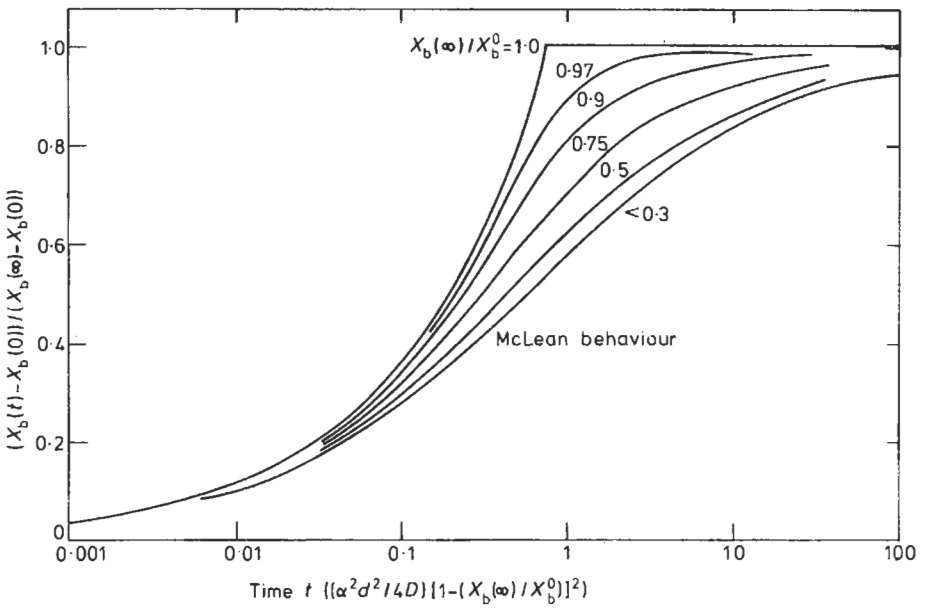


Fig. 25. The kinetics of segregation in binary systems for varying degrees of saturation in the final equilibrium level (after ROWLANDS and WOODRUFF [1979]).

mann's theory to define the grain-boundary concentration. For a ternary system with a dilute impurity and a non-dilute alloying element he finds that eq. (50) is often a good description if the final segregation level is near saturation and similarly eq. (49) is valid if the final level is low. However, for certain critical cases, in which the two interacting species have similar values of $D\beta^{-2}$, the segregation curves exhibit an intermediate plateau.

The effect of site competition and repulsive interaction between solutes on segregation kinetics in multicomponent alloys was analyzed by MILITZER and WIETING [1986]. This effect results in a replacement of a quickly segregating element by another, slowly diffusing but more surface active element. As a result, a non-monotonic course of segregation kinetics appears. Under strong repulsive interaction, the replacement can reach such an extent that an interface miscibility gap occurs. The time necessary for nucleation of new phases is generally only very short and does not delay the kinetics of segregation (MILITZER and WIETING [1987]).

The linear relationship between X_b and $t^{1/2}$ (eq. (50)) was derived assuming the bulk diffusion to be a controlling mechanism of segregation kinetics. In some cases, however, deviations from this behaviour have been observed. The diffusion can be accelerated e.g. via pipe diffusion. Then, the kinetics of interfacial segregation is composed of several contributions, e.g., the bulk diffusion term, $X_v(t)$, the pipe diffusion term, $X_p(t)$, and the term reflecting the enrichment of dislocations, $X_{pv}(t)$,

$$X_b(t) = X_v(t) + X_p(t) + X_{pv}(t). \tag{51}$$

References: p. 1284.

In this way, MILITZER *et al.* [1992] explained a complex segregation behaviour of sulphur at the (001) surface of an Fe–Si alloy.

The process of solute segregation generates composition changes in the vicinity of the interface. The generally encountered situation is illustrated in fig. 26, where this in-depth distribution was calculated for surface segregation of Sn in Cu for different annealing times. The course of the depicted curves was experimentally confirmed by sputter depth profiling (HOFMANN and ERLEWEIN [1978]).

The kinetics of desegregation have, in the past, been considered as much faster than the kinetics of segregation. This should not be so. The relationships are as given in eqs. (49) and (50), with $X_b(0)$ and $X_b(\infty)$ transposed (DU PLESSIS [1990]).

4.7. Non-equilibrium segregation

The above segregation process leads to a well defined equilibrium value of the segregation as time approaches infinity if the temperature and constituents are kept constant. This equilibrium is approached at a rate defined by the relevant diffusion processes as discussed in the previous section.

In non-equilibrium segregation phenomena, the enrichment disappears as time approaches infinity but it may be very significant at intermediate times. These non-equilibrium segregations arise through a number of discrete processes but all depend ultimately on the fact that changes in the materials environment cause the chemical potential of one the constituents to vary from place to place at a given concentration. This constituent then redistributes at a rate governed by diffusion theory, concentrating in certain regions in order to attain a constant chemical potential everywhere. In the meantime the spatial dependence of the chemical potential generally weakens also at a rate governed by diffusion theory, so that the constituent, having segregated, disperses again to become homogeneous throughout the material. This type of segregation is thus transient and is termed non-equilibrium but, in fact, includes many different separate mechanisms. (See also ch. 7, § 8.3.2). Another important point is that each of these mechanisms leads to segregation in a zone of considerably greater width around the appropriate interface than occurred with the equilibrium mechanism discussed earlier. Depending on the mechanism, this zone may range from nm to μm. It should now be recognized that, having established a zone of enrichment, equilibrium partitioning between this zone and the interface will inevitably occur so that the real situation will incorporate contributions from both equilibrium and non-equilibrium segregations. Below, the non-equilibrium aspects are discussed but the above caveat should be borne in mind.

4.7.1. Solute pile-up at growing precipitates

An early form of non-equilibrium segregation concerns the solute pile-up at the face of a growing carbide in steels. Owing to their low solubility in carbides, certain solutes such as Sn and Sb will be rejected as the carbide grows and will pile up on the matrix side of the matrix/precipitate interface. Immediately this happens, the excess atoms start to diffuse into the matrix. Thus, if the carbide grew a constant rate, the solute pile-up would grow asymptotically to a steady level. However, the carbide precipitate growth

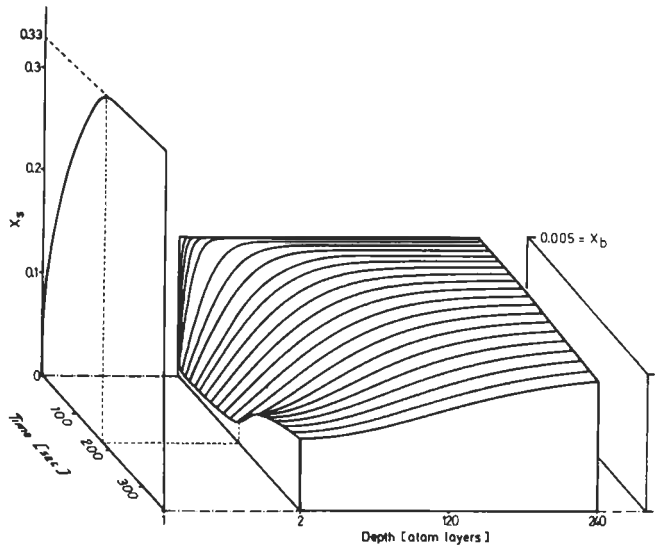


Fig. 26. Calculated time dependence of concentration-depth distribution near the surface for Sn ($X_c = 5 \times 10^{-3}$) at a Cu (111) surface (after HOFMANN and ERLEWEIN [1978]).

rapidly slows down and stops so that the pile-up disperses eventually to zero. In a practical situation, with the carbides growing as the material cools, the pile-up may be "frozen in". Precise calculations are difficult since the extent of the pile-up has been shown, by RELLICK and MCMAHON [1974], to be critically dependent on the cooling rate, the carbide growth rate and the solute partitioning between the carbide and the matrix. With large, rapidly grown carbides, enrichments as high as a factor of 100 may be encountered although much weaker effects are likely to be observed in practice.

4.7.2. Quench-induced segregation

Another form of non-equilibrium segregation which also occurs at interfaces has been analyzed by WILLIAMS *et al.* [1976] on the basis of a model proposed by AUST and WESTBROOK [1965]. In this model, cooling of the material causes the equilibrium vacancy concentration to fall so that vacancies flow to interfaces where they may be readily annihilated. Because solute atoms in the crystal lattice have an associated strain energy, as discussed in § 3.2.1, which may be relieved by pairing the solute atom with a vacancy, the vacancy flow has an associated solute flow. In their calculations for boron segregation during the quenching of 316 steel (17% Cr, 11% Ni stainless steel) from solution temperatures in the range 1173–1633 K, WILLIAMS *et al.* write the vacancy concentration as:

$$v = k_v \exp\left(-E_f/kT\right) \tag{52}$$

where $k_v (= 4)$ contains geometric and entropy terms and E_f, the vacancy formation energy is 1.4 eV (1 eV per atom = 96.5 kJ/mol). The concentration of boron–vacancy com-

plexes, v_B, obeys a similar relation but also depends on v and the boron bulk concentration, B,

$$v_B = Bvk_c \exp(E_B/kT), \tag{53}$$

where $k_c(= 12)$ contains geometric and entropy terms and E_B, the vacancy–solute binding energy, is 0.5 eV. In this system a significant population of complexes occurs at 1623 K but not at 1173 K and, since the diffusivity of the boron–vacancy complexes is higher at all temperatures than that of the vacancies or the boron atoms, significant boron segregation from the bulk level of 90 at ppm could be observed by autoradiography, as shown in fig. 27, for the higher temperature quenches at 50 K/s. These results were confirmed later by KARLSSON *et al.* [1988], who found in a similar material (17% Cr–13% Ni, 316L stainless steel) that the amount of segregated boron increases with increasing starting temperature and is larger for lower cooling rates (~10 K/s). It was also detected after the highest cooling rate 530 K/s. This was also confirmed by computer simulations (KARLSSON [1988]). Non-equilibrium segregation is observed after cooling from temperatures above 1350 K whereas equilibrium segregation dominated after rapid cooling from 1073 K. Effects of non-equilibrium segregation seems to be independent for the majority of grain boundaries but no segregation was detected at coherent twin boundaries (KARLSSON and NORDÉN [1988]). To differentiate between equilibrium and non-equilibrium segregation at grain boundaries of a Ni–In alloy, MUSCHIK *et al.* [1989] found AES depth profiles of the shape for intermediate time in fig. 26 to be indicative for non-equilibrium segregation above 1000–1100 K quenching temperatures.

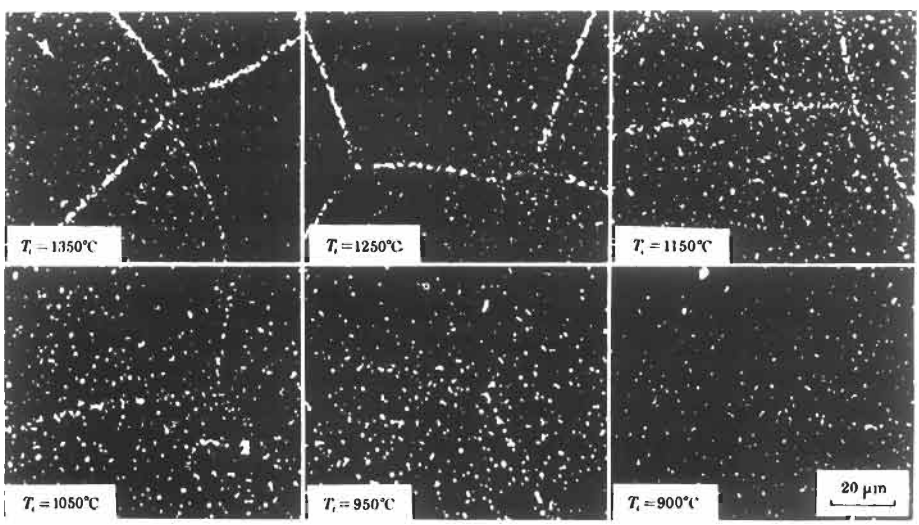


Fig. 27. Optical autoradiograph of austenitic 316 stainless steel containing 90 ppm boron solution heated at 1623 K for 1/2 hour, furnace-cooled to the temperature T_i in the range 1173 to 1623 K, held for 1/2 hour and cooled at 50 K/s to room temperature. Strong boron non-equilibrium segregation is visible for the higher temperature quenches (after WILLIAMS *et al.* [1976]).

Detailed calculations for Cr segregation in $2\frac{1}{4}$Cr1Mo steel by DOIG and FLEWITT [1981] show that, in that case, narrow distributions for Cr enrichment are obtained around the grain boundary for quench rates in excess of 2000 K/s with significant segregation occurring for quenches from temperatures in the range 1300–1400 K. Experiment and theory agree in indicating an excess Cr level of some 2% over a zone of 5 nm on either side of the boundary. The limiting conditions for the quench-induced segregation are presented in fig. 28 as a function of the initial temperature before quench and the quench rate. At quench rates above line 1 there is insufficient time for the solute–vacancy pairs to diffuse an adequate distance, whereas at quench rates below line 2 there is sufficient time for the solute to diffuse away from the segregation region and so outside these lines little segregation will be observed. Again, at initial temperatures below line 3 there is neither sufficient kinetics or sufficient vacancy population for the segregation to be significant. No high temperature limit is appropriate if the quench goes through the temperature range to line 3. Observations of significant quench-induced segregations are reviewed by HARRIES and MARWICK [1980].

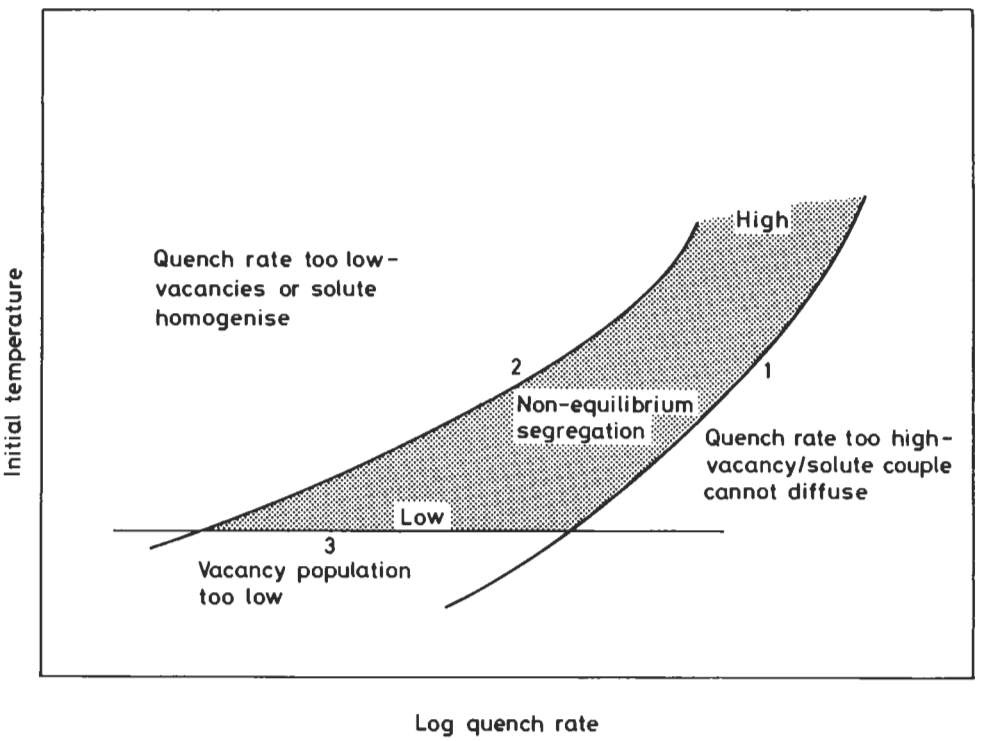


Fig. 28. Diagram of the limiting conditions for appreciable non-equilibrium segregation as a function of quench rate and temperature prior to quenching.

References: p. 1284.

4.7.3. Stress-induced segregation

A third form of non-equilibrium segregation occurs when the vacancy flux is not produced from a change in temperature but from an applied stress at constant temperature. The vacancy flux is directed by the stress to relieve the strain and consequently it in turn directs excess solute atoms, by the vacancy–solute coupling discussed above, to grain boundaries parallel to the stress axis. The source of the solute atoms is thought to be the other grain boundaries rather than the lattice, since the activation energy for grain-boundary diffusion is much lower than that for lattice diffusion. In steels the problem is complicated by the stress-assisted dissolution and reprecipitation of grain-boundary carbides which produces the first non-equilibrium segregation mentioned above due to pile-up. Changes in grain-boundary phosphorus segregation as a function of tensile and compression stresses in $2\frac{1}{4}$Cr steels have been observed and interpreted by SHINODA and NAKAMURA [1981] although, at present, the interpretative model is not fully complete.

4.7.4. Segregation at moving grain boundaries

It has been known for a long time that an atmosphere of solute atoms exerts a drag on a moving grain boundary (ch. 28, § 3.4.1). However, up to now only very little is know on the reverse effect, i.e. on the non-equilibrium segregation of solutes at a moving grain boundary during recrystallization. This is despite the fact that the solute enrichment effects can be much stronger than those of equilibrium segregation. For example, a non-equilibrium segregation of boron was observed on moving grain boundaries in an Fe–3 mass%Si alloy ($\beta_b = 1.6$) although no equilibrium segregation was detected at static boundaries under the same conditions (ZHANG *et al.* [1994a]). Boron segregation at moving grain boundaries is intensified with continuing recrystallization and persists with the new boundaries moving forward. With decreased velocity of moving boundaries at the end of the recrystallization process boron enrichment at grain boundaries decreases and eventually disappears. Obviously, a property responsible for segregation changes due to grain-boundary movement is the width of the boundary.

The non-equilibrium segregation is facilitated by increasing width of a moving interface. The broadening of the boundary results from its interaction with lattice dislocations disappearing there during recrystallization. The width δ of the moving boundary can be expressed by:

$$\delta = d_0 + V\tau d_\tau \Delta\rho \tag{54}$$

where d_0 is the width of the static boundary, V is the boundary velocity, τ is the relaxation time for dislocation annihilation by moving grain boundary, d_τ is the increment of average width of unit area of the boundary during τ, and $\Delta\rho$ is the difference of the dislocation density in deformed and new grains. The enrichment ratio at moving boundaries is then

$$\beta_b = 1 + \frac{\delta}{V}\exp\left(\frac{U_0}{RT} - 1\right) \tag{55}$$

where W is the width of grain-boundary area of increased solute concentration, and U_0 is the potential of solute atoms near the grain boundary. According to eqs. (54) and (55),

one obtains $\beta_b = 1.7$ for boron segregation at moving grain boundaries in Fe–3mass. %Si during recrystallization at 1273 K after 20% deformation (ZHANG *et al.* [1994b]).

4.7.5. Radiation-induced segregation

This topic is treated in chapter 7, § 8.3.2 and chapter 18, § 4.7.

5. *Segregation-related physicochemical properties*

The influence of the interfacial microchemical state on physical metallurgical properties, such as creep-rupture life and temper embrittlement, operates through associated effects on certain basic physicochemical properties which in turn determine the metallurgical behaviour. The main physicochemical properties which are influenced by segregation processes fall into three categories: interfacial energetics, kinetics and cohesion. These are considered in turn below. Wherever these basic interfacial parameters are found to be involved with metallurgical phenomena, clearly these phenomena are influenced by segregation even though no explicit mention of the fact may occur in this chapter or, indeed, elsewhere.

5.1. Interfacial energetics

The close association between interfacial free energies and interfacial composition is discussed in § 2 in terms of the Gibbs Adsorption Theorem. We note that the thrust of much of the effort on the determination of interfacial free energies was originally intended to derive the values for interfacial composition. Both this thermodynamic approach and the new direct spectroscopy techniques can now be used interchangeably by way of parameters such as the interfacial activities or enrichment ratio, through which patterns of enrichment behaviour for a wide range of systems and for various interfaces may be compared and interrelated.

However, there are situations where the treatment of a microstructural or metallurgical problem requires a knowledge of the exact interfacial energy values relevant to the conditions which prevail. Examples of this are the nucleation of precipitates or second phases at grain boundaries, where a determining parameter is the actual interfacial free energy between the boundary plane and the second phase that is being nucleated. Again, in the study of the nucleation and stabilization of cavities formed along grain boundaries during creep, an important quantity is the value of the surface free energy pertaining to the cavity wall. The surface and grain-boundary free energies of most pure metals have been measured experimentally by the techniques discussed earlier and fair assessments have been made for others, on the basis of a knowledge of bulk thermochemical properties with which empirical relationships exist. These have been summarized elsewhere (HONDROS [1969, 1978]). We note, as expected from eqs. (25) and (27), that $\gamma_s = 0.16 H_{sub}$ at around 1500 K, where γ_s and H_{sub}, the heat of sublimation, are measured per mole (OVERBURY *et al.* [1975]). At 0 K γ_s increases to 0.25 H_{sub}, as expected (TYSON [1975]).

In multicomponent crystalline solids, the main determinants of the interfacial energies are the presence of segregants, the temperature, and the crystalline orientation. The

evaluation of the interfacial free energy relevant to the modelling of a metallurgical event is difficult because of its sensitivity to the above parameters, in particular, adsorption of solute species. The direct experimental determination of that quantity is technically often very difficult and indirect.

For binary dilute alloys, reasonable experimental data exist for several systems on the change in absolute surface free energies with increasing bulk content of the alloying element: these are summarized in fig. 29 (HONDROS and MCLEAN [1974]). The depression in γ_s up to the limit of solid solubility is system dependent, however, the estimate usually adopted is about 40%. Furthermore, we note that the data were measured on polycrystalline samples and therefore they refer strictly to the average surface free energies. The anisotropy of surface energy is small for pure metals, but it may be quite significant for alloys in which surface enrichment is involved.

In some cases, extremely low quantities of very surface-active elements in the vapour phase may produce significant effects on the interfacial free energies. Figure 30 shows as an example, the sensitivity of the surface free energy of Fe–3mass%Si at two temperatures to the potential of oxygen in the vapour phase. In order to obtain and hold such low values of oxygen potential, it is required to work with complex gas phase mixtures involving H_2O in H_2. Furthermore, these experiments were conducted on foils showing a (100) texture so that the values refer substantially to this crystal plane. (See also ch. 28, § 4.4.2).

In a subsequent study aimed at providing fundamental data on the surface-energy-driven secondary recrystallization in sheet alloy, MILLS *et al.* [1973] showed that the surface energies of crystals exposing the principal poles differed in a complex but significant manner as a function of temperature and oxygen adsorption, so that by manipulating these variables, each of the textures (111), (110) or (100) could be controllably grown.

For the effect of temperature on the interfacial free energy, considering a single component isotropic solid, eq. (68), ch. 5 reduces to:

$$-\frac{d\gamma}{dT} = S^{\sigma}, \tag{56}$$

where S^{σ} is the interfacial entropy. Thus for pure elements, the temperature dependence of the interfacial free energy is negative. Theoretically, this is computed to be about –0.5 mJ/m^2K which is close to the value conventionally adopted from experimental data, –0.33 mJ/m^2K (HONDROS [1969]).

In contrast to the above, fig. 30 clearly shows that in a system of more than one component the temperature dependence of the surface free energy can be positive. This is consistent with the thermodynamics of surfaces. By differentiating eq. (68), ch. 5 with respect to T, a complex expression may be obtained containing two terms which describe the temperature dependence of the interfacial free energy in a binary system. The first term, relating to entropy, is always negative, while the second term relates to adsorption and is positive. (The latter term refers to the effect of "boiling off" of a segregate with increasing temperature, where it is expected that the amount of adsorption would decrease with increasing temperature, i.e., $d\Gamma/dT$ is negative.) This shows that in a system of more

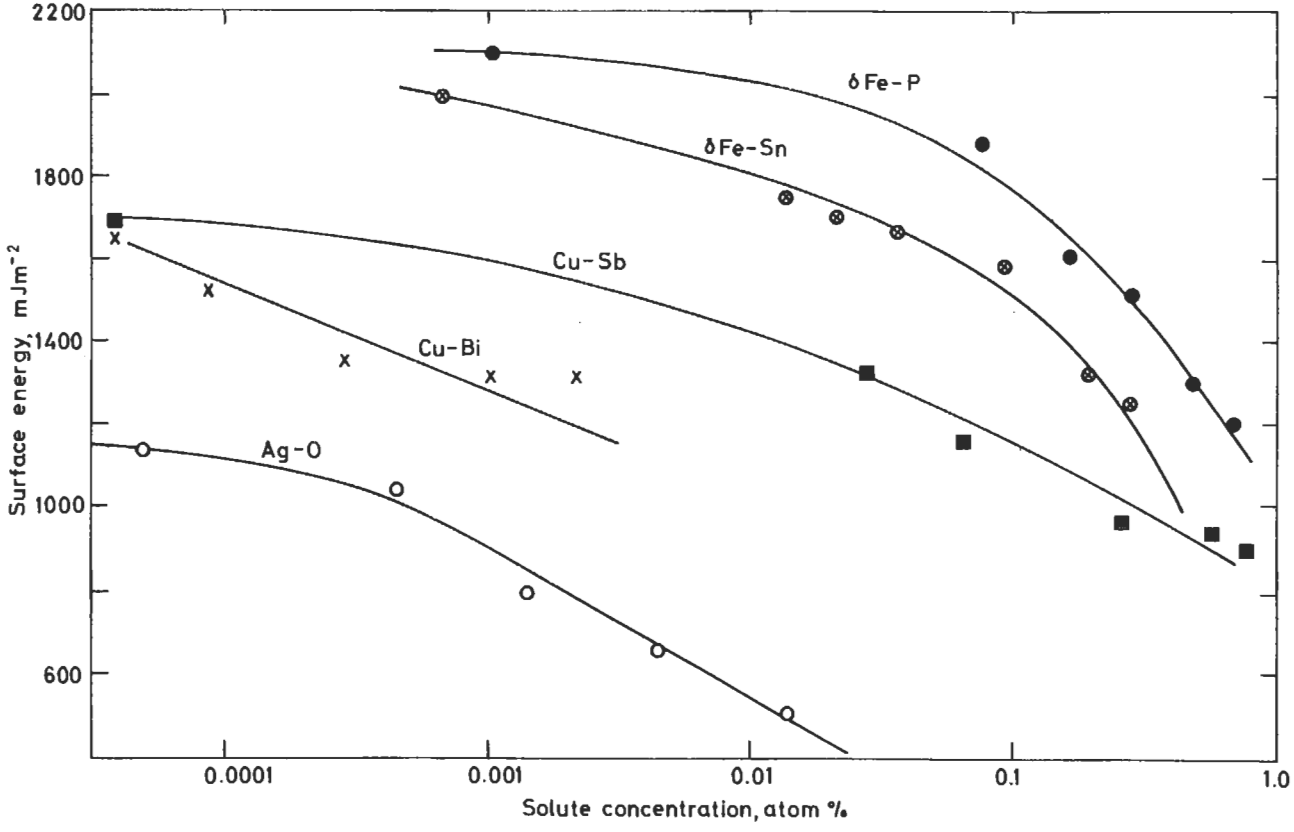


Fig. 29. Dependence of the surface free energy on bulk composition for a number of dilute systems: the host element is shown first in each pair.

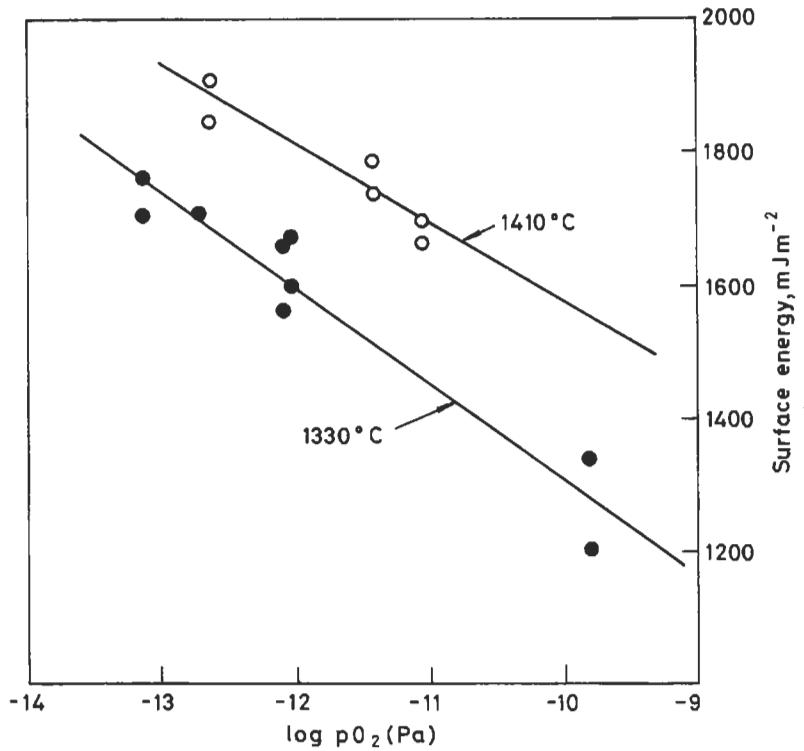


Fig. 30. Dependence of the surface free energy of the (100) plane in Fe–3%Si on oxygen potential and temperature (after HONDROS and STUART [1968]).

than one element, the interfacial energy may *increase* or decrease with increasing temperature, depending on the relative strengths of the two terms of entropy and adsorption.

The above discussion indicates the complexities involved in assessing the precise interfacial free energy value for a given physicochemical situation. There are very few systems for which this quantity has been determined at the level of rigour shown in fig. 30. For more complex interfaces, such as those between a second phase and the matrix, the experimental problems are even more severe. As a general guideline for binary systems, a knowledge of the amount of surface excess of an identified species from a know bulk concentration provides an estimate of the associated change in interfacial energy, through the application of the Gibbs Adsorption Theorem in a reversed manner.

Where the form of the adsorption is known, an approach that has been used for both free surfaces and grain boundaries is to integrate the isotherm. Thus, if the Langmuir isotherm holds, this may be combined with the Gibbs Adsorption Theorem, which for a dilute system yields the equation for grain-boundary free energy:

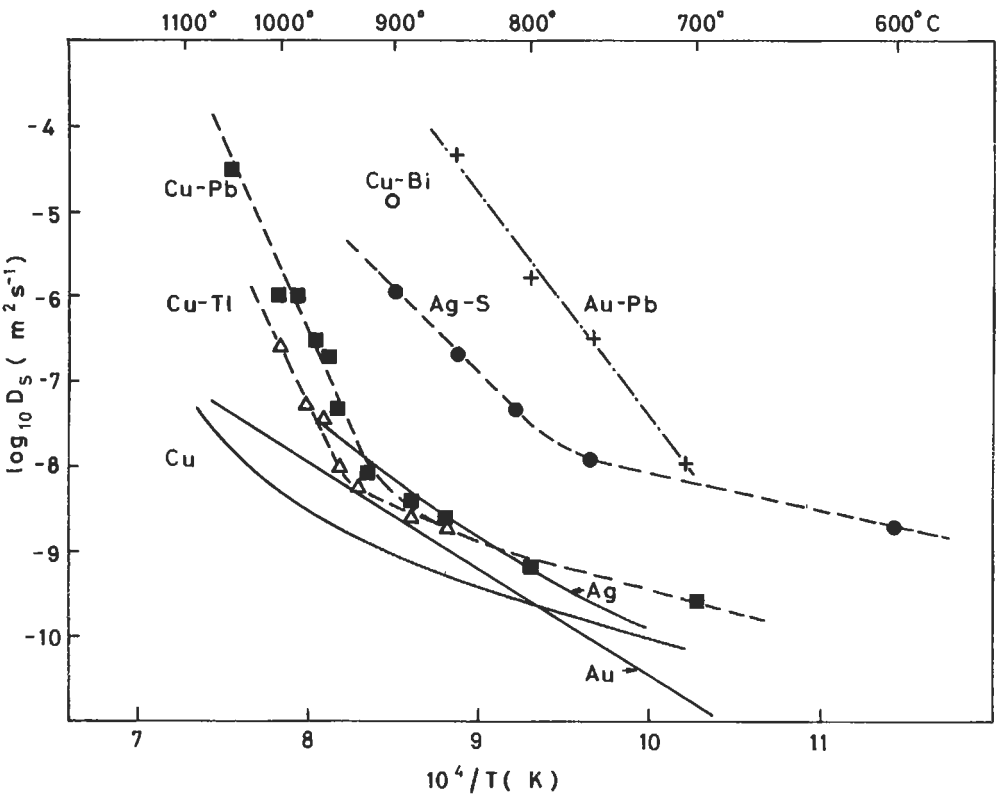


Fig. 31. Surface self-diffusion in Cu, Au and Ag (solid curves) and the effects of adsorption of a number of species (broken curves) (after DELAMARE and RHEAD [1971]).

$$\gamma_b = \gamma_b^0 + RT\Gamma_b^0 \ln\left(1 - \frac{\Gamma_b}{\Gamma_b^0}\right), \tag{57}$$

where the quantities in Γ are defined as above and γ_b^0 is the grain-boundary energy of the pure metal. This approach was used by PETCH [1956] in a similar form (but retaining the gas pressure term), in order to determine the effect of adsorbed hydrogen in lowering the surface energy of iron in a theory of hydrogen-induced fracture, and also by SEAH [1980c], in a study of impurity-aided creep cavitation embrittlement. The enrichment quantity Γ_b required for eq. (57) may be obtained directly by spectroscopic measurements or assessed through the enrichment factor β_b and its dependence on the inverse of the solid solubility (see eq. (15) and fig. 14). This yields a change in free energy $\Delta\gamma$ which in many modelling studies is often the quantity required: an absolute value of γ for the appropriate conditions is assessed by referring to the standard state value for the energy of the pure component.

References: p. 1284.

5.2. Surface and grain-boundary kinetics

The free surfaces of solids offer rapid diffusion paths for the transport of matter. For example, the grooves that are formed at the intersection of grain boundaries with free surfaces when metals are held at high temperature (commonly referred to as "thermal etching" (see fig. 5)), owe their origin to the interchange of matter by surface diffusion. In general, this mechanism predominates in solid state phenomena where the microstructural features involve dimensions of the order of μm. Thus surface diffusivity in an important mechanism in the early stages of sintering, where the necking of the spherical particles takes place.

Considering the effects of surface-adsorbed elements, a good body of evidence affirms that adsorption will generally accelerate the surface diffusivity, while in some specific cases suppression will occur (GJOSTEIN [1966]). The interesting data in fig. 31 for example, compiled by DELAMARE and RHEAD [1971], show the effects of a number of impurities, either arriving from the vapour phase or segregating from the bulk, on the surface diffusivities of copper, gold and silver. The Arrhenius plots for the pure metals are shown as solid curves, whereas each of the broken curves refers to the metal substrate containing an adsorbed layer of impurity under controlled conditions. The second element of each pair denotes the adsorbed species. Clearly, in some cases such adsorbed layers induce spectacular increases in the self-diffusivity of the substrate metal and at the highest temperatures, these values can approach a factor 10^4 above that of the clean substrate. As yet there is no agreement on the mechanism for this remarkable effect: one theory proposes that the uppermost atomic layer of the surface consists of a mixture of adsorbate and adsorbent, constituting in effect a two-dimensional compound with liquid-like properties. From this theory, it follows that adsorbates which form compounds with low melting points should lead to this accelerated surface diffusivity. An earlier theory proposed a "skating" mechanism in which the substrate atom is thought to climb on top of the adsorbed layer, where it may migrate freely with long jump distances of the order of 100 atom spacings, a hypothesis which could explain increases in surface diffusivities of the order of 10^4. This overlayer model of adsorption diffusion now appears to be superseded by the mixed layer or two-dimensional liquid model, which is seemingly supported by Low-Energy Electron Diffraction (LEED) studies on adsorbed monolayers of lead and bismuth on copper.

Whereas the copious studies on surface self-diffusion (ch. 7, § 7.4) fall within the context of a fundamental understanding of mechanisms probably relevant to the important process of catalysis, the somewhat fewer studies of the effect of segregation on grain-boundary self-diffusivity have a more immediate applicability to physical-metallurgical phenomena. We now turn to consider the latter.

Grain boundaries are paths for rapid self-diffusion and it is commonly understood that the transport of matter along such interfaces should be rate-controlling in a number of important metallurgical processes, in particular in temperature regimes of about $0.5T_m$ or below. In general, grain-boundary diffusivity data which refer to the pure elements have been employed in the modeling of various processes in metallurgy. Clearly, the use of such readily available data could introduce serious distortions if transport along grain boundaries

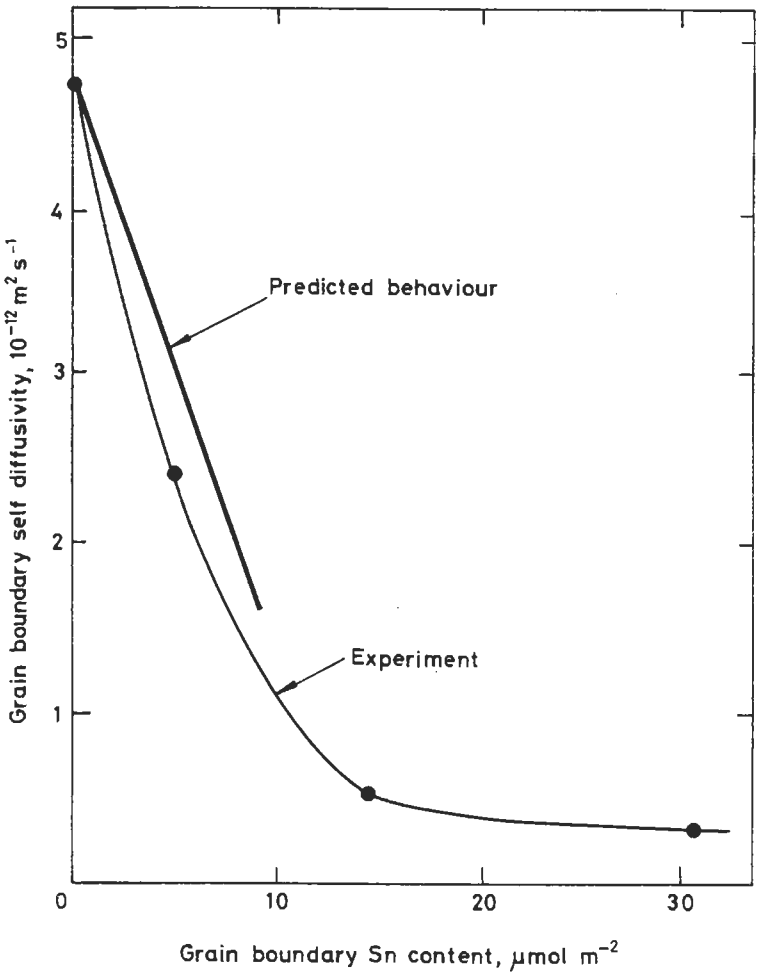


Fig. 32. The measured and predicted dependence of the grain-boundary self-diffusion in Fe–Sn alloys at 888 K on the amount of tin segregated at the boundary (after BERNARDINI *et al.* [1982]).

were to be sensitive to the microchemical aspects of the interface. Below we examine the theory and evidence for the role of solute segregation on grain-boundary diffusivities.

A pure grain boundary is characterized by a high free energy, being a structural zone with a low volume density of matter compared with the bulk lattice. From such simplistic considerations, we might expect that this should lower the activation energy barrier for substitutional self diffusion, so that atom transport is easier along these paths than in the bulk crystalline lattice. This concept has been expressed in a semi-quantitative manner by GIBBS and HARRIS [1969] who related the activation energy for grain-boundary transport Q_b to that for volume diffusion Q_v by the expression

References: p. 1284.

$$Q_b = Q_v - N\alpha'\gamma_b a^2, \tag{58}$$

where a^2 is the cross-sectional area of an atom*, N is Avogadro's number and α' a structure factor which depends on whether diffusion is substitutional or interstitial. This predicts sensibly that increasing γ_b will lower the activation energy for grain-boundary transport and that for a highly coherent interface such as a twin boundary, with a low free energy, the activation energy for diffusion would be little different from that of the lattice.

Consider the presence of a solute in the base solvent with accompanying high grain-boundary enrichment of solute. If the segregation-associated change in grain-boundary energy is $\Delta\gamma_b$, it follows that the activation energy for grain-boundary diffusivity in the presence of segregation, Q_b^{seg}, is related to that for the pristine boundary by:

$$Q_b^{seg} = Q_b + N\alpha' a^2 \Delta\gamma_b. \tag{59}$$

Hence, according to this general model, segregation increases the activation energy for grain-boundary diffusion compared with the pure state. In effect, the segregation permits the sites of high compression and tension to be relaxed, reducing the net free energy of the boundary and increasing the volume density of the boundary by the occupation of vacant sites. Thus, from the diffusional point of view, the material in the boundary zone approaches the character of the material in the bulk lattice. The above approach allows estimates of the magnitude of the effect. Thus, when phosphorus segregates to the grain boundaries of pure iron, the grain-boundary energy is reduced by about 400 mJ/m^2 at its maximum (HONDROS [1965]); substituting in the above equation and including a value of 2 for the structure factor appropriate to vacancy diffusion, the activation energy for grain-boundary diffusion in pure iron is raised from 175 kJ/mol to 200 kJ/mol.

The above account is probably satisfactory for rough estimates of expected effects, and supports the view that the consequence of segregation is the elimination of the boundary sites having a high degree of disorganization, thereby decreasing grain-boundary transport with respect to that in a pure boundary. The subject has been developed analytically by BERNARDINI *et al.* [1982] who derived a more rigorous expression linking the change in diffusivity to the amount adsorbed at the grain boundary. This treatment takes as its point of departure of the theory of BORISOV *et al.* [1964] who derived the basic relationship between grain-boundary energy and grain-boundary diffusivity, given by the expression

$$\gamma_b = \frac{mkT}{2a_{sv}^2}\left[\ln\frac{D_b\delta}{D_v a_{sv}} - \ln m\right], \tag{60}$$

where D_b is the grain-boundary diffusivity, D_v is the volume bulk diffusivity, a_{sv} is the size of the solvent atom, δ is the width of the grain boundary, and $\delta = ma_{sv}$. This equation

* In this chapter the atom size, a, is derived from $\rho Na^3 = M$, where ρ is the bulk density and M the atomic weight of the solid.

is transformed by expressing γ_b in terms of Γ_b, the amount of segregation. By combining eqs. (58) and (60), we effectively extend the Borisov treatment to a binary system in which grain-boundary segregation occurs. This yields:

$$\frac{D_b^*}{D_v^*} = \frac{D_b}{D_v}\left[1 - \Gamma_b/\Gamma_b^0\right]^{2a_{sv}^2/ma_{su}^2}, \tag{61}$$

where D_b^* is the grain-boundary self-diffusion coefficient in the presence of segregation, D_v^* is the coefficient due to the presence of the solute in the bulk, D_b and D_v are the corresponding values in the pure solvent, a_{su} is the atomic size of the solute in the particular solvent, and m is the quantity, retained from the Borisov equation (60), which relates to the number of effective atom layers in the boundary. The above equation has now been confirmed experimentally in iron containing tin, by directly measuring the diffusivity constant using radio-tracer techniques and also by determining the amount of grain-boundary segregation using Auger electron spectroscopy on exposed grain-boundary surfaces. Figure 32 shows how the measured grain-boundary self-diffusivity of iron decreases with increasing levels of tin present at the boundary, at 888 K. Similar data were obtained for a number of temperatures and for both solvent and solute diffusivities along the boundary. In the same figure, the straight line indicates the effect predicted by the application of eq. (61), the treatment necessarily restricting the prediction to the dilute approximation.

Equation (61) may be simplified further by expressing D_v^* as a linear function of the bulk concentration of solute, X_c, following a number of theoretical and experimental investigations on systems involving substitutional self-diffusion, as discussed in ch. 7, § 4.1:

$$D_v^* = D_v(1 + b_v X_c), \tag{62}$$

where b_v is a temperature dependent coefficient representing the variation of D_b with the bulk solute concentration X_c. Thus eq. (61) can now be expressed as

$$D_b^* = D_b\left[1 + (b_v - 2\alpha)X_c\right], \tag{63}$$

where α is a concentration ratio. This defines the quantity actually measured in boundary diffusion measurements in the presence of segregation as $\alpha D_b\delta$, and not the standard $D_b\delta$ where δ is the boundary width. To express eq. (63) in a more accessible form, we introduce the grain-boundary enrichment factor defined earlier, β_b (eq. (15)), which is closely related to α by:

$$\alpha = \beta_b \frac{a_{sv}^2}{ma_{su}^2}. \tag{64}$$

Therefore, the overall effect may be expressed by:

$$D_b^* = D_b\left[1 + \left(b_v - \frac{2\beta_b a_{sv}^2}{ma_{su}^2}\right)X_c\right]. \tag{65}$$

References: p. 1284.

In summary, this equation expresses the effect of segregation (contained in the enrichment factor β_b) on grain-boundary diffusivity. This is offset by the effect of the solute on volume self-diffusivity. The equation then predicts that grain-boundary transport of both solvent and solute atoms may be reduced or increased depending on the balance between the opposing effects implicit in β_b and b_v. Since b_v is generally of the order of 10–100 and β_b for highly surface active species is of the order of 1000–10 000, in general, the effect of the high enrichment factor will dominate in the presence of strongly segregating species and this will reduce the grain-boundary diffusivity. An enhancement of the solute diffusivities is caused by the increased matter transfer due to accumulation of the solute atoms in the grain boundary. It is only an apparent enhancement due to α being greater than unity and is not due to a real increase in D_b. The magnitude of the enhancement depends on the interaction between grain boundary and solute (KAUR and GUST [1988]).

The ability to predict the effect of segregation on grain-boundary diffusivities is a development which permits the quantitative examination of a number of metallurgical phenomena in which these kinetics are known to be rate-determining: some of these applications will be considered in § 6.

5.3. Grain-boundary cohesion

As was noted in the Introduction, historically the most important manifestation of grain-boundary segregation has been the effects on grain-boundary cohesion as expressed in a number of well known forms of intergranular fragility. This continues to provide much of the focus for work in this field. The central problem is the role of solute atoms on atomic cohesion at the interface. This is now recognized to be a wider issue, extending to aspects of *ad*-hesion (i.e., cohesion between dissimilar pieces of matter), as in that between a metal and an oxide or another phase such as a precipitate or even a polymer.

The theoretical and experimental framework for understanding adhesion in such systems and the role of interfacial microchemical interactions is hardly developed and this section will be restricted to grain-boundary cohesion which is now quite well advanced. However, in connection with adhesional phenomena in metals, HONDROS [1980b] has attempted to systematize the data available for metals adhering to oxides and other phases, and it appears that the classical concept of *work of adhesion*, expressed in the formulation of Young and Dupré as the algebraic sum of the two free surface energies and the interfacial energy, may be applied to these solid state systems. This would predict that segregation to metal/oxide interfaces, by decreasing the interfacial energy should increase the work of adhesion but scant data are available to test this hypothesis.

Because of its technological importance, much of the work on grain-boundary cohesion has been concerned with effects of additions in iron, and it is now well established that the elements Cu, Zn, Si, Ge, Sn, P, As, Sb, Bi, S, Se, Te, and Mn all cause intergranular weakness in iron. A similar set also weaken copper and nickel. We now consider the two questions "why do certain elements weaken grain boundaries?" and "what is the relative embrittling potency of these elements?".

A number of theories have been proposed to account for the effect of segregants on grain-boundary cohesion. All have a common approach but differ in their interpretation

of the way to carry out the numerical calculations. In 1957 McLEAN first proposed that in low-temperature intergranular fracture, the total work of fracture is the sum of the ideal work of fracture, γ, representing the energy to disrupt the bonds across the interface, and the irreversible plastic work of deformation, γ_p, which is absorbed in the crystal lattice as dislocations and phonons over a region well away from the boundary. Although γ_p is much larger than γ it is easy to see that if γ is increased, the forces at a crack tip must also be increased to propagate the fracture and hence γ_p will increase. Thus, for a given grain boundary, γ and γ_p are directly related. The precise relation will depend on the orientations of the two grains and the grain boundary to the stress axis as well as on the microstructure. Indeed, analyses by McMAHON and VITEK [1979] and by JOKL *et al.* [1980] show that γ_p is proportional to γ^n where n is in the range 2–5. Thus, to answer the above questions we must understand the effects of segregants on the ideal work of fracture of the grain boundary.

The ideal work of fracture of a clean grain boundary with energy γ_b^0 was given by McLEAN [1957] as:

$$\gamma^0 = 2\gamma_s^0 - \gamma_b^0, \tag{66}$$

where γ_s^0 is the fracture surface energy of the clean material. An analysis of γ_s^0 and γ_b^0 values for many metals by McLEAN and HONDROS (SEAH and LEA [1975]) shows that, typically, $\gamma_b^0 \approx \gamma_s^0/3$ so that the ideal work of fracture of clean grain boundaries is approximately 5/6 of that of an average plane in the crystal. This does not mean that the grain boundaries will always fail in preference to cleavage, since the orientation of the slip planes and the surface energy of the low-index cleavage planes make cleavage generally more likely than intergranular failure in clean bcc metals (HONDROS and McLEAN [1976]).

More important than the choice of brittle fracture modes is the balance between brittle and ductile failure. The fracture behaviour of a 0.2% carbon steel, as a function of the test temperature, is shown schematically in fig. 33. At high temperatures, where dislocations move easily, the yield stress is low and the steel fails in a ductile manner. As the test temperature is reduced, the dislocation movement, governed by a term of the form $\exp(-Q/RT)$, becomes more difficult and the material eventually fails in a brittle manner. The temperature at which this occurs is the ductile–brittle transition temperature (DBTT). If the grain boundaries contain the embrittling segregants, the boundary cohesion may be sufficiently lowered so that the transition temperature is raised and the otherwise ductile material becomes brittle.

An analysis of the ideal work of fracture of grain boundaries with segregants, γ, has been made by HIRTH and RICE [1980] using thermodynamic arguments. They show that the reduction in the ideal work of fracture, where there is no redistribution of the segregant species, is given by:

$$\gamma = \gamma^0 - \int_0^{\Gamma_b} \left[\mu_b(\Gamma) - \mu_s(\Gamma/2)\right] d\Gamma, \tag{67}$$

where $\mu_{b,s}(\Gamma)$ are the chemical potentials of the solute species in equilibrium with a level segregation Γ mol/m^2 at the grain boundary and the free surface, respectively. This relation may be re-expressed:

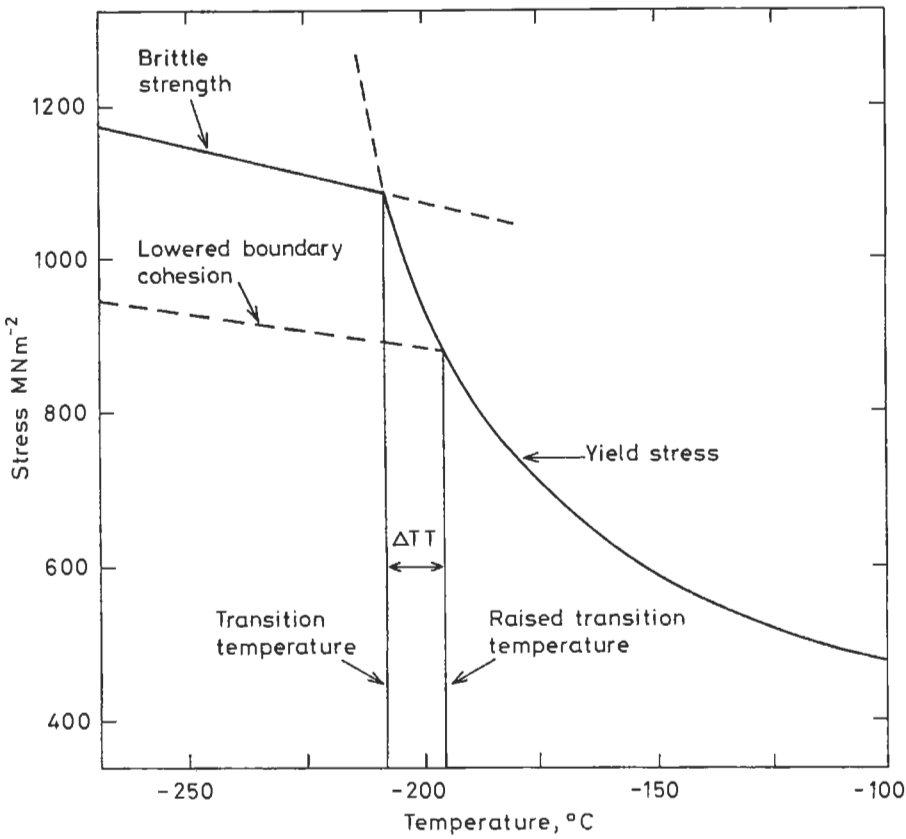


Fig. 33. The effect of temperature on the tensile properties of a 0.2% carbon steel showing, schematically, the lowering of grain-boundary cohesion and the resultant increase in the ductile–brittle transition temperature (after SEAH [1976]).

$$\gamma = \gamma^0 + \Gamma_b(\Delta G_s - \Delta G_b) + RT\int_0^{\Gamma_b} \ln \frac{\Gamma_b^{\max} - \Gamma}{2\Gamma_s^{\max} - \Gamma}\, d\Gamma \tag{68}$$

where Γ_b is the segregation level at the grain boundary, $\Gamma_{b,s}^{max}$ are the maximum segregation levels at the boundary and surface, respectively, and the terms in ΔG refer to the binding energies of segregants, which, it should be noted, are negative quantities (RICE and WANG [1989]). In the dilute limit, the last right-hand term is equal to $-RT \ln 2$ and usually, it can be neglected ($RT \ln 2 = 5.76$ kJ/mol at 1000 K). Thus we see that the ideal work of fracture falls linearly at first with the level of segregation, and that elements which segregate more strongly to free surfaces than to grain boundaries reduce the ideal work of fracture and vice versa. By measuring ΔH_s, ΔH_b and Γ_b under identical conditions in Ni_3Al bicrystals with artificial cavities at the boundary, OTTERBEIN *et al.*

[1993] applied eq. (68) to obtain the decrease of the grain-boundary fracture strength by segregation of sulphur.

An alternative approach which gives the same result in the quasi-chemical pair bonding approximation has been given by SEAH [1980b]. In this approach the actual bond energies across the boundary before and after fracture are calculated in the manner described in § 4.2.2. In this way:

$$\gamma = \gamma^0 + \Gamma_b \left\{ \frac{Z_g}{Z\Gamma_b^0} \right\} \left\{ H_A^{sub*} - H_B^{sub*} - \frac{Z\Omega}{a_A^2} \right\}, \tag{69}$$

where the terms are as described in § 4.2.2. and, in addition Z_g corresponds to Z_v across the grain boundary (approximately $\frac{5}{6}Z_v$), Γ_b^0 is the value of Γ_b at one monolayer, a_A is the atomic diameter of the segregation atom A and H_A^{sub*} is the sublimation enthalpy of A atoms in J/m^2. Calculations may be easily carried out for eq. (69) in the regular solution approximation, but only small errors are involved if Ω is ignored. In this case it is clear that the lower the value of H_A^{sub*} for the solute, the more will a segregant embrittle. Figure 34 shows a compilation of H^{sub*} values for all elements. It is clear that the well recognized segregants, Sb, Sn, S, P, Mn, Si and Cu all with H^{sub*} values below that of iron shown by the dashed horizontal line, will all embrittle iron, whereas Mo and C will improve the cohesion; this is confirmed experimentally (SEAH [1981,1990]).

Detailed studies of the fracture properties of Ni_3Al (LIU *et al.* [1985]) show that B has a strong effect inhibiting intergranular failure. The results of fig. 34 show that, if B segregates to grain boundaries, this is exactly what should occur. WHITE *et al.* [1984] show that B does segregate to the boundaries and that they are strengthened. The theory of HIRTH and RICE [1980] predicts the same result if ΔG_s has a smaller magnitude than ΔG_b. This too is shown by WHITE *et al.*'s [1984] data where no surface segregation is observed.

The effect of individual segregants on boundary cohesion can also be considered from the viewpoint of atomic bonds at an interface. Presence of a foreign element at a grain boundary leads to a redistribution of electron charge density. In the case of electronegative sulphur or phosphorus atoms the charge is drawn from the surrounding metal atoms to the impurity atom and consequently, the adjacent metal–metal bonds are weakened (BRIANT and MESSMER [1982]). The bond parallel to the boundary is stronger than that perpendicular to the interface. This results in a reduction of cohesion across the boundary and thus, in grain-boundary embrittlement. In the case of the less electronegative elements such as boron and carbon, relatively strong covalent-like bonds exist with metal in direction perpendicular to the boundary whereas much weaker bonding occurs with metal atoms in parallel direction thus enhancing grain-boundary cohesion (WU *et al.* [1992]).

As will be shown in the next section, it is not sufficient to consider solely the effect of segregants on the ideal work of fracture, as described above. At a crack tip it is the choice between dislocation emission and bond breaking that decides whether ductility or brittle fracture occurs. It can be shown that embrittling segregants increase the ease of dislocation emission at the crack tip, so offsetting the total embrittling effect discussed

above. The correction is small but allows the correct prediction of the absolute embrittlement of cooper by bismuth (SEAH and HONDROS [1983]).

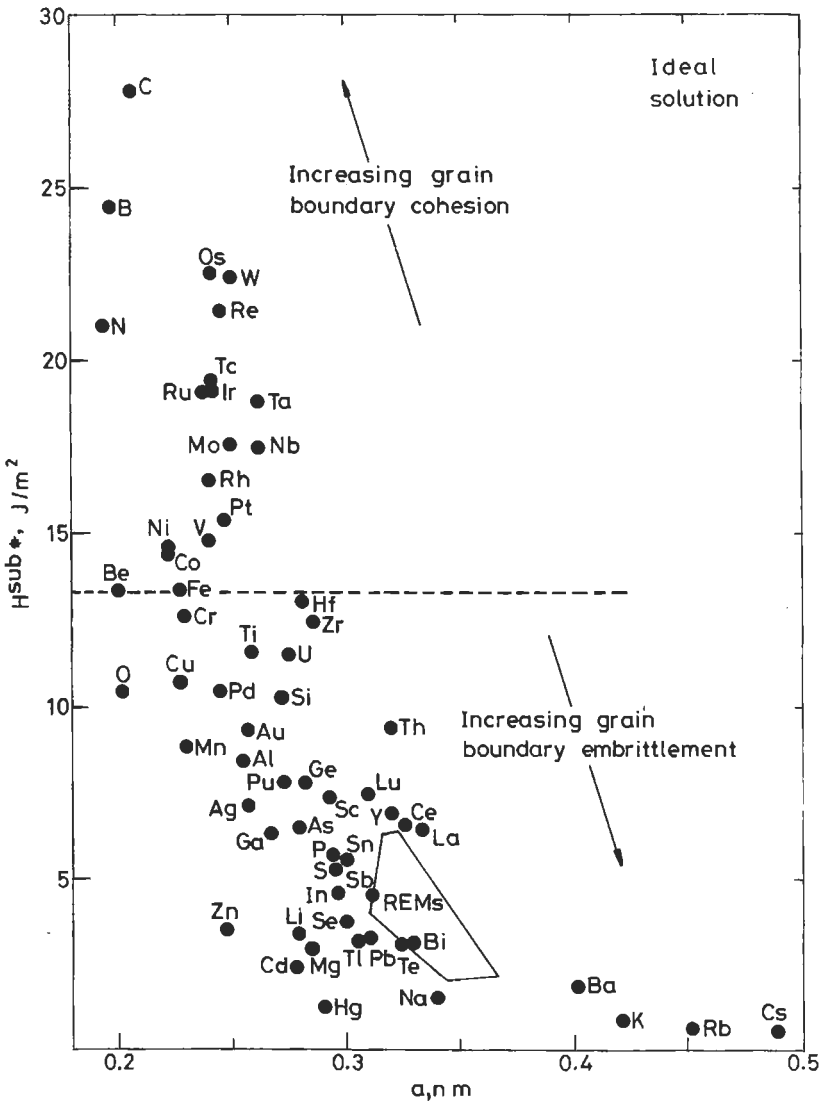


Fig. 34. The general embrittlement–ductility plot for matrix and segregant elements in the ideal solution approximation. Any element may be chosen as matrix and then those elements with higher values of H^{sub*} increase the grain-boundary cohesion and vice versa. As an example iron may be taken as the matrix material (after SEAH [1980b]).

6. Metallurgical phenomena affected by segregation

Interfacial segregation influences metallurgical phenomena through its impingement on several basic physicochemical properties, involving the energetics, kinetics or the cohesion of the interfaces, which were treated in the previous section. This is illustrated schematically in fig. 35. For example, grain-boundary segregation will in general retard the boundary diffusivity and this in turn will affect the rate of diffusion creep or the rate of grain-boundary cavity growth involved in creep embrittlement. Again, because it retards grain-boundary ionic mobility in polycrystalline oxides, this should also inhibit the rate of surface oxide growth. In some instances, the particular metallurgical property may be affected through changes in more than one of the basic physicochemical properties considered: for example, the rate oxidation of nickel alloys containing aluminium may be affected through the kinetics of ion transport across or through the boundaries, and in addition the integrity of the oxide layer is enhanced by improved interfacial adhesion resulting from the presence of segregated species at the metal/oxide interface.

As fig. 35 clearly indicates, numerous metallurgical and microstructural phenomena in alloys are influenced by interfacial microchemistry: in some cases this may be a determining factor while in others it may be simply one of a number of contributory factors. In the space of this chapter, we cannot consider all of these phenomena in detail and so some key references are cited in the caption of fig. 35. Below we consider for each of the important basic physicochemical properties treated in § 5 (i.e., the first three columns in fig. 35) a significant example of its impact in physical metallurgy, following which we mention briefly a number of other metallurgical phenomena for which evidence is accumulating of effects related to microchemistry.

6.1. Surface free energy change: role in creep cavitation

In microstructural terms, the creep-rupture life of alloys serving at high temperature under stress may be controlled by the nucleation, growth and coalescence of microcavities which develop at grain boundaries, in general those normal to the stress axis. A number of contending micromechanisms have been proposed for the nucleation of these cavities, and while there is very strong evidence that in certain engineering alloys these are initiated at sites where slip planes intersect second phase particles such as carbides or manganese sulphide inclusions, cavities have nevertheless been observed in quite pure metals. Many observations in the literature strongly affirm that impurity adsorption at both the grain boundary and the surface of the cavity wall will affect creep-rupture properties. Depending on the temperature and stress regime, grain-boundary segregation may affect the process of cavity growth through its influence on the grain-boundary diffusivity, but in addition, because segregation to the inner surface of the cavity reduces its surface energy, it will also affect the nucleation propensity.

Among the many examples, we cite the work of TIPLER and MCLEAN [1970] who showed metallographically and by creep measurements that small quantities of antimony in cooper increase the amount of grain-boundary cavitation and also lower the creep rupture life compared with that of the pure material. KRAAI and FLOREEN [1964] showed

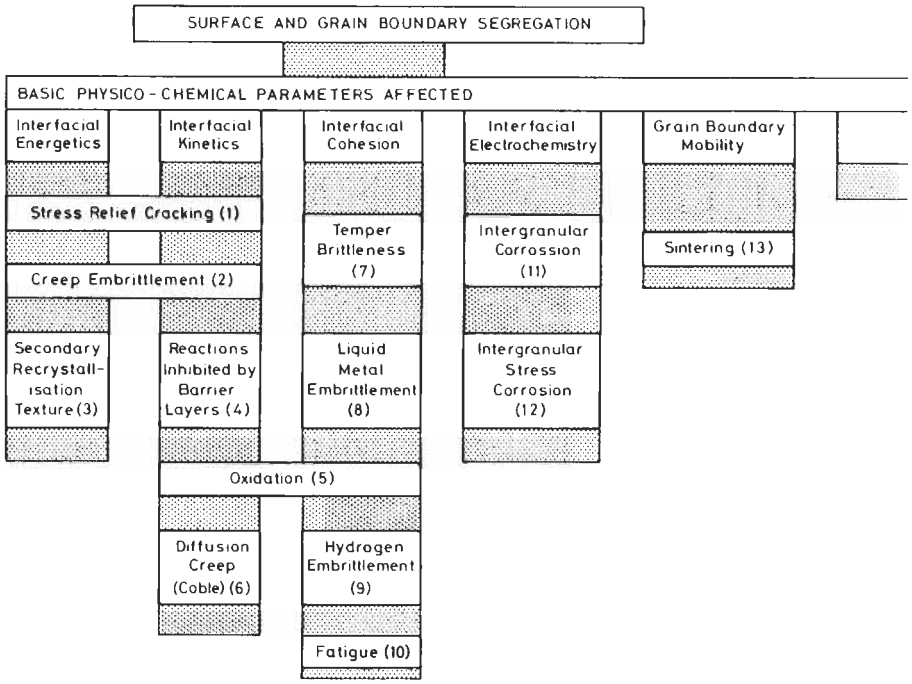


Fig. 35. Some metallurgical phenomena affected by physicochemical parameters influenced by surface- and grain-boundary equilibrium segregation. Key references for the interested reader are: (1) BATTE *et al.* [1980]; MCMAHON [1992]; (2) TIPLER [1980]; GEORGE and POPE [1989]; (3) MILLS *et al.* [1973]; (4) NICOLET [1978]; HAYES and GRIEVESON [1975]; (5) WHITTLE and STRINGER [1980]; (6) HONDROS and HENDERSON [1983]; MATSUKO *et al.* [1991]; (7) BRIANT [1988b]; MCMAHON [1991]; (8) NICHOLAS and OLD [1979]; LYNCH [1989]; (9) LATANISION and OPPERHAUSER [1974]; LYNCH [1988]; (10) SURESH *et al.* [1981]; SUNDARSHAN and LOUTHAN [1987]; (11) KUPPER *et al.* [1981]; WEI and GRABKE [1986]; (12) LEA and HONDROS [1982]; BRUEMMER [1989]; (13) NANNI *et al.* [1976]; GERMAN and D'ANGELO [1984].

that the ductility of nickel is reduced by a factor of ten by the presence of sulphur of up to 20 ppm in the bulk, while THOMAS and GIBBONS [1980], and OSGERBY and GIBBONS [1992] demonstrated the remarkable effects of trace impurities on the creep performance of nickel-base superalloys. Figure 36 shows the substantial reduction of creep rupture ductility in these alloys caused by ppm levels of lead and tellurium. At the maximum impurity level shown in this figure, the deleterious effect of each impurity was associated in every case with a much higher level of grain-boundary cavitation, of the order of 5–10 times greater than that observed in the pure alloy. In Nimonic 105 containing 34 ppm of lead, the amount of grain-boundary damage as measured by the number of cavities per mm^2 was ten times greater than in the alloy with the base composition. Similar effects were observed under presence of bismuth.

In addition to the considerable metallurgical evidence in which impaired creep life is associated with the presence of impurities, which presents circumstantial evidence for an

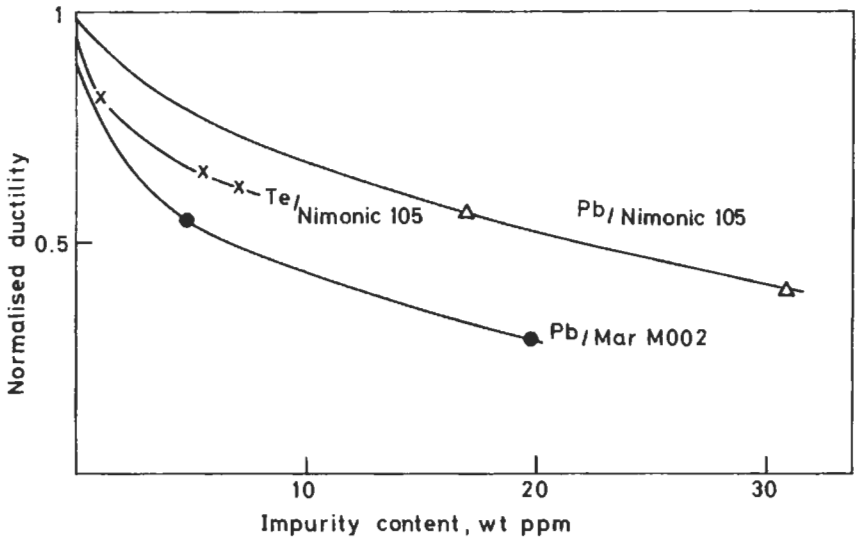


Fig. 36. Reduction in creep rupture ductility in nickel-base superalloy induced by trace levels of impurities.

effect through the interfacial microchemistry, there are important observations using Auger Electron Spectroscopy on the creep grain-boundary fracture surface which show a considerable accumulation of the suspected contaminant. An elegant example of selective segregation through the application of the high resolution (0.2 μm) Auger spectroscopy in shown in fig. 37 (FRANZONI *et al.* [1981]) for a commercial $2\frac{1}{4}$Cr1Mo steel, creep tested at 873 K and subsequently fractured at liquid nitrogen temperature for examination. The figure shows the high magnification scanning electron image of a grain-boundary area and the corresponding Auger spectroscopy maps for the elements Sn and Sb, indicating a simultaneous segregation of these elements only on the cavity surfaces. The non-cavitated areas of the boundary showed only P which was in turn significantly lower within the cavity.

In considering the role of interfacial microchemistry on cavitation, we assume that the cavities have been successfully nucleated while the precise mechanism for nucleation will not concern us. We consider an applied tensile stress, σ, which performs an amount of work $\frac{4}{3}\pi r^3\sigma$ in generating a spherical cavity of radius r. The associated surface energy of the cavity is $4\pi r^2\gamma$ at it is simple to show that as a cavity grows from a very small size, the net energy that has to be supplied first increases and then, after reaching a balance point, decreases, at which critical point the radius is given by r_c, where

$$\sigma = 2\gamma_s / r_c. \tag{70}$$

At this value, the competition between the surface energy which tends to sinter the pore and the applied stress is in balance. When the pore radius reaches a value greater than r_c, energy will be saved as it grows to a microscopical cavity. If the surface segregation of the active species is sufficiently rapid, a pore nucleus with a radius smaller than r_c

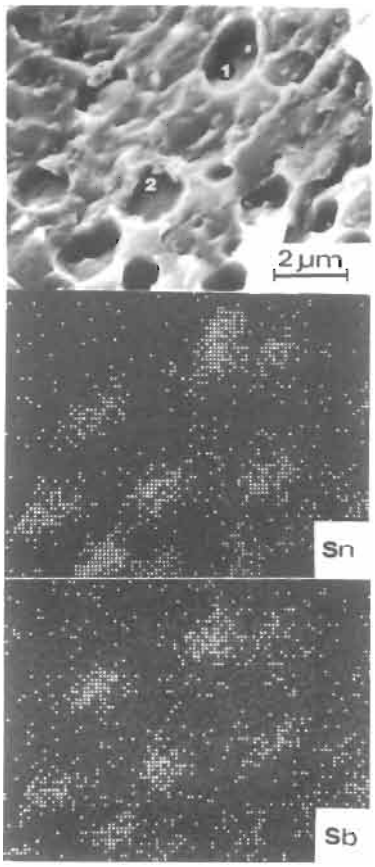


Fig. 37. SEM image of a creep cavitated grain-boundary area in $2\frac{1}{4}$Cr1Mo steel and the corresponding AES maps for Sn and Sb showing segregation on cavity surfaces (after FRANZONI *et al.* [1981]).

may survive according to the above equation, because of the reduction in γ_s, typically by a factor of two. The net effect of this impurity segregation is that a greater density of nuclei per unit area of grain boundary will achieve stability and hence they will be able to grow: the experimental evidence cited above in connection with the incidence of segregation to cavity walls strongly affirms this view.

The detailed mechanism of surface-energy induced cavitation damage has been analyzed by MCLEAN [1981], specifically for the alloy Nimonic 80A. The steep function of the curve reproduced in fig. 38 shows the quantity N, the number of grain-boundary cavities in a cubic millimetre of material per 1% creep strain, as a function of r, the cavity radius. For the pure alloy, substituting for γ_s at a stress of 150 MPa, the critical value of r is 27 nm. At this stress, the cavities which are stable and which will contribute to the creep damage are those which fall to the right of r_c under the solid line, shown as

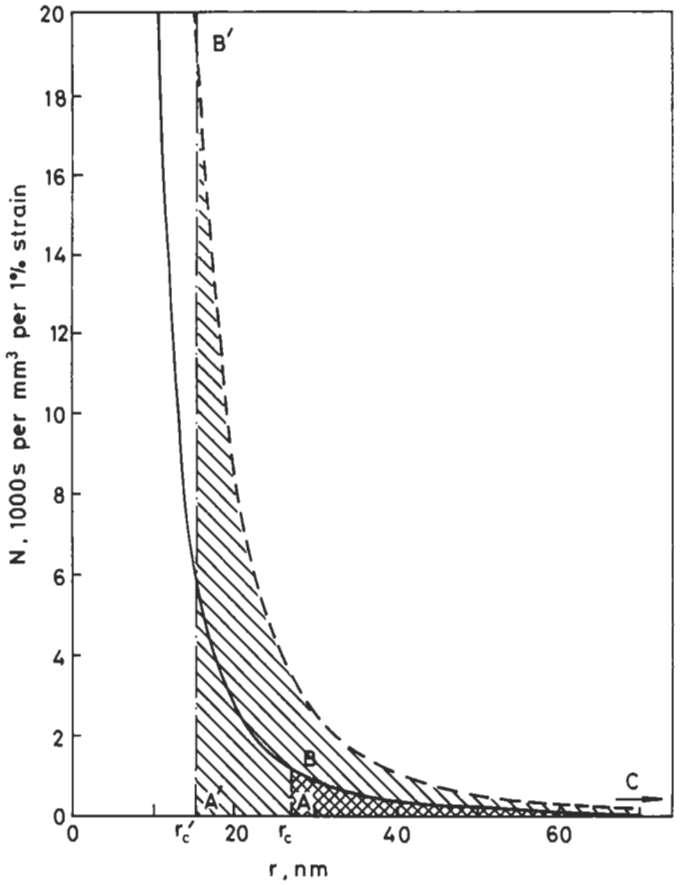


Fig. 38. Theoretical distribution for the density of cavities, N, in Nimonic 80A as a function of the cavity radius, r indicating the areas (hatched) in which cavities will contribute to the creep damage. (Note, there are approx. 100 mm^2 of grain boundary per mm^3 of material here (McLEAN, private commun.)).

the cross-hatched area ABC in the diagram. McLean integrated the equation representing the curve between $r_c = 27$ nm and $r \to \infty$, and found that the number of cavities in this area is about $1.6 \times 10^{13}/\text{m}^3$ per 1% strain. In the first instance, segregation to the cavity wall surface affects this process as we have seen by reducing γ_s, thereby reducing the critical stable pore size to r_c'. This is marked in the diagram at 15.6 nm, corresponding to a 42% reduction in γ_s. There is also a second effect to be considered, associated with segregation to the grain boundary itself. The consequent reduction in grain-boundary energy leads to a reduction in intergranular cohesion (eq. (66)), and, irrespective of the precise mechanism of nucleus generation, a lower stress concentration is required to nucleate a cavity in a segregated boundary than in a clean boundary (HONDROS and McLEAN [1974]) so that more cavities will be generated for a unit strain. Thus the curve

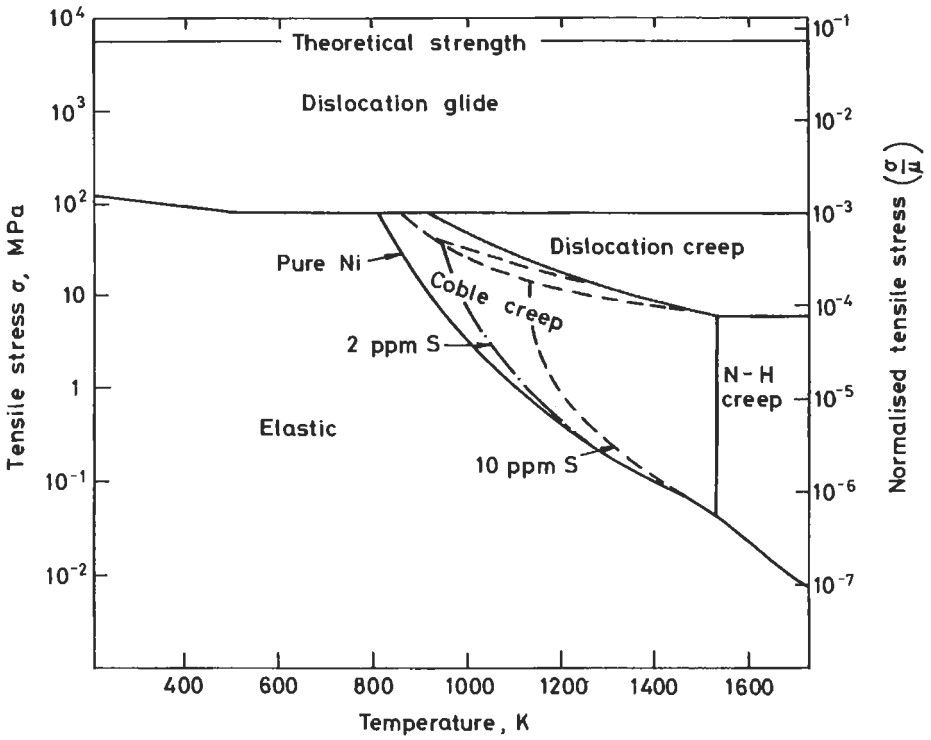


Fig. 39. Theoretical deformation map as a function of tensile stress and temperature where Coble creep dominates and the effect of low levels of sulphur on the boundaries (broken lines), for a grain size of 32 μm, and strain rate of 10^{-8} s.

for nucleation of cavities in the presence of grain-boundary segregation lies at a high value and is shown as the dashed curve in fig. 38. The two effects of segregation are cumulative with respect to cavity stabilization, and all cavities which fall in the hatched area A'B'C to the right of r_c' now contribute to the creep damage. It is difficult to ascribe accurately the position of the dashed curve representing the role of grain-boundary segregation and the ease of decohesion. For an order of magnitude increase in the creep cavity density in the presence of a segregant, the dashed curve corresponds to N values about three times those associated with the solid curve. Thus, while at this stage the model appears somewhat qualitative, the quantities are plausible and consistent with some experimental observations.

6.2. Grain-boundary diffusivity: role in diffusion creep

It is now well established that at lower temperature regimes and under light stresses, the predominant mode of deformation of polycrystals is dominated by diffusional creep, in this case, by mass transport along the grain boundaries. In the course of studying sintering mechanisms in alumina, COBLE [1963] recognized this deformation mode and

accordingly modified the equation for diffusional creep associated with the names of NABARRO [1948] and HERRING [1950] in which volume diffusivity dominates: this is analogous to the Herring equation and is of the form

$$\dot{\varepsilon} = \frac{B_2 \sigma \Omega \delta D_b}{d^3 kT}. \tag{71}$$

Here $\dot{\varepsilon}$ is the strain rate, σ the tensile stress, Ω the atomic volume, d the grain size, and B_2 a numerical constant which varies with grain geometry. It is appreciated that both lattice- and grain-boundary diffusion contribute independently to diffusional creep, and the general equation involving contributions from both volume- and grain-boundary diffusion is given by:

$$\dot{\varepsilon} = \frac{\sigma \Omega}{d^2 kT}\left(B_1 D_v + \frac{B_2 \delta D_b}{d}\right). \tag{72}$$

Clearly, because the activation energy for boundary diffusion is less than that for lattice diffusion and because of a greater sensitivity to grain size, the *Coble creep* mechanism predominates at the lower temperatures and in fine grained material. ASHBY [1972] showed that a generalized deformation mechanism map may be produced by plotting stress as a function of homologous temperature (see ch. 22, § 11). Applying a version of the above constitutive equation, the *Ashby map* indicating the fields of predominance for various deformation mechanism is reproduced in fig. 39 and shown as solid lines, for nickel. This indicates clearly the important tracts where Coble creep predominates. Such maps have been computed on a basis of diffusional data derived from pure metals. If impurity species in the bulk alter the lattice diffusivity, or more importantly, if by enriching at the grain boundaries they affect the boundary transport, this should be reflected in significant shifts in the boundaries on such maps. We examine here to what extent the Coble creep mechanism is affected by the presence of interfacially active impurities and how this is reflected in the positions of the boundaries on the Ashby map.

The predictive eq. (65) developed earlier shows the dependence of grain-boundary diffusivity on grain-boundary segregation as expressed in the enrichment factor β_b. We insert this into eq. (72) above and limit it to the deformation regime where grain-boundary transport is rate-determining.

From these equations we compare the strain rates $\dot{\varepsilon}$ and $\dot{\varepsilon}^*$ resulting from a change of D_b to D_b^* and D_v to D_v^*; hence:

$$\frac{\dot{\varepsilon}^*}{\dot{\varepsilon}} = 1 + b_v X_c - \frac{2a_{sv}^2}{ma_{su}^2}\beta_b X_c, \tag{73}$$

where the symbols are as defined in § 5.1. This may be simplified further: noting that in general b_v has a value of 10^2 and considering very dilute concentrations of highly surface-active species (X_c of the order of 10^{-4}), the second term containing b_v may be ignored in this instance. Hence in a binary system in which the second component is very dilute and also very surface-active, the grain-boundary diffusion creep rate at a given temperature is related to the diffusional creep rate of the pure solvent in the first approximation by:

$$\frac{\dot{\varepsilon}^*}{\dot{\varepsilon}} = 1 - \frac{2a_{sv}^2}{ma_{su}^2}\beta_b X_c, \tag{74}$$

There exist practical systems in which the dilute approximation implicit in the above equation should apply. For example, consider the effect of small impurity additions of sulphur on the nickel deformation map shown in fig. 39 in connection with the boundaries of the Coble creep mechanism. For the application of eq. (74), it is required to know the solid solubility of the impurity in the parent lattice at the temperature of interest, and also the relevant atomic sizes. Thus the quantity β_b is derived by substitution in the solubility law of eq. (15) described earlier. In the case of nickel, a number of dilute impurities at the ppm level may introduce significant effects on Coble creep rates. The computed effect of trace additions of sulphur on the extent of the Coble creep regime in nickel is superimposed on the Ashby map in fig. 39 as the dashed lines, here specifically for a grain size of 32 μm and a critical strain rate of 10^{-8}/s. This shows that by introducing quantities as low as 2 and 10 ppm of sulphur in nickel, the boundaries of the Coble creep regimes are shifted substantially. These large effects predicted for sulphur, magnesium and other dilute impurities (HONDROS and HENDERSON [1982]) demonstrate the important role of interfacial microchemistry on grain-boundary diffusion phenomena and how in this particular case, the impurities induce a beneficial effect by inhibiting grain-boundary transport.

6.3. Interfacial cohesion: role in temper-brittleness

Temper-brittleness is one of the oldest known forms of brittle intergranular failure. It occurs in low alloy steels incorporating alloying elements such as Cr, Ni and Mn when they are slowly cooled from temperatures above 773 K. High-purity alloys do not exhibit this embrittlement and it has been shown that the cause can generally be associated with the adventitious levels of P, Sn, Sb and As impurity, which are typically in the 50–200 ppm range. It has been recognized that Cu is also deleterious and this is important since much present day low alloy steel has copper contents as high as 0.15%. Analyses of commercial steels show that temper-brittleness is, in practice, mainly caused by the grain-boundary segregation of phosphorus whereas in nickel-containing steels, tin also contributes. The phosphorus comes from the original steel making, and is expensive to remove to very low levels. The tin arises from recycled scrap plated steel. The role of individual elements in temper embrittlement is summarized by BRIANT [1988b] and MCMAHON [1991].

It is instructive to consider the analysis of one very well documented instance of temper brittleness investigated at the Watertown Arsenal (CARR and *et al.* [1953]). As we have mentioned earlier, as segregation increases in a steel, the ductile–brittle transition temperature (DBTT) also increases until it rises above the ambient temperature and brittle failure will occur in service. Studies of the DBTT through Charpy impact measurements therefore provide a characterization of the heat treatments that will cause temper brittleness. Figure 40 shows the DBTT measurements for SAE 3140 steel (1.26% Ni, 0.77% Cr, 0.39% C, 0.015% P) heated to 948 K and water-quenched, followed by iso-

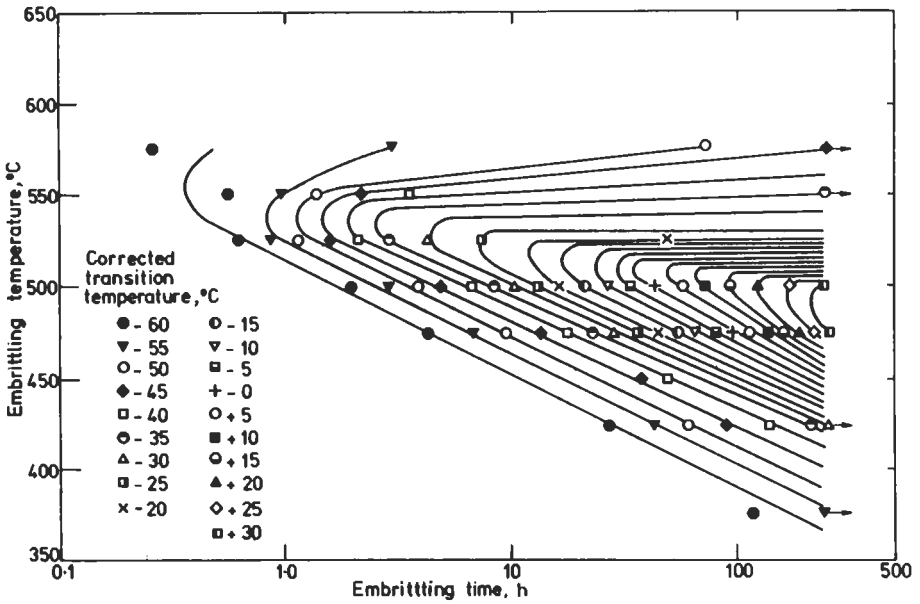


Fig. 40. The time–temperature diagram for embrittlement of SAE 3410 steel corrected to constant hardness (after CARR *et al.* [1953]).

thermal heat treatments at increasing times in the critical temperature range 648–848 K.

In fig. 40 we observe several of the characteristic features of temper-brittleness. The embrittlement occurs most rapidly around 773–823 K, and reheating above 873 K rapidly removes the embrittlement. Material quenched through the temperature range 673–873 K will not be embrittled, whereas large-section castings or forgings necessarily cooled

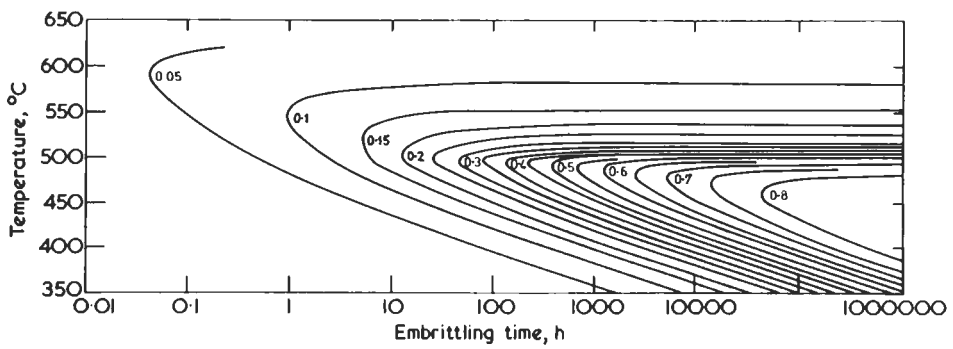


Fig. 41. The time–temperature diagram for the segregation of P in the SAE 3410 steel of CARR *et al.* [1953]. The numbers against the curves denote the P segregation occurring as a result of the embrittling treatment, i.e., in excess of the 0.06 monolayers segregated during 948 K (675°C) temper. (after SEAH [1977]).

References: p. 1284.

slowly through this range will be severely embrittled. The embrittlement is often called *reversible temper-brittleness* since samples embrittled trough slow cooling can be de-embrittled by reheating above 873 K followed by a quench, and then again embrittled by a slow cooling treatment.

The reconstruction of fig. 40 has been made by a calculation of the segregation of phosphorus in the steel using Guttmann's modified theory as described at the end of § 4.4. SEAH [1977] used this approach with $\Delta G^0_{Ni} = -11.5$, $\Delta G^0_P = -44.0$ and $\alpha'_{NiP} = -26.0$ kJ/mol to predict a low phosphorus equilibrium segregation above 800 K but rapidly rising to near the saturation of one monolayer at 748 K. Combined with this prediction are the kinetics of segregation discussed in § 4.5 with $D_P = 1.58 \times 10^{-4} \times \exp(-219000/RT)$ m^2/s. These permit the generation of the segregation contours of fig. 41 which correlate very closely with the DBTT curves. This correlation shows the dramatic DBTT shift of 275 *K* per monolayer of segregated phosphorus.

An alternative to the above interpretation of temper-brittleness via the alloy–impurity coupling model of Guttmann, has been proposed. Studies of phosphorus segregation in pure iron, and alloy with up to 0.1% carbon and 2.2% chromium by ERHART and GRABKE [1981] show that the phosphorus segregation in pure iron is as high as customarily observed in steels but that carbon, if present, will also segregate, reducing the phosphorus segregation through site competition, and improving the cohesion as described in § 5.3. The action of added chromium is shown to precipitate the carbon, so that it no longer segregates and hence the phosphorus segregation returns to the initial high level. Thus, in steels, the presence of chromium appears to enhance the phosphorus segregation, but in reality it is the carbon which the chromium influences and this, in turn, affects the phosphorus segregation. This view is supported by BRIANT [1981] who also finds that the alloying elements have no effect on phosphorus segregation in 0.3% carbon steels. If this model is correct, the lower part of fig. 41, below 723 K, should remain largely unaltered but the top part above 773 K, should arise as a result of carbon competitive segregation deriving from, say, dissolution of carbides at temperatures above 773 K. This is clearly not the situation in figs. 40 and 41 since the Cr can only precipitate a small fraction of the C in the steel. We conclude that, for the SAE 3140 steel, we must be concerned with Guttmann's alloy–impurity coupling. However, it is possible that in higher-Cr steels such as $2\frac{1}{4}$Cr1Mo, with lower carbon contents, the Erhart and Grabke mechanism may operate. In other steels both mechanisms may be important.

Finally, such steels as the Mo-containing $2\frac{1}{4}$Cr1Mo and $3Cr\frac{1}{2}$Mo alloys incur additional effects in that Mo and P react readily to form $(FeMo)_3P$ precipitates. Thus, in untempered steels quenched from above 1123 K, there is little P segregation and most of the Mo is in solution. As a function of time in the tempering range 723–923 K, the Mo that is free in solution couples with the P and according to the Guttmann theory this leads to high equilibrium segregation. However, the coupling is strong, causing the formation of Mo_3P microprecipitates so that free P level is low, thereby considerably retarding the rate of build-up to high segregation at grain boundaries. This is clear from eqs. (49) and (50) where the initial rate of segregation is proportional to the free solute content. Mo retards the diffusivity of P in pure iron by an order of magnitude at these temperatures. Thus Mo initially retards the segregation of P. With time at temperature,

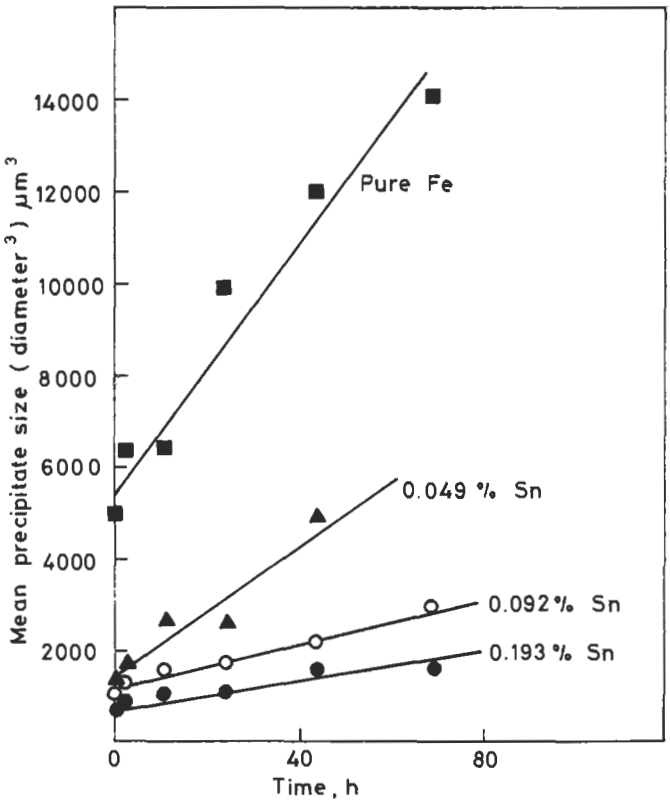


Fig. 42. The effect of increasing bulk tin content on the rate of coarsening of nitride precipitates in iron at 643 K.

the Mo nucleates Mo_2C carbides which precipitate and release the locked-up P. Thus, as time evolves the P segregation may eventually return to the level expected in Mo-free steels, as observed in practice (McMAHON *et al.* [1977]). The initial retarding effect of Mo improves as the Mo content of the steel increases but reaches an optimum in the range 0.4–0.7%, after which the embrittlement deteriorates again (DUMOULIN *et al.* [1980], YU and McMAHON [1980]). Part of the initial improvement is, as discussed above, due to the reduced P segregation. An additional effect arises through the improvement in cohesion caused by Mo segregation predicted in fig. 34 and observed by DUMOULIN *et al.* [1980]. However, as the Mo content increases, the propensity to Mo_2C precipitation increases and above 0.7% content the remedial effect disappears. Later experiments of MENYHARD and McMAHON [1989] suggest that the entire molybdenum effect could be that of cosegregation with carbon which offsets the embrittling effect of phosphorus, and the role of molybdenum as acting as traps for phosphorus seems to be only additional.

References: p. 1284.

6.4. Further examples of metallurgical phenomena influenced by microchemical processes

The case studies of metallurgical properties affected by microchemical processes presented above are being increasingly supported by a growing body of theoretical and experimental studies. Some of these are now fairly mature examples. As fig. 35 (above) indicates, there exist many other examples of metallurgical phenomena at various stages of understanding and different levels of significance, which are being progressively scrutinized in terms of the contribution of the interfacial microchemistry. Below we consider, in brief, a selection of case studies which are being currently expanded by work in several laboratories and which represent, to a certain extent, fresh directions in the understanding of traditional metallurgical problems.

6.4.1. Microchemical barrier layers

There is a category of metallurgical phenomena in which interfacial microchemistry induces its effect through an impedance layer across the interface. The phenomena refer mostly to ripening or microstructural shape modification processes (ch. 9, § 4.2.2.; ch. 15, § 2.4.2), which depend on the ease with which atoms can transfer across the interface between solid phases, but they may also operate at a free surface (as in nitridation or carburization from gas phase equilibria), involving atom transfer across the free surface from the vapour to the interior of the solid.

Consider second phase particles or precipitates in a solid matrix at service or processing temperatures. There exists an inherent tendency for such particles or precipitates to coarsen in order to reduce the total interfacial free energy associated with them. The rate-determining step in normal ripening processes is the volume diffusivity, however in the presence of segregants at this type of interface, the atom transfer step may become rate-determining. This may lead to a significant retardation in the kinetics of ripening.

There exists much indirect evidence for the segregation of impurities to these interfaces. BOYD and NICHOLSON [1971] showed that the addition of 0.1% Cd to an Al–4%Cu alloy produced a remarkable reduction in the ageing kinetics of θ' precipitates at 473 K as indicated by the rate of growth of the mean particle size with time. In addition, they were able to measure by isothermal calorimetry a considerable reduction in interfacial free energy, from 1530 mJ/m^2 to 250 mJ/m^2, produced by the cadmium additive. This is clear evidence for the segregation of the cadmium to the interface. This subject has been considered in most detail for the ripening of γ'-Fe_4N precipitates in an Fe–N alloy. Low bulk additions of antimony, oxygen, phosphorus and tin produce a remarkable inhibition in the rate of coarsening of the nitride. This is illustrated in the work of MILLER [1979] in fig. 42 in which the increasing bulk tin content is seen to reduce the mean nitride particle coarsening rate by an order of magnitude compared with the behaviour in pure iron.

In considering the theory of the segregation-induced retardation of ripening, the starting point is the *Lifshitz–Wagner equation* (LIFSHITZ and SLYOZOV [1961]; WAGNER [1961]) for volume-diffusion controlled growth of a dispersion of particles of initial mean radius r_0:

$$r^3 - r_0^3 = \frac{B_1 \gamma_i V^2 C_0 D_v}{RT}(t - t_0), \tag{75}$$

where r is the mean particle radius after a time $(t - t_0)$, γ_i the interfacial free energy, V the molar volume of precipitates, C_0 the molar concentration of solute in equilibrium with a particle of infinite size and B_1 is a numerical parameter which can vary according to the shape of the precipitates. Segregation to such interfaces will reduce γ_i typically by about 50%. However, much larger inhibitions in coarsening rates have been observed, and a reduction in interfacial energy cannot in itself account for the observed behaviour. It is believed that the most likely mechanism is that the segregate inhibits the rate of transfer of solute atoms across the interface. In this situation the relevant ripening equation becomes:

$$r^2 - r_0^2 = \frac{B_2 \gamma_i V^2 C_0 Q}{RT}(t - t_0), \tag{76}$$

where Q is a reaction rate parameter associated with the ease of transferring a species across the interface. Thus, according to this theory, the mean particle size varies as $t^{1/3}$ for volume-diffusion control and with $t^{1/2}$ for interface-reaction control. In studies of this nature it has not been easy to make a clear distinction in the mechanisms on the basis of this time-dependent growth behaviour because these functions both lie within the experimental scatter. Thus, although these studies do not give a clear confirmation on this mechanism, neither do they contradict it. MILLER [1979] has re-examined the theory and has shown that it predicts a time-invariant size distribution profile depending on the above growth mechanisms and that the data shown in fig. 42 can be fitted more closely to the particle-size profile predicted for interface control.

The surface-active elements which inhibit the ripening of precipitates are now shown to be able to form barrier layers at iron surface, thereby strongly retarding nitridation and carburization reactions. The inhibition can be quite significant as demonstrated by GRABKE *et al.* [1977] with sulphur impurity adsorption up to a level of a monolayer. In this type of atom transfer process, the alloying species is initially present in a molecular form, as CO_2, CH_4 or NH_3, in the vapour phase, and for its entry into the metal the basic mechanism is the stepwise catalytic dissociation of the surface-adsorbed molecule, allowing the C or N ions to penetrate into the interior. The strongly bound impurities at the surface impair the dissociation step, as in a poisoning reaction. In the reverse process, the surface denudation of nitrogen or carbon in iron alloys has also been shown to depend strongly on the existence of chemical barrier layers which inhibit the formation of the molecular species at the free surface. The precise mechanisms appear to be strongly related to the catalytic poisoning of surfaces by impurities, which operate in these instances in a manner that effectively retards the rate of ingress or expulsion of the elements.

6.4.2. Creep-embrittlement

It was noted earlier that as a consequence of surface segregation, the rate of grain-boundary creep cavity generation is increased, leading to a deterioration in creep rupture life. This pertains to typical weldable creep-resistant steels of the Cr–Mo–V series where

the impurities present in the bulk exercise their role in the heat-affected zone of the weld, following a stress-relief heat treatment at temperatures of the order of 973 K. Thus in a stress-relief operation in the average commercial UK steel of the type $\frac{1}{2}Cr\frac{1}{2}Mo\frac{1}{4}V$ (impurities in ppm: P, 170; Sn, 120; Sb, 30; As, 160; Cu, 1200; Ni, 1200), the most active embrittling impurities are Sn, Cu, As, Sb and P in that order with relative fragility potencies given by the ratios 3:2.5:2:1:1, respectively (LEA *et al.* [1980]).

We now consider these alloys in service conditions, where stress and temperatures are lower than in the above, and under longer-term creep durations. Analysis of existing data by SEAH [1980c] reveals a different effect. Here it appears that the rate-determining mechanism is the transport of matter or vacancies along the grain boundaries in order to feed the growing cavities. For a given strain associated with cavity growth, a high density of small cavities is much more harmful than a lower density of large cavities. The formed condition may be promoted either by increasing the cavity nucleation rate (which we have seen is induced by surface segregation to the cavity walls) but it may also be promoted by retarding the growth rate. These effects are noted in the analysis of creep cavitation by SKELTON [1975], and are described by the equation

$$\varepsilon_r = \frac{0.65}{d}\left[\frac{2\pi D_b \delta \sigma \Omega}{NkT}\right]^{1/5}, \tag{77}$$

where ε_r is the strain to rupture, d is the average grain diameter, D_b is the grain-boundary diffusivity and N the continuous rate of creation of cavities; the other symbols have been defined earlier.

A regression analysis on the effects of various impurities on the rupture ductilities during long-term creep at 823 K on $\frac{1}{2}Cr\frac{1}{2}Mo\frac{1}{4}V$ commercial steels has shown a decrease in the rupture ductility with increasing impurity content as indicated in fig. 43 (SEAH [1980c]). The dependence of rupture ductility upon impurities is given by:

$$-\Delta(\text{rupture ductility}) \propto 1.5P + 0.36Cu + 1.6Sn + 1.2S, \tag{78}$$

where the element symbols refer to solute concentrations in the bulk. For the above typical commercial-purity steel, the relative damaging potencies of the impurities are different from that of the stress-relief case of embrittlement mentioned above: the order is now Cu, P, Sn and S with relative fragilities of 2.5:1.5:1:1, respectively.

Thus, in both forms of creep embrittlement pertaining to the generation and growth of grain-boundary cavities, such as those illustrated by stress-relief cracking or long-term creep-embrittlement, the impurity sensitivity arises from different modes of interfacial segregation. This difference appears to be associated with the very different strain rates are experienced.

6.4.3. Intergranular stress-corrosion cracking

The notions of interfacial microchemistry and the effects on interfacial cohesion are being applied currently to explain various forms of intergranular cracking in the presence of both stress and a corrosive environment. Thus, pure iron is resistant to stress-corrosion cracking in nitrate environments, however in commercial carbon steels, failure is

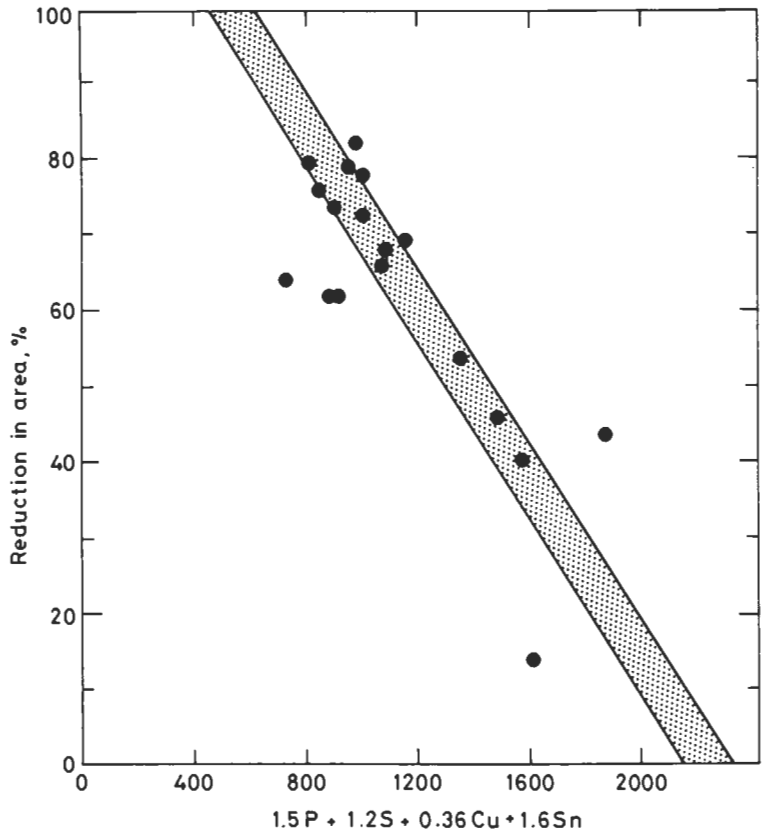


Fig. 43. Regression analysis of creep ductility in $\frac{1}{2}Cr\frac{1}{2}Mo\frac{1}{2}V$ steels as a function of a combined impurity parameter involving P, S, Cu and Sn (after SEAH [1980c]).

invariably intergranular. Many attempts have been made to trace the role of alloying elements, for example carbon, aluminium, and chromium, but there is seemingly little systematic behaviour. The importance of the impurity contents of these commercial steels has been recognized. LEA and HONDROS [1982] prepared samples from high-purity mild steel stock, each sample containing a single impurity at a dilute level comparable with that found in commercial material. Here the free residual sulphur, which is very active, was first rendered harmless by the addition of manganese. Each impurity element was added separately at a bulk level which was predicted to give rise to a grain-boundary segregation content of about 20% of a monatomic layer following an isothermal heat treatment. The samples were tested to failure in a controlled environmental cell containing ammonium nitrate, at a constant strain rate. In each case, a control sample was tested under exactly the same conditions except for a neutral solution. Here it was seen that every specimen exhibited stress corrosion in the nitrate environment, failing in a shorter time than that in the control environment. The copious data shown diagram-

matically in fig. 44 refer to the time for intergranular failure for alloys stress-corroded at a constant strain rate, noting that the data on the right-hand side of the diagram refer to the tests in the controlled solution, here paraffin. In the nitrate solution the fracture was almost always entirely intergranular, while in the control solution ductile fracture predominated. The figure shows that in the aggressive environment, with all other conditions identical, the reduction in life depends on the impurity present. Taking the time to failure under constant strain-rate testing, normalized to the value in a benign environment as the measure of the reduction in the resistance of the steel to cracking, the data can be presented in the form

$$-\Delta R = 20\mathrm{P} + 1.9\mathrm{Cu} + 1.0\mathrm{Sn} + 0.9\mathrm{Sb} + 0.9(\text{others}). \tag{79}$$

Here the element symbols represent the bulk content in wt%. The changes in R thus provide a ranking order of the potent segregating impurities which affect the stress-corrosion propensity. We note that in commercial steels S, Ca and Al are mainly present as precipitates. To optimize the performance of carbon steel in nitrate environments, the clearest metallurgical guideline from this is a reduction in bulk impurity levels, particularly that of residual phosphorus.

Although the detailed micromechanism for this effect awaits elaboration, the principal observations can be described fairly adequately by a dissolution model in which the grain-boundary material is preferentially removed, the driving force for this being the divergence of the segregant's electrochemical potential from that of iron. In an oxidizing

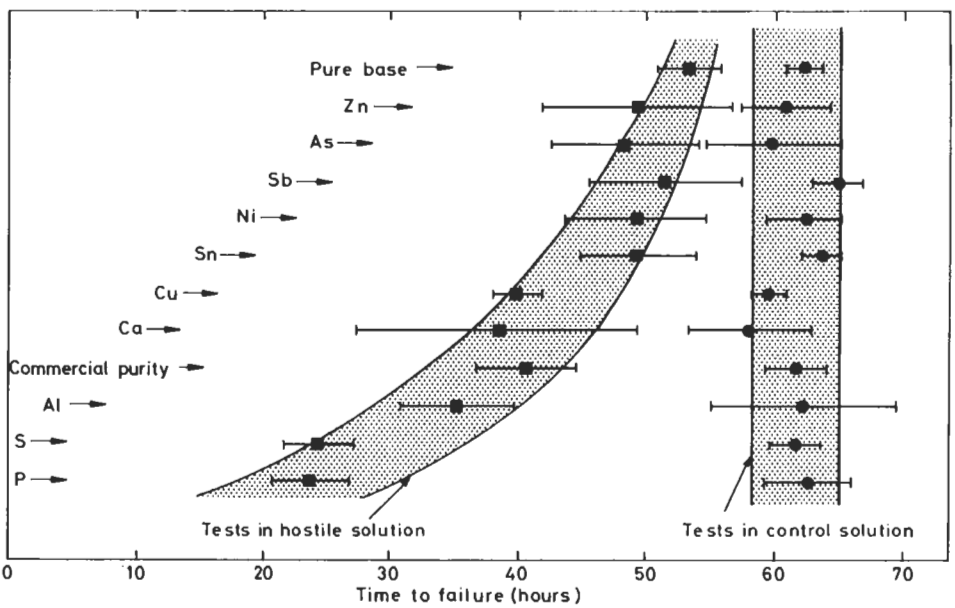


Fig. 44. The times to failure of 11 specimens of carbon steel, each containing a different impurity element, stress-corrosion tested in ammonium nitrate solution at a constant strain rate. The control tests in a neutral environment are shown to the right (after LEA and HONDROS [1982]).

environment such as aqueous ammonium nitrate, the driving force for atomic dissolution is related to the potential difference between the matrix and the segregants, which form a galvanic cell. Thus for segregants which have a more positive potential than iron, these atoms themselves are dissolved more rapidly and a very sharp grain-boundary crack is exposed for further dissolution. On the other hand, for segregants having a more negative potential than iron, atoms adjacent to the boundary are dissolved, thus enhancing the intergranular corrosion but with a different crack tip profile. Grain-boundary segregation during stress-corrosion cracking furthermore promotes dissolution by reducing the rate of repassivation after slip has ruptured protective oxide films and crack tips (LYNCH [1989]).

6.4.4. Intergranular hydrogen-embrittlement

An aspect of the ubiquitous problem of the hydrogen-embrittlement of steels is of interest here. This is the intergranular mode of failure, which has been considered recently in relation to the microchemical nature of the grain boundaries. Thus LATANISION and OPPERHAUSER [1974] proposed that hydrogen penetration along the grain-boundary regions should be promoted by hydrogen-recombination poisons which are present as grain-boundary segregants. These elements, such as P, S, As and Sb, which have a strong tendency to segregate to grain boundaries in steel also have a stronger intrinsic binding to hydrogen than does iron (LYNCH [1989]). A probable effect of this is that following the dissociation of a hydrogen molecule which arrives from the vapour or from solution, the stronger bonding available at the solute atom sites causes the hydrogen to reside at these sites for much longer times than at normal iron sites, and therefore there is a greater chance for the hydrogen to be absorbed into the lattice. Hence the role of these solute species is effectively to increase the amount of hydrogen occluded along the grain-boundary zones. If the mechanism of hydrogen-embrittlement requires the concentration of hydrogen in the metal just ahead of the crack tip, the above grain-boundary segregants should therefore enhance this embrittlement.

6.4.5. Inhibition of surface oxidation on alloys

In certain alloys, especially those containing aluminium, surface oxidation can be inhibited by the introduction of dopants which both reduce the rate of grown of the oxide and also inhibit "breakaway" oxidation by improving the scale adhesion at the metal/metal-oxide interface. This complex subject has been reviewed in detail, for example by WHITTLE and STRINGER [1980] who assessed various mechanisms for the improvement in performance. Several microchemical phenomena have been proposed to account for this. One proposal, which requires experimental verification but which is consistent with the analysis given in § 5.2, is that the dopant segregates to the oxide grain boundaries, thereby reducing the boundary diffusivity. In their study of oxidation behaviour of γ–γ'-Cr_3C_2 eutectic composite containing additions of yttrium at a low level, BULLOCK *et al.* [1974] observed a remarkable improvement (i.e., reduction) in oxidation weight gain at 1100°C and in addition showed by AES that the yttrium level increased with distance in the oxide bulk and reached a peak level at a depth which suggested that the interface between the oxide and the alloy was enriched in yttrium. In addition to the probable improvements in growth kinetics by segregation to the boundaries, it was

References: p. 1284.

considered that the formation of a microchemical barrier layer at the metal/metal-oxide interface improved the adhesion of the alumina in the mixed metal-oxide scale to the metal substrate and hence increased the resistance to oxide scale sparling.

A third possibility for inhibited high-temperature corrosion in through the formation of a microstructural barrier layer at the metal-oxide/vapour interface. The presence of boron in the vapour phase is known to improve greatly the oxidation resistance of chromium-containing steels by increasing the time to "breakaway" oxidation. LEA [1979] showed by AES that the boron was present essentially only at the outer oxide surface layer and it was also accompanied by a high chromium enrichment. The mechanism for this type of inhibition thus has similarities with the processes considered above (§ 6.4.1). In particular, it was considered that the introduction of O–B–O covalent bonds in the outer oxide surface structure would impede the transfer of O ions onto the lattice, thus forming a film which in turn allows slowly diffusing protective chromium from the alloy to reach the free surface (ch. 14, § 3).

7. *Interfacial microchemistry and materials design theory*

All the important themes of physical and chemical metallurgy contribute collectively to the theory of materials design — including the thermochemistry of phase equilibria, crystal defect theory, fracture theory, oxidation and corrosion resistance and so on. This chapter has placed emphasis on the importance of the relatively new subject of interfacial microchemistry for a wide diversity of metallurgical properties, and the insight and information now made available should be viewed in terms of its contribution to materials design. Here we mention two basic approaches to the materials design theory, the alloy design and the grain-boundary design. In the former one we distinguish between deliberate additions of specific elements which procure beneficial effects on certain properties, and the planned control of known elements which are detrimental. In the latter approach, effort is focused on tailoring the structure of a polycrystalline material with improved properties using a specific arrangement of appropriate grain boundaries. The advanced knowledge in these fields provides many opportunities to manipulate material properties, often in the context of "fine tuning" but equally often to gain significant improvements in service life. In the past two decades, several lines of advance have been made in this subject matter which are valuable to alloy design: (i) the identification of potent impurity species (both harmful and beneficial); (ii) the basic mechanisms (sometimes only partly understood) by which they operate; (iii) the enrichment levels as well as bulk contents at which the effects are significant. As for grain-boundary design approach, progress has been made in: (i) identification of the types of grain boundaries characterized by different properties; (ii) processes for controlling the grain-boundary character distribution in polycrystalline materials.

This chapter provides several guidelines. Thus, on the basis of a retardation of grain-boundary transport, or on improved interfacial cohesion, certain types of elements can be beneficial — as in the improvement of oxidation characteristics or in the rate of Coble creep. Certain elements can be predicted to offer an improvement in sintering behavior

through a segregant stabilization of fine grain structure. In an example of benefit to mechanical properties, WHITE *et al.* [1981] demonstrated that the room-temperature strength and ductility of a Pt/Rh/W alloy can be improved by the addition of 75 wt ppm of B which was shown to enrich, in a typical manner, the grain boundaries.

Where the impurities produce deleterious effects, clearly the *alloy design* aspects involve the elimination of the impurities or, more practicably, their control to within tolerance levels. The identification of potent surface-active species and the assigning of bulk tolerance levels have been the specific objectives of many studies in this field. The information is couched often in formulae of the type presented above for intergranular stress-corrosion cracking (eq. (79)) or creep rupture ductility (eq. (78)). This permits the design of alloys of closely controlled compositions for improved performance in the defined conditions of use (HONDROS [1986]).

In a previous discussion (HONDROS and SEAH [1977b]), we presented a number of procedures which might offer a control over the effects of harmful impurities — including the judicious use of the principles of competitive segregation, the selection of remedial segregants, the thinning out of the embrittling species by spreading them out over a large interfacial area, i.e., a very fine grain size, reducing the kinetics of the embrittling species, and the used chemical fixation or precipitation reactions. In the meantime, at least one of the above approaches has been explored in depth, in connection with affording a simple and economical control over temper-embrittlement. SEAH *et al.* [1979] demonstrated that temper-embrittlement induced by P and Sn in $2\frac{1}{4}$Cr1Mo and $3\frac{1}{2}$Ni–Cr–Mo–V steels can be reduced by additions of La or Ce at a predicted level. The selection of La, Ce and mischmetal (approx. 50% Ce, 25% La, 25% other rare-earth metals) are based on a comprehensive thermochemical assessment of the possibility of forming stable microprecipitates in order to reduce the bulk levels of active P and Sn. The optimum addition of La required to remove the embrittling effects of P and Sn is given by the formula

$$\mathrm{La} = 8.7\mathrm{S} + 2.3\mathrm{Sn} + 4.5\mathrm{P} \tag{80}$$

where the element symbols represent the wt% of the element in the steel. Figure 45, taken from this work, shows the high DBTT for a pure $2\frac{1}{4}$Cr1Mo steel with 0.017% P added. The further addition of various levels of La lower the transition temperature significantly, to the limit of a high purity material.

In the future, the increasing knowledge of structure/property relationships for grain boundaries (including anisotropy of grain-boundary segregation and thus, for instance, anisotropy of brittle intergranular fracture) may be practically used to control materials properties. The differences in grain-boundary character and geometric configuration enhance the heterogeneity of all types of deformation and fracture in polycrystals. This effect is schematically shown in fig. 46. The presence of low-energy grain boundaries (Σ-boundaries in the figure) in a polycrystal will not contribute much to the deformation and fracture while the presence of general grain boundaries (R-boundaries) will exert a significant contribution. Therefore, a heterogeneity of the fracture process occurs which can either lead to an easy fracture (figs. A, B, D or F) or to a difficult fracture (figs. C and E) under the same stress. On the basis of these considerations, WATANABE [1984]

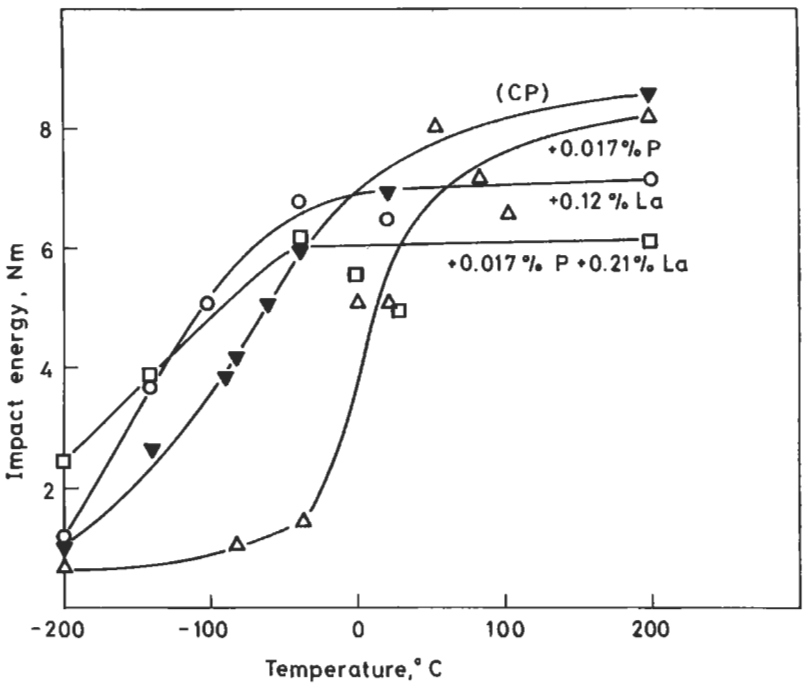


Fig. 45. Improvement in the ductile–brittle transition temperature in a steel susceptible to temper embrittlement, resulting from low level additions of lanthanum which combat the effect of phosphorus (after SEAH *et al.* [1979]).

pointed out the importance of the effect of the type and frequency of grain boundaries to the behaviour of a polycrystal and proposed the concept of *grain-boundary design*. The basic idea of grain-boundary design is to produce polycrystalline materials with grain boundaries of such character and distribution as will suppress a "detrimental effect" of the grain boundary and conversely, increase its "beneficial effect" as much as possible. In this way, the properties of a polycrystal could be improved to a significant extent.

To fabricate a polycrystalline material by means of grain-boundary design, two fundamental steps have to be performed. The first step consists in prediction of the performance and thus, in the proposal of character and distribution of grain boundaries for the given purpose. The other one represents the processing method to achieve the predicted grain-boundary character distribution and configuration in order to obtain the proposed structural material. Practically, of course, the latter step seems to be the limiting one because no method has commonly been developed to produce such structures till now. Thermomechanical processing based on working and annealing has widely been applied to control microstructure of polycrystals such as grain size, second phase dispersion, and texture. Unfortunately, these processes are not developed to such an extent at present that they completely control the character and distribution of grain

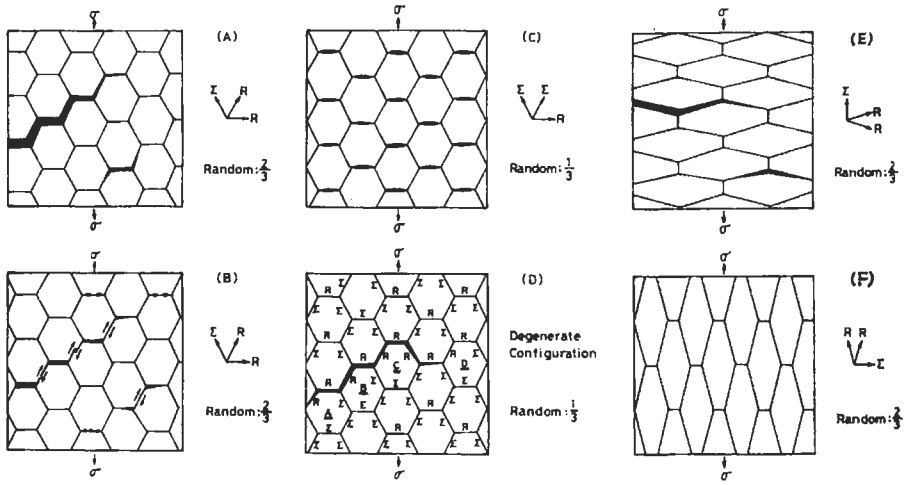


Fig. 46. Intergranular fracture processes in polycrystals with different grain-boundary character distributions and boundary configurations (after WATANABE [1993]).

boundaries. From this point of view, magnetic annealing of ferromagnetic materials seems to be promising to introduce specific types of grain boundaries with a characteristic frequency in a given material. Of course, novel procedures have to be developed in future (WATANABE [1993]).

There is an aspect of the general subject matter of this chapter which has wider ramifications in the context of metals recycling, and which implies that it will be a subject of growing importance in the future. It is clear that because of geo-political, economic and social factors there will be in the future a greater emphasis on the recycling of used materials in industrialized countries. This will reduce their vulnerability to external sources of raw materials and also, there are inherent energy savings in reclaiming metals from used articles rather than obtaining them from raw supplies of ore. Because of the complex mechanical systems used by industry and society and the wide combination of metals involved, inevitably metals derived from recovered scrap will contain certain unwanted residuals which bypass the refinement process (HONDROS [1980b]). This chapter has underlined that such small quantities of noxious elements can provoke remarkable effects on physical and mechanical properties, and more specifically the accumulation of residual elements, such as copper and tin in steel from scrap, will have effects on the performance. Thus, an increased effort devoted to the recycling of metals will, in general, exacerbate the effects of unwanted tramp elements on the properties of alloy aimed for particular end uses. On the basis of a better metallurgical understanding, there will be effort devoted to a control tramp element effects, involving remedial measures, of the type discussed above.

However, we envisage the need for a broader perspective, leading to a new metallurgical or engineering subject matter encompassing the theory of design involving recycling. In this global approach, we consider the final end use of the material as the primary

References: p. 1284.

parameter, and recyclability should be built in to the design of the final material. Thus we may plan for the evolution of secondary alloy products which may not be adequate for the original property requirements, but which may be suitable for other applications. An interesting example of this is a carbon steel with high levels of tin and cooper arising from the use of scrap containing much tin plate — this has been demonstrated to have an improved corrosion resistance associated with the presence of these elements in sub-scale layers. Such an alloy may not have adequate mechanical properties for the original purposes, however it can be used in situations where the improved corrosion resistance is advantageous. Thus, in this theory of design with recycling we can envisage recovery epicycles in which certain grades of alloy might be blended together to produce new alloys which are more suitable for specific applications. The science of interfacial microchemistry will make basic contributions in the future in connection with material design in the context of recycling, improved reliability and advances in performance.

References

ASHBY, M.F., 1972, Acta Metall. **20**, 887.

AUST, K.T., and J.H. WESTBROOK, 1965, Lattice Defects in Quenched Metals (Academic Press, London) p. 771.

BATTE, A.D., J.M. BEAR, S.R. HOLDSWORTH, J. MEYERS and P.R. REYNOLDS, 1980, Phil. Trans. Roy. Soc. **A295**, 253.

BENNINGHOVEN, A., J. GIBER, J. LASLO, M. RIEDEL and H. WERNER, eds., 1982, Springer Series in Chemical Physics, **19**, Secondary Ion Mass Spectroscopy, SIMS III (Springer, Berlin).

BERNARDINI, J., P. GAS, E.D. HONDROS and M.P. SEAH, 1982, Proc. Roy. Soc. **A379**, 159.

BISCONDI, M., 1982, J. Phys. France **43**, C6-293.

BORISOV, V.T., V.M. GOLIKOV and G.V. SHCHERBEDINSKIY, 1964, Phys. Met. Metallogr. **17**, 80.

BOUCHET, D., and L. PRIESTER, 1986, Scripta Metall. **20**, 961.

BOUCHET, D., and L. PRIESTER, 1987, Scripta Metall. **21**, 475.

BOYD, J.D., and R.B. NICHOLSON, 1971, Acta Metall. **19**, 1379.

BOZZOLO, G., J. FERANTE and J.R. SMITH, 1992, Phys. Rev. B **45**, 493.

BOZZOLO, G., B. GOOD and J. FERANTE, 1993, Surf. Sci. **289**, 169.

BRIANT, C.L., 1981, Scripta Metall. **15**, 1013.

BRIANT, C.L., 1988a, Mater. Sci. Technol. **4**, 956.

BRIANT, C.L., 1988b, in: Auger Electron Spectroscopy, eds. C.L. Briant and R.P. Messmer. Treatise on Materials Science and Technology, Vol. 30 (Academic Press Boston) p. 111.

BRIANT, C.L., 1990, Metall. Trans. **21A**, 2339.

BRIANT, C.L., and R.P. MESSMER, 1982, Acta Metall. **30**, 1811.

BRIGGS, D., 1982, Surf. Interface Anal. **4**, 151.

BRIGGS, D., and M.P. SEAH, eds., 1990, Practical Surface Analysis. 2nd ed., Vol. 1, Auger and X-ray Photoelectron Spectroscopy (Wiley and Sons, Chichester).

BRIGGS, D., and M.P. SEAH, eds., 1992, Practical Surface Analysis. 2nd ed., Vol. 2, Ion and Neutral Spectroscopy (Wiley and Sons, Chichester).

BRUEMMER, S.M., 1989, Mater. Sci. Forum **46**, 309.

BRULEY, J., I. TANAKA, H.-J. KLEBE and M. RÜHLE, 1994, Anal. Chim. Acta **297**, 97.

BRUNAUER, S., L.S. DEMING, W.E. DEMING and E.J. TELLER, 1940, J. Am. Chem. Soc. **62**, 1723.

BULLOCK, E., C. LEA and M. MCLEAN, 1979, Met. Sci. **13**, 373.

CARR, F.L., M. GOLDMAN, L.D. JAFFE and D.C. BUFFUM, 1953, Trans. AIME **197**, 998.

COBLE, R.L., 1963, J. Appl. Phys. **34**, 1679.

DEFAY, R. L., I. PRIGOGINE, A. BELLOMANS and D. H. EVERETT, 1966, Surface Tension and Adsorption. (Longmans, London).
DELAMARE, F., and G. E. RHEAD, 1971, Surf. Sci. **28**, 267.
DOIG, P., and P. E. J. FLEWITT, 1981, Acta Metall. **29**, 1831.
DOIG, P., D. LONSDALE and P. E. J. FLEWITT, 1981, Metall. Trans. **12A**, 1277.
DUMOULIN, P. H., M. GUTTMANN, M. FOUCAULT, M. PALMIER, M. WAYMAN and M. BISCONDI, 1980, Met. Sci. **14**, 1.
DU PLESSIS, J., 1990, Sol. State Phenom. **11**, 1.
EGERT, B., and G. PANZNER, 1982, Surf. Sci. **118**, 345.
ERHART, H., and H. J. GRABKE, 1981, Met. Sci. **15**, 401.
ERLEWEIN, J., and S. HOFMANN, 1977, Surf. Sci. **68**, 71.
EWING, R. H., 1971, Acta Metall. **19**, 1359.
FISCHMEISTER, H., 1985, J. Phys. France **46**, C4-3.
FRANZONI, U., H. GORETZKI and S. STURLESE, 1981, Scripta Metall. **15**, 743.
FUKUSHIMA, H., and H. K. BIRNBAUM, 1984, Acta Metall. **32**, 851.
GAS, P., M. GUTTMANN and J. BERNARDINI, 1982, Acta Metall. **30**, 1309.
GEORGE, E. P., and D. P. POPE, 1989, in: Interfacial Structure, Properties and Design, eds. M.H. Yoo, W.A.T. Clark and C.L. Briant. MRS Symp. Proc. Vol. 122 (MRS, Pittsburgh) p. 391.
GERMAN, R. M., and K. A. D'ANGELO, 1984, Int. Met. Rev. **29**, 249.
GIBBS, G. B., and J. E. HARRIS, 1969, In: Interfaces, ed. R.C. Gifkins (Butterworths, London) p. 53.
GIBBS, J. W., reprinted 1948, Collected Works, Vol. 1 (Yale University Press) p. 219.
GJOSTEIN, N. A., 1966, Proc. 13th Sagamore Army Materials Research Conference, Raquette Lake, NY, eds. J.J. Burke, N.L. Reed and V.Weiss (Syracuse University Press).
GRABKE, H. J., W. PAULITSCHKE, G. TAUBER and H. VIEFHAUS, 1977, Surf. Sci. **63**, 377.
GUGGENHEIM, E. A., 1950, Thermodynamics, 2nd Ed. (North-Holland, Amsterdam).
GUTTMANN, M., 1975, Surf. Sci. **53**, 213.
GUTTMANN, M., 1976, Met. Sci. **10**, 337.
GUTTMANN, M., 1980, Phil. Trans. Roy. Soc. **A295**, 169.
GUTTMANN, M., and D. MCLEAN, 1979, Proc. ASM Materials Science Seminar, Interfacial Segregation, eds. W.C. Johnson and J.M. Blakely (ASM, Metals Park, OH) p. 261.
HARRIES, D. R., and D. MARWICK, 1980, Phil. Trans. Roy. Soc. **A295**, 197.
HASHIMOTO, M., Y. ISHIDA, R. YAMAMOTO and M. DOYAMA, 1984, Acta Metall. **32**, 1.
HAYES, P. and P.GRIEVESON, 1975, Met. Sci. **9**, 332.
HERRING, C., 1950, J. Appl. Phys. **21**, 437.
HIRTH, J. P. and J. R. RICE 1980, Metallurg. Trans. **11A**, 1501.
HOFMANN, S., and J. ERLEWEIN, 1976, Scripta Metall. **10**, 857.
HOFMANN, S., and J. ERLEWEIN, 1978, Surf. Sci. **77**, 591.
HOFMANN, S., and P. LEJČEK, 1991, Scripta Metall. Mater. **25**, 2259.
HOFMANN, S., P. LEJČEK and J. ADÁMEK, 1992, Surf. Interface Anal. **19**, 601.
HONDROS, E. D., 1965, Proc. Roy. Soc. **A286**, 479.
HONDROS, E. D., 1969, in: Interfaces, ed. R.C. Gifkins (Butterworths, London) p. 77.
HONDROS, E. D., 1970, Surface Energy Measurements, in: Physicochemical Measurements in Metals Research, Techniques in Metals Research Vol. IV, ed. R.A. Rapp (Wiley, New York) Part 2, p. 293.
HONDROS, E. D., 1976, in: Grain-boundary Structure and Properties, eds. G.A. Chadwick and D.A. Smith (Academic, London) p. 265.
HONDROS, E. D., 1978, Precipitation Processes in Solids, Proc. TMS–AIME Meeting 1976, Eds. K.C. Russell and H.I. Aaronson (Metallurg. Soc. of AIME, Warrendale, PA) p. 1.
HONDROS, E. D., 1980a, Scripta Metall. **14**, 345.
HONDROS, E. D., 1980b, Phil. Trans. Roy. Soc. **A295**, 9.
HONDROS, E. D., 1986, in: Grain-boundary Structure and Related Phenomena (Proc. 4th JIM Ins. Symp.). Trans. JIM Suppl. **27**, 393.
HONDROS, E. D., and P. J. HENDERSON, 1983, Metallurg. Trans. **14A**, 521.
HONDROS, E. D., and D. MCLEAN, 1974, Phil. Mag. **29**, 771.

HONDROS, E.D., and D. MCLEAN, 1976, in: Grain-boundary Structure and Properties, eds. G.A. Chadwick and D.A. Smith (Academic, London) p. 353.
HONDROS, E.D., and M.P. SEAH, 1977a, Metallurg. Trans. **A8**, 1363.
HONDROS, E.D., and M.P. SEAH, 1977b, Int. Met. Rev. **22**, 262.
HONDROS, E.D., and L.E.H. STUART, 1968, Phil. Mag. **17**, 711.
HULTGREN, R., P.A. DESAI, D.T. HAWKINS, M. GLEISER and K.K. KELLY, 1973, Selected Values of the Thermodynamic Properties of Binary Alloys (ASM, Metals Park, OH).
JOHNSON, W.C., N.G. CHAVKA, R. KU, J.L. BOMBACK and P.P. WYNBLATT, 1978, J. Vac. Sci. Technol. **15**, 467.
JOKL, M.L., V. VITEK and C.J. MCMAHON, Jr., 1980, Acta Metall. **28**, 1479.
JONES, R.H., S.M. BRUEMMER, M.T. THOMAS and D.R. BAER, 1981, Metallurg. Trans. **12A**, 1621.
JUHAS, M., L. PRIESTER and M. BISCONDI, 1994, Mater. Sci. Eng. A **185**, 71.
KALDERON, D., 1972, Proc. Inst. Mech. Eng. **186**, 341.
KARLSSON, L., 1988, Acta Metall. **36**, 25.
KARLSSON, L., and H. NORDÉN, 1988, Acta Metall. **36**, 13.
KARLSSON, L., H. NORDÉN, and H. ODELIUS, 1988, Acta Metall. **36**, 1.
KAUR, I., and W. GUST, 1988, Fundamentals of Grain and Interphase Boundary Diffusion (Ziegler, Stuttgart).
KITTEL, C., 1986, Introduction to Solid State Physics, 6th edition (Wiley, New York).
KOZAKEVITCH, P., and G. URBAIN 1961, Mem. Sci. Rev. Metallurg. **6**, 401.
KRAAI, D.A., and S. FLOREEN, 1964, Trans. AIME **230**, 833.
KUMAR, A., and B.L. EYRE, 1980, Proc. Roy. Soc. **A370**, 431.
KUPPER, J., H. ERHART and H.J. GRABKE, 1981, Corros. Sci. **21**, 227.
LATANISION, R.M. and H. OPPERHAUSER, 1974, Metallurg. Trans. **5**, 483.
LEA, C., 1979, Met. Sci. **13**, 301.
LEA, C., and E.D. HONDROS, 1982, Proc. Roy Soc. **A377**, 477.
LEA, C., and M.P. SEAH, 1977, Phil. Mag. **35**, 213.
LEA, C., M.P. SEAH and E.D. HONDROS, 1980, Mater. Sci. Eng. **42**, 233.
LEJČEK, P., 1994, Anal. Chim. Acta **297**, 165.
LEJČEK, P., J. ADÁMEK and S. HOFMANN, 1992, Surf. Sci. **264**, 449.
LEJČEK, P., and S. HOFMANN, 1991, Acta Metall. Mater. **39**, 2469.
LEJČEK, P., and S. HOFMANN, 1993, Interface Sci. **1**, 161.
LEJČEK, P., and S. HOFMANN, 1995, Crit. Rev. Sol. State Mater. Sci. **20**, 1.
LEJČEK, P., A.V. KRAJNIKOV, YU.N. IVASHCHENKO, M. MILITZER and J. ADÁMEK, 1993, Surf. Sci. **280**, 325.
LIFSHITZ, I.M., and V.V. SLYOZOV, 1961, Phys. Chem. Solids **19**, 35.
LIU, C.T., C.L. WHITE and J.A. HORTON, 1985, Acta Metall. **33**, 213.
LUZZI, D.E., 1991, Ultramicroscopy **37**, 180.
LUZZI, D.E., M. YAN, M. ŠOB and V. VITEK, 1991, Phys. Rev. B **67**, 1894.
LYNCH, S.P., 1988, Acta Metall. **36**, 2639.
LYNCH, S.P., 1989, Mater. Sci. Forum **46**, 1.
MARWICK, A.D., 1981, Nucl. Instr. and Meth. **182/183**, 827.
MASSALSKI, T.B., J.L. MURRAY, L.H. BENNETT and H. BAKKER, 1986, Binary Alloy Phase Diagrams (ASM, Metals Park, Ohio).
MASUDA-JINDO, K., 1986, phys. stat. sol. (b) **134**, 545.
MATSUKO, T., K. NAKAJIMA, Y. TERADA and M. KIKUCHI, 1991, Mater. Sci. Eng. A **146**, 261.
MCLEAN, D., 1957, Grain boundaries in Metals (Oxford Univ. Press).
MCLEAN, D., 1981, Metals Forum **4**, 45.
MCMAHON, C.J., JR., 1991, Mater. Charact. **26**, 269.
MCMAHON, C.J., JR., 1992, in: Clean Steels Technology, ed. R. Viswanathan (ASM, Metals Park) p. 3.
MCMAHON, C.J., JR., and V. VITEK, 1979, Acta Metall. **27**, 507.
MCMAHON, C.J., JR., A.K. CIANELLI and H.C. FENG, 1977, Metall. Trans. **8A**, 1055.
MENYHARD, M., B. BLUM, and C.J. MCMAHON, JR., 1989, Acta Metall. **37**, 549.
MENYHARD, M., and C.J. MCMAHON, JR., 1989, Acta Metall. **37**, 2289.
MENYHARD, M., B. ROTHMAN, C.J. MCMAHON, JR., P. LEJČEK and V. PAIDAR, 1991, Acta Metall. **39**, 1289.

MEZEY, L.Z., J. GIBER, W. HOFER and P. VARGA, 1990, Surf. Sci. **234**, 197.
MEZEY, L.Z., and W. HOFER, 1992, Surf. Interface Anal. **19**, 618.
MILITZER, M., YU.N. IVASHCHENKO, A.V. KRAJNIKOV, P. LEJČEK, J. WIETING, and S.A. FIRSTOV, 1992, Surf. Sci. **261**, 267.
MILITZER, M., and J. WIETING, 1986, Acta Metall. **34**, 1229.
MILITZER, M., and J. WIETING, 1987, Acta Metall. **35**, 2765.
MILLER, D., 1979, Doctor of Philosophy Thesis, Univ. of Strathclyde.
MILLER, M.K., 1987, Int. Mater. Rev. **32**, 221.
MILLS, B., M. MCLEAN and E.D. HONDROS, 1973, Phil. Mag. **27**, 361.
MUKHERJEE, S., and J.L. MORÁN-LÓPEZ, 1987, Surf. Sci. **188**, L742.
MUSCHIK, T., W. GUST, S. HOFMANN and B. PREDEL, 1989, Acta Metall. **37**, 2917.
NABARRO, F.R.N., 1948, Report of Conference on Strength of Solids, (Phys. Soc., London) p. 75.
NAJAFABADI, R., H.Y. WANG, D.J. SROLOVITZ and R. LESAR, 1991, Acta Metall. Mater. **39**, 3071.
NANNI, P., C.T.H. STODDART and E.D. HONDROS 1976, J. Mater. Chem. **1**, 297.
NICHOLAS, M.G., and C.F. OLD 1976, J. Mater. Sci. **14**, 1.
NICOLET, M.A., 1978, Thin Solid Films **52**, 415.
NIEHUS, H., W. HEILAND and E. TAGLAUER, 1993, Surface Sci. Rep. **17**, 252.
OSGERBY, S., and T.B. GIBBONS, 1992, Mat. Sci. Eng. A **157**, 63.
OTTERBEIN, U., S. HOFMANN, and M. RÜHLE, 1993, in: High-Temperature Ordered Intermetallic Alloys V, eds. I. Baker, R. Darolia, J.D. Whittenberger and M.H. Yoo (MRS Proc. Symp. Vol. 288; MRS, Pittsburgh) p. 183.
OVERBURY, S.H., P.A. BERTRAND and G.A. SOMORJAI, 1975, Chem. Rev. **75**, 547.
PAIDAR, V., 1987, Acta Metall. **35**, 2035.
PETCH, N.J., 1956, Phil. Mag. **1**, 331.
PICHARD, C., M. GUTTMANN, J. RIEU and C. GOUX, 1975, J. Physique **36**, C4-151.
RELLICK, J.R., and C.J. MCMAHON, JR., 1974, Metall. Trans. **5**, 2439.
RICE, J.R., and J.-S. WANG, 1989, Mater. Sci. Eng. A **107**, 23.
ROBERTS-AUSTEN, W.C., 1888, Phil. Trans. Soc. **A179**, 339.
ROWLANDS, G., and D.P. WOODRUFF, 1979, Phil. Mag. **40**, 459.
RÜDENAUER, F.G., 1994, Anal. Chim. Acta **297**, 197.
SEAH, M.P., 1975, Surf. Sci. **53**, 168.
SEAH, M.P., 1976, Proc. Roy. Soc. **A349**, 535.
SEAH, M.P., 1977, Acta Metall. **25**, 345.
SEAH, M.P., 1979, J. Catal. **57**, 450.
SEAH, M.P., 1980a, J. Phys. **F10**, 1043.
SEAH, M.P., 1980b, Acta Metall. **28**, 955.
SEAH, M.P., 1980c, Phil. Trans. Roy. Soc. **A295**, 265.
SEAH, M.P., 1980d, J. Vac. Sci. Technol. **17**, 16.
SEAH, M.P., 1981, Mater. Sci. Club. Bull. **64**, 2.
SEAH, M.P., 1990, in: Practical Surface Analysis, eds. D. Briggs and M.P. Seah (Wiley and Sons, Chichester). 2nd ed., Vol. 1, ch. 7.
SEAH, M.P., and E.D. HONDROS, 1973, Proc. Roy. Soc. **A335**, 191.
SEAH, M.P., and E.D. HONDROS, 1983, in: Atomistics of Fracture, ed. R.M. Latanision (Plenum, New York), p. 855.
SEAH, M.P., and C. LEA, 1975, Phil. Mag. **31**, 627.
SEAH, M.P., P.J. SPENCER and E.D. HONDROS, 1979, Met. Sci. **13**, 307.
SEIDMAN, D.N., 1991, Mater. Sci. Eng. A **137**, 57.
SEIDMAN, D.N., 1992, in: Materials Interfaces: Atomic-Level Structure and Properties, eds. D. Wolf and S. Yip (Chapman and Hall, London) p. 58.
SEKI, A., D.N. SEIDMAN, Y. OH and S.M. FOILES, 1991, Acta Metall. Mater. **39**, 3167.
SHINODA, T., and T. NAKAMURA, 1981, Acta Metall. **29**, 1631 and 1637.
SKELTON, R.P., 1975, Met. Sci. **9**, 192.
SMITH, C.S., 1948, Trans. AIME **175**, 15.

SPENCE, J.C.H., and J. TAFTØ, 1983, J. Microscopy **130**, 147.
SUNDARSHAN, T.S., and M.R. LOUTHAN, JR. 1987, Int. Mater. Rev. **32**, 121.
SURESH, S., G.F. ZAMISKI and R.O. RITCHIE, 1981, Metall. Trans. **12A**, 1435.
SUTTON, A.P., and R.W. BALLUFFI, 1987, Acta Metall. **35**, 2177.
SUZUKI, S., K. ABIKO and H. KITAMURA, 1981, Scripta Metall. **15**, 1139.
SUZUKI, S., and M. OKU, 1990, Appl. Surf. Sci. **44**, 151.
SWIATNICKI, W., S. LARTIGUE, M. BISCONDI and D. BOUCHET, 1990, J. Phys. France **51**, C1-341.
TAUBER, G., and H.J. GRABKE, 1978, Ber. Bunsen Gesell Phys. Chem. **82**, 298.
THOMAS, G.B, and T.B. GIBBONS, 1980, Proc. Seven Springs Conference, eds. G.K. Tien *et al.* (ASM, Metals Park, OH) p. 99.
TIPLER, H.R., 1980, Phil. Trans. Roy. Soc. **A295**, 213.
TIPLER, H.R., and D. MCLEAN, 1970, Met. Sci. **4**, 103.
TSAI, N.H., G.M. POUND and F.F. ABRAHAM, 1977, J. Catal. **50**, 200.
TSONG, T.T., 1990, Atom-Probe Field-Ion Microscopy (Cambridge Univ. Press, Cambridge).
TYSON, W.R., 1975, Can. Met. Quart. **14**, 307.
TYSON, W.R., 1978, Acta Metall. **26**, 1471.
VITEK, V., and G.J. WANG, 1982, J. Phys. France **43**, C6-147.
WAGNER, C., 1961, Z. Electrochem. **65**, 581.
WANG, H.Y., R. NAJAFABADI, D.J. SROLOVITZ and R. LESAR, 1992, Phil. Mag. A **65**, 625.
WATANABE, T., 1984, Res Mechan. **11**, 47.
WATANABE, T., 1993, Mater. Sci. Eng. A **166**, 11.
WATANABE, T., S. KITAMURA and S. KARASHIMA, 1980, Acta Metall. **28**, 455.
WEI, W., and H.J. GRABKE, 1986, Corrosion Sci. **26**, 223.
WESTBROOK, J.H., 1964, Met. Rev. **9**, 415.
WHITE, C.L., J.R. KEISER and D.N. BRASKI, 1981, Metallurg. Trans. **12A**, 1485.
WHITE, C.L., R.A. PADGETT, C.T. LIU and S.M. YALISOVE, 1984, Scripta Metall. **18**, 1417.
WHITTLE, D.P., and J. STRINGER, 1980, Phil. Trans. Soc. **A295**, 309.
WILLIAMS, F.L., and D. NASON, 1974, Surf. Sci. **45**, 377.
WILLIAMS, T.M., A.M. STONEHAM and D.R. HARRIES, 1976, Met. Sci. **10**, 14.
WOLF, D., 1992, in: Materials Interfaces: Atomic-Level Structure and Properties, eds. D. Wolf and S. Yip (Chapman and Hall, London) p. 1.
WOLF, U., F. ERNST, T. MUSCHIK, M.W. FINNIS and H.F. FISCHMEISTER, 1992, Phil. Mag. A **66**, 991.
WU, R., A.J. FREEMAN and G.B. OLSON, 1992, J. Mater. Res. **7**, 2403.
WYNBLATT, P., and R.C. KU, 1979, Interfacial Segregation, eds. W.C. Johnson and J.M. Blakely (ASM, Metals Park. OH) p. 115.
YAN, M., M. ŠOB, D.E. LUZZI, V. VITEK, G.J. ACKLAND, M. METHFESSEL and C.O. RODRIGUEZ, 1993, Phys. Rev. B. **47**, 5571.
YU, J., and C.J. MCMAHON, 1980, Metallurg. Trans. **11A**, 277.
ZHANG, S.H., X.L. HE and T. KO, 1994a, J. Mater. Sci. **29**, 2655.
ZHANG, S.H., X.L. HE and T. KO, 1994b, J. Mater. Sci. **29**, 2663.
ZHOU, Y.-X., S.-C. FU, and C.J. MCMAHON, JR., 1981, Metall. Trans. **12A**, 959.

Further reading

BRIGGS, D., and M.P. SEAH, eds., 1990, Practical Surface Analysis. 2nd ed., Vol. 1, Auger and X-ray Photoelectron Spectroscopy (Wiley and Sons, Chichester).
BRIGGS, D., and M.P. SEAH, eds., 1992, Practical Surface Analysis. 2nd ed., Vol. 2, Ion and Neutral Spectroscopy (Wiley and Sons, Chichester).
DOWBEN, P.A., and A. MILLER, eds., 1990, Surface Segregation Phenomena (CRC, Boca Raton).
JOHNSON, W.C., and J.M. BLAKELY, eds., 1979, Interfacial Segregation (ASM Metals Park, Ohio).
KELLY, A., D.W. PASHLEY, E.D. HONDROS and C. LEA, eds., 1980, Residuals, Additives and Materials Properties. Proc. Conf. (The Royal Society, London); also appears as Phil. Trans. Roy. Soc. **A295**, 1–341.

McLEAN, D., 1957, Grain boundaries in Metals (Oxford Univ. Press).
RIVIÈRE, J. C., 1990, Surface Analytical Techniques (Clarendon, Oxford).
SUTTON, A. and R. W. BALLUFFI, 1995, Interfaces in Crystalline Materials (Oxford Univ. Press, Oxford).
WOLF, D., and S. YIP, eds., 1992, Materials Interfaces: Atomic-Level Structure and Properties (Chapman and Hall, London).

CHAPTER 14

OXIDATION, HOT CORROSION AND PROTECTION OF METALLIC MATERIALS

S. R. J. SAUNDERS

Division of Materials Metrology,
National Physical Laboratory,
Queens Road, Teddington,
Middlesex, TW11 0LW, UK

J. R. NICHOLLS

School of Industrial and Manufacturing Science,
Cranfield University,
Cranfield,
Bedfordshire, MK43 0AL, UK

R. W. Cahn and P. Haasen†, eds.
Physical Metallurgy; fourth, revised and enhanced edition

1. Introduction

1.1. Definitions

Oxidation and hot corrosion of metallic materials are oxidation reactions (in the chemical sense) that result in the conversion of metals into compounds such as oxides, sulphides or carbides when exposed to high temperatures above about 400°C. Oxidation processes as the name implies refer to formation of oxide layers and occur primarily during exposure to air or steam. Hot corrosion is the more general term where products other than oxides can form during exposure to air containing contaminants such as sulphurous gases from the combustion of fossil fuel or specific process conditions in petrochemical plants where high levels of carbon activity are common. In some cases exposure to molten salts occurs either directly or by condensation reactions, and in that case the term hot-salt corrosion is used.

1.2. General

Protection of material exposed to aggressive environments at high temperatures depends on the properties of the corrosion product which should act as a diffusion barrier to prevent further attack by the atmosphere. Thus the layer should be chemically stable in the environment, prevent diffusion of both the reacting gases and the metal species and finally, of course, in order to maintain its effectiveness as a diffusion barrier, the layer should be mechanically stable, that is it should not crack or spall. The importance of each one of these properties will become clear as various forms of attack are discussed in this chapter.

Oxidation and hot corrosion are important forms of materials degradation that often limit process conditions, which can have serious implications for plant reliability and efficiency. For example, power station boilers, petrochemical plant, gas and steam turbines, waste incinerators, diesel and internal combustion engines and nuclear reactors are all prone to degradation by oxidation and hot corrosion. The current drive toward energy saving by improved efficiency while at the same time minimising emission of pollutants places a heavy demand on the materials scientist, since in most cases these requirements mean higher operating temperatures and exposure to more aggressive environments. Clearly the high temperature behaviour of materials, therefore, can be a major factor inhibiting progress, and it is for this reason that the topic has such importance for our society.

Clean coal technology is a good example of the sort of developments in process technology that have important implications for the construction materials of the plant. Alternative processes to conventional combustion have been developed which involve fluidised bed combustion with additions of limestone as a sulphur capture agent, or gasification processes where materials are exposed to reducing environments which are relatively rich in sulphurous gases. In the case of fluidised bed combustion, heat exchange tubes must be resistant to the erosive action of particulates in the fluidised bed (sand, coal ash and limestone) and, since the sulphur is retained within the bed, also resistant to attack by the calcium sulphate formed by reaction of the limestone and the

sulphurous gases released during the combustion of coal. For coal gasification, on the other hand, process conditions are reducing to promote formation of CH_4, H_2 and CO from the coal/steam mixture, so that the metallic components only form a protective oxide layer if highly alloyed materials are used, and, furthermore, there is a heightened risk of sulphidation because of the relatively high sulphur potential. Again the process designer wants high operating temperatures to achieve high overall efficiency.

1.3. Outline

The thermochemistry of oxidation is considered and illustrated by means of Ellingham diagrams. Kinetics of oxidation of pure metals is developed and discussed in relation to the mechanisms of oxide formation. Alloy oxidation allows for the possibility of selective oxidation, and the criteria controlling the rate controlling processes are described. The importance of the strain tolerance of the protective oxide is then considered in relation to the mechanisms that lead to stress in the layer and the possibility of stress relief. Here the effects of thermal cycling and the mismatch in thermal expansion coefficient of the oxide layer and the metallic substrate are dominant.

Materials are not always exposed in simple oxidising environments, so that there is the possibility of competing reactions where oxide, sulphide or carbide may form. The conditions controlling these reactions are described. Again, simple gaseous attack can be further complicated by the presence of a deposit which can be molten or solid; here the chemical stability of the first formed layer may be inadequate, so that, for example, the oxide can dissolve when in contact with a molten layer.

Test and measurement methods are an important aspect of any technology, so that a brief introduction is given to this topic before considering life prediction and modelling.

Finally, protection is considered where alloy design is inadequate, or where the properties of the bulk (creep strength, for example) are incompatible with the corrosion resistance. Here a range of options is available, and the different processes for the application of protective coatings are described. Coating design is considered: mechanical and chemical properties must be matched with those of the substrate. Design philosophies have evolved that exploit graded coatings to optimise both surface and interface properties.

For a more detailed account of oxidation and hot corrosion the reader should consult reference books on the topic: KOFSTAD [1988a]; BIRKS and MEIER [1983]; KUBASCHEWSKI and HOPKINS [1966]; HAUFFE [1967], RAPP [1983] and EVANS [1960].

2. *Fundamentals of oxidation*

2.1. Thermodynamics

Classical thermodynamics can be used to determine the temperature and gas pressure required to cause the following reaction to take place:

$$M + \tfrac{1}{2}O_2 \rightarrow MO \qquad (1)$$

The Gibbs free energy for this reaction is:

$$\Delta G = \Delta G^{o} + RT \ln \frac{a_{MO}}{a_M \cdot a_{O_2}^{1/2}} \tag{2}$$

where: $\Delta G^{o} = -\Delta G^{o}_{M} - \frac{1}{2}\Delta G^{o}_{O_2} + \Delta G^{o}_{MO}$
a = activity
ΔG^{o} = standard free energy

If $\Delta G^{o} \geq 0$, spontaneous forward reaction will occur. At equilibrium, $\Delta G = 0$ and equation (2) reduces to the Van 't Hoff isotherm:

$$\Delta G^{o} = -RT \ln \frac{a_{MO}}{a_M \cdot a_{O_2}^{1/2}} \tag{3}$$

Since the standard free energy of elements is zero, equation (2) becomes:

$$\Delta G^{o}_{MO} = -RT \ln \frac{a_{MO}}{a_M \cdot a_{O_2}^{1/2}} \tag{4}$$

Equations (3) and (4) describe the conditions at equilibrium where both the metal and oxide are stable, or, in other words, where the formation of oxide on a metal just begins. The equation can be further simplified since the activity of a material in the pure state is unity and under conditions of low pressure and high temperature, the activity of oxygen can be approximated to its partial pressure, thus:

$$\Delta G^{o}_{MO} = \frac{RT}{2} \ln p_{O_2} \tag{5}$$

Therefore, knowing the Gibbs standard free energy of oxide formation at a given temperature, we can readily calculate the oxygen partial pressure required for oxidation to occur. Richardson and Jeffes have compiled the standard free energies of formation of oxides as a function of temperature, and fig. 1 shows these data. These diagrams are commonly known as *Ellingham diagrams* and allow the oxygen pressure for oxide formation at any required temperature to be determined directly. This is done by aligning the point O at the upper left of the diagram with the reaction and temperature of interest, and reading off the pressure on the scale on the right of the diagram. For example, for Al_2O_3 at 1000°C, align the point O with the line representing the reaction $4/3Al + O_2 = 2/3Al_2O_3$ to intersect it for 1000°C and read off the $p(O_2)$ of 10^{-37} atm on the $p(O_2)$ scale at the bottom of the diagram. Also included on this diagram are scales which give similar information about oxidation in two commonly encountered gas mixtures, CO/CO_2 and H_2/H_2O; in these cases, however, the critical gas mixture is given by aligning the temperature and reaction of interest with the points C and H, respectively. Similar diagrams can be constructed that indicate conditions for formation of sulphides, carbides, halides or any other reaction product of interest. (See also ch. 5, § 5).

In considering the oxidation of an alloy it is useful to decide which components oxidise preferentially. This may be determined by comparing the free energies of oxide formation; the oxide with the most negative free energy of formation forms preferential-

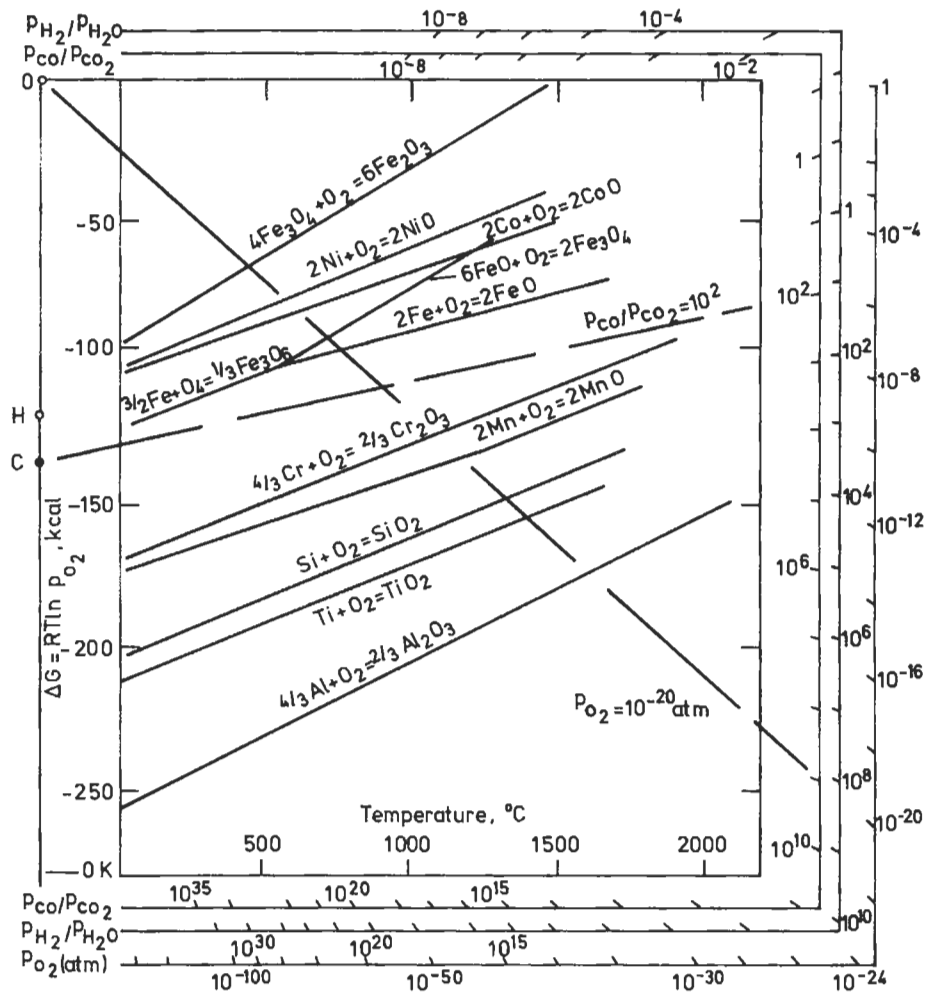


Fig. 1. Ellingham diagram for some oxides of importance in the high temperature oxidation of metals and alloys.

ly. Whether a separate phase forms depends upon the concentration of the components and the mutual solubility of the various oxide phases. If no new phases form, the oxidation rate may be changed by solution of the oxides, and as will be discussed later, this alters the number of defects in the oxide, thereby changing diffusion rates.

References: p. 1357.

2.2. Oxide structure

It is necessary to have an understanding of oxide structure before proposing models for the formation and further growth of the oxide layer. An oxide growing on a metal or alloy may be amorphous or crystalline; generally the very thin films are amorphous and thicker films crystalline. As will be shown later, oxide growth involves the movement of ions and electrons and it is important therefore to understand the restrictions imposed by the structure of amorphous or crystalline oxides upon the movement of ions and electrons.

2.2.1. Amorphous oxides

These oxides can be classified according to their bond strengths; i.e., glass or network formers with covalent bond strengths greater than 300 kJ mol^{-1}, intermediates with bond strengths between 300 and 200 kJ mol^{-1} and network modifiers with ionic bond strengths of less than 200 kJ mol^{-1}. ZACHARIASEN [1932], proposed certain rules for glass-forming oxides, which are useful although not universally applicable because of kinetic considerations. Oxides of the form MO and M_2O do not form networks; M_2O_3 can form a network if the oxygen ions form triangles around each metal atom; MO_2 and M_2O_5 form networks if oxygen tetrahedra form around each metal atom; MO_3, M_2O_7 and MO_4 are also probable network formers. The cations are tightly bound in a network former because of the high bond strengths. A large ion moves more easily than a small ion in such an open network because the small ions are mainly incorporated in the network so that anion transport would be expected in network formers. Since network modifiers do not form extensive networks and the bond strengths are generally low, cation transport is expected in these oxides.

2.2.2. Crystalline oxides

Crystalline oxides which form as the result of oxidation of a metal are generally non-stoichiometric (KOFSTAD [1972]) and thus exhibit both ionic and electronic conductivity. The non-stoichiometric oxide may be either metal-rich or metal-deficit. Oxides having an excess of metal ions maintain electrical neutrality by incorporating additional electrons; this type of compound is known as an n-type semiconductor. Oxides deficient in metal maintain electrical neutrality by the formation of positive holes; a positive hole occurs where the valence state of the cation in the oxide is higher than usual; these compounds are p-type semiconductors. The non-stoichiometry of these oxides can be accommodated in a number of ways. Figure 2 depicts an n-type semiconductor with the excess metal

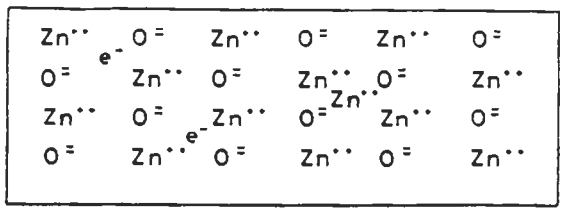


Fig. 2. Zinc oxide as an example of a metal excess (n-type) semiconductor.

located at interstitial positions in the lattice. Alternatively, the oxides may have oxygen ion vacancies. Figure 3 illustrates the structure of a p-type semiconducting oxide which can in principle have both oxygen interstitials and metal vacancies. Generally only metal vacancies are found, although in compounds of the CaF_2 type interstitial anions occur, for example ThO_2, UO_2, etc. The defect structure that has been described allows easy movement of ions, cations via interstitials or vacant cation lattice sites and anions via vacant anion sites or occasionally via interstitials.

It can be seen, therefore, that an n-type semiconducting oxide may allow anion or cation transport, whereas p-type oxides generally permit only cation transport. Table 1 classifies some oxides according to their semiconducting nature, where amphoteric oxides may have either metal excess or metal deficit (n- or p-type).

2.3. Kinetics

An essential concept in oxidation of metals and in hot corrosion generally is that the reaction product can act as a barrier between the metal and the environment thereby controlling rates of attack. Thus, the properties of the reaction product often determine subsequent rates of reaction and the mechanism of the oxidation process is reflected in the kinetic laws that are followed. By accurately determining the kinetics of the process from relatively short-term experiments it should be possible to determine long-term behaviour by extrapolation. However, this can only be done if there is no change in the process and, as will be discussed subsequently, discontinuities are frequently observed in oxidation kinetics. There is greater certainty about the reliability of such an extrapolation if the initial data represent a significant fraction of the lifetime to be estimated.

The movement of ions within an oxide layer, necessary for oxide growth, may involve a thermally activated diffusion process, or assistance by an electric field may reduce the overall activation energy to zero. The two basic mechanisms of oxide growth conveniently divide the subject into (i) the thin oxide film region (i.e., growth by a field assisted mechanism), and (ii) the thick film region (i.e., growth by thermal activation). The transition between the two processes usually occurs at a limiting thickness of about 100 nm.

In general, the kinetics of oxidation can be described by four laws, viz., logarithmic, cubic, parabolic and linear (fig. 4). All four kinetic laws are found in the thin film region. Linear for the first few monolayers of oxide growth, logarithmic up to about 50–100 atom layers of oxide and cubic or parabolic at still greater thicknesses. In the thick-film region a parabolic law is commonly encountered as well as linear growth laws. Often a transition

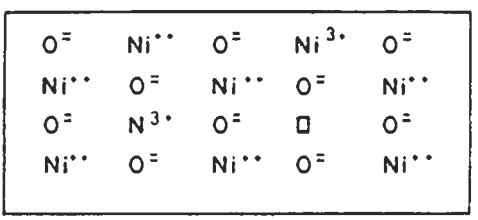


Fig. 3. Nickel oxide as an example of a metal deficit (p-type) semiconductor.

References: p. 1357.

Table 1
Electrical conductivity of oxides (after KUBASCHEWSKI and HOPKINS [1966])

Metal-excess semiconductors (n-type)
BeO, MgO, CaO, SrO, BaO, CeO_2, ThO_2, UO_3, U_3O_8, TiO_2, ZrO_2, V_2O_5, Nb_2O_5, Ta_2O_5, MoO_3, MnO_2, Fe_2O_3, $MgFe_2O_4$, $NiFe_2O_3$, $ZnFe_2O_4$, WO_3 $ZnCo_2O_4$, ZnO, CdO, Al_2O_3, $MgAl_2O_4$, $ZnAl_2O_4$, Tl_2O_3. SiO_2, SnO_2, PbO_2
Metal-deficit semiconductors (p-type)
UO_2, Cr_2O_3, (< 1250°C), $MgCr_2O_4$, $FeCr_2O_4$, $CoCr_2O_4$, $ZnCr_2O_4$, MnO, Mn_3O_4, Mn_2O_3, FeO, NiO, CoO, PdO, Cu_2O, Ag_2O, $CoAl_2O_4$, $NiAl_2O_4$
Amphoteric conductors
TiO, Ti_2O_3, VO, Cr_2O_3 (> 1250°C), MoO_2, (IrO_2), RuO_2

from parabolic (protective oxidation) to linear (breakaway oxidation) is observed and this is normally associated with mechanical failure of the protective oxide.

2.3.1. Thin film region

At low temperatures, electrons from the metal are able to pass through an oxide film to the oxygen adsorbed at the outer surface; electrons do this by the tunnel effect (a quantum-mechanical process by which electrons of lower maximum energy than that of a barrier possess a finite probability of penetrating that barrier). Electrons pass through the initial chemisorbed film to the outer surface and form oxygen anions, and metal cations form at the metal/oxide interface. It has been shown that the potential drop across the film is approximately 1 V so that with a film thickness of 5 nm, the field strength would be of the order of 10^7 V cm^{-1}. The strong electric field generated is able to pull ions through the film.

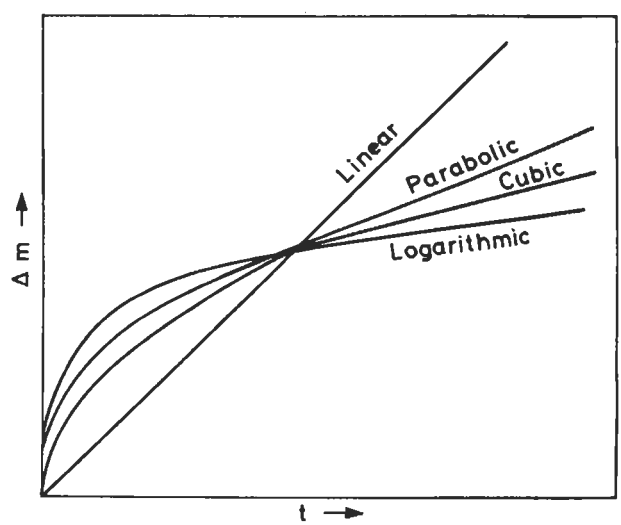


Fig. 4. Oxidation kinetics.

Logarithmic kinetics have been interpreted as resulting from the rate-controlling step being either anion movement or the electron tunnelling process which creates the field across the oxide layer initially (MOTT [1947]; HAUFFE and ILSCHNER [1954]), and an inverse logarithmic law can result when cation transport is the slow step (HAUFFE [1966]; EVANS [1960]; CABRERA and MOTT [1949]). Cubic or parabolic kinetics follow from the prediction that the rate of diffusion of ions is proportional to the magnitude of the electric field, which under constant potential decreases with increasing thickness; a parabolic law follows for p-type semiconductors and a cubic law for n-type semiconductors (CABRERA and MOTT [1949]; ENGELL *et al.* [1954]). Linear kinetics are observed in the formation of the first few monolayers of oxide where the growth process is controlled by surface diffusion and adsorption.

2.3.2. Thick film region

In the thick film (thermally activated) region, parabolic kinetics are derived by assuming that film thickening is controlled by diffusion in a concentration gradient ($x^2 = 2k_p't = k_pt$ where w is the oxide thickness, t is the time and k_p' (or k_p) is the parabolic rate constant). Linear kinetics are observed when the barrier provided by the oxide layer fails, so that either the oxide fails completely and there is direct contact between the metal and the gas, or partial mechanical failure occurs so that a constant thickness of protective oxide remains. Most practically important oxidation problems relate to the thick oxide film region and thus most emphasis will be given to the consideration of oxidation processes in this technically important region:

a) Wagner's theory

WAGNER [1975] derived an expression for the rate of thickening of the oxide expressed in terms of the specific conductivity of the oxide, the transport numbers of cations, anions and electrons and the free energy decrease during the oxidation reaction. This simplified derivation assumed that the conductivity is independent of the oxygen pressure and in fact this is not the case, so that a more complex expression results. There are only three metal oxide systems that accurately follow Wagner's parabolic equation; viz. Cu/Cu_2O, Co/CoO and Fe/FeO, and the very good agreement between the calculated and observed rate constants is evident from Table 2.

Wagner's theory of parabolic oxidation refers only to monophase scales, but in practice most oxides layers are multiphase, for example FeO, Fe_2O_3 and Fe_3O_4 all occur on iron when oxidised at temperatures above 600°C. It has been shown, however, that the rate of growth of each phase obeys parabolic kinetics and that their relative thickness is independent of time but temperature-dependent.

b) Deviations from Wagner's theory

One of the main assumptions in Wagner's model is that the scale is compact, but frequently scales are observed to be porous, and therefore under these conditions deviations from Wagner's theory are to be expected. A porous scale can form during cationic or anionic growth. For example, consider the case of nickel where upon oxidation a compact outer layer of NiO and a porous inner layer, also of NiO, are

Table 2
Measured and calculated rational rate constants*

$2Cu + O_2 \rightarrow Cu_2O$		
Temp. 1000°C Pressure, atm	k_r calc. eq cm^{-2} sec^{-1}	k_r obs. eq cm^{-2} sec^{-1}
8.3×10^{-2}	6.9×10^{-9}	6.2×10^{-9}
1.5×10^{-2}	4.8×10^{-9}	4.5×10^{-9}
2.3×10^{-3}	3.4×10^{-9}	3.1×10^{-9}
3.0×10^{-4}	2.1×10^{-9}	2.2×10^{-9}
$2Co + O_2 \rightarrow 2CoO$		
Pressure, 1 atm Temp. °C	k_r calc. eq cm^{-2} sec^{-1}	k_r obs. eq cm^{-2} sec^{-1}
1000	1.25×10^{-9}	1.05×10^{-9}
1148	5.15×10^{-9}	1.30×10^{-9}
1350	3.15×10^{-8}	4.65×10^{-9}
$2Fe + O_2 \rightarrow 2FeO$		
Pressure, 1 atm Temp. °C	k_r calc. eq cm^{-2} sec^{-1}	k_r obs. eq cm^{-2} sec^{-1}
800	2.5×10^{-9}	2.6×10^{-9}
897	1.1×10^{-8}	1.2×10^{-8}
983	2.8×10^{-8}	3.3×10^{-8}

* The rational parabolic rate constant, k_r, is used when the reaction rate is expressed in terms of equivalents of M_aO_b per square centimetre per second.

formed. A single layer of nickel oxide forms during the early stages of growth of the oxide, but as the layer thickens a duplex structure develops, which consists of an inner region of equiaxed, fine-grained crystallites and an outer region of large columnar crystals. The inner layer is generally more pronounced on less pure materials, and is believed to be due to the presence of impurities segregating to the grain-boundaries, thereby inhibiting grain growth. Tracer diffusion studies have shown that the outer layer grows by movement of nickel vacancies along the grain-boundaries, and the inner layer by molecular oxygen penetration along microcracks and fissures which are present in the outer layer due to the build-up of stress in that layer. In the case of nickel oxide, compressive stresses result because of the constraints imposed on the oxide layer by the receding metal.

A dissociative mechanism has been proposed by MROWEC [1967] for the formation of scales of this type. Metal atoms enter the scale at the metal/oxide interface and diffuse through the existing scale and react with oxygen at the oxide/oxygen interface to form new oxide. There must be corresponding movement of cation vacancies in the opposite direction so that a cationic vacancy flux arrives at the metal/oxide interface and must

somehow be removed. Whilst the film is thin, plastic deformation of the oxide is sufficient to allow continued adherence to metal, but as it thickens the oxide becomes stronger and can no longer deform, so that longitudinal cracks form at the metal oxide interface and the rate of transport of metal from the substrate to the oxide decreases. However, because of the chemical potential gradient in the scale, outward diffusion of metal can still occur. This disturbs the equilibrium at the inner surface of the scale which leads to an increase in chemical potential of oxygen at that surface and eventually oxide decomposes to form molecular oxygen, cations and electrons. The growth of an inner layer due to the presence of molecular oxygen is then possible and the processes described are illustrated in fig. 5 which shows the dissociative mechanism for double layer formation in a single-phase scale. This explanation is disputed by KOFSTAD [1985] who proposed that pores form due to plastic deformation and creep caused by growth stress in the scales. For scale growth controlled by diffusion, a gradient in defects will be developed through the scale such that a lower concentration of defects will occur in the outer layers of the scale; this will have the effect of promoting faster diffusional creep in these outer layers, so that there will be greater stress relief there and will thus avoid the development of the porous structure observed in the inner layers. KÜPPENBENDER and SCHÜTZE [1994] have demonstrated that "pseudo"-plasticity can occur by a micro-cracking and healing mechanism. Pseudo-plasticity requires compressive deformation in one or two directions and the possibility of oxide grain movement in at least one direction perpendicular to the compressive deformation axis.

The other major assumption often used in applying Wagner's theory is that bulk lattice diffusion processes occur. This is only valid at high temperatures ($> \sim 0.6\ T_{mp}$), and most industrially relevant materials are used at lower temperatures. Thus, for polycrystalline oxides, grain-boundary diffusion becomes the dominant process. This is shown very clearly by referring to fig. 6 which shows the parabolic rate constant for the oxidation of nickel as a function of temperature compared with calculated lattice and grain-boundary

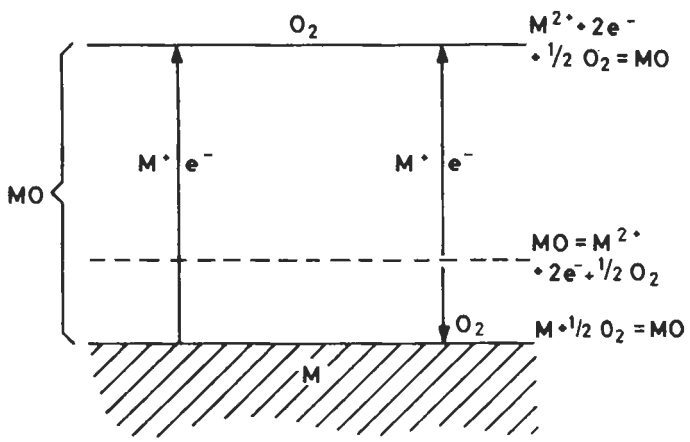


Fig. 5. Schematic representation of the dissociative mechanism.

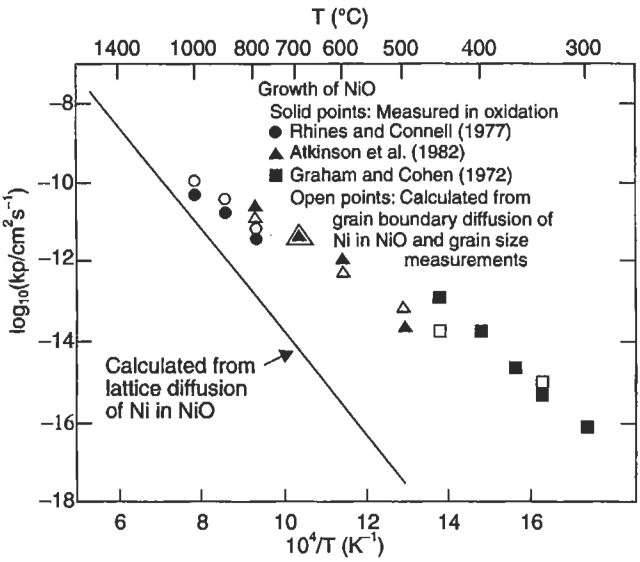


Fig. 6. Comparison of the oxidation rate of nickel with lattice and grain-boundary diffusion of Ni in NiO (ATKINSON [1987]).

diffusion constants (ATKINSON [1987]). By extrapolation of the experimental data shown in fig. 6, it can be seen that it is only at temperatures of about 1200°C and above that the activation energies for lattice diffusion and oxide formation would be similar (230–250 kJ mol^{-1}). At the lower temperatures the rate of oxidation is increasingly greater than would be predicted assuming control by bulk diffusion in the oxide lattice, activation energies being reported in the range 155–170 kJ/mol. The effect of prior cold-work in the nickel is to increase the oxidation rate, but the observed rate law is usually less than parabolic. Both these observations suggest that the rate of oxidation is controlled by grain-boundary diffusion in the oxide; the less than parabolic rate observed in the cold-worked material occurs because the initially fine-grained oxide coarsens during the oxidation process, thereby eliminating some short-circuit diffusion paths. Models have been developed (e.g., SMELTZER *et al.* [1961]) to describe the oxidation reaction where the rate is controlled by dual lattice and grain-boundary diffusion in which the effective diffusion coefficient is given by:

$$D_{eff} = D_l + 2(D'\delta)/g \tag{6}$$

where g is the grain size normal to the growth direction, δ is the grain-boundary width, and D' and D_l are the diffusion coefficients of the boundary and the lattice, respectively. ATKINSON [1987] used this model to calculate parabolic rate constants for the oxidation of nickel where grain-boundary diffusion dominates, and there was good agreement between the calculated values and those obtained experimentally, as shown in fig. 6.

2.4. Properties of oxide layers

It will have become clear from the preceding discussion that the electrical and mechanical properties of an oxide layer are of paramount importance in controlling the course of the oxidation reaction. We will now consider how these properties may be modified in order to control the rate of oxidation of metals.

2.4.1. Electrical properties (diffusion)

The defect structure of an oxide may be modified by a small addition of other elements and, if the oxidation process proceeds by the Wagner mechanism, one might expect changes in the oxidation rate to occur. For example, the addition of a monovalent metal to NiO, a cation-deficit p-type semiconductor, would be expected to reduce the cation vacancy concentration and hence also decrease the oxidation rate. Conversely, if ions of a higher valency are added to a p-type oxide, then the cation vacancy concentration will increase and thus also the oxidation rate. The reverse situation would be expected for n-type oxides.

Similarly, if the anion sub-lattice is altered, for example, by the incorporation of N^{3-}, this would increase the number of vacant anion sites in an n-type semi-conductor and hence the oxidation rate would be increased if anion transport were the rate controlling step. This mechanism has been proposed to explain the differences in oxidation rates of titanium in pure oxygen (where a low value is observed) and in air (a high value).

For a dopant to be effective the solubility, which is determined principally by ionic size, must be sufficiently high. A further probable restriction is that for growth processes controlled by cation diffusion the dopant should be added from the gas phase since new oxide is formed at this interface, and for anion diffusion the dopant should be added from the alloy.

Thus it can be seen that the addition of minor amounts of other elements can cause large differences in the oxidation rate where the Wagner mechanism is operating. It has already been noted that the Wagner mechanism has only been quantitatively established in three cases. However, there are many more examples where qualitative effects of the type just described confirm the validity of this mechanism.

In the case where grain-boundary diffusion is rate-controlling, the effect of impurities can be to alter the diffusion rates *per se* or to influence the microstructure. In the latter case segregation effects are likely to dominate, leading to a reduction in grain size and a consequent increase in the oxidation rate. Very few data exist to indicate which is the controlling effect, but, in the case of nickel, theoretical modelling of the grain-boundary structure by DUFFY and TASKER [1986] and experimental measurements of grain-boundary diffusion rates (ATKINSON and TAYLOR [1986]) indicate that impurities often decrease grain-boundary diffusion rates (fig. 7). Thus it would appear that the effects of impurities in increasing the oxidation rate of nickel most probably results from a reduction in the oxide grain size with a consequent increase in the number of short-circuit diffusion paths.

References: p. 1357.

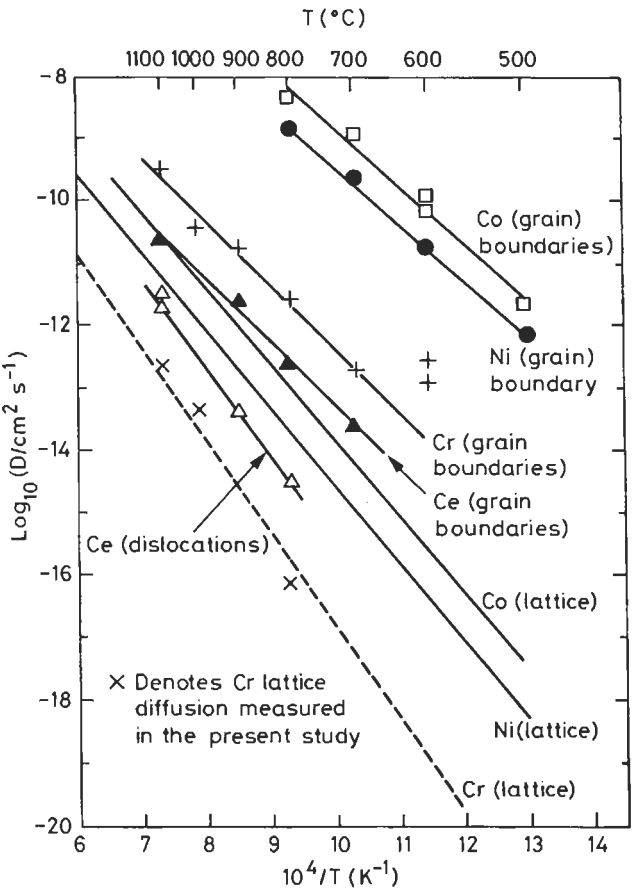


Fig. 7. Arrhenius plot summarising the diffusion coefficients for various impurity cations in the NiO lattice, grain-boundaries and dislocations (ATKINSON and TAYLOR [1986]).

2.4.2. Mechanical properties (stress generation and relief)

Very recently a critical review of this topic was undertaken in a Workshop on the Mechanical Properties of Protective Oxide Scales which was published in Volume 12 of the Journal *Materials at High Temperatures*, 1994.

It was pointed out some time ago by PILLING and BEDWORTH [1923] that, depending upon the growth processes involved, an oxide layer could develop internal stresses as a result of the growth process. If, for example, cation transport is the rate-controlling process, new oxide would form at a free surface, that is, the oxide/gas interface, and thus no stress would be imposed. However, if the oxide grows by oxygen transport, new oxide would form at the metal/oxide interface and the metal lattice would then have to expand or contract to accommodate the newly formed oxide. If the volume of oxide were

larger than the volume of metal consumed, the oxide would be in compression, and, if vice versa, the oxide would be in tension. The ratio, oxide volume/metal volume, is called the *Pilling–Bedworth ratio,* and for most metals it is greater than one. Growth stresses have also been observed in oxides which form by cation transport, so that other factors need consideration. It is clear that unless some stress-relief mechanism operates, build-up of growth stresses will ultimately result in cracking or blistering of the oxide layer, and, as was discussed earlier, this would result in a change of kinetics from parabolic (protective oxidation) to linear (non-protective or breakaway oxidation).

In addition to stresses arising from the oxide volume/metal volume ratio, other mechanisms for *stress generation* have been proposed and recently reviewed by STOTT and ATKINSON 1994. These include: epitaxial stresses, recrystallisation within the scale, the composition of the metal or oxide may change during oxidation, point defect stress, oxide growth within the oxide layer, coulombic forces due to the electric field produced by differing mobilities of the ions, and finally thermal stress (i.e., stress arising from thermal cycles where a thermal expansion coefficient mismatch exists between the oxide and substrate) which is usually by far the most important cause of stress generation in oxides formed on engineering iron- and nickel-base alloys.

That zirconium oxidises to produce a coherent oxide is surprising, since anion transport is favoured in the oxide and, with a large Pilling–Bedworth ratio, one would expect considerable stresses to be developed which could cause the oxide to crack or spall. It is evident, therefore, that various mechanisms which allow *stress relief* by plastic deformation must operate. The following processes have been proposed; dislocation glide, grain-boundary sliding, mechanical twinning, viscous flow in amorphous oxides, Herring–Nabarro or diffusion creep, and finally, dislocation climb processes are often observed and are probably the most generally applicable mechanism for stress relief in oxides.

Where plastic deformation of oxides is insufficient for benign stress relief, catastrophic stress relief will take place by spallation and scale cracking. EVANS [1994] has developed spallation maps which can be used to predict regions of oxide stability as a function of the temperature drop during thermal cycling. These maps are based on two possible modes of scale spallation in which buckling can occur when the interface is weak and the oxide relatively strong (usually when the oxide is thin) or the wedging mode of failure where the interface is strong and the oxide is relatively weak (usually for thick oxides). Figure 8 is an example of such a spallation map, calculated on the basis of measured oxide properties for alumina formed on a ferritic steel.

In simplistic terms, scale failure occurs when the *in-situ* stress exceeds some critical value (σ_c). The critical stress can be evaluated using simple fracture mechanics considerations to predict (EVANS and CANNON [1987]; HANCOCK and NICHOLLS [1988]), whether through-thickness cracking (tensile failure) or scale spallation (compressive failure) results. For the case of tensile failure, the critical stress is given by:

$$\sigma_c = K_{1c}\left(\pi \bar{a}\right)^{-1/2} \tag{7}$$

where K_{1c} is the fracture toughness and $\bar{a}$ is a measure of the defect size present in the oxide. K_{1c} is related to the energy of fracture by $K_{1c}^2 = 2\ E_o \gamma_o$ where E_o is the oxide

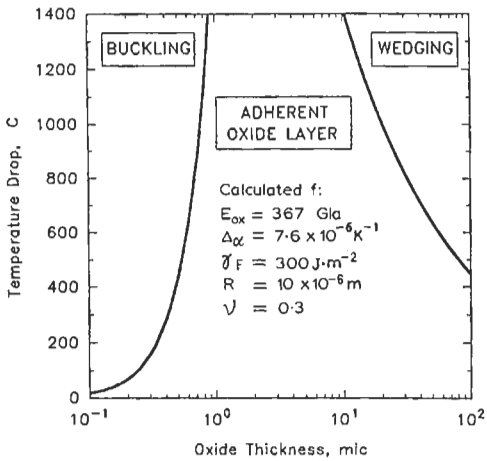


Fig. 8. Notional spallation map for alumina on Fecralloy steel (EVANS [1994]).

elastic modulus and γ_0 is the energy required to form a unit area of new oxide surface. Hence fracture mechanics analysis (EVANS and CANNON [1987]; HANCOCK and NICHOLLS [1988]; NICHOLLS and HANCOCK [1989]) and the assessment of fracture strain energies (EVANS [1988]) are alternative approaches to solving the same problem.

The *in-situ* stress acting at a defect, within the oxide or at the oxide/metal interface, is the summation of all stresses acting on the defect. This stress will include growth stresses, thermally induced stresses and directly applied stresses which may be tensile, compressive or bending. Hence to quantify oxide fracture, the state of stress, fracture toughness and defect distributions present within the oxide or at the metal/oxide interface (depending on the mode of failure) must be capable of being assessed. This topic was recently reviewed in depth by HANCOCK and NICHOLLS [1994].

3. *Oxidation of alloys*

The same principles apply to the oxidation of alloys as have been discussed for pure metal oxidation. The problem is more complex however, because there are at least two constituents which can be oxidized. The added complexity makes prediction of alloy oxidation rates extremely difficult.

3.1. Selective oxidation

Selective oxidation of a component of an alloy will occur when the free energy of formation of its oxide is more negative than that of the oxide of the other alloy constituents. Dependent upon the difference in free energies, temperature, oxygen pressure and concentration, it is possible for the oxide of the minor constituent to form to the exclusion of all other oxides. WAGNER [1956] derived an expression for determining the

minimum atomic fraction $N_{A(min)}$ in an alloy AB for exclusive formation of a coherent layer of the oxide AO, given that the AO will form in preference to BO:

$$N_{A(min)} = \frac{1}{16 Z_A c}\left(\frac{\pi k_p}{d}\right)^{1/2} \tag{8}$$

where d is the interdiffusion coefficient in the alloy, Z_A is the valency of A in the oxide, k_p is the parabolic rate constant for growth of oxide AO, and c is the number of gram atoms of metal per unit volume.

This equation should be treated as predicting only the necessary minimum concentration; it is not a sufficient condition in itself. For example, for the copper-aluminium system at 800°C the calculated value $N_{A(min)} = 0.004$ (0.0008 wt% Al); however, in practice even when $N_A = 0.11$ (5 wt% Al) some copper oxides still form. Examination of the compositional changes that occur at the alloy/oxide interface as a result of selective oxidation shows that the concentration profile is determined by two opposing processes, consumption by the oxidation process and replenishment by diffusion within the alloy. Clearly, if a protective oxide is removed for any reason and a depleted alloy is exposed to the oxidising environment, the protective layer may not be able to reform; it is this situation which is usually responsible for breakaway oxidation. Exclusive formation of Al_2O_3, Cr_2O_3 or less frequently, SiO_2, is most commonly used as a means of protecting many industrially important alloys from oxidation. Generally, the oxidation rate of

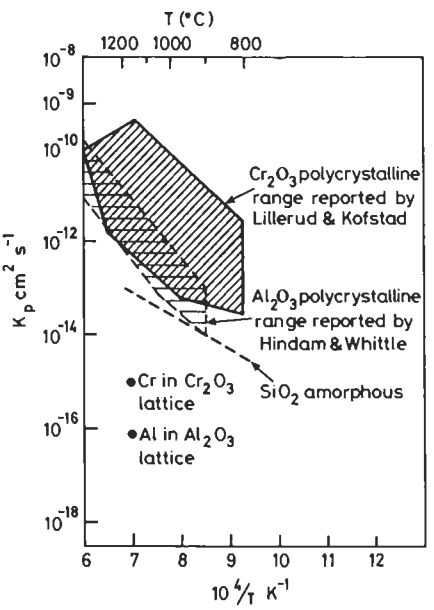


Fig. 9. Observed parabolic rate constants for alloys forming alumina and chromia scales compared with oxide lattice diffusion rates and the rate constant for the formation of silica on silicon; the solid circles refer to the lattice diffusion rates of Cr and Al in Cr_2O_3 and Al_2O_3, respectively, at 1180°C.

References: p. 1357.

alumina-forming alloys is lower than that of chromia-forming alloys, but the lowest rates of all result when silica scales form (fig. 9). In the latter case this is due to the fact that silica is amorphous at temperatures up to about 1200°C and thus grain-boundary diffusion is no longer a factor controlling mass transport.

In cases where several different oxides form, the scale forms in such a way that the oxide with the highest dissociation pressure (most positive ΔG°) forms at the oxide/gas interface and that with the lowest dissociation pressure at the metal/oxide interface. In other cases selective oxidation results in an enrichment of an alloy constituent in the oxide layer relative to the alloy concentration without exclusive formation of its oxide. If the oxides are mutually soluble the major constituent of the scale will have its defect structure altered in a manner previously described by incorporation of the solute. Where the oxides do not form solid solutions, composite oxide layers result. A commonly encountered situation is the formation of complex oxides (double oxides, spinels, etc.). This may happen when oxides of the alloying components react with one another to form a more stable oxide. Ni–Cr, Ni–Al and Fe–Cr alloys provide good examples of this situation where spinel layers result. It is often observed that diffusion rates are much lower in spinels than in normal oxides, and protective oxide scales may consist of these complex oxides. WOOD and STOTT [1983] have summarised the oxide growth mechanisms for these cases and the various possibilities are illustrated schematically for the nickel–aluminium system as shown in fig. 10, where alloys rich in aluminium eventually form an inner layer of alumina, or for more dilute alloys where the alumina forms only as an internal oxide. In the first case, after simultaneous formation of both oxides, the slower growing oxide (Al_2O_3) forms in the inner regions by the displacement reaction

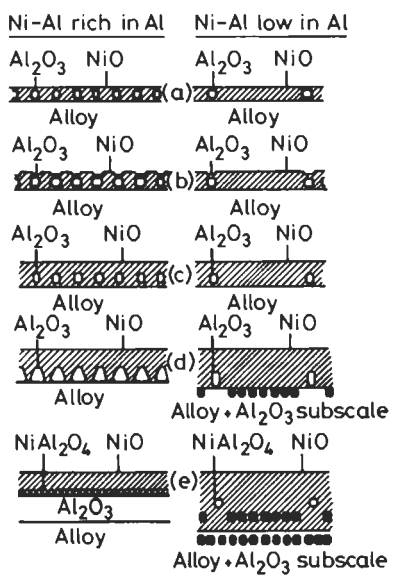


Fig. 10. Schematic representation of transient oxidation of Ni–Al alloys (WOOD and STOTT [1983]).

$3NiO + 2Al^{3+} \rightarrow Al_2O_3 + 3Ni^{2+}$, while at the interface between the Al_2O_3 and the NiO layers a solid solution reaction occurs to allow formation of the complex oxide $NiAl_2O_4$. The case of internal oxidation is discussed in more detail in the next Section.

3.2. Internal oxidation

When oxygen dissolves in the alloy, the more active component (most negative ΔG°) may form a precipitate of oxide particles within the alloy. This phenomenon is known as *internal oxidation* or sub-scale formation. Internal oxidation requires a higher diffusion rate of oxygen in the alloy than that of the alloying component. This means that an oxygen gradient is established and the dissolved oxygen reacts with the most reactive constituent to form the oxide. WAGNER [1959] studied the transition from internal to external oxidation and has derived a model based on the fact that the cross-section available for oxygen diffusion is reduced by the presence of internal oxide. Above a certain alloy concentration, oxygen diffusion slows down to the extent that only an external oxide forms. Lowering the oxygen partial pressure also encourages external oxidation since the solubility of oxygen in the alloy is reduced. If internal oxidation occurs, the oxidation rate is usually high because a protective oxide layer is not established. The transition to external oxidation will result in a reduction of the oxidation rate since now a complete barrier between the alloy and the environment exists.

3.3. Intermetallic alloys

Intermetallic materials are increasingly considered for engineering applications because of their high temperature strength and low density, and a comprehensive review has recently been published by DOYCHAK [1994]. Generally these materials are based on the aluminide or silicide systems where protective alumina or silica scales can form. A particular problem relating to intermetallics, especially $NbAl_3$ and $MoSi_2$, has been the so-called "pest", where accelerated oxidation occurs at an intermediate temperature (550–850°C) after a brief incubation period. Early work on the mechanisms of this process proposed that it was stress-induced, but more recent work (GRABKE *et al.* [1991]; DOYCHAK *et al.* [1991]) suggests that the rapid degradation of $NbAl_3$ is the result of inward diffusion of oxygen through an outer layer of Nb_2O_5, continued inward diffusion of oxygen along the aluminium-depleted intermetallic grain-boundaries and precipitation of internal alumina within the intermetallic in the region unaffected by oxidation, as illustrated schematically in fig. 11.

3.4. Scale adhesion

As was mentioned in the introductory section, a major problem that must be addressed once a protective film has been established is the control of adhesion or mechanical integrity. It has already been pointed out that the protective oxide layer can be subjected to stresses dependant upon the operating growth mechanisms, which usually results in compressive stresses. Additionally, however, thermal cycles (cooling to room temperature) also result in compressive stress being applied to the oxide scale since most

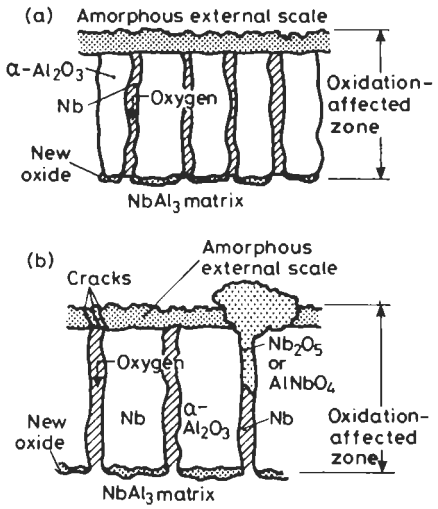


Fig. 11. Schematic of a) the accelerated oxidation mechanism on $NbAl_3$ at 727°C and b) the initiation of pest degradation (DOYCHAK and HESBUR [1991]).

oxides have a lower coefficient of thermal expansion than the metal or alloy. The additive effect of these compressive stresses is often sufficient to cause scale spalling.

Small additions of "*active elements*" (i.e., elements with a high affinity for oxygen) and notably the rare earths, are known to be very effective in promoting the formation of an adherent oxide layer that is resistant to thermal cycles. The active element can be added in elemental form or as an oxide dispersoid. In the latter case, a novel series of alloys produced by mechanical alloying — the so-called *oxide dispersion strengthened* (ODS) materials which have been developed primarily for enhanced high-temperature strength, also show good oxidation resistance (MICHELS [1977]). The alloy MA6000 (Ni-15 Cr, 4.5 Al, 4 W, 2.5 Ti, 2 Mo, 2 Ta, 1.1 Y_2O_3, by weight) is now being used in some gas turbine applications. Early versions of this type of material, e.g., thoria-dispersed (TD) alloys, were evaluated some time ago and STRINGER *et al.* [1972] reported that TD–Ni20Cr had excellent resistance to spalling and reduced oxidation rates compared with the simple binary alloy.

As well as improving adhesion of the scale, the active element addition also reduced the growth rate and the concentration of chromium or aluminium required for preferential formation of the scale, particularly for the chromia-forming alloys. WHITTLE and STRINGER [1980] reviewed the various theories that have been proposed to account for this effect, which include enhanced scale plasticity, formation of a graded seal, modification to the oxide growth process, stronger chemical bonding at the interface, elimination of voids by inert oxide particles acting as vacancy sinks and, finally, oxide protrusions into the alloy which act as "pegs" to improve adhesion. More recently, LUTHRA and BRIANT [1986] SMEGGIL *et al.* [1984] and LESS [1987] have proposed that segregation of sulphur to the scale/alloy interface is responsible for the poor adhesion of the oxide, and

that the effect of the "active element" is to scavenge the sulphur present in the alloy and so restore the intrinsically strong bond between the oxide and the substrate. LUTHRA and BRIANT [1986] have been unable to confirm this effect experimentally, but both FUNKENBUSCH *et al.* [1985] and SMIALEK [1987] have reported results that indicate that for high-purity alloys with sulphur contents of less than about 10 ppm, adherent oxide scales were formed on alumina-forming nickel-base alloys without rare earth alloying additions. PIERAGGI and RAPP [1994], proposed that the active element blocks active sites, i.e., interfacial dislocations, that support diffusion growth at the scale/alloy interface, and thereby alter the growth mechanism and hence also the adhesion of the scale. All these various mechanisms were recently reviewed by STRAWBRIDGE and HOU [1994], and it is clear that no one theory can satisfactorily explain all the experimental observations.

4. Multi-component atmospheres

4.1. Phenomenology

In many industrially relevant cases, alloys are exposed to environments containing a mixture of gases so that there exists the possibility of forming phases other than oxides. Commonly sulphides, carbides and to a lesser extent chlorides are formed, and in almost every case formation of these phases results in greatly increased rates of attack. This is because diffusion rates in sulphides and carbides are higher than in the corresponding oxide, and chlorides are often volatile, thus providing no barrier to further reaction. However, in most practical cases, alloys are used that contain reactive elements such as aluminium or chromium and there is usually sufficient oxygen present in the environment to allow formation of an initially protective oxide layer. Nevertheless, ingress of sulphur, carbon or chlorine through this protective layer can occur, leading to formation of the sulphide, carbide or chloride of the oxide-forming constituent within the alloy so that it becomes depleted in that constituent. Thus, should that oxide layer be damaged by growth stresses or differential contraction induced by thermal cycles, the alloy would be unable to reform a protective oxide. Increased rates of attack may not always be entirely due to alloy depletion effects. For example, oxide doping could occur, and it has been suggested that in the case of Cr_2O_3, S^- is introduced which leads to an increase in the chromium vacancy concentration and a corresponding increase in diffusion rates. In addition, the size of the impurity ion may cause a dilation of the crystal lattice and hence also encourage diffusion. Since most metallic sulphides, carbides or halides are less stable than the oxide, continued reaction results in the conversion of these second phase particles to the oxide which does not form a coherent protective layer, as shown, for example, in the following reaction:

$$2MS + O_2 \rightarrow 2MO + S_2 \tag{9}$$

It can be seen from this reaction that sulphur is released and is then available for further attack upon the alloy. This form of cyclic sulphidation/oxidation/sulphidation is frequently observed and leads to catastrophic rates of attack.

References: p. 1357.

The process by which these aggressive species penetrate the protective oxide layer is often greatly disputed in specific cases, but in principle there are two possibilities: (a) bulk diffusion or (b) short-circuit diffusion via grain-boundaries or microcracks within the oxide layer. It will be shown in a following section that the nature of the diffusion path can sometimes be determined from the reaction products.

4.2. Prediction of reaction products

A powerful method of determining the nature of the product when an alloy reacts with a complex gas mixture is to construct the relevant phase stability diagram, and with a knowledge of the composition of the gas mixture the thermodynamically stable phases are indicated. For example, the phase stability diagram for the Fe–O–S system at 800°C plotted as a function of the partial pressures of oxygen and sulphur is shown in fig. 12. Thus, for a typical combustion gas mixture containing CO_2, O_2, N_2, H_2O and about 400 ppm SO_2 the composition can be determined in terms of $p(O_2)$ and $p(S_2)$ and is shown as "X" in the diagram. It can be seen that Fe_2O_3 is the stable phase. However, if $p(O_2)$ is reduced and $p(S_2)$ is increased as, for example, in a coal gasification environment, the composition of this gas mixture given as "Y" in the diagram indicates that FeS would be the stable phase. It should be noted that the axes selected for these diagrams do not relate to the major constituents of the gas mixture, but because equilibrium is assumed in the calculation, the same result would be obtained no matter what oxygen- or sulphur-containing species were used as a variable. For simplicity, oxygen and sulphur are usually chosen when plotting diagrams of this nature.

The phase stability diagram can be constructed by considering all possible equilibria in the three component system and writing equations in terms of these components, for example:

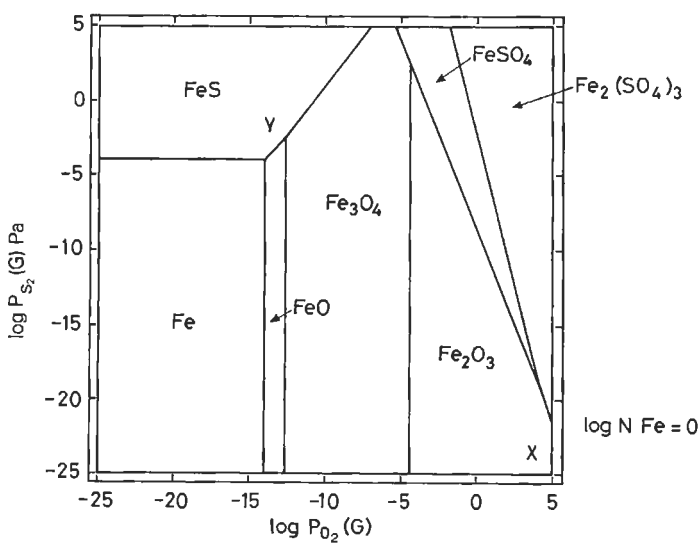


Fig. 12. Phase stability diagram for the Fe–O–S system at 800°C (partial pressures in pascals).

$$Fe + \tfrac{1}{2}O_2 \rightarrow FeO$$

where the equilibrium constant,

$$K = \frac{a_{FeO}}{a_{(Fe)} \cdot \sqrt{P_{O_2}}} = 10^9 \text{ at } 800°\,C \tag{10}$$

thus, knowing the equilibrium constant, the partial pressure of oxygen at the Fe/FeO phase boundary can be calculated and plotted on the diagram. Similarly all other relevant equilibria can be calculated at the temperature of interest. Use of these diagrams clearly shows the benefits of addition of elements such as aluminium and chromium in high-temperature corrosion-resistant alloys. Figures 13 and 14 are phase stability diagrams of Cr–O–S and Al–O–S systems, respectively, for the same conditions as were plotted in fig. 12 for Fe–O–S. It can be seen that oxide formation covers a much greater range of conditions, so that for an iron–aluminium or iron–chromium alloy, protective oxides could form in cases where, in the absence of these alloying additions, iron sulphide would appear. For example, the region delineated by a rectangle in figs. 13 and 14 defines typical conditions encountered in a coal gasification system, where normally iron sulphide would form (point "Y" in fig. 12).

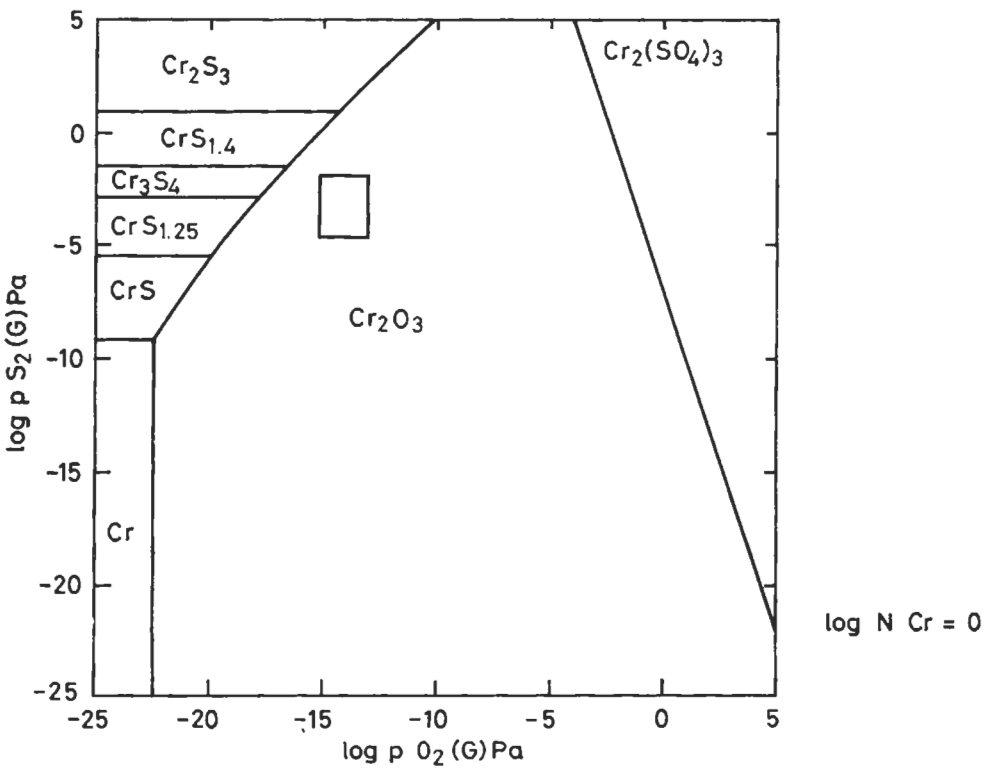


Fig. 13. Phase stability diagram for the Cr–O–S system at 800°C (partial pressures in pascals).

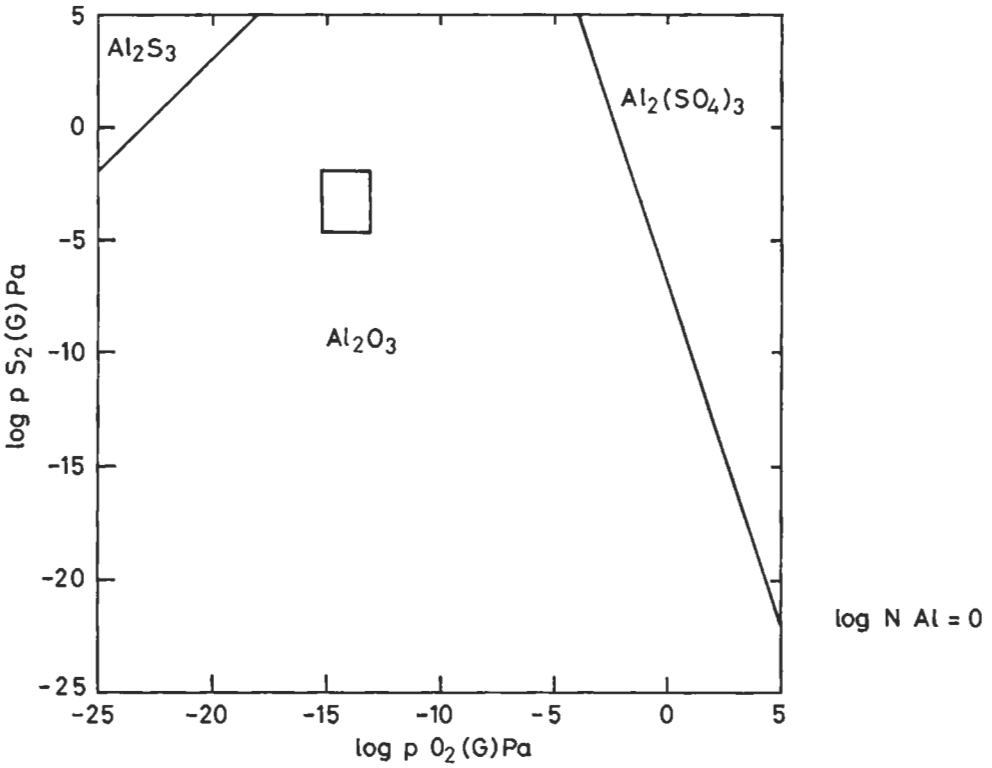


Fig. 14. Phase stability diagram for the Al–O–S system at 800°C (partial pressures in pascals).

An even more complex problem, however, is to consider relevant phase diagrams of an *alloy* in a multi-component gas mixture where oxygen-, carbon- and sulphur-containing species may result in oxide, carbide and sulphide formation. Commonly all the phase stability diagrams for individual elements are superimposed so that an approximate description is obtained. Clearly, a complete representation of the equilibria could only be depicted by using a three-dimensional diagram. These are very difficult to calculate and even more difficult to construct. Approximations, therefore, are required, and it is convenient to consider the behaviour of an element in an alloy by plotting diagrams of a single element using a reduced activity. This approximation neglects the possibility of mixed or complex oxide formation, but it is now possible to carry out the necessary calculations to produce phase stability diagrams for mixed oxide systems (BARRY and DINSDALE [1994]; PELTON *et al.* [1979]). If the gas mixture contains more than two active components a useful approach is to maintain one component at a constant level whilst varying the other two and then systematically alter components of interest to cover the range required. Clearly these are approximations and this needs to be borne in mind when reviewing the information available from these diagrams.

It must be remembered, however, that these diagrams will only indicate the presence of a phase by assuming equilibrium conditions, but kinetic factors are also very important. For example, in some conditions it has been observed that oxide phases only form *exclusively* at partial pressures of oxygen about 10^4 greater than the values for the thermodynamic boundary for transition from sulphide to oxide formation, giving rise to the so called "kinetic boundary" (NATESAN [1983]). Figure 15 shows the position of the kinetic barrier for an austenitic stainless steel (AISI 310) at 650°C on the superimposed phase stability diagrams for iron, chromium and nickel, together with the region defining typical coal gasification atmospheres shown as a rectangle in the figure. Two points of interest arise from this diagram. Firstly, that the kinetic barrier for this alloy lies within the composition used in coal gasification so that in many plants this alloy would be unsuitable, and secondly, the position of the line is also very close to the FeS/FeO phase boundary. This latter observation may explain the origin of the kinetic barrier with chromia-forming iron-based alloys, since if iron diffusion occurs, sulphides would then form on the surface and promote sulphur migration through the chromia layer.

4.3. Reaction path

If one assumes that an oxide layer forms in a multi-component atmosphere and aggressive species such as carbon-, sulphur- or chlorine-containing gases are also present and are able to penetrate the oxide, then the nature of the corrosion product may depend upon the diffusion path in the oxide. If only bulk diffusion of the aggressive species occurs, then, since diffusion can only take place in a concentration gradient, the partial pressure of, for example, sulphur would be lower at the metal/oxide interface than at the oxide/gas interface. If, on the other hand, the oxide contained microcracks and the gas were able to penetrate the oxide, the gas equilibrium would no longer be controlled by the bulk gas composition because the partial pressure of oxygen would be set by the metal/metal oxide equilibrium and not by that in the gas phase (STROOSNIJDER and QUADAKKERS [1986]). In most practical cases where protection is conferred by formation of alumina or chromia, the partial pressure of oxygen at the metal/oxide interface is much lower than that in the gas phase; thus for a gas containing SO_2, for example, this could result in a large increase in the partial pressure of sulphur at the metal/oxide interface compared with that in the bulk gas mixture. The effect of reducing the partial pressure of oxygen is clearly seen by considering the following equilibrium:

$$SO_2 \rightleftharpoons O_2 + \tfrac{1}{2}S_2$$

where the equilibrium constant

$$K = \frac{P_{O_2} \cdot \sqrt{P_{S_2}}}{P_{SO_2}} = 10^{-14} \; at \; 800°\,C \tag{11}$$

In the case of a nickel chromium alloy exposed to a combustion gas mixture containing 400 ppm SO_2 at 800°C, the $p(S_2) = 10^{-21}$ and $p(O_2) = 10^3$ Pa. However, at the alloy/oxide interface where Cr_2O_3 forms, the partial pressure of oxygen would equal 10^{-22} Pa, and this in turn would cause $p(S_2)$ resulting from 400 ppm SO_2 at the metal/oxide interface

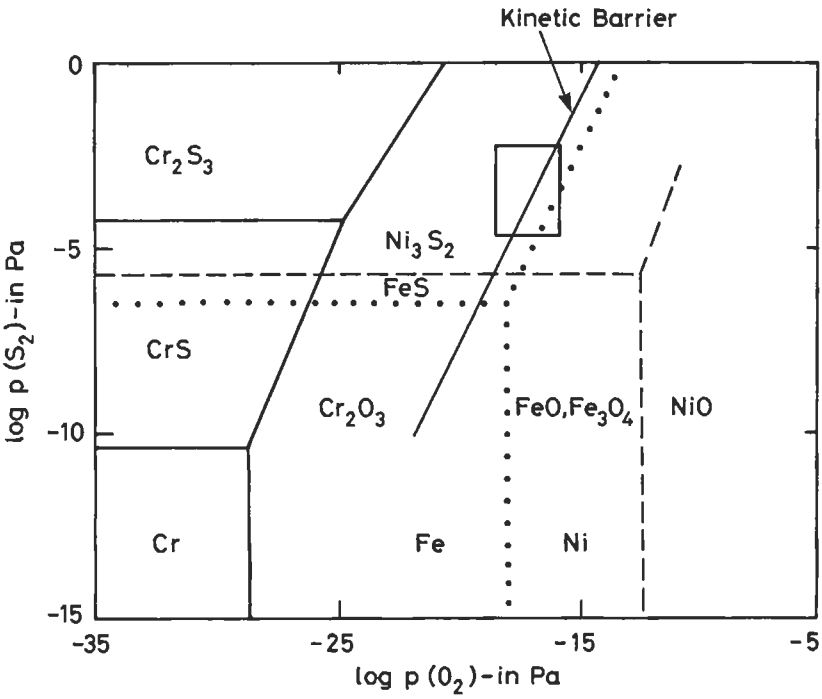


Fig. 15. Phase stability diagrams for the Cr–O–S, Fe–O–S and Ni–O–S systems at 650°C, indicating the range of compositions found in coal gasification atmospheres and the kinetic barrier for exclusive formation of chromia observed for AISI 310 exposed in these conditions.

to equal 10^{15} Pa. In many ways this is a somewhat unrealistic calculation since in most cases there may not be complete access to the gas at the metal/oxide interface, and only a reduced concentration of SO_2 may be present. It is useful to consider the minimum concentration of SO_2 at that interface that is necessary for sulphide formation to occur. This can be shown to be defined by:

$$P_{(SO_2(min))} = \exp \Delta G^{\circ}_{MS} + 2\Delta G^{\circ}_{MO} - \Delta G^{\circ}_{SO_2}/RT \tag{12}$$

Figure 16 is a schematic phase stability diagram in which three possible conditions are illustrated for penetration of the oxide by either the bulk diffusion or the microcrack mechanism, and the limiting sulphur dioxide pressure is marked by a dashed line. In region 1 the $p(SO_2)$ is lower than the minimum value for sulphide formation so that sulphide cannot form by either mechanism. In region 2 transport of molecular sulphur dioxide through the scale would result in sulphide formation for $p(SO_2)$ greater than the minimum, but not by bulk diffusion processes, whereas in region 3 bulk diffusion of ions would result in sulphide formation. Table 3 shows values for $p(SO_2)$(min) for various oxides as a function of temperature, where it can be seen that the presence of highly stable oxides (large negative ΔG°) results in very low values of $p(SO_2)$(min).

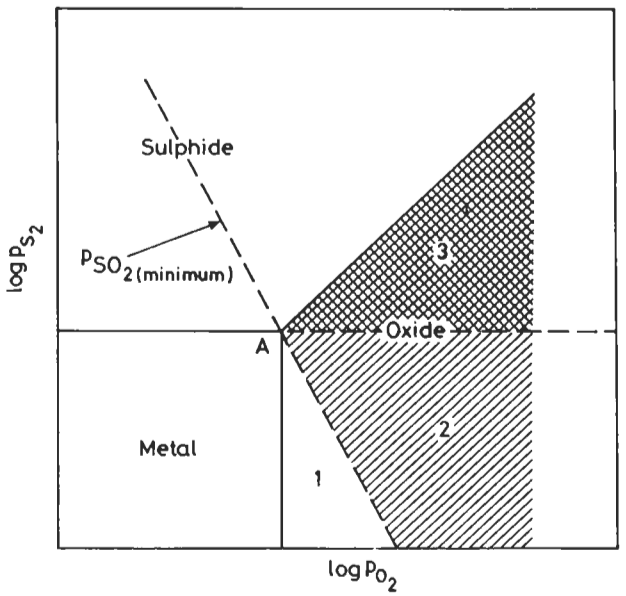


Fig. 16. Schematic thermodynamic phase stability diagram for M–O–S systems, indicating the minimum $p(SO_2)$ for sulphide formation (see text for details).

It is important to realise that if the sulphidising gas does not contain oxygen (if it is H_2S for example), then no mechanism exists for increasing the $p(S_2)$ at the oxide/alloy interface, so that this rather dangerous condition which promotes rapid sulphidation would not apply.

5. *Hot-salt corrosion*

Hot-salt corrosion is a form of high-temperature attack on metals and alloys that occurs in the presence of a molten deposit; metalwork in power station boilers, waste incinerators, gas turbines and diesel engines is susceptible to this degradation process. Most work in this area has been related to the gas turbine but the underlying principles can be readily extended to other spheres. We will consider the environment, the phenomenology of the corrosion process, the mechanisms proposed to account for the corrosion behaviour and, finally, briefly consider the implications of new types of fuel for the gas turbine.

5.1. The environment

Molten deposits containing largely Na_2SO_4 and/or V_2O_5 form on turbine blades and nozzle guide vanes in the hot section of the engine. The deposits form as a consequence of combustion of sulphur- and vanadium-containing fuels which may also contain alkali

References: p. 1357.

Table 3
Minimum values of $p(SO_2)$ in the atmosphere for sulphide formation at the metal/scale interface by molecular transport of SO_2 through the scale.

Temperature, °C	$p(SO_2)$ (Pa)			
	Ni	Co	Fe	Cr
400	1.6×10^{-7}	1.0×10^{-8}	2.0×10^{-14}	1.3×10^{-33}
600	2.5×10^{-2}	6.3×10^{-4}	1.6×10^{-8}	6.3×10^{-23}
800	7.9×10^{1}	6.3×10^{-1}	4.0×10^{-5}	3.2×10^{-16}
1000	6.3×10^{3}	7.9×10^{-1}	7.9×10^{-3}	4.0×10^{-11}
1200	–	–	4.0×10^{-1}	2.5×10^{-8}

metal impurities which are usually sodium-rich, either as NaCl or sodium naphthenate. Alternatively, the alkali metal may be ingested as an impurity in the air used for combustion, and again NaCl (seawater) is generally the main impurity, particularly for the marine gas turbine. Since the air:fuel ratio is about 50:1, it is clear that air quality is very important in determining the quantity of sodium sulphate formed in the combustion gases. The following equation indicates the process by which sodium sulphate forms:

$$2NaCl + H_2O + SO_2 + \tfrac{1}{2}O_2 + Na_2SO_4 + 2HCl \tag{13}$$

The combustion gases containing these impurities are in the temperature range 1100–1400°C so that these compounds are usually in the vapour phase. However, when the combustion gas reaches the turbine blades and nozzle guide vanes where metal temperatures may be about 850°C, condensation will occur if the vapour pressure of Na_2SO_4 is greater than the equilibrium value for that temperature, i.e., if the dewpoint is reached. It is generally accepted that there will be severe corrosion whilst the deposit is molten. If the temperature is lowered so that the deposit solidifies, the rates of attack that are observed are greatly reduced.

A problem of particular importance in gas turbines is the rate at which the sodium sulphate-forming reaction proceeds. It should be realised that the gas velocity in a turbine is about 300 m s^{-1} and the distance between the combustion chamber and the blade can be as little as 0.3 m or for larger industrial engines 2–3 m, so that the sulphate forming reaction must be complete in about 1–10 ms. In the case of NaCl particles in the gas stream, clearly the size of the particle will affect the conversion process. It has been calculated, for example. that at 1300°C and 10 atmospheres (typical flame conditions) a 10 μm NaCl particle would only be 10% converted to sulphate in 5 ms (McCreath [1983]). In normal turbine operation particle sizes would be 10 μm or less, and therefore when NaCl is the main contaminant some unreacted NaCl particles would be expected to hit the blades before evaporation or conversion processes are complete.

Typically then, the environment in which gas turbine blades find themselves is one in which molten Na_2SO_4 and/or V_2O_5 can deposit from the vapour, and also, depending upon the operating conditions of the turbine, there is the possibility of impaction by NaCl (seawater) particles; in addition NaCl vapour and HCl gas may also be present.

5.2. Phenomenology of hot-salt corrosion

Two types of attack can be observed. The more usual form occurs in the temperature range 800–950°C, known as *Type I hot-corrosion* and is characterised by the formation of a thick porous outer oxide layer, an intermediate layer of internal oxide particles mixed with depleted alloy and an inner region containing internal sulphide particles which leads the attack; fig. 17 is a micrograph illustrating this type of attack. It is generally observed that a maximum in rate of attack occurs at about 850–900°C; at higher temperatures, deposition of molten salts is reduced since the vapour pressure is insufficient for condensation.

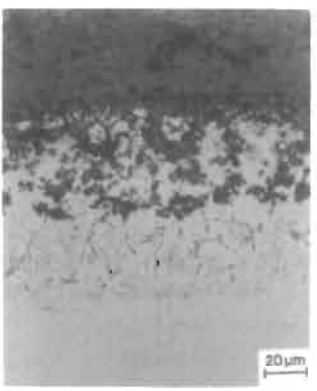
The other form of attack is known as *low-temperature corrosion* or *Type II* and is usually observed in marine gas turbines in the temperature range 650–750°C. In this type of attack large oxide and sulphide-filled pits form, but there are no internal sulphide particles, and fig. 18 is a micrograph illustrating this type of attack. Owing to the large temperature gradient that exists along the length of a turbine blade, it is possible for both types of hot-salt corrosion to occur on one component.

5.3. Mechanism of attack

We will consider in detail the more usual form of Na_2SO_4-induced hot-salt corrosion, Type I, and the principles applied will provide a basis for interpreting the special case of low temperature hot-corrosion or Type II. A brief description of vanadate-induced attack will also be given.

5.3.1. Na_2SO_4-induced attack

No single mechanism can be proposed to explain all observations but there are three main theories, viz., a) sulphidation, b) fluxing theory and c) pseudo-scale formation, and we will briefly consider the relevance of each.



a

Fig. 17. Optical micrograph of Nimonic 105 blade after service in a gas turbine, showing conventional hot-salt corrosion attack, Type I hot-corrosion.

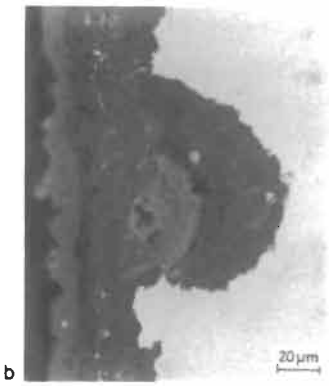


Fig. 18. Optical micrograph of IN738LC blade after service in a gas turbine, showing low temperature pitting attack, Type II hot-corrosion.

a) Sulphidation

One of the first theories to explain Na_2SO_4 induced corrosion was proposed by SIMONS *et al.* [1955] who suggested that sulphidation of the alloy occurred, as indicated by the following reactions:

Triggering phase:	$Na_2SO_4 + 3R \rightarrow Na_2O + 3RO + S$	(14)
	$M + S \rightarrow MS$	(15)
Autocatalytic phase:	$Na_2SO_4 + 3MS \rightarrow 4S + 3MO + Na_2O$	(16)
	$4M + 4S \rightarrow 4MS$	(17)
Enhanced oxidation phase:	$2M + 2MS \rightarrow 2M.MS$	(18)
	$2M.\,MS + O_2 \rightarrow 2MO + 2MS$	(19)

The "R" in the triggering phase is a reducing agent and this is usually the alloy, but pyrolytic carbon resulting from poor combustion of the fuel may also act as a reducing agent.

The sulphidation reaction proceeds in two phases, a triggering phase followed either by the autocatalytic phase or the enhanced oxidation phase dependent upon the thickness of the melt present. With thick Na_2SO_4 layers, where there is limited oxygen access, the autocatalytic reaction is observed, but with thin layers we find the enhanced oxidation phase where the eutectic M.MS phase formed is oxidised to MO and more sulphide. The enhanced oxidation phase should be self-sustaining once the initial sulphide has formed with oxidation of the sulphide releasing sulphur for further sulphide formation. Pre-sulphidation of many alloys often leads to continuous sulphide formation by this mechanism when the material is subsequently oxidized in air or oxygen.

(b) Fluxing theory

The fluxing model was initially proposed by BORNSTEIN and DECRESCENTE [1969], and developed by GOEBEL and PETTIT [1970], and can best be illustrated by considering the decomposition of Na_2SO_4 into its acidic and basic components.

$$Na_2SO_4 \rightarrow Na_2O \text{ (base)} + SO_3 \text{ (acid)} \quad (20)$$

With basic fluxing of nickel oxide the following reaction then takes place:

$$NiO + Na_2O \rightarrow Na_2NiO_2 \quad (21)$$

and for acidic fluxing:

$$NiO + SO_3 \rightarrow NiSO_4 \quad (22)$$

The processes involved in basic fluxing are schematically represented for the case of nickel in fig. 19. It can be seen that basic fluxing is favoured by processes that increase the oxide ion activity in the Na_2SO_4 melt. Figure 20 is a phase stability diagram for the Al–O–S system which illustrates acidic and basic fluxing of alumina. The formation of oxide under the molten salt would result in a decrease in the $p(O_2)$ at the melt oxide/ interface and a consequent increase in the $p(S_2)$ so that there is now sufficient driving force for diffusion of sulphur into the oxide, thereby decreasing the $p(SO_3)$ and promoting basic fluxing. Acidic fluxing will be favoured by low temperatures because of the increasing stability of SO_3, or by reactions of alloying constituents with Na_2SO_4 that deplete the oxide ion concentration by, for example, complex oxide formation.

$$MoO_3 + Na_2O \rightarrow Na_2MoO_4 \quad (23)$$

RAPP and GOTO [1981], pointed out that in order to sustain the dissolution reaction, a solubility gradient must be present in the molten salt layer. Thus oxide can dissolve at the oxide/melt interface, migrate down a concentration gradient to a site of lower solubility where precipitation occurs, and this is believed to be the origin of the external porous oxide layer found on samples exposed to this form of attack. Type I hot-corrosion is generally believed to be the result of basic fluxing or of direct attack on the substrate by the melt through a cracked oxide layer, and Type II is the result of acidic fluxing by

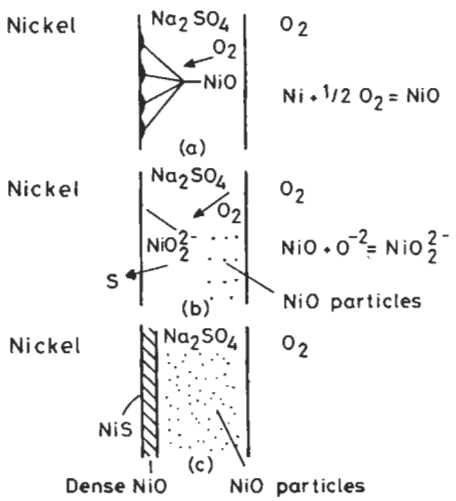


Fig. 19. Schematic illustration of the basic fluxing reactions of Na_2SO_4 with pure Ni.

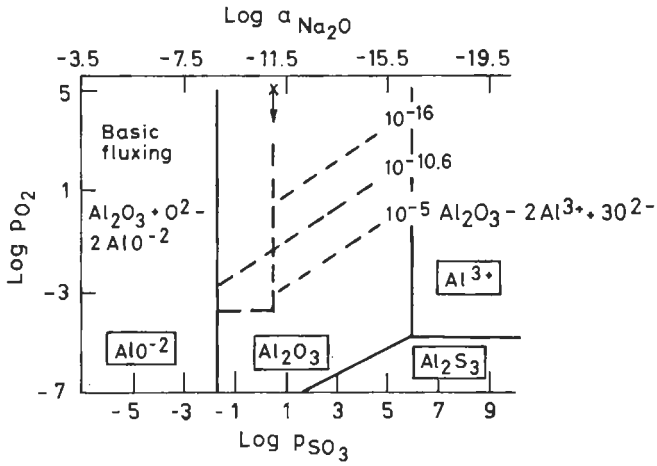


Fig. 20. Phase stability diagram of the Al–O–S system at 1000°C: ..... indicates S_2 isobars, → compositional charges in the Na_2SO_4 melt at the melt oxide interface with increasing time.

a low-melting eutectic of, for example, 38% $NiSO_4.Na_2SO_4$ (m.p. 681°C) which are stabilised by a high p_{SO_3} (LUTHRA and SHORES [1980]).

The solubility of oxides relevant to the hot-salt corrosion of superalloys has been measured by Rapp and co-workers, and fig. 21 is taken from a review of their work, (RAPP [1987]). It is noteworthy that there is a difference of about six orders of magnitude in basicity (i.e., a_{Na_2O}) between the solubility minima for the most basic oxides

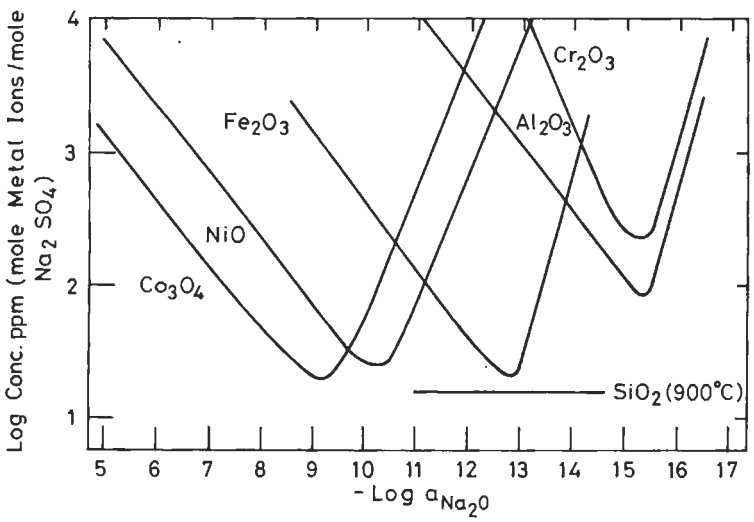


Fig. 21. Oxide solubility in Na_2SO_4 at 927°C as a function of oxide ion activity (RAPP [1987]).

and the protective alumina, chromia and silica which are all acidic oxides. Chromia scales are known to be more resistant than alumina to acid fluxing. It is suggested that chromia scales are able to support a buffering action and resist large changes in the basicity of the melt, and that the solubility minimum corresponds to the conditions in the gas turbine. The solubility data do not immediately suggest the cause of the superiority of chromia relative to alumina, but RAPP [1987], suggests that this might be the result of the faster growth of chromia. Another possibility is that the alumina scale is more vulnerable to mechanical damage, which would permit direct attack on the alloy.

(c) Pseudo-scale

In this theory (REISING and KRAUSE [1974]) the Na_2SO_4 acts merely to transport oxygen to the alloy surface and the Ni^{2+} ions produced by the oxidation process are mobile in the liquid medium and move away to the liquid/gas interface where NiO particles form, producing a porous scale. Rapid oxidation ceases when the porous scale has absorbed all the Na_2SO_4 removing it from contact with the alloy surface and hence a protective oxide layer can be re-established. Experimental observations indicate that the $NiO : Na_2SO_4$ ratio is usually constant and about 50 : 1, whereas the fluxing theory which has the following overall reaction:

$$2Na_2SO_4 + 9Ni \rightarrow 2Na_2O + 6NiO + Ni_3S_2 \qquad (24)$$

suggests that the $NiO : Na_2SO_4$ ratio should be 3 : 1. However, the effects of alloy chemistry and corrosion are probably better explained by the fluxing theory than by the "pseudo-scale" model. The fluxing model is now largely used in interpreting effects of hot-salt corrosion and has gained wide acceptance.

5.3.2. Vanadate-induced attack

Residual fuels contain relatively large amounts of sodium and especially vanadium. Figure 22 shows the phase diagram for the Na–V–O system which indicates formation

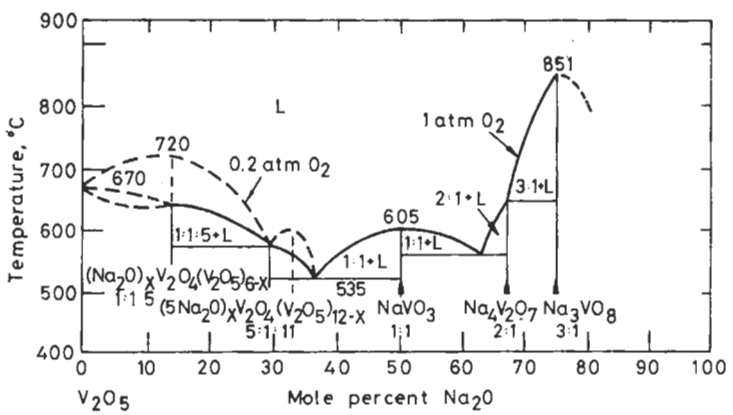
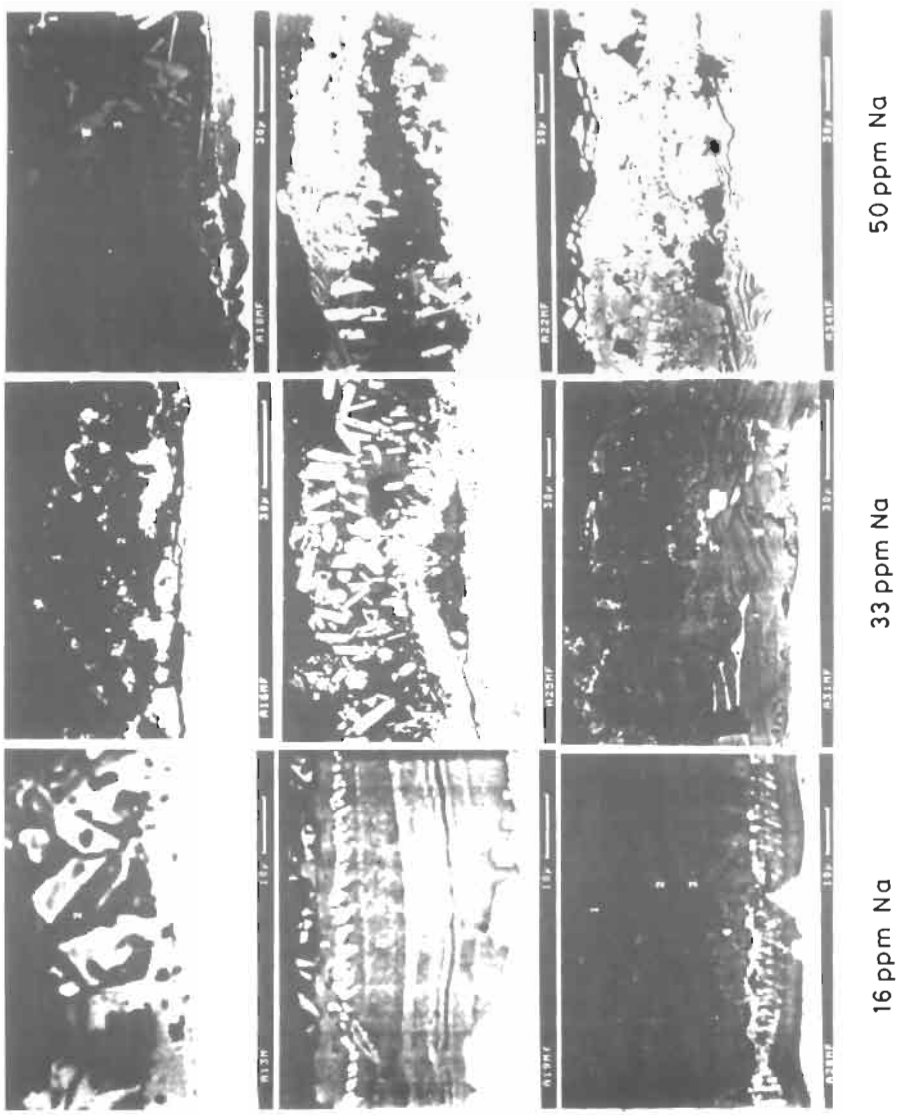


Fig. 22. The quasi-binary Na_2O–V_2O_5 diagram (KERBY and WILSON [1973]).

of complex sodium vanadates with melting points as low as 535°C, although experimentally even lower temperatures have been found. The presence of these compounds in the combustion gas effectively limits the operation of diesel engines using these fuels to exhaust valve temperatures of about 450–500°C. Higher temperatures would promote



(b)

Fig. 23. (a) Backscattered electron images of Stellite 6 samples after exposure in a burner rig at 600°C, as a function of sodium and vanadium concentrations in the fuel. (b) Backscattered electron images of Stellite 6 coated exhaust valve, after exposure in a marine slow-speed diesel engine. Compare with the micrograph showing attack with 150 ppm V and 16 ppm Na in (a).

acid fluxing by the molten vanadates. As with the previous case of sulphate corrosion, alloys and coatings with high chromium contents are used to resist attack when molten deposits form.

A study of the effects of varying sodium and vanadium ratios on the corrosion of diesel exhaust valve alloys (SAUNDERS *et al.* [1990]) showed that corrosion product morphology found in service could be simulated in the laboratory (fig. 23), and this was used to predict exhaust valve lives that agreed well with service experience.

5.4. Coal-fired gas turbines

In conventionally fired gas turbines, levels of 0.6 ppm Na in the fuel result in very severe corrosion of Nimonic 105, a material of moderate corrosion resistance. As mentioned previously with an air-to-fuel ratio of about 50 : 1, air purity is particularly important. If, however, the gas turbine were fired with a gas derived from coal where the calorific value of the gas is such that the air-to-fuel ratio would be about 3 : 1 then *fuel* quality is considerably more important for this application. Generally, the sulphur level in coal-derived fuels will be somewhat lower than in currently used oils. However, sodium and chlorine levels would be at least an order of magnitude greater than in conventionally fired gas turbines and also potassium will be present in significant quantities. The effect of potassium is not clearly defined but it might be expected that the lower melting point of sodium and potassium sulphate mixtures would encourage hot-salt corrosion at even lower temperatures. The presence of increased HCl, formed from combustion of the chlorine-containing compounds, would cause NaCl to form as a stable phase on the blades and, even in coals with relatively low chlorine contents, a 1% NaCl

solution in Na_2SO_4 may result which would be expected to increase corrosion rates. Clearly this is a complex situation and much work is required to understand the chemistry of these Na–K–Cl–S–O molten salts.

In addition to the greatly increased contaminant levels in the fuel, coal-fired gas turbines will be subjected to erosion from ash particles which despite filtration may remain entrained in the combustion gas. Some erosion tests have shown that coal ash particles 2–20 μm in diameter are capable of removing large amounts of metal in very short periods. Some testing has been carried out, however, and although severe corrosion has been observed in some cases, the catastrophic rates that might have been predicted have not been observed. The blades quickly became coated with a deposit which was largely composed of coal ash and it is possible that the "pseudo-scale" theory may explain the somewhat better corrosion behaviour. The coal ash could be acting as a porous sponge, soaking up the sulphate melt and removing it from contact with the underlying alloy.

6. *Test and measurement methods*

The performance of materials at high temperatures is dominated by their resistance to mechanical deformation and attack by the environment. Thus, in attempting to select materials for service in these conditions it is important to have reliable means of predicting their behaviour, and whilst some information can be obtained by using the experience gained in the operation of existing plant, introduction of new materials or of different operating conditions requires data to be produced from laboratory testing procedures. High temperature corrosion can involve attack in the gas phase or the combined effects of hot gases and molten or solid deposits. As might be expected, if solid deposits can form, an increase in gas stream velocities would result in erosion rather than deposition but this is beyond the scope of this chapter. The mechanical behaviour of oxide layers is an extremely important factor determining the overall performance in corrosion and erosion processes, so that the available methods of ascertaining the mechanical properties of thin layers will also be considered.

6.1. Monitoring oxidation processes

The methods used to measure the rates of oxidation, i.e., processes forming only oxide layers, are considered in this section but, as will be evident, many of the techniques are applicable to the other reactions and processes discussed in this chapter. In the first place, consider the simplest process in which a test is carried out isothermally.

6.1.1. Isothermal testing

The purpose of tests of this type is to determine the rate at which the specimen is converted to oxide, and the most commonly used procedure is the gravimetric method in which the change in mass of the sample is recorded. Usually a mass gain corresponding to uptake of oxygen is observed, but if oxide spalling occurs this mass gain is reduced

by the amount of metallic oxide lost, and in extreme cases net mass losses can be recorded. The simplest form of the experiment is to place the sample directly into the furnace, withdrawing it periodically and measuring the change in mass. This experiment clearly imposes temperature cycles, which are discussed later, but in some cases, particularly for thin adherent oxide layers, similar results are obtained with experiments carried out continuously. Figure 24 shows schematically a typical arrangement for conducting such experiments with continuous weighing of the sample using a microbalance.

Other methods involving measurement of changes in pressure or volume, or of the optical properties of the surface, such as ellipsometry are also used, but a complete description of these and other methods is unfortunately beyond the scope of the present chapter. The reader is referred to KUBASCHEWSKI and HOPKINS [1967] for full details.

Oxidation rates also have been determined by measurement of the natural frequency of a freely suspended specimen by BRUCE and HANCOCK [1969]; HANCOCK [1970]; HURST *et al.* [1975], and as will be shown later in the section on mechanical stability of oxides, cracking and spalling of the scale can also be detected. The characteristic frequency of the first mode of vibration, f, is related to the dimension and the elastic modulus, E, of the composite rod as follows:

$$f^2 = \frac{(4.73)^4}{16\pi l^3 (m_o + \Delta m)} \left[E_m c^4 + E_o \left(b^4 - c^4 \right) \right] \tag{25}$$

where E_m and E_o are the elastic moduli of the metal and oxide, m_o is initial mass of the rod, Δm is the mass gain, c is the radius of metal and b is the radius of the composite.

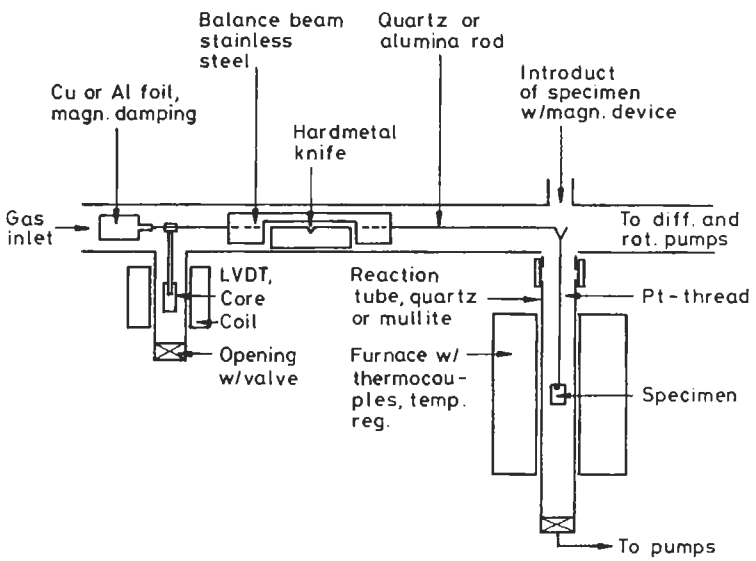


Fig. 24. Schematic drawing of an automatic recording thermobalance.

Thus, by solving the above equation, mass changes can be derived, together with the elastic properties of this composite, as follows:

$$f^o - f_o^2 = \Delta\left[\frac{2\beta M_m}{M_o - M_m} - \frac{M_o - M_m}{M_o + M_o}\right] \tag{26}$$

where: Δ is defined by $m = m_0(1+\Delta)$
f_o is frequency of the unoxidised rod
M_m, ρ_m, E_m are the atomic weight, density and elastic modulus of the metal
M_o, ρ_o, E_o refer to an oxide
and $\beta = (E_o \rho_m M_o)/(E_m \rho_o M_o)$

6.1.2. Cyclic oxidation

The superimposition of thermal cycles in an oxidation experiment, *cyclic oxidation*, is used to encourage scale failure due to spallation (LOBB [1984]; BARRETT and LOWELL [1975]; BARRETT *et al.* [1978]). Hence cyclic oxidation tests are used to monitor both scale adherence and the ability of the scale to successfully repair after repetitive scale failure.

The performance of materials is generally monitored gravimetrically, either continuously (LOBB [1984]; BARRETT and LOWELL [1975]) or in a discontinuous manner (BARRETT and LOWELL [1975]; BARRETT *et al.* [1978]) although other monitoring methods have been used. For example, measurement of the resonant frequency of the specimen during thermal cycling has been successfully used to detect (HANCOCK [1970]) the onset of scale failure, before gross spallation, and hence mass loss was observed. Figure 25 illustrates typical data produced by gravimetric tests. Curve (a) represent good performance with little scale failure; curve (b) is intermediate behaviour, that is initially protective but after a limited test duration becomes non-protective, while curve (c) shows poor behaviour with scale spallation occurring from the onset of the test. Curve (b) is characteristic of materials that have a limited reserve of stable scale forming elements, such as chromium, aluminium and silicon, so that after repeated scale failure and repair the morphology and composition of the scale changes, possibly forming less protective spinels, for example. These less protective scales spall more readily and rapid mass losses are recorded (BARRETT *et al.* [1978]).

6.1.3. Mechanistic studies

Various experimental procedures have been devised specifically to obtain information about the mechanism of oxidation and should be included in any discussion of test methods. When considering mechanisms of oxide growth it is important to know the transport path (lattice or grain-boundary diffusion), where oxide growth is occurring, i.e., anion or cation movement, the diffusion coefficients of the more mobile ions and the equilibrium defect concentration. Tracer and marker techniques can be used to determine the nature of mobile species and transport paths and the ROSENBURG [1960] method, using interrupted kinetics experiments, gives information about diffusion coefficients and defect concentrations. The methods most commonly used are marker and tracer and these are discussed below.

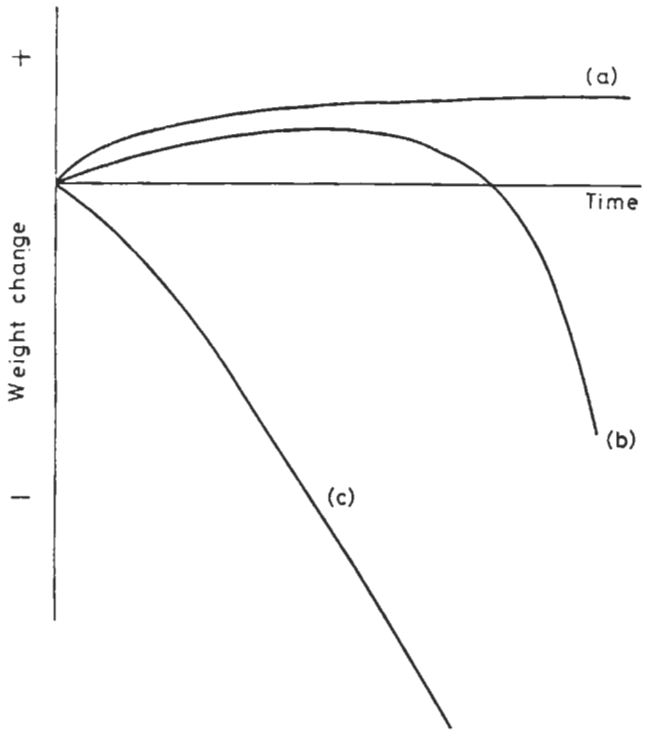


Fig. 25. Schematic diagram of typical mass changes versus time during cyclic oxidation: (a) protective, (b) protective initially, (c) spalling.

i) Marker methods

If inert material, platinum or a stable oxide, is placed on the surface of unoxidised metal, its final position in the oxide scale will give information about the transport processes that have taken place, as is illustrated in fig. 26. If the scale is pore-free the final position of the marker ideally reflects the transport processes that have taken place. However, the presence of pores can result in the marker being displaced. For example, in cases where an outer dense layer forms over an inner porous region, the marker is

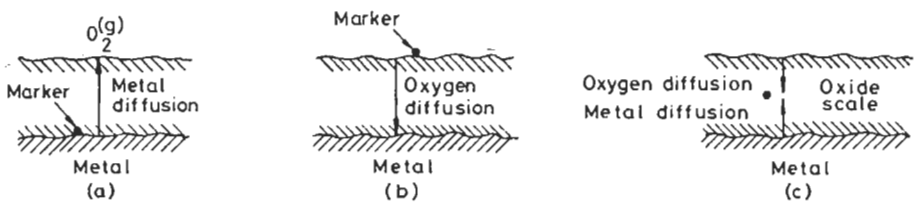


Fig. 26. Ideal location of inert markers after lattice diffusion-controlled growth of oxide scales: (a) metal diffusion predominant; (b) oxygen diffusion predominant; (c) simultaneous metal and oxygen diffusion.

References: p. 1357.

frequently found at the interface between the porous and dense layers even though cation lattice diffusion is faster than anion diffusion. In this case, the most likely mechanism is that molecular oxygen has penetrated the outer layer as the result of cracking of the scale or of the development of microchannels due to build-up of stress, so that the final position of the marker reflects movement of material but not solid-state diffusion processes. It is also suggested that with large markers, the marker itself can partially obstruct the diffusion process and some plastic flow can occur around the marker.

ii) Tracer methods

In this method, a radioactive form of one of the reactants in the oxidation process is used. This can be done by either depositing a thin layer of radioactive metal on the surface of the specimen or by carrying out the experiments with radioactive oxidant at some stage in the experiment; ^{18}O is commonly used in oxidation experiments. ATKINSON *et al.* [1979] have calculated idealised profiles for various limiting cases which are reproduced in fig. 27, where it can be seen that for the case of a scale, MO, growing by vacancy diffusion the profile depends on the effective charge, α, of that vacancy. Where grain-boundary diffusion takes place it has been assumed that there is no exchange of metal between the grain-boundary and the oxide lattice and thus the cations are immobilised. In fact, some exchange does occur and this would have the effect of broadening the profile. In most practical cases, some combination of the limiting cases referred to above is usually found, so that the resultant profile will be an appropriate mixture of the simple cases.

6.2. Mechanical failure of oxide scales

By careful design of a cyclic oxidation experiment the susceptibility of a material to spalling can be assessed. Hence cyclic oxidation tests provide a qualitative method of assessing scale adherence. To quantify scale adherence, and the onset of spalling, controlled cooling experiments are necessary. Continuous monitoring permits the onset of first fracture to be detected such that from a knowledge of the temperature drop the fracture energy for scale spallation can be calculated. This approach has been used by EVANS and LOBB [1984] to evaluate the interfacial fracture energy of chromia-rich scales formed on a 20Cr/25Ni/Nb-stabilised stainless steel after oxidation in CO_2/2%CO at 930°C.

6.2.1. Internal stress measurements

Two methods have been widely used to determine the internal or growth stresses present within oxide scales; these are X-ray stress analysis (e.g., LUTHRA and BRIANT [1986]) and methods based on the deflection of a thin metallic strip oxidising on one surface only.

X-ray stress measurement techniques are usually evaluated at room temperature, although some facilities are available with a high temperature capability (STOUT *et al.* [1989]). Hence the internal stress determinations include both growth stresses and thermally induced stresses during the cooling cycle. The X-ray determination of stresses is based on the measurement of lattice strains in a set of lattice planes of known

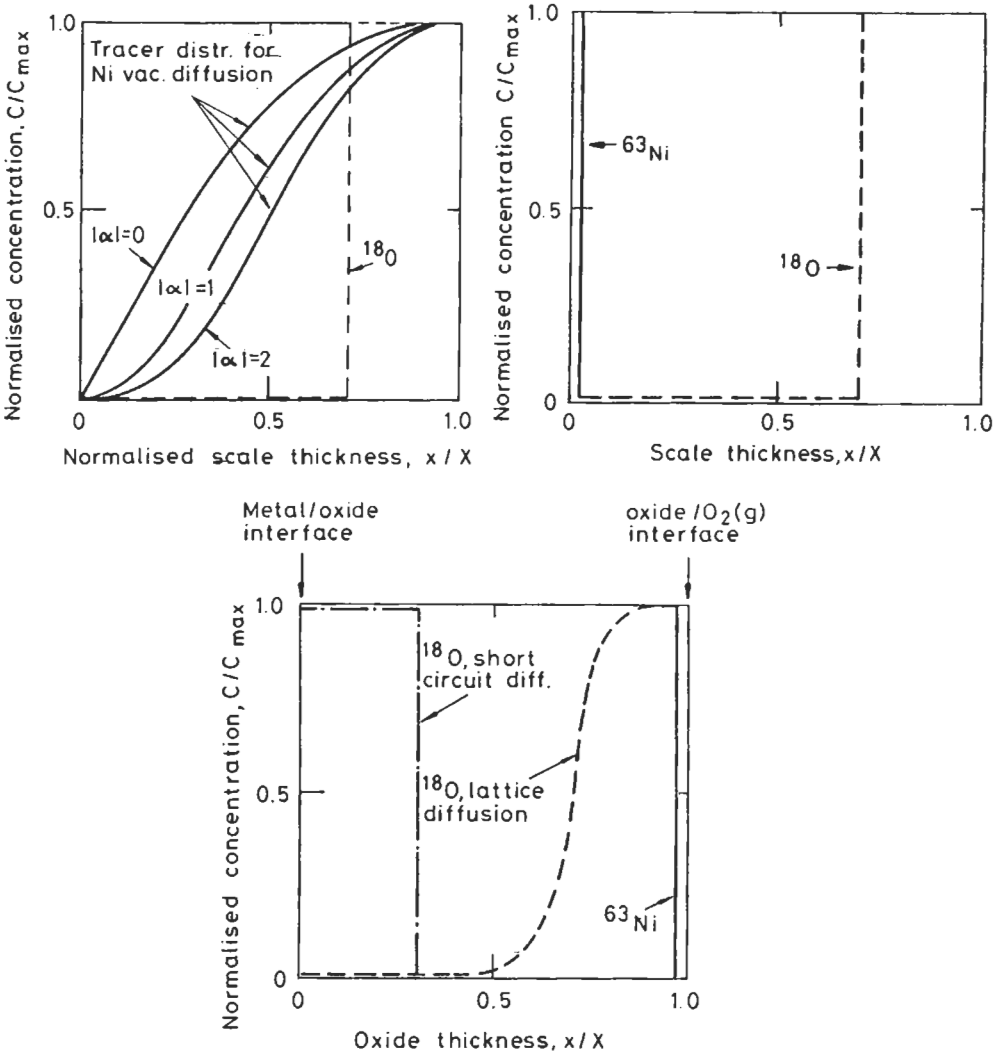


Fig. 27. (a) Metal tracer in a scale of MO growing by metal vacancy diffusion only. (b) Distribution of M and ^{18}O tracers in a scale growing by short circuit diffusion. (c) Distribution of M and ^{18}O tracers in a scale growing by short circuit diffusion and lattice diffusion (ATKINSON *et al.* [1979]).

orientation and then correlating these lattice strains with the strain calculated from elastic theory (LUTHRA and BRIANT [1986]). This approach is known as the $\sin^2\psi$ method.

The determination of internal and growth stresses by direct deflection measurements on a thin metal beam at temperature (ROY and BURGESS [1970]) has been widely used by Huntz and co-workers (DELAUNAY *et al.* [1980]; HUNTZ [1988]). Figure 28 schematically illustrates this method of measurement. The thin metallic strip under test is only

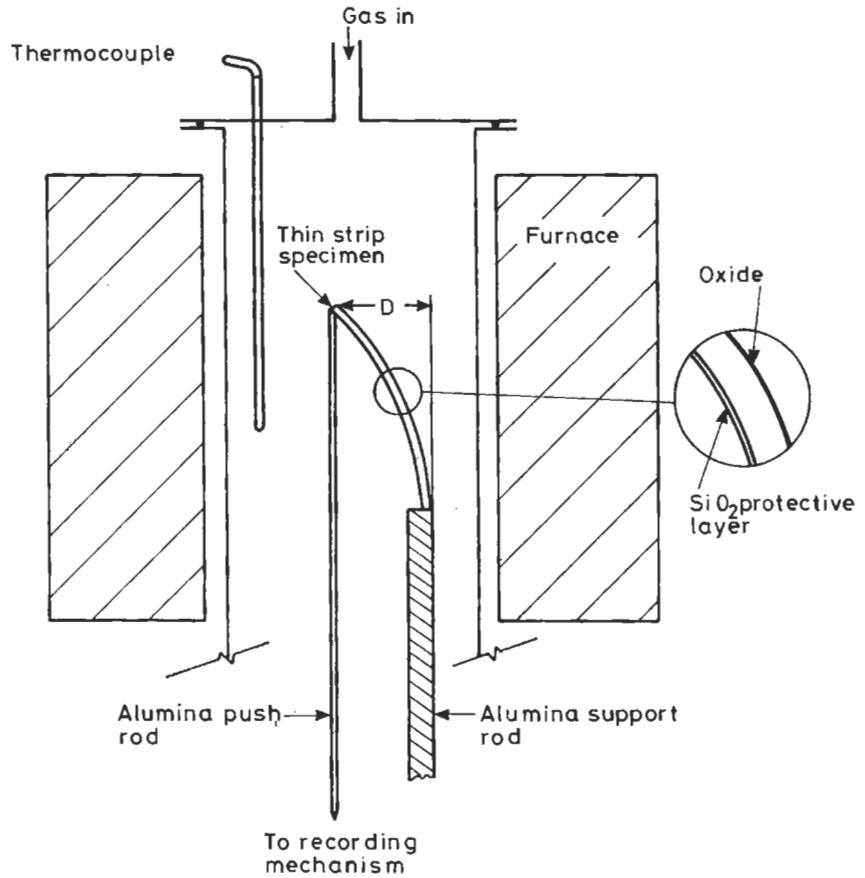


Fig. 28. Schematic diagram of the deflection study apparatus.

permitted to oxidise on one side by coating the second, usually with silica. Oxidation causes the strip to bend with the degree of deflection, a measure of the stress generated within the oxide as defined by:

$$\sigma_o = \frac{E_m X_m^2 D}{3L^2 X_o} \tag{27}$$

where D is the beam deflection, L is the beam length, X_m and X_o are thickness of the metal and oxide, respectively and E_m is the elastic modulus of the metal. This measured stress is only a part of the total stress due to oxidation, as a residual stress component remains after beam deflection due to the tensile elastic strains generated within the metal beam. This method does not take account of plastic deformation of the substrate by creep, for example, and recently a detailed analysis has been carried out EVANS [1995] in which an analytical solution was developed from a relationship derived from the

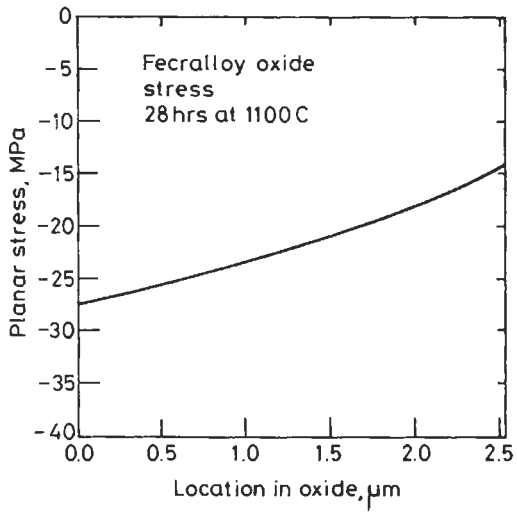


Fig. 29. Stress distribution in an alumina scale formed on Fecralloy at 1100°C (EVANS [1995]).

balance of forces and the balance of momentum. Figure 29 is an example of the results obtained using this method for the case of a Fecralloy steel oxidised at 1100°C.

6.2.2. Detection of scale failure

Three methods have been used to date to detect scale cracking both during thermal cycling and under conditions of applied load. As discussed in the section on cyclic oxidation measurements, controlled cooling experiments can reveal the critical temperature drop necessary to cause through-scale cracking (increased oxidation rates are observed) or scale spallation (mass losses are recorded) (LOBB [1984]; EVANS and LOBB [1984]; BAXTER *et al.* [1984]; ROLLS and NEMATOLLAHI [1983]). However, this method is restricted to metal/oxide systems where mass changes induced by scale cracking can be readily detected. The vibration technique (as described in section 6.1.1) can also be used to detect the onset of scale cracking. It is very sensitive to crack formation within the oxide which results in a reduction of the stiffness of the composite system and hence a discontinuity in resonant frequency is observed (see fig. 30) when the oxide scale cracks. Continued cracking would result in a continual decrease in the recorded resonant frequency. Should oxide repair occur then a frequency transition is observed with the frequency recorded after repair approaching that observed before the cracking event.

More recently, acoustic emission has replaced the vibration technique as a route to detect *in situ* early scale failure, and its use in the determination of oxide fracture was recently reviewed by SAUNDERS *et al.* [1994]. This technique is not limited to particular specimen geometries, has similar sensitivities to scale cracking and can be monitored automatically. Acoustic emission has found wide usage in monitoring cracking during thermal cycling (KHANNA [1985]; CHRISTL *et al.* [1987]), superimposed stress (SCHÜTZE [1985]) and during isothermal oxide growth (PRICE and BENNETT [1989]). Figure 31 shows

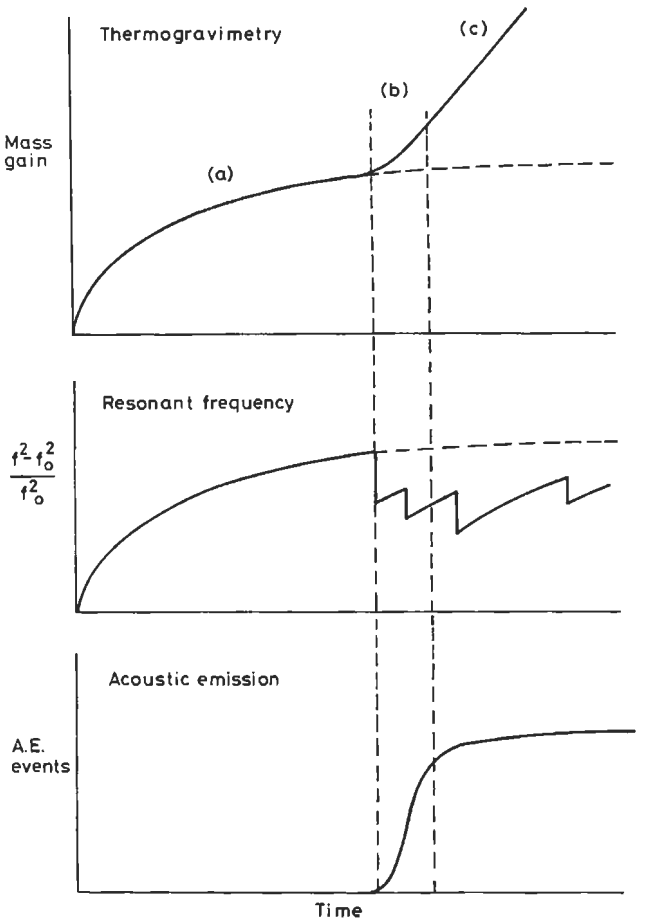


Fig. 30. Schematic response to breakaway oxidation, monitored using three methods.

the typical coupling arrangement of an acoustic transducer to a specimen through the use of an acoustic waveguide. Figure 31 specifically relates to monitoring acoustic events during the cracking of oxides under applied tensile loads (SCHÜTZE [1985]), however, similar geometries can be used to monitor scale failure under cyclic conditions or during isothermal oxidation by combining acoustic emission with thermogravimetric studies (BENNETT *et al.* [1989]). Acoustic microscopy in which an acoustic signal is focused onto a sample using water as a coupling medium can also be used to detect scale cracking.

6.2.3. Measurement of the macro defects (cracks, voids and pores) present in an oxide scale

As discussed earlier in this section, scale failure results when the *in situ* stress exceeds some critical value necessary to propagate a defect. Hence from a knowledge of the

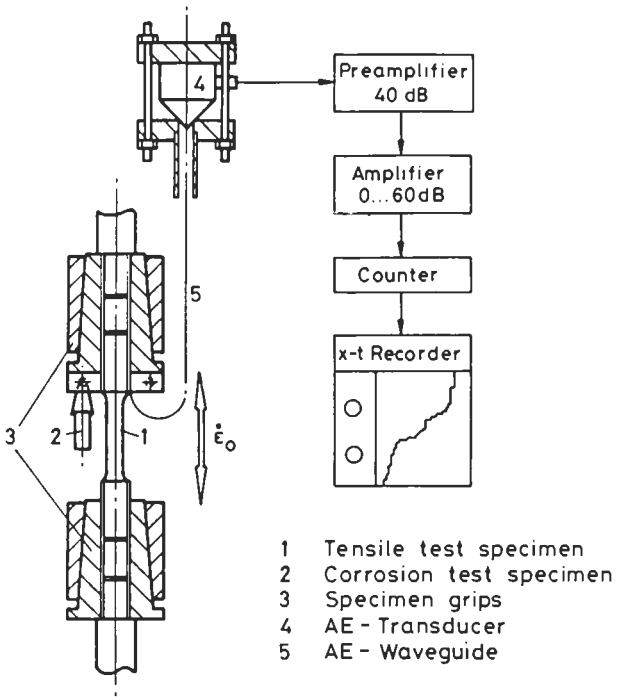


Fig. 31. Schematic of an acoustic emission test arrangement.

stress to cause fracture, the growth stress and the defect distribution in the scale, a measure of the scale fracture toughness and scale fracture energy can be determined. HANCOCK and NICHOLLS [1988] proposed a method for assessing the composite defect size in an oxide containing multiple voids and pores. The analysis follows guidelines proposed for the assessment of defects in welds, and requires that each defect is idealised and compared with its neighbours to assess if any interaction occurs. Interacting defects are considered as a single defect and the assessment procedure continues, until a single composite equivalent defect results. Using this defect size, values of fracture toughness can be evaluated. Evidently, this process requires detailed metallographic analysis. This approach, however, has been applied (NICHOLLS and HANCOCK [1989]) to a number of scales containing active elements and has clearly demonstrated that a major role of active element additions is to reduce the macro defects present within the oxide scales with consequent benefit for scale adherence and spalling resistance.

6.3. Mixed oxidant tests

Many industrial applications involve exposure of high-temperature components to complex gas mixtures, and it is necessary to devise relevant laboratory test procedures. In designing the test it is important to consider whether the atmosphere being simulated

is equilibrated or unequilibrated, and, also in the context of the laboratory environment, whether adequate safety precautions exist since frequently the gas mixtures to be used contain toxic or combustible components. In this section consideration will be given to factors controlling the composition of the atmosphere and the experimental arrangements required to monitor gas composition, including any modification to experimental arrangements necessary to allow safe operation and introduction of samples.

6.3.1. Control of gas composition

Normally, equilibrated gas mixtures are required, and therefore, it is common practice to assume that the main oxidising (in the chemical sense) species controls reaction rates, so that the partial pressures of oxygen (p_{O_2}), sulphur (p_{S_2}), halogenic gases (for example, p_{Cl_2}) and the activity of carbon (a_c) are controlled by establishing relevant gas equilibria. It is relatively straightforward to carry out calculations to determine the partial pressure in these simple binary gas mixtures, but in multi-component atmospheres calculation of the equilibrated gas composition requires the use of iterative routines such as the MTDATA (DAVIES and BARRY [1989]). The composition of the gas used in the test will, of course, control the type of scale that forms (see section 4.2).

The choice of gas mixture is usually determined by a direct simulation (for industrial atmospheres or environments containing high partial pressures of reactive species) or by using a gas with an equivalent partial pressure of active species for low p_{O_2}, p_{S_2}, etc. KOFSTAD [1989] has questioned this latter approach, since he considers that the specific components of a gas mixture can individually control scale defect concentrations or microstructure thereby altering rates of attack. For example, it is known that H_2/H_2O or CO/CO_2 mixtures of equivalent $p(O_2)$ yield different rates of attack. Representative testing, therefore, should be carried out with exact simulations of the relevant gas mixture where possible.

In flowing gas streams, which are usually used to avoid depletion effects of active species (often present in only small quantities), it may be necessary to preheat the gas by passing it over a heat exchange system or in a serpentine path through the hot-zone of the furnace before contacting the samples. Some gas mixtures are slow to equilibrate so that use of catalysts is necessary, and a good example of this is the SO_2/SO_3 equilibrium which is readily established over a platinum catalyst (LUTHRA and WORRELL [1976]). In

Table 4
Gas equilibria

Gas equilibrium	Partial pressures and activities
H_2/H_2O	$p(O_2)$
CO/CO_2	$p(O_2)$ and a_c
H_2/H_2S	$p(S_2)$
SO_2/O_2	$p(S_2)$ and $p(O_2)$
CH_4/H_2	a_c
H_2/HCl	$p(Cl_2)$

other cases the presence of the metal or oxide on the specimen surface is an effective catalyst, for example, iron, nickel and iron oxide. However, this catalytic activity may be poisoned by components of the gas mixture such as hydrogen sulphide (GRABKE and WOLF [1987]). The nature of a hot corrosion experiment requires that the gas mixture flows in a thermal gradient, so that care is required to ensure that thermal diffusion effects do not cause unmixing of the gas by differential diffusion of the components in the temperature gradient (JOST [1952]). Generally, the heavier components will tend to concentrate in the cooler regions of the furnace for high convection velocities and vice versa for low velocities.

6.3.2. Experimental procedures

A paramount consideration is safety. Acceptable toxicity levels are given in the Registry of Toxic Effects of Chemical Substances, and the laboratory should be regularly checked to ensure that safe levels exist in the ambient atmosphere.

Monitoring the gas composition is, of course, desirable and gas chromatography and mass spectrometry are commonly used to determine the composition at ambient temperature, but it is not usually possible to measure gas compositions at the test temperature by these methods. Use of solid electrolyte zirconia probes to measure $p(O_2)$ of the gas mixture within the hot-zone is being used increasingly (KIUKKOLA and WAGNER [1957]). This method relies upon the establishment of the following electrolytic cell:

$$Pt/O_2\ (air)/ZrO_2/Pt\ (O_2\ unknown) \tag{28}$$

Here the e.m.f. generated is directly proportional to the difference between the reference (for example, air) and unknown $p(O_2)$ values. These probes are available commercially or can be relatively easily made in the laboratory. There have been attempts to develop probes for the measurement of sulphur and carbon activities but these are still in their infancy (WORRELL [1975]; TANIGUCHI *et al.* [1988]).

6.4. Attack by molten salts

The case of gas turbine operation will be considered as an example. In designing a suitable test it should be borne in mind that in the turbine the salt deposit is constantly being replenished, and that owing to the action of centrifugal forces the deposit is also being removed to the blade tips and thus is usually quite thin. When a porous oxide layer or deposits form on a blade then the molten salt can also be trapped within pores. This is a complex situation which most researchers acknowledge cannot easily be realistically simulated in laboratory tests, so that various simplified test methods have been devised to allow some information about reactions kinetics and mechanisms to be obtained. SAUNDERS and NICHOLLS [1984] reviewed the various test methods that have been used and have classified these into crucible tests, furnace tests and burner rig tests.

There is general agreement that burner rig tests gives the best simulation of operating experience, and corrosion product morphologies obtained as a result of burner rig testing agree well with those found on blades taken from operating engines. However, furnace tests with careful control of gas and deposit chemistry and deposit replacement can give a close simulation of dynamic corrosion that can be expected in service, at a very

References: p. 1357.

modest cost. Considerable disagreement exists about the quantitative comparison of corrosion rates from different rigs. A major factor in the confusion is that in the past operators of rigs have attempted to compare results taken from rigs where the *concentration* of contaminants was controlled to that expected in the turbine. Since it has been established (SAUNDERS and NICHOLLS [1984]) that deposition rate is the primary factor controlling corrosion rates, only rigs operating at the same velocity as the turbine would be expected to produce comparable rates of attack. It has been pointed out (HANCOCK [1982]; SAUNDERS *et al.* [1982]) that contaminant *flux rate* (CFR) is a better parameter to control deposition and allows low velocity burner rigs to operate at comparable deposition rates to high velocity rigs and gas turbines. Recently, an international round-robin was carried out which established the validity of this approach (SAUNDERS and NICHOLLS [1995]).

6.5. Attack by solid deposits

In cases where solid deposits form it is not only important to ensure that the amount and composition of the deposit corresponds with that found in service, but, because of porosity effects in altering gas composition, the structure of the deposit must also be controlled.

In recent years considerable interest has developed in fluidised-bed combustion of coal, where deposit-induced corrosion can occur. In this case the deposit consists of a mixture of $CaSO_4$ (derived from sulphur capture processes in the bed) and coal ash and it is believed that the active species in the corrosion process is $CaSO_4$. Tests were carried out in which samples were buried in various $CaSO_4$-containing mixtures ($CaSO_4 + C$, $CaSO_4 + CaO$ etc.) with or without control of the gas atmosphere. While at high temperatures these tests generally reproduced corrosion found in service they did not reproduce attack observed at temperatures below about 800°C. This is believed to be due to inadequate simulation of the pore structure of the deposit. Slurries of $CaSO_4$-containing material have been used by NATESAN [1982], while SAUNDERS and SPENCER [1987], attempted to overcome the problem by using powder compaction techniques. In ash deposition conditions typical of those found within a fluidised bed, one cannot discount the rôle of particle impact in modifying deposit morphologies and hence corrosion (STEPHENSON and NICHOLLS [1992]).

7. *Life prediction modelling*

There is considerable interest in the ability to predict when oxide scales fail and their consequence on increased corrosion rates. This section will review "lifing" models that have been developed based on oxide failure, and hot-salt corrosion mechanisms.

7.1. Oxidation models

The development of spallation maps (see fig. 8) and their application to the prediction of scale failure was was discussed in section 2.4.2. From this work it was shown that

scale thickness in conjunction with temperature drop determines when the scale fails. This concept has been extended further to predict the lives of thermal barrier coatings (TBC) on hot-path components. The spallation of the ZrO_2-8wt%Y_2O_3 TBC can be modelled on the basis of the growth of the alumina at the bond coat/TBC interface, with the number of cycles to failure, N_f, determined from a knowledge of the oxide thickness (h), the thickness of the oxide to fail in one cycle (h_c) through a Manson-Coffin type relationship:

$$N_f = \left[\left(\frac{\Delta\varepsilon_f}{\Delta\varepsilon}\right)\left(1-\frac{h}{h_c}\right)^c + \left(\frac{h}{h_c}\right)^c\right]^b \tag{29}$$

where $\Delta\varepsilon_f$ is the strain required to cause failure in one cycle with no oxidation damage and $\Delta\varepsilon$ is the tensile component of the in-plane elastic strain range in the oxide at the interface, b and c are constants (MEIER *et al.* [1991]; MEIER *et al.* [1992]).

7.2. A probabilistic model of corrosion loss

Hot corrosion processes invariably proceed in two distinct stages; an initiation stage, where breakdown of the protective surface oxide occurs, albeit by chemical or mechanical means, followed by a propagation stage where deposited contaminants have direct access to the substrate and corrosion rates can be considered catastrophic.

To model localised maximum corrosion attack (e.g., Type II, vanadic attack and other more generalised corrosion morphologies) measurement methods must be developed that not only account for the rate of growth of the pits, but also the dispersed nature of this form of attack. The basic concept behind these models was developed by one of the authors in the early 80s and has been used to predict corrosion rates within the marine gas turbine (NICHOLLS and HANCOCK [1983]; NICHOLLS and HANCOCK [1987]) and within diesel engines (NICHOLLS and TRINER [1990]; NICHOLLS and SAUNDERS [1990]). When modelling localised corrosion the common practice of measuring mean section loss and/or maximum observed corrosion penetration clearly does not apply, and hence statistical method for evaluating metal loss due to corrosion must be used (NICHOLLS and HANCOCK [1983]).

The statistical evaluation of metal loss requires that many readings of section loss (a minimum of 24 readings is recommended) be taken randomly along a corroding surface and from one component section to another. These are then evaluated using statistical techniques, for example fig. 32. In this figure, remnant coating thickness was determined at 2.5 percentiles, i.e., 39 using data points per curve. The probability of exceedance is calculated by ordering the section loss data from highest loss to lowest. Then if i is the ranking of the corrosion loss reading x_i, the probability of exceedence is given by $i/(n+1)$, where n is the total number of readings. The statistical model that is most appropriate (a normal or log normal distribution for general corrosion) can be determined from the probability plot. For example in fig. 32, the region of low corrosion loss approximates to a normal distribution. However, a few localised deep pits are observed, with a morphology typified by fig. 18. A method for modelling this extreme corrosion is required.

References: p. 1357.

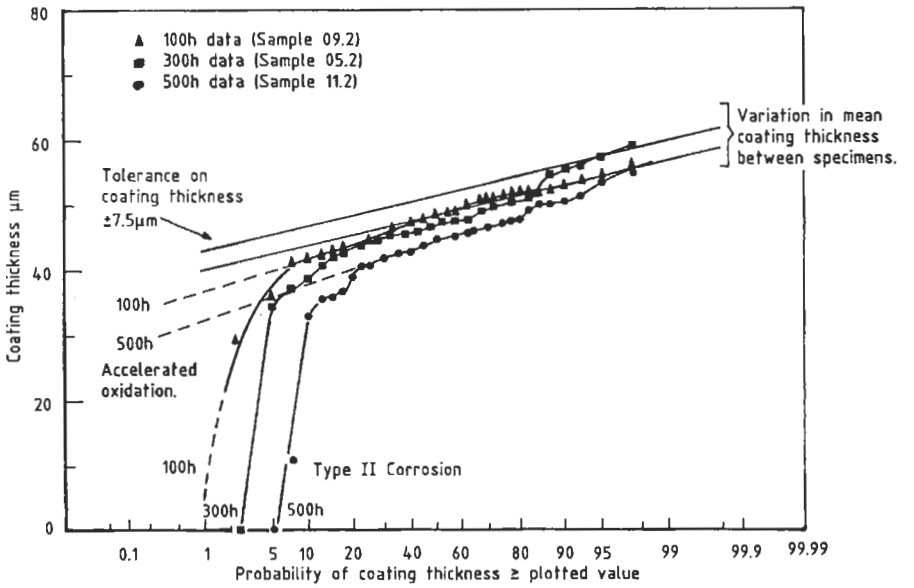


Fig. 32. Probability plot of pack aluminide coating thickness following burner rig testing at 700°C.

7.3. Modelling extreme corrosion

Here, the life-limiting rate of corrosion is associated with the growth of the deepest pits, that is to say the extreme corrosion data. In fig. 32, and from similar plots of section loss, extreme corrosion can be seen to occur for the upper 10% of metal loss readings. For the data included in fig. 32 where each point represents a 2.5 percentile, four data points can be considered extreme. This should be the minimum number of data points necessary to model extreme corrosion. However, by increasing the number of sections examined it is possible to increase the number of extreme data values available. A limited number of data points taken from multiple sections would provide better estimates of the extreme corrosion if multiple sections are available for analysis.

This extreme corrosion data can be modelled using "extreme value statistics"; the best model being a Gumbel (BERRY [1975]) extreme value model of maxima (NICHOLLS and HANCOCK [1983]; NICHOLLS and HANCOCK [1987]).

A necessary condition for the use of this model is that "the upper tail of the original distribution decreases at least as rapidly as the exponential function". This conditions is met for data that follow either a normal or log-normal distribution, and hence a Gumbel extreme value model can be used to model extreme rates of hot-salt corrosion. However, this condition should be validated (through the use of probability plots) when applying extreme value statistics to other localised corrosion phenomena.

The probability density function f(x), of the Gumbel extreme value model of maxima is given by:

$$f(x) = (1/\sigma_e) \cdot \exp\left[1 - (x - \mu_e)/\sigma_e\right] \cdot \exp\left[-\exp(-x - \mu_e/\sigma_e)\right] \tag{30}$$

where μ_e and σ_e are the distribution parameters

The cumulative distribution function, $F(x)$, is given by:

$$F(x) = \int_{-\infty}^{x} f(x)dx = \exp\left(-\exp(x - \mu_e)/\sigma_e\right) \tag{31}$$

The expected value, E(x), is given by:

$$E(x) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \mu_e + \sigma_e \cdot \gamma \tag{32}$$

where γ is Euler's constant (=0.5772), and the variance of x, V(x), is given by:

$$V(x) = \pi^2\sigma_e^2/6 \tag{33}$$

by using the log-likelihood function, L(x) given by:

$$L(x) = n \cdot \log(\sigma_e) - \sum_{i=1}^{n}(x_i - \mu_e)/\sigma_e - \sum_{i=1}^{n}\exp\left[-(x_i - \mu_e)/\sigma_e\right] \tag{34}$$

the maximum likelihood estimates of μ_e and σ_e can be evaluated by solving the simultaneous equations:

$$n = \sum_{i=1}^{n}\exp\left[-(x_i - \mu_e)/\sigma_e\right] \tag{35a}$$

$$n = \sum_{i=1}^{n} x_i\left[1 - \exp\left(-(x_i - \mu_e)/\sigma_e\right)\right] \tag{35b}$$

These equations have to be solved iteratively.

Figure 33 shows an example of the relationship between the expected arithmetic mean (E(x)) and the expected standard deviation ($\sqrt{V(x)}$) and those measured experimentally (NICHOLLS [1993]). In both cases the degree of fit is very good, confirming that these extreme corrosion data may be modelled using a Gumbel extreme model of maxima.

7.4. Development of a life prediction model

Since Type II corrosion is associated with the growth of pitting morphologies, only the extreme corrosion data from these burner rig tests is considered in developing the type II hot corrosion life model.

The most likely "extreme" value of corrosion (x_e) can be calculated and its value, of x_e, that has only a 5% chance of exceedance, is given by:

$$x_e = \mu_e = \sigma_e \ln\left[-\ln(0.95)\right] \tag{36a}$$

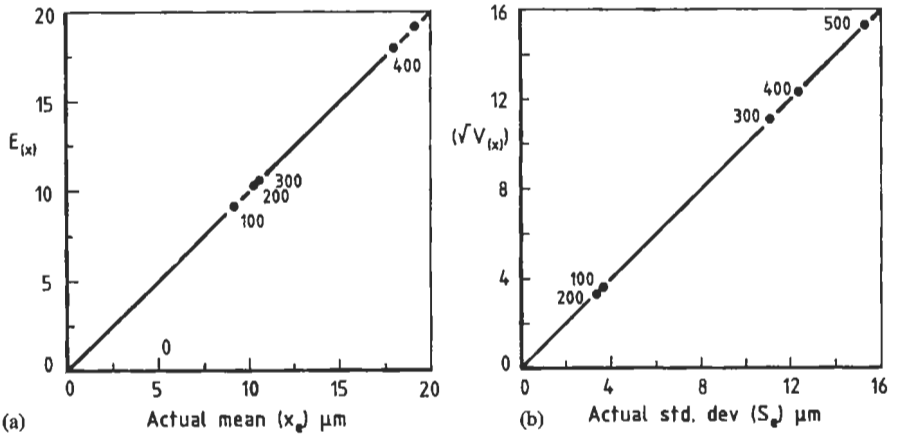


Fig. 33. Relationship between (a) expected (E_x) and actual mean, (b) expected ($\sqrt{V}_x$) and actual standard deviation of the extreme values.

and, more generally, at any probability of exceedance (α):

$$x_e = \mu_e = \sigma_e \ln[-\ln(1-\alpha)] \tag{36b}$$

By modelling μ_e and σ_e as a function of exposure time a corrosion life model can be developed which has the form for any probability,

$$x_{e(1-\alpha)} = A + B\sqrt{t} + Ct\ln[(1-\alpha)] \tag{37}$$

where A, B and C are constants dependant on the corrosion process, temperature etc., t is the exposure time in hours, x_e is the maximum expected Type II hot-corrosion pit depth, and α is the risk that the pit depth will exceed x_e within the exposure time (t).

Clearly, the constants A, B and C depend on the rate of arrival of contaminants, i.e., the contaminant flux rate. For low contaminant flux rates it can be shown that the observed corrosion rate increases with contaminant flux and follows a power law. Incorporating contaminant flux dependence into equation 37, one obtains:

$$x_{e(1-\alpha)} = X_{Na}^{0.8}\left\{A + B\sqrt{t} + Ct\ln(1-\alpha)\right\} \tag{38}$$

where A, B and C are redefined as flux independent corrosion rate constants, and take the values of 15.9, 2.2 and –0.12 respectively, for the model reproduced in fig. 34.

Two alternative failure criterion exist. Figure 34 plots the predicted maximum corrosion rates, calculated using equation (38), for two sodium contaminant flux rates (0.05 and 0.13 mg/cm^2.h) and for two confidence limits (95 and 99%). Also included in the figure are the curves for the most likely extreme value of Type II corrosion (i.e., the most likely Type II corrosion pit depth) at these two sodium fluxes. As can be seen, these models are capable of predicting the corrosion expected under engine service trials. It is also possible to predict the point at which corrosion penetration through the coating occurs, or when the corrosion pit is sufficient to cause fatigue failure (NICHOLLS *et al.* [1993]).

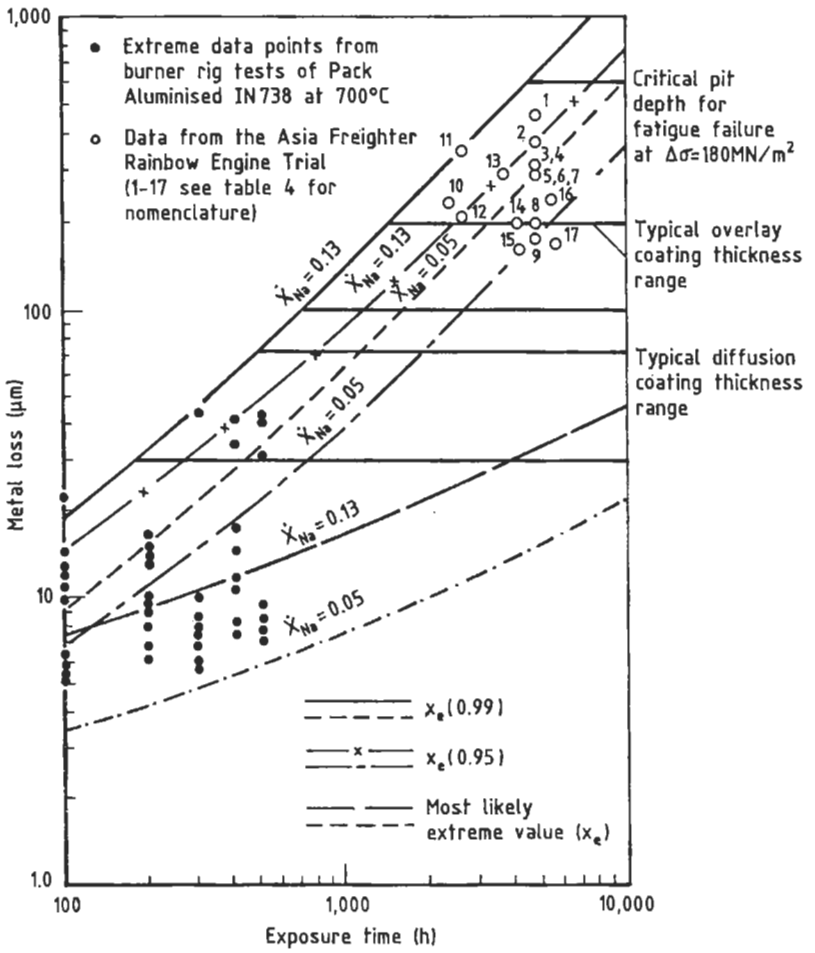


Fig. 34. Prediction of Type II corrosion at 700°C for sodium contaminant fluxes of 0.05 and 0.13 mg cm^{-2} h^{-1}.

8. *Developments in coating technology*

In discussing coating properties, performance will be assessed with reference to the desirable properties of a high-temperature coating system (table 5). In general, two types of coating process can be identified. The first requires that the composition of the surface of an alloy should be altered by diffusion, while the second involves the deposition of a metallic or ceramic overlay with improved oxidation resistance. Hence, these two processing routes are known as diffusion and overlay coating, respectively.

Diffusion coatings provide a surface enrichment of aluminium, chromium and/or silicon which results in the formation of protective surface oxides of Al_2O_3, Cr_2O_3 or SiO_2.

References: p. 1357.

Table 5
Requirements of a high-temperature system

Coating property	Requirement	Location		
		Coating Surface	Bulk	Coating/substrate interface
Oxidation/corrosion resistance	Low rates of scale formation.	X	...	...
	Uniform surface attack.	X	...	...
	A thermodynamically stable surface oxide.	X	...	...
	Ductile surface scales.	X	...	...
	Adherent surface scales.	X	...	...
	High concentration of scale forming elements within the coating to act as reserve for scale repair.	X	...	...
		X	X	...
Interface stability	Low rate of diffusion across interface at operating temperatures.	...	...	X
	Limited compositional changes across interface.	...	...	X
	Absence of embrittling phase formation during service.	...	...	X
Good adhesion	Matched coating and substrate properties to minimise coating mismatch and stress generation at coating/substrate interface.	...	X	X
	Optimum surface condition before coating.	...	...	X
	Growth stresses during coating formation should be minimised.	...	X	X
Mechanical strength	Coating must withstand all stress (creep, fatigue, and impact loading) that is generated at component surface during service.	...	X	...
	Well matched thermal expansion coefficients between coating and substrate to minimise thermal stressing and thermal fatigue.	...	X	X

Metallic *overlay coatings* are a specifically designed corrosion resistant alloy that are deposited onto a component surface and are typified by the M–Cr–Al–X series where M can be Fe, Ni and/or Co and X is an oxygen active metal, i.e., Y, Ce, Hf, Si or a combination of these. For this class of coating, interaction with the substrate is generally negligible during deposition. In addition, thicker coatings may be produced which can give extended service lives. The reduced substrate interactions, choice of corrosion resistant

alloy and the ability to deposit thicker coatings results in greater design flexibility.

Ceramic *barrier coatings* are a class of overlay coating, where corrosion protection (and thermal protection) are provided by an impervious ceramic barrier. For oxidising environments the ceramic barrier should be a stable oxide, and SiO_2, Al_2O_3 and ZrO_2 are the prime contenders, while nitride ceramics, TiN and Si_3N_4 have been examined for reducing conditions. In addition to their role as a corrosion barrier, ceramic overlay coatings may be engineered to provide additional thermal protection, the thermal barrier concept, where a cooled component, for example a turbine blade is coated with a layer of ceramic (typically 0.3–0.4 mm thick) having low thermal conductivity, thereby retarding the rate of heat flux into the component surface. Several types of thermal barrier system have been developed with varying degrees of success. The preferred family of ceramics used for proprietary coatings is based on stabilised zirconias, and in particular yttria partially stabilised zirconia (Y-PSZ).

8.1. Diffusion-coating processes

Diffusion-coating processes have been applied for many years to improve the environmental resistance of a base alloy by enriching the surface in Cr, Al, or Si. Pack chromising has been in widespread use since the early 1950s (SAMUEL and LOCKINGTON [1951]; SULLY and BRANDES [1967]) to increase the oxidation/corrosion resistance of low alloy steels. In the late 1960s aluminising was first used for the protection of superalloy gas turbine aerofoils (GOWARD and BOONE [1971]; GRISAFFE [1972]). There was renewed interest in siliconising and silicon-modified diffusion coatings for high temperature service in the early 1970s (FELIX and ERDOS [1972]; BAUER *et al.* [1989]) when novel solutions to the low temperature corrosion problems, associated with contaminants in industrial turbine plant burning impure fuels, were required. Siliconising had previously been dismissed as a major contender because of the strong embrittling effect on high-temperature alloys and the tendency to form many low melting point metallic and oxide eutectics (FELIX and ERDOS [1972]).

Diffusion coatings can be applied to hot-gas components using several techniques, including pack cementation, slurry cementation, and metallising. A comprehensive review of the methods of deposition of diffusion coatings was in MEVRET *et al.* [1986]. Fluidised bed techniques can also be used to deposit diffusion coatings on a laboratory scale (DURET *et al.* [1982]). This latter approach would permit the coating of large components with a close tolerance on coating thickness, as a result of the uniform temperatures and, therefore, chemical activities possible within a fluidised bed.

In the pack cementation process, components to be coated are buried in a "pack", contained in a sealed retort (fig. 35). The retort is heated to the desired processing temperature under either an inert gas or hydrogen atmosphere to prevent oxidation. The exact process cycle, time, and temperature are dependent on the required coating, coating thickness, and subsequent substrate heat treatment. The pack contains three elements: i) a donor alloy that releases solute material at a known rate and hence determines the pack activity, ii) a halide activator that dissociates during the process cycle and acts to

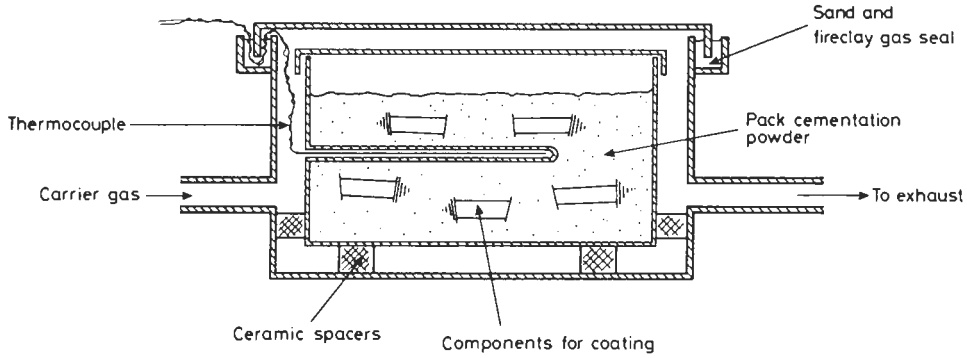


Fig. 35. Schematic diagram of pack cementation retort.

transport solute material from the pack to the component to be coated, and iii) an inert oxide diluent to prevent pack-sintering.

Typical pack compositions used to produce a range of metallic coatings are given in table 6. Of these the two diffusion-coating processes that are most widely used are "aluminising" and "chromising". The pack-aluminising process will be used as an example to illustrate a typical process cycle. During the aluminising process material from the pack is transferred to the component surface through the formation of intermediate volatile aluminium monohalide gas and therefore the coating process is probably more accurately described as a chemical vapour deposition process. Interdiffusion between the depositing aluminium and the substrate alloy results in the formation of the intermetallic coating, primarily NiAl or CoAl, depending on the alloy base, but containing to a degree most of the elements present in the base alloy either in solution or as dispersed phases.

The deposition rate and morphology of the coating depend on pack activity, process time and temperature. Coatings are classified as either "low activity", when outward diffusion of the base alloy occurs, or "high activity" when inward diffusion of aluminium

Table 6
Typical pack compositions and deposition temperatures for halide activated pack cementation

Coating	Pack composition, wt%	Deposition temperature, °C
Al	1.2–2.7Al, 0.8–11.2NaF, bal. Al_2O_3	900
Cr	48Cr, 48Al_2O_3, 4NH_4Cl	850–1050
Ti	77Ti, 20TiO_2, 3NH_4Cl	...
Si	5Si, 3NH_4Cl, bal. Al_2O_3	...

occurs. In the latter case a surface layer of Ni_2Al_3 forms and a further heat treatment is required to convert this brittle phase to NiAl. This step is usually combined with the heat treatment required to recover substrate properties. Figure 36 illustrates a typical aluminide coating deposited onto a nickel based superalloy using a high activity coating process. The coating was deposited using an aluminising pack containing 2%Al at 900°C, and was heat-treated for 2 h at 1120°C, then 24 h at 845°C.

Clearly, the properties of the aluminide coating (or for that matter any diffusion coating) depend upon the process methodology, the substrate composition and the subsequent heat treatment. Typically, aluminide coatings contain in excess of 30 wt%Al and are deposited to thicknesses between 30–100 μm depending on the type of aluminide formed. They offer satisfactory performance for many aviation, industrial and marine engine applications. Under severe hot corrosion conditions, or at temperatures above 1100°C, aluminide coatings offer limited protection. Hence modified aluminide coatings have been developed that offer improved corrosion resistance.

8.2. Modified aluminide coatings

Modified aluminides have been fabricated using one of the following techniques (DURET *et al.* [1982]):

i) Co-deposition of elements from the pack or slurry;
ii) pretreatments of the superalloy before pack aluminising, for example, chromising prior to pack aluminising;
iii) deposition of a metallic layer using electroplating or PVD (physical vapour deposition) techniques. For example, a platinum aluminide coating is formed by depositing platinum onto the superalloy prior to aluminising.

Alloying additions include Cr, Si, Ta, the rare earths and precious metals. Many of these coatings are now commercially available. Probably, the most significant advance in this area was the development of the platinum modified aluminide class of coating. Figure 37

Fig. 36. Backscattered electron micrograph of a fully processed high activity pack aluminide coating on IN 738LC.

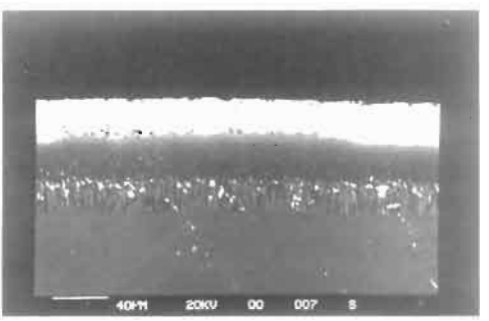


Fig. 37. Backscattered electron micrograph of a platinum aluminide (RT22) coating on IN 738LC.

References: p. 1357.

illustrates the microstructure of the RT22 variant, marketed by Chromalloy. The SEM micrograph is a backscattered electron image, the platinum-rich aluminide phase within the coating being shown in light contrast.

8.3. Overlay coatings processes

Diffusion coatings, by the nature of their formation, imply a strong interdependence on substrate composition in determining both their corrosion resistance and mechanical properties, hence, the possibility of depositing a more "ideal" coating, with a good balance between oxidation, corrosion and ductility has stimulated much research interest. Early coatings of this type were alloys based on cobalt (Co, Cr, Al, Y) containing chromium additions in the range 20–40%, aluminium additions between 12–20% and yttrium levels around 0.5%, with the most successful coating being Co25Cr14Al0.5Y (TALBOOM *et al.* [1970]). Recent coatings are more complex and are based on the M–Cr–Al–X system, where M is Ni, Co, Fe or a combination of these and X is an active element, for example Y, Si, Ta, Hf etc. (GOWARD [1983]; CLARKE [1979]; PRATER *et al.* [1981]; GOEBEL *et al.* [1981]). The composition of the M–Cr–Al–X system is selected to give a good balance between corrosion resistance and coating ductility, while the active element addition(s) can enhance oxide scale adhesion and decrease oxidation rates. Current thinking suggests that a combination of active elements is beneficial in reducing coating degradation through their synergistic interaction. Overlay coatings have been deposited using a range of techniques including plasma spraying and physical vapour deposition (PVD). Table 7 summarises some typical commercial overlay coatings and their methods of deposition.

The earliest production method was electron-beam PVD (TALBOOM [1970]). However, because of the high capital cost in setting up a commercial EB-PVD plant, plasma spray methods have found wide acceptance, particularly the argon-shrouded and vacuum-

Table 7
Some typical commercial overlay coatings

	Composition	Specification	Deposition Process
Co, Cr, Al, Y	Co–23Cr–12Al–0.5Y	ATD2B	EB PVD*
	Co–18Cr–11Al–0.3Y	ATD5B	EB PVD
	Co–18Cr–8Al–0.5Y	LCO29	Argon shrouded plasma spray
Co, Ni, Cr, Al, Y	Co–32Ni–21Cr–8Al–0.5Y	LCO22	Argon shrouded plasma spray
	Co–23Ni–30Cr–3Al–0.5Y	LCO37	Argon shrouded plasma spray
Ni, Co, Cr, Al, Y	Ni–23C0–18Cr–12.5Al–0.3Y	ATD7 PWA270	EB PVD
Ni, Cr, Al, Y	Ni–20Cr–11Al–0.3Y	ATD10	EB PVD

* Electron-beam physical vapour deposition

plasma spray processes (NICOLL [1984]; TAYLOR *et al.* [1985]; RESTALL and WOOD [1986]). More recently, composite electroplating (KEDWARD [1969]; FORSTER *et al.* [1985]; HONEY *et al.* [1986]) and laser fusion (POWELL *et al.* [1985]; STEEN [1987]) methods have been used to deposit overlay coating systems. However, coatings produced by EB PVD processes are still considered the commercial standard against which other process routes are compared.

8.3.1. Physical vapor deposition

Physical vapour deposition processes encompass both electron beam (EB) evaporation and sputtering, either of which may be used in conjunction with ion-plating (see below).

Commercial EB PVD coatings are produced in a vacuum environment, 10^{-2}–10^{-4} Pa, with deposition rates often exceeding 25 μm/min. Components are generally preheated in vacuum to between 800–1100°C, and are rotated within the evaporant cloud during the deposition process (BOONE [1980]). Rotation attempts to ensure uniform coverage as the process is primarily "line of sight" at these pressures. Problems such as columnar grains with unbonded interfaces (known as "leaders") due to shadowing effects can occur as the coating grows. High deposition temperatures result in increased surface diffusivity which reduces the density of these leaders and also permits some minimal interdiffusion between the coating and substrate during the coating process cycle, ensuring good adhesion. Hence coating spallation, a problem with many processing routes is not a problem provided the substrates are properly cleaned prior to coating. Post-coating processing includes glass bead peaning and heat treatment which results in complete closure of any remaining leaders. Figure 38 is a micrograph of a Co,Cr,Al,Y coating (ATD5B) produced by an EB-PVD processing route.

In the sputtering process, positive gas ions (usually argon) produced in a glow discharge or plasma bombard a target of coating material dislodging groups of atoms. These then enter the vapour phase and deposit onto substrates to be coated (BUNSHAH [1980]). Deposition rates are much slower than the EB evaporation route, generally

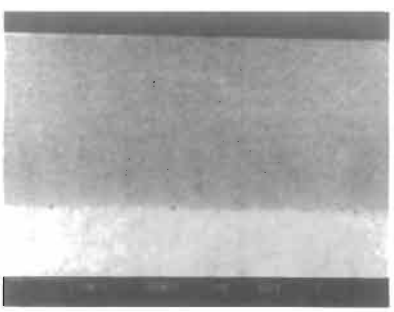


Fig. 38. Secondary electron micrograph of an EB PVD Co–Cr–Al–Y (ATD 5B) coating on Mar M002 – note the fine microstructure compared with that produced by vacuum plasma spraying, cf. fig. 39.

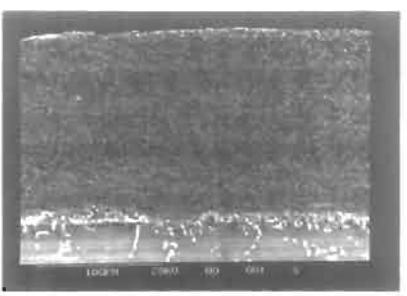


Fig. 39. Backscattered electron micrograph of a vacuum plasma sprayed Co–Ni–Cr–Al–Y (ATD2B) coating on IN 738LC.

10–20 μm/h for a diode system or up to 50 μm/h for a planar magnetron system. Magnetron systems (having magnetic plasma confinement adjacent to the target) usually result in more uniform deposition and can virtually eliminate substrate heating during the deposition process. Working pressures are of the order of $1–10^{-2}$ Pa depending on whether a DC glow discharge or RF plasma is used to generate ion bombardment and hence offers excellent throwing power and good overall coverage of components to be coated. Because of the low process temperature, coatings are invariably heat-treated following deposition to produce the desired properties.

Both EB evaporation and sputtering may be combined with *ion-plating*. Ion-plating is essentially a PVD process in a soft vacuum ($1–10^{-1}$ Pa) with evaporant depositing onto substrates held at a high negative potential (often between 2–5 kV) (TEER [1980]). During ion-plating the components to be coated are initially bombarded with positive ions, formed in the discharge, which remove oxides and other contaminants from the surface. When the surfaces are sufficiently clean the vapour source is energised and metal evaporant enters the discharge and is deposited onto the sample. With EB evaporation and ion-plating, deposition rates are typically 10–20 μm/min.

An interesting combination of sputtering and ion-plating (sputter ion-plating, SIP), using a hollow cathode sputtering target and application of a bias to the substrates held within the hollow target, has been described (COAD and DUGDALE [1979]) and used to deposit coatings on turbine blade materials (COAD and RESTALL [1982]). The system operates at soft vacuum (1–10 Pa). The geometry of the sputter targets and components to be coated offer very good throwing power and therefore uniform coatings. The system is heated to ~300°C during coating with a bias typically of 1 kV. Deposition rates are typical of diode sputtering systems, i.e., ~10 μm/h. Again, heat treatment of the as deposited coating is required to obtain the desired properties.

8.3.2. Spraying processes

Plasma-spraying has the advantage of being able to deposit metals, ceramics, or a combination of these, generating homogeneous coatings with microstructures consisting of fine equiaxed grains, i.e. no columnar/leader defects. High deposition rates are possible with little significant change in composition occurring from the powder feedstock through to the coating, even when the elements in the coating have widely differing vapour pressures. The most obvious limitations of this process are that the coating process is 'line of sight', requiring complexed robotic manipulation for complete coverage, and that the more reactive elements may well oxidise during the spraying process if conducted in air. Porosity problems previously reported for plasma spray coatings can largely be overcome using post coating thermomechanical treatments. Figure 39 shows a fully processed, argon-shrouded plasma-sprayed LCO22 (Co–Ni–Cr–Al–Y) overlay coating.

Carrying out the plasma-spraying process in a closed chamber under reduced pressure ($5–7 \times 10^3$ Pa), known as low-pressure plasma-spraying (LPPS) or vacuum plasma-spraying (VPS), permits close control of the gaseous environment. This removes the possibility of unwanted gas-metal reactions and also permits high particle velocities

giving improved adhesion and higher densities, thus overcoming many of the limitations of plasma-spraying but at an increased capital plant cost.

In the high-velocity oxy-fuel (HVOF) spraying process fuel is burnt with oxygen at a high pressure and generates a high-velocity exhaust jet. Fuel gases may include acetylene, kerosene, propane, propylene, hydrogen, methacetylene-propadiene (MAPP) mixture etc. The ratio of the gas flow rates determines the temperature of the flame. The temperature of oxygen-acetylene reaches a maximum at 3170°C if mixed 1.5 to 1 (by volume) and oxygen-propylene reaches the maximum temperature of 2900°C at the ratio of 4 : 1 (NEIDERBERGER and SCHIFFER [1990]). The velocity of the exhaust jet in the Jet–Kote torch is about 2000 m s^{-1} (KREYE [1991]). The main advantage of this process is the shorter residence time in the flame (lower powder temperature) and higher kinetic energy of the particles impacting. This produces a dense coating with less degradation of this powder during spraying.

The most widely sprayed powders are carbides, although there is considerable interest in using the HVOF process to deposit M–Cr–Al–Y overlay coatings. The bond strength of HVOF-sprayed coatings can be as high as 90 MPa (KREYE [1991]). Porosity is low, less than 1%, and typical coating thicknesses are in the range 100–300 μm.

8.3.3. Laser processes

Lasers can be used to form coatings directly by (i) preplacing a powder onto the substrate, (ii) blowing the powder into a laser-generated melt pool, or (iii) applying the clad in wire or sheet form. Alternatively, coatings produced by other routes such as plasma-spraying or electroplating may be modified by surface treatment with the laser. By careful selection of laser conditions, dilution of the coating by substrate melting can be minimised and large areas can be surfaced by a series of overlapping tracks. Power, beam diameter, and beam velocity are the main process variables with typical values being 1.5 kW, 5 mm, and 10 mm s^{-1}, respectively. As with plasma-spraying it is necessary to apply an inert gas shroud to reduce oxidation of the coating. POWELL *et al.* [1985], suggest that adhesion is improved if some melting of the substrate occurs.

Problems associated with reflectivity of the surface occur particularly with use of wire and sheet, and for the preplaced powder method there are difficulties in maintaining the powder on the surface. Thus the blown powder method is favoured by STEEN [1987]. Porosity in coatings can occur from cavities between overlapping tracks, solidification cavities, or gas evolution, but problems can generally be overcome be careful choice of processing parameters.

Residual stresses in a laser coating are generally tensile and occur from restrained contraction of the prior molten clad. Cracking may result, but can usually be overcome by preheating or pretensioning the substrate before coating (STEEN [1987]).

8.4. Oxidation and hot-salt corrosion resistance of diffusion and overlay coatings

The corrosion performance of a range of diffusion and overlay coatings has been reviewed by MALIK *et al.* [1982] and NICHOLLS *et al.* [1984]. Over the temperature range 700–950°C the platinum-modified aluminides performed exceptionally well. The other

diffusion coatings, chromised and chrome-aluminised coatings, generally fared worse than the conventional aluminide, although chromium-rich coatings have been shown to offer improved corrosion resistance in an industrial turbine environment (MALIK *et al.* [1982]; LUTHRA and LEBLANC [1987]).

Overlay coatings generally perform better at higher temperatures where oxidation is the dominant failure mode (above 900°C) reflecting the good adherence of the thin alumina scales which is promoted by the presence of active elements such as yttrium. However, at lower temperatures, 650–850°C, corrosion rates are relatively high. Methods have been investigated to improve the traditional M–Cr–Al–Y coatings by use of a platinum underlayer and overlayers (PRATER *et al.* [1981]), other additions such as Ti, Zr, Hf, Si and Ta, (DURET *et al.* [1986]; SAUNDERS and NICHOLLS [1984]; BORNSTEIN and SMEGGIL [1982]). Surface modification by CVD (RESTALL and HAYMAN [1984]), PVD (GOEBEL *et al.* [1981]) or slurry cementation techniques, (GOEBEL *et al.* [1981]) and re-processing of the coating surface using laser beams (DURET *et al.* [1982]; BORNSTEIN and SMEGGIL [1982]) have also been considered with varying degrees of success.

Surface modification results in the formation of a duplex coating structure and this can result in improved performance, for example, a pulse-aluminised Co–Ni–Cr–Al–Y coating exhibits superior corrosion resistance at 750 and 850°C compared to its plasma-sprayed counterpart (NICHOLLS *et al.* [1984]). Similarly, silicon modifications to the surface of Co–Cr–Al–Y coatings (GOEBEL *et al.* 1981) have also been shown to improve resistance to low-temperature hot corrosion.

In addition to corrosion resistance, surface coatings can have a major influence on the erosion/corrosion resistance of hot-end components. As mentioned earlier, environmental resistance depends on the formation of thin protective oxide scales and therefore the interaction between an erodent and the oxide scale that is formed on a component is of particular importance in determining the erosion/corrosion behaviour of a particular coating. The effect of particle impingement on degradation rate depends on whether the impact event produces failure within the protective oxide scale. The higher the impact energy (larger particles, higher velocities) the more likely it is that scale failure will occur and this will result in an increase in corrosion rate (HANCOCK *et al.* [1987]; STEPHENSON [1989]) and reduce incubation times to the onset of catastrophic corrosion.

8.5. Thermal stability of diffusion and overlay coatings

Structural stability of coatings is also an important factor if they are to maintain their protective qualities over extended periods of time at high temperatures. Coatings degrade not only by loss of scale-forming elements to the surface, but also by interdiffusion with the substrate. This can result in additional problems such as the formation of topologically close-packed (TCP) phases, e.g., sigma, below the coating, which cause embrittlement of the substrate.

In aluminide diffusion coatings, degradation results in the breakdown of the outer β-NiAl phase by the formation of a γ'-Ni_3Al network which after extended exposure time can completely penetrate through the β-phase. These γ' leaders behave as short-circuit corrosion paths which results in rapid failure of the coating. The thermal stability of diffu-

sion coatings, including pack aluminide, chromium aluminide and platinum aluminide variants were studied in the early eighties (COOPER and STRANG [1982]; LANG and BULLOCK [1982]). These studies have shown that the thermal stability of β-NiAl can be improved by the addition of platinum. The two-layered $PtAl_2$ β structure degrades to single-phased β with no γ' leaders even after extended exposure time, e.g., 18,000 h at 850°C (HANCOCK and NICHOLLS [1984]). This is a major factor which contributes to the improved corrosion resistance of platinum–aluminide coatings compared to conventional aluminides.

Few papers have been published on the metallurgical stability of overlay coatings, probably because diffusion is not a prime factor in the formation of these coatings. Nevertheless, substrate interaction effects are still observed e.g., hafnium from hafnium modified MM200 (GOWARD [1983]) has been found to improve the hot corrosion resistance of Co–Cr–Al–Y coatings. Movement of other elements, e.g., tungsten and molybdenum, from the substrate, may well be detrimental to corrosion resistance. The beneficial effect of a platinum underlayer under marine conditions (CLARKE [1979]) may well be due to a diffusion barrier effect.

One of the most comprehensive studies on the stability of overlays is that of MAZARS *et al.* [1986], which have examined in detail the diffusion degradation of Ni–Cr–Al–Y, Co–Cr–Al–Y, and Fe–Cr–Al–Y systems. From their work they conclude that diffusion stability can be considered a life limiting factor if coatings are operated at temperatures > 1000°C for prolonged periods. However, limited interdiffusion was observed in the Ni–Cr–Al–Y/Ni based, Ni–Cr–Al–Y/Fe based, and Co–Cr–Al–Y/Fe based systems and hence these systems are of particular interest for high temperature applications.

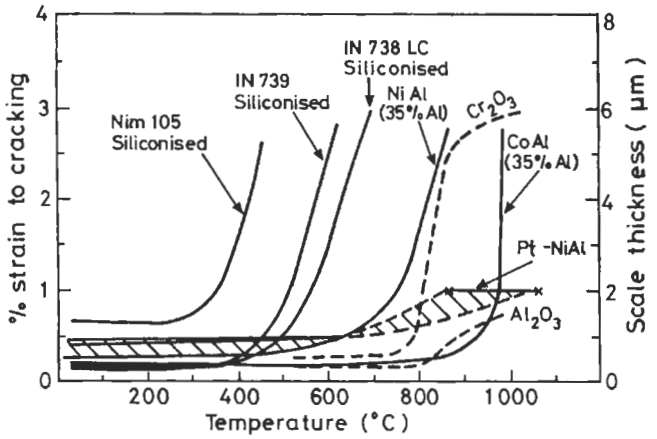
8.6. Mechanical properties of diffusion and overlay coatings

Probably the most important property of a coating for service at elevated temperature is its resistance to cracking by thermally induced stresses. A summary of the ductile-brittle transition temperatures (DBTT) for a range of diffusion and overlay coatings is shown in fig. 40, in which it can been seen that the DBTT of the aluminides is higher than that observed for many of the overlay coatings. Since the peak tensile surface strains on the turbine blades are likely to occur at relatively low temperatures, (GLENNY [1974]) then from a ductility point of view these diffusion coatings will be inadequate or only marginally adequate for relatively high strain applications.

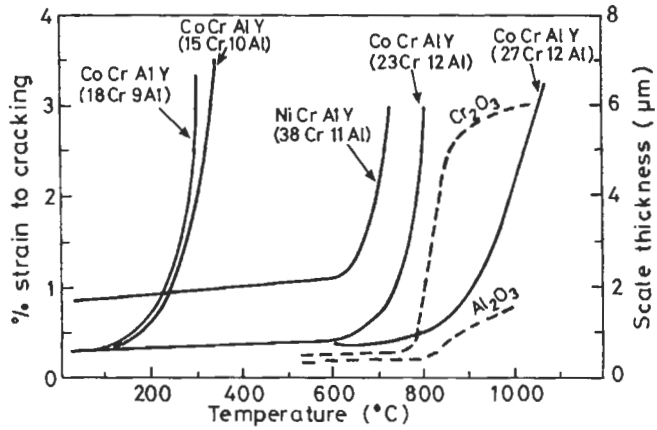
Unlike the aluminides, the ductile-to-brittle transition of overlay coatings, and hence their resistance to cracking, can be modified by varying the coating composition. It is evident from fig. 40 that, for a given Cr and Al level, the ductility of Ni–Cr–Al–Y coatings is significantly better than that of Co–Cr–Al–Y coatings. The ductile to brittle transition temperature is found to increase with Cr and Al level. This is not unexpected since coatings with a high volume fraction of Co–Al or Ni–Al shows ductile to brittle transition temperatures close to those for the diffusion aluminides.

This transition from ductile behaviour to brittle behaviour, as the temperature drops, is of significance when considering the thermo-mechanical cycles seen in service. Should coatings crack on cooling, there is a danger of the crack propagating to the coating/substrate interface. It may arrest there, may propagate along the interface or may run into the substrate component, thereby decreasing the component's life. In an investigation of

References: p. 1357.



Ductile / brittle transition curves for diffusion coatings.



Ductile/brittle transition curves for overlay coatings.

Fig. 40. Ductile/brittle transition curves for a) diffusion coatings and b) overlay coatings.

a plasma-sprayed Ni–Co–Cr–Al–Y coating on a single crystal super alloy at 650°C, a marked reduction in the fatigue life of the coated component at low strain ranges was observed (GAYDA *et al.* [1987]). The mechanical behaviour of coated components is therefore of major importance but is beyond the scope of this chapter. Readers are recommended to consult the papers by WELLS and SULLIVAN [1968]; STRANG and LANG [1982]; GAYDA *et al.* [1987]; and AU *et al.* 1990 for further appreciation of the subject.

8.7. Future trends in overlay coating design

Although coatings are available that offer adequate oxidation/corrosion resistance for many applications, improvements in coating performance can still be made. The drive to operate plant hotter, to increase efficiency, coupled with the desire for increased component

lives, requires that new generations of coatings be developed. As operating temperatures are increased, it no longer becomes possible to achieve the desired service lives using diffusion coatings and overlay coatings therefore are the only possible development route.

For metallic systems, much work is underway to develop overlay coatings containing multiple active elements. This acknowledges that although many of the active elements (Y, Hf, Ta, Si etc.) produce similar improvements in oxidation behaviour, they function by different mechanisms and hence show maximum benefit at different operating temperatures. Since high-temperature components encounter a range of temperatures, clear benefits are possible from the inclusion of multiple active elements in a coating.

But even within the M–Cr–Al–X systems, diffusion of elements between the substrate and coating can have a major influence on coating performance. Therefore, to provide long-term stability it is necessary to develop diffusion barrier coatings to minimise the interdiffusion between the coating and the substrate. Some interdiffusion is of course necessary to give good adhesion, hence the diffusion barriers must be tailored to limit the movement of particular problematic elements.

Figure 41 illustrates the deposition of an intermetallic diffusion barrier at the interface of a high Cr-containing overlay coating and a superalloy substrate (MarMOO2). The diffusion barrier was developed to limit chromium movement from the coating into the substrate (NICHOLLS *et al.* [1988]) and hence reduce the likelihood of forming sigma phase within the alloy below the coating. Chromium diffusion rates were reduced by one to two orders of magnitude over the temperature range 750–1150°C. Diffusion barrier concept based on amorphous silica have also been examined to limit coating oxidation by restricting grain-boundary diffusion (BENNETT *et al.* [1991]; MORSSINKOF *et al.* [1989]).

Once the concept of a diffusion barrier is accepted as a method of providing good interface stability, one is no longer constrained in the design of the best overlay coating. No longer is substrate compatibility a requirement in specifying the overlay coating composition. By removing this constraint it is possible to design overlay coatings with optimised oxidation or corrosion resistance (NICHOLLS *et al.* [1989]; NICHOLLS *et al.*

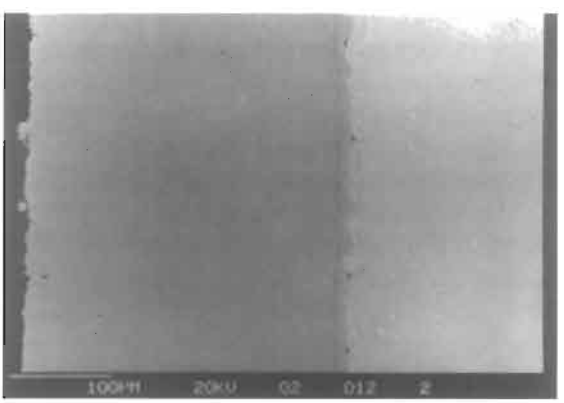


Fig. 41. Graded M–Cr–Al–Y overlay coating and diffusion barrier (located at the coating/substrate interface) system on Mar M002, deposited by EB-PVD.

References: p. 1357.

[1993]). These studies have shown that optimum performance is achieved using different base alloy compositions.

Matching the diffusion barrier coating compositions to those of the optimised corrosion resistant alloys leads to the concept of functionally graded (or multilayered) coatings. The graded coating concept is not new; it was first proposed by Boone in 1979 (BOONE *et al.* [1979]). However, processing technology was not capable of varying the composition of more than one element through the coating at that time. It is now possible to vary the composition of all three base elements, Ni, Cr and Al, and also add up to three active elements as necessary. Figure 41 illustrates this new generation of EB-PVD overlay coating. The coating incorporates a diffusion barrier layer, an interfacial zone low in chromium and a zone graded outward in Cr, with a high Cr, high Al corrosion resistant surface layer (NICHOLLS *et al.* [1993]). For low-temperature hot-corrosion resistance the surface composition should ideally be Ni–33Cr–33Al with the presence of active elements, possibly in the form of additions of yttrium, silicon and hafnium in combination.

Plasma-sprayed thermal barrier coatings (e.g., fig. 42) have been used since the early 1950s and have been effectively used to protect piston crowns and valve faces in diesel engines and flare head and primary zone sections of combustors against the effects of hot spots within turbines, giving considerably improved component lives. Because of this success much attention has been directed towards the use of thermal barriers on gas turbine blades and vanes (DUVALL and RUCKEL [1982]); but only within the last decade have they been used on highly stressed turbine components within commercial gas turbine engines (FAIRBANKS and HECHT [1987]; RHYS–JONES [1990]). Strain tolerant TBC are now being developed which have a columnar grain structure that are produced by EB-PVD processing, and fig. 43 is an example of an experimental coating now being produced at Cranfield.

Extending these concepts further, one can envisage the "ideal" coating for hot gas paths. It would consist of a strain-tolerant thermal barrier coating, deposited onto a bond coat which exhibited good corrosion resistance but, more importantly, matched thermal

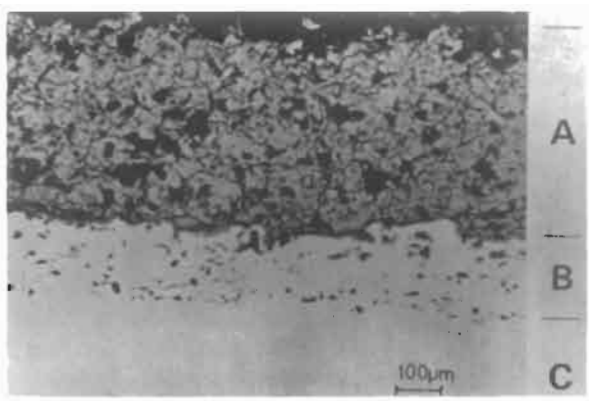


Fig. 42. Air plasma-sprayed thermal barrier coating, A TBC, B bond coat and C substrate.

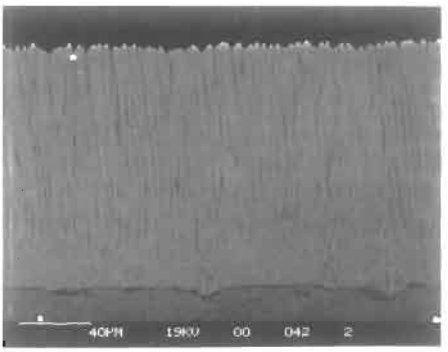


Fig. 43. Secondary electron micrograph of a EB-PVD thermal barrier coating, deposited at Cranfield.

expansion coefficients. Within this bond coat, composition can be *graded* to provide the required corrosion resistance, coating mechanical properties and a diffusion barrier at the interface with the substrate. This conceptual coating is possibly only a few years away, for research is underway into many of the fundamentals necessary to produce is *functionally graded high-temperature* protection system. The studies aimed at developing strain tolerant ceramic layer have been discussed above. Work was carried out in Germany (FRITSCHER [1987]) to develop bond coats with matched thermal expansion coefficients and the possibilities for graded corrosion-resistant coatings and diffusion barrier coatings have already been mentioned.

Acknowledgement

SRJS has been supported by the UK Dept. of Trade and Industry and their support is gratefully acknowledged.

References

ATKINSON, A., R.I. TAYLOR and P.D. GOODE, 1979 Oxid. Metals, **13**, 519.

ATKINSON, A. and R.I. TAYLOR, 1986 J. Phys. Chem. Solids, **47**, 315.

ATKINSON, A., in "Oxidation of Metals and Alloys and Associated Mass Transport", eds. M.A. Dayananda *et al.*, p. 29–47, Met. Soc. of AIME, Warrendale, 1987.

AU, P., R.V. DAINTY and P.C. PATNAIK, Proc. Conf. "Surface modification technologies II" pp. 729–748, Eds. T.S. Sudarshan and D.G. Bhat TMS-AIME (1990).

BARRETT, C.A. and C.E. LOWELL, 1975 Oxid. Met., **9**, 307.

BARRETT, C.A., R.G. BARLICK and C.E. LOWELL, "High Temperature Cyclic Oxidation Data" NASA TM 83655, May 1984.

BARRETT, C.A., J.R. JOHNSTON and W.A. SANDERS, 1978, Oxid. Met., **12**, 343.

BARRY, T.I. and A. DINSDALE, 1994, Mat. Sci. & Tech., **10**, 1090.

BAUER, R., H. W. GRUNLING and K. SCHNEIDER, in Proc. 1st Conf. on 'Advanced materials for alternate fuel capable directly fired heat engines', (eds. J.W. Firbanks and J. Stringer), p. 505; Palo Alto, CA, Electric Power Research Institute (EPRI) (1989).

BAXTER, D. J., R. C. HURST and R. T. DERRICOTT, 1984, Werkstoffe u. Korros., **35**, 266.

BENNETT, M. J., C. F. KNIGHTS, C. F. AYRES, A. T. TUSON, J. A. DESPORT, D. S. RICKERBY, S. R. J. SAUNDERS and K. S. COLEY, Mater. Sci. & Eng. **A139** (1991) 91.

BENNETT, M. J., D. J. BUTTLE, P. D. COLLEDGE, J. B. PRICE, C. B. SCRUBY and K. A. STACEY, Mater. Sci. & Eng., **A120**, (1989) 199.

BERRY, K. V., "Statistical Models in Applied Science", John Wiley, London 1975.

BIRKS, N. and G. H. MEIER, "Introduction to the High Temperature Oxidation of Metals" Edward Arnold, London 1983.

BOONE, D. H., in 'Materials coating techniques' (Conf. Proc. LS-106), Chap. 8; Neuilly-sure-Seine, AGARD (1980).

BOONE, D. H., S. SHEN and MATIALU, Proc. 4th Conf. on 'Gas Turbine Materials in a Marine Environment' p. 739, Annapolis, USA (1979).

BORNSTEIN, N. S. and J. SMEGGIL, in 'Corrosion of Metals processed by Directed Energy Beams', Met. Soc. AIME, p. 147–158 (1982).

BORNSTEIN, N. S. and M. A. DECRESCENTE, 1969, Trans. AIME, **245**, 1947.

BRUCE, D. and P. HANCOCK, 1969, J. Inst. Metals, **97**, 140 and 148.

BUNSHAH, R. F., Proc. Conf. 'Materials coating techniques' Neuilly-sur-Seine, AGARD, (1980).

CABRERA, N. and N. F. MOTT, 1948–49, Rep Prog. Phys., **12**, 163.

CHRISTL, W., A. RAHMEL and M. SCHÜTZE, 1987, Mater. Sci. Eng., **87**, 289.

CLARKE, R. L., 1979, Proc. 4th Conf. on Gas Turbine Materials in a Marine Environment, (Publ. Naval Sea Systems Command, USA), pp. 189–219.

COAD, J. P. and J. E. RESTALL, 1982, Met. Technol., **9**, 499.

COAD, J. P. and R. A. DUGDALE, in Proc. Conf. on 'Ion plating and allied techniques', London, CEP Consultants, p. 186–196 (May 1979).

COOPER, S. P. and A. STRANG, in 'High temperature alloys for gas turbines' Conf. Proc., (eds. R. Brunetaud *et al.*), 249–260, Dordrecht, D. Reidel Publishing Co. (1982).

DAVIES, R. H. and T. I. BARRY, "MTDATA Handbook", National Physical Laboratory, 1989.

DELAUNAY, D., A. M. HUNTZ and P. LACOMBE, 1980, Corros. Sci., **20**, 1109.

DOYCHAK, J., Intermetallic Compounds: Vol 1, Principles Eds. J.H. Westbrook and R.L. Fleischer, 1994 John Wiley, 977.

DOYCHAK, J. and M. G. HESBUR, in Microscopy of Oxidation (eds. M.J. Bennett and G.W. Lorimer) The Institute of Metals, London, 1991, p. 206.

DUFFY, D. M. and P. W. TASKER, Report TP1155 Harwell Laboratory, Oxon. UK, 1986.

DURET, C., A. DAVIN, G. MARRIJNISSEN and R. PICHOIR, in 'High temperature alloys for gas turbines' (Conf. Proc.) (eds. R. Brunetaud *et al.*), p. 53–87; Dordrecht, D. Reidel Publishing Co (1982).

DUVALL, D. S. and D. L. RUCKLE, 1982, ASME paper 82-GT-327.

EVANS, H. E. and R. C. LOBB, Proc. 9th Int. Cong. on Metallic Corrosion, Toronto, 1984, Vol 2, p. 46–53, National Research Council, Ottawa Canada.

EVANS, H. E., 1988, Mater. Sci. Technol., **4**, 415.

EVANS, A. G. and R. M. CANNON, in "Oxidation of Metals and Associated Mass Transport", (eds. M.A. Dayananda *et al.*), 135–160, 1987, Warrendale PA, Met. Soc. of AIME.

EVANS, H. E., 1994, Mater High Temp., **12**, 219.

EVANS, H. E., 1995, submitted to Mat. Sci and Eng.

EVANS, U. R., "The Corrosion and Oxidation of Metals", Matthew Arnold, London, 1960.

FAIRBANKS, J. W. and R. J. HECHT, 1987, Mater. Sci. Eng. **88**, 321.

FELIX, P. and E. ERDOS, 1972, Werkstoffe u. Korros., **23**, 626.

FORSTER, J., B. P. CAMERON and J. A. CAREWS, 1985, Trans. Inst. Metal Finish, **63**, 115.

FRITSCHER, K., 1987, Z. Werkstofftechn **18**, 265.

FUNKENBUSCH, A. W., J. G. SMEGGIL and N. S. BORNSTEIN, Metall. Trans. A, **16A**, (1985) 1164, and SMEGGIL, J. G., A. W. FUNKENBUSCH and N. S. BORNSTEIN, Metall. Trans. A, **17A**, (1986) 923.

GAYDA, J., T.P. GABB, R.V. MINER and G.R. HALFORD, Proc. TMS-AIME Annual Symp. 1987, pp. 217–223. Eds. P. K. Liaw and T. Nicholas, TMS (1987).

GLENNY, R.J.E., in 'High temperature materials in gas turbines', (eds. P.R. Sahm and M.O. Speidel), 257; Amsterdam, Elsevier (1974).

GOEBEL, J.A., C.S. GIGGINS, M. KRASIJ and J. STRINGER, in Proc. 2nd Conf. on 'Advanced materials for alternate fuel capable heat engines', (ed. J. W. Firbanks and J. Stringer), Report no. 2639SR, 7/1; Palo Alto, CA, EPRI (1981).

GOEBEL, J.A. and F.S. PETTIT, 1970, Met. Trans., **1**, 1943 and 3421.

GOWARD, G.W. and D.H. BOONE, 1971, Oxid. Met., **3**, 475.

GOWARD, G.W., in 'High temperature corrosion' (Conf. Proc. NACE-6), (ed. R. A. Rapp), p. 553–560; Houston, TX, NACE (1983).

GRABKE, H.J. and I. WOLF, 1987. Mat. Sci. and Eng., **87**, 23.

GRISAFFE, S.J., in 'The Superalloys' (eds. C.T. Sims and W.C. Hagel), p. 341; New York, John Wiley (1972).

HANCOCK, P., 1982, Corros. Sci., **22**, 51.

HANCOCK, P., J.R. NICHOLLS and D.J. STEPHENSON, 1987, Surface and Coatings Tech. **32**, 285.

HANCOCK, P., and J.R. NICHOLLS, 1988, Mater. Sci. Technol., **4**, 398.

HANCOCK, P. and J.R. NICHOLLS, in 'Coating for heat engines' (Workshop Proc.), (ed. R. L. Clarke *et al.*), 31–58; Washington, DC, US Department of Energy (1984).

HANCOCK, P. and J.R. NICHOLLS, 1994, Materials at High Temps., **12**, 209.

HANCOCK, P., 1970, Werkstoffe und Korros., **21**, 1002.

HAUFFE, K., "Oxidation of Metals", Plenum, New York, 1966.

HAUFFE, K. and B. ILSCHNER, 1954, Z. Elektrochem., **58**, 382 and 467.

High Temperature Corrosion, ed. R.A. Rapp, 1983, Houston Texas, National Association of Corrosion Engineers.

HONEY, F.J., E.C. KEDWARD and V. WRIDE, 1986, J. Vac. Sci. Tech.

HUNTZ, A.M., 1988, Mater. Sci. Technol., **4**, 1079.

HURST, R.C., J.B. JOHNSON, M. DAVIES and P. HANCOCK, in "Deposition and Corrosion in Gas Turbines", eds. A.B. Hart and A.J.B. Cutler, 143–157, Applied Science Publishers, Barking, 1973.

ILETT, C., "Applications of Scanning Acoustic Microscopy in Materials Science", Oxford Univ, 1984. Thesis submitted for D. Phil.

JOST, W., "Diffusion in Solids, Liquids, Gases", 492–501, 1952, New York, Academic Press Inc.

KEDWARD, E.C., 1969, Metallurgia, **79**, 225.

KERBY, R.C. and J.R. WILSON, 1973, Can. J. Chem., **51**, 1032.

KHANNA, A.S., B.B. JHO, B. RAJ, 1985, Oxid. Met., **23**, 159.

KIUKKOLA, K. and C. WAGNER, 1957, J. Electrochem. Soc., **104**, 379.

KOFSTAD, P., "Nonstoichiometry, Diffusion, and Electrical Conductivity of Binary Metal Oxides", Wiley, New York, 1972.

KOFSTAD, P., 1989, Mater. Sci. Eng., **A120**, 25.

KOFSTAD, P., 1985, Oxid. Met., **24**, 265.

KOFSTAD, P., "High Temperature Corrosion", Elsevier Applied Science, London, 1988.

KREYE, H., 'High velocity flame spraying-processes and coating characteristics'. Second Plasma Technik, Symposium, Lucerne, Switzerland 5–7 June (1991), p. 39–47.

KUBASCHEWSKI, O. and B.E. HOPKINS, "Oxidation of Metals and Alloys", Butterworths, London, 1967.

KÜPPENBENDER, I. and M. SCHÜTZE, 1994, Oxid. Met., **42**, 109.

LANG, E. and E. BULLOCK, 'European concerted action, COST 50 — Materials for gas turbines', EUR Report 8242 EN, (1982).

LEES, D.G., 1987, Oxid Met., **27**, 75.

LOBB, R.C., 1984, Thermochemica Acta, **82**, 191.

LUTHRA, K.L. and O.H. LEBLANC, 1987, Mater. Sci. & Eng., **88**, 329.

LUTHRA, K.L. and C.L. BRIANT, 1986, Oxid. of Metals, **26**, 397.

LUTHRA, K.L. and D.A. SHORES, 1980, J. Electrochem. Soc., **127**, 2202.

LUTHRA, K.L. and W.L. WORRELL, in "Properties of High Temperature Alloys with Emphasis on Environmental Effects", (eds. Z.A. Foroulis and F.S. Pettit), 318–330, 1976, Princetown NJ, Electrochem Soc. Inc.

MALIK, M., R. MORBIOLI and P. HUBER, in 'High temperature alloys for gas turbines' Conf. Proc., (eds. R. Brunetaud *et al.*) 87–98, Dordrecht, D. Reidel Publishing Co. (1982).

Mater. Sci. Technol., (1988) **4** (No. 5 Special conf. issue — Oxide/Metal Interface and Adherence).

Materials at High Temps., **12** (1994).

MAZARS, P., D. MARESSE and C. LOPVET, in 'High temperature alloys for gas turbines-1986', (eds. M. Betz *et al.*) 1183–1192; Dordrecht, D. Reidel Publishing Co. (1986).

MCCREATH, C. C., 1983, Corrosion Sci., **23**, 1017.

MEIER, S. M., K. D. SHEFFLER, D. M. NISSLEY and T. A. CRUISE, 'Thermal barrier coating life prediction model development' Transactions of the ASME, Volume 114, p. 258–263, April 1992.

MEIER, S. M., K. D. SHEFFLER and D. M. NISSLEY, 'Thermal barrier coating life prediction model development — Phase 2 — Final Report 2'. Pratt & Whitney Engine Company, NASA Contractors Report, CR-189111, July 1991.

MEVREL, R., C. Duret and R. Pichoir, 1986, Mater. Sci. Technol., **2**, 201.

MICHELS, H. T., 1977, Met. Trans A, **8A**, 273.

MORSSINKOF, R. W. J., T. FRANSEN, M. D. HENSINKVELD and P. J. GELLINGS, 1989, Mater. Sci. Eng. **A121**, p. 449–445.

MOTT, N. F., 1947, Trans. Faraday Soc., **43**, 429.

MROWEC, S., 1980, Werkstoffe und Korrosion, **31**, 371.

MROWEC, S., 1967, Corros. Sci., **7**, 563.

NATESAN, K., in "High Temperature Corrosion", NACE-6, ed. R.A. Rapp, p. 336–344, NACE Houston, 1983.

NICHOLLS, J. R. and P. HANCOCK, in "High Temperature Corrosion", NACE-6, p. 198–210, ed. R.A. Rapp, NACE, Houston, 1983.

NICHOLLS, J. R. and P. HANCOCK, in "Plant Corrosion Prediction of Materials Performance", (eds. J.E. Strutt and J.R. Nicholls), 257–273, Chichester, UK, Ellis Horwood, 1987.

NICHOLLS, J. R. and P. HANCOCK, in "The Rôle of Active Elements in the Oxidation of High-Temperature metals and Alloys", (eds. E. Lang *et al.*), 1989, Elsevier Applied Science, p. 195–223.

NICHOLLS, J. R., K. J. LAWSON, L. N. AL–YASIRI and P. HANCOCK, 1993, Corros. Sci., **35**, 1209.

NICHOLLS, J. R., K. J. LAWSON, G. CHESTER, L. H. AL YASIRI and P. HANCOCK, in European Research on Materials Substitution (Eds. I.V. Mitchell and H. Nosbuch) p. 295–307, Elsevier Applied Science, London (1988).

NICHOLLS, J. R. and S. R. J. SAUNDERS, in "High Temperature Materials for Power Engineering" (Eds. E. Bachelet *et al.*), Kluwer Academic Publishers, Dordrecht, (1990), p. 865–875.

NICHOLLS, J. R., D. J. STEPHENSON, P. HANCOCK, M. I. WOOD and J. E. RESTALL, in Proc. Workshop on 'Gas turbine materials in a marine environment', Bath UK Ministry of Defence, Session 2, Paper 7 (Nov. 1984).

NICHOLLS, J. R. and D. A. TRINER, "Proc. Conf. Diesel Engine Combustion Chamber Materials for Heavy Fuel Operation", Institute of Marine Engineers, London, (1990) p. 121–130.

NICOLL, A. R., in 'Coatings and surface treatment for corrosion and wear resistance', (eds. K.N. Strafford *et al.*), p. 180; Chichester, Ellis Horwood (1984).

NIEDERBERGER, K. and SCHIFFER, B., Eigenschaften verschidener Gase und deren Einfluss (1990), beim thermischen spritzen. Thermische Spritzkonferenz, Essen, Germany, 29–31 August 1990, p. 1–5.

PELTON, A. D., H. SCHMALZRIED and J. STICHER, 1979, J. Phys. Chem. Solids, **40**, 1103.

PIERAGGI, B. and R. RAPP, 1994, Materials at High Temps., **12**, 229.

PILLING, N. B. and R. E. BEDWORTH, 1923, J. Inst. Met., **29**, 529.

POWELL, J., P. S. HENRY and W. M. STEEN, Preprints Conf on 'Surface engineering with lasers', London, May, The Institute of Metals, paper 17 (1985).

PRATER, J. T., J. W. PATTEN, D. D. HAYES and R. W. MOSS, in Proc. 2nd Conf. on 'Advanced materials for alternate fuel capable heat engines', (eds. J.W. Firbanks and J. Stringer), Report No. 2639SR, 7/29–7/43; Palo Alto, CA, EPRI (1981).

RAPP, R. A., 1987, Mat. Sci. & Eng., **87**, 319.

RAPP, R. A. and K. S. GOTO, in "The Hot Corrosion of Metals by Molten Salts", eds. J. Braunstein *et al.*, p. 81, Electrochemical Soc., Pennington NJ, 1981.

REISING, R. F. and D. P. KRAUSE, 1974, Corrosion, **30**, 131.

RESTALL, J. E. and C. HAYMAN, in 'Coatings for heat engines' (Workshop Proc.), (ed. R.L. Clarke *et al.*) 347–357; Washington, DC, US Department of Energy (1984).
RESTALL, J. E. and M. I. WOOD, 1986, Mater. Sci. Tech., **2**, 225.
RHYS-JONES, T. N., in 'Surface Engineering Practice' (Eds. K.N. Strafford *et al.*) Ellis Horwood Ltd, p. 258–269 (1990).
ROLLS, R. and M. NEMATOLLAHI, 1983, Oxid. Met., **20**, 19.
ROSENBURG, A. J., 1960, J. Electrochem. Soc., **107**, 795.
ROY, C. and B. BURGESS, 1970, Oxid. Met., **2**, 235.
SAMUEL, R. L. and N. A. LOCKINGTON: Met. Treat. Drop Forging, **18** (1951) 354–359, 407–415, 440–444, 495–502 and 506.
SAUNDERS, S. R. J., M. M. NAGL and M. SCHÜTZE, 1994, Materials at High Temps. **12**, 103.
SAUNDERS, S. R. J. and J. R. NICHOLLS, 1984, Thin Solid Films, **119**, 247.
SAUNDERS, S. R. J. and J. R. NICHOLLS, 1989, Mat. Sci. and Techn., **5**, 570.
SAUNDERS, S. R. J. and J. R. NICHOLLS, 1995, *Materials at High Temperatures*, in press.
SAUNDERS, S. R. J. and S. J. SPENCER, 1987, Mats. Sci. and Eng., **87**, 227.
SAUNDERS, S. R. J., S. J. SPENCER and J. R. NICHOLLS, "Proc. Conf. Diesel Engine Combustion Chamber Materials for Heavy Fuel Operation", Institute of Marine Engineers, London, (1990) p. 111–120.
SCHÜTZE, M., 1985, Oxid. Met. **24**, 199.
SIMONS, E. L., G. V. BROWNING and H. A. LEIBHAFSKY, 1955, Corrosion, **11**, 505.
SMEGGIL, L. G., A. W. FUNKENBUSCH and N. S. BORNSTEIN, Spring Meeting of the Electrochemical Soc., Cincinnati, May 6–11, 1984, Extended Abstracts 84, p. 27.
SMELTZER, W. W., R. R. HAERING and J. S. KIRKALDY, 1961, Acta Metall. **9**, 880.
SMIALEK, J. L., 1987, Metall. Trans. A, **18A**, 164.
STEEN, W. M., in Proc. Conf. on 'Applied laser tooling', (eds. G.D.D. Soares and M. Perez-Amor), 131–211; Dordrecht, Martinue Nijhoff (1987).
STEPHENSON, D. J., 1989, Corrosion Sci. **29**, 647.
STOTT, F. H. and A. ATKINSON, 1994, Mater High Temp., **12**, 195.
STOUT, J. H., D. A. SHORES, J. D. GOEDJEN and M. E. ARMACANQUI, 1989, Mater. Sci. & Eng., **A120**, 193.
STRAWBRIDGE, A. and P. HOU, 1994, Materials at High Temp., **12**, 177.
STROOSNIJDER, M. J. and W. J. QUADAKKERS, 1986, High Temp. Tech., **4**, 141.
SULLY, A. H. and E. A. BRANDES, 'Chromium', 2edn. Chap. 7; London, Butterworths (1967).
TALBOOM, F. T., R. C. ELAM and L. W. WILSON, 'Evaluation of Advanced Superalloy Protection Systems' Report CR7813, NASA, Houston TX(1970).
TANIGUICHI, M., M. WAKAHARA, T. UCHIDA and K. HIRAKAWA, 1988, J. Electrochem. Soc. **135**, 217.
TAYLOR, T. A., M. P. OVERS, B. J. GILL and R. C. TUCKER, J. Vac. Sci. Tech., **3** (1985) 2526.
TEER, D. G., Proc. conf 'Materials coating techniques' Neuilly-sur-Seine, AGARD (1980).
WAGNER, C., 1956, J. Electrochem. Soc., **103**, 772.
WAGNER, C., 1975, Prog. Solid State Chem., **10**, 3.
WELLS, C. H. and C. P. SULLIVAN, 1968, Trans. ASM, 61, pp. 149–155.
WHITTLE, D. P. and J. STRINGER, Phil. Trans. R. Soc. Lond., 1980, A295, 309–329.
WOOD, G. C. and F. H. STOTT, "High Temperature Corrosion" (ed. R.A. Rapp), 227, 1983, Houston Texas, National Association of Corrosion Engineers.
WORRELL, W. L., in "Metal-Slag-Gas Reactions and Processes", (eds. Z.A. Foroulis and W.W. Smeltzer), 822–833, 1975, Princetown NJ, Electrochemical Soc. Inc.
ZACHARIASEN, W. H., 1932, J. Amer. Chem. Soc., **54**, 3841.

Further reading

See the list of titles at the end of § 1.3.

CHAPTER 15

DIFFUSIVE PHASE TRANSFORMATIONS IN THE SOLID STATE

R. D. DOHERTY

Department of Materials Engineering
Drexel University
Philadelphia, PA 19104, USA

R. W. Cahn and P. Haasen†, eds.
Physical Metallurgy; fourth, revised and enhanced edition

1. General considerations

1.1. Introduction

Solid-state phase transformations are a central topic in physical metallurgy, since almost all industrial alloys are heat-treated after casting to improve their properties. The heat-treatment changes the microstructure of the alloy, either by deformation followed by a recovery and recrystallization process (ch. 25), or by some type of phase transformation. Two main types of phase transformation are found: polymorphous changes and precipitation reactions. In a *polymorphous change*, in for example iron, cobalt or titanium, there is a change of the metal's crystal structure. This affects all the atoms in the alloy and has a tremendous scope for changing the microstructure of the alloy. In a *precipitation reaction*, which is crucial in alloys based on aluminium, copper and nickel, which have the same crystal structure up to the melting temperature, the main method of modifying the microstructure is to alloy with elements that are soluble in the base metal at high temperature but which come out of solution at lower temperatures. The particles of the new phase are called *precipitates*. The name comes from the same reaction in liquid solutions where the resulting solid crystals, being heavier than the mother liquid, do *fall* out of solution. No such movement under gravity is expected in "precipitation" from solid solutions. In a precipitation reaction there is a change from a single phase, α, structure to a two phase, $\alpha + \beta$, structure. The growth of the β precipitate requires both a change of composition of the matrix and, in cases where the two phases have a different structure, an interface reaction in which all the atoms that are to form the β phase undergo this structural change. In a minority of precipitation reactions such as those that precipitate Guinier–Preston (GP) zones, the two phases have the same structure and orientation. In such cases, no separate interface reaction is required to grow the precipitate, merely a change in composition.

In all polymorphous and most precipitation reactions, there is a migration of an interface between two crystalline phases, and there are two possible modes of interface migration. In the first of these modes, atoms make thermally activated random jumps across the interface, a "diffusive" mechanism. In the second mode, the daughter crystal grows into the parent, by a coordinated shear-type (martensitic) motion of all the atoms at the interface. The description "military" suggested by CHRISTIAN [1965] for this second type of transformation is rather appropriate and leads to his equivalent label of "civilian" for the uncoordinated interface motion of atoms in the first type of transformation. In recent literature, for example BHADESHIA and CHRISTIAN [1990], the alternative terms of "reconstructive" and "displacive" are used for the two types of method for transforming one crystal structure to the new one. The present author prefers the label of civilian or reconstructive to diffusive since any reaction such as the formation of bainite in steel (a mixture of low carbon, bcc, α, iron plus iron carbide) from single phase, solid solution of carbon in fcc, γ, austenite requires both a change of composition, that takes place by a true diffusive motion of individual carbon atom jumps and the γ to α interface reaction. The interface process could be either military or civilian. If it were military then bainite would have to be described as being diffusive and non-diffusive!

The civilian or reconstructive types of transformation are the subject of this chapter while the military types are the subject of ch. 15. Bainitic transformations, which appear to have both a diffusive character for change of carbon concentration and a martensitic character for the change of matrix structure, are discussed in § 2.6. The precipitation of close-packed hexagonal, silver-rich, Ag_2Al from an Al–Ag solid solution, which is now universally agreed to take place by a combined diffusive long-range composition change and a displacive interface reactions is discussed in detail in § 2.2.3 and 2.2.4.

The same distinction between a military/displacive and a civilian/reconstructive transformation is seen for the movement of a grain boundary in a single phase structure. Here the military transformation is that of mechanical twinning and the civilian one is the usual mechanism of thermally activated grain boundary migration seen in recrystallization and grain growth.

1.2. Driving forces — free energy changes

All structural transformations are driven, at constant temperature and pressure, by a reduction in Gibbs free energy, G, from the original to the final structure. The solution thermodynamics that are the intellectual base of the analysis of phase transformations are described in detail in ch. 5. The definition of G is:

$$G = H - TS, \tag{1}$$

where T is the absolute temperature, S is the entropy and H, the enthalpy, is given by:

$$H = U + PV, \tag{2}$$

U is the internal energy, P is the pressure, V the volume of the material. In transformations between condensed phases, the second of these terms is negligable with respect to the first term. Under these circumstances the difference between G, the Gibbs, and F, the Helmholtz free-energies becomes insignificant at least as regards work done against external pressure and the symbol F will be used throughout this chapter for the function that is minimized during a structural transformation:

$$F = U - TS, \tag{1a}$$

For a polymorphous transformation from a parent α-phase to a daughter β-phase, the driving "force", in Joules per unit volume, which has the equivalent dimensions as a pressure, in Newtons per unit area, is given in terms of the changes in internal energy, $\Delta U_{\alpha\beta}{}^{v}$, and entropy, $\Delta S_{\alpha\beta}{}^{v}$, by:

$$\Delta F^{v}_{\alpha\beta} = \Delta U^{v}_{\alpha\beta} - T\,\Delta S^{v}_{\alpha\beta} \tag{3}$$

At the equilibrium transformation temperature, T_e, $\Delta F_{\alpha\beta}{}^{v}$ is of course zero. So if $\Delta U_{\alpha\beta}{}^{v}$ and $\Delta S_{\alpha\beta}{}^{v}$ are effectively constant, at a finite undercooling, $\Delta T = (T_e - T)$, the driving force is given by:

$$\Delta F^{v}_{\alpha\beta} = \Delta U^{v}_{\alpha\beta}\left(\Delta T/T_e\right) \approx \Delta H^{v}_{\alpha\beta}\left(\Delta T/T_e\right) \tag{4}$$

For small undercoolings the constancy of $\Delta U_{\alpha\beta}{}^{v}$ and $\Delta S_{\alpha\beta}{}^{v}$ are reasonable assumptions

but for larger undercoolings, a correction should be made to eq. (4) whenever the specific heats of the two phases are different. This is readily carried out using standard thermodynamic methods, see for example SWALIN [1972]. Table 1 gives some typical values of the latent-heat changes, $\Delta H_{\alpha\beta}{}^{v}$, from which it can be seen that the driving force for solid-state polymorphous phase changes is much smaller than for solidification at similar undercoolings.

For a precipitation reaction in a binary alloy AB, such as the one described by the free energy/composition curves in fig. 1, the free-energy changes can be readily determined. The overall free-energy change, *per mole of the alloy*, is I → J, where J lies on the common tangent to the two free-energy curves. However, the free energy change, *per mole of precipitate*, is K → L where K lies on the tangent that touches the α-phase at I, the alloy composition. The intercepts M and N of the tangents at the atomic fraction of B, C_B, of $C_B = I$, are the partial molar free energies, $\bar{F}_B$, also called the chemical potentials, μ_B, of component B in the α-phase at atomic fractions C_0 and C_α, respectively. From standard solution thermodynamics, we have;

$$\bar{F}_B(C_0) = F_B^0 + RT \ln a_B(C_0), \tag{5a}$$

$$\bar{F}_B(C_\alpha) = F_B^0 + RT \ln a_B(C_\alpha), \tag{5b}$$

where $F_B{}^0$ is the free energy of pure B in its standard state, R is the gas constant and a_B is the activity of B in the α-phase at the two compositions C_0 and C_α. M → N is then given by:

$$MN = \bar{F}_B(C_\alpha) - \bar{F}_B(C_0) = RT \ln \frac{a_B(C_\alpha)}{a_B(C_0)}. \tag{6a}$$

If α is a dilute solution with C_0 and $C_\alpha \ll 1$, as in the phase diagram shown in fig. 2 and in most commercial alloys, then the activity coefficient, f_B, of B in the α-phase is constant (Henry's Law) so eq. (6a) becomes:

$$MN = RT \ln(C_\alpha / C_0). \tag{6b}$$

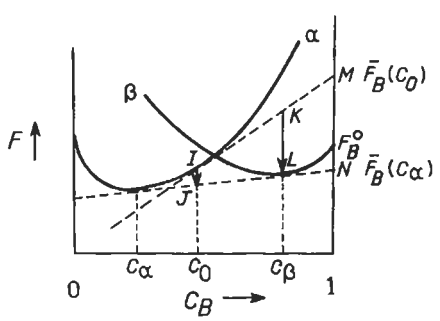


Fig. 1. Free-energy–composition curve of two phases α and β showing conditions for nucleation of β from supernaturated α of composition C_0.

Table 1
Latent heat changes.

Element	Transformation	Latent heat change (kJ/g-mol)	T_e(K)
Fe	liquid to solid	–15.5	1809
Ti	liquid to solid	–18.9	2133
Fe	austenite to ferrite	–0.9	1183
Ti	bcc to hcp	–3.5	1155

since

$$a_B(C_0) = f_B C_0.$$

If, again as in fig. 2, β is a dilute terminal solid solution of A in B, then KL in fig. 1 is effectively MN and the driving force per mole of precipitate, $\Delta F_{\alpha\beta}{}^{m}$, or per unit volume of precipitate, $\Delta F_{\alpha\beta}{}^{v}$, are given by:

$$\Delta F_{\alpha\beta}^{m} = RT \ln(C_\alpha / C_0), \tag{6c}$$

$$\Delta F_{\alpha\beta}^{v} = \{RT/V^m\} \ln(C_\alpha / C_0). \tag{6d}$$

Since $C_\alpha < C_0$ in a supersaturated solution, $\Delta F_{\alpha\beta}{}^{m}$ and $\Delta F_{\alpha\beta}{}^{v}$ are both negative. The change from free energy per mole to free energy per unit volume is achieved through the use of V^m the molar volume of the β-phase. Very often the precipitating β-phase is an intermetallic compound so that C_β is significantly < 1 so that a correction must then be made, giving:

$$\Delta F_{\alpha\beta}^{m} = \frac{C_\beta - C_\alpha}{1 - C_\alpha} RT \ln(C_\alpha / C_0). \tag{6e}$$

For the precipitation of an intermetallic phase such as $MnAl_6$ from an aluminium-rich solid solution, the effect of the modification to eq. (6e) is to decrease the driving force very significantly, by 1/7 for $MnAl_6$. An extreme version of this effect is found for the growth of ferrite, α, from low-carbon austenite, γ, in iron-carbon alloys. Since the precipitate is solute-depleted rather than solute-enriched, the form of the equation derived as here is slightly changed. We now have:

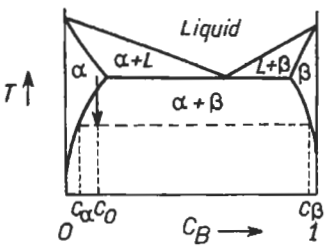


Fig. 2. Simple phase diagram derived from fig. 1.

References: p. 1497.

$$\Delta F_{\gamma\alpha}^{m} \approx \frac{C_\gamma - C_\alpha}{C_\gamma} RT \ln \frac{1 - C_\gamma}{1 - C_0} \tag{7}$$

where the compositions are given in atomic fractions of the solute, which in this case is carbon.

The low driving force for this reaction even when the alloy is cooled well into the two-phase, α plus γ, region for example when the volume fraction of α can be 0.5 or more, makes this reaction very different from most precipitation reactions. The difference arises since the driving force for nucleation of ferrite, $\Delta F_{\alpha\beta}$, is very small while the volume fraction of new phase, $\{(C_\gamma - C_0)/(C_\gamma - C_\alpha)\}$ which is also the supersaturation, Ω, that drives the diffusional growth processes (§ 2.2.2) can be very large. That is, for the formation of ferrite from austenite and in other equivalent matrix phase changes, the nucleation of the new phase has a much lower driving pressure while growth is driven much faster than in most precipitation reactions. This idea is considered in more detail in § 2.2.6.2*.

Finally it should be noted that if the matrix phase is not a dilute terminal solid solution the activity coefficient, f_B, will not be constant. The correction for this effect derived by PURDY [1971] then needs to be included. This gives:

$$\Delta F_{\alpha\beta}^{m} \approx \frac{C_\beta - C_\alpha}{1 - C_\alpha} \frac{RT}{\varepsilon_\alpha} \ln\left(C_\alpha / C_0\right). \tag{8}$$

Here ε_α is the non-ideality factor, $\varepsilon_\alpha = 1 + \mathrm{d}\ln f_B / \mathrm{d}\ln C_\alpha$, derived originally to DARKEN [1948] for diffusion studies. This term can be very significant, $\varepsilon_\alpha \neq 1$ if the matrix phase from which the precipitate is forming is itself an intermediate phase such as β brass in Cu–Zn or Cu–Al or an ordered intermetallic phase such as α_2 in the Ti–Al system.

1.3. Stable and unstable free-energy curves

The free-energy/composition curves shown in fig. 1 are both those of stable or at least metastable phases, since any alloy on a free energy composition curve where the curvature, $\mathrm{d}^2F/\mathrm{d}C^2$, is positive cannot spontaneously reduce its free energy, except by nucleating a distinctly different new phase. In a phase diagram such as that sketched in fig. 3, where an α-phase, below a critical temperature T_c, decomposes into solute-rich, α'', and solute-poor, α', phases, the free-energy/composition curve is unstable below T_c,

* Just above the eutectoid temperature (~1000 K) in Fe–C, where the carbon content of austenite is 0.8 wt%C ($C_\gamma = 0.0017$), C_α is almost zero. So, while the supersaturation for growth, for an alloy with 0.4 wt% C ($C_0 = 0.0008$), Ω is 0.5, the driving force for nucleation, $\Delta F_{\alpha\beta}{}^{m}$, is only -4 J/g-mole. For precipitation of a B-rich phase, such as θ', $CuAl_2$, from an aluminium solid solution at say 190°C with an equilibrium solubility C_α of 0.0002 and for an alloy composition C_0 of 0.02, then while Ω is only 0.06, the driving force for the nucleation of the copper rich precipitate, $\Delta F_{\alpha\beta}{}^{m}$, is 1500 times larger at $-6{,}350$ J/g-mole. As a result of these differences the polymorphous reaction in iron gives coarse grain-boundary nucleated precipitates of ferrite while the Al–Cu reaction gives a very different microstructure — that of very small copper-rich precipitates finely dispersed in the aluminium grains, § 2.1, giving precipitation hardening.

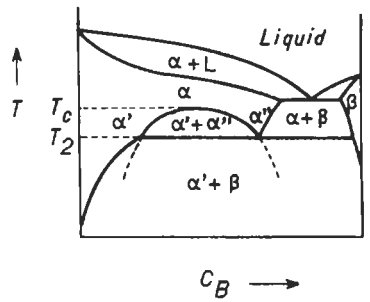


Fig. 3. More complex phase diagram than in fig. 2, showing the development of solid-state immiscibility in the α-phase below a critical temperature T_c.

figs. 3 and 4. This instability arises in the part of the curve where the curvature, d^2F/dC^2, is negative. Within this "*spinodal*" region an alloy of composition i can split into j and k and so reduce its free energy (CAHN [1968]). This can occur not only if α' and α'' are the stable phases but also when as in fig. 4, below T_2, there is an even more stable β-phase. The formation of the β-phase requires the system to overcome a *nucleation* barrier. In this case, the solute enhanced region α'', although metastable with respect to β, will still form when the rate of its formation is faster than the kinetics of the nucleation and growth of the more stable β-phase. Such regions of solute-enriched α'', sharing the same crystal lattice as α', are known as *Guinier–Preston*, GP, *zones*.

1.4. Gibbs's two types of transformation

From figs. 3 and 4, it is possible to see the distinction between the two types of transformation first described by the father of this subject, J. Willard Gibbs. In the first, type I, of these reactions, usually described as a *nucleation and growth reaction*, a small region of a new phase, such as β, forms from within the matrix. The new region, the *nucleus*, has a completely different composition and/or structure from the parent phase, α. The nucleus is separated from the parent matrix by an interface which has an

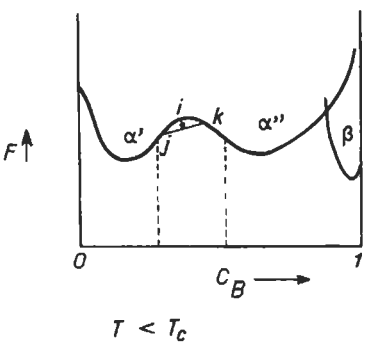


Fig. 4. Free-energy–composition curve for the phase diagram of fig. 3 at a temperature below the eutectoid temperature T_2. The curvature, d^2F/dC^2, is negative in the part between the dashed lines, positive elsewhere.

interfacial energy, σ, whose units are J/m^2. This change is one that can be described as being large in the magnitude of the local change in structure and/or composition but one that is *spatially* localized to the small part of the matrix which has transformed. The majority of the matrix remains completely untransformed at least initially, though subsequent growth of the nucleus into the matrix occurs. Other names for such "nucleation and growth" transformations are "heterogeneous" or "discontinuous" transformations. "Heterogeneous" nucleation and "discontinuous" precipitation, however, have different and quite specific meanings within the area of nucleation-and-growth reactions. As a result, it is better not to use these names for the whole class of Gibbs Type I transformations. Details of heterogeneous nucleation and discontinuous precipitation processes in various types of nucleation-and-growth reactions are discussed in § 2.

The second mode of transformation is shown by a composition fluctuation in an alloy within an unstable region of its free-energy curve. The result of such a fluctuation is that a region of initially uniform composition develops a composition wave whose amplitude grows with time, fig. 5, but with, at least initially, a constant wavelength. This second type of reaction, a Gibbs Type II transformation, where the initial change in composition in any local region is small but where the composition fluctuation is *delocalized* in space, is usually called *spinodal decomposition*, at least when the wavelength is significantly longer than diameter of an atom. When the wavelength is that of an atom, leading towards an ordered solid solution the transformation is then usually called *continuous ordering*. Spinodal decomposition is known to occur in many precipitation reactions, for example in concentrated Al–Zn alloys (RUNDMAN and HILLIARD [1967]). The term *continuous transformations* is now a common name for the Gibbs type II transformation in both precipitation and in ordering and this terminology is used here in § 3.*

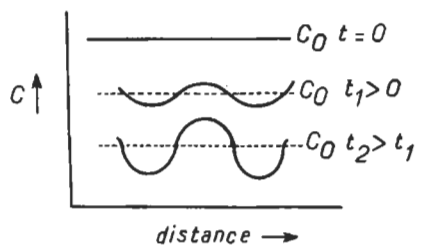


Fig. 5. Variation in solute distribution with time during a continuous reaction, Gibbs type-2: *spinodal decomposition.*

* There is again a possible confusion here since the term *continuous precipitation* is sometimes used for precipitation by a nucleation and growth reaction that does not require a migrating high angle grain boundary. *Discontinuous precipitation* is the nearly universal name for precipitation that does occur behind a migrating high angle grain boundary. This name discontinuous precipitation is so commonly used that it will he used here, § 2.5, even though an alternative name such as a moving two-phase boundary (MTPB) might be less confusing. The context of the discussion will usually avoid any difficulty. However, in this chapter, we will not use the term "*continuous precipitation*" for a nucleation and growth reaction that takes place without the assistance of a moving grain boundary.

1.5. First order and higher order transformations

In most structural transformations in physical metallurgy there is a discontinuous change in the slope of the free energy–temperature curve as the structure changes at the equilibrium temperature (fig. 6). This type of transformation is described as *first order*, since there is a discontinuity in the first derivative, dF/dT, and thus a discontinuous change in the internal energy, giving a finite *latent heat*, EHRENFEST [1933]. In *second order* transformations, there is a discontinuity in the second differential, d^2F/dT^2, and therefore in the first differential of internal energy, which is the *specific heat*. CHRISTIAN [1979] discussed the thermodynamic and mechanistic classifications of structural changes of interest to metallurgists and noted that some apparent second order transformations appear to show an apparent infinity in the specific heat rather than just a discontinuity. Unlike the first order transformations, these apparent second order transformations show an increase of specific heat (i.e., a latent heat) as the temperature approaches the critical temperature.

1.6. Short-range and long-range diffusion

In polymorphous changes, such as solidification and solid state structural transformations in pure (single component) materials, and also during grain-boundary migration in recrystallization and grain growth, the only atomic process is the transfer of atoms across the interface: This is the *interfacial* process. For nucleation-and-growth reactions where the parent and product phases have different compositions, there are then two *successive* processes; firstly, long-range transport by diffusion over distances of many atomic spacings, commonly described as the *diffusional process* and secondly, atomic transport across the interface, normally a thermally activated short-range diffusional process which is again the *interfacial process*. The long-range diffusion only involves a fraction of the total number of atoms in the new phase, those required to change the composition of the matrix to that of the new phase. The fraction may be nearly 100%, for example in the precipitation of a B-rich phase from a supersaturated A-rich solid solution (fig. 2). In other cases, like the precipitation of bcc ferrite from dilute fcc austenite in low-carbon steels where there is only a small change of composition, the fraction of atoms taking part in the diffusional process is very much less than one. By

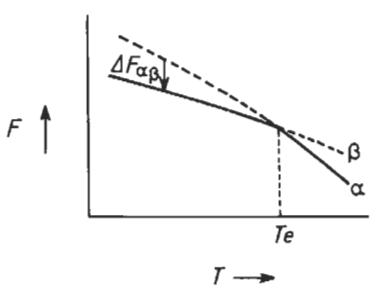


Fig. 6. Variation of free energy with temperature, showing a first-order reaction at an equilibrium transition temperature T_e.

contrast *all* the atoms that are to be part of the precipitate must take part in the interfacial reaction.

Since the two reactions, long-range diffusion and the short-range interfacial step, are *successive* reactions, the *slowest* of the two processes will be rate-controlling (§ 2.3). This may be contrasted with the situation where there are two *alternative* reactions, for example the formation of either (i) the stable phase with the lowest free energy or (ii) a metastable phase with a higher free energy. Where there are two or more *alternative* reactions, then the *fastest* total reaction will determine which reaction occurs and this will, in consequence, determine the rate of the transformation and the resulting microstructure. This situation occurs, for example, in precipitation-hardening reactions at low temperatures and high driving forces, where metastable precipitates with low barriers to nucleation form in preference to the stable precipitate when that is more difficult to nucleate. The initial metastable reaction removes much of the driving force for the formation of the stable phase, so that the metastable product may then remain for the useful lifetime of the sample.

1.7. Techniques for studying phase transformations

Since phase transformations produce a change of microstructure, almost any structural characterisation technique can be used to study a transformation. EDINGTON [1979] reviewed this topic and detailed references to the tremendous range of published work can be found there and in the numerous conference reports on the various structural characterisation techniques.

The techniques that have been used include optical, scanning and transmission electron, and field-ion microscopy. The increased resolution of the last two techniques involves a severe penalty in that the regions studied are highly localized and so the structures observed in great detail can be potentially unrepresentative. Diffraction techniques are also of vital importance, both those involving general diffraction and also those using selected-area diffraction in both transmission microscopy and at even higher spatial resolution, scanning transmission electron microscopy (STEM). At lower spatial resolution, X-ray Kossel methods, electron channelling patterns and more recently Back Scattered Kikuchi Diffraction (BKD) — also called Electron Back Scattered Patterns (EBSP) — have been of great value in investigations involving bulk samples. Of major importance, particularly for quantitative investigations, are chemical analytical studies of highly localized regions. Conventional electron-probe microanalysis allows chemical analysis down to scale of 1–2 μm, analytical electron microscopy by either X-ray emission or by electron energy loss has allowed routine analysis down to a resolution of 10 nm and below. Ultimate or near-ultimate resolution has been achieved by the atom-probe field-ion microscopy technique, the determination of groups of individual atoms by time-of-flight measurements of field evaporated atomic layers (HAASEN and WAGNER [1986], MILLER [1987]. Other successful analytical techniques include high-resolution Auger electron spectroscopy and ion probe mass spectrometry using focussed ion beam erosion. For a detailed comparison of the techniques, a conference on microanalysis at high spatial resolution provides a detailed source of information, LORIMER [1981]. The

older techniques of measuring average solute levels throughout the microstructure should not be neglected since these give additional kinetic information to that provided by the more detailed local measurements of composition. Any physical property that is modified by the solute level in the matrix can be used for general chemical analysis. For magnetic materials the ferromagnetic Curie temperature is often used, at least for studying slow reactions or ones occuring at temperatures well above the Curie temperature. Other methods include studies of lattice parameter by X-ray, electron or neutron diffraction, semi-quantitative measurements of minority phase amounts by Guinier X-ray diffraction in which the white X-ray background has been eliminated, thermoelectric power and some electrochemical methods. The most commonly used however is that of electrical conductivity since, for metals, the presence of solute has the largest effect on electrical resistivity, compared to almost all other structural defects. For ionic crystals, increased solute such as divalent magnesium ions in monovalent lithium fluoride increases the vacancy concentration, so leading to increased ionic conductivity at least in a temperature range where extrinsic conductivity is found. This technique allows very accurate chemical analysis of the solid solutions in ionic compounds.

Many of these modern high-resolution techniques have only recently become available and are usually expensive in equipment costs, so much less use has been made of them than of the very widely used transmission electron microscopy which revolutionized this subject after its introduction in the late 1950s. However, as EDINGTON [1979] has shown, these latter techniques, introduced in the 1970s, have made extensive contributions to the subject and will continue to do so. Further use of these techniques may modify the present situation, which is that in many areas, the *theories* of phase transformations are well ahead of rigorous *tests* of those theories, at least, as far as quantitative experimental studies of the kinetrics of transformation are concerned.

It is to be noted that there are two distinct modes of investigation of microstructural change which are used. In most cases the microstructure is investigated in samples where the reaction has been halted by rapid quenching to room temperature. The study of these static microstructures allows detailed insight of the microstructure at the highest resolution. However, the dynamic development of the reaction with time, in a particular region, cannot be followed by the static technique at least in one sample. A series of different samples quenched after different times must be used and *averaged* structural information obtained. The alternative technique, of dynamic or *in-situ* observation of structural change in a hot-stage microscope or equivalent equipment allows continuous investigation of the transformations in a selected region. This advantage is however, obtained at a severe cost of poorer resolution and with severe time limitations. The structure often changes before it can be fully characterized. An additional problem for the in-situ studies is the worry about modification of the transformation by the presence of near-by surfaces and by the effects of the illumination, usually a beam of high energy electrons. The use of higher voltage electron microscopy can reduce the first of these difficulties, by allowing thicker foils to be used, due to the greater penetration of the higher energy electrons. However, point defects especially vacancies can be induced by the higher-voltage beam and these vacancies if present at higher concentrations than the vacancies in thermal equilibrium can then modify the diffusional kinetics, see for

References: p. 1497.

example (e.g., WEAVER *et al.* [1978] and DOHERTY and PURDY [1983]). In most cases, however, a careful use of a *combination of techniques* allows the potential errors to be recognized and overcome and should enable valid conclusions to be drawn. In general, the static methods provide more accurate quantitative measurements while the dynamic or in-situ methods are better suited for giving qualitative insight into the operating mechanisms.

Other techniques that can give useful dynamic insights into structural changes include calorimetry, if heat is being released or absorbed at a detectable rate, dilatometry where there are significant volume changes and for most reactions neutron diffraction is attractive since large samples can be studied both statically and, most conveniently, dynamically at temperature due to the highly penetrating nature of neutrons. (see ch. 12).

2. *Nucleation-and-growth transformations*

2.1. Theory of nucleation

The basis of the theory of nucleation is that when a new phase forms within a parent phase, an interface is formed between the two phases. The formation of the interface creates a local increase of free energy when the first atoms assemble in the new structure. The theory of this process, which describes the interfacial energy barrier to nucleation, was originally developed for physical chemistry by VOLMER and WEBER [1928], VOLMER and FLOOD [1934] and BECKER and DORING [1935]. The model was applied by TURNBULL and FISHER [1949] to solidification of metals and subsequently solid-to-solid metallic phase transformations, see TURNBULL [1956]. Following these studies the concept has become fundamental to the qualitative and quantitative understanding of structural changes in physical metallurgy. The theory appears to be, qualitatively at least, highly successful both for solidification (ch. 9, § 3) and in successfully describing the nucleation step in most solid-state nucleation-and-growth reactions. The subject of solid-state nucleation has been reviewed on many occasions, for example by KELLY and NICHOLSON [1963], RUSSELL [1970], NICHOLSON [1970], CHRISTIAN [1975], RUSSELL [1980] and HAASEN *et al.* [1984], with similar conclusions to those advanced here. For almost all diffusional solid state reactions the theory is in good agreement, at least qualitatively and semi-qualitatively, with a very large amount of experimental data. There are, however, rather few fully quantitative experimental tests of the theory; these do, however, support the theory at least of homogeneous nucleation to a surprisingly high degree. The one exception where the theory has been found not to apply, even qualitatively, is for the nucleation of new grains in the recrystallization of deformed materials, see for example DOHERTY [1978]. In this case, the nuclei of low dislocation density regions of the correct structure pre-exist in the deformed state and can develop by processes that do not require a local increase in free energy. The nucleation of recrystallization is discussed in ch. 28, § 3.3. For nucleation in martensitic reactions the theory, at least to date, appears to have been somewhat less useful, as is discussed in the following chapter.

The model, for diffusional phase transformations, assumes that a new phase, differing from the parent phase in structure and/or composition, is built up, atom by atom, by

thermally activated atom transfer across the interface. There is, as a result, a local change in the free energy, which is given, in the usual continuum description, as a decrease in the *volume* free energy, due to transfer of atoms from a less stable to a more stable phase, and an increase in the *interfacial* free energy due to the increase of the area of the interface between the two phases. In the conventional model of nucleation, the macroscopic values of these parameters are used even though it is unlikely that a cluster of only a few atoms would have the macroscopic properties that the new phase will show when it has grown to a size much larger than atomic dimensions. The total change of free energy, ΔF_t, when a new phase of volume, V_β, and interface area, A_β, forms, often causing an additional increase in elastic strain energy per unit volume of precipitate, ΔF_E^v, is given as:

$$\Delta F_t = V_\beta \Delta F_{\alpha\beta}^v + A_\beta \sigma + V_\beta \Delta F_E^v, \tag{9}$$

where σ is the specific energy of the α–β interface and $\Delta F_{\alpha\beta}^v$ is the free energy change per unit volume of the new β phase, given by the macroscopic molar free energy change, $\Delta F_{\alpha\beta}^m$, divided by the molar volume, V_m, of the new phase. The first and last terms in eq. (9) are both dependent on the volume of the nucleus and are treated together. Only if the first term is larger (more negative) than the third term, which is always positive, can the reaction proceed. The volume of the nucleus increases as the *third* power of the nucleus size ($4\pi r^3/3$ if the nucleus is a sphere of radius r), while the area term, which is always positive and acts as the barrier to the process, increases only as the *second* power of the nucleus size ($4\pi r^2$ for a sphere). Figure 7 plots eq. (9) for a spherical nucleus, and shows that at small radii the nucleus shows an increase of ΔF_t and so is unstable and tends to redissolve in the matrix. Only when the new phase has reached the *critical radius*, r*, and has acquired the critical free energy increase, ΔF^*, does the nucleus become potentially stable, so that, after the addition of one further atom, $r > r^*$, it can grow with a continuous decrease of free energy. For a spherical nucleus the values of the critical parameters are given by:

$$r^* = -2\sigma / \left(\Delta F_{\alpha\beta}^v + \Delta F_E^v\right), \tag{10a}$$

$$\Delta F^* = \frac{16}{3} \frac{\pi \sigma^3}{\left(\Delta F_{\alpha\beta}^v + \Delta F_E^v\right)^2}. \tag{10b}$$

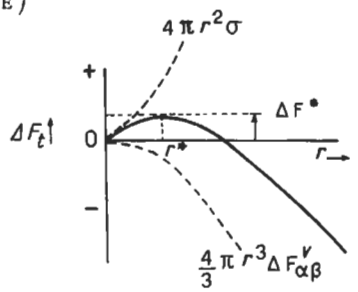


Fig. 7. The nucleation problem; free energy of a sphere of new phase in a supersaturated matrix, as a function of the radius of the sphere.

References: p. 1497.

The required values of $\Delta F_{\alpha\beta}{}^{v}$ for use in these equations can be obtained from $\Delta F_{\alpha\beta}{}^{m}$, for example from eqs. (4), (6c), (6e) (7) or (8), by dividing the value of $\Delta F_{\alpha\beta}{}^{m}$, by the volume of a gram-mole of the new phase, V_m. For most metallic phases, the volume of a gram-mole, containing an Avogado's number of total atoms, of any metallic phase is close to 10^{-5} m^3/g.mole. The problem is to find the appropriate value for $\Delta F_E{}^{V}$; this is discussed in outline § 2.1.2, but for mathematical simplicity for the rest of § 2.1 it will be assumed that the strain term is negligible, as it will be for example in solidification from most metallic melts, at least above the glass transition temperatures In all cases the correct values of r^* and ΔF^* can be recovered by simple addition of ΔF_E^V to $\Delta F_{\alpha\beta}^{v}$.

The rate of nucleation per unit volume, $I_{\alpha\beta}^{v}$, can be obtained from the simple application of statistical mechanics to show that if the critical formation energy of a nucleus is ΔF^* then, at equilibrium, the concentration of critical-sized nuclei, n^* in a unit volume is given by:

$$n^* = N^v \exp(-\Delta F^*/kT), \tag{11}$$

N^v is the number of atomic sites per unit volume on which the assembly of the nucleus could have started, for homogeneous nucleation this is all atom sites. The rate of nucleation is then the product of this concentration of critical nuclei and the rate of atomic addition to the nuclei to make them just supercritical:

$$I_{\alpha\beta}^{v} = N^v A^* \nu \exp(-\Delta F_A/kT) \exp(-\Delta F^*/kT), \tag{12}$$

A^* is the number of atomic sites on the surface of the critical nucleus to which an atom can join after overcoming a growth barrier ΔF_A (§ 2.2.1) and ν is the atomic vibration frequency of an atom in the matrix at the interface, usually assumed to be of the order of 10^{13}/s. In most cases A^* is assumed equal to the total number of atomic sites on the interface but as discussed in § 2.2.3 this may not always be valid and atoms may be able to add to a nucleus only at *ledges* on the interface.

The simple theory above needs to be modified to take into account the expected loss of critical-sized nuclei by their growth, during nucleation, into supercritical regions. This loss reduces the value of N^* to less than the value expected at equilibrium, eq. (11), by a factor, Z, which is usually about 0.05 (RUSSELL [1970]). The near constancy of this correction arises since the rate of formation of critical-sized nuclei and their growth to become supercritical both occur at a similar rate determined by atomic diffusion onto the growing cluster of atoms. Inclusion of a factor of $\exp(-\tau/t)$ to allow the achievement of the expected steady state rate of nucleation, I_s^v, yields the result given by RUSSELL [1970]:

$$\begin{aligned} I_{\alpha\beta}^{v} &= \exp(-\tau/t)\, Z N^v A^* \nu \exp(-\Delta F_A/kT) \exp(-\Delta F^*/kT) \\ &= \exp(-\tau/t)\, I_s^v, \end{aligned} \tag{12a}$$

$$Z = \left\{\Delta F^*/3\pi kT n^{*2}\right\}^{1/2}. \tag{12b}$$

There are many problems with the use of the theory to predict experimentally measured quantities such as the final density of the new regions of the product phase. These

difficulties include the usual lack of knowledge of the interfacial energy σ which plays such a vital part in determining $I_{\alpha\beta}{}^{v}$ ($I_{\alpha\beta}{}^{v}$ is proportional to the exponential of σ raised to the power 3!), the difficulty in calculating the strain energy term in eq. (10b), and the difficulty in calculating how the driving force for the phase change, $\Delta F_{\alpha\beta}{}^{v}$, falls as the reaction proceeds, so that the rate of nucleation can be integrated to give a total number of growing nuclei. Finally, as discussed in § 2.4.2, a significant number of nuclei will dissolve up, by a process of *Ostwald ripening*, during the precipitation reaction itself. These problems are discussed by SERVI and TURNBULL [1966], LANGER and SCHWARTZ [1980], KAMPMANN and WAGNER [1984] and HAASEN and WAGNER [1992]. However, despite these quantitative difficulties the theory is extremely useful in describing a large volume of experimental observations. A major success of this theory is its ability to account for the observation that nucleation increases from a rate that is almost undetectable to rates that are too fast to measure over a narrow range of *undercooling*, ΔT in eq. 4 or *supersaturation*, C_0/C_α in eq. 6d. This is readily shown by use of eq. (11) for a matrix phase change driven by a free-energy function of eq. (4). Under these circumstances the concentration of critical nuclei is given by:

$$\log n^* = 28 - K/\Delta T^2 \tag{13}$$

This arises since the number of atomic sites in a cubic metre is about 10^{28} and K is a constant given by:

$$K = \frac{16}{3}\frac{\pi\sigma^3 T_c^2 V_m^2}{\Delta H_{\alpha\beta}^{v} kT}\,. \tag{14}$$

If the undercooling which gives a density of nuclei of 1 (10^0) m^{-1} is $\Delta T(0)$ and that which gives a density 10^6 times as high is $\Delta T(6)$, then substitution into eqs. (13) shows that:

$$\Delta T(6)/\Delta T(0) = (28/22)^{1/2} = 1.1. \tag{14a}$$

That is, a 10% increase in supercooling will increase the density of critical-sized nuclei, and therefore the rate of nucleation, by a million times. The constant K in eq. (14) can be easily evaluated if the undercooling that gives a measurable rate of nucleation is known, by substitution of a value of N^* of about 10^3 m^{-3} in eq. (13). This approximate value of K allows rather an an accurate estimate of the effective interfacial energy, σ, to be made using eq. (14). In many cases, for example in solidification of pure metals, the determination of the interfacial energy by use of eq. (13) has given reasonable values of this energy even when the undercooling appears to have been significantly underestimated (CANTOR and DOHERTY [1979]). Since eq. 12 also contains the activation energy of diffusion, $\exp(-\Delta F_A/kT)$, the initial increase of nucleation rate with fall of temperature ceases, giving a maximum rate. Below the maximum, the nucleation rate falls rapidly as the temperature falls. This variation of nucleation rate with temperature (together with a similar behaviour for growth) causes the so-called "*C*-curve" behaviour seen in *time–temperature-transformation*, TTT, diagrams, § 2.5.1.

Another major success of the theory of homogeneous nucleation is its ability to

References: p. 1497.

understand a very large range of qualitative experience in which, in many solid state precipitation reactions, the phase that forms is *not* the equilibrium structure. It is frequently found that a metastable phase, characterized by good atomic fit with the matrix and therefore a low value of interfacial energy, σ, is the one that forms. The theory is compatible with this phenomenon since the rate of nucleation is so sensitive to σ; the successful precipitate will be the one that nucleates fastest, and this is usually the precipitate that has the lowest interfacial energy, even though it may not yield the most stable phase. Similarly, nucleation theory accounts for the fact that when a new crystalline phase forms in a matrix of a crystalline parent phase there is almost always a particular orientation-relationship between the phases, see for example BARRETT [1952]. This relationship appears to be the one that minimizes the value of the crucial parameter, the interfacial energy, in eq. 10b. This idea is discussed further in § 2.1.2.

Finally it should be recognized that additional qualitative support for this picture of nucleation is provided by an apparent failure of the model. The failure arises since in many cases nucleation occurs at much smaller undercoolings than expected. In these cases it is found that the new phase forms on some *defect* in the parent phase, for example a dislocation or a grain boundary. The second term in eq. (9) is significantly reduced by the fact that the new phase has consumed some of the existing energy of the defect, thereby reducing the barrier to nucleation. This topic of *heterogeneous nucleation* is discussed in § 2.1.4.

2.1.1. Interfacial structure and energy

A major difference between the interfaces produced by nucleation during solidification from a liquid or glassy phase, and by nucleation in solid-state reactions, is in the possibility of atomic matching across an interface between two crystals. For crystal-crystal interfaces the atomic matching ranges from perfect, in fully *coherent* interfaces, through less perfect matching in *semicoherent* interfaces, to random matching in fully *incoherent* interfaces. The atomic structure of such interfaces is discussed in ch. 9, § 2.3.3.

A simple example of a fully coherent interface is one between two phases that have the same crystal structure, lattice parameter and orientation, giving an orientation relationship often loosely described, for systems with cubic matrix and precipitate, as "cube–cube". For fully coherent interfaces, in a binary AB alloy, the only contribution to the interfacial energy, σ, comes from the higher energy of the unlike A–B bonds across the interface. Since the system is one of limited miscibility between the two phases, it is expected that the unlike bond energy h(AB) will be greater than the average energies of the bonds, $[h(\mathrm{AA}) + h(\mathrm{BB})]/2$, as described in any discussion of the so-called *quasi-chemical* bond model (e.g., SWALIN [1972]).

Such completely perfect matching is unlikely to be found and deviations from coherency can develop as the extreme conditions of perfect matching described above are relaxed. In most reactions there will be a small difference in the lattice parameters of the two phases, a_α and a_β. This gives rise to a misfit, δ:

$$\delta = 2(a_\beta - a_\alpha)/(a_\alpha + a_\beta) \tag{15}$$

For small misfits and very small precipitates, the difference in atomic spacing can be

taken up by elastic strain in the two phases; but with larger precipitates the system can lower the resulting elastic strain energy by concentrating the strain in a cross-grid of edge dislocations spaced a distance d apart (fig. 8). Over the dislocation spacing, d, there are n atomic planes in the phase with the larger parameter and $n+1$ planes in the other phase. When the misfit is much less than unity, then:

$$n = 1/\delta, \tag{16}$$

and

$$d = \left(a_\alpha + a_\beta\right)/2\delta. \tag{17}$$

For such a *semicoherent interface* the interfacial energy has two components; a "chemical" term as discussed for the coherent interface and an additional "structural" term, due to the cross-grid of edge dislocations. For small precipitates, where the diameter of the nucleus may be less than the dislocation spacing, d, the nucleus will remain coherent, though elastically strained, and consequently will have a higher solubility because the elastic energy raises the volume free energy of the precipitate phase (CAHN [1968]).

An incoherent phase boundary, or a phase boundary with poor atomic fit and high energy, will have an atomic structure rather like that of a high-angle grain boundary, whose misorientation is well away from any possible coincidence-site boundary relationships (ch. 9, § 2.2.1.1). Interfaces with poor atomic fit in two phase systems are likely to arise in one of two main ways: The first way is if, after nucleation of the precipitate, the orientation of the surrounding matrix phase is changed by the passage of a high-angle grain boundary in a recrystallization or grain-growth reaction. This phenomenon has been discussed by DOHERTY [1982] with examples of the resulting change in interfacial properties and consequent microstructural changes. The other main origin of incoherent interfaces arises when the two crystals, α and β, do not share a common crystal structure. This is the usual situation in most precipitation reactions and is inevitable in all polymorphous changes. As is discussed in the next section, under these conditions part at least of the interface is almost certain to have very poor atomic

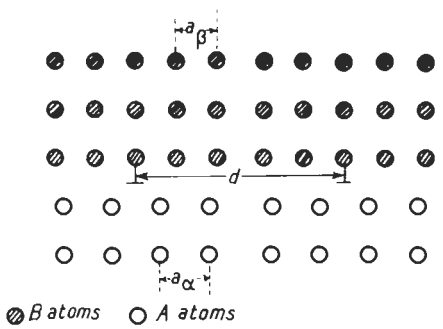


Fig. 8. Semicoherent interface between two phases with the same structure but different atomic spacing and containing a set of edge dislocations a distance d apart.

matching giving an interface that can be regarded as incoherent or at least as having a poor atomic fit and a high specific interfacial energy.

2.1.2. Equilibrium shape

When the nucleus and the matrix have different crystal structures, it is almost invariably found that there is a definite and reproducable orientation relationship between the phases. BARRETT [1952] gave a very complete listing of the orientation relationships for many reactions. The orientation relationship appears to allow a good atomic fit to develop between the two crystals either along a particular plane, WESTMACOTT and DAHMEN [1984], or occasionally along a particular direction, DAHMEN and WESTMACOTT [1986], LUO and WEATHERLY [1987] and DAHMEN [1994]. Typical examples of these orientation relationships are provided by fcc-hcp/matrix-precipitate pairs in systems such as Al–Ag_2Al and in the Cu–Si system. These precipitate matrix pairs show a good fit between (111) in fcc and (0001) in hcp. These precipitates form with these planes parallel and they provide the *habit plane* of the flat plate-like precipitate crystals that form in these reactions (e.g., fig. 24 below). The atomic fit between the two structures is shown in fig. 9, where it can be seen that the two structures, the matrix, α, and the

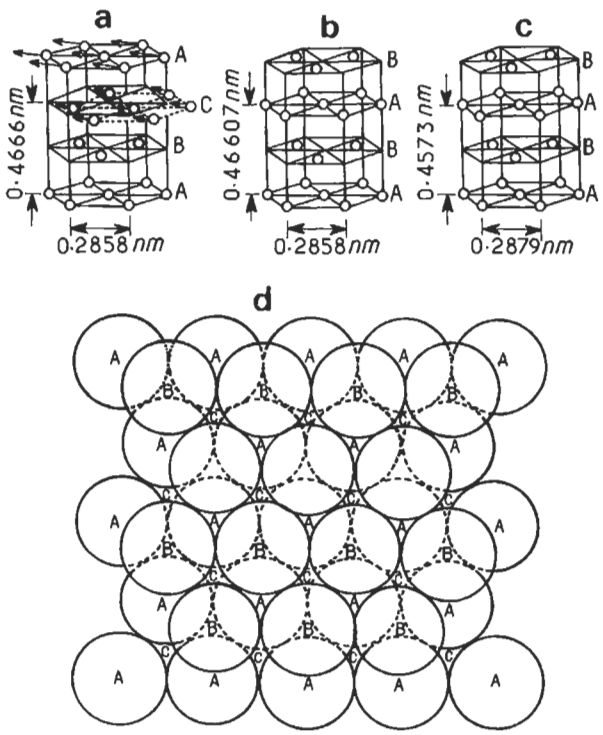


Fig. 9. Atomic structure of fcc–cph matrix–precipitate structures in the Al–Ag_2Al system; (a) the fcc matrix; (b) γ' intermediate precipitate, fully coherent; (c) equilibrium Ag_2Al, partially coherent; (d) the fcc {111} stacking sequence.

metastable precipitate, γ', have identical atomic arrangements on the habit plane, so allowing full coherence in two dimensions. The two equilibrium phases have slightly different lattice parameters but the intermediate metastable precipitate, γ', has the same atomic spacing as the aluminium matrix, in the close-packed plane, and so shows full coherence, and consequently very low energy, for the habit plane (the structural component being zero). Precipitates are, however, three-dimensional and the interface must lie at all possible orientations around the crystal. Consequently, the non-habit plane parts of the interface, the rims of the plate-like crystals, have much poorer atomic matching and higher interfacial energy. In other words, the magnitude of σ will vary significantly around the precipitate. This can be represented in a section through the *σ-plot* as shown in fig. 10. The σ-plot is a radial plot of the interfacial energy as a vector whose length varies as a function of the orientation of the interface of the precipitate. The deep cusps in the plot occur at the orientations of good fit, corresponding to the habit plane. AARONSON *et al.* [1968] calculated that the ratio of energy between the good fit and the rest of the interface is 1 : 2.7 for the equilibrium Ag_2Al phase in Al; similar calculations for the intermediate phase give a higher ratio of 1 : 10 (FERRANTE [1978]). For anisotropic σ-plots such as the one seen in fig. 11, the Gibbs–Wulff theorem (see for example MULLINS [1963] and MARTIN and DOHERTY [1976]) predicts the equilibrium shape of the precipitate, for the orientation relationship that gives the particular anisotropic σ-plot. This shape is the inner envelope of so-called *Wulff planes* that are planes drawn perpendicular to the vector from the origin at the intersection with the σ-plot, fig. 11. The Wulff plane has, of course, the orientation of the interfacial plane whose value of σ is OB. Figure 11 shows as an example the Wulff plane, LBM, for the σ vector, OB, and also shows the equilibrium shape of the precipitate, which provides for facet planes at the cusps, C_1, C_2, C_3 these facet planes are the good-fit interfaces.

The *aspect ratio*, A, of plate like precipitates, defined as the length-to-thickness ratio, is equal to the energy ratio $\sigma_{\text{incoherent}}/\sigma_{\text{coherent}}$. The cusp gives a good-fit *plane* where there is good match of crystal planes as in the fcc-hcp system discussed and in many other precipitation reactions, for example Al-θ' in the aluminium–copper system which has

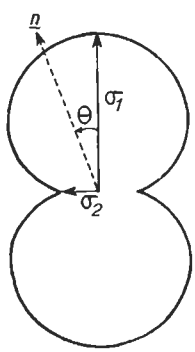


Fig. 10. The σ plot, the variation of interface energy with the orientation of the boundary plane for the particular orientation that allows good fit at the cusp plane, σ_2.

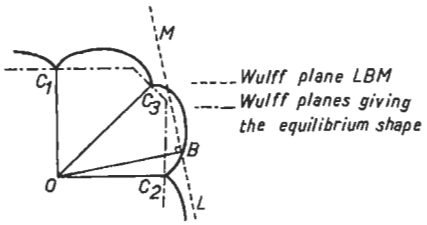


Fig. 11. The Wulff construction. The Wulff plane at B is LBM, that is normal to OB at B. The equilibrium shape is the inner envelope of Wulff phases, in this case, the cusps C_1, C_2 and C_3.

habit planes on the matrix (100) planes (fig. 15, below). For systems like the fcc-bcc system in, for example, the iron–copper system, there is only a good-fit *direction*, which is the close-packed direction <110> in the fcc phase and <111> in the bcc. This form of matching of directions provides the well-developed *needle morphology*, with an aspect ratio of 5 : 1, reported by SPEICH and ORIANI [1965] for the Fe–Cu system. For good-fit directions, the σ-plot, schematically given in two dimensions by fig. 10, must be rotated about the vertical axis, σ_1, to produce the appropriate three-dimensional needle or rod-like form: For plate-like crystals the three-dimensional plot is obtained by rotation about the horizontal axis, σ_2. In other examples the interface between fcc and bcc, for example in steels, remains plate-like, see for example AARONSON *et al.* [1970].

Since the ends of needle crystals, and the rims of plate-like crystals, have poorer atomic fit than the coherent needle axis or habit planes of the plates, it is commonly assumed that these ends and rims are fully incoherent. This may be usually the case but it need not necessarily be true since in the fcc-hcp pair Al–Ag_2Al it has been seen that the rims of the plates are also faceted, FERRANTE and DOHERTY [1979]. This implies that the rim is also coherent, though with a poorer fit made up by the cores of Shockley partial dislocations, HOWE *et al.* [1985]. CHRISTIAN [1994] shows that a precipitate in a matrix cannot be fully or partially coherent on one part of the interface and incoherent elsewhere so the rims of all plate-like precipitates should be considered coherent but to have a high density of dislocations, giving high energy and mobility.

It is usually assumed during nucleation that the nucleus forms with the minimum interfacial energy, since any other assumption increases the free-energy barrier to nucleation and strongly reduces the rate of nucleation. The assumption of minimum interfacial energy requires the nucleus to have the *equilibrium shape*, this will be plate-like or needle-like if the σ plot shows the normal anisotropy for two crystals with different crystal lattices. If the plot is, however, fully isotropic, the expected shape of the nucleus will be a sphere. In two examples of "cube–cube" relative orientations between crystals with the same structure, nearly spherical equilibrium shapes have been reported. In the bcc-bcc beta–gamma brass system studied by STEPHENS and PURDY [1975] and in the all-fcc system of Ni–Ni_3X (where X can denote a mixture of several atoms, usually Al or Ti) (see RICKS *et al.* [1983]), equilibrium shapes that were nearly spherical were reported. In the latter case, it seems that the precipitate develops the usually reported

cuboid morphology only with growth, and in systems when there is some small elastic misfit, δ. Both the brass and the nickel systems can evolve into a dendritic morphology during growth under condition of very low supersaturation but this is a growth form not an equilibrium shape! The shape instability leading to dendritic growth is discussed in § 2.5.

For plate-like precipitates with an equilibrium value of the aspect ratio A, that is, with $A = \sigma_1/\sigma_2$, the critical values of the plate radius, $a(y)^*$ and nucleation barrier, ΔF^*, are found (see AARONSON *et al.* [1978]) to be given by:

$$a(y)^* = -8\delta_1 / \left\{3\left(\Delta F^{v}_{\alpha\beta} + \Delta F^{v}_{E}\right)\right\}, \tag{18}$$

$$\Delta F^* = 256\pi\sigma_1^3 / \left\{27A\left(\Delta F^{v}_{\alpha\beta} + \Delta F^{v}_{E}\right)\right\} \tag{19}$$

σ_1 is the interfacial energy of the poor fit rim and σ_2 that of the good fit habit plane.

2.1.3. Strain energy

Strain energy plays a vital role in the nucleation of solid-state phase changes, as shown by eqs. (9)–(19). There appear to be two different types of strain that can be developed by the formation of a new phase in a matrix phase. The first type of strain occurs only when both phases are crystalline and is that caused by a misfit, due to different lattice parameters — eq. (15), between two coherent phases. The second type of strain is that caused when the new phase occupies a different volume from the region of the matrix that it has replaced. This second type of strain can arise in various ways. One of these is when the volume per atom is different in the two structures and the precipitate grows without a change in the number of atom sites. An example would be the nucleation of the less dense bcc ferrite from austenite in pure iron. A further method for the development of this volume strain is if there is a significant difference in the rates of diffusion of two components of the alloy. An example is when zinc-rich γ-brass precipitates from β-brass when there is a more rapid inward flux of the faster diffusing zinc atoms than the compensating outward movement of copper atoms. This increases the *number of atomic sites* in the region where γ-brass is growing, giving a dilational strain. This effect has been directly observed on the surface of a brass alloy, by CLARK and WAYMAN [1970] who showed that the γ-phase stood above the surrounding surface. In the interior of the sample this displacement would produce an elastic strain unless relieved by plastic deformation. A similar increase in volume arises during growth of Fe_3C in Fe by rapid interstitial diffusion of carbon upto the growing precipitate.

A local strain around a precipitate will produce a large energy increase if the strain is taken up *elastically* since the elastic moduli of metals are large. For aluminium where Young's modulus is 70 GPa, a 1% strain would give an elastic energy of 3.5 MJ/m^3 (35 J/g-mole), a 5% strain would give an energy 25 times larger. In the case of a coherent interface with a finite δ, the long range elastic distortions can be reduced *plastically* by the introduction of dislocations into the coherent interface, but at the cost of raising the interfacial free energy. For the second, dilational, type of strain the elastic energy can be relaxed either plastically, by dislocation motion — usually the glide of prismatic edge

dislocation loops, or by the diffusional motion of vacancies. Both of these release mechanisms, especially the diffusional motion of vacancies, occur more easily at high temperatures.

Two limiting conditions are of interest to consider. The first condition would be for precipitation at low temperatures where there will usually be a high free-energy driving force (eq. 6d), since the equilibrium solubility C_α is usually very small at low temperatures. The second condition occurs for high-temperature reactions, that occur with much lower driving forces. With small driving forces then the elastic strain energy opposing nucleation can easily be larger than the free-energy decrease driving the reaction. For example, precipitation of $AlCu_2$ from an aluminium-copper alloy at 800 K with a typical high-temperature supersaturation ratio $C_0/C_\alpha = 1.1$, gives a $\Delta F^{m}_{\alpha\beta}$ of only about 200 J/g-mole which would be offset by an elastic strain of only 2.5%. For high temperature precipitation it is likely that the volume strain must be accommodated by diffusion, or by plastic accommodation. Both of these processes become faster at higher temperatures. For low-temperature precipitation-hardening reactions, however, the supersaturation ratios are very much larger, so the elastic strain energy is more easily provided by the large free-energy change driving the precipitation. Electron microscopy has shown that such elastic distortion does indeed occur under these low-temperature, high-supersaturation conditions (see KELLY and NICHOLSON [1963]). The resulting elastic strains are a significant part of the strengthening response of such precipitation.

NEMOTO [1974] found, by direct observation using in-situ high-voltage electron microscopy, that during dissolution of cementite by decarburization of an Fe–C alloy at 700°C, with high rates of dissolution, part of the change of volume of the reaction was provided by intense localized dislocation motion. For slower rates of dissolution such dislocation movement was not observed and the strain was assumed to be accommodated by movement of iron vacancies. The analysis of the coupled diffusion of substitutional iron atoms and rapid interstitial carbon atoms to relax localized dilational strain has been provided by ORIANI [1966] for the case of precipitate coarsening in the same, iron-cementite, system. A further example was provided by MAKENAS and BIRNBAUM [1980] who reported that there was very significant plastic deformation during hydride precipitation in the niobium-hydrogen system. They charged hydrogen electrolytically into previously prepared thin foils of niobium and observed precipitation by TEM as the foils were continuously studied during cooling. Precipitation of hydride occurs by rapid interstitial diffusion of hydrogen causing a large expansion ($\Delta V/V$ of 18%) at temperatures too low for diffusional relaxation of the metallic atoms. Around each precipitate there was an intense tangle of dislocations with additional prismatic dislocation loops punched out in the <111> slip directions. During resolution of the hydrides, the plastic strain was not reversible and this led to a large temperature hysteresis between precipitation and resolution. Similar examples of plastic deformation during other low-temperature hydride precipitation reactions were reviewed by MAKENAS and BIRNBAUM [1980]. It would be of interest to know if the interfacial change from the metal structure of niobium to the hydride structure during the low temperature hydride precipitation were civilian or military interface reactions, § 2.6.

LEE *et al.* [1980, 1983] reviewed the current theoretical and experimental results on

the influence of elastic and plastic strains on precipitation in general, and nucleation in particular. For the elastic situation the detailed theoretical analysis shows that the elastic energy is minimized when the precipitates are formed with a thin disc- or plate-like shape. This is the morphology revealed by electron-microscopic studies in all cases where Guinier–Preston (GP) zones form when the solute in the zone has a different atomic size from the matrix atoms that the solute replaces. Disc-like zones are found for example in aluminium-copper and copper-beryllium alloys where there are significant size differences between the atoms. However in systems like aluminium-silver, where the atoms are almost the same size, spherical zones are found (KELLY and NICHOLSON [1963]). The shapes found for the intermediate and equilibrium precipitate in the Al–Ag system are plate-like — but this is due to the fact that the later precipitates have a different, hexagonal, structure from the fcc matrix. The observations on GP zones are therefore much more relevant to understanding the influence of elastic strain on precipitate shape than are observations on shapes of metastable precipitates, where there is a complication due to differences in precipitate crystal structure. Finally it should be noted that, as expected, the habit plane of misfitting disc-like zones appears to be in the matrix direction with the lowest elastic modulus, <100> in most cubic metals (WERT [1976]).

LEE *et al.* [1980] also discussed the expected *plastic* yielding that can occur around a spherical nucleus under the assumption of isotropic elasticity. They assumed that for precipitates larger than about 1 μm the macroscopic yield stress will operate. For smaller precipitates, however, it is likely that the yield stress will be raised by the lack of dislocations in the sub-micron regions. By use of a model proposed by ASHBY and JOHNSON [1969], in which dislocations can be nucleated with the help of a supersaturation of point defects, LEE *et al.* [1980] were able to derive an expected yield stress for flow, as a function of the precipitate radius. Their analysis gave results that agreed well with a range of experimental results, as regards the critical size of precipitate which went from coherent, elastically strained, to semi-coherent, plastically relaxed.

2.1.4. Heterogeneous nucleation

In many examples of nucleation in solids it is found that the nucleation sites are not distributed randomly, but are concentrated at particular sites in the matrix, usually at some type of crystal defect. The defects that act as the sites for *heterogeneous nucleation* include grain boundaries, dislocations, stacking faults or the interfaces of previously formed precipitates. These sites were described in some detail by NICHOLSON [1970]. CAHN [1956, 1957] provided the currently accepted analysis for nucleation of incoherent precipitates on grain boundaries and on dislocations. The analysis is based on the idea that the nucleus grows on the pre-existing defect so that the energy of formation of the nucleus is reduced by that proportion of the defect energy that is consumed by the nucleus as it forms. For nucleation on the boundary between two matrix grains (fig. 12) the grain boundary that has been destroyed is shown as the dashed line within the new crystal. The critical parameter is the semi-angle θ determined by the ratio $\sigma_{\alpha\alpha}/2\sigma_{\alpha\beta}$, where $\sigma_{\alpha\alpha}$ is the grain-boundary energy and $\sigma_{\alpha\beta}$ is the incoherent energy of the interface between β precipitate and the α matrix:

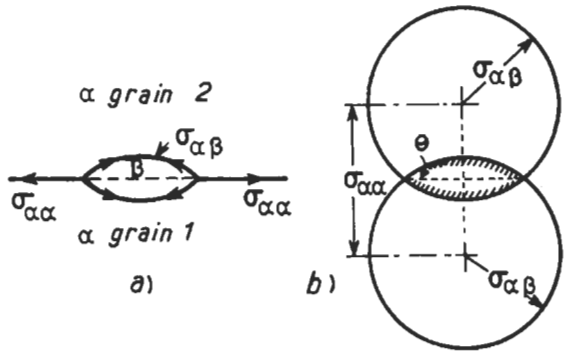


Fig. 12. (a) The double hemispherical cap model for the expected shape of an incoherent precipitate at an α–α grain boundary showing the destroyed α–α grain boundary as the dashed line. (b) The Wulff construction for this situation. (After LEE and AARONSON [1975]).

$$\cos\theta = \sigma_{\alpha\alpha}/2\sigma_{\alpha\beta}. \tag{20}$$

When the new phase "wets" the grain boundary, which occurs when $2\sigma_{\alpha\beta} < \sigma_{\alpha\alpha}$, giving $\cos\theta = 1$, then there is no barrier to nucleation. This often arises when the second phase is a liquid, as occurs for example in the *hot shortness* of steels in the presence of molten iron sulphide, and in other cases of *liquid-metal embrittlement*. In general, however, grain boundaries in metals are not "wetted" by crystalline second phases and so some fraction of the barrier to homogeneous nucleation remains in the case of heterogeneous nucleation: The barrier is given by:

$$\Delta F^*(\text{het}) = \Delta F^*(1 - \cos\theta)^2(2 + \cos\theta)/2. \tag{21}$$

For nucleation at grain *edges* (the lines where three grain boundaries meet) and at grain *corners* (where four grains, and four grain edges, meet) there are an even larger reductions in the barrier to nucleation (CHRISTIAN [1975]). Such reductions in the barrier to nucleation do not automatically lead to a large increase in the nucleation rate, since there is also a dramatic reduction in the number of suitable sites for nucleation to start [the term N_v in eq. (11)]. If the mean grain diameter is d_g, the thickness of the grain boundary is δ_b and the volume of an atomic site is V_a then the number of nucleating sites *per grain* declines approximately as d_g^3/V_a: $d_g^2\delta_b/V_a$: $d_g\delta_b^2/V_a$: δ_b^3/V_a: as we consider homogeneous nucleation, grain-boundary nucleation, grain-edge and grain-corner nucleation, respectively. For a grain size of 50 μm and a boundary thickness of 0.5 nm, the ratios are 10^{15}: 10^{10}: 10^5: 1. These are rather significant reductions in N_v. The result of the two effects, the reductions in both ΔF^* and in N_v, is that for nucleation under a large driving force and with a small interfacial energy, homogeneous nucleation is expected to dominate, giving a high density of grain interior precipitates. With smaller driving forces or with higher values of the interfacial energy (especially if no good atomic matching between the two phases is possible) then the consequent large value of Δf^* in eq. (11) requires heterogeneous nucleation, at suitable sites for any significant rate

of nucleation to be possible. This conclusion is in very good qualitative agreement with a huge amount of experimental evidence for the sites of nucleation in many systems (NICHOLSON [1970]). It is always found, for example, that with a small supercooling into a two-phase field giving a low driving force, nucleation occurs solely at grain corners and edges and only at somewhat higher driving forces is nucleation found at grain-boundary sites and eventually at grain-interior sites. Moreover, it is very commonly found that the grain-interior nucleation is not homogeneous, since dislocations are also strong heterogeneous nucleation sites (CAHN [1957]). It appears that only for very large driving forces indeed, such as by diffusion of a third element into a binary alloy to give, for example, internal oxidation or nitriting (with a solute whose oxide or nitride has extremely low solubility) can homogeneous nucleation be expected to occur as the fastest process for largely incoherent interfaces. Even here it is likely that nucleation on dislocations will often dominate. Only for nucleation at densities that correspond to precipitate spacing significantly less that the typical dislocation spacing of about 1μm (10^{18} precipitates/m^3) is homogeneous nucleation probable. Heterogeneous nucleation at closer spacings is still possible either by closely spaced precipitates strung out along dislocation lines or by some type of autocatalytic nucleation such as has been reported for θ' in Al–Cu (see next section) by PEROVIC *et al.* [1981] where the elastic strain of one precipitate plate facilitates the nucleation of a closely spaced neighbour.

In most cases there is insufficient driving force for homogeneous nucleation except with largely coherent interfaces, as in GP zone formation. Most intermediate precipitates, even with a good fitting plane or direction, appear to nucleate either on dislocations or, in some two-stage heat treatments, on pre-existing GP zones (LORIMER and NICHOLSON [1969] and JACOBS and PASHLEY [1969]. Precipitation of a third phase, γ, at the interface of a matrix, α, and an existing precipitate, β, is very similar to that at an $\alpha\alpha$ grain boundary, fig. 12, except that the interfacial energies, $\sigma_{\alpha\gamma}$ and $\sigma_{\beta\gamma}$, are not equal, so there are two angles, θ_1 and θ_2, given by eqs. 20a and 20b.

$$\sigma_{\alpha\alpha} = \sigma_{\alpha\gamma} \cos\theta_1 + \sigma_{\beta\gamma} \cos\theta_2 , \tag{20a}$$

$$\sigma_{\alpha\gamma} \sin\theta_1 + \sigma_{\beta\gamma} \sin\theta_2 . \tag{20b}$$

The classical theory of heterogeneous nucleation, outlined above, assumed that the energy of the interface between the two phases, $\sigma_{\alpha\beta}$, has a constant value. This is a reasonable assumption for systems where there is no significant atomic matching but this is only likely if one of the two phases is non-crystalline, for example the internal oxidation of copper-silicon to precipitate amorphous silica. In most precipitation reactions, two crystalline phases are involved with a very high likelihood of some atomic matching. It is not known how great difference this can make to the theory of nucleation on dislocations but for nucleation on grain boundaries it is known that a significant difference from the classical theory is produced. This arises since, if the precipitate has an orientation relationship giving good atomic fit with one crystal, then it is very unlikely to have any atomic matching with the other grain at a grain-boundary nucleation site. The assumption of a constant interfacial energy is then invalid.

The situation of grain-boundary nucleation with lattice matching has been discussed

References: p. 1497.

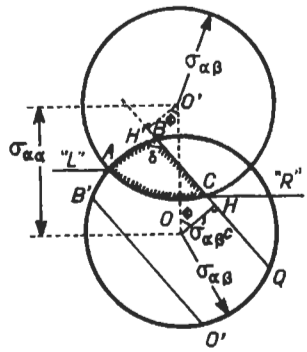


Fig. 13. The Wulff construction for a precipitate with a low-energy coherent interface BH, $\sigma_{\alpha\beta}$c, between the grain-boundary precipitate and the upper grain at a particular angle ϕ between the habit plane and the boundary plane (after LEE and AARONSON [1975]).

by LEE and AARONSON [1975], who developed a modification to the Gibbs–Wulff theorem for grain-boundary nucleation (figs. 13 and 14). They showed that the expected shapes are very different from that of fig. 12 and consequently the effectiveness of grain-boundary nucleation sites is very different from that given by eq. (21). Lee and Aaronson then found that the barrier to nucleation is reduced still further, by factors of upto 100, depending on the ratio $\sigma_{\text{coherent}}/\sigma_{\text{incoherent}}$ and the orientation of the facet plane with respect to the grain-boundary plane. The largest reduction occurs when the facet plane is nearly parallel to the grain boundary, with the angle ϕ tending to 0° (fig. 13).

For heterogeneous nucleation of a β phase on $\alpha\alpha$ grain boundaries, where in almost every case there is no orientation relationship between the α grains, then the new β phase will only have a low energy coherent or semicoherent interface with *one* of the α grains. With the other grain (or grains, if the nucleation site were a three grain edge or a four grain corner) the interface will be a high energy incoherent interface. This distinction between the two types of interface for grain boundary precipitate is not only important for nucleation but is also critical for the subsequent *growth* of the precipitates, § 2.2, 2.2.3, 2.2.4 and 2.6. In some cases, e.g., the grain boundary "allotriomorphs" or the two-phase, pearlite, structure (§ 2.5.1]) in Fe–C alloys, growth is into the grain with which there is no orientation relationship. Here, the interface is high energy but fully

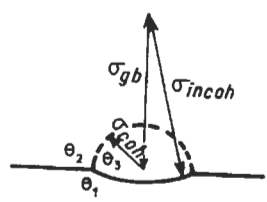


Fig. 14. The equilibrium shape for a precipitate "cube–cube" oriented with the upper grain, with a small value of the coherent energy, σ_{coh}, and a larger value of the incoherent energy, σ_{incoh} (after DOHERTY [1982]).

mobile for a reconstructive interface reaction. In other cases the growth is into the grain with the good fit orientation relationship giving anisotropic energy and mobility and the possibility of some type of displacive interface reaction, § 2.6. This type of growth usually produces Widmanstätten precipitate shapes: plates, laths or needles.

2.1.5. Experimental studies of nucleation

Despite the very extensive range of experimental investigations on precipitation processes in metals, as described, for example, by RUSSELL and AARONSON [1978], there are only a few investigations that have attempted to test the various theories of nucleation *quantitatively*. As previously mentioned, however, there are numerous experimental observations that are in agreement, at least in outline, with the main predictions of the theories. Amongst the successful *qualitative* observations are the following:

(i) There is almost invariably an *orientation relationship* between the a crystalline precipitate and the parent crystal lattice, and this orientation relationship is apparently always one that allows for good atomic matching between the two crystals, so that the interfacial energy will be minimized at least on certain planes.

(ii) At high supersaturations it is commonly found that a *metastable* precipitate structure is formed rather than the stable phase. In every case where this has been studied, the metastable phase has better atomic matching with the matrix than the stable phase has. This has been discussed already for the aluminium–silver system but is perhaps better known for the aluminium–copper system (fig. 15). The phases that are precipitated before the equilibrium θ-phase, $CuAl_2$, include the disc-like GP zones and the tetragonal θ'. The GP zones are fully coherent and the θ' is fully coherent with the aluminium matrix on the habit plane where the basal plane of the precipitate matches the {100} planes of the matrix. This type of pattern of precipitation is found in a wide range of metallic and nonmetallic solid-state precipitation reactions, as reviewed by RUSSELL and AARONSON [1978].

(iii) Owing to the strong barrier to *homogeneous nucleation* provided by the interfacial energy and the potency of many heterogeneous nucleation sites in solids, homogeneous nucleation is only found with either extremely high free-energy driving forces or with precipitates with very low values of the interfacial energy. Typical examples of homogeneous nucleation are found for the precipitation of GP zones or intermediate phases which have the same crystal structure as the matrix. In addition to the many different GP zones formed in precipitation-hardened aluminium alloys (LORIMER [1978]), other examples are fcc cobalt precipitated in copper–cobalt alloys (SERVI and TURNBULL [1966], LEGOUES and AARONSON [1984], WENDT and HAASEN [1985] AL–KASSAB [1987]), ordered γ'fcc precipitates of Ni_3Al in nickel–aluminium alloys (KIRKWOOD [1970], HIRATA and KIRKWOOD [1977] and WENDT and HAASEN [1983]) and ordered Cu_4Ti in copper–titanium alloys (VON ALVENSLEBEN and WAGNER [1984]). These three systems will be discussed in more detail here, since they provided *quantitative* tests of the kinetic theory of homogeneous nucleation. Further discussion of these reactions is given in § 2.4.2, where the systems provide excellent tests of models of the *coarsening* reaction.

(iv) As described in § 2.1.4., the variation of *heterogeneous nucleation* sites with supersaturation seems also to be in good agreement with the predictions of the classical theory of heterogeneous nucleation, but with the exception of Aaronson's investigation of

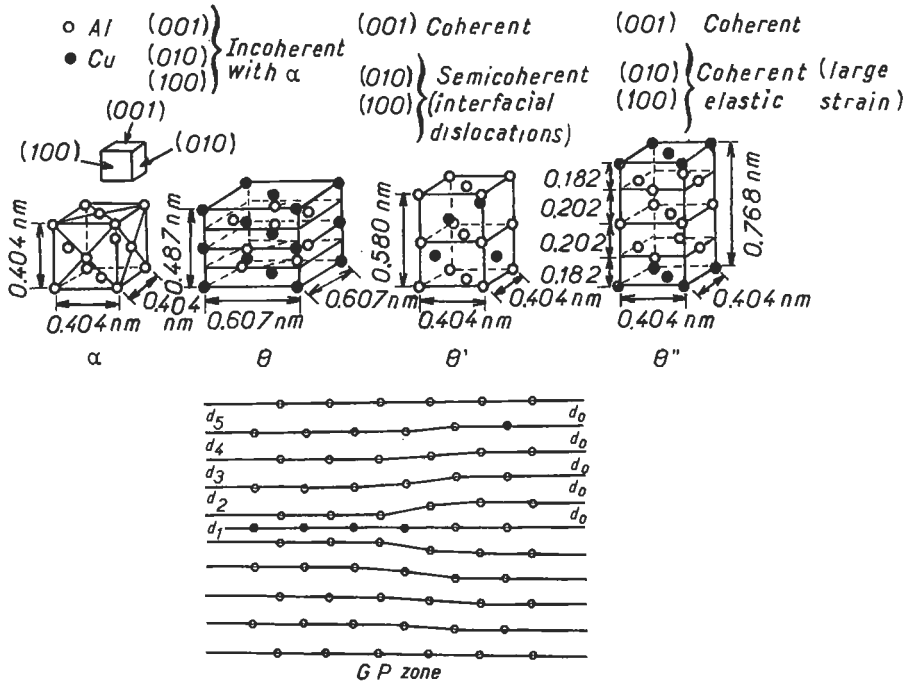


Fig. 15. The atomic fit between fcc aluminium and the equilibrium phase θ and the various intermetallic phases θ' and θ'' and the GP zones in Al–Cu system (after HORNBOGEN [1967]).

grain-boundary nucleation (AARONSON [1979]) there seems to have been almost no published quantitative tests of the theory for grain-boundary nucleation or for nucleation on dislocations.

The classic test of *homogeneous nucleation* theory in metallic solids was provided by SERVI and TURNBULL [1966]. They studied the precipitation of fcc cobalt from copper by an in-situ technique: electrical resistivity measurements on supersaturated copper-cobalt alloys that had been directly quenched to the precipitation temperature after solution treatment. The overall kinetics of the reaction, observed by electrical resistivity changes, were close to those expected for a nucleation-and-growth reaction with nuclei forming only in the earliest stages of the reaction. The bulk of the transformation appeared to show diffusion-controlled growth of a near-constant number of nuclei. Servi and Turnbull reported however that the agreement with the predictions of this model was not exact. As described in various reviews (for example, TURNBULL [1956]), the fraction $f(x)$ of the supersaturated solute that remains in the solution during diffusion-controlled growth with a constant number of growth centres is given by*:

$$f(x) = 1 - \exp(-t/\tau)^n, \tag{22}$$

* It should be noted that the use of eq. 22 is only an approximation for a precipitation reaction. This is discussed further in § 2.4.

the exponent, n, is expected to be 3/2 and the time constant, τ, is determined by the diffusion coefficient, phase diagram solubilities and the density of the growing nuclei.

Over a range of cobalt contents, from 1.0wt% to 2.7wt% Co, the observed value of the exponent n fell from 1.95 to 0.97. The investigators ignored this discrepancy and calculated the density of nuclei in each case from the measured time constant, τ. From these calculated precipitate densities, which were directly confirmed by electron microscopy (TANNER and SERVI [1966]), Servi and Turnbull were able to determine the *initial* rates of nucleation. They were then able to show that these experimental nucleation rates agreed rather closely with the predictions of steady-state nucleation theory [eq. (12)]. The observed apparent interfacial energy was found to be 0.20 J/m^{-2} which compared very closely to a calculated value, based on the quasi-chemical bonding model, of 0.23 J/m^{-2}. They reported a difference in the calculated and observed pre-exponential terms in eq. (12) of a factor of 100, the observed value being smaller than the calculated value. Two corrections to this have been found — a misplaced π in Servi and Turnbull's equation 2 and the incorrect insertion of the atomic fraction of cobalt in the value of N_v, eq. (11), (RUSSELL [1970]). The effect of these corrections is that the steady-state theory apparently underestimated the rate of nucleation by a factor of about 10^5. LEGOUES and AARONSON [1984] studied the nucleation rate in Cu–Co by conventional TEM observations. They showed that their *data* closely matched that of Servi and Turnbull but they claimed to have largely eliminated the significant error found by Servi and Turnbull by a recalculation of the solution thermodynamics and as a result reasonable agreement with classical nucleation theory, eq. (12a) was reported. WENDT and HAASEN [1985] studied nucleation in the same system but at much higher supersaturations using the ultra high resolution technique of field ion microscopy with atom probe chemical analysis (FIM-AP). They reported much lower rates of nucleation (by about 10^5) than predicted by the classic equations but they noted that at the supersaturation they used, the critical size of the nucleus was only about 0.2 nm and ΔF^* was only about kT so that, in such circumstances, the use of macroscopic parameters such as the interfacial energy is unlikely to be valid. HAASEN and WAGNER [1992] have described a new cluster dynamics model for simulating nucleation in Cu–Co in which even at the supersaturations of the early work, comparable to that of Servi and Turnbull, it was found that the standard homogeneous nucleation theory significantly overestimated the total nucleation density expected. However the discrepancy found was only a factor of 30, which for nucleation theory is not a very significant error! It will be of great interest to see the results from future applications of this cluster dynamics method to the modelling of nucleation, since a true atomic model for the formation of critical sized clusters will avoid the limitation of using macroscopic thermodynamics for a process occurring on an atomic scale.

Several tests of nucleation theory have been carried out in the nickel–aluminium system. These include FIM at very high resolution but with low statistical accuracy by WENDT and HAASEN [1983], small angle neutron scattering studies on the same samples by BEDDOE *et al.* [1984] and high resolution TEM studies by XIAO and HAASEN [1991]. These studies took into account the potential loss of precipitates that have nucleated, by

References: p. 1497.

the process of Ostwald ripening (§ 2.4.2) and the data were reviewed in detail by HAASEN and WAGNER [1992]. The earlier studies, whose results are described in § 2.4.2 as figs. 42 and 43, were essentially all obtained when nucleation was complete and the precipitate densities were *falling* due to coarsening. KAMPMANN and WAGNER [1984] (see HAASEN and WAGNER [1992] for details) discussed these results using the model of LANGER and SCHWARTZ [1980] which combined nucleation, growth and coarsening and they also developed an new explicit numerical model for analysing the combined reaction, a model based on eq. (12a) for nucleation. A very reasonable match to the experimentally observed particle densities was obtained with fitted values of the interfacial energy and diffusion coefficient that were also very reasonable (σ of 0.016 J/m^{-2}, and D of 5×10^{-21} m^2s^{-1} at 823 K).

More recent studies by XIAO and HAASEN [1991] gave insight into the earliest stages of precipitation in Ni–Al using high resolution TEM. Their study was for a lower alloy composition, 12at% Al, and at a lower temperature, 773 K. They observed very small ($r\approx 1$ nm) ordered regions that appeared to remain essentially unchanged for a few hours until replaced by larger ($r>2$ nm) steadily-growing precipitates. The latter achieving a maximum density of 10^{24} m^{-3} after about 40 hours. The small particles were assumed to be sub-critical sized "embryos" that should be continuously forming and dissolving while the large particles were the stable, super-critical, precipitates. XIAO and HAASEN [1991] thus appear to be the first to report direct observations on embryos whose predicted existance is essential for the nucleation model. So their direct sighting of these embryos was a rather important observation! The mean radius of the later and larger supercritical precipitates remained essentially constant while their *density* increased. At longer times, $t\geq 40$ hours, the density fell and the mean radius of the precipitates, raised to the power 3, increased *linearly* with time. These growth kinetics are fully characteristic of interfacial energy driven coarsening, § 2.4.2. For the time period from 2 to 20 hours while the density of supercritical precipitates was *increasing*, Xiao and Haasen were able to measure directly a steady state nucleation rate, I_s^v of eq. (12a) of 8×10^{19} nuclei $m^{-3}s^{-1}$, close to the value of 1×10^{19} $m^{-3}s^{-1}$ predicted for the values of σ (0.014 J/m^{-2}) and D of (1.6×10^{-21} m^2s^{-1} at 773 K) obtained from analysis of the coarsening data. The value of interfacial energy gives a predicted critical radius r^* of 1.7 nm matching quite closely the apparent critical radius that divided the ordered solute rich embryos from the growing precipitates. This agreement is very satisfactory — however the extreme sensitivity of nucleation rate to interfacial energy in eq. 12 suggests that predicting a rate matching the theory to better than an order of magnitude error is not very likely.

HAASEN and WAGNER [1992] discussed the application of the numerical modelling technique of KAMPMANN and WAGNER [1984] to the data on nucleation in Cu–1.9at% Ti and as in the analysis of the earlier Ni–12at% Al data they managed to get a very good fit to the data on nucleation and coarsening. For Cu–Ti, the fitted values of the interfacial energy and the diffusion coefficient were 0.067–0.07 J/m^{-2} and 2.5×10^{-19} m^2s^{-1} at 623 K.

An earlier successful test of nucleation theory was reported by STRUTT [1974] for nucleation of a coherent metastable phase in the LiF–Mgf_2 ionic system studied by ionic

conductivity. RUSSELL [1970] described a similar successful test of nucleation theory, in a precipitation reaction in an oxide glass system. The theory of homogeneous nucleation has also been well tested for metallic solidification (chapter 8, § 4.3; see also CANTOR and DOHERTY [1979]). It would appear that despite early reports, for example by HEADY and CAHN [1973], of an apparent marked failure of the theory in a case of liquid unmixing, the model of homogeneous nucleation is in essentially very good shape quantitatively as well as qualitatively as discussed earlier. LANGER and SCHWARTZ [1980] were able to show that the basic theory of nucleation remained valid for liquid unmixing. In the unmixing experiments in liquids close to the unmixing temperature the experimental problem appeared to be that of actually observing an initial nucleation rather than the onset of a significant amount of unmixing. Observable unmixing, the onset of "cloudiness", involves diffusion close to the unmixing temperature and this is an inherently slow process. Optical scattering only becomes significant for droplets sized and/or spaced apart by distances of the order of the wavelength of light, 0.5 μm.

The only reports, known to the author of this chapter, on the quantitative testing of the theory of *heterogeneous nucleation* are those of LANGE [1979] as presented by AARONSON [1979] and of PLICHTA *et al.* [1980]. The first investigation was of the grain-boundary nucleation rate of ferrite from austenite in high-purity iron–carbon alloys. Considerable care was taken to distinguish grain-edge from grain-boundary nucleation: only data for nucleation on the boundaries between two austenite grains were considered. In order to make a comparison with the experimental results, an equilibrium shape for the ferrite nucleus on the grain boundary was needed. The only model that was at all successful was a pill-box nucleus shape for ferrite on the boundary between two austenite grains, where all the interfaces were partially or fully coherent. It was reported that all other nucleus shapes predicted nucleation rates orders of magnitude smaller than the rates observed. As Aaronson has pointed out, a pillbox model requires the habit plane of the precipitate to be accurately parallel to the appropriate plane in both the austenite grains, a rather unlikely situation. However, since austenite is unstable in the iron–carbon system, room-temperature crystallographic studies to test this hypothesis could not be carried out.

The other study on heterogeneous nucleation kinetics was the observation by PLICHTA *et al.* [1980] of nucleation of the massive transformation in two Ti alloys. In the *massive transformation*, discussed further in § 2.2, the driving force for nucleation is determined from the under-cooling, ΔT, below the temperature, usually described as T_0, where two phases with the same composition have the same free energy. As in the earlier study in iron, the spherical cap model underestimated the observed nucleation rates by, in this case, thousands of orders of magnitude! Agreement could again be achieved only by use of the pillbox model with a low-energy coherent interface, again with only a small fraction of potential grain-boundary nucleation sites appearing to be active.

Two ideas seem to follow from these two sets of very interesting results. Firstly, there is the need for experiments on a system where the orientations of both phases can be studied after precipitation in a system with reliable thermodynamic data, for example precipitation in a dilute alloy. Secondly, there is the possibility that heterogeneous nucleation at grain boundaries may not depend on the general grain-boundary structure

References: p. 1497.

but may occur mainly at grain-boundary defects such as the intrinsic and extrinsic dislocations known to exist in metallic grain boundaries. It has been reported, for example by DESALOS *et al.* [1981], that after controlled rolling of low-alloy steels there is an enhanced density of ferrite grains nucleating on the austenite grain boundaries (fig. 16). This strongly indicates that nucleation is enhanced by dislocations absorbed on austenite grain boundaries.

PLICHTA and AARONSON [1980] attempted to provide an answer to the first problem by observing the orientation relationships for boundary-nucleated massive precipitates in Ag–26at% Al which has the same bcc–hcp pair of phases as in the Ti alloys studied. Plichta and Aaronson showed that 46 out of 47 boundary precipitates did indeed have the "Burgers" relationship, $(0001) /\!/ \{110\}$, $[1120] /\!/ \langle 111 \rangle$, with one of the two matrix grains, but rarely with both. They were able to provide evidence that for about half of the non-Burgers-related orientations some indication of good atomic fit was possible. Unfortunately no nucleation-rate experiments were carried out in the crystallographically investigated system.

Additional evidence for the failure of the classical model of boundary nucleation has been described by BUTLER and SWANN [1976]. They measured, by TEM, the density of precipitates heterogeneously nucleated on grain boundaries in Al–Mg–Zn alloys. They studied precipitates forming under higher driving forces, giving much higher precipitate densities than studied in the observations of LANGE [1979] and PLICHTA *et al.* [1980]. Butler and Swann showed direct evidence for preferential nucleation on defects at boundaries, notably at grain-boundary steps (fig. 17) and at dislocations in low-angle boundaries. It is therefore rather likely, in experiments such as those described by LANGE [1979] and PLICHTA *et al.* [1980]), that the precipitates seen may indeed have come from *defects on the grain boundaries* and that this could account for the high rates of nucleation seen. Other aspects of Butler and Swann's observations were in qualitative agreement with the ideas of LEE and AARONSON [1975]. Butler and Swann showed that the density of precipitates varied from boundary to boundary and along an individual boundary when the boundary plane varied. The latter observation appears to be due to the change of angle ϕ between the boundary plane and the precipitate habit plane, fig. 13. In some cases, as the boundary plane varied a different precipitate orientation was found. In most TEM studies of grain-boundary precipitation, as for example that of WILLIAMS and EDINGTON [1976], the precipitates at a boundary are always found to be one of the variants of the precipitates nucleated inside one of the two grains that join at the boundary.

In the absence of additional evidence, particularly on the role of grain-boundary

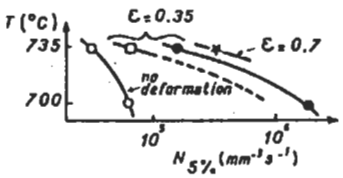


Fig. 16. Observed rate of nucleation, after 5% transformation of γ to α in a low-carbon steel, with no deformation and with two levels of prior hot deformation in the austenite (after DESALOS *et al.* [1981]).

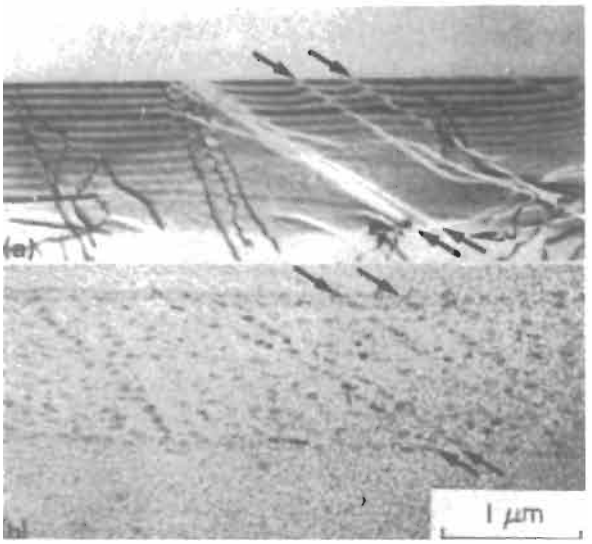


Fig. 17. Precipitation on a grain boundary in Al–Zn–Mg: (a) Grain-boundary ledges, arrowed, and dislocations in the as-quenched alloy; (b) the same area after 75 s ageing at 240°C, with preferential nucleation on the ledge and on some dislocations. (From BUTLER and SWANN [1976].)

dislocations, it is probably safest at present to regard the existing models of heterogeneous nucleation as being not yet fully established. However, the analysis of Lee and Aaronson seems likely to be valid for general boundary nucleation but to need revision to take into account the higher level of heterogeneities on the grain boundary itself. These heterogeneities probably will dominate the reaction at low driving forces where nucleation is studied by optical microscopy.

In addition, there is a tremendous lack of reliable interfacial-energy information needed to make use of the theory (HONDROS [1978]). BUTLER and SWANN [1976], for example, report evidence that higher solution-treatment temperatures that will reduce segregation of impurities and so raise the boundary energy, caused an increase of nucleation rate. Such qualitative ideas, though valuable, do not of course substitute for quantitative evaluation of the interfacial energy under various conditions.

2.2. Growth processes

A critical distinction in solid-state, growth processes, where the two crystals have different crystal structures, arises from the presence or absence of a good atomic fit across the interface. If there is no good atomic fit, as is expected, for example, for one at least of the interfaces for β nucleated at $\alpha\alpha$ grain boundaries, then rather easy atom transfer can be expected across the incoherent interface by the thermally activated jump process described in the following section. This type of interfacial growth process appears very similar to the movement of high-angle grain boundaries in recrystallization and grain growth and can be readily classified as a civilian or reconstructive process. In

solid-state phase transformations, examples include the "massive" reaction (a polymorphous change of crystal structure at a constant composition), the growth of grain-boundary nucleated "allotriomorphs", the growth of most eutectoidal two-phase structures such as "pearlite" in Fe–C alloys and the process of discontinuous precipitation, all of these reactions appear to nucleate at the parent grain boundaries.

A good atomic fit across an interface of a growing precipitate (i.e., a *coherent* interface) is, however, strongly promoted by the need to minimize the interfacial energy barrier to nucleation, § 2.1. The growth of such a coherent β precipitate into an a matrix with a different crystal structure faces a significant barrier discussed in detail in § 2.2.3 and 2.2.4. The growth barrier arises from the difficulty of finding atomic sites at the good-fit interface for atoms to leave one structure and join the other. Transfer apparently can only occur at growth *steps* or *ledges* that, in most cases, appear to have a dislocation character. The forward motion of the good-fit coherent interface appears to require the nucleation of these growth ledges and their lateral motion by a process that, at least in some cases, is equivalent to that of dislocation glide. In the case of the fcc/cph interface, particularly, as studied in Ag_2Al precipitation from silver-supersaturated aluminium solid solutions (HOWE *et al.* [1985]), this is known to be the operating mechanism. Such a mechanism is truly that of a military or displacive transformation, one of whose identifying characteristics is the production of a shape change seen as surface relief of an invariant plane strain type, as has been reported for Al–Ag by LIU and AARONSON [1970]. The mechanism of this type of growth reaction in transformation where there is a composition change and observable surface relief has become perhaps the most contentious area in diffusional phase transformations and was the subject of two recent conferences: INTERNATIONAL CONFERENCE ON BAINITE [1990] and ROLE OF SHEAR AND DIFFUSION IN PLATE SHAPED TRANSFORMATION PRODUCTS [1994]. The dispute will be described in § 2.6. Here it is sufficient to note that the growth of *coherent precipitates* is very common and occurs in almost all precipitation reactions where the precipitate grow in the grains in which they were nucleated, forming plate-like, lath-like or needle-like Widmanstätten precipitates. A plate-like precipitate is one whose three dimensions, a, b and c, have $a = b \gg c$, for a lath $a > b > c$ and for a needle $a \gg b = c$. It appears to be generally agreed that the coherent plate and lath precipitates have growth ledges and often have shape changes, but it is not agreed if all, some or only a few, grow by a displacive mechanism. There are only a few reports, see § 2.6, of the presence or absence of surface relief effects for the majority of Widmanstätten plate-like precipitates for example those that form in precipitation-hardening aluminium and copper alloys. For needle-like precipitates, for example the fcc Cu rods in bcc Fe whose coarsening was studied by SPEICH and ORIANI [1965] there appears to have been no reports of either growth ledges or surface relief shape changes.

2.2.1. Growth without change of composition

When the growing precipitate phase and the matrix have the same composition, then the growth process merely requires atom transfer across the interface. The reconstructive mechanism for this is equivalent to that seen in the movement of a high-angle grain boundary in pure metals (ch. 28, § 3.4). One complication in polymorphous phase changes, as with grain boundary migration, is the possibility of *solute drag* (CAHN [1963]

and LÜCKE and STUWE [1963]). The solute drag phenomenon arises if a trace of an alloy addition is adsorbed onto the interface between the growing and shrinking phases. The rate of motion of the interface can then be greatly retarded unless the reaction is sufficiently fast for the interface to "escape" from the solute atmosphere. Solute drag on grain boundaries slows down, but does not stop, boundary migration. For grain-boundary movement it is difficult to acquire a sufficient driving force for escape from the solute, but in solid-state phase changes the driving force is likely to be sufficiently large for solute drag not to be a significant problem in many situations, for example with supercoolings of more than about 50 K [eq. (4)]. See § 2.6, however, for a discussion on the possible role of solute drag in halting bainite growth above a critical temperature in some alloy steels.

For a phase transformation where the growing α crystal is a different phase from the β matrix, a new situation arises which is not present when the two crystals are the same phase, as in grain-boundary migration. The distinction is that any impurity will almost always prefer to concentrate in one or other of the two phases, that is, the partition coefficient,

$$k_{\alpha\beta} = C_\alpha / C_\beta, \tag{23}$$

will not be unity. This is illustrated by fig. 18 which shows a phase transformation in a dilute binary alloy where the solute segregates to the high-temperature β-phase and $k_{\alpha\beta} < 1$.

The free-energy-composition curve at a temperature, T_o, below the equilibrium transformation temperature for the pure matrix, T_e, is seen in fig. 18b, and at this temperature a critical composition $C_{(To)}$ is indicated. For an alloy with *less solute* than this critical concentration, the β-phase can transform directly, with a reduction of free energy, to the α-phase *without any change of composition*. This situation occurs through-

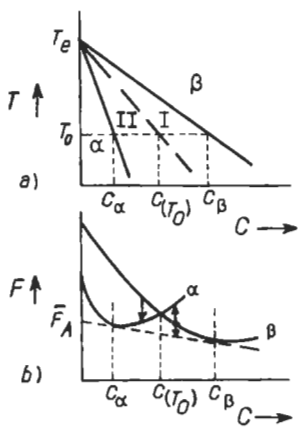


Fig. 18. (a) Phase diagram showing equilibrium below an equilibrium transition temperature, T_e; (b) the free-energy–composition curve to show the possibility of a diffusionless transformation, β to α, in region II. In region I only a reaction to give solute-depleted α can lower the free energy.

out region II in fig. 18. It is clear that a *diffusionless*, or better a *partitionless*, transformation to α of the same composition is not the most stable product in region II, but is a metastable product. The metastable reaction is only expected if it can occur significantly faster than the formation of the most stable product, solute-depleted α of composition C_α. Since partitionless transformation does not require long-range solute transport, it will generally occur faster than any reaction requiring long-range diffusion (§ 2.2.2). In region I of the two-phase region, formation of solute-depleted α is the only possible reaction. At temperatures and at compositions in the α-phase below the two-phase field, the partitionless product of homogeneous α of the initial alloy composition is both the fastest forming and also the most stable product. Partitionless transformation is known to occur in the class of transformations described as *massive*, after the shapes of the product grains often found (MASSALSKI [1970]). Since no material can ever be absolutely pure, any polymorphous reaction, if it is to occur without long-range diffusion of the impurity, must also be treated as an example of a massive transformation. In the following simple analysis it is assumed that the product and parent phases have the same composition and moreover that there is no solute build-up or depletion at the interface, in other words there is no solute drag effect considered.

Figure 19 shows the expected variation of free energy per atom across an $\alpha\beta$ interface below a transition temperature, when the low-temperature α-phase is more stable. There is a difference in free energy per atom, $\Delta F^{a}_{\beta\alpha}$, between the two phases and an activation energy barrier, ΔF_a, to atomic jumping across the interface. In the treatment given here, it is assumed that the interface is incoherent or at least has no atomic matching, as in most high angle grain boundaries. This assumption has two important consequences: Firstly that the barrier to atomic transport is likely to be of the order of the activation energy for *grain-boundary diffusion* rather than the higher activation energy for bulk diffusion. The second consequence is that there should be no difficulty for atoms

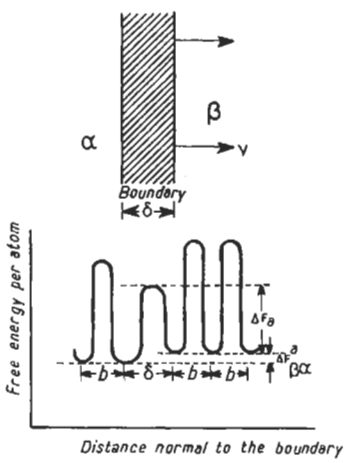


Fig. 19. Thermally activated boundary motion; there is an activation barrier ΔF_a for atoms to cross the boundary and a driving force, $\Delta F^{a}_{\beta\alpha}$ per atom, for growth of α at the expense of β.

to find a site in either the α or β phase for them to "join" after crossing the interface. Situations where this is may not be the case are discussed in § 2.2.3.

The fluxes of atoms, per atom site, from β to α, $J_{\beta\alpha}$, and the reverse flux, $J_{\alpha\beta}$, are given by:

$$J_{\beta\alpha} = \nu A_{ac} \exp(-\Delta F_a/kT), \tag{24}$$

$$J_{\alpha\beta} = \nu A_{ac} \exp\left\{-\left(\Delta F_a + \Delta F^a_{\beta\alpha}\right)/kT\right\}, \tag{25}$$

where ν is the atomic vibration frequency of atoms at the interface, typically about 10^{13} s^{-1}, and A_{ac} is the *accommodation coefficient*. A_{ac} is the probability that an atom, having left one of the crystals, can find a site in the other crystal across the interface. A_{ac} for an incoherent interface is usually assumed to be *a constant of the order of unity*. The *net* flux of atoms is given by the difference of the two fluxes, $J_{\beta\alpha}$ and $J_{\alpha\beta}$:

$$\begin{aligned} J_{\beta\alpha}(\text{net}) &= J_{\beta\alpha} - J_{\alpha\beta} \\ &= \nu A_{ac} \exp(-\Delta F_a/kT), \left[1 - \exp\left(-\Delta F^a_{\beta\alpha}/kT\right)\right] \end{aligned} \tag{26}$$

For small undercoolings, where $\Delta F^a_{\beta\alpha} \ll kT$, the difference term in eq. (26) becomes $\Delta F^a_{\beta\alpha}/kT$. The velocity of the interface, v, is given as the product of the net flux per atom site and the displacement per atom, which is the atomic spacing, b, in the growing α-phase:

$$v = bJ_{\beta\alpha}(\text{net}) = M_{\alpha\beta}\ \Delta F^a_{\beta\alpha}. \tag{27}$$

The interfacial mobility is $M_{\alpha\beta}$ and is determined from eqs. (26) and (27) as:

$$M_{\alpha\beta} = (b\nu A_{ac}/kT) \exp\left(-\Delta F_a/kT\right). \tag{28}$$

There are alternative definitions of interfacial mobility based on the replacement of the free-energy difference, $\Delta F^a_{\beta\alpha}$, in eq. (27) by either a temperature difference, ΔT_i, an interface supercooling, or a composition difference, ΔC_i, in a phase change involving composition changes (§ 2.2.2).

The simple linear relationship between the interface velocity and the driving force [eq. (27)] fails if the accommodation coefficient, A_{ac}, *is not constant*. This may occur for certain facetted interfaces where there is a difficulty in starting a new atomic layer. This is found in solidification, for materials with high-entropy of freezing. For solid-state transformations the same situation may arise for coherent interfaces between phases of different structure or orientation (§ 2.2.3). In both cases, growth occurs only at growth ledges, § 2.2.3, and since the density of these ledges will vary, for example with undercooling, then the linear relationship of eq. (27), which requires a constant mobility, is not expected. This problem of the general form of growth kinetics of facetted interfaces has been little addressed analytically in solid-state reactions. In solidification, A_{ac} and thus the mobility are found to vary linearly with undercooling in systems growing by use of a screw dislocation spiral. The expected growth velocity then increases with the *square* of the undercooling. The mobility of solid-state ledged

interfaces then requires a knowledge of ledge spacings and thus models of ledge nucleation frequencies. The nucleation of growth ledges is currently, however, not understood at all, as described by DOHERTY and RAJAB [1989], see § 2.2.4.

2.2.2. Transformations involving long-range diffusion

In region I of the phase diagram in fig. 18, the formation of the low-temperature phase requires the diffusion of solute *away* from the low-solute α-phase. On the other hand, for the typical precipitation reaction, such as the growth of solute-rich β from supersaturated α (fig. 2), the reaction proceeds by the diffusion of solute towards the growing β-precipitates. The two situations are shown in fig. 20. The composition of the matrix phase at the interface is C_i, which is shown as lying somewhere between the equilibrium value, C_β for the polymorphous reaction, or C_α for the precipitation reaction, and the alloy composition, C_o, which is usually the bulk matrix composition at some distance from the new phase. Growth of a spherical, product phase of radius r, is determined by the composition gradient outside the precipitate which is a function of the radial distance R and the time, t, since the nucleation of the precipitate. The total flux, J_R, of solute atoms (in atoms/s^{-1}) across the surface of a sphere of radius R ($R > r$, the radius of the precipitate) is given by:

$$J_R = -4\pi R^2 D(\mathrm{d}C'/\mathrm{d}R)_R \tag{29}$$

The concentration, C', must be in the units of atoms of solute per unit volume. The relationship between C' and the atomic fraction of solute C is:

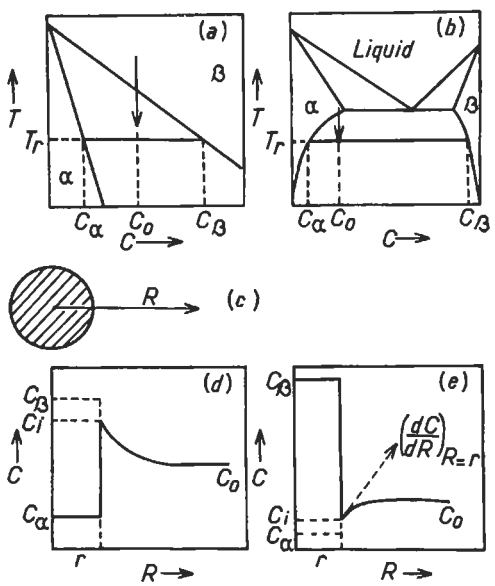


Fig. 20. Diffusion around a growing precipitate. In the matrix phase-change (a and d), solute is diffusion away from the low-solute α, while in (b) and (e) solute is diffusing towards the solute-rich β.

$$C' = C/V_a \tag{30}$$

so

$$J_R = -4\pi R^2 D\left(\mathrm{d}C/\mathrm{d}R\right)_R/V_a \tag{29a}$$

V_a is the volume of one atom of the alloy, which for simplicity is assumed independent of composition. The volume of the precipitate is $\frac{4}{3}\pi r^3$, and developing the model for the case of precipitation, the *increase* in the number of solute atoms, Δn_B, in the precipitate above that of the matrix from which it forms, is given by:

$$\Delta n_B = (\tfrac{4}{3}\pi r^3)(C'_\beta - C'_i) = (\tfrac{4}{3}\pi r^3/V_a)(C_\beta - C_i) \tag{31}$$

$$\mathrm{d}\Delta n_B/\mathrm{d}r = 4\pi r^2 (C_\beta - C_i)/V_a \tag{32}$$

The total flux to the precipitate, $\mathrm{d}\Delta n_B/\mathrm{d}t$ is then given by:

$$\frac{\mathrm{d}\Delta n_B}{\mathrm{d}t} = \frac{\mathrm{d}n_B}{\mathrm{d}r}\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{4\pi r^2}{V_a}(C_\beta - C_i)\frac{\mathrm{d}r}{\mathrm{d}t} \tag{33}$$

This flux will be that provided by diffusion [eq. (29a)], with $R = r$. The supersaturation for diffusion is usefully defined as Ω_D, eq. (34).

$$\Omega_D = (C_0 - C_i)/(C_\beta - C_i). \tag{34}$$

Flux-balance from eqs. (29a) and (33) then gives:

$$4\pi R^2 D(\mathrm{d}C/\mathrm{d}R)_R = 4\pi r^2 (C_\beta - C_i)(\mathrm{d}r/\mathrm{d}t) \tag{35}$$

The change of sign in the flux equation (29) arises since the negative flux is one in the $-R$ direction, which is a *positive* flux for precipitate growth. Provided that there is a low supersaturation, $\Omega_D < 0.1$, then during the time that it takes a solute atom to diffuse across the solute-depleted region around the precipitate, the solute distribution and the instantaneous growth rate, $\mathrm{d}r/\mathrm{d}t$, can both be regarded as essentially constant. Rearranging eq. (35) for the condition that $\Omega_D < 0.1$, yields:

$$\frac{\mathrm{d}R}{R^2} = \frac{D\,\mathrm{d}C}{r^2\left(C_\beta - C_i\right)\left(\mathrm{d}r/\mathrm{d}t\right)}. \tag{36}$$

Equation (36) can then be integrated, from $R = r$, $C = C_i$, to $R = \infty$, $C = C_o$, assuming that $\{r^2(C_\beta - C_i)(\mathrm{d}r/\mathrm{d}t)\}$ is essentially constant in the time it takes solute to diffuse from $R \gg r$ upto $R = r$. Integration of eq. (36), under these conditions, yields:

$$\frac{1}{r} = \frac{D\left(C_0 - C_i\right)}{r^2\left(C_\beta - C_i\right)\left(\mathrm{d}r/\mathrm{d}t\right)}. \tag{37}$$

Equation (37) can now be rearranged to give the growth rate, $\mathrm{d}r/\mathrm{d}t$:

$$dr/dt = \frac{D}{r}\frac{(C_0 - C_i)}{(C_\beta - C_i)} = \frac{D\Omega_D}{r}. \tag{38}$$

The growth rate, dr/dt, clearly does vary with the precipitate radius, r, (eq. 38) but at small supersaturations this can be ignored for the small change in radius that occurs during the time period in which a solute atom moves across the solute depleted region. This simplifying assumption is known as the *constant size approximation.*

The growth rate can also given by a relationship derived from the local departure from solute equilibrium $(C_i - C_\alpha)$, by use of a slightly modified definition of mobility $M_{\alpha\beta}'$ rather than $M_{\alpha\beta}$ in eq. (28)*. The velocity is then given by:

$$dr/dt = M'_{\alpha\beta}\frac{(C_i - C_\alpha)}{(C_\beta - C_i)}. \tag{39}$$

The two equations for the growth rate, eq. (38) and eq. (39), can be equated and rearranged to determine a dimensionless parameter, $rM_{\alpha\beta}'/D$, which determines the unknown composition, the interface composition in the α matrix, C_i:

$$rM'_{\alpha\beta}/D = \frac{(C_0 - C_i)}{(C_i - C_\alpha)}. \tag{40}$$

The dimensionless parameter, $rM_{\alpha\beta}'/D$ is the compositional equivalent to the Nusselt number in heat flow. When $rM_{\alpha\beta}'D \gg 1$, then $C_i \to C_\alpha$, giving the condition of *local equilibrium* (HILLERT [1975]). This condition also gives what is called *diffusion-controlled growth.* Diffusion-controlled growth arises when the interface reaction is so fast that any further increase of mobility does not significantly accelerate the overall reaction. The opposite extreme, when $rM_{\alpha\beta}'/D \ll 1$, leads to $C_i \to C_o$. This arises when the solute diffusion is so rapid and atom transport across the interface so slow, that the solute-depleted region around the precipitate is effectively eliminated by rapid diffusion. We then have the condition described as *interface-controlled growth.* In either extreme case, the composition difference driving growth, $C_o - C_\alpha$, is applied fully to the slowest, most resistive step, either the diffusion in the matrix or the interface reaction. The intermediate case, where $rM_{\alpha\beta}'/D \approx 1$ is usually described as one of *mixed control,* where the composition difference driving the reaction is applied to *both* processes. Here, an increase of *either* the diffusion coefficient *or* the mobility will accelerate the reaction.

Diffusion control. With local equilibrium, $C_i = C_\alpha$, eq. (38) becomes:

$$dr/dt = \frac{D}{r}\frac{(C_0 - C_\alpha)}{(C_\beta - C_\alpha)} = \Omega D/r \tag{41}$$

* There is a straightforward relationship between the two mobilities, which can be obtained by relating the local departure from solute equilibrium at the interface, $C_i - C_\alpha$, to the free energy per atom (KL in fig. 1), this is:

$$M'_{\alpha\beta} = M_{\alpha\beta}\, kT\, \{(C_\beta - C_i)/(C_i - C_\alpha)\}\, \ln(C_i/C_\alpha)$$
$$\approx M_{\alpha\beta}\, kT\, (C_\beta - C_i)/C_\alpha \quad \text{when } (C_i - C_\alpha)/C_\alpha \ll 1.$$

where Ω is the dimensionless supersaturation, $(C_o - C_\alpha)/(C_\beta - C_\alpha)$. Ω is also, by the *lever rule* of phase equilibria, the atomic fraction of the β precipitate phase at equilibrium. If, as was assumed above, that the volume per atom is constant then Ω is also the *volume fraction* of β. It is easy to use eq. (41) to demonstrate the conditions that allowed the simple integration of eq. (36), that is that $r^2(dr/dt)$ is essentially constant over the time that solute diffuses across the solute depleted region. With $C_i = C_\alpha$, and using eq. (41), eq. (35) yields the concentration profile in the matrix for $R \geq r$ as:

$$C = C_0 - \frac{(C_0 - C_\alpha)r}{R}.$$

Differentiation of this expression shows, at $R = r$, that the concentration gradient at the interface, $(dC/dR)_{R=r}$, is given by:

$$(dC/dR)_{R=r} = (C_0 - C_\alpha)/r \tag{35a}$$

when $R = 2r$, $C = C_o - \{(C_o - C_\alpha)/2\}$. This result means that half the difference in composition in the profile occurs over a distance r from the precipitate interface. The time, Δt, that solute takes to diffuse over a distance r, when the diffusion coefficient is D, is given by:

$$\Delta t \approx r^2/D. \tag{41b}$$

With $dr/dt = D\Omega/r$, during a time Δt, the radius will grow by:

$$\Delta r \approx \Delta t(dr/dt) = \Omega r.$$

So for $\Omega < 0.1$ there is near constancy of r, and thus of dr/dt, *during the time that solute traverses the composition profile*. This result validates the assumptions used in the integration of eq. 36.

Equation (41) can be integrated from $t = 0$ when $r = 0$, to give the square-root of time relation expected for diffusion control:

$$r = (2\Omega Dt)^{1/2}. \tag{42}$$

Interface control. From eq. (39) with $C_i = C_o$, we obtain the linear relationship:

$$r = M'_{\alpha\beta}\Omega t. \tag{43}$$

It should be clear from the role of the parameter $rM_{\alpha\beta}'/D$ in controlling the reaction, that with $r \approx 0$ immediately after nucleation, the situation should always tend towards interface control. This may only be detectable, however, for very small values of the *critical radius* of nucleation, r^* [eq. (10a)]. Diffusional growth cannot occur until a critical nucleus has formed by thermal fluctuation. Even after nucleation, precipitates close to the critical radius will have their growth conditions modified by interfacial effects, with the solubility $C\alpha$ increased by the Gibbs–Thomson equation (§ 2.2.6). The only conditions for which interface control, eq. (43), or even mixed control can be realistically expected is for *very low values of mobility*. This can occur under conditions of faceted or ledged growth (§ 2.2.3).

References: p. 1497.

Diffusion control — General. More general solutions for diffusion-controlled growth of both spheres and flat plates have been provided by AARON *et al.* [1970], including the exact solution for the general case without the approximations involved in eq. (35). For a sphere the radius is given by:

$$r = (\alpha D t)^{1/2}. \tag{44}$$

The dimensionless coefficient α is defined by:

$$\Omega = (\alpha/2)\exp\alpha\left[\exp(-\alpha) - (\alpha\pi)^{1/2}\operatorname{erf}\left(\alpha^{1/2}\right)\right]. \tag{45}$$

The approximate relationship of eq. (42) is valid with a low supersaturation $\Omega < 0.1$, for which $\alpha \approx 2\Omega$. This is the usual situation for many precipitation reactions. For the growth of plates, the half-thickness a_x is given by:

$$a_x = 2\alpha'(Dt)^{1/2} \tag{46}$$

and

$$\Omega = \alpha'\pi^{1/2}\exp\left(\alpha'^2/2\right)\operatorname{erfc}\left(\alpha'/2\right). \tag{47}$$

For oblate and prolate spheroids, that are reasonable approximations to the frequently found plate-like and needle-like precipitate shapes, HAM [1958] and HORVAY and CAHN [1961] have shown that the half-thicknesses, a_x, and half-lengths, a_y, are given by:

$$a_x = 2(\beta D t)^{1/2}, \tag{48}$$

$$a_y = 2A(\beta D t)^{1/2}, \tag{49}$$

The parameter, β, is obtained from the supersaturation, Ω, and the aspect ratio A ($A = a_y/a_x$) by equations of a form similar to eqs. (45) and (47) which cannot be inverted to give β as explicit functions of A and Ω (HORVAY and CAHN [1961]). HORVAY and CAHN [1961] provide figures showing how β increases with the supersaturation at different values of the aspect ratio, A. The equations for β for a particular values of A and Ω are, of course, readily solved numerically. These equations are different for plate-like and needle-like precipitates. The assumptions that lead to eqs. (48) and (49) were that the precipitates were formed with a given shape and from then on, growth occurred with diffusion control and local equilibrium, all round the precipitate, that is, with $C_i = C_\alpha$, at all positions on the interface. The analysis then predicts that the shape is preserved, that is, the precipitate grows with a constant aspect ratio.

The prediction of a constant shape was challenged by the successful application of stability theory to the growth process (MULLINS and SEKERKA [1963]; § 2.2.6). However, as will be discussed in § 2.2.6, the stability criticisms of the Ham, Horvay and Cahn analysis do not rule out its usefulness for the highly elongated "Widmanstätten" shapes. In fact, the Ham, Horvay and Cahn analysis appears to have very wide validity for the growth of Widmanstätten precipitates, except at high Ω ($\Omega \geq 0.5$), when linear lengthening kinetics, eqs. (59b and 64), as distinct from the square root of time lengthening kinetics eq. (49), are found.

2.2.3. Role of interface structure in growth processes

AARONSON [1962] pointed out that for the plate-like *Widmanstätten precipitates* in systems like Al–Ag_2Al (fig. 9, above), the broad faces of the precipitates have a very good atomic fit (§ 2.1.2). This is because the atomic positions on (0001) hexagonal planes are the same as on the {111} fcc planes in the matrix. For growth of either phase by single-atom transfer to the other structure, a very considerable rise in energy results (fig. 21a, b). This continues with a second atom (fig. 21c) and only becomes a reasonable structure with either three, five or six atoms added (fig. 21d, e). The obvious conclusion of this idea is that it will be *very difficult* for individual atoms to join the precipitate on the habit plane: In other words the accommodation coefficient, A_{ac}, in eqs. (24) and (25) will be close to zero for precipitate thickening, giving very low values of the mobilities, $M_{\alpha\beta}$ or $M_{\alpha\beta}'$.

The solution to this problem appears to be the formation of a *ledge* or *step* on the broad face of the precipitate. Atoms can then add to the *riser* of the ledge which will be essentially incoherent or at least ill fitting, so the interface moves forward by *sideways* motion of the ledges. The structure of the riser of the ledge is that of the core of a dislocation, which for the fcc/cph interface is a Shockley partial dislocation: $a/6\langle 112\rangle$ with respect to the fcc crystal. When the model of growth by ledge migration was put forward by AARONSON [1962] there was little direct evidence to support his suggestion though a strong need, from the observed shapes of the Widmanstätten precipitate, for an explanation of the very high aspect ratios, A, found; these are often of the order of 100 or more (see for example RAJAB and DOHERTY [1989]). Such values of the aspect ratio are *very much higher* than the values of A expected for the *equilibrium shapes* of the precipitates, § 2.1.2. AARONSON [1962] and AARONSON *et al.* [1970] argued that the high aspect ratio of the Widmanstätten plates and of needles was due to the need for ledge growth on the good fit parts of the precipitate interface. Following the work of WEATHERLY and SARGENT [1970] and WEATHERLY [1971] it became clear that the growth ledges postulated by Aaronson do indeed actually occur on the broad faces of Widmanstätten precipitates. Furthermore, the ledges have dislocation character and that the good fit areas

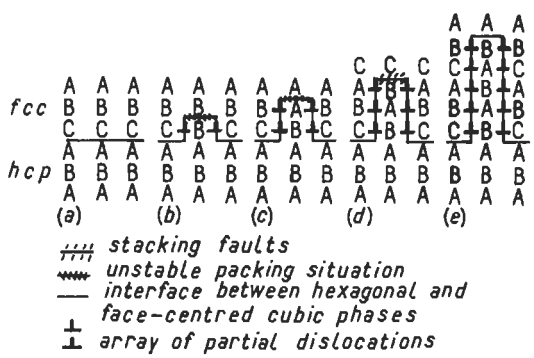
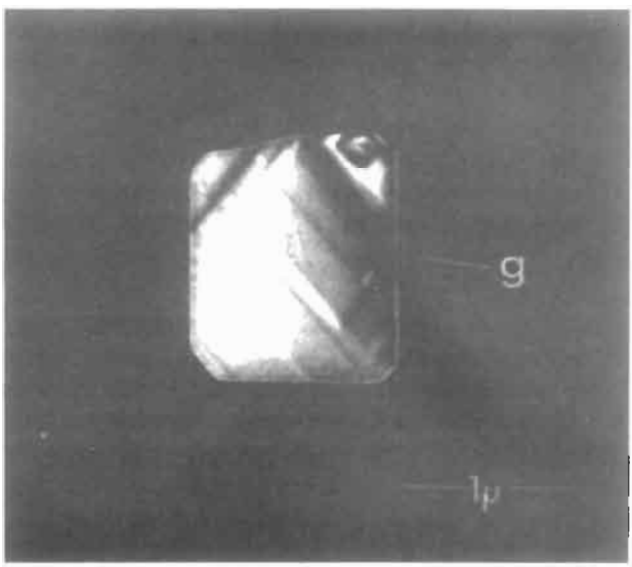


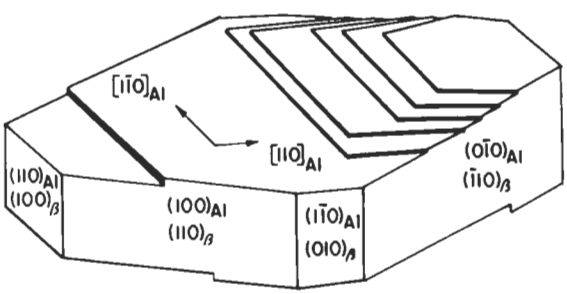
Fig. 21. Problem of growth of the lower hcp phase from the upper fcc phase by random atomic jumps: intermediate structures (b) and (c) are very unstable compared to (d), a three-layer ledge, or (e), a five- (or six-)layer edge. (After MARTIN and DOHERTY [1976].)

of these interfaces do often show reduced mobility compared to ill-fitting parts of the interface (e.g. PURDY [1978], RAJAB and DOHERTY [1989]). Figure 22 shows the dislocation-like contrast shown by the ledges on faceted Mg_2Si precipitates in an aluminium alloy. HOWE *et al.* [1985] directly imaged the dislocation ledges in an fcc/cph system by high resolution TEM and showed that the steps occurred in clusters of 1, 2 or 3 Shockley partial dislocations with 2, 4 or 6 atomic layers per step.

It was implicitly assumed by Aaronson that ledged interfaces would always show interface-controlled or mixed controlled growth, that is, ledged interfaces would be *inhibited* and grow more slowly than expected for diffusion control. This does not follow automatically however, as will be discussed in § 2.2.4.



(a)



(b)

Fig. 22. Growth ledges on a Mg_2Si plate-like precipitate seen by dark-field electron microscopy (from WEATHERLY [1971]).

The motion of growth ledges, on good-fit interfaces, need not be based on the ledge moving by individual random, thermally activated atomic jumps at the growing phase at the risers. As was originally suggested by NICHOLSON and NUTTING [1961] and by HREN and THOMAS [1963] and confirmed by HOWE *et al.* [1985], growth of a hexagonal precipitate such as Ag_2Al from a fcc matrix can occur by *glide* of an $a/6\langle 112\rangle$ dislocation on a {111} habit plane (fig. 23). The glide of a single such dislocation will cause two atomic layers of the matrix fcc structure to transform to the hexagonal structure, and to cause *shear* of the matrix structure. This idea raises a fascinating problem, since this type of transformation when it occurs in cobalt and in cobalt–nickel alloys (GAUNT and CHRISTIAN [1959]) is fully *martensitic* (see ch. 16). However, in dilute aluminium–silver alloys it is not possible, thermodynamically, for diffusionless (partitionless) precipitation of the hexagonal phase to occur, since the fcc solid solution does not contain enough silver for a diffusionless transformation. There is, in addition, overwhelming experimental evidence that the precipitates are indeed silver-rich compared to the solid solution (e.g., LAIRD and AARONSON [1969], FERRANTE and DOHERTY [1976] and RAJAB and DOHERTY [1989]). Further evidence for the dual martensitic (displacive) and diffusional character of the fcc-to-hcp transformation in aluminium–silver alloys was provided by the observations of AARONSON *et al.* [1977] on the heterogeneous nucleation of the intermediate γ'-phase on dislocations. Previously LIU and AARONSON [1970] had shown that these precipitates produced, plane-strain, *surface relief* during precipitation. The precipitates were observed to form by the glide of a partial Shockley dislocation, away from the other Shockley partial, which together with the first partial had made up a perfect lattice dislocation in the supersaturated solid solution before precipitation (fig. 24). By assuming that the transformation involved a change of composition, that is, that the precipitates were indeed the expected silver-rich metastable γ'-phase, the authors were able to predict successfully the required undercooling for this type of nucleation to occur. It is clear that this reaction has the shear characteristics of a martensitic (displacive) reaction, while also involving a diffusive change of composition. This combination seems perfectly acceptable for a case where the reaction requires a diffusional change of composition but where the interface process can occur by a shear process. It is tempting to use the term "bainitic" to describe a transformation that requires long-range solute transport by diffusion to give the necessary change in

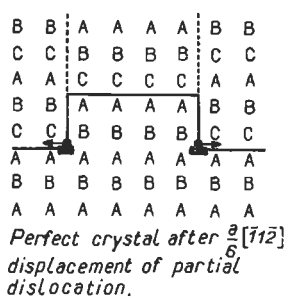


Fig. 23. Growth of hcp phase from fcc by glide of partial dislocation: two-layer ledge. (After DOHERTY [1982].)

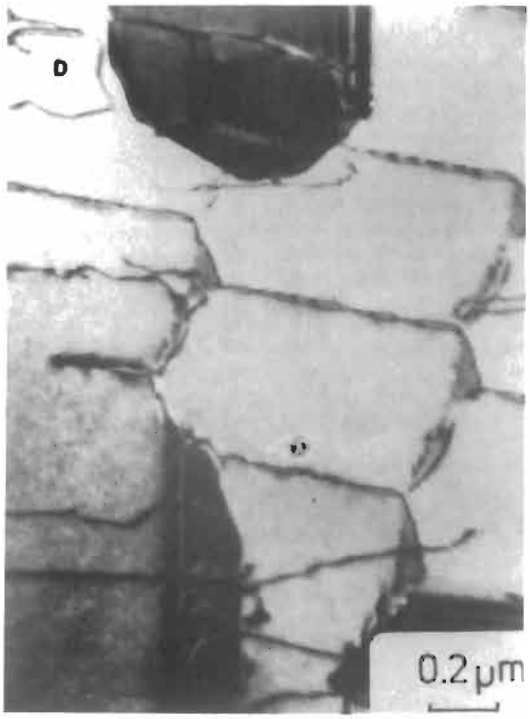


Fig. 24. Al–15 wt% Ag, solution-treated at 500°C, directly quenched to 360°C and held there for 60 s, showing the γ' plates forming behind moving partial dislocations (from AARONSON *et al.* [1977]).

composition (C_α to C_β) but where the change from α to β crystal structures occurs by a displacive reaction. The problem with this name is that it is not yet fully agreed that the *bainitic transformation* in the fcc-to-bcc decomposition of austenitic iron occurs by these two processes. Shear by dislocation glide parallel to the habit plane certainly gives the invariant-plane strain characteristic of the martensitic transformations (GAUNT and CHRISTIAN [1959]). So even if bainite itself has not been proved to form by the required combination of solute diffusion and martensitic like crystal structure change a "bainitic" type of reaction certainly does occur in the Al–Al_2Ag precipitation reaction.

Irrespective of how the mechanism of ledge movement across the coherent interface is described, the occurrence of ledged interfaces (see for example the review by AARONSON [1974]) indicates that that this is an important phenomenon and that the diffusional mobility of these interfaces will be controlled by the nucleation and growth of the ledges (§ 2.2.4).

It is important to note, however, that growth ledges of any form are only expected for coherent interfaces between phases which have *different* structures and/or orientations. The difficulty identified in fig. 21 only arises since the atomic fit is restricted to isolated atomic planes. If the two phases have identical structures and orientations as in, for example, the so-called "cube-cube" relationship, then there is no difficulty in adding

atoms onto the growing phase — in fact there is no interface step at all, since the arrival of the required solute at the interface automatically changes the structure from that of the matrix to that of the precipitate. Examples of this type of cube-cube relationship, already cited in § 2.1.2, are γ precipitates in β phases, in Cu–Zn, Cu–Al or Ag–Zn, Ni_3Al precipitates in Ni alloys, and Al_3Sc and Al_3Li precipitates in Al alloys. In addition, the formation of Guinier–Preston zones, solute rich clusters in heavily solute supersaturated alloys, are further examples of the formation of precipitates with the same structure and orientation as the matrix from which they grow. As expected, on the basis of the ideas above, there are no unambiguous reports of ledge structures at the coherent interfaces between any cube-cube matrix-interface pairs. This failure to see the ledges in these systems must be considered significant, given the large quantity and the high quality of the experimental observations, particularly in the nickel-based alloys.

2.2.4. Growth of ledged interfaces

JONES and TRIVEDI [1971, 1975] analyzed the growth kinetics expected for a ledged interface. They considered ledges of height h and spacing λ. If the ledges have an instantaneous velocity *along* the interface of v_s then the overall *normal* velocity of the interface, v_i is:

$$v_i = v_s h/\lambda. \tag{50}$$

Jones and Trivedi analysed the case of an isolated ledge moving with a constant velocity and determined the steady state diffusion field that would describe the solute field feeding solute to the ledge. They produced a relationship between the dimensionless Peclet number, P, ($P = v_s h/2D$) and the dimensionless supersaturation, Ω, for an isolated ledge. Their solution, for the case when there is diffusion-controlled atomic addition to the riser of the ledge, is:

$$\Omega = 2P\alpha(p) \tag{51}$$

The function $\alpha(P)$ can be evaluated numerically from data provided by JONES and TRIVEDI [1971]. Typical values of P are about 0.03 at $\Omega = 0.1$, 0.08 at $\Omega = 0.2$ and 0.24 at $\Omega = 0.4$. It was reported by JONES and TRIVEDI [1971] that the solute field about a moving ledge has a finite extent (fig. 25), and that this field becomes more localized around the ledge as the supersaturation increases and the ledge velocity rises. An immediate consequence of their analysis, for multi-ledged interfaces, is that if ledges are more widely spaced than the extent of the diffusion field, then each ledge can move *independently* at the steady state velocity. As a result, the interface velocity v_i will have a *constant* value, given by:

$$v_i = 2DP/\lambda. \tag{52}$$

DOHERTY and CANTOR [1982] drew attention to the problem raised by this result when compared to the expected velocity for the *diffusion*-controlled growth of a plate, which is obtained, *assuming a negligible barrier from the interface process*, by differentiating eq. (46), to give the normal interface velocity as:

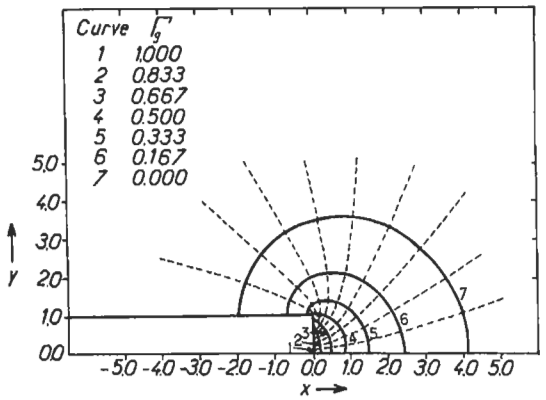


Fig. 25. Analytical model of ledge growth: Predicted contours of equal solute content about a ledge moving with a constant velocity. Contours are for different values of the normalized supersaturation, $\Gamma_g = 1-\Omega/\Omega_0$. With $\Omega_0 = (C_0-C_\alpha)/(C_\beta-C_\alpha)$ the initial supersaturation and $\Omega = (C-C_\alpha)/(C_\beta-C_\alpha)$ the remaining supersaturation, the value of Γ_g is 0.0 for zero supersaturation (equilibrium solubility), and 1.0 for full supersaturation. (After JONES and TRIVEDI [1971].)

$$v_i = da_x/dt = \alpha'(D/t)^{1/2} \tag{53}$$

The α' of eqs. (46 and 53) has no relationship with the $\alpha(P)$ of eq. 51. The difficulty is that since eq. (53) contains time in the denominator of the velocity function, the velocity of the mobile, diffusion-controlled interface falls steadily with time, eventually becoming smaller than the velocity predicted for the ledged interface, of spacing, λ, however small the ledged velocity it might be, provided only that λ is finite. Moreover, the velocity advantage of the ledged interface continues to grow with longer and longer times, becoming, at infinite time, infinitely larger. Although fixed ledge spacings and infinite growth times are not experimentally realistic the "thought experiment" described above indicates a major difficulty with the Jones and Trivedi analysis. DOHERTY and CANTOR [1982] were unable to find the error in the analysis but tested the results against their numerical simulations of the diffusion field around a pair of ledges. The simulation assumed the ledges to be nucleated in the fully supersaturated matrix and to grow by solute flux to the ledge riser. There was no barrier to solute addition at the *risers* of the ledges, which were assumed to show local equilibrium, $C_i = C_\alpha$. The rest of the interface was, however, not allowed to accept solute; that is, its mobility was set to zero. There was, by the geometric "point effect of diffusion", a higher flux to the top of the ledge than to the base. However, the ledge shape was held stable by distributing the arriving flux *uniformly* along the riser: This model made the implicit assumption of rapid interface diffusion along the riser interface. The riser interface will have, in the solid state case, fig. 23, the structure of the *core of a partial dislocation* and so would be expected to show rapid solute transport down the riser.

For a ledged precipitate growing from a liquid in solidification, or from a vapor, then such rapid solute transport will not be expected and so an alternative model is needed.

A suitable model for growth from a liquid will be the one initially used by Jones and Trivedi and subsequently by ENOMOTO [1987] in which the ledge or step velocity is determined by the velocity of the *base of the step*, where local equilibrium was assumed. With this boundary condition, higher parts of the riser retain some supersaturation which feeds, by matrix diffusion, to the base of the step.

The results of the Doherty–Cantor finite-difference computer simulation were significantly different from the analytical results of Jones and Trivedi. As expected, the high local flux from a fully supersaturated matrix gave a high initial ledge velocity but this velocity fell steadily with time. For widely separated ledges, the total ledge displacement, X_s, showed an initial time exponent of about 0.76. This increased to 0.88 at longer times for supersaturations up to 0.1 (fig. 26). The origin of this ever-decreasing ledge velocity was the steady movement of isoconcentrates away from the ledge as time increased (fig. 27). The only condition for which a constant velocity, v_s, of the *ledge* was found by DOHERTY and CANTOR [1982], was for high supersaturations, $\Omega = 0.5$, and quite long times. Here the solute distribution was highly asymmetric (fig. 27). The asymmetry appears to be physically reasonable since the ledge is growing rapidly and therefore "catching up" with the isoconcentrates in front of the ledge but leaving a huge "trail" of depleted matrix behind it. Subsequent simulations by ENOMOTO [1987] that were run for longer times than those of DOHERTY and CANTOR [1982] found that isolated ledges eventually achieved constant (steady-state) ledge velocities *at all supersaturations* but again with *asymmetric solute distributions* particularly at high supersaturations. In all cases there was an initial transient time, that decreased with higher supersaturations, during which, as seen by Doherty and Cantor, the growth velocity fell. After the transient period the isolated ledge reached a steady velocity that was faster the higher the supersaturation. ATKINSON [1981, 1982] presented an improved analytical model for the steady state growth of an isolated ledge at a constant ledge velocity which avoided the error in the original Jones and Trivedi treatment. Atkinson's analytical results matched closely the steady state velocities for isolated ledges produced by Enomoto's simulations.

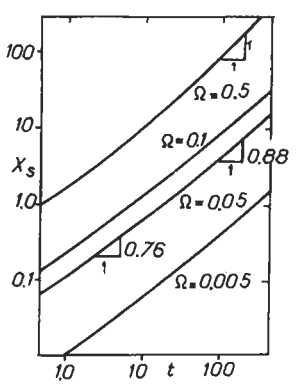
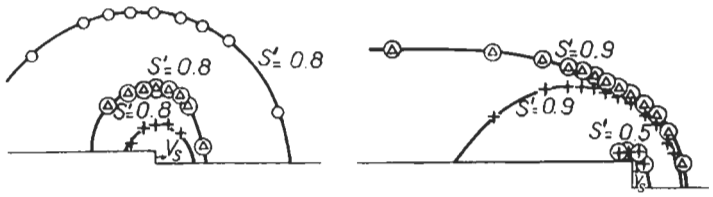


Fig. 26. Computer simulation of ledge growth. The total ledge displacement X_s in units of ledge height h, for dimensionless time of growth, $t = h^2/D$, for various supersaturations. Only for $\Omega = 0.5$ was a constant velocity found at long times. Isolated ledge. (After DOHERTY and CANTOR [1983].)

References: p. 1497.

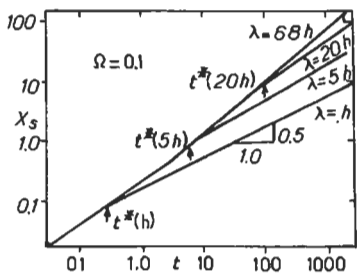


	V_s		t	V_s	
	0.056	+	6.4	1.04	
$\Omega = 0.05$	0.042	△	51.0	0.996	$\Omega = 0.5$
	0.034	○	409.0	0.995	

Fig. 27. Computer simulation of ledge growth. Predicted contours of equal solute supersaturation at growing ledge for three values of t in units of h^2/D, with the corresponding values of the ledge velocity, V_s, in units of D/h, for two different supersaturations, 0.05 and 0.5. S' a normalized supersaturation, defined by $S' = (C - C_\alpha)/(C_0 - C_\alpha)$; for the full initial supersaturations $S' = 1$. Isolated ledges. (After DOHERTY and CANTOR [1983].)

Both Atkinson and Enomoto assumed the *solidification* boundary condition for the velocity of ledge, described above, rather than the *solid state* condition used by Doherty and Cantor. The result of the different assumed boundary conditions was that the velocities found by Enomoto were slightly smaller than those in the simulations of Doherty and Cantor (see DOHERTY and CANTOR [1988, 1989]). Apart from the minor difference in the boundary conditions at the riser, the simulations run by the two studies appeared to match rather well — despite some rather fierce attacks by ENOMOTO and AARONSON [1989] on the initial simulation of Doherty and Cantor.*

The critical difference between the computer models and the earlier analytical result of Jones and Trivedi was found when the interledge spacing, λ, was reduced to allow diffusional interaction between adjacent ledges (see fig. 28a). When the growth times reached and exceeded the values of $\lambda^2/4D$, where the diffusion fields overlapped, the time exponent of the *ledge* displacement fell to 0.5, which, by eq. (52), produced the



* Since the author of this review was involved in this and an earlier dispute on ledge growth kinetics, DOHERTY *et al.* [1977, 1978b], AARONSON [1977], SANKARAN and LAIRD [1977, 1979], readers may wish to read the original papers before deciding whether or not to accept the opinion presented here.

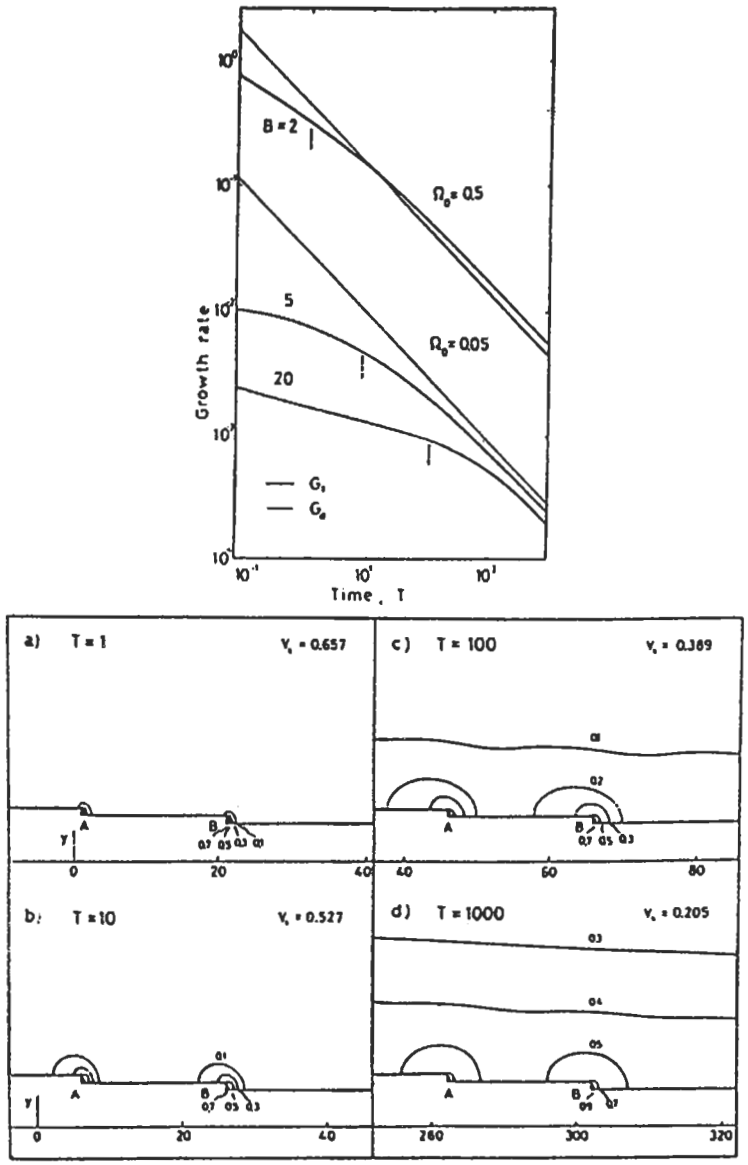


Fig. 28. Computer simulation of ledge growth. (a) Total ledge displacement X_s against time for closely spaced ledges, $\lambda = h$, $\lambda = 5h$, $\lambda = 20h$ and $\lambda = 68h$. For $t > t^*$ ($t^* = \lambda^2/D$) the diffusion fields overlap and an $X_S/t^{0.5}$ relationship is found. (After DOHERTY and CANTOR [1983]). (b) Precipitate growth (thickening rate) as a function of time for the ledged interface, G_l, (thick line) and the volume diffusion limited rate, G_d, (thin line) for two supersaturations, $\Omega_o = 0.5$ and 0.05, and the ledge spacings B (λ/h) = 2, 5 and 10. The arrows indicate the values of $t^* = \lambda^2/4D$. (c) The isoconcentrates, here $(1-\Omega/\Omega_o) = 1-S'$, at a supersaturation, Ω_o, of 0.5 are shown for various dimensionless times T (Dt/h^2), with multiple ledges, at a spacing λ of 20h, after being displaced by the growth distances in units of h and with the dimensionless step velocities V_S (vh/D) indicated. t^* is 100. (After ENOMOTO [1987]).

References: p. 1497.

same time exponent for the interface displacement, a_x. This is, of course, the time exponent for full diffusion-controlled growth of the interface [eq. (46)]. The simulation results, both those of Doherty and Cantor and of Enomoto, showed, after the overlap of the diffusion fields from adjacent ledges, that the interface displacement not only had the diffusion limited time *exponent* but also, to within a few percent, the correct value [$2\alpha' D^{1/2}$, eq. (46)] for the diffusion-controlled *slope* of a plot of the half thickness, a_x versus $t^{1/2}$. Such plots from the simulations, had time delays corresponding to the initial periods of slow growth, where solute close to the interface could not join the precipitate except at the ledges. At times somewhat after the diffusion fields overlapped, the *kinetic* behaviour was as though the interface could accept solute everywhere. This occurs since the diffusion fields around interacting ledges, apart from the perturbation close to the step, are similar to the fields about a flat interface that is accepting solute at all points along the interface. Enomoto's simulations found very similar results. Figure 28b shows his computer predicted growth rate for the ledged interface, G_l, and that expected for diffusion limited thickening by eqs. (46 and 47), G_d, at two supersaturations and for two dimensionless spacings B ($B=\lambda/h$). The critical time for diffusional interaction, t^*, is indicated by the vertical arrows. For times upto t^* the growth rate was clearly inhibited by the ledged structure but after times of the order of $4t^*$ the ledged interface moved at the rate limited by diffusion. The solute isoconcentrates are shown in fig. 28c, the contours are given as $(1-\Omega/\Omega_o)$. For times less than t^* ($t=400$) much of the interface is almost fully supersaturated, $(1-\Omega/\Omega_o)\to 0$. But when $t=1000$ the isoconcentrates, $(1-\Omega/\Omega_o)$, from 1 to 0.4 are parallel to the interface. Under these last conditions, *any solute atom that enters the solute-depleted region is almost certain to be captured by the interface* by one of the moving ledges.

Both sets of computer simulations clearly support the physically reasonable assumption that once the reaction time is significantly larger than the interaction time, $\lambda^2/4D$, between ledges the ledged interface should show diffusion limited growth kinetics. This was the assumption used implicitly in the successful application of the Ham, Horvay and Cahn growth kinetics, eqs. (48 and 49) to match the behaviour of Widmanstätten plate-like precipitates in aluminium alloys discussed in 2.2.5.2.*

There is no analytical theory that describes the situation for diffusionally interacting ledges — this is still a situation that can only be handled by numerical simulation. However a rather clear pattern for ledge growth now appears established by DOHERTY and CANTOR [1982] and ENOMOTO [1987], at least for constant ledge spacings. A set of ledges, at a given fixed distance apart, grows initially in an isolated manner with a steadily falling ledge velocity that, if the ledge spacing is large enough, will achieve a constant step velocity and thus a constant interface velocity. The interface thickening rate for such isolated, non-diffusionally interacting ledges is slower than expected for volume diffusion control. Interaction between the ledges, however, starts to occur at times of the

* The assumption was also used explicitly in an early dispute about this analysis, DOHERTY *et al.* [1977, 1978b], AARONSON [1977], SANKARAN and LAIRD [1977, 1979], and the computer simulations of DOHERTY and CANTOR [1982] were run, in part, to test this assumption.

order of $\lambda^2/4D$, and, at times somewhat larger than this, there is a transition to volume controlled thickening. Once this situation is reached then the analytical equations of Ham and Horvay and Cahn become applicable and the precipitates should continue to grow, maintaining the aspect ratios that they had acquired at the time of the transition to diffusion controlled thickening. This does indeed appear to be the experimental situation, see FERRANTE and DOHERTY [1979] and RAJAB and DOHERTY [1989].

In contrast to the situation for the *growth* of ledges, there is almost no understanding at all for ledge *formation* or *nucleation*. WEATHERLY [1971] applied conventional nucleation theory to the formation of a "pillbox" disc on the surface of a precipitate, but the predicted rates of nucleation were too low for any reasonable rate of ledge formation. This happened even during initial growth when the interface has full solute supersaturation and therefore a high driving force for nucleation. At later stages of growth, and especially during Ostwald ripening (§ 2.4), there is so little supersaturation that "homogeneous" ledge nucleation appears impossible (see DOHERTY and RAJAB [1989]). Various possible heterogeneous nucleation sites have been described such as those at emerging screw dislocations and at the intersections of precipitates on different habit planes (AARONSON [1974]). The beautiful experimental observations of HOWE *et al.* [1985] show direct evidence for heterogeneous nucleation of ledges on the broad faces of a precipitate caused by the *intersection of other precipitates* forming on *other variants of the habit plane*. RAJAB and DOHERTY [1989] provided indirect evidence from growth kinetics, § 2.2.5.2, for ledge formation by this method of precipitate intersection. However DOHERTY and RAJAB [1989] showed that this effect could not be accounted for using conventional heterogeneous nucleation theory. At the point of intersection, where heterogeneous nucleation of a ledge using the interfacial energy of the high energy rim of the intersecting plate might have been expected, and appears to actually occur, there is no supersaturation at all since the rim of the other plate is, from its growth kinetics, known to be a perfect sink for solute. The critical problem of understanding the mechanism and thus the kinetics of *ledge nucleation* in the thickening of Widmanstätten precipitates is therefore a major unsolved problem in the field of diffusional phase transformations.

2.2.5. Quantitative experimental observations of growth rates

As with nucleation studies there is a considerable lack of quantitative data on growth rates, despite the large number of qualitative observations on growth of product phases by diffusional processes. There are a few such quantitative measurements however and these will be briefly described.

2.2.5.1. Interface-controlled growth rates, without change of composition. The best data for this type of reaction are probably those for grain-boundary migration in high-purity alloys (ch. 28, § 3.4.1) It appears from the current state of knowledge, from the work of AUST and RUTTER [1963] and the review by HAESSNER and HOFMANN [1978], that the single-atom model of eq. (27) does appear to describe the process very well, with activation energies of the order of those expected for grain boundary diffusion. However, grain-boundary migration experiments are often complicated by solute-drag effects and also by the problem of knowing what is the magnitude of the free energy

driving force. Solute drag effects (Chapter 28) arise from the usual desire of solute to adsorb preferentially on a grain boundary particularly when the solute and solvent atoms have different sizes. When the grain boundary tries to migrate the solute can be partially left behind and thereby exert a drag on the boundary which therefore moves more slowly than in the case of very high purity materials when the boundary can quite easily escape the solute field. Despite some uncertainties about the details of solute drag model, it seems safe to conclude from the present state of knowledge that the basic model for mobility of an incoherent boundary is in quite good shape. Low-angle grain boundaries and *coherent* twin boundaries show very low mobility (ch. 28, §§ 2.4.2 and 3.4.2). The low mobility of the twin boundaries appears to arise from the need for a ledge mechanisms for migration combined with a low density of the required growth ledges.

For polymorphous phase transformations there are considerable difficulties in determining the growth rates, since in many cases the kinetics of the process are apparently controlled by nucleation. A clear example of this is provided by the numerous studies of the fcc–bcc phase transition in iron with a range of trace alloy additions. WILSON *et al.* [1979] briefly reviewed the field and presented additional data. They report a series of thermal arrests at different cooling rates for high-purity iron and mild steel; similar results have been previously found. Two transformations occurring at the highest temperatures, approximately 840 and 735°C, are both described by the authors as occurring by the massive transformation mechanism, the product in each case being apparently ferrite of the same composition as the austenite. The distinction is due apparently to different nucleation mechanisms. The high-temperature transformation product is described as equiaxed, having a polygonal shape, and was treated by Wilson *et al.* as occurring with the volume stress relaxed, while the lower-temperature product is described as massive, and consists of small grains with an irregular, presumably facetted, outline. HILLERT [1995] reports four plateaux in very rapid cooling of pure iron: the first seen at the highest temperature is attributed to equiaxed (massive) growth of iron, the second at faster cooling and lower temperatures is attributed to plate-like (Widmanstätten) growth, the next is lath (dislocated) martensite and the one at the fastest cooling rates and at the lowest temperature is plate (twinned) martensite. The distinction between the first two is apparently that the massive reaction takes place by migration of the incoherent interface of the bcc ferrite into the fcc austenite grain with which it does *not* have a good fit orientation relationship. The plate-like growth of Widmanstätten ferrite would then be the growth of grain boundary nucleated ferrite into the fcc austenite with which it *does* have the good fit orientation relationship. It would seem that the plateau growth, where the growth rate becomes faster at essentially constant supercooling, should in both these cases be *partitionless*, certainly with respect to substitutional impurities if not the fast diffusing interstitial impurities. The two martensitic reactions will be displacive and partitionless with respect to all impurities. The fascinating distinction is that between plateaux #2 and #3 Widmanstätten ferrite and lath martensite in nominally pure iron. Since there are two plateaux there has to be a distinction between them — this could be either that the Widmanstätten ferrite is not displacive (§ 2.6) or that it is not partitionless with respect to the rapidly diffusing interstitial impurities. (It could, of course, also be different in both respects.) It was reported by HILLERT [1995]

that when trace quantities of carbon, upto 0.02wt%, were added to the nominally pure iron, there was a more rapid fall in the plateau temperatures for reactions #1 and #2 than for the lath martensite reaction #3. This result was analysed by Hillert as showing growth of the first two reaction under local equilibrium, diffusion control, for carbon. That is, there was a "spike" of carbon ahead of the growing interface, giving a carbon concentration of the ferrite equal to that of the alloy, with the austenite being carbon enriched, but having the same carbon *activity*, as the ferrite across the interface. The width of the carbon enrichment spike in the austenite would be of the order of $D_\gamma(C)/v$. $D_\gamma(C)$ is the carbon diffusion coefficient in austenite and v is the interface growth velocity. Since the diffusion coefficients of the substitutional elements are orders of magnitude smaller than that of interstitial carbon their spikes will be of negligible widths — indicating true partitionless growth of ferrite for the substitution impurities. However additions of nickel had been previously shown by SWANSON and PARR [1964] to have equivalent effects in lowering the plateau temperatures of reactions #1 and #2. HILLERT [1995] takes this result as confirmation of a solute drag effect in the massive and Widmanstätten ferrite growth reactions, see § 2.6. It is not clear however whether the effect of nickel was not merely the slope of the T_0 line, fig. 18, for Fe–Ni.

A detailed general review of the topic of *massive transformations* was provided by MASSALSKI [1970]: Details of the various structural features in this type of transformation can be obtained from this source. One early investigation that did provide quantitative data on the diffusionless type of transformation is the study of the massive transformation in copper-zinc alloys by KARLYN *et al.* [1969]. The authors quenched bcc brass containing 38at% Zn to room temperature and observed the transformation products on briefly up-quenching samples by pulse-heating. They only obtained the diffusionless fcc product when the sample was held in the single-phase α region of the phase diagram and it was not found when samples were held in region II of the phase diagram (fig. 18, above) where it is expected that the diffusionless transformation *should* occur. The explanation offered for this failure to see the transformation in region II of the two-phase field, when it is thermodynamically possible was that of a nucleation difficulty. It was hypothesized that only pre-existing nuclei would have had time to develop, and since these pre-existing nuclei had been held in the two-phase region, they were all surrounded by rejected solute. The effect of this rejected solute would be to cause the local concentration at the interface to be outside region II of the phase diagram. When the samples were heated into the single-phase field, however, the nuclei would, after a brief time delay to reabsorb the excess zinc solute, be able to grow diffusionlessly. This was what actually was seen to happen and the authors were able to describe the growth rate of the α grains by eq. (27) with the low diffusional activation energy (for copper alloys) of only about 60 kJ/mole. This value is well below that expected for bulk diffusion of copper (200 kJ/mole) and lower even than that expected for grain-boundary diffusion (80–100 kJ/mole).

PLICHTA *et al.* [1978] compared the growth rates found in a series of massive reactions in titanium binary alloys with previous studies of the massive transformation in other systems. They showed, for all the systems, that the activation energy was significantly less than expected for bulk diffusion. This result was obtained by use of

eq. (27) and again provides evidence that this equation does indeed adequately describe interface-controlled partitionless transformations.

2.2.5.2. Reactions involving long-range solute diffusion. The range of quantitative experimental data available prior to 1968 was reviewed by AARONSON *et al.* [1970]. They reported that in most cases the experimental studies showed good agreement with diffusion-controlled growth models, such as were described in § 2.2.2. These authors also considered additional mechanisms such as the growth of grain-boundary precipitates, the theories for which were not included in § 2.2.2. In this modification a precipitate on a grain boundary is fed solute by diffusion over a short distance in the matrix to the grain boundary followed by rapid grain-boundary diffusion to the precipitate. This "collector plate" mechanism gives faster growth than expected for solute transport solely by volume diffusion. For several important examples of thickening of Widmanstätten plates, for example in Fe–C, Cu–Si and Al–Cu, AARONSON *et al.* [1970] reported evidence for ledged growth with the growth data showed signs of growth slower than expected for diffusion control. Later investigations of precipitation of the hexagonal phase from the fcc matrix phase in Cu–Si by KINSMAN *et al.* [1971, 1973] and by TIEN *et al.* [1971, 1973] again showed in this case apparent control not by silicon diffusion to the growing hexagonal phase but control by interface processes that allowed the ledge mechanism to operate. Tien *et al.* used an electrical-resistivity technique for measuring solute content in the matrix and at small driving forces suggested a dislocation pole mechanism similar to that used to interpret twinning in fcc metals (e.g., KELLY and GROVES [1970]). The twinning dislocation was an $a/6$ <112> partial. The Cu–Si phase diagram is an interesting example of the type of system which has a very low driving force for *nucleation* since the two phases, cubic α and hexagonal κ, are very similar in composition [eq. (6d)]. The microstructure of the precipitated two-phase structure shows the hexagonal plates running completely across the matrix grains. As a result this is a rather unusual precipitation reaction, though somewhat similar to Widmanstätten ferrite in the iron-carbon system in having a low driving force for nucleation but a large supersaturation for growth. The precipitates in Cu–Si nucleate apparently at stacking faults while the iron system requires grain-boundary nucleation. The formation of precipitates on a single variant of the <111> habit planes means that there will be no intersections with other precipitates. Such intersections, as noted in § 2.2.4, appear to allow the formation of thickening ledges. Lack of facility in forming the growth ledges will account for growth at interface-limited rates slower than expected for full diffusion control. Recent observation on the growth of face-centered tetragonal γ-phase from hexagonal α in Ti–Al by SADLER–HOLM [1994] has shown that in this system there is also a lack of precipitate intersections. The hexagonal phase has only the single close packed (00.1) habit plane so no intersections are possible for fcc precipitates that form on this single variant cph plane. In the resulting so-called "lamellar" structure the inhibited thickening of the γ precipitate phase lead to retained aluminum supersaturation in the hexagonal phase particularly after it has transformed to the ordered α_2 structure.

In the more usual precipitation reaction where a high density of intersecting plate-like precipitates form, for example the θ'-phase in Al–Cu or γ' Ag_2Al in Al–Ag, there has been a considerable amount of work and resulting controversy. SANKARAN and LAIRD

[1974] extended the observations of AARONSON and LAIRD [1968] and claimed that the results showed inhibited thickening compared to the rates predicted by the diffusion-control model. They also reported that the lengthening reaction was faster than that expected for linear unstable lengthening (§ 2.2.6). CHEN and DOHERTY [1977] reanalyzed the earlier data on the growth kinetics of θ' in Al–Cu and suggested that the results appeared to fit quite well to the HAM [1958, 1959] and HORVAY and CAHN [1961] diffusion-controlled model for growth with a constant aspect ratio [eqs. (48 and 49)]. This led, not unexpectedly, to a considerable controversy between the various authors (SANKARAN and LAIRD [1977, 1978], AARONSON [1977] and DOHERTY *et al.* [1977, 1978b]). It is probably impossible for someone deeply involved in a scientific argument to give a really objective account of the situation but it appears that it is reasonable to say that the data presented agree surprisingly well with the diffusion-controlled model, that is, with no significant inhibition of the growth rate, but with very large aspect ratios. These high length-to-width shapes then require some significant *initial* inhibition of thickening at the early stages of precipitate growth.

The situation is perhaps easier to understand in the studies of precipitate growth in the Al–Ag system (LAIRD and AARONSON [1969], FERRANTE and DOHERTY [1976, 1979], and RAJAB and DOHERTY [1982, 1989]). Figure 29 shows the observed growth kinetics in a particular experiment and indicates that the results show the square root of time relationship expected for diffusion control. Moreover, the *slopes* of the size/$t^{1/2}$ plots are close to those predicted by HAM [1958, 1959], and by HORVAY and CAHN [1961] for diffusion control [eqs. (48) and (49)]. However, in this and other experiments, the observed value of the aspect ratio, A, varied widely with the precipitate growth conditions and, in every case, A was much greater than the value expected for precipitates with the equilibrium shape determined by interfacial energies.

Ferrante and Doherty postulated that the aspect ratio was increased by immobility in the thickening reaction at the *earliest stages of the reaction*, before the measurements shown by fig. 29 could be made. Under these conditions the aspect ratio would increase rapidly, as had been previously argued by Aaronson.

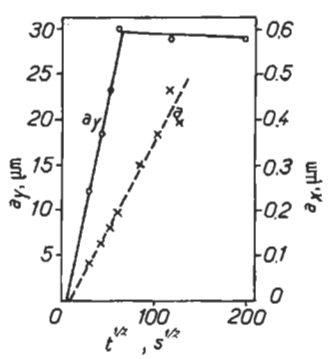


Fig. 29. Half-lengths, a_y, and half-thickness, a_x, of plate-like precipitates of Ag_2Al growing in aluminium, as a function of time$^{1/2}$ (after FERRANTE and DOHERTY [1979]).

References: p. 1497.

The details of this hypothesis are easily given: The diffusion-controlled thickening rates are given by:

$$da_x/dt = (D\beta/t)^{1/2}, \tag{54}$$

$$da_y/dt = A(D\beta/t)^{1/2}, \tag{55}$$

These are the derived equations for growth *rates* that when integrated, give the growth equations (48) and (49), quoted above.

The rate of change in aspect ratio, dA/dt is easily derived from the definition of A:

$$A = a_y/a_x, \tag{56}$$

$$dA/dt = \frac{1}{a_x}\left(\frac{da_y}{dt} - A\frac{da_x}{dt}\right). \tag{56a}$$

Substitution of eqs. (54) and (55) into eq. (56a) shows, as had been pointed out by HAM [1958], that with diffusion-controlled lengthening and thickening starting from an initial aspect ratio A, *that shape will be preserved* ($dA/dt = 0$). However, if the thickening reaction is inhibited by some factor f, with $f < 1$, eq. (54) becomes:

$$da_x/dt = f(D\beta/t)^{1/2}, \tag{54a}$$

So, under this condition of inhibited thickening there will be an *increase* of the aspect ratio, given by:

$$dA/dt = \frac{A}{a_x}\left(\frac{D\beta}{t}\right)^{\frac{1}{2}}(1 - f). \tag{57}$$

FERRANTE and DOHERTY [1979] showed direct evidence for this effect by TEM observations at the earliest stages of precipitation, at shorter times than appear on fig. 29. The observations were of thickening slower than predicted, $F < 1$, and a steadily rising value of the aspect ratio, A [eq. (57)]. RAJAB and DOHERTY [1982, 1989] extended the observations in the Al–Ag system to a wider range of temperatures and supersaturations. They also found, for most of the precipitation reaction, diffusion-controlled growth rates. In addition, they found that the experimental value of the aspect ratio, A, though constant for a particular reaction condition, varied systematically with the *density* of precipitates. The aspect ratio *fell* as the density of precipitates *increased*. The authors argued that this variation in A was due to variation in the duration of the period of initial thickening inhibition, with $f < 1$, caused by a shortage of growth ledges. However, when precipitates on different habit planes intersected each other, which occurred earlier in the process, and so at smaller values of A, for *higher densities of precipitates*, then inhibition ceased and the precipitates *retained the value of A at the time of intersection*. The microstructural origin of this effect was assumed to be easier ledge nucleation at the intersection of two precipitates. Microstructural studies by RAJAB [1982], RAJAB and DOHERTY [1989] and particularly those of HOWE *et al.* [1985] gave experimental support for this hypothesis.

The current position on observation of precipitate growth kinetics suggests therefore that the theoretical ideas described in § 2.2 appear to be well founded. That is, most reactions are diffusion-controlled except those with a coherent interface, between different structures, when a ledge mechanism is required. Even with a ledged growth, however, diffusion-controlled growth is possible provided that there is a sufficiently active source of ledges. The major problem that remains is the lack of understanding about the mechanisms and kinetics of ledge formation. A similar conclusion appears to apply also for Ostwald ripening of plate-like precipitates (§ 2.4).

Finally, mention should be made again of the studies on nucleation and growth in the copper–cobalt system by SERVI and TURNBULL [1966]. Their observations showed clear evidence for diffusion-controlled growth of metastable fcc cobalt precipitates in the fcc matrix. This is a "cube-cube" orientation relationship between fully coherent precipitates which should not need growth ledges. The agreement with diffusion control in this system is very satisfactory confirmation of the mobility of interfaces between coherently connected structures when they have the same orientation and structure. The precipitate growth was studied in this system from the *earliest* stages of growth, at the smallest values of radius r, by virtue of the electrical resistivity technique used. Later studies on other systems with a very high density of homogeneously nucleated precipitates, such as those of XIAO and HAASEN [1992], while not indicating any significant interface control do not find the r^2/t growth kinetics of diffusion control. This result seems reasonable, since the r^2/t kinetics are only predicted for *independent* growth of *diffusionally isolated* precipitates. If the spacing Δ_3 between precipitates, after nucleation at a time t is less than $(Dt)^{1/2}$, then such isolated growth kinetics is not to be expected. The successful numerical analysis of nucleation, growth and coarsening by KAMPMANN and WAGNER [1984] and HAASEN and WAGNER [1992] in Cu–Ti and Ni–Al alloys were based on diffusion-controlled growth and the success of their models supports the expected (§ 2.2.3) lack of any interface immobility between the coherent (cube–cube) nuclei in both alloys.

2.2.6. Growth instabilities

This section is divided into two parts; firstly is a consideration of the onset of instability, and secondly is a discussion of the *linear* growth expected for well-developed instabilities, the lengthening of dendrites and Widmanstätten plates and needles occurring with *a constant tip or rim radius.*

2.2.6.1. Initial instability. For a spherical precipitate growing in a solute field (fig. 20, above) with either full diffusion control or even with mixed control, $C_\alpha \leq C_i \leq C_o$, the solute gradient sets up a potentially unstable situation, first analyzed by MULLINS and SEKERKA [1963]. If the sphere is perturbed by any oscillation, described mathematically as a spherical harmonic of amplitude δ_S, then those parts of the sphere with a locally larger value of the "radius", the distance from the center of the original sphere, (but a smaller radius of curvature, ρ) will penetrate into more supersaturated matrix and grow faster, while parts of the sphere with a smaller value of the radius but with a larger radius of curvature will grow more slowly, see fig. 30. As a result, the instability will grow forming a cellular or dendritic interface. This situation had been long recognized

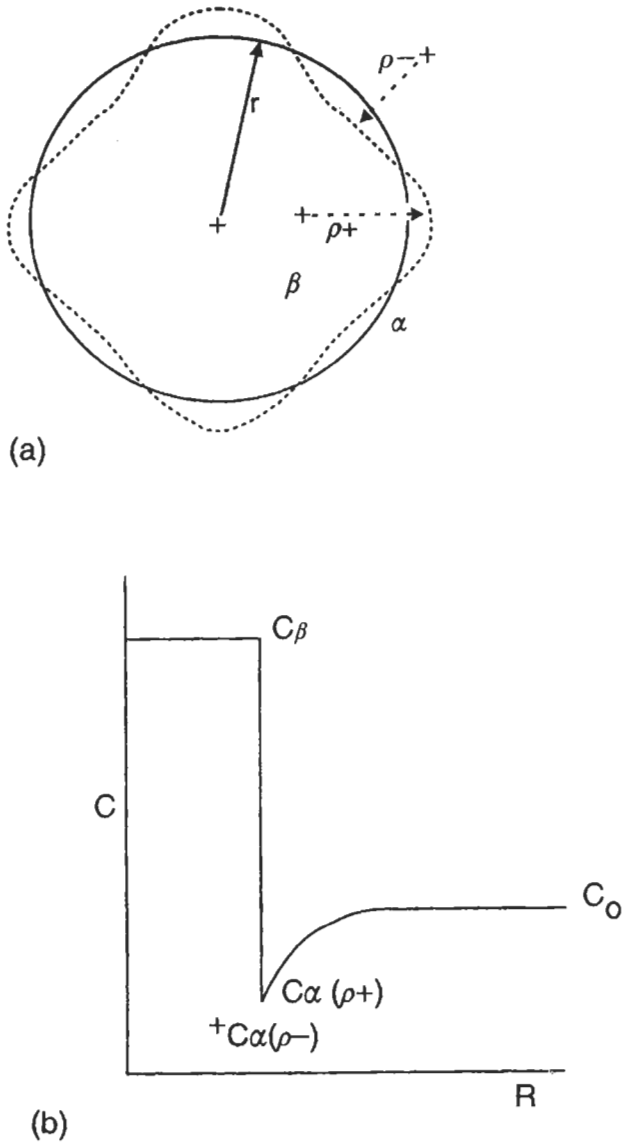


Fig. 30. (a) A sphere of unperturbed radius r (full line) perturbed (dotted line) by a spherical harmonic of $1=6$ in six "<100>" directions, four of which are shown. At the maxima the local radius of curvature has decreased to $\rho+$ while at the minima it has increased (and changed sign) to $\rho-$. The matrix solubility has increased to $C_\alpha(\rho+)$ at the maxima and decreased to $C_\alpha(\rho-)$ at the minima. (b) Solute C as a function of the radial distance R from the center of the β precipitate to the α matrix through one of the maxima with solubility ($C_\alpha(\rho+)$. $C_\alpha(\rho-)$, the solubility at the minima, is also shown at its radial distance from the origin.

in solidification and had been given a classical analysis by TILLER *et al.* [1963] as constitutional supercooling (ch. 8, § 6.4.1). In solidification, a temperature gradient provides some stability against growth of perturbations into the commonly observed growth form, a highly branched *dendrite*. For an *isothermal* precipitation reaction no such stabilizing temperature gradient exists. Mullins and Sekerka showed, however, how the interfacial energy will partially stabilize the growing sphere.

In the following discussion, only the case of the growth of a solute-rich precipitate by flux towards the growing precipitate will be considered. The opposite case, of growth of a solute-poor phase, for example, the growth of alpha-iron in austenite in Fe–C alloys, merely requires an appropriate change of notation.

To describe this stabilization, the Gibbs–Thomson equation for the increase of solubility of a solute-rich spherical precipitate of radius r is needed. This is given in the general case by PURDY [1971] (see also MARTIN and DOHERTY [1976, 1995])* as:

$$C_\alpha(r) = C_\alpha\left(1 + \frac{(1 - C_\alpha)}{(C_\beta - C_\alpha)}\frac{2\sigma V_m}{RTr\varepsilon_\alpha}\right), \tag{58}$$

where ε_α is the Darken non-ideality factor (§ 1.2). For dilute terminal solid solutions this reduces to the usual form:

$$C_\alpha(r) = C_\alpha\left(1 + \frac{2\sigma V_m}{RTr}\right). \tag{58a}$$

The effect of surface energy on solubility is such that when the precipitate nucleates at the critical radius r^*, $C_i = C_\alpha(r) = C_0$, and so, by eq. 38, there is zero net supersaturation, and the growth rate, dr/dt, is zero. Eq. 38 becomes:

$$dr/dt = \frac{D}{r}\frac{C_0 - C_\alpha(r)}{C_\beta(r) - C_\alpha(r)}. \tag{59}$$

Substitution of eq. (58a) into eq. (59) gives:

$$dr/dt = \frac{2D\sigma V_m C_\alpha}{(C_\beta - C_\alpha)RT}\frac{1}{r}\left(\frac{1}{r^*} - \frac{1}{r}\right). \tag{60}$$

Here $C_\beta(r) - C_\alpha(r) \approx C_\beta - C_\alpha$ since both compositions shift in the same direction with the Gibbs–Thomson effect.

This growth rate is plotted out in fig. 31, which shows that the growth rate rises to a maximum value at $r = 2r^*$, and then declines, becoming that given by eq. (41) only for $r \gg r^*$.

Mullins and Sekerka showed that for this type of diffusion-controlled growth, the

* Figure 1 used to describe the driving force for precipitation is readily converted to be used to derive the Gibbs–Thomson equation. For the Gibbs–Thomson equation, C_0 becomes $C_\alpha(r)$ and $L \rightarrow K$ becomes $2\sigma V_m/r$, the increase in free energy per mole ($dG/dn = dG/dr.\ dr/dn$) due to the curvature of spherical precipitates of radius r.

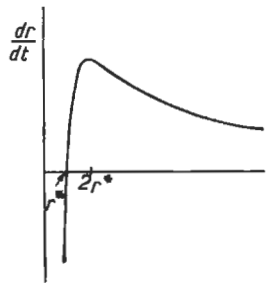


Fig. 31. Variation of growth rate of a spherical precipitate radius at a fixed supersaturation. At $r = r^*$, the critical radius for nucleation, the growth rate is zero.

spherical shape was stable against all perturbations until it reached a limiting radius for absolute stability, $r_1 = 7r^*$. The physical origin of the stability is that while an unperturbed sphere has a constant radius and thus solubility, at the maximum of the perturbation in fig. 30, the radius of curvature ρ, which is the radius to be used in eq. (58) for non spherical precipitates, *decreases*, below the radius, r, of the unperturbed spherical precipitate while at the minimum in the perturbation the radius of curvature *increases*. As a result of the Gibbs–Thomson effect, the solubility rises at the maxima and decreases at the minima. This creates a concentration difference that opposes the development of the perturbation. (The radius of curvature of a surface is the mean of any two radii that are perpendicular to each other — $1/\rho = 1/r_1 + 1/r_2$.)

The condition of *absolute instability* is that a perturbation of amplitude δ_S will grow, that is $d\delta_S/dt > 0$. A spherical harmonic with two maxima is the smallest to show this type of instability. A spherical harmonic with two maxima, $l = 2$, is a sphere perturbed to an oblate spheroid – an ellipsoid of revolution†. This has a shape similar to that of either a rugby or an american football. Higher spherical harmonics are spheres distorted with increasing numbers of maxima — given by the parameter l such that with $l = 6$ a distortion is capable of giving maxima in 6 directions equivalent to the cubic crystallographic directions <100>, fig. 30. Growth of this will give the shape seen in cubic solidification dendrites. CORIELL and PARKER [1967] extended Mullins and Serkerka's analysis and showed that absolute stability was less important than relative stability. The criterion for a sphere to show *relative instability* is that

$$\frac{1}{\delta_S}\frac{d\delta_S}{dt} \geq \frac{1}{r}\frac{dr}{dt}. \tag{61}$$

This criterion arises since if the relative increase, $(1/\delta_S)(d\delta_S/dt)$ is > 0 but $< (1/r)dr/dt$, then although the sphere is unstable, the observed shape will become apparently *more spherical*, rather than less, with time. The smallest radius that shows relative instability is the $l = 3$ spherical harmonic with three maxima; this occurs when $r \geq 36r^*$. When $l = 6$,

† $l = 1$ merely displaces the sphere.

giving the 6 <100> orthogonal dendritic arms, $r \geq 66r^*$.*

The analysis above is a limiting case since, as discussed by SHEWMON [1965], there are other stabilizing factors that must also be considered, particularly for solid-state reactions. Shewmon developed his analysis since he could find no examples of solid-state dendritic growth, and wished to understand why. However, dendritic growth shapes are now known, and have been studied, in a few metallic solids by MALCOLM and PURDY [1967], BAINBRIDGE and DOHERTY [1969], BAINBRIDGE [1972], FAIRS [1977], RICKS *et al.* [1983] and BAUMANN and WILLIAMS [1985]. Despite these examples of sold state dendritic growth, the problem was only changed to — why were there so *few* examples of solid-state dendrites?

The additional stabilizing factors that inhibit growth perturbations include the following:

(i) strong anisotropy in the interfacial energy — for a facet plane;
(ii) interface immobility;
(iii) interfacial diffusion;
(iv) diffusion within the precipitate;
(v) highly anisotropic precipitate shape — high initial aspect ratios;
(vi) closely spaced precipitates;

For metallic *solidification*, where dendrites are ubiquitous, none of the first five stabilizing influences are significant, while most solid-state diffusional reactions do have some or all of these stabilizing factors.

The first four of these effects were discussed by SHEWMON [1965] and each was shown to be strongly stabilizing. No quantitative analysis was made for (i) the stability of an interface at an energy cusp orientation giving a facetted plane (figs. 10 and 11). The physical basis of the enhanced stability due to anisotropy is easily seen, however. Any shape perturbation of a facetted plane would have to increase not only the interfacial *area*, the usual problem for a perturbation, but a perturbation at a cusp *also* requires the interface to rotate away from the cusp plane and thereby greatly increase the specific interfacial *energy*. Partial interface immobility (ii), giving mixed control, will also stabilize the precipitate since it reduces the concentration gradients of solute away from the interface, these gradients cause the instability. High diffusivity at the interface (iii), or within the precipitate (iv), relative to the diffusivity in the matrix phase is stabilizing

* MULLINS and SEKERKA [1963] had suggested that their analysis indicated that the growth equations of Ham, and Horvay and Cahn, eqs. 48 and 49, would be unlikely to apply since the spheroid shapes described should be absolutely unstable. As a result the length/$t^{1/2}$ kinetics should transform into the linear kinetics discussed in § 2.2.6.2. This is clearly experimentally incorrect, § 2.2.5.2. The origin of the mistake by Mullins and Sekerka is not obvious but appears to arise for at least two reasons. The first is the relative stability analysis of CORIELL and PARKER [1967]. The plate like precipitates of Ham, Horvay and Cahn are approximately highly distorted spheres with $l = 2$ — such shapes though absolutely unstable are relatively stable. As a result, even if capillarity is added to Ham, Horvay and Cahn analysis, the predicted shape stability may not change. In addition, as discussed below, there is no good analysis of shape stability in the presence of high anisotropy of the interfacial energy. The approximate analysis offered by DOHERTY [1982], see § 2.2.6.2., suggests that high aspect ratio precipitates should be stable against linear lengthening except under high supersaturations, $\Omega \geq 0.5$, much higher than the typical values of $\Omega \leq 0.1$, found in precipitation reactions.

References: p. 1497.

since solute flux along these pathways can reduce the interfacial area but unlike diffusion in the matrix it does not contribute to precipitate growth. An analysis of surface smoothing, in the absence of any perturbing gradients, was given by MULLINS [1959] for the various diffusion pathways. On any surface, variations in curvature are removed by diffusion in both phases and along the interface unless the interfacial energy stabilization is overcome by the gradient instability.

The effect of an anisotropic precipitate shape (v) on shape stability has not been formally analyzed with a perturbation technique. DOHERTY [1982], however, indicated why a shape with a high aspect ratio should be stable by use of the equations developed in § 2.2.2 for stable growth by comparison with the expected linear growth rate of the end of a Widmanstätten plate. The analysis is given in the discussion of linear growth models (§ 2.2.6.2). Linear growth of Widmanstätten plates (PURDY [1971]) as in the linear growth of dendrites, MALCOLM and PURDY [1967], DOHERTY *et al.* [1978] is the usual growth consequence for the tips of the instability. No full analysis of the transition from developing instability to linear dendritic growth appears yet been made. The current models of dendritic growth, due to DOHERTY *et al.* [1978], LANGER and MULLER–KRUMBHAAR [1978], HUANG and GLICKSMAN [1981], KURZ and FISHER [1989], find that the tips of dendrites have a radius of curvature that is determined by an instability condition.

The final stabilizing influence of closely spaced precipitates (vi) is easily understood, though there is apparently no proper analysis of the stability of closely spaced precipitates. The effect arises since closely spaced precipitates show an overlap of the diffusion fields of adjacent particles (the *soft impingement* effect). This effect, for two precipitates spaced less than $(Dt)^{0.5}$ apart, reduces the concentration gradients about each precipitate, see fig. 20, as $C_0 = C_\infty \rightarrow C_\alpha$. Under these circumstances the composition gradient driving the instability falls, and the closely spaced spherical precipitates will be stabilized. The result of this is seen for solidification either whenever there is a high rate of primary nucleation (KATTAMIS *et al.* [1967]) or by secondary nucleation in stir-casting (see VOGEL [1977], and DOHERTY *et al.* [1984]). Under these conditions of closely spaced growth centres, the solidification growth form is found to be spherical, not dendritic. This was nicely illustrated for the Al–Al_3Li solid state case studied by BAUMANN and WILLIAMS [1985]. With high supercooling below the solvus, the coherent ordered Al_3Li precipitates, with a high spatial density; were found to be spherical. With lower supercoolings, giving more widely spaced precipitates, these showed a clear dendritic growth instability.

The combination of all these effects appears to explain the lack of shape instabilities, at least those that might give true dendritic shapes, in all but a few solid-state reactions. The few dendritic shapes that have been found occur with cube-cube oriented precipitates that have *low and isotropic interfacial energies* and *low interfacial diffusivities*; in addition, the precipitates all have ordered structures, which have low rates of diffusion *within the precipitates* and the examples giving dendritic shapes all occur with *low densities of precipitates*. The clearest examples were of γ-brass in β-brass (MALCOLM and PURDY [1967] and BAINBRIDGE and DOHERTY [1969]) and γ' in γ, Ni_3Al in Ni (RICKS *et al.* [1983]) and the equivalent structure of δ' in α, Al_3Li in Al, BAUMANN and

WILLIAMS [1985] the last two at small undercoolings. These examples all have, coherent low diffusivity interfaces, ordered low diffusivity precipitates and low precipitate densities. Deliberate attempts to produce shape instabilities in other cube-cube systems, for example for silver precipitates in copper (DOHERTY [1976]), failed, presumably owing firstly to the high rates of interfacial diffusion expected in the semicoherent interface containing a high density of edge dislocations between silver and copper, where there is a *large misfit* between the two phases, and secondly to the significant rate of diffusion within the *disordered* silver rich precipitates.

2.2.6.2. Linear growth models. For both dendrites and plate-like and needle-like Widmanstätten growth forms, there are well developed models for the linear growth of the tips based on the assumption that the *tip radius does not change*. For the tip of a needle, the radius is the same in all directions, while for the rim of the plate the constant radius is the small one between the two parallel habit planes and not the large radius of the growing cylindrical plate. The earliest model for lengthening of a plate was given by ZENER [1946] and modified by HILLERT [1957]. The essential features of the model are the assumption of a constant tip radius which will give, for a needle crystal, a growth rate given by a slightly modified form of eq. (41)

$$\mathrm{d}r/\mathrm{d}t = 2\frac{D}{r}\frac{C_0 - C_\alpha}{C_\beta - C_\alpha} = \frac{2D\Omega}{r} \tag{41a}$$

The extra factor of two in the eq. 41a, compared to eq. 41, comes from the difference in growth geometry. For the growth of a sphere (eq. 41) the flux of solute, $D(\mathrm{d}C/\mathrm{d}R)_{R=r}$, is applied over the whole surface area, $4\pi r^2$, and all of the surface is growing *normal* to the surface. For the linear lengthening of a circular shaped needle crystal the flux is delivered to one side, so the solute flux is fed onto an area of $2\pi r^2$, but *the area that is fed by the flux is only the cross sectional area of the needle*, πr^2, fig. 32, so the growth rate is thereby doubled.

In the solidification literature, see for example, DOHERTY [1980] or KURZ and FISHER [1989], the velocity, $v = \mathrm{d}r/\mathrm{d}t$ is made dimensionless as the Peclet number P in eq. 41b.

$$P = \frac{vr}{2D} = \frac{C_0 - C_\alpha}{C_\beta - C_\alpha} = \Omega \tag{41b}$$

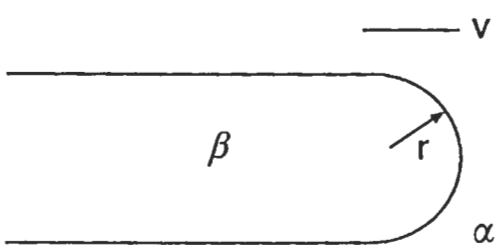


Fig. 32. Linear lengthening at a constant growth velocity v, and constant tip radius of a cylindrical needle crystal of β with a hemispherical tip of radius r growing in and a supersaturated α matrix.

However, the effect of the surface energy has also to be included as in eq. 59 and so we obtain:

$$dr/dt = \frac{2D}{r}\frac{C_0 - C_\alpha(r)}{C_\beta(r) - C_\alpha(r)}. \tag{59b}$$

This gives the same plot as that of fig. 31 apart from the factor of two. For plate-like crystals, the equations are slightly different, but the physical reasoning and the resultant form of the relationship, between velocity and plate-tip radius are very similar to that seen for the simple needle crystal.

Having produced such relationships as those of eq. 59b etc., it is necessary for the system, and thus any theory of the process, to select the particular value of the radius for the plate-like or needle-like crystal. Any value of r is an equally valid solution to a diffusion equation, such as eq. (59b). For convenience and for lack of any other simple choice, the initial theory selected the maximum-velocity criterion (HILLERT [1957]). As explained in reviews for plate-lengthening, for example by HILLERT [1975], and for needle and dendritic growth, for example by DOHERTY [1980], the simple theory has proved inadequate and has had to be modified. Two major modifications have been required: The first is for a more accurate diffusion model and the second is the adoption of a more successful determining principle for tip radius. Both of these modifications are described below.

(1) *A better diffusional model.* For the *needle-like* crystal the required diffusional solution, neglecting interfacial energy considerations, is that due to IVANTSOV [1947]. The Peclet number P is given by:

$$\Omega = P\exp(P)E_1(P). \tag{62}$$

$E_1(P)$ is the integral exponential function, see for example KURZ and FISHER [1989]. For a *plate-like* crystal the equivalent expression, IVANTSOV [1947], is:

$$\Omega = (\pi P)^{1/2}\exp(P)\operatorname{erfc}\left(P^{1/2}\right). \tag{63}$$

This gives a unique value of P ($P = vr/2D$) for a given supersaturation Ω but still leaves the choice of the radius r and thus the growth velocity v undetermined. TRIVEDI [1970a] solved the problem after introducing the surface energy and thus the Gibbs–Thomson modification of solubilities and selecting a tip radius that yielded the maximum growth rate. Trivedi's rather complex equation can be reduced at low-to-medium values of the supersaturation Ω by the the approximation due to BOSZE and TRIVEDI [1974]:

$$v = \frac{27D\Omega_{BT}^3}{256\pi r^*}. \tag{64}$$

The Bosze and Trivedi supersaturation, Ω_{BT}, is related to the supersaturation, Ω, by:

$$\Omega_{BT} = \Omega/\left(1 - \frac{2\Omega}{\pi} - \frac{\Omega^2}{2\pi}\right). \tag{64a}$$

r^* is the critical radius of nucleation, eq. 10a, where $C_o = C_\alpha(r)$ and is often related to the supersaturation by:

$$r^* = \Gamma/\Omega. \tag{64b}$$

The capillarity constant, Γ, is given for precipitation by:

$$\Gamma = \frac{2\sigma V_m C_\alpha (1 - C_\alpha)}{RT(C_\beta - C_\alpha)^2 C_\beta}. \tag{64c}$$

HILLERT [1975 and 1994] has shown that eq. (64) gives good agreement with the observed rates of linear lengthening of Widmanstätten α-iron plates in Fe–C alloys and PURDY [1971] showed equivalent agreement for the needle-like α-brass precipitates forming from β-brass in Cu–Zn alloys. In both these cases the precipitates grow from grain-boundary nucleation sites since, the composition differences between the two phases being small, there is insufficient driving force for nucleation within grains [eqs. (6d) and (7)].

(2) *Determining principle for tip radius.* Despite the obvious attraction of analytical simplicity of choosing the radius that gives the maximum growth rate, there is no reason why nature should share this outlook. Various other determining principles have been suggested in the past, mainly for the linear growth of two-phase products, eutectics, eutectoids and cellular precipitation reactions, see § 2.5 (PULS and KIRKALDY [1972] and LIVINGSTON and CAHN [1974]). In her studies, on solid state dendritic growth of ordered γ dendrites in β-brass, FAIRS [1977] measured both the linear growth rate and tip radius for a range of precipitation conditions. She observed that the tip radius was larger than that expected to give the maximum growth rate. As a result the growth rate was, of course, less than the expected maximum velocity. These experimental solid state observations were matched by the detailed studies by GLICKSMAN *et al.* [1976] on the *solidification* of thermal dendrites in undercooled succinonitrile, a low entropy of melting, transparent, organic analogue for metals. In their succinonitrile study, the authors reported that the growth velocities were also significantly less than expected for the maximum velocity. The data on succinonitrile were reanalyzed by DOHERTY *et al.* [1978a] who showed that the effect observed by FAIRS [1977] was occurring in the organic solidification dendrites: The tip radius was significantly larger than that which would give the maximum velocity but the data lay on the plot predicted by the Trivedi equation (TRIVEDI [1970b]) for dendritic or needle growth.

Fairs showed both in her solid-state dendrites and in the organic dendrites that the operating tip radius seems to be determined by a *stability* criterion. This was originally thought to be relative stability but, as argued by HUANG and GLICKSMAN [1981], absolute stability is more likely to be appropriate for a perturbation on a tip with a *constant tip radius*. It is, however, still not intuitively obvious why stability rather than the maximum-velocity criterion controls the tip radius and therefore the growth velocity. LANGER and MULLER–KRUMBHAAR [1978] developed a detailed theory for dendritic growth with the tip radius determined by a stability criterion and their general theory appears to be very successful. According to this description, the dendritic or needle growth rate is determined from the Ivantsov equation for a needle crystal, eq. (62). Peclet number, P, determined by eq. (62) is determined by:

$$P = vr/2D = (v\Gamma/2D\sigma'). \tag{65}$$

References: p. 1497.

Here σ' is a stability constant that is found to be 0.025. A similar analysis should be feasible for determining the operating rim radius for the growth of plate-like precipitates — however no such analysis appears to have been published — nor have any detailed experimental studies been published aimed at determining the operating conditions for the linear growth of plate-like precipitates. However OLSON *et al.* [1990] have described, see § 2.6, an interesting new approach to this problem for bainite growth.

In the current analysis of dendritic growth, there is still no adequate treatment showing how an initial instability in a spherical precipitate develops into a branched dendrite, whose linear growth can be described by stationary diffusion fields around the constant tip shape [eqs. (64) and (65)]. Since the tip radius in both cases is determined by stability — relative stability for the initial perturbation and absolute stability for the linear growing needle — it is possible that a nearly constant tip radius applies through out the transition. However there is no theoretical or experimental support for this speculation. An even more striking problem is the lack of a stability theory to describe how a plate-like Widmanstätten precipitate initially growing, like a sphere with a size-root-time relationship, can first evolve a perturbation at the rim, and then subsequently adopt the constant growth rate and tip radius experimentally found (HILLERT [1975]). The present author (DOHERTY [1982]) suggested a possibility for what it might be that determines the change from growth with a stable shape, with a constant A in eqs. (49) and (55), to an unstable *linear* growth, given for example by eq. (63), with an increasing A. Doherty suggested that the change might be expected to occur when the linear growth rate becomes *faster* then the instantaneous lengthening of the stable shape. Eq. (55) shows that the instantaneous growth rate of a stable plate-like precipitate falls with increased time and thus precipitate size. If the hypothesis is correct, then the value of the aspect ratio, A, will be crucial since large values of A, that promote stable growth, will greatly delay the transition to linear lengthening. The other factor of crucial importance is the supersaturation, Ω. Table 2 gives the half-length a_y^* for which the Bosze and Trivedi relationship [eq. (63)] predicts the same lengthening rate ($v = da_y/dt$) as the stable Horvay and Cahn result. The results in table 2 were obtained for a typical value of the aspect ratio seen for precipitates nucleated within a grain, $A = 100$, and a capillary constant, Γ, of 0.6 nm, needed to determine the critical radius of nucleation r^*. For grain-boundary nucleated precipitates a smaller value of supersaturation will arise if

Table 2
Plate half-length for which eqs. (49) and (63) predict the same lengtening rate.

Supersaturation Ω	Supersaturation Ω^*	Precipitate tip radius r, nm	Growth parameter β	Plate half-length a_y^*, μm
0.01	0.01	$1060r_c = 6.0\times10^4$	7×10^{-5}	2.3×10^{-6}
0.03	0.0306	$306r_c = 7\times10^3$	3×10^{-4}	0.12×10^{-6}
0.1	0.107	$100r_c = 6\times10^2$	4×10^{-3}	1.2×10^{-4}
0.3	0.38	$28r_c = 60$	5×10^{-2}	1.2×10^{-3}
0.6	1.07	$10r_c = 10$	1×10^{-1}	50

For $a_y > a_y^*$ linear lengthening would be expected.

smaller initial values of A were used.

The huge values of a_y* at the low supersaturations typical of age-hardening precipitations of solute-rich precipitates from solution-treated alloys, $\Omega \leq 0.3$, suggest why linear lengthening is not seen in such precipitation reactions as Ag_2Ag from Al and θ' (Al_2Cu) in Al–Cu (§ 2.2.5). However, matrix phase transformations, such as the formation of ferrite from austenite, usually do occur at large supersaturations, $\Omega \geq 0.6$, but with low nucleation densities (§ 1.3). In such a case, the values of a_y* become smaller than the matrix grain size so the observation of linear unstable growth of Widmanstätten plates in this system is then understandable. This can only be a tentative description, since the analysis is approximate and the operating value of A in the stable growth regimes of Widmanstätten ferrite is not known. Experimentally, however, observation on the early stages of the formation of Widmanstätten plates at grain boundaries should be relatively straightforward. In addition, although the Trivedi analysis for linear growth rates have been corrected for interfacial energy effects, the growth of the stable precipitates in eq. 49 has not been corrected for capillarity. In the case of the most common examples of linear growth of Widmanstätten precipitates, in bcc iron precipitating from fcc iron in Fe–C alloys and in fcc α-brass forming from bcc β-brass, the plates develop from grain-boundary nucleated films.

TOWNSEND and KIRKALDY [1968] observed the development of this instability and suggested that it could arise as seen in fig. 33. It is assumed that the bcc ferrite forms in (a) on a grain boundary with the good fit plane of the precipitate nearly parallel to the matching plane in the upper grain. Growth continues along the grain boundary in (b) but when the curved grain boundary moves away from the good fit interface, further growth, (c to e), allows instabilities to develop allowing the plates to form with the low energy habit planes shown. This model nicely matched the microstructures seen. Development of instabilities in this form should allow linear growth to occur at smaller supersaturations and sizes and in the analysis of growth from a sphere (table 2). TOWNSEND and KIRKALDY [1968] and HILLERT [1975] both find linear lengthening at smaller values of Ω than the minimum of about 0.6 suggested by the simple analysis of table 2, using the value of A of 100.

2.3. Precipitate dissolution

There has been much less interest in the dissolution of precipitates than in their growth for the obvious reason that, provided an alloy is heated into a single-phase region, all the precipitates will dissolve, given sufficient time. As a result, the microstructure of the final heat treated material does not carry with it the microstructural consequences of the kinetics of dissolution. This is not the case for precipitation where the microstructure is determined by the kinetics of the reaction. However, for precipitates *that are not completely soluble* after an increase of temperature, there will be a period of dissolution that will affect the microstructure because of the different shrinkage rates of different precipitates. In addition, observation of precipitate dissolution can provide insight into the amount of diffusion control. This idea was exploited by HALL and HAWORTH [1970] and ABBOTT and HAWORTH [1973] in studies of the dissolution of

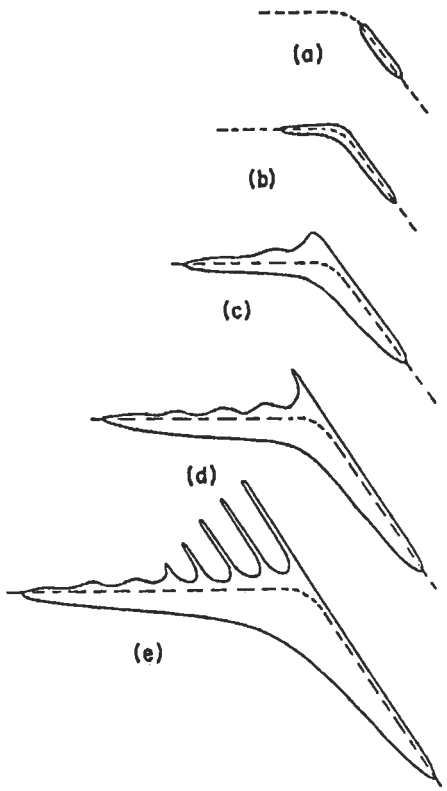


Fig. 33. Development of grain-boundary-nucleated Widmanstätten plates: (a) Grain-boundary allotriomorph, with good fit plane, with the upper grain, almost parallel to the grain boundary; (b) Precipitate growth around a curved grain boundary; (c) to (e) Growth of instabilities evolving into linear lengthening Widmanstätten plates. After TOWNSEND and KIRKALDY [1968].

plate-like precipitates in the Al–Cu and Al–Ag systems by microprobe analysis of the solute fields in the matrix normal to the habit plane of the precipitates. In the first example, apparent diffusion-controlled dissolution was found, while in the Al–Ag alloy some evidence for mixed control was reported with the solute content in the matrix at the interface, C_i, being less than the equilibrium value, C_α. For such systems, where a ledge mechanism would be expected, it should be remembered that for *dissolution* of isolated plates the incoherent rims of the plates are good sources of ledges (MARTIN and DOHERTY [1976]).

A detailed analysis of precipitate dissolution for spherical precipitates and for diffusion away from the broad faces of plate-like precipitates was published by AARON and KOTLER [1971]. Their review provides a very detailed summary of the theory and need not be reproduced here. Only some of their major conclusions will be described. Aaron and Kotler compared the processes of growth and dissolution and showed that

latter is not the simple opposite of the former, particularly for spherical precipitates. The distinction arises since the rate of dissolution, for a diffusion-controlled reaction, is initially infinitely fast and then falls as the concentration gradient in front of the interface becomes shallower. However as the process continues, the radius of the sphere shrinks and the volume releasing the solute falls so the dissolution rate rises again. The effect is due to the spherical symmetry of the matrix absorbing the solute. Plate-like dissolution by *thinning* however does not show this acceleration and the rate of thinning continues to decline (fig. 34). Aaron and Kotler also point out that the details of the diffusional analysis are more difficult for dissolution than for growth since, in the dissolution case, the source of solute, the interface, is retreating away from the surrounding matrix so that the solute distribution close to the interface undergoes a more complex behaviour during dissolution than during growth (fig. 35). This problem becomes very severe for the case of dissolution of a plate-like or needle-like precipitate growing back from the tip. In this case the tip will move into a region of matrix already enriched with solute due to the thinning process. No analytical solution of the problem appears to be available.

In the analysis provided by AARON and KOTLER [1971], it was shown that, although capillarity effects are expected to increase the rate of dissolution of spherical particles owing to the rise of solubility at small radii, these modifications appear to be small for the 1 μm-sized and larger particles of interest in dissolution reactions above the equilibrium solvus. Although this was not considered by Aaron and Kotler, the acceleration due to capillarity will be of much greater significance for the dissolution of the very much smaller precipitates of concern in precipitation *reversion*. Reversion is the dissolution of very fine metastable precipitates when the ageing temperature is raised during a precipitation-hardening reaction. SARGENT and PURDY [1974] pointed out the possibility that a fine dispersion of precipitates may redissolve completely even when the raised temperature remains within the two-phase field, so long as the radii of the precipitates are less than the critical radius for nucleation of the alloy at the higher temperature. The critical radius is the value of r in eq. (58) at which $C_\alpha(r)$ is equal to the alloy content, C_o. A clear example of this effect is seen in Al–Mn alloys in which fine precipitates of Al_6Mn formed at lower temperature are seen dissolve up on heating, *even though the alloy is still supersaturated with the solute*. The transformation is readily observed using electrical conductivity on samples quenched during processing. On

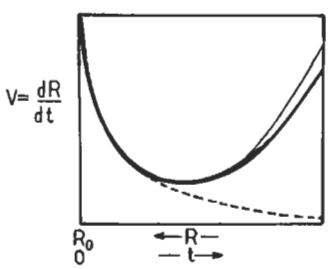


Fig. 34. Dissolution of an isolated sphere (heavy solid line) and a plate (dashed line) by diffusion-controlled dissolution. The influence of the Gibbs–Thomson effect on the dissolution rate of small spheres is shown by the thin solid line. (After AARON and KOTLER [1971].)

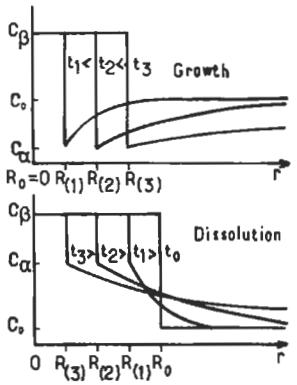


Fig. 35. Distinction between growth and dissolution of a solute-rich β precipitate: the composition is C_α at the interface and C_0 in the matrix a long way from the precipitate. (After AARON and KOTLER [1971].)

heating, the conductivity rises as Mn is initially precipitated, however above about 800 K the conductivity starts to fall as heating continues — indicating net Mn dissolution. On halting the heating, and holding the sample at a constant temperature, the dissolution quickly ceases and precipitation resumes — confirming that *the previous dissolution was taking place in a supersaturated solid solution.* The low solute content of the precipitate, $C_\beta = 1/7$, greatly promotes the increase of Gibbs–Thomson solubility in this system — see eq. (58). This type of precipitate dissolution is exactly the effect predicted by Sargent and Purdy. An alternative type of reversion also occurs. Here the temperature is required to exceed the *solvus of the metastable precipitate* even though the temperature is below the solvus temperature of more stable precipitates. Reversion of copper-rich GP zones, formed by room temperature "natural" ageing of Al–Cu alloys, on heating to a temperature, 450 K, at which the θ' phase subsequent precipitates during artificial age hardening is a well known example of this second type of reversion. The sample initially softens, as the zones dissolve, before rehardening with subsequent precipitation of θ'. Reversion of Al–Li was used by VASUDEVAN and DOHERTY [1987].

2.4. Competitive growth

There are two subsequent stages to the free growth of isolated precipitates described in § 2.2. The first of these occurs when the diffusion fields around adjacent precipitates start to overlap, but while the matrix around the precipitate remains supersaturated with solute. This is briefly discussed in § 2.4.1. Its essential feature is that, for almost all of the particles, the composition in the matrix exceeds the Gibbs–Thomson solubility, $C_\alpha(r)$ given by (eq. 58). As a result almost all of the particles survive so that the density of precipitates (the number per unit volume — N_v) either continues to increase if nucleation is occurring or remains essentially constant if nucleation is essentially complete. The second stage of competitive growth — usually called *Ostwald ripening* — occurs when a significant fraction of the nucleated precipitates have a Gibbs–Thomson solubility,

$C_\alpha(r)$ greater than the average matrix composition, so that they then dissolve and the microstructure shows a decrease in the precipitate density, N_v. Ostwald ripening is discussed in § 2.4.2.

2.4.1. Growth from a supersaturated matrix after "soft" impingement

The situation when the diffusion fields around adjacent precipitates start to overlap is a process usually described as *soft impingement*. With soft impingement, the concentration away from a solute-rich precipitate falls below the initial alloy composition, C_o, and the effective supersaturation, Ω, falls with time causing the rate of growth given by (eq. 43) to fall. For growth of solute *depleted* precipitates, fig. 20d, the solute in the matrix rises, but the value of the supersaturation falls. After soft impingement, not only will the growth rate fall but also the basic r^2/t relationship will change; giving a larger time exponent than the value of 2 derived for independent precipitate growth. Care is needed in experimental studies of precipitation to distinguish a fall of the exponent caused by soft impingement, from an apparent fall in the exponent caused by measuring the *mean* radius, as $\bar{r}^2/t$. With true soft impingement, there is a failure in $r_{max}{}^2/t$ where the size measured is that of the largest precipitates — ones that nucleated at the start of the reaction. Continued nucleation of precipitates, even if the younger precipitates still grow as $r^2/(t-t_n)$, will give a failure in the expected kinetics if these are determined from the *average* precipitate size.

It is probably useful to think of the regime of growth, as distinct from the regime of nucleation, as being characterized by *an almost constant density* of precipitates (the number per unit volume, N_v). During the nucleation regime the density of precipitates *rises* and during competive coarsening § 2.4.2, the density *falls*.

As described in various reviews of overall transformation kinetics, for example by TURNBULL [1956], FINE [1964] and CHRISTIAN [1970, 1975], the rate of the precipitation reaction will be given by eq. (22) during the growth period that includes soft impingement but not during competive coarsening. The references cited here include a general discussion of what is now called the Johnson–Mehl, Avrami, Kolmogorov (JMAK) equation, eq. (22). In the general case that includes precipitation where the volume fraction of the product does not tend to 1, the equation gives the amount of new phase formed but normalized by the amount of the new phase present at equilibrium. This ensures that the predicted reaction parameter, $f(x)$, tends to unity as the reaction is completed. The values of the exponent n are related, see table 3, to the particular nucleation and growth conditions. Table 3 was originally published by CHRISTIAN [1970]. The time constant τ, is determined from the nucleation and growth rates and the form of the equation, $1-\exp(-t/\tau)^n$, comes from competition between the growth centers, either by "soft" impingement of diffusion fields or by "hard" impingement between growing grains. A full discussion of this topic has not been provided here, firstly since the earlier reviews are rather complete and, secondly, because (with a few exceptions) use of the overall reaction kinetics seems to have played little role in understanding the mechanisms of transformations. The failure arises since similar values of n can he obtained for different mechanisms. In the opinion of the present author, the direct determination of the separate nucleation and growth rates is of much greater value in understanding the

Table 3
Value of n in kinetic law $f = 1 - \exp(-kt^n)$.

(a) Polymorphous changes, discontinuous precipitation, eutectoidal reactions, interface controlled growth, etc.

Conditions	n
Increasing nucleation rate	> 4
Constant nucleation rate	4
Decreasing nucleation rate	3–4
Zero nucleation rate (saturation of point sites)	3
Grain-edge nucleation after saturation	2
Grain-boundary nucleation after saturation	1

(b) Diffusion-controlled growth (early stages of reaction only)

Conditions	n
All shapes growing from small dimensions, increasing nucleation rate	$> 2^{1/2}$
All shapes growing from small dimensions, constant nucleation rate	$2^{1/2}$
All shapes growing from small dimensions, decreasing nucleation rate	$1^{1/2}$–$2^{1/2}$
All shapes growing from small dimensions, zero nucleation rate	$1^{1/2}$
Growth of particles of appreciable initial volume	1–$1^{1/2}$
Needles and plates of finite long dimensions, small in comparison with their separations	1
Thickening of long cylinders (needles), e.g. after complete edge impingement	1
Thickening of very large plates, e.g. after complete edge impingement	1/2
Segregation to dislocations (very early stage only)	2/3

reaction mechanisms than is use of eq. (22). FINE [1964] points out that the JMAK equation is, in fact, only an approximation to the expected kinetics for precipitation with soft impingement. However, the equation is of value in providing a simple picture of the kinetics of the overall process and in confirming that individual nucleation and growth rates have been correctly measured. A nice illustration of the use of the equation is its application in biology to the two-dimensional spread of colonies of lichen on rocks or of oysters in freshly established oyster beds. The equation is also commonly used in describing the rates of recrystallization in deformed materials (chapter 28, § 3.2) and also in studies of isothermal crystallization of glasses particularly of metallic glasses (chapter 19, § 4.3). All these cases are for "hard" impingement between growing colonies or grains for which the assumptions of the JMAK equation are valid. ROLLETT *et al.* [1989] provide a recent review of the JMAK equation and the conditions needed for its applicability. They showed that one of the critical features required for the JMAK equation to be valid, is the spatial uniformity *of the driving force for the reaction*. When the driving force varies, for example between different grains in a deformed material, the impingement correction that yields the JMAK equation is no longer valid. This is, however, not likely to be a problem in structural changes driven by chemical free energy changes which should not vary from grain to grain.

2.4.2. Competitive coarsening: Ostwald ripening

Following the soft impingement of diffusion fields, the precipitates continue to grow, draining the remaining solute from the matrix and approaching the equilibrium volume fraction of the new phase. The approach to the equilibrium volume fraction does not mark the end of the reaction, however, since the microstructure with a distribution of fine precipitates has a high area of interface, and this raises the free energy of the system. The process of reduction of this interfacial energy involves the growth of larger precipitates at the expense of smaller ones. In a similar way, after primary recrystallization of deformed metals the resulting grain structure usually has a significant amount of grain-boundary area that drives the process of grain growth.

The process of competitive precipitate growth in a near-equilibrium matrix composition is commonly described as *Ostwald ripening* after the discoverer of the same process for salt crystals precipitating from aqueous solutions (OSTWALD [1900]). The mechanism of this process is provided by the Gibbs–Thomson equation, eq. (58). This equation shows that, with a range of precipitates sizes, the solubility will be least with the large precipitates and highest with the small precipitates. This difference in solubility produces a solute gradient from small precipitates to large ones and diffusion of solute between the precipitates then causes the smallest precipitates to shrink and disappear. The solute from the shrinking precipitates feeds the larger ones and so, for both reasons the mean precipitate size rises. The rate of growth is described by fig. 30 but with the critical radius of nucleation now replaced by the mean radius, $\bar{r}$; $\bar{r} = r^*$.

The topic has been reviewed by for example MARTIN and DOHERTY [1976], by MARTIN *et al.* [1995] and the topic is also discussed in chapter 9, § 4.2.2, so full details of the process will not be presented here. For diffusion-controlled growth the classic (LSW) theory of the process due to WAGNER [1961] and LIFSHITZ and SLYOZOV [1961], predicts that, even starting from a narrow size distribution, the ripening process will yield a steady-state distribution of particle sizes. After achieving this steady state size distribution, the mean precipitate radius will coarsen from an initial value, $\bar{r}_0$, to $\bar{r}_t$ with time, t, as:

$$\bar{r}_t^3 = \bar{r}_0^3 + \frac{8}{9}\frac{\sigma D C_\alpha V_m t}{RT}. \tag{66}$$

It should be noted that this commonly quoted equation was derived for a particular set of assumptions. These include that the solubility in the matrix, C_α, is given as atomic fraction, that both matrix and precipitate are terminal and dilute solid solutions, so $C_\alpha \to 0$ and $C_\beta \to 1$, and that the matrix has a constant activity coefficient, γ, or equivalently has a second differential of the free energy with composition of the matrix, f''_α, which is that of an ideal solution. In addition the volume fraction of the precipitates is assumed to be vanishingly small. The general expression, see for example CALDERON *et al.* [1994] or MARTIN *et al.* [1996], is:

$$\bar{r}_t^3 = \bar{r}_t^3 + \frac{8}{9}\frac{\sigma D C_\alpha (1 - C_\alpha) V_m t}{F_\alpha'' (C_\beta - C_\alpha)^2}$$

$$\text{or} \quad \bar{r}_t^3 = \bar{r}_0^3 + \frac{8}{9}\frac{\sigma D C_\alpha (1 - C_\alpha) V_m t}{RT\left(1 + \dfrac{d\ln\gamma}{d\ln C_\alpha}\right)(C_\beta - C_\alpha)^2} \tag{66a}$$

The changes in eq. (66a) can be very significant. This influence of having a value of the precipitate atomic fraction, C_β, less than 1 is particularly important. If the precipitate is an intermetallic phase A_nB in A then this leads to a predicted increase of the coarsening rate by approximately $(n+1)^2$. For Ni_3Al in Ni this is an increase of 16 and for Al_6Mn in Al the increase is almost 50 times!

The general result of eq. 66 has been extensively tested both qualitatively and quantitatively many times and, in general, the model appears to be very successful, as shown for example by ARDELL [1987 and 1990], XIAO and HAASEN [1991] and MARTIN *et al.* [1995]. However, there remain a number of problems. The main problem is that although almost all tests on spherical and near-spherical precipitates give the expected $\bar{r}_t^3/t$ relationship, the observed *rates* of coarsening, the gradients of the $\bar{r}_t^3/t$ plots, are usually *larger* than expected. In addition, there is a corresponding somewhat *wider* size distribution than predicted by the LSW analysis. A further point is that the simple diffusion theory [eq. (60) and fig. 30] predicts that the fastest growth rate will be for particles *twice* the size of the average, yet the LSW-predicted spread of particle sizes for diffusion-limited growth cuts off at a maximum particle size of only *1.5r*. It is not immediately obvious why particles are theoretically limited to this size, since fig. 30 suggests that after extended growth the maximum precipitate size is likely to be *2r*, GREENWOOD [1956]. Experimental size distributions, even after very long periods of coarsening, do indeed show precipitates up to *2r* in size, see for example fig. 36 from BOWER and WHITEMAN [1969]. That is, the experimental results almost appear to fit better with the simple analysis than with the rigorous LSW analysis! It seems likely that

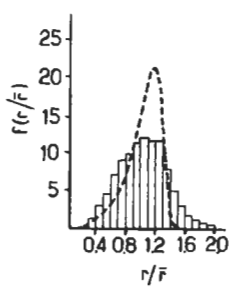


Fig. 36. Experimental results for the distribution of precipitate reduced radii ($r/\bar{r}$) against reduced radius after extensive diffusion-controlled coarsening of Fe_3Si in Fe at 700°C, aged 312 h. The dotted line shows the expected distribution in the limit of a low volume fraction ($f \to 0$). (After BOWMAN and WHITEMAN [1969].)

the origin of the effect arises from the fact during the initial precipitation reaction the particles acquire a wider size distribution than the predicted steady-state size distribution of coarsening. Not all precipitates nucleate at the same time, nor since the sites of nucleation are not ordered in space do the precipitates have equal volumes of supersaturated matrix to draw the solute from. So the initial size distribution will usually be rather wide.

The rates of change of initial particle size distributions with coarsening have been examined (i) by computer simulations, for example by VOORHEES and GLICKSMAN [1984c] and BROWN [1989, 1990a and b, 1992a and b], (ii) analysed theoretically by HILLERT *et al.* [1989 and 1992] and (iii) experimentally studied by FANG and PATTERSON [1993]. It seems established that, in the final limit of very extensive coarsening, the LSW size distribution is the only distribution that can, mathematically, maintain a steady state form (see for example HILLERT *et al.* [1989], MULLINS and VINALS [1989] and HOYT [1990]). BROWN [1989] found, by numerical simulation of coarsening, that particles initially given a wide size distribution maintained a significant number of particles in the size range above $1.5\bar{r}$ for extended times. VOORHEES and GLICKSMAN [1984c] also noted a similar effect. FANG and PATTERSON [1993] measured the effect of different initial particle size distributions on coarsening. They used tungsten powders in liquid nickel–iron matrix, albeit at a rather high volume fraction of solid particles ($f_s = 0.6$) see below. They found, as might be expected, that the rate of coarsening was controlled by the size distribution (fig. 37). Narrow distributions of particles showed *slow* coarsening and wide distributions coarsened *more quickly*. The size distributions changed rapidly at the earliest times, as shown by changes in the natural logarithm of the standard deviation (fig. 38). During the time interval when the size distributions were changing *rapidly*, the coarsening equation did not show the expected cube law for diffusion limited coarsening (fig. 37b). The kinetics however quickly achieved the expected cube-law after rather short periods of non steady-state coarsening. However, the distributions, particularly the wide distribution, did not reached a really steady-state form even after extensive coarsening. During the period when the distributions were changing only *slowly*, however, the cube law fitted the data to a rather high precision. This result is important in showing that a good experimental fit to the cube law does not guarantee that the kinetic coefficient will have the value expected for true steady-state coarsening.

Another problem with the LSW analysis is that it was developed for *point* particles, that is, for volume fractions of second phase that are close to zero, so that, in the LSW analysis, individual particles are only aware of the *average* solute content, $\bar{C} = C_\alpha(\bar{r})$, and they are unaware of the local concentrations introduced by neighbouring particles. ARDELL [1972] introduced an initial approximate, but physically reasonable, correction for the effect of finite volume fractions and his initial analysis showed that when the volume fraction of second phase was 0.25, the rate of coarsening implied by eq. (66) was expected to increase *tenfold*. Subsequent analyses such as that BRAILSFORD and WYNBLATT [1979] suggest that the acceleration of coarsening with volume fraction that though still significant is much smaller than predicted by Ardell. The computer simulations of VOORHEES and GLICKSMAN [1984c] suggest only a doubling of the coarsening rate at a volume fraction of 0.25 and a fourfold increase at a volume fraction of 0.5. For

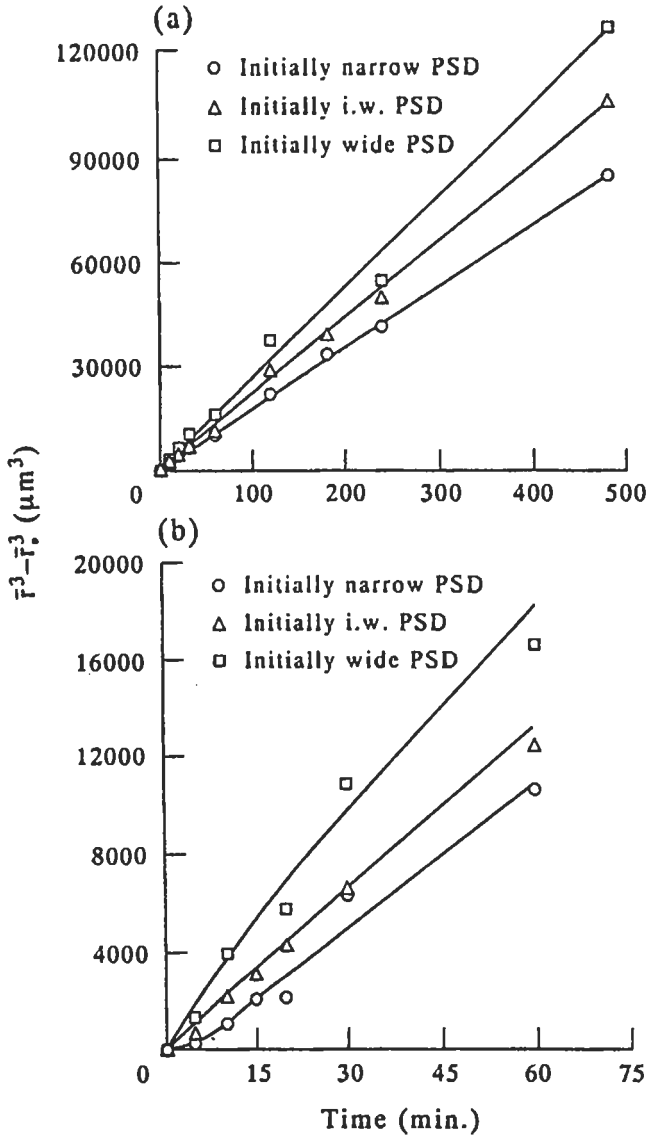


Fig. 37. Plots of $\bar{r}^3-\bar{r}_o^3$ against coarsening time, experimental results for W-14Ni-6Fe showing the effect of the initial width of the particle size distribution on the rate of coarsening. (a) The entire coarsening period (b) the initial coarsening period. After FANG and PATTERSON [1993].

several sytems of solid particles of transition metals in liquid copper as reviewed by ANNAVARAPU and DOHERTY [1995] and for two solid/solid systems, Al_3Li in Al and Co in Cu reviewed by ARDELL [1987, 1990] showed clear evidence of a significant increase in coarsening rate as the volume fraction increased. For the solid–solid cases, corres-

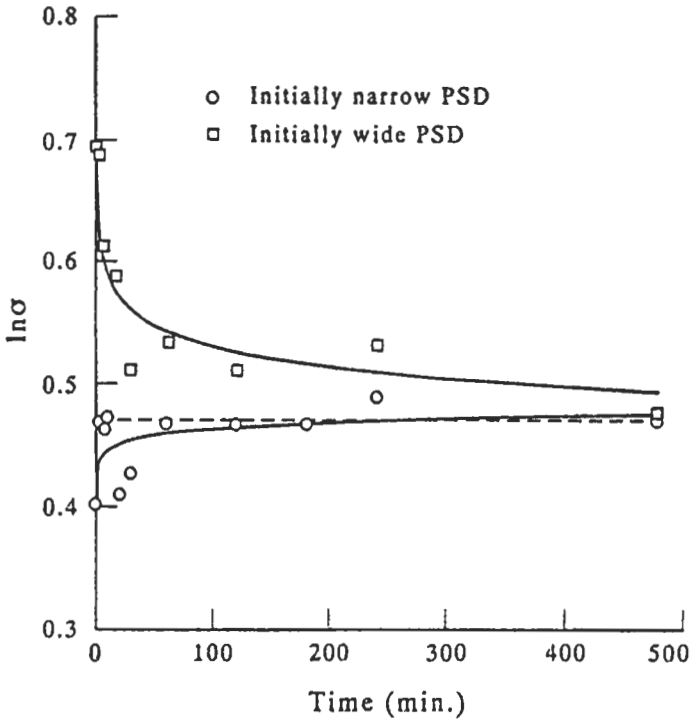


Fig. 38. Experimentally measured change of the standard deviation of the log normal distribution, ln σ, during coarsening of W particles in the W-14Ni-6Fe alloy. After FANG and PATTERSON [1993].

ponding increases in the apparently steady state width of particle size distribution was also found. However Ardell's reviews clearly found that although there were considerable differences in the rate constants found by different investigators for the system of Ni_3Al in Ni there was almost no evidence, in any of the studies, of any acceleration of the coarsening rate with increase of the volume fraction of Ni_3Al in Ni. The reasons for both the scatter in the data and for the lack of any variation in coarsening rate with volume fraction are not at present clear. The recent, very detailed study by XIAO and HAASEN [1991] on coarsening kinetics in a single alloy of Ni–12at% Al is worthy of note. They obtained, from careful measurements of precipitate sizes, precipitate volume fraction and the precipitate densities, data that, as discussed in § 2.1.5, matched classical nucleation theory and also fitted the classical LSW coarsening theory (as modified by ARDELL [1990]) very closely both as regards the coarsening kinetics, the growth of volume fraction precipitated and the decrease in precipitate density. These agreements, described below, were obtained *despite a significant volume fraction* of the precipitates in the system (about 15%) and of course the LSW theory was developed for negligable volume fractions. So for the Ni_3Al/Ni system the volume fraction effect found in other systems is not apparently occurring. This is a major problem in current phase transformation studies.

References: p. 1497.

XIAO and HAASEN [1991] made two independent determinations of the interfacial energy, σ, and the operating diffusion coefficient D. σ was found to be 0.014 J/m^{-2} while D was found to be significantly larger (4 times) than expected from diffusion studied by more conventional means, the values of σ and D determined from the various coarsening tests matched the values obtained from nucleation data astonishingly well — § 2.1.5.

The expected variation in volume fraction, $f(t)$, with time is given by eq. (66b)

$$f_\infty - f(t) = k_5\left(t - t_1\right)^{-1/3} \tag{66b}$$

$$\text{with} \quad k_5 = \left(C_\beta - C_0\right)C_\alpha \frac{\left(\frac{9}{4}\frac{V_m\sigma}{R^2T^2C_\alpha D}\right)^{\frac{1}{3}}}{\left(C_\beta - C_\alpha\right)^2}$$

t_1 is the time (45 h in this experiment) at which "late stage" coarsening starts. The condition for late-stage coarsening is that $\bar{C} \rightarrow C_\alpha(\bar{r})$. That is, when the only remaining supersaturation arises from the increase in solubility due to the interfacial energy. $(t-t_1)$ was the time used in eq. (66a) to obtain the slope of the $\bar{r}^3$/time plot.

The expected variation for the number density N_v as a function of time, again $(t-t')$, is given by:

$$N_v(t) = k_3(t - t_1)^{-1} - k_6(t - t_1)^{-4/3} \tag{66c}$$

The rate constants, k_3 and k_6, are given by:

$$k_3 = \frac{27\left(C_0 - C_\alpha\right)RT}{128\pi\left(C_\beta - C_\alpha\right)C_\alpha V_m\sigma D}$$

$$k_4 = \frac{27RT\left(9RT/4C_\alpha V_m\sigma^4 D\right)^{1/3}}{128\pi\left(C_\beta - C_\alpha\right)C_\alpha V_m\sigma D}$$

The first term in eq. (66c) is that from the original LSW model and did not include the change in number density arising from the increase in volume fraction of precipitates arising from the decrease in solubility of the precipitates with time as the particles coarsen — that is from the Gibbs–Thomson effect. The second term in eq. (66c) is the correction, derived by ARDELL [1987, 1990], that includes this effect.

The start of late-stage coarsening, $t_1 = 45$ h, was close to the time (40 h) at which the maximum value of $N_v(t)$ was found, that is when the net rate of nucleation, that is the rate of nucleation less the rate of particle dissolution, reached zero. At this time, the diffusion distance $(Dt)^{1/2}$ was about 15 nm, three times larger than the mean spacing, $\lambda \approx (N_v(t))^{-1/3}$, between the precipitates. The very high density of precipitates achieved in this system with a very low interfacial energy suggests that precipitates will interact diffusionally from early in the process — nucleation can only continue in regions with a locally lower density of precipitates while coarsening is occurring in regions with a

locally higher density of precipitates. Two results appear to follow directly from this result. The first result is that late-stage coarsening will be expected to follow soon after the maximum density of precipitates was seen. The second result is that only a very limited period of free diffusional growth with a near constant density of precipitates should be expected. Both of these results were reported in the study. The number density showed a continuous change with the density increasing to a maximum and then falling again with no significant period of near constant density. It was reported that at no time was the expected relationship for free growth of a constant density of precipitates, $\bar{r}^2$ increasing as t, found. As noted, in § 2.4.1, such kinetics would not be expected if nucleation was continuing. For a high rate of homogeneous nucleation such as was found in the system even $r_{max}{}^2$ might not be expected to increase linearly with t. With a very high rate of nucleation from the alloy composition, isolated precipitates surrounded by matrix at the alloy composition will not be found — the matrix under these circumstances will be nucleating more precipitates — thereby depleting the solute. Xiao and Haasen reported that during the period of nucleation there was only a small increase in the mean precipitate radius — see similar earlier results shown in fig. 43 in the next section. It was also reported that after extensive coarsening from 45 to 256 h, during which time $\bar{r}^3$ increased tenfold, the size distribution did approached the LSW form with the maximum radius of 1.4 $\bar{r}$. This match with the expected LSW distribution was found despite the observations being made in a system with a significant, equilibrium, volume fraction of precipitates $f = 0.12$.

For a full discussion of the various other theoretical modifications to the initial LSW theory, to take account of both interface immobility, and alternative diffusion paths along grain-boundaries dislocations, etc., the reviews by MARTIN and DOHERTY [1976] or MARTIN *et al.* [1995] can be consulted. Developments in both the theory and the experiments of coarsening are also discussed in chapter 9. One recent development due to YANG *et al.* [1992], HIGGINS *et al.* [1992] and ANNAVARAPU and DOHERTY [1995] should however be noted. This is coarsening by coalescence of particles at grain boundaries caused by *particle drag*. Particles at grain-boundary triple points and at grain corners can remove grain-boundary curvature and thus halt grain growth but the particles then experience a net unbalanced force. For particles on small grains this leads to the migration towards each other of the particles, causing both the loss of the smallest grain and the coalescence of the set of particles to a single larger particle. The rate determining step in this process is solute diffusion within the particles rather than between then so for small volume fractions of particles this leads to a rapid acceleration of the coarsening rate. YANG *et al.* [1992] showed this model gave good agreement with their experiments in a solid solid system (Ni–Ag) and ANNAVARAPU and DOHERTY [1995] were able to use a modification of their model to account successfully for the rapid acceleration of coarsening in high purity solid–liquid mixtures as the volume fraction of solid approached 1.

2.4.3. During initial nucleation and growth

It was initially pointed out by KAMPMANN and KAHLWEIT [1967, 1970] that competitive dissolution of small precipitates will not be restricted to periods *after* precipitation was complete. This is because during all stages of precipitation the average

solute content of the matrix will be falling steadily owing to nucleation and growth of precipitates — for example as measured by the electrical resistivity studies of cobalt precipitation in Cu–Co alloys (SERVI and TURNBULL [1966]). The fall in solute content causes the critical radius of nucleation, r^*, to rise, and any precipitate whose radius falls below this growing size will redissolve. The process is shown schematically in fig. 39. The alloy is steadily cooled below the solubility limit, becoming supersaturated at $t = t_s$. The supersaturation, S, here defined as as $\bar{C}/C_\alpha$, first rises with time (because C_α is falling as the temperature falls) so r^* initially falls with time. The first nucleus forms at t', when S is still rising and r^* falling; When, at t'', the last nucleus forms S is then falling and consequently, $r^*(t)$ is rising. The smallest precipitates, whose growth rate is less than dr^*/dt, then have too high a solubility and will dissolve up even though precipitation, measured by a fall in S, is still continuing. The critical result is not that coarsening, as measured by a fall in precipitate density, is occurring while the supersaturation, $S = \bar{C}/C_\alpha$, is falling. The significant feature is that there can be a fall of the precipitate density even though $\bar{C} > C_\alpha(\bar{r})$. There is always a finite Gibbs–Thomson supersaturation, $S_{GT} = C_\alpha(\bar{r})/C_\alpha$, for precipitates of a finite mean radius. With $\bar{C} = C_\alpha(\bar{r})$ there is still precipitation occurring given by eq. (66b) above.

Numerical simulations of combined precipitation and coarsening were published by KAMPMANN and KAHLWEIT [1970], assuming diffusion-controlled growth and dissolution, using homogeneous nucleation theory and assuming that the supersaturation builds up steadily with time. Figure 40 shows the result of one set of calculations with fig. 40a showing the change of supersaturation, S; curve 1 is the change of S that would occur in the absence of precipitation and curve 2 the actual change of supersaturation that is determined by the nucleation and growth rates. The total number of precipitates larger than the critical size is seen in fig. 40b as the number of precipitates, Z/N_t, normalized by dividing by the final solute atomic concentration. The number of precipitates rises during the precipitation reaction to a maximum value and then starts to

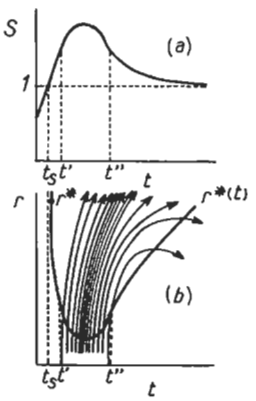


Fig. 39. Competitive coarsening during precipitation. An alloy is quenched becoming supersaturated at $t = t_s$. The first nucleation occurs at t' and the last t''. As the supersaturation, $S = \bar{C}/C_\alpha$, falls, $r^*(t)$ rises and small precipitates redissolve. (After KAMPMANN and KAHLWEIT [1970].)

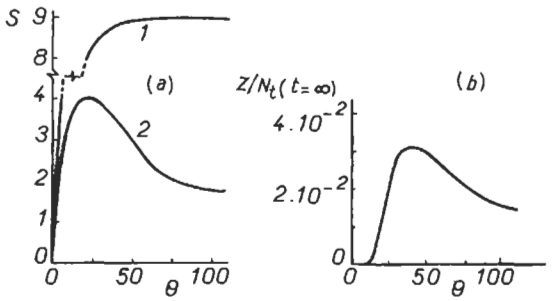


Fig. 40. Numerical solution to competitive coarsening: (a) the variation in supersaturation S with time θ, curve 1 with no precipitation, curve 2 with precipitation; (b) the variation in the density of the precipitates Z, scaled with the equilibrium solute content N_t ($t=\infty$). (After KAMPMANN and KAHLWEIT [1970].)

fall while the supersaturation is still large, approximately 4. Figure 41 shows the computed change of i^*, the critical precipitate size in terms of the number of atoms of solute, during coarsening; this is compared to the "steady-state" loss of precipitates predicted by an analysis equivalent to the LSW analysis described in § 2.4.1, shown as the dashed line. KAMPMANN and KAHLWEIT [1970] suggested that the discrepancy is probably within the experimental error of the numerical analysis used. The dimensionless time parameter, θ, scales with the diffusion coefficient, so that with the diffusion coefficient expected for liquids $D=10^{-9}$ m^2/s, $\theta=10$ corresponds to only 10^{-6} s, but with a substitutional diffusion coefficient of $D=10^{-16}$ m^2/s this would be 100 s. It is clear from this model, therefore, that the number of growing precipitates found in the later stages of precipitation should be considerably smaller than the original density of growing nuclei, for example in the studies of the nucleation kinetics in the experiments

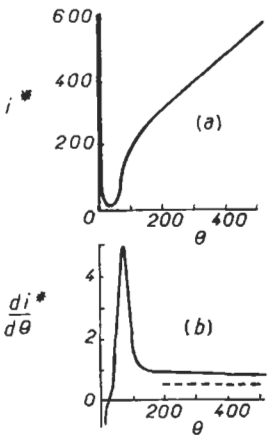


Fig. 41. Competitive coarsening during precipitation: (a) the variation in the critical size i^* with coarsening time, where i^* is the number of atoms in the critical cluster; (b) the rate of change of the critical size with time. (After KAMPMANN and KAHLWEIT [1970].)

of Servi and Turnbull discussed in § 2.1.5. LANGER and SCHWARTZ [1980] analytically evaluated the effect of interfacial energy driven coarsening on precipitate densities. The results reported show similar trends to those of the numerical models described here but no detailed comparison between the evaluations was made.

There are several experimental studies that have been made to test the idea of competive precipitate dissolution during precipitation in the Ni–Ni_3Al system. These experiments used conventional transmission electron microscopy (TEM), KIRKWOOD [1970] and HIRATA and KIRKWOOD [1977], field ion microscopy/atom probe techniques, WENDT [1981] and recently high resolution TEM, XIAO and HAASEN [1991]. The last study was discussed in § 2.4.1 above. In all the studies a fall of at least an order on magnitude in precipitate density was found while the precipitation was continuing. The early results of WENDT [1981] are shown in figs. 42 and 43. The supersaturation, $\Delta C = \mathfrak{C} - C_\alpha$, fell with time and approached the initially expected $t^{-1/3}$ kinetics of eq. 66b after about 50 minutes. The initial number density fell approximately linearly after this time. However this is not the correct kinetics which should be those of eq. (66c). Figure 44 from the recent study of XIAO and HAASEN [1991], shows that the product of $(t-t_1)$ and $N_v(t)$ was not constant but fell linearly with $(t-t_1)^{-1/3}$, exactly as predicted by eq. (66d), obtained from eq. (66c),

$$N_v(t)(t-t_1) = k_3 - k_6(t-t_1)^{-1/3}. \tag{66d}$$

The recent studies of XIAO and HAASEN [1991], described above, show very clearly how both the nucleation and coarsening processes can strongly overlap — especially with very high densities of precipitates. It seems likely that this will be a general result for systems showing essentially homogeneous nucleation. For the reasons discussed in § 2.4.1, any system with a high rate of homogeneous nucleation will very quickly achieve soft impingement — that is when the spacing between precipitate, approximately $(N_v)^{-1/3}$, becomes smaller than the diffusion distance, $(Dt)^{1/2}$. So free growth with $(r_{max})^2$ increasing with time is unlikely to be seen. The only conditions where free growth with $(r_{max})^2$ increasing with time should be expected is for systems showing heterogeneous nucleation so that not only are the precipitates widely spaced, allowing $(N_v)^{-1/3} \gg (Dt)^{1/2}$, but on a microstructural level the full supersaturation is expected to persist in regions of the matrix *between the effective heterogeneous nucleation sites*. Homogeneous nucleation will occur there at negligable rates at supersaturations where heterogeneous nucleation and growth dominates. KAMPMANN and WAGNER [1984] and HAASEN and WAGNER [1992] show this result rather clearly through the Kampmann and Wagner "numerical" model of combined nucleation, growth and coarsening at high precipitate densities. In support of this idea is the observation that experimental studies showing $(a_{max})^2$ increasing with time appear to be only found in studies for heterogeneously nucleated precipitates, see, for example, the results discussed in § 2.2.3 for *plate-like* precipitates, where the precipitate and matrix have different structures. Homogeneously nucleated precipitates seem limited to systems where the interfacial energy is low and isotropic giving very high densities so preventing free growth from occurring at the undercoolings usually studied. If the density of precipitates is greatly reduced then the free growth of spherical precipitates with $(r_{max})^2$ increasing with time is still not seen since under these circumstances the spheres

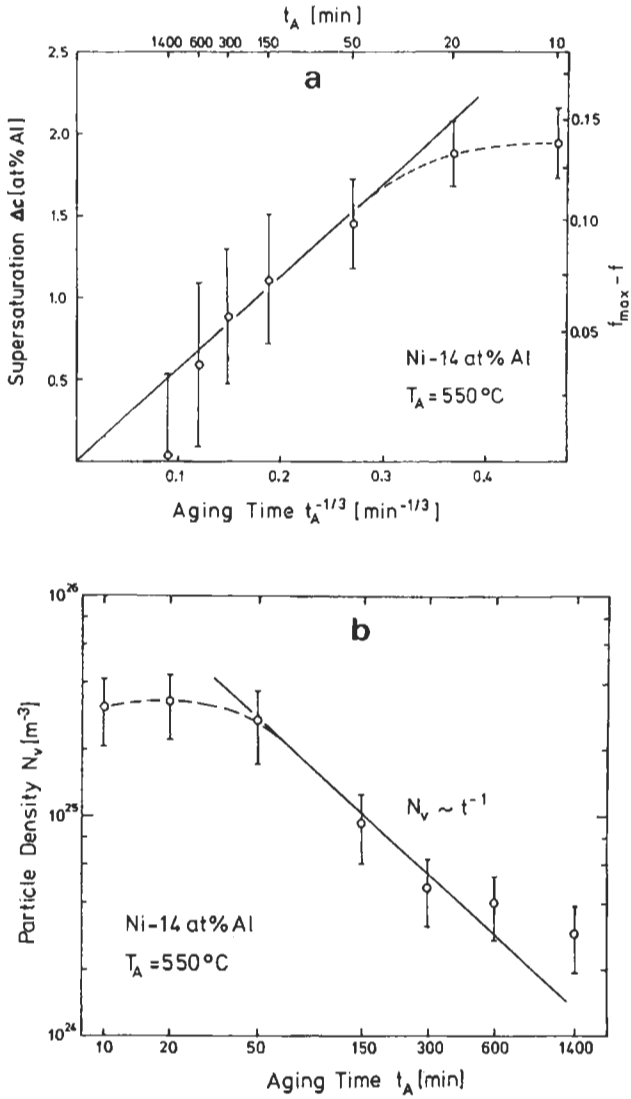


Fig. 42. Precipitation in a Ni–14 at% Al alloy, aged at 550°C: (a) super-saturation, Δc, versus ageing time, t_A (from right to left); f is the precipitated volume fraction; (b) particle density, N_v, versus ageing time, compared with the theoretical slope (solid line). (From WENDT [1981].)

become unstable and grow as dendrites with a constant tip growth velocity as discussed in § 2.2.6. Examples of homogeneously nucleated dendrites are reported in the Ni_3Al–Ni system (RICKS *et al.* [1993]) and in the similar Al_3Li–Al system (BAUMANN and WILLIAMS [1985]). An example of a system showing heterogeneously nucleated dendrites is the γ-brass/β-brass system (MALCOLM and PURDY [1967] and DOHERTY [1982]).

References: p. 1497.

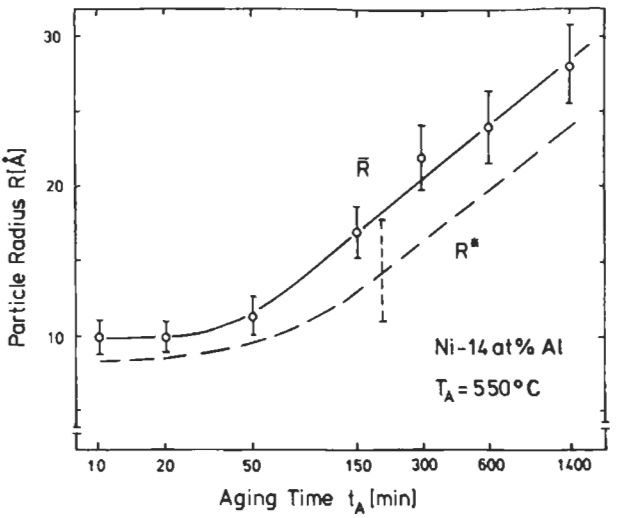


Fig. 43. Measured mean particle radius, $\bar{r}$, and calculated radius of critical nuclei, R^*, versus ageing time for the alloy of fig. 37 (from WENDT [1981]).

2.4.4. Coarsening of Widmanstätten precipitates

One of the early studies of precipitate coarsening was of the growth of the needle-like precipitates of fcc copper in bcc iron (SPEICH and ORIANI [1965]). Despite the anisotropic shape, arising from the good atonic fit along the close-packed directions <110> fcc and <111> bcc, the precipitates showed apparent diffusion limited coarsening with a constant precipitate aspect ratio of about 4. The authors used a modified version of the HAM [1958, 1959] analysis to model the diffusion controlled growth. They obtained

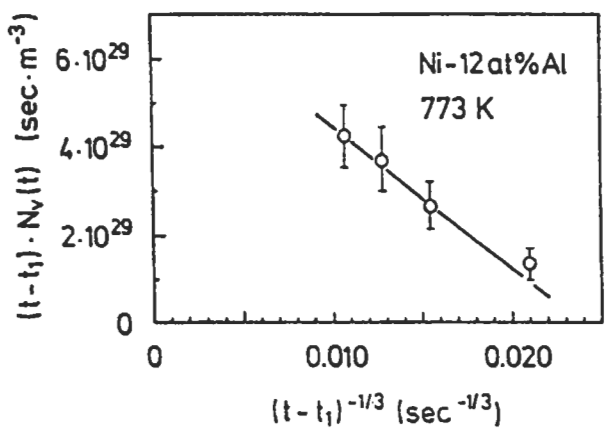


Fig. 44. Experimental values of the product of $(t-t_1)N_v(t)$ as a function of $(t-t_1)^{-1/3}$ for coarsening of Ni-12 at% Al at 773K, eqs. (66c and d). After XIAO and HAASEN [1991].

reasonable agreement with their experimental results. FERRANTE and DOHERTY [1979], as part of an investigation of the role of ledged interfaces in the growth of plate-like precipitates in Al–Ag, studied the coarsening of the precipitates after the completion of the initial precipitation. The plate-like precipitates in this system had *much larger* aspect ratios than expected for the equilibrium shapes: $A > 100$ while $A_{eq} \approx 3$ (AARONSON *et al.* [1968]). Ferrante and Doherty showed that with $A \gg A_{eq}$ the modified Gibbs–Thomson equations for the solubility at the incoherent rim, C_r, and that at the semicoherent facet plane, C_f, are given by:

$$C_r = C_\alpha \left[1 + \left(\frac{1 - C_\alpha}{C_\gamma - C_\alpha} \right) \left(1 + \frac{A}{A_{eq}} \right) \frac{\sigma_r V_m}{RTa_y} \right], \tag{67a}$$

$$C_f = C_\alpha \left[1 + \left(\frac{1 - C_\alpha}{C_\gamma - C_\alpha} \right) \frac{2\sigma_r V_m}{RTa_y} \right]. \tag{67b}$$

Here, as in § 2.2.2, a_y is the half length (equivalent to the radius) of a cylindrical, plate-like precipitate and σ_r is the interfacial energy of the incoherent rim of the cylindrical precipitate. Despite the difference in solubility between the rim and the facets of the precipitate, with $A \gg A_{eq}$, the plates showed a reluctance to equilibrate towards the expected equilibrium shape, even with very extended coarsening times. RAJAB and DOHERTY [1989] reported similar behaviour in a further study of the same system. The conclusion reached by both studies was that this reluctance of plate-like precipitates in the Al–Ag_2Al system to change shape was due to the thickening reaction under the supersaturation arising from $C_f < C_r$, *being inhibited by a lack of growth ledges.* An unexpected result from the initial coarsening study in 1979 with these out-of-shape Ag_2Al precipitates was that despite inhibited thickening, the coarsening reaction, measured by the rate of *lengthening*, was very fast. However, further analysis of the results by use of the Ham, Horvay and Cahn analysis, supplemented by use of eqs. (67a and 67b), showed that the intuitive expectation, that inhibited thickening would slow coarsening, had been wrong. With inhibited thickening, so that A remains $\gg A_{eq}$, there is a much higher driving force for *coarsening*, as the only means of ridding the system of the excess interfacial energy, eq. (67a). Under the highly approximate condition of a constant aspect ratio, the analysis by FERRANTE and DOHERTY [1979] yielded the following prediction for the rate of growth of the mean half-length of the precipitate:

$$\overline{a_y}^3(t) = \overline{a_y}^3(o) + \frac{3AD\sigma_r V_m}{2\pi RT} \left(1 + \frac{A}{A_{eq}} \right) \frac{C_\alpha (1 - C_\alpha)}{(C_\gamma - C_\alpha)^2} \cdot t. \tag{68}$$

Eq. (68) gave surprisingly close agreement with the experimental results, despite the fact that while the mean aspect ratio remained constant, the individual values of A were found to show a strong correlation with precipitate size. The correlation between A and a_y indicated that there was a mean solute concentration determined by the Gibbs–Thomson solubility at the mobile ends or rims of the plates, eq. (67a). In the initial study by

Ferrante and Doherty a fine microstructure had been produced by low temperature ageing followed by coarsening at a higher temperature. In the later study by RAJAB and DOHERTY [1989] coarsening observations were carried out isothermally at both high and low temperatures by first precipitating and then coarsening Ag_2Al at the same temperature. Under these circumstances a slow decrease of the mean aspect ratio was found, but at rates very much slower than expected for diffusion-controlled transport from the rims of the precipitates to the facets of the same precipitate. Again a lack of thickening ledges appeared to be the reason — despite the observation that precipitate/precipitate intersections were still present. These intersections had been found to be effective sources of growth ledges under the larger supersaturations of the initial precipitation. That is, the supersaturations found during growth appeared to be sufficient to support nucleation of growth ledges while the smaller supersaturations found in coarsening appeared to be insufficient to allow a sufficiently density of ledges to form.

Detailed investigations of the coarsening of plate-like precipitates with $A \gg A_{eq}$ were provided by MERLE and FOUQUET [1981] and MERLE and MERLIN [1981]. They studied θ' plates in Al–Cu and found, in that system, that the precipitates formed with high and variable aspect ratios. As in the study by RAJAB and DOHERTY [1989] in Al–Ag, the highest aspect ratios of the θ' plates were found for those conditions with the lowest precipitate densities. During coarsening the aspect ratio fell towards what appears to be the equilibrium value of $A_{eq} \approx 20$. The rate of fall was *fastest* in those microstructures which had the *highest* precipitate density, see fig. 45. The implication of these results is clearly that the highest precipitate densities lead to the greatest chance of precipitate intersection and thus a source of growth ledges. MERLE and MERLIN [1981] and subsequently MERLE and DOHERTY [1982] showed that the kinetics of the change of shape could be described by a ledge model with solute transport at 225°C by *interface diffusion* around the incoherent edge of the precipitate. However, the agreement required the *experimental* measurement of ledge spacing and as noted in the discussion of growth of ledged precipitates § 2.2.4. there remains no way, at present, of predicting the ledge spacings, owing to the lack of a successful model for ledge nucleation. As noted by DOHERTY and RAJAB [1989] the inhibition of thickening during coarsening was more marked in Al–Ag than in Al–Cu for reasons that were not known. It would be good to

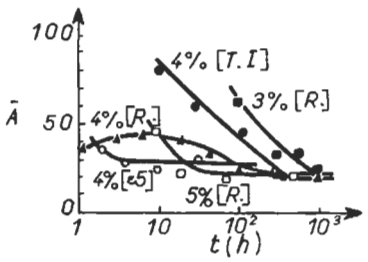


Fig. 45. The variation of the mean aspect ratio of the plate-like θ' precipitates in Al–Cu at 225°C. T.I. is an interrupted quench from the solution-treatment temperature, *R* is a room-temperature quench followed by the 225°C age, and ε5 is a 5% room-temperature strain prior to ageing at 225°C. 3%, 4% and 5% are the Cu contents. Note that $\bar{A}$ falls *faster* with the *highest density* of precipitates. (After MERLE and FOUQUET [1981].)

have more studies of the coarsening of plate like precipitates since these have the most common precipitate shape in successful precipitation hardening systems!

2.5. Discontinuous reactions: moving two-phase boundary MTPB reactions

The reactions discussed in this section are all those occurring at a *moving* interface, giving a *two-phase*, usually lamellar or rod-like, product. The moving interface may be an interphase boundary, as in eutectic and eutectoidal decomposition, or it may be a matrix grain boundary as in the precipitation reaction conventionally described as *discontinuous precipitation* (GUST [1979]). The moving interface is usually incoherent. This name, as previously noted in § 1.4, is an unfortunate one since the general name for the second class of transformations considered by Gibbs, those that do *not* involve nucleation and growth, is *continuous transformations* (SOFFA and LAUGHLIN [1983]). The term "discontinuous" should then apply to *all* Gibbs type-I, nucleation-and-growth, transformations and not just the subset of them that occur at a moving boundary giving a two phase product. (Or a two phase product reacting at a moving interface to give a single phase product as in discontinuous dissolution.) The term "discontinuous transformations" will be used here in the narrow meaning even though an alternative term, such as "moving two-phase boundary" (MTPB) reactions would be clearer. This review is not the place to make this unilateral change of notation, but care has been taken to avoid the expression "continuous precipitation" for reactions that either are nucleated entirely within a matrix grain or for those that are, nucleated at grain boundaries and then grow away from the boundary, for example Widmanstätten ferrite. The term *continuous transformations* will be restricted to reactions which do not involve nucleation and growth (§ 3).

2.5.1. Eutectoidal decomposition

In *eutectoidal decomposition*, the general reaction involves the change $\gamma \rightarrow \alpha + \beta$. The most important example is the "pearlite" reaction, the decomposition of austenite, γ, in an iron–carbon alloy containing 0.8wt% C, to give ferrite, α, and cementite, Fe_3C: $\gamma \rightarrow \alpha + Fe_3C$. The name *pearlite* comes from the similarity between the optical effect produced by light diffraction from the lamellar spacing (which is comparable wavelength of visible light) of the two-phase eutectoid and the same optical effect produced by "mother of pearl", a biological composite material.

The process can be understood from the schematic microstructure of fig. 46, for the transformation occurring isothermally at an undercooling ΔT below the eutectoid temperature of fig. 47. The extrapolations in the phase diagram, in fig. 47, give the compositions of the γ in equilibrium with α and β as $C_\gamma(\alpha)$ and $C_\gamma(\beta)$; the equivalent values in the α phase are $C_\alpha(\gamma)$ and $C_\alpha(\beta)$. If the main diffusion path is along the moving boundary, as it often is metal–metal eutectoids such as Cu–Al, then the compositions in the boundary are given as $k^b_{\alpha\gamma}\, C_\gamma(\alpha)$ and $k^b_{\beta\gamma}\, C_\gamma(\beta)$ where $k^b_{\alpha\gamma}$ and $k^b_{\beta\gamma}$ are the partition coefficients for solute between the appropriate phase boundary and the decomposing austenite. Since the separate values of the different coefficients are rarely known, an average value, k^b_γ, is often used.

References: p. 1497.

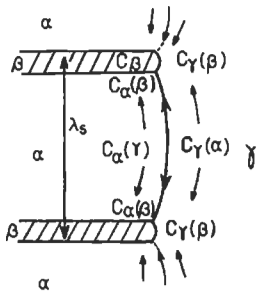


Fig. 46. Growth of an $\alpha+\beta$ lamellar structure, showing the variation in composition, $C_\gamma(\alpha)$, $C_\gamma(\beta)$ in the γ-phase, and $C_\gamma(\gamma)$, $C_\alpha(\beta)$ in the α-phase, that drive the reaction.

From figs. 46 to 48 it can be seen that there is a composition difference ΔC_γ within the γ-phase, $\Delta C_\gamma = C_\gamma(\alpha) - C_\gamma(\beta)$, and an equivalent composition differences ΔC_α in the α-phase and ΔC_b in the boundary. In the initial treatment of the diffusional problem given by ZENER [1946], it was assumed that diffusion took place down a solute gradient between the centers of adjacent lamellae. The mean gradient, dC/dy, is approximately $2\Delta C_\gamma/\lambda_s$ in the austenite, where λ_s is the lamellar spacing. The solute diffusion was assumed to occur within a thickness of approximately $\lambda_s/2$ away from the interface, where the gradient is assumed to have the mean value, with no gradient further away from the interface. By a mass balance of the diffusing solute needed for the reaction, the following relationship between the growth velocity, v, and the spacing was obtained:

$$v = \frac{2D_\gamma \Delta C_\gamma}{f_\alpha f_\beta} \cdot \frac{1}{C_\beta - C_\alpha} \cdot \frac{1}{\lambda_s}, \tag{69}$$

where f_α and f_β are the relative widths of the two phases, assuming that the molar volumes are the same in both phases (HILLERT [1972]). Eq. (69) fails to take into account, however, the interfacial energy, $\sigma_{\alpha\beta}$, of the interface between the product

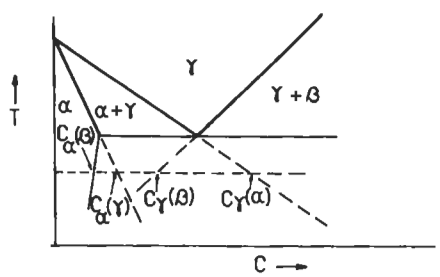


Fig. 47. The equilibrium diagram showing the origin of the composition differences seen in fig. 46.

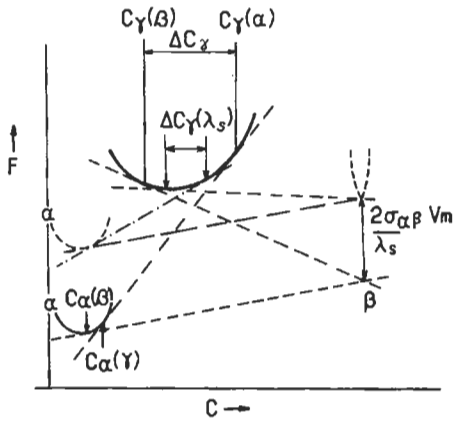


Fig. 48. The free-energy–composition diagram that gives the diagram seen in fig. 41. The composition difference ΔC_γ is reduced to $\Delta C_\gamma(\lambda_s)$ by the rise in free energy per mole of $\alpha+\beta$, $2\sigma_{\alpha\beta}V_m/\lambda_s$, caused by the lamellar spacing.

phases.* The origin of the lamellar equivalent to the Gibbs–Thomson effect is shown in fig. 47. The interfacial energy per unit volume of the lamellar product is $2\sigma_{\alpha\beta}/\lambda_s$ and so the two-phase product is raised in molar free energy by $2\sigma_{\alpha\beta}V_m/\lambda_s$ (J/mole^{-1}) with a consequent reduction of ΔC_γ to ΔC_γ (λ_s). If the critical spacing that reduces ΔC_γ to zero, λ_s(rev), is included, eq. (69) can be changed to give the velocity corrected for the interfacial energy effects. The description of λ_s(rev) is used since at that spacing, both the two-phase product with this spacing and the high-temperature austenite phase have *the same free energy* and so the reaction would be reversible at $\lambda_s=\lambda_s$(rev). With ΔC_γ changed to $\Delta C_\gamma(\lambda_s)$, the velocity becomes:

$$v=\frac{2D_\gamma\Delta C_\gamma}{f_\alpha f_\beta\left(C_\beta-C_\alpha\right)}\frac{1}{\lambda_s}\left(1-\frac{\lambda_s(\text{rev})}{\lambda_s}\right). \tag{70}$$

In deriving these diffusional equations the compositions are used in units of atoms per unit volume. However since the composition terms occur as $\Delta C'/\Delta C''$, use of atomic fractions will introduce only very small errors arising from small changes in the atomic volumes.

In the iron–carbon system, the diffusion coefficient in the α phase product is significantly faster than in the close-packed austenite, γ, so diffusion in α, is an alternative pathway which can be described by appropriate modification to eq. (70). For solute transport by diffusion in the interface the analysis by TURNBULL [1955] showed

* In the simple derivation of this Zener model, it is easiest, as here, to use the $\alpha\beta$ interface. For the point of changing the concentrations at the moving interfaces the energies of the $\alpha\gamma$ and $\alpha\beta$ interfaces are critical. The parameters are interrelated by the triple-point angles where the three phases meet, see JACKSON and HUNT [1966] and KURZ and FISHER [1989].

that not only must the compositions and the diffusion coefficient be changed but the diffusion thickness is reduced from $\lambda_s/2$ to the boundary thickness, δ_b, so the resultant equation is:

$$v(b) = \frac{8D_b k_\gamma^b \Delta C_\gamma \delta_b}{f_\alpha f_\beta (C_\beta - C_\alpha)} \frac{1}{\lambda_s^2} \left(1 - \frac{\lambda_s(\text{rev})}{\lambda_s}\right). \tag{70a}$$

HILLERT [1969, 1972, 1982] has discussed these equations and the more rigorous solutions to the diffusion equations in these various cases. He also considers the possibilities that there may be some significant supersaturation left in the product phases and that some free energy, or equivalently composition difference, may be required to drive the interface. That is, the reaction front may move faster than the rate required to achieved full equilibrium and that there may some element of interface control. For full analytical details and analysis of experimental results the reviews by Hillert, quoted above, should be consulted.

It is apparent from eqs. (70) and (70a) that the problem is not fully specified, even in the simple treatments, until some means of determining *the operating lamellar spacing* can be found. Experiment shows that there is a constant spacing at an given undercooling and that this spacing decreases as the undercooling increases and consequently $\lambda_s(\text{rev})$ decreases. For a eutectoid reaction with a molar entropy change ΔS_m, at an undercooling ΔT below the eutectoid temperature, $\lambda_s(\text{rev})$ is given by:

$$\lambda_s(\text{rev}) = \frac{2\sigma_{\alpha\beta} V_m}{\Delta S_m \Delta T} \tag{71}$$

A discussion of the choice of spacing is given in § 2.5.4. For the usually adopted assumption of a maximum velocity, we find $\lambda_s = 2\lambda_s(\text{rev})$. This gives the usual expression for volume diffusion control of:

$$\begin{aligned} v &= \frac{D\gamma \Delta C\gamma}{f_\alpha f_\beta (C_\beta - C_\alpha)} \frac{1}{\lambda_s} \\ &= \frac{D\gamma \Delta C\gamma}{f_\alpha f_\beta (C_\beta - C_\alpha)} \frac{1}{2\lambda_s(\text{rev})} \end{aligned} \tag{70b}$$

There is a simple relationship between ΔC_γ and undercooling, ΔT:

$$\Delta C_\gamma = \Delta T \frac{m_\alpha + m_\beta}{m_\alpha m_\beta}$$

Here m_α and m_β are the slopes of the γ solubility with temperature $m_\alpha = dT/dC_\gamma(\alpha))$. So combining these equations together we find the often quoted form of this valid for all undercoolings:

$$v\lambda_s^2 = \frac{4D_\gamma \sigma_{\alpha\beta} V_m}{\Delta S_m f_\alpha f_\beta (C_\beta - C_\alpha)}. \tag{70c}$$

The equivalent equation derived by JACKSON and HUNT [1966] has the same form as this but is expressed in terms of the interfacial energies, $\sigma_{\alpha\gamma}$ and $\sigma_{\beta\gamma}$, and the triple point angles, of the moving interfaces with the parent γ matrix.

The growth velocity, eq. 70b, initially increases with undercooling (as ΔT^2) due to the effect of temperature on the spacing and ΔC_γ. However there is the usual exponential fall of the diffusion coefficient with temperature, so that although the eutectoid growth velocity initially *increases* with undercooling, it quickly reaches a maximum possible velocity and then the growth velocity decreases rapidly in the temperature range where diffusion becomes very slow. This qualitative pattern is always found and is of great importance in the heat-treatment of steels. In outline, the carbon-containing austenite needs to be quenched to form the high hardness martensite and to do this, the diffusional decomposition to pearlite, and the other diffusional products, allotriomorphic and Widmanstätten ferrite and also bainite (§ 2.6), must be avoided. To achieve this, either very rapid quenching is required or for samples too thick to be rapidly quenched, the transformation to pearlite needs to be slowed down by inhibiting either the nucleation of pearlite or its growth. The experimental diffusional kinetics are usually reported on time-temperature-transformation (TTT) diagrams showing the time for isothermal transformation; these diagrams normally show "*C*-curve" behaviour (fig. 49), the "nose" of the *C*-curve being that for the maximum overall rate of transformation which has the usual

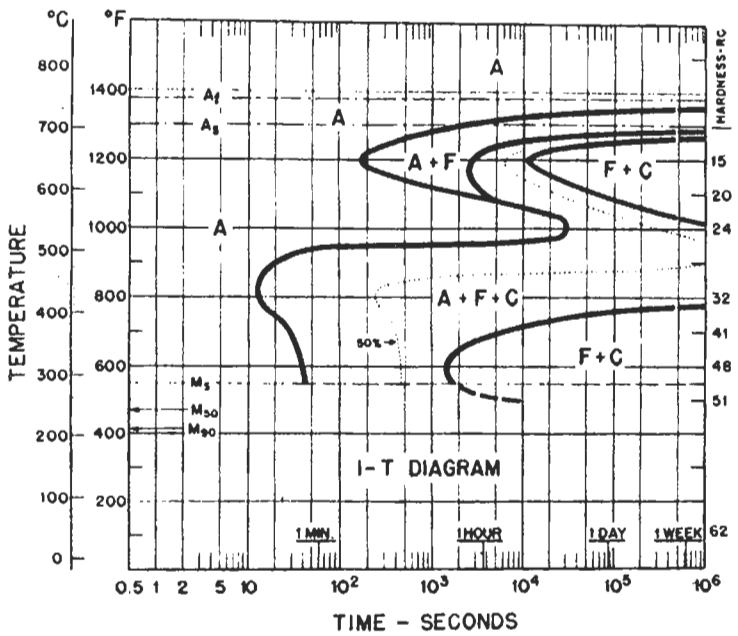


Fig. 49. Isothermal TTT diagram for a low-alloy steel, containing 0.4% C, 0.3% Mo, 0.8% Cr, 1% Mn and 1.8% Ni. The high-temperature transformation is firstly to precipitate ferrite (F) followed by pearlite (F+C), while the lower-temperature diffusional reaction is the formation of bainite (§ 2.6). (From the Atlas of Isothermal Transformation and Cooling Curves, ASM, Metals Park, OH, 1977.)

References: p. 1497.

"*S*" kinetics of a JMAK type of nucleation and growth reaction (§ 2.4.1). In TTT diagrams, of course, the *x*-axis is not velocity but its effective reciprocal, the isothermal holding time to achieve a certain fraction transformed. The kinetics of the experimental TTT curve depend not only on the variation of growth rate with temperature but also the equivalent variation of nucleation rate. For the same reasons, the nucleation behaviour also show *C*-curve variation with temperature, § 2.1. The same type of *C* curve behaviour, for the equivalent reasons, is found for the diffusional precipitation of single phase a iron α (ferrite "F" in fig. 49) from γ austenite ("A") and the lower reaction of austenite to "bainite" (§ 2.6). These reactions as well as the formation of pearlite are seen in fig. 49.

Although the transformation can be slowed down by an increase in austenite grain size, which increases the spacing between the grain-boundary nucleation sites, this is not the best solution, for reasons related to optimizing mechanical properties. The best method of improving the "hardenability" of steels is by *slowing* down the pearlite *growth rate*. This is achieved by addition of small amounts of substitutional metallic contents (nickel, manganese, chromium, etc.) which are soluble in the austenite but which at equilibrium will partition between the two product phases, but do so by *slow substitutional diffusion*. This idea the basis of *low-alloy* steels. At high temperatures the solute must partition since pearlite with uniformly distributed solute will have a higher free energy than the parent austenite. At lower temperatures, and higher undercoolings, there will be sufficient free-energy difference to allow transformation with "no partitioning" (HULTGREN [1947]) but with an inevitable decrease in the free energy driving force and therefore an increase in the critical spacing, λ_s(rev). The "no partitioning" of ternary solute in eutectoids is comparable to the diffusionless precipitation of single-phase products discussed in § 2.2.1. The reduction in growth velocity caused by these ternary solute effects increases with the amount of the low-alloy solute additions.

There has been considerable experimental interest in studies of solute partitioning in steels (see, for example, WILLIAMS *et al.* [1979], RIDLEY and LORIMER [1981], SMITH *et al.* [1981] and RICKS [1981]). Although it is well established that at high temperatures there is strong solute partitioning in low-alloy steel pearlites, and that with a fall in temperature the partitioning becomes incomplete, experimental difficulties appear, so far, to have prevented demonstration of the full no-partitioning of solute during the pearlite transformation. The difficulty is spatial resolution of the analysis, since at higher undercoolings the experimental spacings become rather too fine for the current analytical resolution. For the bainite reaction in steels there seems little doubt that there is no partitioning of the substitutional metallic elements between the ferrite, α, and the carbide phases § 2.6.

2.5.2. Discontinuous precipitation — MTPB precipitation

An equation similar to eq. (70) describes the growth of a two-phase product behind a moving matrix grain boundary that relieves a supersaturated α matrix, the reaction being $\alpha \rightarrow \alpha' + \beta$. This process occurs in many low-temperature precipitation reactions and arises after grain-boundary nucleation of the precipitate phase β which can then be supplied with solute by grain-boundary diffusion at temperatures where the competitive processes involve solute transport by *very much slower bulk* diffusion. The solute can be supplied from regions adjacent to the boundary — a process that is called the "collector

plate mechanism" if the boundary does not move. However, if the boundary can be induced to migrate into one of the grains, it can then collect and transport the solute from the supersaturated grain interior that it passes through, without requiring any bulk diffusion. The precipitates initially nucleated on the boundary can then grow while remaining attached to the moving boundary giving the characteristic lamellar structure. GUST [1979] published a very detailed review of the literature on this topic, starting from the original observation of the precipitation of copper at moving grain boundaries in Ag–Cu alloys by AGEEW and SACHS [1930]. There are several points of interest in this type of transformation. They include the problem of how and why the reaction initiates, that is why the boundary starts to move, and how it is possible for different parts of the same boundary to migrate in different directions, as reported for example by WILLIAMS and EDINGTON [1976]; the question of what determines the residual supersaturation in the product, since in many cases this is reported to be appreciable (e.g., PORTER *et al.* [1974]); there is also the general problem of determining which reaction, the discontinuous one or growth by volume diffusion, will occur under any given circumstances. Studies by SULONEN [1964] and more recently by CHUNG *et al.* [1992] have shown that many of these questions may be answerable in terms of an *elastic stress model* for boundary migration in discontinuous precipitation. This is equivalent to the process driving the process usually called diffusion (or chemically) induced grain boundary motion — DIGM discussed in § 2.5.5. The role of elastic stress in both discontinuous, MTPB, precipitation and DIGM is considered in § 2.5.5.

There is extensive evidence for significant residual supersaturation in the products of discontinuous precipitation. This comes from lattice-parameter measurements and from very high resolution microanalysis by electron energy loss techniques applied to Mg–Al alloys (PORTER *et al.* [1974]) and to Al–Li alloys (WILLIAMS and EDINGTON [1976]). Figure 50 gives the chemical analysis results across a matrix grain boundary in Mg–Al. This shows (i) a discontinuous change at the boundary, characteristic of a process proceeding by boundary diffusion, and (ii) that the resulting solid solution is still supersaturated in the aluminium solute.

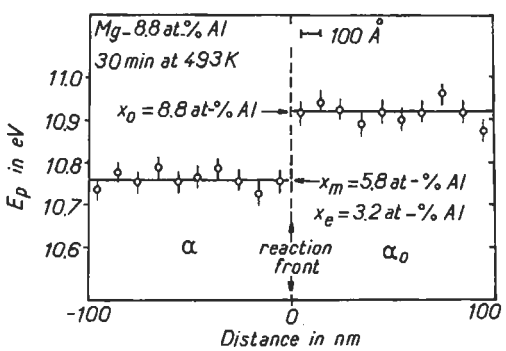


Fig. 50. Microanalysis by electron energy loss, across the migrating grain boundary during discontinuous precipitation in a Mg–Al alloy (after PORTER *et al.* [1974]).

The initiation of the discontinuous reaction was studied by TU [1972] in the Pb–Sn, solder, system at room temperature. His micrographs show regular grain-boundary nucleation of a colony of Sn-rich precipitates which grow apparently with an orientation relationship with the grain that is *not* consumed, when the boundary moves. Once the boundary starts to move, further precipitates form to give the operating lamellar spacing.

An important question concerns the conditions required for the discontinuous reaction to dominate. This will presumably require that the discontinuous process occurs faster than any alternative reactions such as *grain-interior* precipitation (usually of an easy-to-nucleate, usually metastable phase) or the growth of a boundary nucleated phase *into* the grain away from the stationary boundary. Both of these processes occur by mechanisms whose rates are determined by bulk diffusion. It might be expected, then, that the discontinuous process will take over from growth of boundary precipitates as the temperature falls, since the relative advantage of boundary diffusion over bulk diffusion increases as the temperature falls. Boundary diffusion has a lower activation energy. Results reviewed by GUST [1979] confirm this description with the discontinuous process in, for example, copper-rich Cu–In (PREDEL and GUST [1975]) dominating the precipitation process as the temperature of the reaction fell below 300°C. However, if the temperature falls very far, general grain-interior nucleation dominates most precipitation reactions, at least in alloys selected for precipitation hardening (KELLY and NICHOLSON [1963]). For such alloys, high temperatures of precipitation frequently cause discontinuous precipitation of the stable phase in place of the general precipitation of the strengthening metastable phases. On this basis, there should be only a narrow "window" of conditions that allows the discontinuous reaction to proceed. This is just as well since discontinuously precipitated alloys do not have good mechanical properties; having neither the high strength of precipitation-hardened alloys nor the good toughness of single-phase alloys. The studies by SULONEN [1964a and b] and by CHUNG *et al.* [1992], discussed in § 2.5.5, show clearly that, in addition to these requirements, there must be an elastic stress effect due to a change in the lattice parameter of the α matrix by changes in the solute content is also required. It appears as if the elastic stress in the consumed grains drives the process so it should be possible to develop a model for the kinetics of this type of discontinuous, MTPB, precipitation based on the elastic stress model. Such a model, however, does not yet appear to have been published.

2.5.3. Discontinuous (MTPB) coarsening

The ordinary coarsening process, described in § 2.4, involves the reduction of interfacial energy by growth of large precipitates, at the expense of smaller ones, by diffusion either through the lattice or — for a set of precipitates all linked by grain boundaries — by boundary diffusion in *stationary* boundaries. LIVINGSTON and CAHN [1974] discovered a *discontinuous* (MTPB) *coarsening* reaction in polycrystalline eutectoids in Co–Si, Cu–In and Ni–In, when the microstructures were annealed at temperatures close to the eutectoid temperature. Figure 51 shows a typical microstructure in which the upper grain is growing into the lower grain on the right while the opposite is happening on the left. The grain boundary is only visible in the grey Co_2Si phase. The driving force for this reaction certainly involves the reduction in interfacial area between

Fig. 51. Discontinuous coarsening of eutectoid in Co–Si alloys; a polycrystalline sample annealed at 1000°C for 96 h. × 1875. (From LIVINGSTON and CAHN [1974].)

the two phases. The reaction kinetics were successfully analyzed by LIVINGSTON and CAHN [1974] on the basis of boundary diffusion to give a predicted boundary velocity:

$$v = \frac{8C_b D_b \delta_b}{f_\alpha^2 f_\beta^2 \left(C_\beta - C_\alpha\right)} \frac{\sigma_{\alpha\beta}}{\lambda_2^2 \lambda_1 RT} \left(1 - \frac{\lambda_1}{\lambda_2}\right), \tag{72}$$

where λ_1 and λ_2 are the spacings before and after coarsening, C_b is the solute composition in grain boundary $\sigma_{\alpha\beta}$ is the energy of the interface between the two phases and the other terms have their previous meanings. This analysis was consistent with all the observations made. Of particular interest was the clear demonstration of which way the boundary moves. As seen in fig. 51, the grain whose lamellae lie nearly parallel to the initial boundary plane is able to grow, since it automatically has a larger effective spacing along the boundary. The only difficulty that arises here is the inevitable change of interfacial plane that occurs with this process, even though the two phases have the same relative orientations; the properties, particularly the energy of the interface, $\sigma_{\alpha\beta}$, may vary with boundary orientation unless the interface is fully incoherent.

A further very important point, discussed by Livingston and Cahn, was the magnitude of the spacing ratio, λ_2/λ_1, which was shown to be between 5 and 7 for Cu–In and Co–Si and between 10 and 20 for Ni–In. This is discussed in more detail in § 2.5.4. At present there is no evidence if discontinuous coarsening required movement of the grain boundary is, or is not, driven by elastic stress effects.

CHUANG *et al.* [1988] studied the discontinuous coarsening reaction in an Ni–7.5at% In alloy where the initial lamellar structure had formed by discontinuous precipitation.

They showed that the initial lamellar structure retained, as it commonly does, a significant solute supersaturation in the matrix phase and that during coarsening this residual supersaturation was essentially removed. In this case, both the residual supersaturation and the interfacial energy were driving the reaction. The authors reported that the transport mechanisms, in both the discontinuous precipitation and the subsequent coarsening reaction, were by grain-boundary diffusion. Somewhat smaller increases (by 3 to 5 times) of the lamellar spacing occurred here than in the coarsening reaction studied by Livingston and Cahn whose starting structure had been produced by eutectoid decomposition.

Other examples of discontinuous coarsening had been published by GUST [1979], though without discussion. One further example has already been mentioned, the work of WILLIAMS and EDINGTON [1976] using Al–Li alloys. At low ageing temperatures close to room temperature, the discontinuous reaction apparently involved formation by *precipitation* of the two-phase lamellar structure from a fully supersaturated solid solution, but at higher ageing temperatures the discontinuous reaction was one of *coarsening* when a lamellar, $Al + Al_3Li$, cube-cube oriented, product grew out from grain boundaries to consume the fine coherent precipitate nucleated in the adjacent grain. The reaction halted when the spherical Al_3Li distribution had coarsened by the usual LSW process, and also presumably when the aluminium supersaturation had been removed by completion of the grain-interior reaction. An interesting result reported by Williams and Edington was that grain boundaries close to coincidence-site orientations (ch. 9, § 2.2) were less able to show the discontinuous reaction than general high-angle grain boundaries.

DOHERTY [1982] discussed the formation of similar coarsened lamellar structures by deformation-induced boundary migration (recrystallization) in nickel alloys containing coherent Ni_3Al precipitates. Here the driving force for boundary movement was clearly the stored energy of the dislocations.

An interesting final reaction in this series is one described by NAKKALIL and GUPTA [1987]: this was the discontinuous (MTPB) *dissolution* of discontinuously precipitated Al–20% Zn when annealed close to the solvus temperature. Of interest here is that the spacing, of the dissolving two phase structure is fixed. Does the reaction have time to reach equilibrium giving a homogeneous solid solution and what role, if any, does the coherency stress play?

2.5.4. Determination of lamellar spacing in discontinuous (MTB) reactions

In the initial ZENER [1946] analysis it was assumed that the system adopted the spacing that allowed the interface to migrate at the *maximum velocity*, which for the volume-diffusion mechanism gives $\lambda_s = 2\lambda_s(\text{rev})$ and for interface diffusion mechanism, $\lambda_s = 1.5\,\lambda_s(\text{rev})$. PULS and KIRKALDY [1972] discussed alternative optimization criteria, including the spacing that gives the *maximum rate of entropy production*. This yields somewhat larger optimum spacings, which are $3\lambda_s(\text{rev})$ and $2\lambda_s(\text{rev})$, for volume and interface diffusion mechanisms, respectively. There is, however, considerable difficulty in estimating the appropriate values for $\lambda_s(\text{rev})$ since its evaluation requires knowledge of the interfacial energy as well as the free energy driving the eutectoid reaction. Free-energy calculations for discontinuous precipitation will be easier, particularly in dilute

alloys. However, as LIVINGSTON and CAHN [1974] pointed out, the discontinuous coarsening reaction does not have either of these difficulties, since the relevant parameter is the spacing ratio which can be easily *directly* measured, this is not true however if the initial lamellar structure still retains some solute supersaturation. The spacing ratios in coarsening were found by Livingston and Cahn to be considerably bigger than predicted by either of the proposed optimization criteria, maximum velocity or maximum rate of entropy production. The coarsening results therefore suggest growth well away from the maximum-velocity condition, at larger spacings. The same conclusion is also suggested by various experimental studies on discontinuous precipitation (RUSSEW and GUST [1979] and SPEICH [1968]), results reviewed by HILLERT [1982]. It was shown by Hillert that the experimental results lay close to the line, predicted by his diffusional model, of a plot of the residual supersaturation against lamellar spacing, but at *larger values of the spacing* than was predicted for the maximum-velocity criterion.

By comparison with the equivalent problem in dendritic growth, it would appear that there are *experimental* reasons, firstly, for expecting growth on the *large* side of the spacing predicted by the maximum-velocity criterion, and secondly, for the potential successful application of some type of stability criterion for lamellar spacing. The choice of spacing will also need to be analysed in the context of the insights into the role of elastic stress described below.

2.5.5. Diffusion-induced grain-boundary migration (DIGM)

In discontinuous, MTPB, reactions the available free energy of the precipitation, for example IJ in fig. 1, might be applied solely to the growing β-phase alone and the growth of the β-phase then may pull the depleted α matrix along. Alternatively the energy could be applied to *both phases*. HILLERT [1972, 1982] in his analysis of discontinuous precipitation has used this idea of the partition of the free energy between the two phases. In order to try to study this problem more fully, a simpler experimental situation was devised by HILLERT and PURDY [1978]. Their experiment identified a process that is now called by the title of this section, *diffusion-induced grain-boundary migration*, DIGM. CIGM, for chemically induced boundary migration, is sometimes used as an alternative name for the same process. DIGM can be achieved by alloying, or dealloying, a polycrystalline sample from the surface at a temperature at which bulk diffusion has been almost "frozen out", but where boundary diffusion still takes place at a significant rate. HILLERT and PURDY [1978] introduced zinc into thin polycrystalline iron samples from a vapour source, an Fe–11.3wt% Zn alloy, at temperatures between 545 and 600°C. They observed boundary migration which left a layer of zinc-enriched solid solution behind the migrating boundary. They treated the driving force for this reaction is the free energy of mixing per atom, ΔF_m. The effective or apparent mobility, M_a, of the boundaries was then estimated from

$$v = M_a \Delta F_m. \tag{73}$$

Apparent mobilities of grain boundaries in iron, at 580°C, of M_a of about 10^{-17} m^4/J s were found which were much smaller than the values found from ordinary boundary-mobility experiments or from the work of SPEICH [1968] on discontinuous precipitation

in iron-rich Fe–Zn alloys ($M = 10^{-5}$ and 10^{-12} m^4/J s respectively). In studies of discontinuous precipitation in Cu–Co by PEROVIC and PURDY [1981] values of M_a lower by 10^4 than boundary mobilities measured in recrystallization were also reported. It was not clear, at that time, whether the discrepancy arises from extraneous sources such as solute drag or if the use of eq. (73) to determine velocity is not justified. The process of DIGM had been previously seen, for example by RHINES and MONTGOMERY [1938] and by DEN BRODER [1972] and was subsequently identified in solid-liquid mixtures by YOON and HUPPMANN [1979]. The paper of HILLERT and PURDY [1978] created a great deal of interest in the process and numerous subsequent studies of the phenomenon followed as described by the reviews of HANDWERKER [1989] and YOON [1989].

CAHN *et al.* [1978] and BALLUFFI and CAHN [1981] reviewed the phenomenon of DIGM and noted that it occurred under a whole range of conditions, including both alloying and dealloying in systems with either a positive or a negative deviation from ideal solution behaviour, that is, with positive or negative heats of mixing. CAHN *et al.* [1978] draw attention to the critical theoretical problem in DIGM and equivalently in discontinuous precipitation. This is how the overall free-energy decrease can *couple* with the individual atomic motions required for boundary migration. The problem is shown in the schematic migrating-boundary structure in fig. 52. The boundary contains solute in equilibrium with the solute content of the growing grain on the *left* but, apart from the boundary plane, there is no solute in the grain on the right. If an atomic layer of solute-rich alloy is formed on the grain on the left by atom transfer from the right, together with solute addition *down* the boundary, a decrease of free energy of mixing will have occurred. However, this difference in free energy is apparently *not available* to the atoms in the right grain, that must jump across the boundary. CAHN *et al.* [1978] proposed a model for the process based on *differences in the rate of boundary diffusion*, the grain-boundary equivalent of the Kirkendall effect. This has the result that extra lattice sites

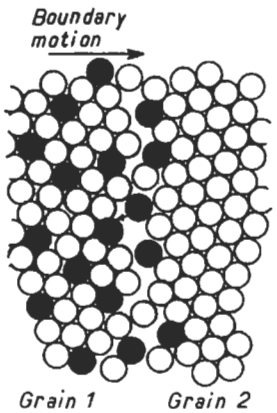


Fig. 52. Diffusion-controlled grain-boundary motion; solute atoms, solid circles, have diffused down the boundary and are causing growth of grain 1, which is being solute-enriched and therefore causing a reduction in free energy by the free energy of mixing. What is the driving force causing a net transfer of atoms across the boundary?

must be created, or lost, at the boundary which provides, via the climb of grain-boundary edge dislocations, a means of moving the grain boundary. There is clear evidence, for example in the Fe–Zn alloy experiments of Hillert and Purdy, that there is surface tilting at the migrating grain-boundary regions, characteristic of volume changes predicted by the boundary Kirkendall effect.

The topic of DIGM has been reviewed several times recently, by HANDWERKER [1989], YOON [1989] and DOHERTY [1992] and it now seems clearly established that the driving force that moves the grain boundary is the *elastic coherency stress due to the change of lattice parameter in the matrix* in the layer of solute enriched material immediately ahead of the moving grain boundary. An example of this is monolayer of such enrichment shown in fig. 52. If the grain boundary is moving with a velocity v and the bulk diffusion coefficient is D, then the thickness, δ_s, of the solute-enriched layer, in the grain being consumed, is given by:

$$\delta_s = D/v \quad \text{when } D/v > b \tag{74a}$$

$$\delta_s = b \quad \text{when } D/v < b. \tag{74b}$$

b is the atomic diameter of the matrix atoms. CHUNG *et al.* [1992] suggested that eq. (74a) would be expected to apply even when $D/v < b$, but this fails to take into account that diffusion of solute down a grain boundary does not occur in a channel of atom sites *within* the boundary but occurs by atoms which are *part of the surface layer of both grains*. Figure 52 sketches the situation expected of eq. (74b). The large driving force of the free energy of mixing will ensure that atoms in a grain boundary should always be part of the surface layers of both grain in a boundary and this idea is clearly confirmed by the *experimental* results reported by CHUNG *et al.* [1992]. In the situation where η, $\eta = (1/a)(\mathrm{d}a/\mathrm{d}C) \neq 0$, then the presence of such a solute-rich layer, of composition C_s different from the matrix composition C_0 with either $\delta_s = b$ or $\delta_s > b$, will give rise to an elastic energy per atom, ΔF_{Ea}. This occurs since the solute modified layer which unstressed will have a different lattice parameter from the unsolute modified part of the grain but is coherently stressed to that lattice parameter. This elastic coherency energy is given by:

$$\Delta F_{Ea} = V_a Y(\mathbf{n})\delta^2 \tag{75a}$$

Here V_a is the volume per atom, $Y(\mathbf{n})$ is the elastic modulus that varies with the surface normal $\mathbf{n}$ and misfit δ is the strain given by:

$$\delta = \eta(C_s - C_0) \tag{75b}$$

What appears to happen during DIGM is that the solute diffuses down the initially stationary grain boundary and enters the grains on both sides of the boundary to a depth that is approximately $(Dt)^{0.5}$, where D is the bulk diffusion coefficient and t the time of diffusion. There will be a steady build up of elastic strain until on *one* side of the boundary the strain is *plastically relaxed* by the formation of a cross-grid of edge dislocations. These dislocations take up the change of lattice parameter between the

solute enriched region and the original composition. Such dislocation arrays, at a position close to the original position of the grain boundary, are commonly reported, see for example PAN and BALLUFFI [1982]. Once the elastic energy is plastically relaxed on one side of the boundary, then there is a driving pressure to cause migration of atoms *from the grain that is still elastically strained.* The velocity of the migration is given by equation (73) but with the driving force that of the elastic coherency energy, ΔF_{Ea}, not the full free energy of mixing, ΔF_m. The free energy of mixing drives the diffusion process down the grain boundary and into the grains so giving rise to the elastic strain energy. Since $\Delta F_m \gg \Delta F_{Ea}$ the low apparent value of boundary mobility in DIGM noted above arises since the incorrect larger driving force was used for its evaluation.

The elastic stress energy explanation for DIGM was demonstrated by several studies from Yoon and his colleagues using liquid–solid mixtures, as reviewed by YOON [1989] For example, RHEE and YOON [1989], using liquid phase sintered Mo–Ni established a uniform Ni content in the Mo powder particles which, after sintering, produced fully and partially liquid wetted grain boundaries. Liquid wetted boundaries were ones where the grain boundary energy, σ_{gb}, was larger than twice the solid liquid energy, σ_{sl}. The partially wetted boundaries were ones with $\sigma_{gb} < 2\sigma_{sl}$. On adding either Co or Sn to the liquid, there was rapid liquid and grain-boundary diffusion of the ternary solutes but slow bulk diffusion into the solid particles — causing elastic coherency strain in the surface layers on both sides of the boundary. When this strain was plastically relaxed by an array of edge dislocations on one side of the boundary, it caused DIGM of partially liquid wetted boundaries and an equivalent process of liquid film migration (LFM) of fully wetted grain boundaries. The authors were able to vary the strain energy driving the reaction by use of different combinations of Co and Sn: One solute, Co, gave a negative strain and the other a positive strain. Figure 53 shows the predicted *parabolic* relationship between the migration distance, in a fixed diffusion time, and the calculated coherency strain, ΔF_{Ea}, from eq. (75a) being used as e the driving pressure in eq. (73), with $v = M_a \Delta F_{Ea}$. The study showed the same predicted type of parabolic dependence on coherency strain for migration of grain boundaries ("CIGM" for chemically-induced grain-boundary migration) as for grain boundaries wetted by liquid films. The kinetics of LFM are apparently slightly faster than that of DIGM/CIGM. The coherency energy model of DIGM is also confirmed by the initial observation by HILLERT and PURDY [1978] that the position from which the grain boundary has moved in DIGM is marked by an array of dislocations that have relaxed the coherency strain in the grain that subsequently grew into the opposite grain. This plastic relaxation on the growing side of the boundary is what lowers the coherency energy of one grain and allows it to grow into the grain that is still elastically strained. It is commonly seen that after awhile DIGM stops — again a not unexpected feature if the grain which is being consumed also manages to relax the coherency strain plastically.

The coherency strain energy model had been previously identified, at least qualitatively, as controlling discontinuous precipitation by the elegant experiments of SULONEN [1964a and b]. In these experiments the velocities of moving grain boundaries, giving discontinuous precipitation, were modified by the application of unidirectional tensile or compressive stresses. Boundaries parallel or normal to the stress axis had their velocities

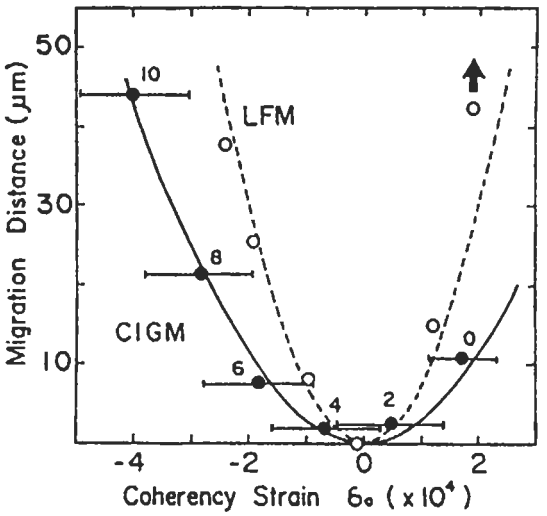


Fig. 53. The observed variation of the average migration distance for DIGM (labelled CIGM) and Liquid Film Migration (LFM) for Mo-Ni alloys embedded in Mo-Ni-Co-Sn on holding at 1460°C for 2 hours as a function of the coherency strain. The numbers above the filled circles are the values of the increased Co content. The Solid and dotted lines are the parabolic fits to the experimental data. After RHEE and YOON [1989].

increased or decreased in a predictable manner in a series of binary alloys dependent on the signs of the applied stress and the local elastic coherency strain in a range of systems. HILLERT [1972] provided a clear analysis of the thermodynamics of the situation. Interestingly, despite the availability of this critical evidence of the role of coherency strain energy in discontinuous precipitation, the insights were largely overlooked for nearly 30 years in much, but not all, of the subsequent literature of discontinuous precipitation and DIGM. It is pleasing to note that the good work of SULONEN [1964a and b] was eventually recognized. CHUNG *et al.* [1992] repeated the Sulonen experiments in an Al–21.8at% Zn alloy that after solution treatment was held under different tensile stresses at both high, 215°C, and low, 75°C, temperatures. The changes in boundary velocities of discontinuous precipitation were determined as a function of applied stress. Figure 54 shows the averaged growth rates of the two types of boundaries, parallel and transverse to the stress, as a function of applied tensile stress. The predicted transverse and parallel velocities, v_t and v_p, using the elastic energy relationships, derived by HILLERT [1972], were quoted by CHUNG *et al.* [1992] as:

$$v_t = M_a\left(\frac{Y\delta^2}{1-\nu} - \frac{2\delta\sigma\nu}{1-\nu}\right) \tag{76a}$$

$$v_p = M_a\left(\frac{Y\delta^2}{1-\nu} + \delta\sigma\right) \tag{76b}$$

The values of the apparent mobility, M_a, and the misfit, δ, were calculated by applying

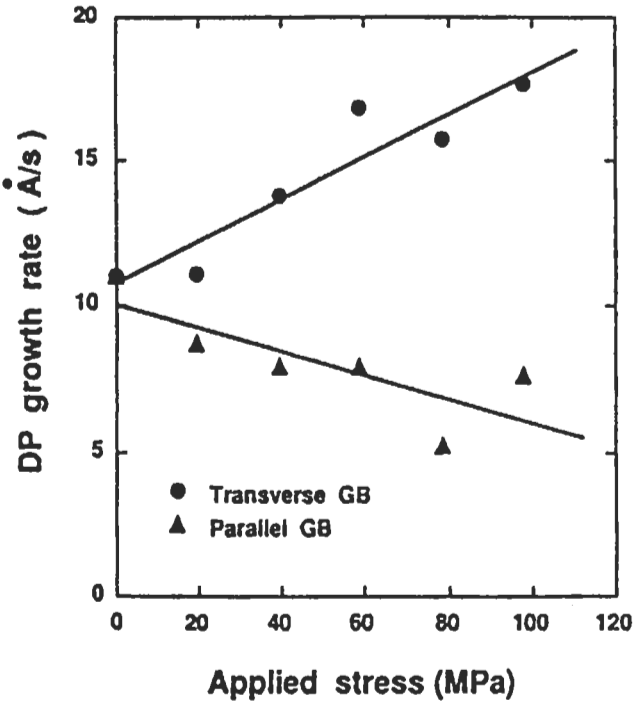


Fig. 54. The average discontinuous precipitation growth rate as a function of applied tensile stress, for boundaries normal (transverse) to the stress and boundaries parallel to the stress in Al-22at % Zn held at 75°C for 1 hour. After CHUNG, SHIN and YOON [1992].

these equations to the the slopes and intercepts of the two plots in fig. 54. The values found were $M_a = 4.1 \times 10^{-15}\ m^4\ J^{-1}\ s^{-1}$ and $\delta = -1.6 \times 10^{-3}$ for the transverse boundaries and $M_a = 2.5 \times 10^{-15}\ m^4\ J^{-1}\ s^{-1}$ and $\delta = -2.0 \times 10^{-3}$ for the parallel boundaries. The estimated value of the strain, δ, due to the misfit with $C_o = 21.8$at% Zn, the alloy composition, and $C_s = 12.8$at% Zn, the equilibrium solute content at the transformation temperature of 75°C was, from eq. (75b), $\delta = -2.5 \times 10^{-3}$. It is likely that the value of C_s will not fall all the way to the equilibrium solubility since most studies of discontinuous precipitation indicate that the product matrix retains some supersaturation. The very close qualitative and quantitative agreement between theory and experiment here is, however, very satisfactory and clearly indicates that the analysis appears valid. In other words, the identification by SULONEN [1964] of coherency stress as the central mechanism in discontinuous precipitation, that he had qualitatively demonstrated, is now quantitatively confirmed. It might be noted that exactly similar coherency effects should occur in eutectoid transformations and thus influence the migration kinetics. However the author of this review knows of no theoretical or experimental investigation of this idea.

CHUNG *et al.* [1992] were disturbed that the analysis worked as well for the conditions at the low temperature, 75°C, where D/v was very much less than an atomic

diameter, b, than at the higher temperature where D/v was much larger than an atomic diameter. However, as noted above, this result appears very reasonable since the surface layer of a grain boundary exposed to solute *should* contain a high solute concentration. In fact, it is clear that the results from CHUNG *et al.* [1992], together with similar results that they quote from earlier studies, demonstrate that the solute must penetrate the surface layer during boundary diffusion even when the *macroscopically predicted penetration*, D/v, fails to indicate this. In other words the analysis of Chung *et al.* shows clearly that, during grain boundary diffusion, the outer layer of both grains carries the solute content of the grain boundary and changes the equilibrium lattice parameter of the solute modified layer even for conditions (such as $D/v \ll b$) in which continuum modeling predicts no solute penetration.

The clear demonstration of the role of elastic strain energy in driving boundary migration in discontinuous precipitation immediately explains why the resulting matrix may be supersaturated. The boundary migrates at a rate determined by the product of the boundary mobility and the elastic energy and not at rates determined directly by the precipitation. So the boundary migration could easily be *too fast* to allow the full release of the matrix supersaturation. The result also allows an easy prediction as to the likelihood of the discontinuous precipitation process to occur. The larger the values of η and of $(C_o - C_s)$ the more likely is the discontinuous process to occur. BOHM [1961] had earlier noted the importance of the minimum magnitude of η in determining the onset of discontinuous precipitation in a range of copper alloys. A significant value of the elastic energy is only a necessary condition but it is not always a sufficient condition for this mode of precipitation since supersaturation may be released even faster by a high rate of nucleation and growth of grain interior usually metastable precipitates — with a low interfacial energy barrier to nucleation.

An important phenomenon, described by CHONGMO and HILLERT [1981], remains however a major mystery. In their study of DIGM in Fe–Zn they observed a phenomenon now called *diffusion-induced recrystallization* (DIR) on the surface of their samples. In some circumstances new small grains were *nucleated* at the surface of the sample in *regions away from prior grain boundaries*. The term "recrystallization" refers to the formation of new grains, usually as a result of annealing *plastically deformed material*. The process of nucleation of recrystallization, discussed in chapter 28, requires high angle grain boundaries to give the required mobility for new grains to grow. In plastically deformed materials the required misorientations are known to be produced by plastic deformation, see for example the review by DOHERTY [1978]. How solute entry into the surface of the material with its associated elastic coherency energy is able to produce the required large local misorientations remains a unsolved problem in DIGM. The phenomenon of DIR is reported as occurring more easily with rougher surfaces than with electropolished surfaces and it occurs more often at higher temperatures. Similarly, though the mechanisms of growth in DIGM now appears to be well understood, there is less insight into which grain DIGM moves into at arbitrary grain boundaries and what causes the common observation, see for example HILLERT and PURDY [1978] and CHONGMO and HILLERT [1981], that in some grain boundaries, migration of different parts of the same boundary occurs at different places into one grain and elsewhere on the

same boundary into the other grain. As noted above similar two-way boundary migration can happen in discontinuous precipitation. That is, the nucleation stage of DIGM, discontinuous precipitation and particularly DIR is not yet understood.

2.5.6. Experimental results on discontinuous eutectoidal reactions

PULS and KIRKALDY [1972] reviewed in detail the experimental results of eutectoidal decompositions. For the eutectoidal reaction in binary iron–carbon alloys, the reactions appear to go at the rates expected for volume diffusion of carbon, though with some indication of rate enhancement by additional boundary diffusion. In the observations on eutectoid growth in binary substitutional alloys several studies, such as the fcc Al–78% Zn eutectoid (CHEETHAM and RIDLEY [1971]) showed *boundary* diffusion to be dominant while in the Cu–Al eutectoid where the decomposing "austenite" phase is bcc with a high bulk diffusivity, *volume* diffusion control was reported (ASUNDI and WEST [1966])*. In all these eutectoid systems, as in eutectic solidification where the reviewed results were considered along with solid-state reactions by KIRKALDY and SHARMA [1980], the results were tested against models which assumed a maximum-velocity, or the equivalent, a minimum-undercooling criterion, and showed physically sensible behaviour.

From their study on Fe–C–X ternary alloys, PULS and KIRKALDY [1972] suggested that the data are apparently explicable on the reasonable basis that there is no partitioning of the substitutional solute and hence carbon diffusion control operates at low temperatures: At high temperatures, however, there is evidence for a shift to control by boundary diffusion of the substitutional elements, which are then partitioning.

Unfortunately there is lack of full information on the crucial parameters, such as the surface energies and the relevant diffusion coefficients, to allow a really satisfactory test of the theories to be made. However, at the moment, the overall situation appears to be one of residual doubt about the operating criterion that determines the spacing, but apart from this problem the theories appear in good shape. However it would appear that a general account of all moving two-phase boundary transformations: eutectoid decomposition, discontinuous (MTPB) precipitation, discontinuous (MTBB) coarsening and dissolution, DIGM and LFM is needed that includes *the role of coherency strain energy*.

2.6. Bainitic transformations

In the isothermal TTT diagrams for iron-carbon alloys there is an important transformation, seen in fig. 49, at temperatures below the "nose" of the pearlite reaction. The product of this lower-temperature transformation is called bainite and usually appears to be a mixture of α-Fe (ferrite) and finely dispersed cementite, Fe_3C. Two varieties of bainite, known as *upper bainite* and *lower bainite*, have been identified, corresponding to

* This result is interesting for the theory of precipitate shape stability (§ 2.2.6). The γ-phase in Cu–Zn and in Cu–Al precipitated from supersaturated β-phase as solid-state dendrites and it was argued in § 2.2.6.1 that a high rate of bulk diffusion relative to interface diffusion is required for such instabilities to form. That the eutectoidal decomposition of bcc β in Cu–Al is by bulk rather than interface diffusion is in agreement with this hypothesis.

their respective temperatures of formation. Upper bainite has carbides which appear to have formed with the usual orientation relationship found for carbides forming in the austenite matrix. Lower bainite appears to have carbides with an orientation relationship with the growing ferrite phase indicating that they precipitated within the ferrite, *after* it was formed. In some steels, particularly those with high levels of silicon, carbide free bainite is formed, at least initially. The subject has been reviewed and discussed many times by AARONSON [1969], HEHEMANN [1970], HEHEMANN *et al.* [1972] BHADESHIA and EDMONDS [1980], CHRISTIAN and EDMONDS [1984], AARONSON *et al.* [1987–89], BHADESHIA [1985, 1987] CHRISTIAN and EDMONDS [1988/9], AARONSON *et al.* [1990], BHADESHIA and CHRISTIAN [1990], AARONSON and HALL [1994a], AARONSON *et al.* [1994b], HILLERT [1994]. The last four references come from two recent conferences: INTERNATIONAL CONFERENCE ON BAINITE [1990] and ROLE OF SHEAR AND DIFFUSION IN PLATE SHAPED TRANSFORMATION PRODUCTS [1994] devoted solely to the topic of bainite and related transformations. There are difficulties in making a clear interpretation; the following represents a summary of what appear to be significant results reviewed in the light of the ideas presented in this chapter. For full details of the experimental results and differing views on the transformation the papers at the conferences and their cited references should be consulted. It should also be noted that the author of this review has not carried out any experimental or theoretical investigations in the field of bainite.

There are several interpretations of the experimental results on bainite, all of these disputes are connected with the mechanism of the *interface reaction* as the plate-like product, bcc ferrite in Fe–C bainite, grows from the matrix, fcc austenite in Fe–C bainite. They concern the nature of the interfacial reaction, the carbon supersaturation in the bainitic ferrite and the mechanism of the incomplete reaction in bainite.

i) *Bainite: displacive or reconstructive?*

One interpretation is that the formation of bainite plates in bainite occurs by a coordinated atomic motion, a mechanism similar to that seen in martensite. This is described as a "military" or a "displacive" reaction (CHRISTIAN [1965], BHADESHIA [1985] see §§ 1.1 and 1.6). The alternative viewpoint is that the transfer of atoms occurs by random, thermally-activated atomic jumps across the interface a mechanism described as "civilian" or as "reconstructive". The term diffusive for the civilian interface reaction will not be used here since, at some stage, in the formation of low carbon ferrite from austenite with a higher carbon content, there is a composition change occurring by the diffusional motion of carbon atoms (This is why the reaction is discussed in the present chapter). For clarity, the two models will be described here as either *displacive* or *reconstructive*. One of the difficulties of interpretation concerns the significance of the *surface relief* that accompanies the formation of plates of bainite. This type of surface relief is a defining experimental characteristic of a displacive reaction, see the following chapter, and the observation by KO and COTTRELL [1952] of surface relief in the bainite reaction was for most, but not all, students of the subject, the critical observation supporting the displacive mechanism.

The mechanism of the bainite reaction in iron-carbon alloys needs to be compared to the mechanisms of other austenite decomposition reactions in iron–carbon alloys. These other reactions include:

(i) The formation of carbon-depleted, grain-boundary "allotriomorphic" α ferrite or, at higher carbon contents, iron carbide, Fe_3C. The name allotriomorph comes from the non plate-like shapes of the grain-boundary products that form at high temperatures and low supersaturations.

(ii) The precipitation of non-carbon partitioned ferrite, the "massive" reaction, § 2.2.5.2, forming from very low carbon austenites.

(iii) The growth of lamellar sheets of α plus carbide, the "pearlite" reaction, § 2.5.1 and 2.5.6.

(iv) The growth of coherent "Widmanstätten" plates of either carbon-depleted ferrite or equivalently of iron carbide, § 2.2.6.2.

(v) The formation of "lath" martensite (with a high dislocation density) and, at higher carbon concentrations, "plate" (internally twinned) martensite in rapidly quenched iron carbon alloys. (These displacive reactions are discussed in the next chapter).

The first three reactions are agreed to be reconstructive. There is no surface relief reported during their growth. In addition, it is generally accepted that these grain-boundary nucleated structures grow into the austenite grain with which the product or products *have no good-fit orientation relationship*. That is, the interface has the same type of incoherent structure as is found for a high-angle grain boundary. The martensite reactions are agreed to be displacive and to occur with, characteristic surface relief, no partitioning of carbon at least during the initial growth of the martensite plates. The growth of martensite can occur at rates that approach the speed of sound in iron ($\approx$ 1 km/s). The growth of plates of carbon depleted Widmanstätten ferrite (iv) also has some aspects of an apparent displacive reaction, notably its surface relief effects, as described by CLARK and WAYMAN [1970] and by WATSON and MCDOUGAL [1973], and so the interface reaction in Widmanstätten ferrite is thus a central part of the bainite dispute. As previously discussed in § 2.2.6, the Widmanstätten plates result from growth into the parent austenite grain with which the product does have a good atomic fit orientation relationship. The habit plane of the precipitate is the interface plane that shows the good atomic fit. The growth rate or at least the lengthening of a Widmanstätten plate is known to be constant at a fixed undercooling and to occur at a rate matching that expected for carbon-diffusion control, § 2.2.6.2.

PURDY and HILLERT [1984] have argued for the close morphological similarity between the formation of Widmanstätten ferrite and bainite and between Widmanstätten ferrite and martensite by some simple and elegant metallographic experiments. Bainite plates in partially transformed Fe–C alloys were shown, by up-quenching, to continue their growth as Widmanstätten plates while maintaining the *same morphology they had as bainite*. Similarly, continued growth of Widmanstätten plates as martensite plates *parallel to the original ferrite plates* was shown by down-quenching partially transformed structures.

A mark of the essential crystallographic basis of the growth of both Widmanstätten ferrite and of bainite is that no demonstration of either structure growing across an austenite grain boundary appears to have been reported. KO and COTTRELL [1952] initially noted that bainite plates did not cross austenite grain boundaries. The crystallographic model for martensite growth, as for mechanical twinning, forbids the continued

growth of a twin or a martensite plate across a grain boundary into a crystal with which it does not have the required orientation relationship. Precipitates growing by random, thermally activated, atomic motion (the "civilian transformation", CHRISTIAN [1965]) can, and certainly do, grow across matrix grain boundaries, an important example being pearlite, as reported by MEHL [1947].

There are other precipitation reactions which are also described as "bainitic" by virtue of showing both *diffusional changes in composition* as well as *surface relief* effects. Examples include Cu–Zn (GARWOOD [1954], Cu–Al (SMITH [1962]), Cu–Sn (DE BONDT and DERUYTERRE [1967]), Ag–Cd (KOSTIC and HAWBOLT [1975]) in which plates of fcc (α) precipitation from bcc (b) all growing with significant supersaturation showed surface relief, as discussed by CLARK and WAYMAN [1970]. LIU and AARONSON [1970] reported surface relief effects for the Widmanstätten precipitation of Ag_2Al plates from supersaturated Al. Other reactions, in which there is a change of chemical composition and surface relief effects during precipitation, include: Cr–Cu, (HALL *et al.* [1975]), Ti–Cr, (LEE and AARONSON [1988]) and Cr–Ni (CLIFF *et al.* [1990]). In all these reactions there is long-range diffusion of *substitutional* solute combined with the interface reaction that gives rise to the surface relief. CLIFF *et al.* [1990] recently demonstrated that at the earliest stages of growth of these substitutional bainites that there was a measurable concentration difference between the two phases, so confirming that the reactions are not fully martensitic. Surface relief combined with long range diffusion of *interstitial* solute has been reported for oxide (LANDUYT and WAYMAN [1968]), hydride (BOWLES *et al.* [1977], CASSIDY *et al.* [1977]), and nitride (DAHMEN *et al.* [1987] precipitation as well as for the bainite and Widmanstätten ferrite reactions in Fe–C alloys.

For the precipitation of Ag_2Al, previously discussed in § 2.2.3 and 2.2.5.2, it is now agreed by the leading protagonists of the reconstructive model for bainite growth, AARONSON *et al.* [1990], that the Ag_2Al reaction does indeed occur by a combined diffusive flux of solute to the precipitate and a displacive interface motion. In this case the displacive reaction takes place by the glide of Shockley partial dislocations on every second {111} plane, giving a purely displacive transformation of the fcc matrix to the cph precipitate, CHRISTIAN and EDMONDS [1988/9]. These partial dislocations are the growth *ledges* discussed in § 2.2.3 and, as previously mentioned, the system shows surface relief effects. It appears therefore that a combined reaction showing long range substitutional solute transport by diffusion with an interfacial reaction that is completely described by dislocation shear model can and does indeed occur. It remains to be settled however if such a description applies to all, or even some, of the other reactions forming plate-like products with a good matching habit plane. Only for the Ag_2Al reaction does there appear to be full agreement on the structure of the interface. It might be noted that the martensitic reaction between fcc and cph is unusually simple in that the so-called "Bain-Strain" plus rigid body rotation that transforms, with minimum strain, one crystal structure into the product is the required "Invariant Plane Strain" for minimum distortion needed for the "Phenomenological Theory of Martensite" (see the next chapter for details of this theory). Most other martensite reactions are ones in which some other *lattice invariant shear*, usually dislocation slip or twinning, is needed, in combination with the Bain–Strain plus rigid body rotation, to produce the required *invariant plane strain*. In

these other martensitic reactions the interfacial structures needed for glissile displacive interface migration are much more complicated than that between fcc and cph.

KO and COTTRELL [1952] first observed the characteristic surface relief of the formation of bainite and SRINIVASAN and WAYMAN [1968] showed that the form of the surface displacements was the invariant plane strain characteristic of a martensitic reaction. There is also clear evidence from CLARK and WAYMAN [1970], WATSON and MCDOUGAL [1973] and BHADESHIA [1981]) that Widmanstätten ferrite also has a displacive character, with an invariant plane strain, demonstrated by the surface relief of the ferrite plates. Frequently the surface relief is described as being "tent-shaped" as though there were *two* back to back plates shearing in *opposite directions*, though both shears still have an invariant plane strain form. Such plates, with the tent shaped relief, are commonly reported to be mono-crystalline, HALL and AARONSON [1994], and thus they have not formed by the growth of two *differently oriented* product crystals. That is, the tent shaped surface relief is caused by growth of a single crystal plate showing opposite shear as the two opposite plate surfaces grow apart. In ferrite, there is also clear evidence for a high dislocation density in the product plates (both bainitic and Widmanstätten) just as is found in lath martensite formed from low carbon austenite. The dislocation density is found to be lower in ferrite plates formed at higher temperatures (see for example BHADESHIA and CHRISTIAN [1990]). One explanation of the surface relief effects and the dislocations is that the bcc ferrite grows from fcc austenite by a displacive mechanism. A reasonable explanation for the lower dislocation densities might be faster dislocation recovery at higher temperatures. Another view, see for example by AARONSON *et al.* [1990, 1994], is that the surface relief is caused by the glide of growth ledges with dislocation character along the ferrite/austenite interface. It is not clear, to the author of this review, what distinction there can be between a true displacive transformation and one occurring by such glide of interface dislocations. If the structural transformation from one structure to the other can be accomplished solely by the glide of the growth ledges that are dislocations (as in the fcc–cph transformation) then the reaction is displacive in character even if accomplished by ledge migration, just as in the agreed case of Ag–Al. The question that appears to remain unsettled, however, is if the glide motion of the dislocation/growth ledges, on the bcc/fcc interface in Widmanstätten ferrite and in bainite, can *fully* accomplish the transformation in structure or if some additional, thermally activated, atomic motion, not present in the fully glissile martensite reaction, is required in Widmanstätten ferrite and in bainite. It is perhaps significant, as noted by WAYMAN [1994], that the current phenomenological theory of martensite, described in the following chapter, does not as yet fit two the types of martensite formed in Fe–C alloys, in low carbon, dislocated lath martensite and in, medium carbon, plate martensite with near {225} habit planes. Given this uncertainty about two closely related martensite reaction, it is perhaps not a critical failure of the displacive mechanism for bainite and Widmanstätten ferrite that a full description of the process exactly matching the phenomenological theory is not yet available.

There has been, in the recent conferences on this topic, considerable analysis of the crystallography of the reactions involving plate precipitates showing surface relief. These

analyses include those of AARONSON *et al.* [1994], AARONSON and HALL [1994], CHRISTIAN [1990 and 1994], DAHMEN [1994], HIRTH [1994], HOWE [1994] and MUDDLE *et al.* [1994]. It seems rather well established that the motion of growth ledges which have dislocation character in the fully or partially coherent precipitates can give rise to a *lattice site correspondence* even if the occupation of these sites may change, becoming more solute rich as solute rich phase grows into the low solute matrix. That is, there is be atomic *site correspondence* but not *atom correspondence.* In other words, as the solute, needed for the change in chemical composition between the two phases, reaches the growth ledges, then these move forward with site correspondence by glide of the Shockley partial dislocations, structurally the riser of the ledge, and this causes two {111} atomic layers to change from fcc to cph *in structure and in composition.* Thermally activated atom exchange between the sites as they displace appears to be required to allow the thermodynamically required composition change to occur. For a Widmanstätten or bainitic ferrite plate to grow from an interstitial rich austenite the two coupled processes, on the displacive model, are the rapid diffusion away of interstitial atoms combined with the ferrite growth by lateral glide of the dislocation like growth steps — giving the observed shape change and a true *atomic correspondence for the substitutional atoms.* However, this view though widely is not fully accepted, see for example AARONSON *et al.* [1994].

An interesting question that appears not to have been addressed in the debate on the displacive or reconstructive mechanism of bainite is the role of *austenite grain size* in helping or hindering the reaction. For most reconstructive reactions, a finer grain size promotes the reaction by providing heterogeneous nucleation sites, however for martensite reactions a fine grain size is known to hinder the reaction, see for example INOKUTI and CANTOR [1982]. A fine grain size is known to degrade hardenability but it is not clear if this is due solely to the known acceleration of the high temperature reconstructive transformations giving allotriomorphic ferrite and pearlite or if a *fine grain size can also accelerate the formation of bainite.* Experiments to answer this question would be of considerable interest.

ii) *Carbon supersaturation in growing bainite plates*

The second dispute concerns the carbon supersaturation in bainite. HILLERT [1975, 1994] has demonstrated, using the Trivedi analysis for the linear lengthening of both Widmanstätten ferrite plates and bainite sheaves that the lengthening reaction is rather well described by near carbon diffusion control local equilibrium. Figure 55 gives, as the solid line, the predicted kinetics based on a simplified model of carbon diffusion-controlled lengthening. The predictions match the experimental points rather well despite the uncertainties in much of the data used. However, the agreement is based on a rather low value of the interfacial energy — only 0.06 J/m^{-2}. A higher value would lower the theoretical line and would certainly increase the small deviations seen for some alloys especially at the *lowest transformation temperatures* of 650 K. The result of carbon diffusion-controlled lengthening appears to be fully accepted for Widmanstätten ferrite, seen at the higher temperatures in fig. 54, but this result is not universally accepted for bainitic ferrite, for example by BHADESHIA and CHRISTIAN [1990] or by OLSON *et al.*

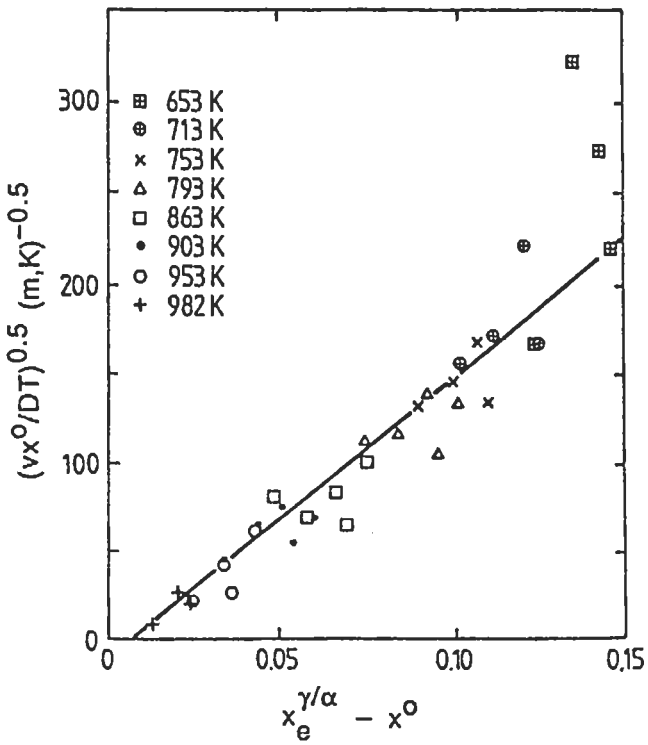


Fig. 55. Experimental lengthening rates of Widmanstätten and bainitic α-ferrite in iron–carbon alloys as a function of the carbo supersaturation at different temperatures. The line is the rate predicated for carbon diffusion control that is, local carbon equilibrium at the α/β interface. D is the diffusion coefficient of carbon in austenite, v is the growth velocity, T the absolute temperature x^o the carbon content of the alloy and $x_e^{\gamma/\alpha}$ the equilibrium carbon content of γ in equilibrium with α. After HILLERT [1994].

[1989, 1990].*

HILLERT [1994, 1995] has pointed out that given the known growth rate of bainitic ferrite combined with the high mobility of carbon at the temperatures of bainitic growth it is difficult to understand how there could be a significant difference of carbon activity across a slowly moving ferrite austenite interface. Local equilibrium for carbon across the ferrite and austenite interface does not, of course, mean that *both* phases may not be supersaturated with carbon with respect to the precipitation of carbide. Bainite formation typically occurs in a temperature well below that which at pearlite forms and for pearlite formation both austenite and ferrite are supersaturated with carbon, so this will be

* The bainite reaction in Fe–C alloy steels is known as one in which there appears to be no partitioning of the substitutional, slow-diffusing solutes, (Cr, Ni, Mo, Si, etc.), see for example BHADESHIA and CHRISTIAN [1990] and STARK *et al.* [1990]. The lack of measurable partitioning of the substitutional solutes during the growth of bainite is generally accepted.

certainly true for bainitic ferrite. The precipitation of carbide either at the austenite ferrite interface (upper bainite) or partially within ferrite for lower bainite is an essential feature of most but not all bainite reactions.* Precipitation of carbides during bainite growth has the important effect of allowing bainite plates to grow in "sheaths" of numerous adjacent plates. Growth of multiple plates in contact is not possible for Widmanstätten

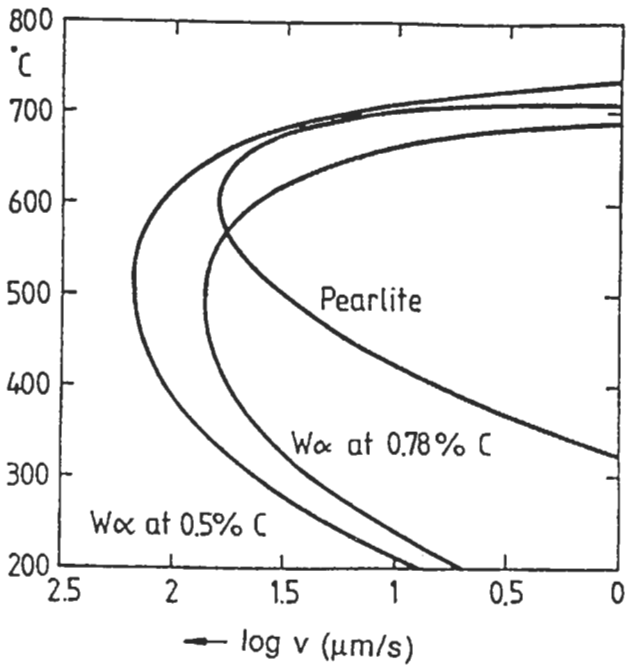


Fig. 56. Experimental lengthening rates of Widmanstätten α ferrite in two iron–carbon alloys (from fig. 55) and the experimental growth velocity of pearlite. After HILLERT [1994].

* HILLERT [1994] has shown using models based on the experimental growth rates for pearlite and diffusion limited linear lengthening rates for Widmanstätten and bainitic iron, that in plain carbon steels at temperatures below about 550°C that Widmanstätten ferrite lengthens faster than pearlite grows (fig. 56). His explanation of this faster low temperature growth of ferrite is in terms of the asymmetry of the extrapolated austenite solubilities below the eutectoid temperature. This explanation, however, seems rather incomplete given the differences between the theoretical kinetics for pearlite growth, § 2.5.1 (eq. 70), and that for linear lengthening of Widmanstätten ferrite, eq. 64 § 2.2.6.2. The pearlite grows with a high energy ferrite/austenite interface with no orientation relationship, while the Widmanstätten ferrite grows with a good fit, *low energy*, interface. The energy of the interfaces partially determines the required carbon diffusion distances in both reactions and this may be a dominating effect at temperatures below the nose of the TTT diagram where the activation energy of diffusion dominates the kinetics. An increased driving force would be the dominant effect above the nose in the curve since above the nose, an lowering of transformation temperature increases the driving force and accelerates the reaction. However, as noted by HILLERT [1994], the asymmetry which makes austenite at lower temperatures more supersaturated with respect to ferrite than to cementite is an *additional* reason why ferrite plates can grow faster than pearlite at low temperatures.

References: p. 1497.

ferrite since here the excess carbon rejected by one plate will, by decreasing the supersaturation of iron, fig. 20a, prevent growth of an adjacent plate in the carbon enriched austenite next to the first plate. ENOMOTO [1994] has, by computer simulation, modelled the growth of ferrite plates as two adjacent ledges of height equal to a ledge height and again shown that plates of bainitic and Widmanstätten ferrite lengthen at rates controlled by carbon diffusion control. His numerical results support Hillert's calculations.

In contrast to this picture of apparent diffusion-controlled growth, OLSON *et al.* [1989, 1990] have modelled the bainite transformation as a displacive reaction. Their rather elegant models were for the velocity of a martensite-like plate growing with *partial supersaturation of carbon* (the interface reaction) matching the growth limited by removal of the additional carbon by the standard Ivantsov analysis, eq. 63. The essential feature of Olson *et al.*'s model is that carbon diffusion, away from the lengthening plate, reduces the barrier for the subsequent martensitic growth. Low carbon austenite can grow martensite at a higher temperature than can higher carbon austenite. The theory for the martensite reaction mobility was obtained from an earlier dislocation model of martensitic growth by GRUJICIC *et al.* [1985]. In the 1989 paper, the usual arbitrary, maximum velocity, assumption was needed to provide the determining criterion for growth, § 2.2.6.2. In the later paper, however, the need for this additional criterion was avoided by use of a model for the velocity-modified change of *carbon partitioning* developed by AZIZ [1982] for rapid solidification. The results of these calculations are shown in fig. 57. Figure 57a gives the calculated residual supersaturation in the bainitic ferrite in Fe–0.4wt% C for both plate nucleation and growth. Fully supersatured *nucleation* leads to martensite so the M_s temperature is predicted to be close to 400°C, compare with fig. 49 for an alloy steel with 0.4wt% C. At higher temperatures, carbon depletion was needed for nucleation but fully supersatured *bainite* growth was predicted at the tip of the growing plates. At even higher temperatures, there was significant residual supersaturation even for growth and this supersaturation approaches zero, local equilibrium for carbon, only at about 550°C. At this stage, however, growth of bainite ceases, on the Olson *et al.* model, at the bainite start temperature. The calculated TTT curve for bainite with a B_s temperature of about 550°C is shown in fig. 57b. The conflict between the two calculated TTT diagrams for plane carbon steel in figs. 56 and 57 nicely illustrates the alternative models for the transformation. The obvious experimental solution to this conflict, is the experimental determination of *bainite* growth rates in plain carbon steel at 500 to 600°C. Unfortunately this growth is hidden from experimental examination by the competing pearlite reaction. In alloy steels however, where the pearlite reaction has been slowed down, there is indeed the well established "bay" in the TTT diagrams, see fig. 49, so here comparison is possible. This bay would appear to provide solid evidence in support of the displacive mechanism for bainite growth as predicted by the model of OLSON *et al.* [1990]. However competing mechanisms giving rise to the known inhibited decomposition of austenite in low alloy steels, below the allotriomorphic ferrite and pearlite reaction temperatures and above those of bainite is then the subject of the third dispute in this field.

iii) *Mechanism of the incomplete reaction in Fe–C bainites.*

In many alloy steels there appears to be a well established upper temperature for the

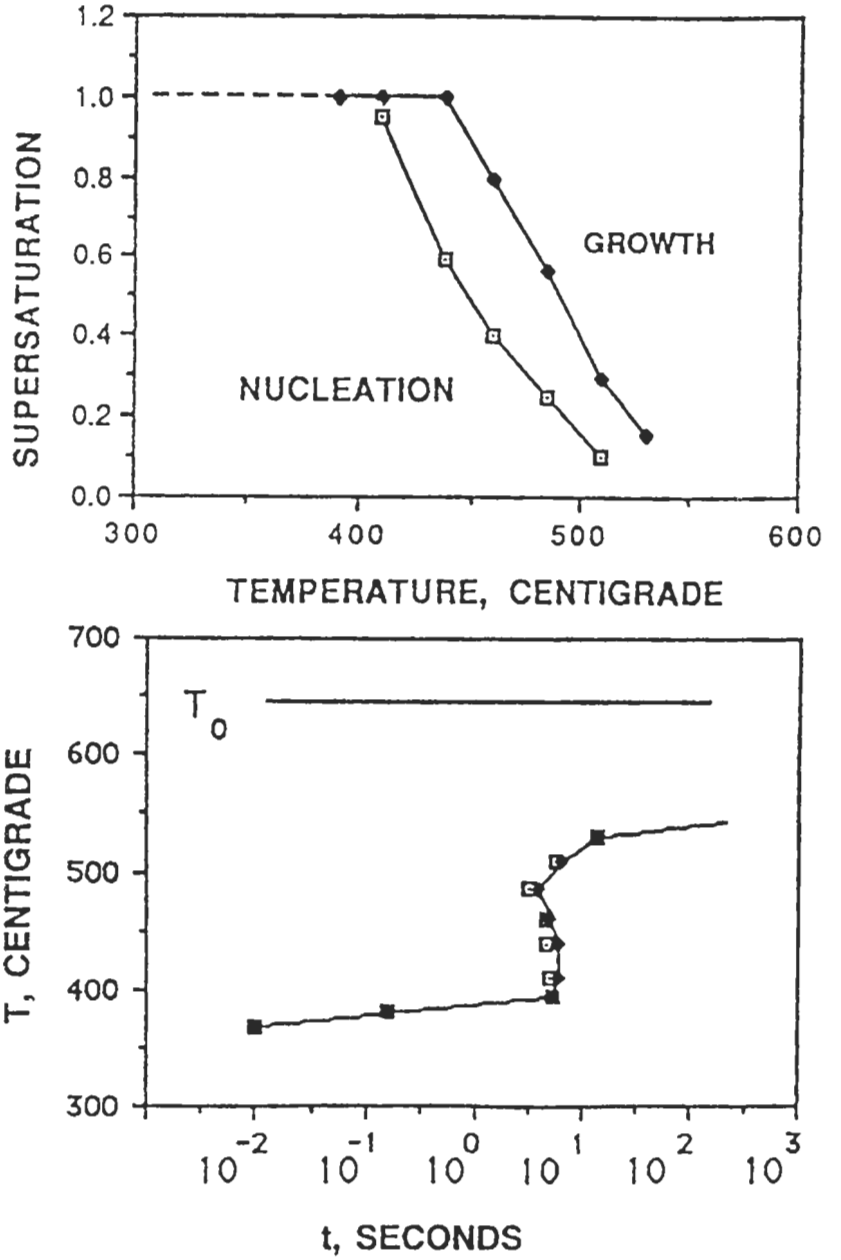


Fig. 57. (a) Plot of the calculated normalized supersaturation (x^{α}/x^{O}) of the carbon in ferrite as a function of the transformation temperature for an iron-0.4wt% carbon alloy for the situations of initial nucleation and for linear growth on a model for displacive interface mobility, carbon diffusion and solute trapping. (b) Calculated isothermal time temperature transformation diagram for the iron-0.4wt% carbon alloy filled points. The open points are from the authors' earlier maximum velocity analysis. After OLSON *et al.* [1990].

References: p. 1497.

formation of bainite, the bainite start temperature, B_s, and below this temperature although the reaction starts, it is commonly reported as ending with only a *partial transformation*, see fig. 49. At B_s, the fraction untransformed tends to 100%. This was, for example, found experimentally by BHADESHIA and EDMONDS [1979] who noted, in an alloy of Fe–0.4% C–3% Mn and 2% Si, that at 450°C, 5% of the structure transformed to bainite formed after only 100 s, but there was little further transformation of austenite until, after 32 days, the remaining austenite then decomposed to *pearlite*. This type of behaviour appears to arise rather naturally from the models of OLSON *et al.* [1990] as seen in fig. 57b. However there is an alternative explanation offered by proponents of the reconstructive viewpoint. This explanation is that the metallic alloying elements added to Fe–C to improve "hardenability", manage to inhibit bainite growth by a "solute drag" mechanism. This has been argued recently for example by AARONSON *et al.* [1990], AARONSON and HALL [1990], who cite the earlier references, and by HILLERT [1994, 1995]. Solute drag is well established for its role in inhibiting grain boundary motion in recrystallization, where very large reductions in grain boundary mobility are found for additions of a few parts per million of substitutional impurities (ch. 28, § 3.4.1). The drag is proposed to occur both directly, as in recrystallization, and in an additional way, in the decomposition of austenite, by changing the activity coefficient of the interstitial solute. However it is not immediately obvious how it can cause the apparent complete stasis of the bainite reaction while, at a higher temperatures the mobility inhibiting (hardenability) effects of the solute merely slow down but do not halt the growth of ferrite and pearlite. It is also not clear how solute drag can allow a certain fraction of bainite to form quickly but then to stop further bainite nucleation and growth. In addition, it might be noted that current models of solution hardening by mobile solute atoms, such as Mg in Al, see for example SCHMIDT and MILLER [1982] suggest that solute drag effects are responsible for inhibiting the movement of dislocations and thus promoting solution hardening and in particular inhibiting dynamic recovery during working and thus promoting strain hardening. If solute can inhibit dislocation migration under the influence of mechanical stress, solute does seem capable of exerting equivalent solute drag on at least slowing moving *displacive* interfaces. Martensitic interfaces may move too quickly for any significant solute drag effects which might account for the lack of significant solute hardenability effects of the formation of lath martensite in steel, beyond the solute effects associated with the thermodynamic stabilization of austenite by, for example C, Ni and Mn. All current models of solute drag can *slow* interface movement but they have not been shown to *halt* the movement.

Bainite is clearly still a topic in which there is no universally accepted qualitative model for the basic mechanism of its formation. For the reactions described as bainitic, the evidence of surface relief is, for most students of the subject, critical evidence for some type of displacive interface reaction. For the group of investigators associated with Aaronson, a model of the surface relief based on glide of interface dislocations associated with growth ledges is argued as being, in some way, different from a true displacive reaction. It is agreed, for at least some reactions showing surface relief, the substitutional bainites, Cu–Zn, Ag–Cu, Cr–Ni etc., Widmanstätten ferrite, high-temperature bainite and in particular Ag_2Al from Al, that surface relief is combined with a composition change

so that there is at least some reactions that do have these two characteristics, and which are currently describable as "bainitic". What is clearly not yet established is (i) if the dislocation glide causing the surface relief is the only atomic processes needed for the structural transformation and (ii) what is the level of solute supersaturation in many of the reactions, with the largest discrepancy for the lower bainite in Fe–C forming at temperatures close to the martensite start temperature. There is also disagreement on the origin of the incomplete bainite transformations at high temperatures.

Unusually for a topic in metallic solid-state phase transformations, the problem is here a lack of agreed theory rather than a shortage of quantitative data. It might be noted that there appears to be no disagreement that there can be:

(i) Fully reconstructive reactions with both a change of composition and a civilian change of structure, for example growth of grain boundary ferrite or pearlite growing behind interfaces with no lattice coherency.

(ii) Fully martensitic reactions with no change of composition (i.e., partitionless) and with a military or displacive change of structure. Examples are the lath and plate martensite reactions in Fe–C and Fe–Ni alloys.

(iii) The so-called "massive" reactions. These are the formation of a different structure or orientation by a reconstructive interface process with no change of composition, that is they are as in (ii) partitionless. Examples of the latter include the massive reaction in iron and many other metallic systems, described by MASSALSKI [1970]. In addition of course are the recrystallization and grain growth involving grain boundary migration in non supersaturated alloys.

There appears to be no reason why the final combination cannot be found. This is:

(iv) A diffusional change of composition with a displacive lattice transformation. This is agreed to occur in the Widmanstätten precipitation of Ag_2Al in AlAg.

The question that remains to be settled is: Does bainite and or Widmanstätten ferrite growth actually occur in this manner? There is an related and rather interesting, unanswered, question: Do all plate-like products, growing coherently in a crystalline matrix, have a displacive interface mechanism?

Although it is not fully accepted that bainite does form by a displacive lattice transformation, in the absence of any satisfactory alternative explanation of the surface relief seen in bainite the name "bainitic" for a displacive transformation occurring with some change of chemical composition seems currently satisfactory. It would therefore be appropriate to describe the precipitation of silver-enriched cph γ from fcc α in Al–Ag, where the interface reaction is established to be displacive, as a further example of a bainitic reaction. If in the future, a non displacive origin was ever established for the surface shear in the bainite reaction in steel, then a new name would be needed for the reaction identified in Al–Ag as occurring by a combination of a change of composition and a martensitic, fully displacive, interface process. There seems, however, no need to invent a new name, at least at present. That is, in the terms used by AARONSON *et al.* [1990, 1994], though clearly not with their approval, this reviewer has adopted the "surface relief" definition of bainite for reactions involving a change of composition as well as surface relief.

References: p. 1497.

3. *Continuous transformations*

Although the vast majority of phase transformations in physical metallurgy are of Gibbs type I-involving nucleation and growth, Gibbs second mode of transformation, type II, now usually called *continuous transformations*, are important and interesting diffusive transformations. A continuous transformation can only occur if the initial and final structures share either a common crystal lattice or, as in liquids and glasses, a lack of a lattice.

The initial example of a continuous transformation was *spinodal decomposition*, the development of a composition fluctuation on a scale larger than that of the interatomic spacing. This transforms a phase, α, with an initially uniform solute composition into two phases of the same lattice type, α' and α'', one solute depleted, the other enriched. The reaction is solely a *diffusion* reaction involving atoms exchanging lattice sites. Although binary equilibrium diagrams indicate only a few potential examples of *equilibrium spinodal decomposition* in metals, there are numerous metastable reactions that could occur by the continuous mechanism (see CAHN [1968] and SOFFA and LAUGHLIN [1983]). A second example is the process of *continuous ordering*, see COOK *et al.* [1969] and DE FONTAINE [1975]. This consists of a composition modulation but with a wavelength of the interatomic spacing, transforming a disordered solid solution into an ordered one.

3.1. Spinodal decomposition

Two excellent early reviews of this subject were provided by CAHN [1968] and HILLIARD [1970]; The reviews gave clear accounts of the basic theory and they reviewed the relevant early experimental observations. The following will therefore give an outline of the initial theory, following CAHN [1968], and will then try to bring out the important advances made subsequently. A later review of the topic was provided by SOFFA and LAUGHLIN [1983] who considered the not uncommon reaction in which an ordered coherent *precipitate* forms from a disordered solid solution.

In a region of a single phase where the curvature of the free-energy composition plot, d^2F/dC^2, is negative (fig. 4) *uphill* diffusion can occur, as this lowers the free energy. In the literature on spinodal decomposition, the discussion is usually in terms of this second differential, while in the literature on diffusion, the discussion is usually based on the "Darken term", $1 + d\ln f_\alpha / d\ln C_\alpha$, derived by DARKEN [1948] in his analysis of diffusion, ch. 7, § 5.3.1.2). If D_A^* and D_B^* are the tracer diffusion coefficients of the two components of atomic fractions C_A and C_B, then the interdiffusion coefficient, $\tilde{D}$, is given by:

$$\tilde{D} = \left(C_A D_B^* + C_B D_A^*\right) \frac{1 + d\ln f_\alpha}{d\ln C_\alpha} \tag{77}$$

It is readily shown, using standard solution thermodynamics, e.g., by MARTIN and DOHERTY [1976], that:

$$\frac{d^2F}{dC^2} = RT\left(\frac{1}{1 - C_B} + \frac{1}{C_B}\right)\left(\frac{1 + d\ln f_\alpha}{d\ln C_\alpha}\right) \tag{78}$$

From here onwards, the atomic fraction C will stand, as elsewhere in this chapter, for the atomic fraction of the second component, B. So $C_A = 1 - C_B = 1 - C$.

From eqs. (77) and (78), we obtain:

$$\tilde{D} = (1 - C)D_B^* + CD_A^* \frac{C(1 - C)}{RT} \frac{d^2F}{dC^2} \tag{79}$$

$$\tilde{D} = M_D \frac{d^2F}{dC^2} \tag{79a}$$

Equations (79) and (79a) define the diffusional mobility, M_D.*

As a consequence of the sign of D being determined by the sign of d^2F/dC^2, a homogeneous solid solution with $d^2F/dC^2 < 0$, is *unstable* in the presence of an initial infinitesimal fluctuation (fig. 5). So any such sinusoidal perturbations of initially minimal amplitude can grow in amplitude. The crucial problem concerns the *wavelength* of the perturbations that can grow. To discuss this, the thermodynamics of inhomogeneous solid solutions are needed (CAHN and HILLIARD [1958]). If there were no restrictions as to a minimum wavelength, then the growth of short wavelengths, requiring diffusion over the shortest distance, would grow fastest, leading to the microstructure with the shortest possible wave length, an interatomic spacing. Unfortunately this yields an ordered *solid solution*. However spinodal decompositions only occur in systems that wish to *phase-separate*: They need to *maximize* the number of like neighbour bonds. An ordered solid solution however *minimizes* the number of like neighbour bonds. Some additional thermodynamic feature is needed to avoid this failure of an over-simple model.

For an homogeneous AB alloy in which we exchange a mole of A for a mole of B, the change of free energy, ΔF, is:

$$\Delta F = \frac{\mu_B - \mu_A}{V_m} = \frac{dF^v}{dC} \tag{82}$$

Here μ_A is the chemical potential of A, $\mu_A = \{dF/dn_A)_{T, p, C_A, C_B}$, where n_A is the number

* It has been recently pointed out, MARTIN *et al.* [1995] that there is a simple physical explanation for the need to use the interdiffusion coefficient, $\tilde{D}$, rather than D_B, where $D_B = D_B^* (1 + d \ln f\alpha / d \ln C_\alpha)$, in the second law of diffusion, eq. 80.

$$(\delta C'/\delta t)_x = D(\delta^2 C'/\delta x^2)_t \tag{80}$$

C' is the volume concentration of B atoms, ($C' = C/V_a$ where V_a is the volume per atom) so the use of D_B might have been expected. In the presence of a Kirkendall effect, where $D_A^* \neq D_B^*$, and where as a result there is also a change of volume, the concentration of B atoms in a small region, $(\delta C/\delta t)_x$, is affected not only by the flux of B in and out of the region but also by the change of volume of the region. Rederiving the second law in the standard way but in the presence of a Kirkendall effect, yields either eq. 80, DARKEN [1948], or a modified second law, eq. 80a

$$(\delta C'/\delta t)_x = D_B(\delta^2 C'/\delta x^2)_t + C'(\delta v/\delta x)_t \tag{80a}$$

Here v is the velocity of interface planes, as measured by the velocity of inert Kirkendall markers. The velocity v is given by DARKEN [1948] as:

$$v = V_a(D_B - D_B(dC'/dx) = (D_B - D_B)(dC/dx). \tag{81}$$

References: p. 1497.

of moles of A, μ_B is the chemical potential of B and F^v is the free energy of unit volume of the homogeneous solution. In this simple treatment it is assumed that the atoms are of equal size; a correction to account for the change of lattice parameter with composition in a real system is introduced later.

The diffusive flux, J, is then given by the usual equation, the first law of diffusion:

$$J = -\tilde{D}\nabla C' \tag{83}$$

with $\tilde{D}$ given by eq. (79), and C' is C/V_m, the concentration expressed in atoms of B per unit volume.

From the thermodynamics of inhomogeneous solid solutions, a correction for inhomogeneity is introduced as:

$$\left(\frac{dF^v}{dC}\right)_{inh} = \frac{dF^v}{dC} - \frac{2K\nabla^2 C}{V_m} \tag{84}$$

K is a positive (for systems that wish to un-mix) *gradient-energy coefficient* which is determined by the difference in number of like atomic neighbours between an atom in a homogeneous alloy and an atom in an alloy which has a variation in composition. CAHN [1968] pointed out that the first differential, ∇C, will have no effect on the free energy since with a linear gradient, ∇C, any excess of one component in one direction will be balanced by an equal depletion of that component in the opposite direction. The first term in the series to produce an effect is $\nabla^2 C$. The value of K is given by:

$$K \approx N^v k T_c \psi^2, \tag{85}$$

where T_c is the critical temperature, below which homogeneous alloys will wish to *unmix* into A-rich and B-rich regions, and ψ is the chemical "interaction distance" of atoms (CAHN [1968]). In regions where $\nabla^2 C > 0$ the effect of K is to lower the free energy of B atoms since they sense a *higher* density of like atoms than they would do in a homogeneous solid solution. Where $\nabla^2 C < 0$ the effect of K is to raise the free energy of B atoms since they sense a *lower* density of like atoms than they would do in a homogeneous solid solution.

Comparison of a solution with a sinusoidal variation of composition about an average composition C_0 with the same alloy with a uniform composition C_0 shows, fig. 5, that the microstructure with the composition modulation has a *higher free energy* with $K > 0$. This occurs since when $C > C_0$ $\nabla^2 C < 0$, and with $C < C_0$, $\nabla^2 C > 0$, so there are *more atoms having their free energy increased* than there are atoms having their free energy decreased. The total free energy increase, due to the gradient energy term, with $K > 0$, increases both with an increase of amplitude and with a *decrease in wavelength*. Both of these changes increase the magnitude of the composition curvature, $\nabla^2 C$. This extra K term is the new thermodynamic parameter needed in the analysis of continuous transformations in general and spinodal decomposition in particular. Physically, it acts in a similar way as does the interfacial energy in nucleation and growth reactions.

Substitution of eqs. (79a) and (84) into the diffusion equation, eq. (83) gives:

$$J = \frac{-M_D\, d^2F}{dC^2 \nabla C'} + 2M_D K \nabla^3 C'. \tag{86}$$

The change of composition with time, dC/dt, is obtained in the normal way for the derivation of the second law of diffusion (e.g., SHEWMON [1963]), as:

$$\frac{dC}{dt} = \frac{V_m dC'}{dt} = \frac{M_D\, d^2F}{dC^2 \nabla^2 C} - 2M_D K \nabla^4 C. \tag{87}$$

CAHN [1961] showed that this differential equation has the following solution.

$$C = C_0 + \exp\big(R(\beta)\, t\big) \cos(\beta r). \text{ const.}, \tag{88}$$

with C_o = initial composition; β = wave number, $2\pi/\lambda_w$; λ_w = wavelength of the particular fluctuation. The "amplification" factor $R(\beta)$ is given by:

$$R(\beta) = -M_D\, \beta^2 \left(\frac{d^2F}{dC^2} + 2K\beta^2 \right). \tag{89}$$

Since the diffusional mobility, M_D, is inherently positive, it can be seen, from the term in brackets in eq. (89), that in a system showing unmixing, when K is positive, short-wavelength fluctuations will decay since $K\beta^2 > -d^2F/dC^2$ for large values of β, but fluctuations below a critical wave number, β^*, that is with wavelengths longer than a critical wavelength, λ^*, can grow:

$$\beta^* = \left(-\frac{d^2F}{dC^2} \frac{1}{2K} \right)^{1/2}. \tag{90}$$

The corresponding critical wavelength, λ^*, is $2\pi/\beta^*$.

The fastest growing wavelength, λ_{max}, is determined from eqs. (88) and (89) as the wavelength having the maximum amplification factor, $R(\beta)$. This occurs at $\beta^*/\sqrt{2}$, owing to the double effect of an increasing wave number in (i) reducing the *diffusion distance*, the term outside the bracket in eq. (89), as well as its effect in reducing the *driving force*, the terms inside the brackets. This is a similar result to that previously seen in other areas of phase transformation by the balance between interfacial energy requiring large spacings and rapid diffusion requiring short ones (§ 2.2.6 and § 2.5). In the present case, all wavelengths will be found to be present, but with initially very small amplitudes. These arbitrary amplitudes provide the undefined "const" in eq. (88). This result comes from the mathematical idea that any perturbation can be represented as a sum of sine or cosine waves. In addition, at very small amplitudes, all wavelengths are able to grow *independently*, so the fastest-growing wavelength will be present and will be able to grow faster than all other wavelengths and so must dominate the decomposition at least initially when the theory, given above, applies.

An important addition to the theory comes from considerations of elastic strains that arise when there is a change of lattice parameter with a change of alloy content. With η defined as the unit strain per unit composition difference, $\eta = (da/dC)/a = (d \ln a/dC)$, Y

is Young's Modulus and ν_p is Poisson's ratio, eq. (89) is changed by the elastic strain to:

$$R(\beta) = -M_D\beta^2\left(\frac{d^2F}{dC^2} + \frac{2\eta Y}{1-\nu_p} + 2K\beta^2\right). \tag{89a}$$

The strain term, in eq. (84a), acts, in addition to the gradient-energy term, to *inhibit* the reaction. It is often more convenient, however, to consider the second term together with the first term, as defining the *coherent spinodal* region. This is a region, in the phase diagram, in which the sum of the first two terms in the parentheses of eq. (89a) is negative. Only within this coherent spinodal region can fluctuations develop while the crystal remains fully coherent, but elastically distorted, as the solute-rich and solute-poor regions, with different values of the unstrained lattice parameter, a, interacting with each other within an elastically distorted lattice.

RUNDMAN and HILLIARD [1967] tested the model for spinodal decomposition by small-angle X-ray scattering experiments with Al–Zn alloys. Al–Zn has a phase diagram very similar to that shown in fig. 3. Their results, fig. 58, show the behaviour expected for spinodal decomposition. The alloy, Al–22at% Zn, had been quenched from the single-phase, all fcc, region and annealed at 65°C for the times indicated. The critical wave number, β^*, where there is no change of scattered intensity with longer aging time is seen, as the"cross-over" point. Also seen is the maximum rate of increase of intensity at a wave number about 0.7 of β^*. The interpretation of small-angle diffraction is discussed in ch. 12, § 5.1. Other experimental demonstration of the importance of the gradient term in solid solutions with short wavelength composition fluctuations were diffusion experiments in stable solid solutions, made inhomogeneous at a very short wavelengths by repeated depositions of alloys of different compositions, see in particular COOK and HILLIARD [1969] and PHILOFSKY and HILLIARD [1969]. Other early experimental studies

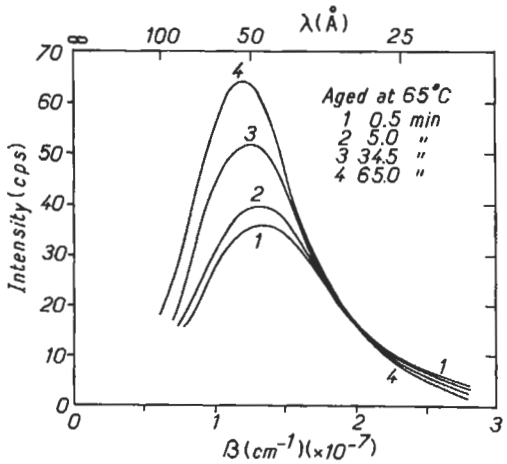


Fig. 58. Small-angle X-ray spectra for an Al–22 at% Zn alloy quenched from 425°C and annealed at 65°C for the times indicated (after RUNDMAN and HILLIARD [1967]).

that supported the linear spinodal model described by CAHN [1968] were reviewed by HILLIARD [1970].

Following the initial development of Cahn's theory of spinodal decomposition, there have been various modifications to the theory to try to deal with later stages in the reaction when the linear model of independent growth of all fluctuations is no longer valid. A further important modification was consideration of the influence of *random thermal fluctuations*, "Brownian Motion", on the process. Random thermal fluctuations will give local increases of energy paid for by the increase of entropy that comes from the "disorder" introduced by the small groups of atoms having the increased energy. In nucleation theory this process gives the vital eq. (11) for the probability of creating a nucleus with a *local* increase of energy Δf^*. COOK [1970] introduced this idea of random solute movements into a treatment of spinodal decomposition and he showed that the effects seen in experiments on small-angle scattering could be fitted to the changes proposed by his model. The effect of the thermal fluctuations is essentially that for alloys close to the spinodal boundary the distinction between spinodal decomposition and the nucleation and growth of coherent "zones" becomes much less clear-cut than in the original model of spinodal decomposition.

LANGER [1975] discussed these ideas in more detail and showed that the initial model of spinodal decomposition, as described above, predicts that at the spinodal point, when $d^2F/dC^2 = 0$, the critical wavelength becomes infinite and the only mechanism for transformation will then be by the nucleation of solute-rich "Guinier–Preston zones". However, nucleation of such zones will occur frequently with very small critical radii and with the very low-energy, "diffuse" interface expected in these circumstances*. In other words, nucleation will result in a structure very much like that of spinodal decomposition. The same picture is given in recent models of spinodal decomposition modifed by thermal fluctuations. Figure 59a shows the structure factor S computed by LANGER *et al.* [1975] as a function of q, a temperature-modified wave number and τ, a modified annealing time, for an alloy held at the spinodal boundary. It can be seen that a spinodal-like fluctuation builds up even with $d^2F/dC^2 = 0$, but the maximum intensity shifts to larger wavelengths with increased reaction times.

The other modification to the theory deals with the later time development of the fluctuations. The results, also computed by LANGER *et al.* [1975] for the expected development of the structure factor for an alloy at the centre of the spinodal region are shown in fig. 59b; $q = 1$ corresponds to the critical wave number β^*. The value q_p is $\beta^*/\sqrt{2}$, the fastest growing fluctuation in the linear theory. It can be seen that for short times, this is where the maximum growth of the intensity actually appears. However, at longer times, the peak in the structure factor, the wavelength with the strongest intensity

* Simple application of the "quasi-chemical bond" model to a coherent interface between solute-rich and solute-poor zones shows that the interfacial free energy, σ, is reduced at relatively high temperatures by making the interface extend over several atomic layers at some increase in energy but with a compensating increase in entropy CAHN and HILLIARD [1958]. Atom-probe FIM has shown, however, that at temperatures much lower than the critical temperature, T_c, where the solid solution becomes unstable, the coherent interface (in Cu–1.9 at% Ti) was atomically sharp (VON ALVENSLEBEN [1982]).

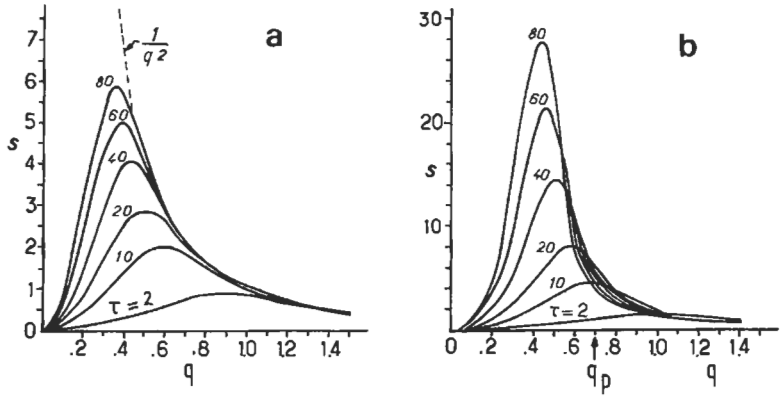


Fig. 59. Computed structure factor as a function of modified wavenumber for increasing times, τ, for (a) an alloy at the edge of the spinodal region and (b) an alloy at the centre of the spinodal region. (After LANGER *et al.* [1975].)

moves to smaller wave numbers, larger wavelengths. In addition, the peak intensity does not continue to grow *exponentially* with time. These changes appear to indicate the movement towards the LSW-type *coarsening*, with larger wavelengths growing at the expense of shorter wavelengths. Within spinodal decomposition, a process that is equivalent to interfacial energy-induced coarsening appears to be occurring *during* the decomposition and is not restricted to the last stages of the reaction when the solute-rich and solute-poor regions have discrete interfaces and are close to the equilibrium form of α' and α''. Coarsening is expected with spinodal decomposition even before any well defined interfaces have formed. It arises directly from the gradient energy term, K. As described above, the increase of energy due to the K term in eq. 84, for a composition fluctuation of a given amplitude increases as the curvature rises due to shorter wavelengths and higher composition amplitudes. So competition between wavelengths will favour the development of longer wavelengths at the expense of shorter ones, with the same composition amplitude. This is a true *coarsening* process.

Experimental data on the later stages of spinodal decomposition are shown in fig. 60 for Al–22at% Zn (HILLIARD [1970]). The data here are for a higher temperature, 150°C, than in fig. 57 which was obtained at 65°C, so the diffusional unmixing is much more advanced than in the earlier study. It is seen in fig. 59 that the peak of intensity is shifted to much smaller wave numbers as the reaction proceeds; this is exactly the effect produced in the computer model of LANGER *et al.* [1975]. Assuming an activation energy for diffusion of Zn in Al of 120 kJ/g-mole gives the ratio of diffusion coefficients at the two temperatures, $D(150°C)/D(65°C)$, as 5000. The much higher intensities in fig. 51 than in fig. 49 can be seen, confirming that the reaction is, indeed, much further advanced.

TSAKALAKOS [1977] and TSAKALAKOS and HILLIARD [1980] provided some analytical insight into the later stages of spinodal decomposition, when the compositional amplitude of the fluctuation is no longer small, but begins to approach the difference in composition $\Delta C_{\alpha'\alpha''}$ between the solute-rich and solute-poor regions of the two phase systems

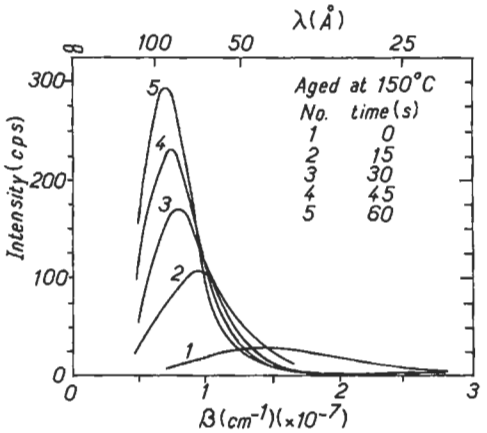


Fig. 60. Experimentally observed small-angle X-ray spectra for an Al–22 at% Zn alloy annealed at 150°C for the times indicated (after RUNDMAN *et al.* [1970]).

(fig. 3). The difficulty is readily seen from the free-energy-composition curve (fig. 4), since, as the composition fluctuation reaches the spinodal points, $d^2F/dC^2 = 0$, the driving force for further unmixing by spinodal decomposition then vanishes. It should be noted that at the spinode points where, $d^2F/dC^2 = 0$, solute transfer will occur easily across the "interface" since the chemical potential of the solute, the intercept of the tangent to the free energy curve is higher at the low composition spinode point than at the high-composition spinode point. DITCHEK and SCHWARTZ [1980] discussed the theory and extend it beyond the single wavelength considered by Tzakalakos, to a range of wavelengths. For a single wavelength the amplitude grows until it reaches a *critical* wave form which has a $\tanh(\beta r)$ function (CAHN and HILLIARD [1958, 1959]).

The system can then continue to lower its free energy by perturbing the wavenumber to smaller values; this will *reduce* the gradient energy which is opposing further decomposition. Ditchek and Schwartz's extension, to consideration of a range of wavelengths, allows the reaction to continue by the growth in amplitude of waves with smaller wavenumbers. In the application of the analysis to experimental results in spinodally decomposing alloys, several fitting parameters are required. These take into account the initial composition waves produced during the quench after the solution treatment in the single-phase region. This is a weakness in their test of the theory, but nevertheless the comparison of the experimental results and the theory, shown as the points and the solid lines, respectively, in fig. 61, is rather satisfactory. The results were obtained for a Cu–10.8at% Ni–3.2at% Sn alloy, solution-treated at 800°C, quenched to room temperature and then aged at 350°C for various times, before examination of the composition fluctuation; $\varepsilon = 0.015$ corresponds to the metastable equilibrium value of the composition difference of about 2.4at% Sn between the two phases; the modulation is almost entirely in the tin content. Figure 61 shows satisfactory agreement with the theory and, in addition, various features of the theory and the results can be seen. These features include the initial growth of amplitude at a fixed wavelength, $\lambda_w = 5$ nm (50 Å); the early

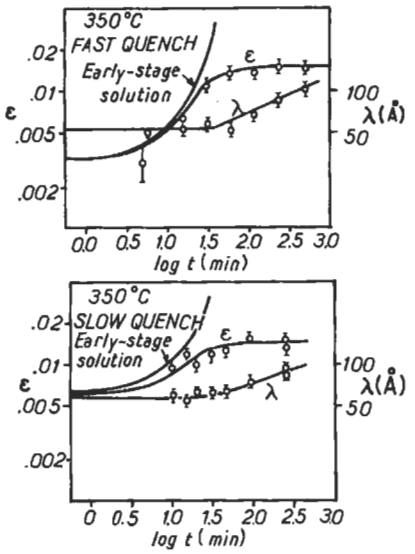


Fig. 61. Experimental results (points) and theoretical predictions (solid lines) for the amplitude, ε, and dominant wavelength, λ, for the spinodal decomposition of Cu–Ni–Sn after slow and fast quenches (after DITCHEK and SCHWARTZ [1980]).

departure from exponential growth of the amplitude, which occurs before the onset of the increase in the dominant wavelength. Finally it can be seen that the faster quench has given an initially smaller modulation in composition than does the slower quench. In the Cu–Ni–Sn alloys the tin-rich regions show an ordering reaction to give a DO_{22} superlattice at later stages of the decomposition.

In a high-resolution study using the field-ion microscopy/atom-probe technique, BIEHL and WAGNER [1982] were able to study the details of spinodal decomposition in an alloy of Cu–2.7at% Ti. They reported a steady rise in the titanium content of the Ti-rich regions as the reaction progressed at 350°C indicating a spinodal process. This rise in solute content, saturated at 20at% Ti, the composition of the ordered phase, Cu_4Ti. The dominant wavelength grew with time both during, and after, the time that the modulation was increasing in magnitude, with the dominant wavelength of the modulation increasing with reaction time to the $\frac{1}{4}$ power. The diameter of individual titanium-rich clusters also grew, in this case with time to the $\frac{1}{3}$ power, the time exponent expected for LSW coarsening both *during* and after the change in solute content.

It is striking that in the study of decomposition of Ni–14at% Al by WENDT [1981], previously discussed in § 2.2.6, it was clear that the reaction in the nickel alloy was much better described as a *nucleation-and-growth* reaction, rather than as a spinodal reaction. This was seen since the solute-rich regions, when first detected, had the equilibrium composition of Ni_3Al. Wendt had used the same techniques, FIM/AP, as had BIEHL [1980]. It seems surprising that an alloy whose composition, 14at% Al, is quite close to the composition of the final precipitate, should show a clear nucleation behaviour, while

the Cu–2.7at% Ti alloy that was much further from the composition of the ordered phase, Cu_4Ti, should show all the indications of spinodal decomposition. The difference may be due to there being a very different form of the free-energy-composition curve in Cu–Ti at 350°C than in Ni–Al at 550°C (their homologous temperatures, T/T_m, are 0.46 and 0.48, respectively); this is discussed further in § 3.2. VON ALVENSLEBEN [1982] also found by FIM/AP that a Cu–1.9at% Ti alloy decomposes by nucleation and growth. So the spinodal point appears to be be between 1.9 and 2.7at% Ti at 350°C. Subsequent reanalysis by VON ALVENSLEBEN and WAGNER [1984] showed, however, that while Cu–1.9at% Ti certainly decomposed by a nucleation, growth and coarsening mechanism, § 2.1.5., but it was less certain that decomposition of the 2.7% Ti alloy was by spinodal decomposition. SOFFA *et al.* [1984] described FIM/AP studies of spinodal decomposition in Fe–Cu at times where the metastable copper rich regions remained bcc and in all bcc Fe–Cr alloys.

The classic electron microscopy studies of spinodal decomposition were provided by BUTLER and THOMAS [1970] and by LIVAK and THOMAS [1971] using ternary Cu–Ni–Fe alloys. In the first investigation, a symmetrical composition (51.5at% Cu, 33.5at% Ni, 15at% Fe) near the centre of the spinodal region, was investigated. In the second study, an asymmetrical composition (32at% Cu, 45.5at% Ni, 15at% Fe) nearer to the edge of the spinodal region was used. In both studies, Curie-temperature measurements, which are very sensitive to the composition of the Fe–Ni rich phase, supplemented the microscopy. In the "as-quenched" alloys no sign of decomposition could be seen but, since the scattering factors of the components are similar in the Cu–Ni–Fe alloys, any initial perturbations will be very difficult to detect, unlike the case in the Cu–Ni–Sn alloy used by DITCHEK and SCHWARTZ [1980]. Butler and Thomas showed that the waves developed along <100>, which are the elastically softest directions. They also found that the two-phase structure was initially composed of rod-like particles with diffuse interfaces, but the "precipitates" developed planar interfaces with extended coarsening. At long times of ageing, the interfaces lost their coherency by the usual formation of a cross grid of edge dislocations.

BUTLER and THOMAS [1970] reported that the wavelength of the modulation coarsened, with a time to the power $\frac{1}{3}$ dependency, both during the later stages of the reaction and also earlier while the Curie temperature was varying. The change in the Curie temperature indicated a variation in composition of the copper-depleted regions, that is, spinodal decomposition was still occurring. For alloys aged at 625°C, these simultaneous changes can be clearly seen, fig. 62a, and these results confirm that the coarsening reaction is occurring during the initial decomposition as well as after the decomposition. In their study on the asymmetrical Cu–Ni–Fe alloys, LIVAK and THOMAS [1971] gave similar results, but there were some important distinctions. One of these distinctions was the slower development of the composition fluctuations; the change of Curie temperature continued for 100 h at 625°C for the asymmetrical alloy (fig. 62b) as distinct from only 1–5 h (fig. 62a) for the symmetrical alloy. This distinction is to be expected since the value of d^2F/dC^2 is *smaller* in alloys away from the symmetry point. In the asymmetrical alloy there was, in addition, no sign of growth in the dominant wavelength for nearly 10 h at 625°C (fig. 61b) so that coarsening appears also to develop

References: p. 1497.

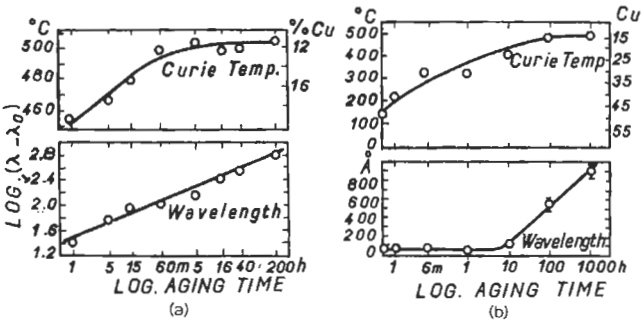


Fig. 62. Variation of the Curie temperature and the dominant wavelength with time on ageing at 625°C, for two Cu–Ni–Fe alloys: (a) the symmetrical alloy (after BUTLER and THOMAS [1970]); (b) the symmetrical alloy with 32 at% Cu (after LIVAK and THOMAS [1971]).

later in the asymmetrical alloy than in the symmetrical alloy.

In an important extension of the application of TEM to spinodal decomposition, SINCLAIR *et al.* [1976] and WU *et al.* [1978] used high resolution *lattice imaging techniques* to demonstrate the change in lattice parameter on a very fine scale, produced in spinodally decomposing alloys. Figure 63 shows the measured variation in lattice spacing in a Cu–29at% Ni, 3at% Cr alloy aged at 700°C for 10 min. The apparent wavelength of the fluctuation, 4.8 ± 0.8 nm, was very close to that obtained from electron diffraction, 5 ± 0.5 nm, in the same sample. In this high resolution study, it was clearly shown that, as would be expected, the interface was *diffuse* at the early stages of the reaction and became much sharper at the later stages; the lattice was then no longer continuous.

3.2. Continuous ordering

In the discussion of nucleation-and-growth reactions in § 2, no specific discussion was offered for the precipitation of *ordered phase* from a disordered matrix, since there seems no real distinction between growth of an ordered phase from a disordered matrix and the nucleation and growth of any other phase, in a Gibbs type-I transformation. For continuous reactions, however, continuous ordering is distinctly different from spinodal decomposition and will therefore be briefly considered here as separate reaction. Much of what follows is derived from the original paper by COOK *et al.* [1969] and the detailed review by SOFFA and LAUGHLIN [1983]; reference should also be made to reviews by DE FONTAINE [1975, 1981] and by COOK [1976] for more detailed accounts of the topic.

It has already been mentioned that in Cu–Ti and in Cu–Ni–Sn the end-result of spinodal decomposition can be the formation of an *ordered* structure in the solute rich regions of the microstructure. That is, spinodal decomposition of the solid solution is the first continuous reaction and then order later develops as the composition increases towards the ideal value for a particular long-range-ordered structure. However, continuous ordering can occur without change of composition if the alloy is close to the correct composition. This was analyzed by COOK *et al.* [1969].

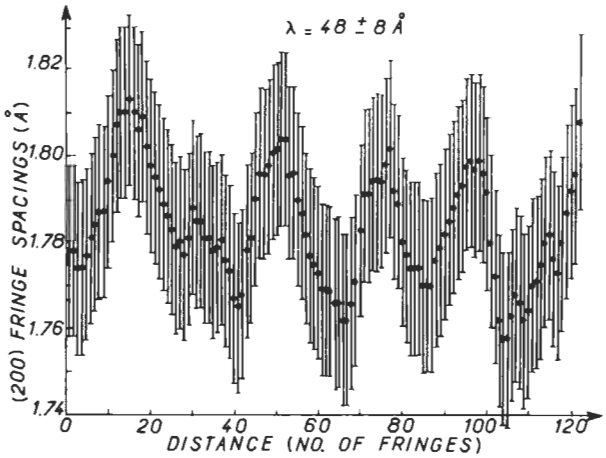


Fig. 63. The variation in fringe spacing from lattice images of a spinodally decomposed Cu–Ni–Cr alloy, aged at 700°C for 10 min (after Wu *et al.* [1978]).

The spinodal decomposition model discussed in § 3.1 was based on a *continuum* analysis (Cahn [1961, 1962]). However, Hillert [1956, 1961] had previously considered a *discrete lattice* model for decomposition, albeit in only one dimension. Hillert's approach was extended by Cook, de Fontaine and Hilliard [1969] to a three-dimensional, discrete lattice model. Their model considered not only spinodal decomposition, the unmixing reaction with a negative d^2F/dC^2 and a positive gradient energy parameter K, as discussed previously, but also the opposite case, a strongly *positive* d^2F/dC^2, and a *negative* K, a situation expected for a system that can go, without a change in composition, from a disordered high-temperature structure to an *ordered* low-temperature structure, as the temperature falls below a critical temperature T_c.

The results of the discrete lattice calculation are very similar to the continuum model in the spinodal case. Figure 64 shows the amplification factor, here called α, as a function of a/λ, where a is the simple cubic lattice parameter, for three situations. Figure 64a is for a spinodal decomposition, that is, with negative d^2F/dC^2 and positive K. The critical wavelength, λ_c, and the fastest-growing wavelength, λ_m, are shown. Figure 64b shows the amplification factor for a *continuous ordering reaction*, with positive d^2F/dC^2 and negative K, where the maximum amplification occurs at $a/\lambda = 1$, that is, the maximum and minimum in composition occur on *adjacent lattice sites*, giving an ordered structure. Figure 64c shows the situation with a negative amplification factor at all values of a/λ, which will occur above the critical instability temperature T_{instab}. This temperature is predicted by this treatment to occur when:

$$\frac{d^2F}{dC^2} = \frac{-32K}{a^2}. \tag{91}$$

In a real situation, in any one region of the lattice, perturbations of different "phase" will conflict with each other, but eventually one perturbation will dominate the rest and

References: p. 1497.

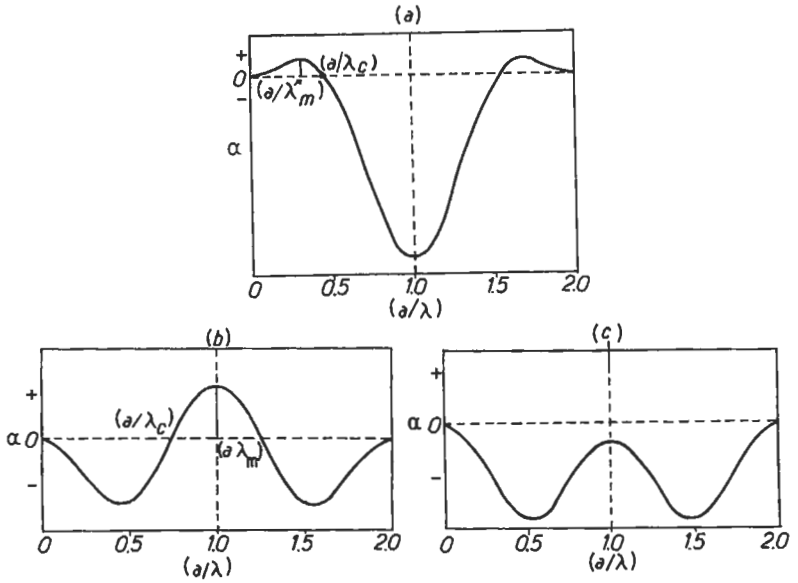


Fig. 64. The amplification factor, α, for the discrete lattice model of continuous reactions: (a) for spinodal decomposition; (b) for ordering below the instability temperature; (c) above the instability temperature. a is the lattice parameter of the simple cubic structure. (After COOK *et al.* [1969].)

establish a particular *domain* of order in that region. The order in each domain will then steadily increase. Different regions will be dominated by waves with differing *phases* of modulation, giving the multiple domain structure of the ordered alloy. (For a discussion on domain structure see ch. 4, § 4.2) COOK *et al.* [1969] extended the analysis to bcc and fcc lattices. As in the original continuum model of spinodal decomposition, the analysis of COOK *et al.* [1969] is restricted to small perturbations where d^2F/dC^2 and K remain constant, so again only the early stages of the reaction are described by this theory. Extensions to the initial discussion on continuous ordering have been provided by DE FONTAINE [1975, 1981] and by COOK [1976], who also reviewed some of the early experimental investigations on continuous ordering.

SOFFA and LAUGHLIN [1983] reviewed the considerable amount of published work on continuous ordering reactions but they concentrated mainly on the more commonly studied case of decomposition of metastable solid solutions in which composition fluctuations occur on a large scale, spinodal decomposition, with ordering developing in the solute-rich regions. They discuss only a few studies of continuous ordering occurring without any change of compositions. However, the two-stage reactions of the spinodal decomposition to give an ordered precipitate discussed by Soffa and Laughlin are probably more important technologically.

CHEN and COHEN [1979] investigated a process of continuous ordering without changes of composition, by studies on the intensities of scattered X-rays at superlattice positions in various alloys at temperatures *above* the ordering temperatures. For Cu_3Au,

apart from some additional, heterophase, fluctuations just above the critical temperature, the data behave in the expected way for this type of theory (COOK [1976]). Extrapolation indicated an instability temperature, below which continuous ordering is expected. Between T_c and T_{instab}, ordering is expected to occur but by a *nucleation and growth* mechanism. Similar results were reported by Chen and Cohen for a higher copper alloy, for Co–Pt and for Fe–Al. Continuous ordering has also been demonstrated in β-brass (KUBO *et al.* [1980]) and in Ni–Mo alloys as reviewed by DE FONTAINE [1975, 1981]. It appears from the present data that continuous ordering can occur in the manner proposed and that to date the proposed model appears to be reasonably successful.

Figure 65, from SOFFA and LAUGHLIN [1983], illustrates various possible free energy-composition curves that can give ordered second phases by continuous reactions, involving a change of composition. In fig. 65a, the disordered curve is unstable against ordering at all compositions. An alloy of composition C_0 can lower its free energy by ordering, whereupon the ordered phase can then spinodally decompose. S_1 and S_2 are the spinodal points for the ordered structure. This type of process is described as a *conditional spinodal* (ALLEN and CAHN [1976]). An alloy of lower solute content, C_0', that falls outside the spinodal region of the ordered structure will after ordering not then decompose spinodally but will precipitate the ordered phase via a nucleation and growth process. In fig. 65b, however, the disordered curve also shows an instability so that compositions γ_0 can initially decompose in the normal spinodal manner, with the solute-rich cluster becoming ordered in the normal course of the reaction when the composition of the solute-rich regions has reached a value where the ordered phase has a lower free energy than the disordered phase. The distinction between a C_0 and a γ_0 alloy could be the difference between the systems, Ni–Al and Cu–Ti, discussed in § 3.1. It was reported there, that experiments had shown that in the nickel alloy the ordered phase was precipitated with the correct composition and order right from the start of the decomposition while in the copper alloy there was steady increase in the titanium content of the titanium-rich regions for $C_0 > 1.9$at% Ti. The schematic diagrams of fig. 64 appear to describe, qualitatively, the processes that can take place in various important alloys that precipitate ordered phases either by nucleation and growth or by spinodal decomposition. LAUGHLIN *et al.* [1984] discussed the decomposition of more concentrated Cu–Ti alloys

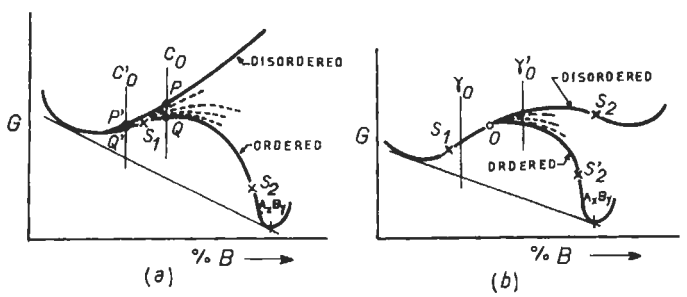


Fig. 65. Free-energy–composition curves of alloys that show various types of spinodal decomposition and continuous ordering, see the text (after SOFFA and LAUGHLIN [1983]).

References: p. 1497.

(5at% Ti) than were also studied by VON ALVENSLEBEN and WAGNER [1984] and also decomposition in Fe–Be. They showed that both of these alloy systems first underwent homogeneous ordering before subsequently spinodally decomposing, eventually forming, in the high solute regions, the ordered phase, Cu_4Ti in Cu–Ti.

4. *Application of phase transformation theory to specific alloy systems*

The application of phase transformation theory to specific systems is a huge topic and can only be treated in the briefest outline here. The main alloy system is the one based on iron which takes advantage of the allotropic transformation between the high temperature fcc (austenite) phase and the low temperature bcc (ferrite) phase since fifty times as much steel is used compared to te second most important alloy system, that based on aluminium. However the growth rate of aluminium use is faster than that of steel (FINNISTON [1978]). Chapter 17 of this volume deals specifically with the physical metallurgy of steels and gives the most appropriate discussion of the topic, and further references. For completeness however, two additional references to this topic can be given: PICKERING [1978] and HONEYCOMBE [1981].

For aluminum, an excellent review of the various alloy reactions is provided by POLMEAR [1981] whose book also includes a discussion of the phase transformations in magnesium and titanium alloys. A recent discussion of the physical metallurgy, processing and properties of aluminum alloys is the reviewed edited by VASUDEVAN and DOHERTY [1989]. The review by MONDOLFO [1976] provides exceptional detail on the equilibrium diagrams and microstructures of the aluminum alloys in use at that time. LORIMER [1978] reviewed the physical metallurgy of aluminum precipitation reactions. The same volume on precipitation processes, edited by RUSSELL and AARONSON [1978], includes accounts of precipitation in nickel alloys by MERRICK [1978], in titanium alloys by WILLIAMS [1978] as well as EDMONDS and HONEYCOMBE [1978] on iron. For the nickel alloys, SIMS and HAGEL [1972] and BETTERIDGE and HESLOP [1974] provide detailed additional information and for titanium alloys there are also the conference reports on titanium published by Pergamon Press [1971] and Plenum Press [1973]. The recent series of International Conferences on Aluminum Alloys listed, in further reading, give a wide range of recent industrial applications of diffusive phase transformations, often coupled with thermomechanical processing.

As discussed in § 1 for metals such as iron and titanium, the change of matrix crystal structure offers tremendous scope for many heat-treatments while in the alloys with a single matrix crystal structure, aluminium, copper, nickel, magnesium etc., only precipitation is readily available as a means of modifying the microstructures. However, although this is not discussed in this chapter, various additional microstructural modifying processes are available and are of increasing importance; these include rapid quenching from the liquid (ch. 19) which gives a much wider range of starting microstructure for subsequent treatment to work on than is provided by the conventional ingot casting route; the production of composite structures by powder metallurgy (ch. 31) and also the wide control of microstructure made possible by thermomechanical

treatments in which phase transformations are combined with deformation. This has been widely exploited for grain-size control in low-carbon alloys of iron and for ausformed iron, see ch. 17, but is now of increasing interest for modification of microstructure by precipitation during deformation. HORNBOGEN and KÖSTER [1978] have shown something of the interesting results obtained by cold deformation, followed by precipitation, of supersaturated solid solutions. The next step is the combination of warm deformation with precipitation, an area only just being scientifically explored in alloys other than iron. The process occurs every day of course during current warm-rolling of alloys that are subsequently precipitation strengthened. An important review of the field was provided in the book edited by MORRIS [1979], on thermomechanical processing of aluminium alloys. There are also excellent reviews by FERTON [1981] on thermomechanical processing of aluminium alloys and by LUETJERING and PETERS [1981] on thermomechanical processing of titanium alloys. Recent conference reports on aluminum alloys and nickel-based superalloys update the situation with respect to all aspects of the physical metallurgy of these industrial alloys.

5. *Problems in phase transformations*

A very interesting and unusual approach in materials science and technology was the challenge offered by MADDIN [1976] in the 25th volume of the journal, Materials Science and Engineering to 45 scientists to write a brief account of what was *not known* in the 45 topics of major interest to the authors. Many of the topics were from the area of phase transformations. 40 of those invited accepted the challenge and produced an important series of articles. However, most of the reviews were conventional accounts of what was established and sadly did not really address the challenge of discussing what was *not known* or *not understood*. From the previous pages in this chapter, some of the important questions raised but not satisfactorily answered can be listed and there are no doubt more not described in this chapter because of the authors insufficient knowledge. Questions and uncertainties that seem most important, or at least interesting, include the following:

(1) More experimental data to test all the theories of the transformation but in particular quantitative data are particularly needed on the rates of heterogeneous nucleation on both grain boundaries and dislocations, to answer the questions raised by *coherency* of the interface and the higher-level heterogeneities in grain boundaries.

(2) Extension of atomic modeling of nucleation such as the cluster model described by HAASEN and WAGNER [1992].

(3) What are the mechanisms of nucleation and the nucleation kinetics of the growth ledges on coherent interfaces between phases of different structures, both during precipitation and during coarsening?

(4) The rationalization of the role and significance of surface relief in transformations showing a change of composition, particularly the formation of Widmanstätten ferrite and the bainite reaction in steels, the extent of the departure for local equilibrium in the solute contents in these bainitic products and the origin of the incomplete formation of bainite in some alloy steels.

References: p. 1497.

(5) Do all, or only some, Widmanstätten plates grow by a displacive interface mechanism, giving surface relief?

(6) What determines if precipitates that only have a good fit direction with the matrix (e.g., bcc/fcc) grow as needles or as plates?

(7) Understanding the determining principle of the lamellar spacing in all types of discontinuous reactions, and the tip radius in linear lengthening of Widmanstätten plates and needles.

(8) The problem of shape stability for non-spherical precipitates.

(9) The problem of understanding the effect, if any, of the volume fraction of second-phase particles on the rate of Ostwald ripening, especially in the Ni–Ni_3X systems.

(10) What are the conditions in which discontinuous precipitation at grain boundaries occurs in preference to nucleation and growth without the interference of a moving grain boundary?

(11) To what extent does the coherency stress, known to be the controlling mechanism of diffusion-induced grain-boundary migration, play in the growth kinetics of all discontinuous (moving two-phase boundary) reactions?

(12) How does diffusion-induced recrystallization manage to produce the high angle grain boundary between the new grain and the old grain into which it grows?

(13) The need for a great deal more data on free energies, interfacial energies and the role of impurities in modifying interfacial energies, needed for accurate prediction of phase-transformation kinetics and thus of resultant microstructures.

(14) The problem of the complications to all aspects of phase transformations produced by external variables such as elastic stress, plastic deformation, radiation (HUDSON [1978]) and temperature gradients (MCLEAN [1979]) and the effects of these variables both on microstructures and on the phase transformations that produce these microstructures.

It would appear from this list that despite the great successes achieved in the past 50 years in understanding phase transformations in metals, much remains unknown and therefore fascinating in this field.

Acknowledgements

The author is very grateful to many scientists whose discussions have provided a great deal of such insight as this chapter contains. In particular I would like to acknowledge Prof. H. I. Aaronson, Dr. J. W. Cahn, Prof. R. W. Cahn, Dr. B. Cantor, Prof. J. W. Christian, Prof. M. Hillert, Dr. J. W. Martin, Prof. G. R. Purdy, Dr. R. Shuey, Dr. A. K. Vasudevan as well as my former doctoral students, especially Dr. B. G. Bainbridge, Dr. S. J. M. Fairs, Dr. M. Ferrante, Dr. K. Rajab, Dr. J. Liu, Dr. A. D. Rollett, Dr. J. T. Staley and the late Dr. R. P. Singh. The mistakes are however all my responsibility. The author also wishes to thank Prof. Erik Nes for his hospitality at the Norwegian Institute of Technology that has enabled the revision of this chapter to be finally completed.

References

AARON, H. B., and G. R. KOTLER, 1971, Metallurg. Trans. **2**, 393.

AARON, H. B., D. FAINSTEIN and G. R. KOTLER, 1970, J. Appl. Phys. **41**, 4404.

AARONSON, H. I., 1962, in: The Decomposition of Austenite by Diffusional Processes (Interscience, New York) p. 387.

AARONSON, H. I., 1969, in: The Mechanisms of Phase Transformations in Crystalline Solids (The Metals Society, London) p. 270.

AARONSON, H. I., 1974, J. Microsc. **102**, 275.

AARONSON, H. I., 1977, Scripta Metall. **11**, 731 and 741.

AARONSON, H. I., 1979, in: Phase Transformations, York Conference (The institution of Metallurgists, London) p. II–I.

AARONSON, H. I., and C. LAIRD, 1968, TMS AIME **242**, 1437.

AARONSON, H. I., J. B. CLARK and C. LAIRD, 1968, Met. Sci. **2**, 155.

AARONSON, H. I., C. LAIRD and K. R. KINSMAN, 1970, in: Phase Transformation (ASM, Metals Park, OH) p. 313.

AARONSON, H. I., K. C. RUSSELL and G. W. LORIMER, 1977, Metallurg. Trans. **8A**, 1644.

AARONSON, H. I., 1977, Scripta Metall. **11**, 731, 741.

AARONSON, H. I., J. K. LEE and K. C. RUSSELL, 1978, in: Precipitation Processes in Solids, eds. K. C. Russell, K. C. and H. I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 31.

AARONSON, H. I., and W. T. REYNOLDS, Jr., 1987, Scripta Metall. **21**, 1985.

AARONSON, H. I., and W. T. REYNOLDS, Jr., 1988, Scripta Metall. **22**, 567 and 575.

AARONSON, H. I., J. M. RIGSSEE, T. FURUHAWA, N. PRAHSU, W. T. REYNOLDS, Jr. and J. M. HOWE, 1989, Scripta Metall. **23**, 279.

AARONSON, H. I., W. T. REYNOLDS, Jr., G. J. SHIFLET and G. SPANOS, 1990, Metallurg. Trans. **21A**, 1343.

AARONSON, H. I., and M. G. HALL, 1994, Metallurg. Mater. Trans. **25A**, 1769.

AARONSON, H. I., J. P. HIRTH, B. B. RATH and C. M. WAYMAN, 1994, Metallurg. Mater. Trans. **25A**, 2655.

ABBOTT, K., and C. HAWORTH, 1973, Acta Metall. **21**, 951.

AGEEW, N., and G. SACHS, 1930, Z. Phys. **66**, 293.

AL–KASSAB, T., 1987, Diploma thesis, University of Göttingen, Quoted by HAASEN and WAGNER [1992].

ALLEN, S., and J. W. CAHN, 1974, Acta Metall. **24**, 425.

ANNAVARAPU, S., and R. D. DOHERTY, 1995, Acta Metall. and Mater. **42**, 3207.

ARDELL, A. J., 1972, Acta Metall. **20**, 61.

ARDELL, A. J., 1987, in Phase Transformations 87, ed. G. Lorimer, The Institute of Metals, London, p. 485.

ARDELL, A. J., 1990, Scripta Metall. and Mater. **24**, 343.

ASHBY, M. F., and L. JOHNSON, 1969, Phil. Mag. **20**, 1009.

ASUNDI, M. P., and D. R. F. WEST, 1966, J. Inst. Metals **94**, 327.

AUST, K. T., and J. W. RUTTER, 1963, in: Recovery and Recrystallization of Metals, ed. L. Himmel (Interscience, New York) p. 131.

AZIZ, M. J., 1982, J. Appl. Phys. **53**, 1158.

BAINBRIDGE, B. G., and R. D. DOHERTY, 1969, in: Quantitative Relationship between Properties and Microstructure, eds. D. G. Brandon and A. Rosen (Israel Univ. Press, Jerusalem) p. 427.

BAINBRIDGE, B. G., 1972, D. Phil Thesis, University of Sussex, U. K.

BALLUFFI, R. W., 1982, Metallurg. Trans. **13A**, 2069.

BALLUFFI, R. W., and J. W. CAHN, 1981, Acta Metall. **29**, 493.

BARRETT, C. S., 1952, Structure of Metals, 2nd Ed. (McGraw-Hill, New York) p. 548.

BAUMANN, S. F., and D. B. WILLIAMS, 1985, Metallurg. Trans. **16A**, 1203.

BECKER, R., and W. DORING, 1934, Ann. Phys. **24**, 719.

BEDDOE, R. E., P. HAASEN and G. KOSTORZ, 1984, in Decomposition of Alloys: The Early Stages; eds., P. Haasen, V. Gerold, R. Wagner and M. F. Ashby. Pergammon Press Oxford, p. 233.

BETTERIDGE, W., and J. HESLOP, eds., 1974, The Nimonic Alloys (Crane and Russak, New York).

BHADESHIA, H. K. D. H., 1980, Acta Metall. **28**, 1103.

BHADESHIA, H. K. D. H., 1981, Acta Metall. **29**, 1117.
BHADESHIA, H. K. D. H., and D. V. EDMONDS, 1980, Acta Metall. **28**, 1265.
BHADESHIA, H. K. D. H., 1980, Acta Metall. **28**, 1103.
BHADESHIA, H. K. D. H., 1985, Prog. in Mater. Sci. **29**, 321.
BHADESHIA, H. K. D. H., 1987, Acta Metall. **21**, 1605.
BHADESHIA, H. K. D. H., and J. W. CHRISTIAN, 1990, Metallurg. Trans. **21A**, 767.
BIEHL, K. E., 1980, Ph.D. thesis, Univ. of Göttingen.
BIEHL, K. E., and R. WAGNER, 1982, in: Solid-State Phase Transformations, the Pittsburgh Conference, ed. H. I. Aaronson et al., (Met. Soc. AIME, Warrendale, PA) p. 185.
BOHM, H., 1961, Z. Metallkde. **52**, 564.
BOSZE, W. R., and R. TRIVEDI, 1974, Metallurg. Trans. **5**, 511.
BOWER, E. N., and B. M. WHITEMAN, 1969, in: The Mechanisms of Phase Transformations in Crystalline Solids (The Metals Society, London) p. 119.
BOWLES, J. S., B. C. MUDDLE and C. M. WAYMAN, 1977, Acta Metall. **25**, 513.
BROWN, L. C., 1989, Acta Metall. **37**, 71.
BROWN, L. C., 1990a, Scripta Metall. and Mater. **24**, 963.
BROWN, L. C., 1990b, Scripta Metall. and Mater. **24**, 2231.
BROWN, L. C., 1992a, Scripta Metall. and Mater. **26**, 1939.
BROWN, L. C., 1992b, Scripta Metall. and Mater. **26**, 1945.
BRAILSFORD, I. N., and P. WYNBLATT, 1979, Acta Metall. **27**, 489.
BUTLER, E. P., and P. SWANN, 1976, Acta Metall. **24**, 343.
BUTLER, E. P., and G. THOMAS, 1970, Acta Metall. **18**, 347.
CAHN, J. W., 1956, Acta Metall. **4**, 441.
CAHN, J. W., 1957, Acta Metall. **5**, 168.
CAHN, J. W., 1961, Acta Metall. **9**, 795.
CAHN, J. W., 1962, Acta Metall. **10**, 179.
CAHN, J. W., 1963, Acta Metall. **11**, 789.
CAHN, J. W., 1968, TMS AIME **242**, 166.
CAHN, J. W., and J. E. HILLIARD, 1958, J. Chem. Phys. **28**, 259.
CAHN, J. W., and J. E. HILLIARD, 1959, J. Chem. Phys. **31**, 3 and 788.
CAHN, J. W., J. D. PAN and R. W. BALLUFFI, 1978, Scripta Metall. **13**, 2069.
CALDERON, H. A., P. W. VOORHEES, J. L. MURRAY and G. KOSTORZ, 1994, Acta Metall. and Mater. **42**, 991.
CANTOR, B., and R. D. DOHERTY, 1979, Acta Metall. **27**, 33.
CASSIDY, M. P., B. C. MUDDLE, T. E. SCOTT, C. M. WAYMAN and J. S. BOWLES, 1977, Acta Metall. **25**, 829.
CHEETHAM, D., and N. RIDLEY, 1971, J. Inst. Metals **99**, 371.
CHELLMAN, D. J., and A. J. ARDELL, 1974, Acta Metall. **22**, 577.
CHEN, H., and J. B. COHEN, 1979, Acta Metall. **27**, 603.
CHEN, Y. H., and R. D. DOHERTY, 1977, Scripta Metall. **11**, 725.
CHONGMO, A., and M. HILLERT, 1981, Acta Metall. **29**, 1949.
CHRISTIAN, J. W., 1965, in: The Physical Properties of Martensite and Bainite (The Metals Society, London) p. 1.
CHRISTIAN, J. W., 1970, in: Physical Metallurgy, 2nd Ed., ed. R. W. Cahn (North-Holland, Amsterdam) p. 471.
CHRISTIAN, J. W., 1975, The Theory of Transformations in Metals and Alloys, 2nd Ed. (Pergamon Press, Oxford).
CHRISTIAN, J. W., 1979, in: Phase Transformations, York Conference (The Institution of Metallurgists, London) p. 1.
CHRISTIAN, J. W., and EDMONDS, D. V., 1984, in: Phase Transformations in Ferrous Alloys, eds A. R. Marder and G. I. Goldstein, TMS-AIME, Warrendale, PA, p. 293.
CHRISTIAN, J. W., and EDMONDS, D. V., 1988, Scripta Metall. **22**, 573, 577.
CHRISTIAN, J. W., and EDMONDS, D. V., 1989, Scripta Metall. **23**, 285.
CHRISTIAN, J. W., 1990, Metallurg. Trans. **21A**, 799.
CHRISTIAN, J. W., 1994, Metallurg. Mater. Trans. **25A**, 1821.
CHUNG, Y. H., M. C. SHIN and D. Y. YOON, 1992, Acta Metall. and Mater. **40**, 2177.

CHUANG, T.H., R.A. FOURNELLE, W. GUST and B. PREDEL, 1988, Acta Metall. **36**, 775.
CLARK, H.M., and C.M. WAYMAN, 1970, in: Phase Transformations ASM, Metals Park, OH) p. 19.
CLIFF, G., F. HASON, G.W. LORIMER and M. KIKUCHI, 1990, Metallurg. Trans. **21**A, 831.
COOK, H.E., 1970, Acta Metall. **18**, 297.
COOK, H.E., 1976, Mater. Sci. Eng. **25**, 127.
COOK, H.E., and J.E. HILLIARD, 1969, J. Appl. Phys. **40**, 2192.
COOK, H.E., D. DE FONTAINE and J.E. HILLIARD, 1969, Acta Metall. **17**, 765.
CORIELL, S.R., and R.L. PARKER, 1967, in: Crystal Growth, ed. H.S. Peiser (Pergamon Press, Oxford) p. 703.
DAHMEN, U., and K.H. WESTMACOTT, 1986, Acta Metall. **34**, 475.
DAHMEN, U., P. FERGUSON and K.H. WESTMACOTT, 1987, Acta Metall. **35**, 1037.
DAHMEN, U., 1994, Metallurg. Trans. **25**A, 1857.
DARKEN, L.S., 1948, Trans. Met. Soc. AIME **174**, 184.
DE BONDT, M., and A. DERUYTERRE, 1977, Acta Metall. **25**, 993.
DEN BRODER, F.J.A., 1972, Acta Metall. **20**, 319.
DE FONTAINE, D., 1975, Acta Metall. **23**, 553.
DE FONTAINE, D., 1981, Metallurg. Trans. **12**A, 559.
DESALOS, Y., A. LEBON and R. LOMBRY, 1981, in: Les Traitements Thermomechaniques, Comptes Rendus, 24e Colloque de Metallurgie Saclay, France (CEN-Saclay) p. 137.
DITCHEK, B., and L.H. SCHWARTZ, R.D., 1980, Acta Metall. **28**, 807.
DOHERTY, R.D., 1976, unpublished research.
DOHERTY, R.D., M. FERRANTE and Y.H. CHEN, 1977, Scripta Metall. **11**, 733.
DOHERTY, R.D., 1978, in: Recrystallization in Metallic Materials, ed. F. Haessner (Riederer-Verlag, Stuttgart) p. 23.
DOHERTY, R.D., B. CANTOR and S.J.M. FAIRS, 1978a, Met. Trans. **9A**, 621.
DOHERTY, R.D., M. FERRANTE and Y.H. CHEN, 1978b, Scripta Metall. **12**, 885.
DOHERTY, R.D., 1980, in: Crystal Growth, ed. B. Pampkin (Pergamon Pres, Oxford) p. 485.
DOHERTY, R.D., 1982, Met. Sci. **16**, 1.
DOHERTY, R.D., and B. CANTOR, 1982, in: Solid-State Phase Transformations, the Pittsburgh Conference, ed. H.I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 547.
DOHERTY, R.D., and G.R. PURDY, 1982, in: Phase Transformations, the Pittsburgh Conference, ed. H.I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 561.
DOHERTY, R.D.., M. FERRANTE and Y.H. CHEN, 1977, Scripta Metall. **11**, 733.
DOHERTY, R.D., B. CANTOR and S.J.M. FAIRS, 1978a, Metallurg. Trans. **9**A, 621.
DOHERTY, R.D., M. FERRANTE and Y.H. CHEN, 1978b, Scripta Metall. **12**, 885.
DOHERTY, R.D., E.A. FEEST and H.I. LEE, 1984, Mater. Sci. and Eng. **65**, 181.
DOHERTY, R.D., and B. CANTOR, 1988, Scripta Metall. **22**, 301.
DOHERTY, R.D., and B. CANTOR, 1989, Scripta Metall. **23**, 1989.
DOHERTY, R.D., and K. RAJAB, 1989, Acta Metall. **37**, 2723.
DOHERTY, R.D., 1992, Encylopedia of Mater. Sci. and Eng., Supplementary Vol. 3, Ed. R.W. Cahn, Pergamon Press, Oxford, p. 1695.
EDINGTON, J.W., 1979, in: Phase Transformations, York Conference (The Institution of Metallurgists, London) p. 1–1.
EDMONDS, D.V., and R.W.K. HONEYCOMBE, 1978, in: Precipitation Processes in Solids, eds. K.C. Russell and H.I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 121.
EDMONDS, D.V., and R.W.K. HONEYCOMBE, 1979, in: Phase Transformations, York Conference (The Institution of Metallurgists, London) p. 121.
EHRENFEST, P., 1933, Proc. Acad. Sci. Amsterdam **36**, 153.
ENOMOTO, M., 1987, Acta Metall. **35**, 935 and 947.
ENOMOTO, M., and H.I. AARONSON, 1989, Scripta Metall. **23**, 1983 and 1995.
FAIRS, S.J.M., 1977, D. Phil. thesis, Susses Univ.
FANG, Z., and B.R. PATTERSON, 1993, Acta Metall. and Mater. **41**, 2017.
FERRANTE, M., 1978, D. Phil. thesis, Sussex Univ.
FERRANTE, M., and R.D. DOHERTY, 1976, Scripta Metall. **10**, 1059.

FERRANTE, M., and R. D. DOHERTY, 1979, Acta Metall. **27**, 1603.
FERTON, D., 1981, in: Les Traitments Thermomechaniques, Comptes Rendus, 24~ Colloque de Metallurgie, Saclay, France (CE~-Saclay) p. 243.
FINE, M. E., 1964, Introduction to Phase Transformations in Condensed Systems (Macmillan, New York).
FINNISTON, M., 1978, Metallurg. Trans. **9A**, 327.
GARWOOD, R. D., 1954, J. Inst. Metals **83**, 64.
GAUNT, P., and J. W. CHRISTIAN, 1959, Acta Metall. **7**, 529.
GLICKSMAN, M. E., R. J. SCHAEFER and A. J. AYERS, 1976, Metallurg. Trans. **7A**, 1747.
GRUJICIC, M., G. B. OLSON and W. S. OWEN, 1985, Metallurg. Trans. **16A**, 1713, 1723 and 1735.
GREENWOOD, G. W., 1956, Acta Metall. **4**, 243.
GUST, W., 1979, in: Phase Transformations, York Conference (The Institution of Metallurgists, London) p. 11–27.
HAASEN, P., V. GEROLD, R. Wagner and M. F. ASHBY, 1984, Decomposition of Alloys: The Early Stages, Pergammon Press, Oxford.
HAASEN, P., and R. WAGNER, 1986, Ann. Rev. Mater. Sci. **15**, 43.
HAASEN, P., and R. WAGNER, 1992, Metallurg. Trans. **23A**, 1901.
HAESSNER, F., and S. HOFMANN, 1978, in: Recrystallization in Metallic Materials, ed. F. Haessner (Riederer-Verlag, Stuttgart) p. 63.
HALL, M. G., and C. HAWORTH, 1970, Acta Metall. **18**, 331.
HALL, M. G., H. I. AARONSON and G. W. LORIMER, 1975, Scripta Metall. **9**, 533.
HALL, M. G., and H. I. AARONSON, 1994, Metallurg. Mater. Trans. **25A**, 1923.
HAM, F. S., 1958, J. Phys. Chem. Solids **6**, 335.
HAM, F. S., 1959, Quart. Appl. Maths. **17**, 135.
HANDWERKER, C. A., 1988, In Diffusion Phenomena in Thin Film and Microelectronic Materials, Ed. D. Gupta and P. S. Ho Noyes, Park Ridge, NJ. 245.
HEADY, R. B., and J. W. CAHN, 1973, J. Chem. Phys. **58**, 896.
HEHEMANN, R. F., 1970, in: Phase Transformations (ASM, Metals Park, OH) p. 379.
HEHEMANN, R. F., K. R. KINSMAN and H. I. AARONSON, 1972, Metallurg. Trans. **3**, 1077.
HIGGINS, G. T., WIRYOLUKITO, S. and NASH, P., 1992, In Grain Growth in Polycrystalline Materials, Eds. G. Abbruzzese and P. Brosso, Materials Science Forum (Trans Tech Publications Zurich) 94–96, 671.
HILLERT, M., 1956, D. Sc. Thesis MIT.
HILLERT, M., 1957, Jernkontorets Ann. **141**, 757.
HILLERT, M., 1961, Acta Metall. **9**, 525.
HILLERT, M., 1969, in: The Mechanisms of Phase Transformation in Crystalline Solids (The Metals Society, London) p. 231.
HILLERT, M., 1972, Metallurg. Trans. **3**, 2729.
HILLERT, M., 1975, Metallurg. Trans. **6A**, 5.
HILLERT, M., 1982, Acta Metall. **30**, 1689.
HILLERT, M., and G. R. PURDY, 1978, Acta Metall. **26**, 333.
HILLERT, M., O. HUNDERI and N. RYUM, 1989, Scripta metall. **23**, 1979.
HILLERT, M., O. HUNDERI and N. RYUM, 1992, Scripta metall. **26**, 1933.
HILLERT, M., O. HUNDERI and N. RYUM, 1989, Scripta metall. **26**, 1943.
HILLERT, M., 1994, Metallurg. Mater. Trans. **25A**, 1957.
HILLERT, M., 1995, Private Communication.
HILLIARD, J. E., 1970, in: Phase Transformations (ASM, Metals Park, OH) p. 497.
HIRATA, T., and D. H. KIRKWOOD, 1977, Acta Metall. **25**, 1425.
HIRTH, J. P., 1994, Metallurg. Mater. Trans. **25A**, 1885.
HONDROS, E., 1978, in: Precipitation Processes in Solids, eds. K. C. Russell and H. I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 1.
HONEYCOMBE, R. W. K., 1981, Steels: Microstructure and properties (Edward Arnold, London).
HORNBOGEN, E., 1967, Aluminium **43**, 115.
HORNBOGEN, E., and U. KOSTER, 1978, in: Recrystallization in Metallic Materials, ed. F. Haessner (Riederer-Verlag, Stuttgart) p. 159.

HORVAY, G., and J. W. CAHN, 1961, Acta Metall. **9**, 695.
HOWE, J. M., H. I. AARONSON and R. GRONSKY, 1985, Acta Metall. **33**, 649.
HOWE, J. M., 1994, Metallurg. Mater. Trans. **25A**, 1917.
HREN, J. A., and G. THOMAS, 1963, TMS AIME **221**, 304.
HUANG, S.-C., and M. E. GLICKSMAN, 1981, Acta Metall. **29**, 701.
HUDSON, B., 1978, in: Precipitation Processes in Solids, eds. K. C. Russell and H. I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 284.
HULTGREN, A., 1947, Trans. ASM **39**, 815.
INOKUTI, Y., and B. CANTOR, 1982, Acta Metall. **30**, 343.
INTERNATIONAL CONFERENCE ON BAINITE, 1990, Metallurg. Trans. **21A**, 767-893 and 1343-1545.
IVANTSOV, P. G., 1947, Dokl. Akad. Nauk. SSSR **58**, 567.
JACKSON, K. A., and J. D. HUNT, 1966, Met. Trans. AIME. **236**, 1129.
JACOBS, M. H., and D. W. PASHLEY, 1969, in: The Mechanisms of Phase Transformation in Crystalline Solids (The Metals Society, London) p. 43.
JONES, G. J., and R. TRIVEDI, 1971, J. Appl. Phys. **42**, 4299.
JONES, G. J., and R. TRIVEDI, 1975, J. Cryst. Growth **29**, ISS.
KAMPMANN, L., and M. KAHLWEIT, 1967, Ber. Bunsen Gesell. Phys. Chem. **71**, 78.
KAMPMANN, L., and M. KAHLWEIT, 1970, Ber. Bunsen Gesell. Phys. Chem. **74**, 456.
KAMPMANN, R., and R. WAGNER, 1984, in Decomposition of Alloys: The Early Stages, eds. P. Haasen, V. Gerold, R. Wagner and M. F. Ashby, Pergammon Press Oxford, p. 91.
KAMPMANN, R., H. ECKERLEBE and R. WAGNER, 1984, In Phase Transformations in Condensed Systems — Experiment and Theory, Materials Reseearch Society, Pittsburgh, PA 15237.
KARLYN, D., J. W. CAHN and M. COHEN, 1969, TMS AIME **245**, 1971.
KATTAMIS, T. Z., U. T. HOLMBERG and M. C. FLEMINGS, 1967, J. Inst. Metals **99**, 343.
KELLY, A., and G. W. GROVES, 1970, Crystallography and Crystal Defects (Addison-Wesley, Reading, MA) p. 310.
KELLY, A., and R. B. NICHOLSON, 1963, Prog. Mater. Sci. **10**, 151.
KELLY, P. M., and J. NUTTING, 1961, J. Iron Steel Inst. **197**, 199.
KINSMAN, H. A., H. I. AARONSON and E. EICHEN, 1971, Metallurg. Trans. **2**, 1041.
KINSMAN, H. A., H. I. AARONSON and E. EICHEN, 1973, Metallurg. Trans. **4**, 369.
KIRKALDY, J. S., and R. C. SHARMA, 1980, Acta Metall. **28**, 1009.
KIRKWOOD, D. H., 1970, Acta Metall. **18**, 563.
KO, T., and S. A. COTTRELL, 1952, J. Iron Steel Inst. **172**, 307.
KOSTIC, M. M., and E. B. HAWBOLT, 1975, Metallurg. Trans. **10A**, 165.
KURZ, W., and D. J. FISHER, 1989, Fundamentals of Solidification, Trans. Tech Publications, PO Box 10, Aedermannsdorf, Switzerland.
KUBO, H., I. CORNELIS and C. M. WAYMAN, 1980, Acta Metall. **28**, 395.
LAIRD, C., and H. I. AARONSON, 1969, Acta Metall. **17**, 505.
LANDUYT, and WAYMAN, 1968, Acta Metall. **16**, 803.
LANGE, W. F., 1979, Ph.D. Thesis, Michigan Technol. Univ., Houghton.
LANGER, J. S., 1971, Ann. Phys. **65**, 53.
LANGER, J. S., 1975, in: Fluctuations, Instabilities and Phase Transitions, ed. T. Riste (Plenum, New York) p. 19.
LANGER, J. S., and H. MULLEK KRUMBHAAR, 1978, Acta Metall. **26**, 61.
LANGER, J. S., and A. J. SCHWARTZ, 1980, Phys. Rev. **A21**, 948.
LANGER, J. S., M. BAR–ON and H. D. MILLER, 1975, Phys. Rev. **A11**, 1417.
LUO, C. P., and G. C. WEATHERLY, 1987, Acta Metall. **35**, 1963.
LAUGHLIN, D. E., K. B. ALEXANDER and L. L. LEE, 1984, in Decomposition of Alloys: The Early Stages, eds. P. Haasen, V. Gerold, R. Wagner and M. F. Ashby. Pergammon Press Oxford. p. 221.
LEE, J. K., and H. I. AARONSON, 1975, Acta Metall. **23**, 799 and 809.
LEE, J. K., and W. C. JOHNSON, 1983, in: Solid-State Phase Transformations, the Pittsburg Conference, ed. H. I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 127.
LEE, J. K., Y. Y. EARMME, H. I. AARONSON and K. C. RUSSELL, 1980, Metallurg. Trans. **11A**, 1837.

LEE, H. J., and H. I. AARONSON, 1988, Acta Metall. **36**, 787.
LEGOUES, F. K., and H. I. AARONSON, 1984, Acta Metall. **32**, 1855.
LIFSHITZ, I. M., and V. V. SLYOZOV, 1961, J. Phys. Chem. Solids **19**, 35.
LIU, Y. C., and H. I. AARONSON, 1970, Acta Metall. **18**, 841.
LIVAK, R. J., and G. THOMAS, 1971, Acta Metall. **19**, 497.
LIVINGSTON, J. D., and J. W. CAHN, 1974, Acta Metall. **22**, 495.
LORIMER, G. W., 1978, in: Precipitation Processes in Solids, eds. K. C. Russell and H. I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 87.
LORIMER, G. W., ed., 1981, Quantitative Analysis with High Spatial Resolution (The Metals Society, London).
LORIMER, G. W., and R. B. NICHOLSON, 1969, in: The Mechanisms of Phase Transformations in Crystalline Solids, (The Metals Society, London) p. 36.
LÜCKE, K., and H. P. STUWE, 1963, in: Recovery and Recrystallization of Metals, ed. L. Himmel (Wiley, New York) p. 171.
LUETJERING, G., and M. PETERS, 1981, in: Les Traitements Thermomechaniques, Comptes Rendus, 24e Colloque de Metallurgie, Saclay, France (CEN-Saclay) p. 225.
MADDIN, R., 1976, Mater. Sci. Eng. **25**, 1.
MALCOLM, J. A., and G. R. PURDY, 1967, TMS AIME **239**, 1391.
MAKENAS, B. J., and H. K. BIRNBAUM, 1980, Acta Metall. **28**, 979.
MARTIN, J. W., and R. D. DOHERTY, 1976, 1996, Stability of Microstructure in Metallic Systems (Cambridge Univ. Press). (New edition in press, 1996).
MARTIN, G., and A. BARBU, 1984, in Decomposition of Alloys: The Early Stages; eds, P. Haasen, V. Gerold, R. Wagner and M. F. Ashby, Pergamon Press Oxford, p. 70.
MASSALSKI, T. B., 1970, in: Phase Transformations (ASM, Metals Park, OH) p. 433.
MCLEAN, M., 1979, in: Phase Transformations, York Conference (The Institution of Metallurgists, London) p. 11–54.
MEHL, R. F., 1947, J. I.S.I. **159**, 113.
MERLE, P., and R. D. DOHERTY, 1982, Scripta Metall. **16**, 357.
MERLE, P., and F. FOUQUET, 1981, Acta Metall. **29**, 1919.
MERLE, P., and J. MERLIN, 1981, Acta Metall. **29**, 1929.
MERRICK, H. F., 1978, in: Precipitation Processes in Solids, eds. K. C. Russell and H. I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 161.
MILLER, M. K., 1987, International Materials Reviews **32**, 221.
MONDOLFO, L. F., 1976, Aluminium Alloys: Structures and Properties (Butterworths, London).
MORRIS, J. G., ed., 1979, Thermomechanical Processing of Aluminum Alloys (AIME, New York).
MUDDLE, B. C., J. F. NIE and G. R. HUGO, 1994, Metallurg. and Mater. Trans. **25A**, 1841.
MULLINS, W. W., 1959, J. Appl. Phys. **30**, 77.
MULLINS, W. W., 1963, in: Metal Surfaces (ASM, Metals Park, OH) p. 17.
MULLINS, W. W., and R. F. SEKERKA, 1963, J. Appl. Phys. **34**, 323.
NAKKALIL, R., and S. P. GUPTA, 1987, Acta Metall. **35**, 2157.
NEMOTO, M., 1974, Acta Metall. **22**, 847.
NICHOLSON, R. B., 1970, in: Phase Transformations (ASM, Metals Park, OH) p. 269.
NICHOLSON, R. B., and J. NUTTING, 1961, Acta Metall. **9**, 331.
OLSON, G. B., H. K. D. B. BHADESHIA and M. COHEN, 1989, Acta Metall. **37**, 381.
OLSON, G. B., H. K. D. B. BHADESHIA and M. COHEN, 1990, Metallurg. Trans. **21A**, 805.
ORIANI, R. A., 1966, Acta Metall. **14**, 84.
OSTWALD, W., 1900, Z. Phys. Chem. **34**, 495.
PERGAMON PRESS (Oxford), 1971, Science, Technology and Application of Titanium.
PAN, J. D., and R. W. BALLUFFI, 1982, Acta Metall. **30**, 861.
PEROVIC, A., and G. R. PURDY, 1981, Acta Metall. **29**, 53.
PEROVIC, V., G. R. PURDY and L. M. BROWN, 1981, Acta Metall. **29**, 889.
PHILOFSKY, E. M., and J. E. HILLIARD, 1969, J. Appl. Phys. **40**, 2198.
PICKERING, F. B., 1978, Physical Metallurgy and the Design of Steels (Appl. Science Publ., London).
PLENUM PRESS (New York), 1973, Titanium Science and Technology.

PLICHTA, M. R., and H. I. AARONSON, 1980, Acta Metall. **28**, 1041.
PLICHTA, M. R., H. I. AARONSON and J. H. PEREPEZKO, 1978, Acta Metall. **26**, 1293.
PLICHTA, M. R., J. H. PEREPEZKO, H. I. AARONSON and W. F. LANGE, 1980, Acta Metall. **28**, 1031.
POLMEAR, I. J., 1981, Light Alloys (Edward Arnold, London).
PORTER, D. A., D. B. WILLIAMS and J. W. EDINGTONS, 1974, in: Electron Microscopy 1974, vol. 1, eds. J. V. Sanders and D. J. Goodchild (Australian Academy of Science, Canberra) p. 656.
PREDEL, B., and W. GUST, 1975, Mater. Sci. Eng. **17**, 41.
PULS, M. P., and J. S. KIRKALDY, 1972, Metallurg. Trans. **3**, 2779.
PURDY, G. R., 1971, Met. Sci. **5**, 81.
PURDY, G. R., 1978, Acta Metall. **26**, 477.
PURDY, G. R., and M. HILLERT, 1984, Acta Metall. **32**, 823.
RAJAB, K., 1982, D. Phil. thesis, Univ. of Sussex.
RAJAB, K., and R. D. DOHERTY, 1982, in: Solid-State Phase Transformations, the Pittsburgh Conference, ed. H. I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 555.
RAJAB, K., and R. D. DOHERTY, 1989, Acta Metall. **37**, 2709.
RHEE, Y. H., and D. N. YOON, 1989, Acta Metall. **37**, 221.
RHINES, F. N., and A. M. MONTGOMERY, 1938, Nature **141**, 413.
RICKS, R. A., 1981, in: Quantitative Analysis with High Spatial Resolution, ed. G. W. Lorimer (The Metals Society, London) p. 85.
RICKS, R. A., A. J. PORTER and R. C. ECOB, 1983, Acta Metall. **31**, 43.
RIDLEY, N., and G. W. LORIMER, 1981, in: Quantitative Analysis with High Spatial Resolution, ed. G. W. Lorimer (The Metals Society, London) p. 80.
ROLE OF SHEAR AND DIFFUSION IN PLATE SHAPED TRANSFORMATION PRODUCTS, 1994, Metallurg. Mater. Trans. **25A**, 1785-2016 and 2552-2673.
ROLLETT, A. D., R. D. DOHERTY, D. J. SROLOVITZ and M. P. ANDERSON, 1989, Acta Metall. **37**, 627.
RUNDMAN, K. B., and J. E. HILLIARD, 1967, Acta Metall. **15**, 1025.
RUNDMAN, K. B., D. DE FONTAINE and J. E. HILLIARD, 1970, quoted by HILLIARD [1970].
RUSSELL, K. C., 1970, in: Phase Transformations (ASM, Metals Park, OH) p. 219.
RUSSELL, K. C., 1980, Adv. Colloid and Interface Sci. **13**, 205.
RUSSELL, K. C., and H. I. AARONSON, eds., 1978, Precipitation Processes in Solids (Met. Soc. AIME, Warrendale, PA).
RUSSEW, K., and W. GUST, 1979, Z. Metallk. **70**, 523.
SADLER–HOLM, P., 1994, M. S. Thesis, Drexel University, Philadelphia, PA 19104.
SANKARAN, R., and C. LAIRD, 1974, Acta Metall. **22**, 957.
SANKARAN, R., and C. LAIRD, 1977, Scripta Metall. **11**, 385.
SANKARAN, R., and C. LAIRD, 1978, Scripta Metall. **12**, 877.
SARGENT, C. M., and G. R. PURDY, 1974, Scripta Metall. **8**, 569.
SCHMIDT, C. G., and A. K. MILLER, 1982, Acta Metall. **30**, 615.
SERVI, I. S., and D. TURNBULL, 1966, Acta Metall. **14**, 161.
SHEWMON, P. G., 1963, Diffusion in Solids (McGraw-Hill, New York).
SHEWMON, P. G., 1965, TMS AIME **233**, 736.
SIMS, C. T., and W. C. HAGEL, 1972, The Superalloys (Wiley, New York).
SINCLAIR, R., R. GRONSKY and G. THOMAS, 1976, Acta Metall. **24**, 789.
SMITH, C. S., 1962, in: Decomposition of Austenite by Diffusional Processes, eds. V. F. Zackay and H. I. Aaronson (Interscience, New York) p. 237.
SMITH, G. D. W., A. J. GARRATT–REED and J. B. VANDERSANDE, 1981, in: Quantitative Analysis with High Spatial Resolution, ed. G. W. Lorimer (The Metals Society, London) p. 238.
SOFFA, W. A., and D. E. LAUGHLIN, 1983, in: Solid-State Phase Transformations, the Pittsburg Conference, ed. H. I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 159.
SOFFA, W. A., S. S. BRENNER and M. K. MILLER, 1984, in Decomposition of Alloys: The Early Stages, eds. P. Haasen, V. Gerold, R. Wagner and M. F. Ashby. Pergammon Press Oxford p. 227.
SPEICH, G. R., 1968, TMS AIME **242**, 1359.
SPEICH, G. R., and R. A. ORIANI, 1965, TMS AIME **233**, 623.

SRINIVASAN, G. R., and C. M. WAYMAN, 1968, Acta Metall. **16**, 609 and 621.
STARK, I., G. D. W. SMITH and H. K. D. H. BHADESHIA, 1990, Metallurg. Trans. **21A**, 831.
STEPHENS, D. E., and G. R. PURDY, 1975, Acta Metall. **23**, 1343.
STRUTT, J., 1974, D. Phil thesis, Univ. of Sussex.
SULONEN, M. S., 1964a, Acta Metall. **12**, 749.
SULONEN, M. S., 1964b, Acta Polytech. Scand. **28**, 3.
SUNDQUIST, B. E., 1973, Metallurg. Trans. **4**, 1919.
SWALIN, R. A., 1972, Thermodynamics of Solids, 2nd Ed. (Wiley, New York).
SWANSON, W. D., and J. G. PARR, 1995, J. Iron and Steel Inst. **204** 104.
TANNER, L., and I. S. SERVI, 1966, Acta Metall. **14**, 231.
TIEN, J. K., P. G. SHEWMON and J. S. FOSTER, 1971, Metallurg. Trans. **2**, 1193.
TIEN, J. K., P. G. SHEWMON and J. S. FOSTER, 1973, Metallurg. Trans. **4**, 370.
TILLER, W. A., K. A. JACKSON, J. W. RUTTER and B. CHALMERS, 1953, Acta Metall. **1**, 428.
TRIVEDI, R., 1970a, Metallurg. Trans. **1**, 921.
TRIVEDI, R., 1970b, Acta Metall. **18**, 287.
TOWNSEND, R. D., and J. D. KIRKALDY, 1968, Trans ASM. **61**, 605.
TU, K. N., 1972, Metallurg. Trans. **3**, 2769.
TURNBULL, D., 1955, Acta Metall. **3**, 55.
TURNBULL, D., 1956, in: Solid State Physics, eds F. Seitz and D. Turnbull (Academic, New York) vol. 3, p. 226.
TURNBULL, D., and J. C. FISHER, 1949, J. Chem. Phys. **17**, 71.
TZAKALAKOS, T., 1977, Ph.D. thesis, Northwestern Univ.
TZAKALAKOS, T., and J. E. HILLIARD, 1980, quoted by DITCHEK and SCHWARTZ [1980].
VAN HORN, K., ed., 1967, Aluminum, vols. 1–3 (ASM, Metals Park, OH).
VASUDEVAN, A. .K., and R. D. Doherty, 1987, Acta Metall. **35**, 1193.
VASUDEVAN, A. K., and R. D. DOHERTY (Eds.), 1989, "Aluminum Alloys — Contemporary Research and Applications", Treatise on Materials Science and Technology, Vol. 31, Academic Press.
VOGEL, A., 1977, D. Phil thesis, Univ. of Sussex.
VOLMER, M., and A. WEBER, 1926, Z. Phys. Chem. **119**, 277.
VOLMER, M., and H. FLOOD, 1934, Z. Phys. Chem. **170**, 273.
VON ALVENSLEBEN, L., 1982, Dipl. thesis, Göttingen.
VON ALVENSLEBEN, L., and R. WAGNER, 1984, in Decomposition of Alloys: The Early Stages, eds. P. Haasen, V. Gerold, R. Wagner and M. F. Ashby. Pergammon Press Oxford, p. 143.
VOORHEES, P. W., and M. E. GLICKSMAN, 1984a, Metal. Trans. **15A**, 1081.
VOORHEES, P. W., and M. E. GLICKSMAN, 1984b, Acta Metall. **32**, 2013.
VOORHEES, P. W., and M. E. GLICKSMAN, 1984c, Acta Metall. **32**, 2013??.
WAGNER, C., 1961, Z. Elektrochem. **65**, 581.
WATSON, J. D., and P. G. MCDOUGAL, 1973, Acta Metall. **21**, 961.
WEATHERLY, G. C., 1971, Acta Metall. **19**, 181.
WEATHERLY, G. C., 1, and C. M. SARGENT, 1970, Phil. Mag. **22**, 1049.
WEAVER, L., B. HUDSON and J. NUTTING, 1978, Met. Sci. **12**, 257.
WENDT, H., 1981, Ph.D. thesis, Univ. of Göttingen.
WENDT, H., and P. HAASEN, 1983, Acta Metall. **31**, 1649.
WENDT, H., and P. HAASEN, 1985, Scripta Metall. **19**, 1053.
WERT, J. A., 1976, Acta Metall. **24**, 65.
WESTMACOTT, K. H., and U. DAHMEN, 1984, in Decomposition of Alloys: The Early Stages, eds. P. Haasen, V. Gerold, R. Wagner and M. F. Ashby, Pergammon Press Oxford, p. 204.
WILLIAMS, D. B., and J. W. EDINGTON, 1975, Met. Sci. **9**, 529.
WILLIAMS, D. B., and J. W. EDINGTON, 1976, Acta Metall. **24**, 323.
WILLIAMS, J. C., 1978, in: Precipitation Processes in Solids, eds. K. C. Russell and H. I. Aaronson (Met. Soc. AIME, Warrendale, PA) p. 191.
WILLIAMS, P. R., M. K. MILLER, P. A. BEAVAN and G. D. W. SMITH, 1979, in: Phase Transformations, York Conference (The Institution of Metallurgists, London) p. II–98.

WILSON, E. A., S. M. C. VICKERS, C. QUIXALL and A. BRADSHAW, 1979, in: Phase Transformations, York Conference (The Institution of Metallurgists, London) p. II–67.
WU, C. K., R. SINCLAIR, and G. THOMAS, 1978, Metallurg. Trans. **9A**, 381.
XIAO, S. Q., and P. HAASEN, 1991, Acta Metall. and Mater. **39**, 691.
YANG, S. C., HIGGINS, G. T. and NASH, P., 1992, Mater. Sci. and Tech. **8**, 10.
YOON, D. N., and W. J. HUPPMANN, 1979, Acta Metall. **27**, 973.
YOON, D. N., 1989, Ann. Rev. Mater Sci. **19**, 43.
ZENER, C., 1946, TMS AIME **167**, 550.
ZETTLEMOYER, A. C., 1969, Nucleation (M. Dekker, New York).

Further reading

The Mechanism of Phase Transformations in Crystalline Solids (The Metals Society, London, 1969).
A. C. Zettlemoyer, Nucleation (M. Dekker, New York, 1969).
Phase Transformations (ASM, Metals Park, OH, 1970).
G. A. Chadwick, The Metallography of Phase Transformation (Butterworths, London, 1972).
H. Warlimont, ed., Order-Disorder Transformations in Solids (Springer, Berlin, 1974).
J. W. Christian, The Theory of Transformation in Metals and Alloys. 2nd Ed., 1975; also 1st Ed., 1965 (Pergamon Press, Oxford).
M. Hillert, Diffusion and Interface Control of Reactions in Alloys, Metallurg. Trans. 6A (1975) 5.
J. W. Martin, and R. D. Doherty, The Stability of Microstructure in Metallic Systems (Cambridge Univ. Press. 1976). New Edition in Press (with B. Cantor as co-author)
K. C. Russell, and H. I. Aaronson, (eds.), Precipitation Processes in Solids (Met. Soc. AIME, Warrendale, PA, 1978).
Phase Transformations, York Conference (The Institution of Metallurgists, London, 1979).
I. N. Easterling, and O. A. Porter, Phase Transformations in Metals and Alloys (Van Nostrand, London, 1981).
H. I. Aaronson, (ed.), Solid-State Phase Transformations, the Pittsburgh Conference (Met. Soc. AIME, Warrendale, PA, (1983).
J. E. Hatch, (ed) Aluminum: Properties and Physical Metallurgy, ASM (Metals Park, Ohio) (1984)
P. Haasen, V. Gerold, R. Wagner and M. F. Ashby, (eds.) Decomposition of Alloys: The Early Stages, Pergamon Press, Oxford. (1984)
E. A. Starke, and T. H. Sanders, (eds.) Aluminum Alloys: Their Physical and Mechanical Properties (ICAA 1) (EMAS, Warley, UK) (1986)
G. W. Lorimer, (ed.), Phase Transformations '87., The Institute of Metals, London. (1987)
A. K. Vasudevan, and R. D. Doherty. (eds.) "Aluminum Alloys — Contemporary Research and Applications" Treatise on Materials Science and Technology Vol 31. Academic Press 1989.
C. Q. Chen, and E. A. Starke (eds) Aluminum Alloys: Their Physical and Mechanical Properties (ICAA 2) International Academic Publishers (1990)
International conference on bainite, Metallurg. Trans., 21A 767-893 and 1343-1545, (1990).
H. D. K. H. Bhadeshia, Bainite in Steels (London, Institute of Materials, 1992).
P. Haasen, and R. Wagner, High-resolution Microscopy and Early-stage Precipitation Kinetics, Metallurg. Trans., (1992) 23A, 1901.
L. Arnberg, O. Lohne, E. Nes and N. Ryum (eds,); Aluminum Alloys: Their Physical and Mechanical Properties (ICAA 3) The Norwegian Institute of Technology, Trondheim, N7034, Norway (1992)
T. H. Sanders, and E. A. Starke (eds.), Aluminum Alloys: Their Physical and Mechanical Properties (ICAA 4) Georgia Institute of Technology School of Materials Science and Engineering Atlanta GA USA. (1994)
Role of shear and diffusion in plate shaped transformation products, Metallurg.Mater. Trans., 25A 1785-2016 and 2552-2673 (1994).
W. Soffa, and D. Laughlin (eds.), Phase Transformations 94, Met. Soc. AIME, Warrendale, PA, (1994).

CHAPTER 16

PHASE TRANSFORMATIONS, NONDIFFUSIVE

C. M. WAYMAN

Department of Metallurgy and Mining Engineering
and Materials Research Laboratory
University of Illinois at Urbana–Champaign
Urbana, IL 61801, USA

H. K. D. H. BHADESHIA

University of Cambridge
Department of Materials Science and Metallurgy
Pembroke Street,
Cambridge CB2 3QZ, U.K.

R. W. Cahn and P. Haasen†, eds.
Physical Metallurgy; fourth, revised and enhanced edition

1. Overview

This chapter deals with phase changes where there are no long range movements of atoms in the sense of, say, their diffusional flow down a concentration gradient according to Fick's Law. Indeed, in most cases to be considered the atoms move less than an interatomic distance and retain the relationship with their neighbors during the phase change. The prototype of this behavior is a martensitic transformation, which is equivalent to the deformation of the parent crystal lattice into that of the product (martensite). Because of their nature, martensitic transformations are frequently referred to as "displacive", "shear-like" or "diffusionless". A vast amount of geometrical or crystallographic information on martensitic transformations has accrued over the years, and we shall see that all martensitic transformations have certain common crystallographic characteristics. These characteristics help distinguish martensitic reactions and have led to a successful phenomenological theory of martensite crystallography, which will be discussed in this chapter, along with a consideration of the thermodynamics and kinetics of the transformation. As will be seen, there are other phase changes, which although not commonly recognized as martensitic (e.g., bainite formation), are described well by the martensite crystallography theory. Finally, some other types of diffusionless, displacive solid-state transformations such as omega phase and charge-density wave formation are considered, both of which involve lattice distortions with correlated atomic displacements.

2. Martensitic transformations

2.1. Introduction and general characteristics

The name *martensite* is after the German scientist Martens. It was used originally to describe the hard microconstituent found in quenched steels. Many materials other than steel are now known to exhibit the same type of solid-state phase transformation, known as a *martensitic transformation* frequently also called a *shear* or *displacive transformation*. Martensite occurs in, for example, nonferrous alloys, pure metals, ceramics, minerals, inorganic compounds, solidified gases and polymers (table 1).

It has long been recognized that martensitic transformations are diffusionless; this is why martensite can grow at temperatures below 100 K where the mobility of atoms is negligible. It is not implied that all martensitic transformations occur at low temperatures; indeed, many of them occur at comparatively high temperatures (table 1), but these too are diffusionless. Martensite can also grow at speeds in excess of 1000 m s^{-1}, speeds which are inconsistent with the diffusion of atoms. In most cases, the amount of martensite that is obtained depends on the transformation temperature rather than the time at that temperature, in which case the transformation is called *athermal*. As pointed out by CHRISTIAN [1970], the overall kinetics of a martensitic reaction depend on both the nucleation and growth stages and will largely be dominated by the slower of the two. For example, slow thermal nucleation may give rise to isothermal transformation characteristics. As will be seen later, the martensitic transformations do not obey classical nucleation theory.

Table 1
The temperature M_S at which martensite first forms on cooling, and the approximate Vicker's hardness of the resulting martensite for a number of materials.

Composition	M_S/K	Hardness HV
ZrO_2	1200	1000
Fe–31Ni–0.23C wt.%	83	300
Fe–34Ni–0.22C wt.%	< 4	250
Fe–3Mn–2Si–0.4C wt.%	493	600
Cu–15Al	253	200
Ar–40N_2	30	

The interface between the martensite and its parent has to have a structure which is consistent with the absence of diffusion and with the rapid rate of growth. A glissile interface like this does not require significant thermal activation in order to move, and must be coherent or semi-coherent, depending upon the crystallography of the particular material undergoing transformation. In most cases, e.g., the formation of martensite in ferrous alloys, the interface is semi-coherent and the parent and product lattices are coherently accommodated only over local regions of the boundary. The accumulating misfit is then relieved periodically by auxiliary deformation processes in such a way that the interface remains glissile. On the other hand, e.g., the fcc → hcp transformation in Co and its alloys, the martensite-parent interface is fully coherent, and the deformation it causes when it moves (normal to itself) is sufficient to accomplish the required lattice change. In contrast, the macroscopic deformation caused by the movement of a semi-coherent martensite interface is related in a more complex way to the strain which converts the parent to the product lattice. More discussion of interfaces will follow later, but for now it is emphasized that both semi-coherent and coherent martensite-parent interfaces must be glissile.

Martensitic transformations are most readily distinguished from other solid-state phase changes on the basis of their crystallographic characteristics, which imply a "military" or coordinated transfer of atoms from the parent to the product phase. In a "civilian" transformation the atoms move across the phase boundary in an uncoordinated manner by diffusion. They feature a coordinated structural change involving a lattice correspondence and an ideally planar parent-product interface which during movement (transformation) produces an invariant-plane strain shape deformation. A fine-scale inhomogeneity in the martensite, such as slip, twinning or faulting, is usually observed at the electron microscope scale. This secondary deformation, an intrinsic part of the transformation process, produces the invariant-plane condition at the macroscopic scale and provides a semi-coherent glissile interface between the martensite and the parent phase. Crystallographic features between the martensite and the parent phase such as the habit (invariant) plane and orientation relationship are usually not expressible in terms of exact relations involving integral Miller indices. The various crystallographic features of

References: p. 1552.

martensitic transformations will be discussed and illustrated with representative experimental examples. After this, a brief account of the development of the phenomenological crystallographic theory will be presented, followed by an algebraic analysis.

2.2. Experimental observations of crystallographic features

Figure 1 is an optical micrograph of a polycrystalline Fe–24.5Pt at.% alloy taken after transformation of the parent phase into martensite by cooling. The specimen was polished flat prior to transformation to martensite. The observed contrast results purely from deformation caused by the formation of martensite. The light and dark shades correspond to regions of martensite which have undergone distortion in different senses with respect to the initial surface. This surface tilting or macroscopic distortion is known as the *shape deformation* or *shape strain*. It can be demonstrated by serial sectioning that the martensite has a three-dimensional shape which is that of lenticular plates. An X-ray analysis would reveal that the plates showing different optical contrast also feature different lattice orientations with respect to the initial parent grain. X-ray analysis would also show that a structural change had occurred, which for this Fe–Pt alloy is an fcc → bcc transformation*. The differently oriented plates are different crystallographic variants of the habit plane and orientation relationship.

Referring again to fig. 1, if straight scratches were purposely abraded on a flat specimen still in its parent phase, then the formation of martensite would displace the

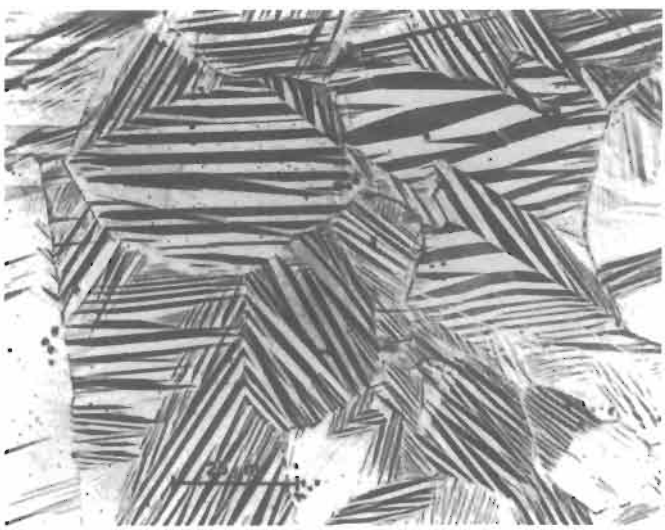


Fig. 1. Optical micrograph showing surface relief due to a martensitic transformation in an Fe–24.5% Pt alloy. The specimen was polished to a flat condition at room temperature and then cooled to produce the martensite.

* The terms fcc, bcc, bct and hcp are abbreviations for face-centered cubic, body-centered cubic, body-centered tetragonal and hexagonal close-packed, respectively.

scratches in a characteristic manner. After analyzing several nonparallel scratches crossing a given plate and noting the initial position of the same scratches in the untransformed parent phase, a distortion matrix which describes the shape deformation can be derived, which shows that a deformation which is homogeneous on an optical scale has taken place. This shape deformation resembles a simple shear, but in general is better described as an invariant-plane strain (fig. 2) in which one plane is left undistorted and unrotated; this invariant plane is also known as the *habit plane*.

An invariant-plane strain is a homogeneous distortion such that the displacement of any point is in a common direction. The magnitude of the displacement is proportional to the distance from a fixed plane of reference, which is the invariant-plane (itself unaffected by the strain). In most martensitic transformations a volume change accompanies the structural change which produces a normal component δ to the invariant-plane strain.

It is well established that the shape deformation caused by the growth of martensite is an invariant-plane strain. Nevertheless, it is also true that for many martensitic transformations, an invariant-plane strain cannot convert the crystal structure of the

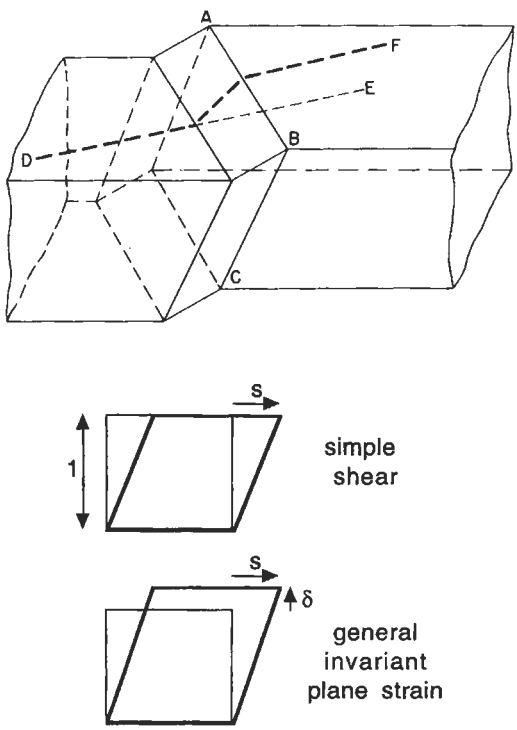


Fig. 2. (a) An invariant-plane strain characteristic of martensitic transformation. The initially straight line DE is displaced to the position DF when the martensite plate with habit plane ABC is formed. The plane ABC is invariant (unrotated and undistorted) as a result of the formation of martensite. (b) An illustration of the difference between a simple shear and a general invariant-plane strain. s is the shear component of the strain and δ the dilatational component of strain. The dilatational strain defines the volume change of transformation.

References: p. 1552.

parent into that of the product phase. In iron alloys the martensitic transformation converts an fcc parent into a bcc (or bct) product, but the measured shape deformation matrix when applied to the parent fcc will not produce a bcc structure. This apparent inconsistency will be explained later.

Martensite plates in a given alloy usually possess a unique habit plane, as shown for example in fig. 3. The stereographic projection shows the experimentally determined habit plane poles for different martensite plates, which, as can be seen, cluster near the plane in the parent phase with Miller indices $(3\ 10\ 15)_P$ (subscripts P and M, respectively, are used to designate the parent and martensite).

Early work suggested that, apart from a small relative rotation of corresponding unit cells in the parent and product phases, a homogeneous, lattice distortion would account for the known structural change in martensitic transformations. As early as 1924, BAIN suggested that the austenite (parent phase) → martensite transformation in steels could be explained by a homogeneous "upsetting" of the parent fcc lattice into the required bcc (or bct) lattice, as shown schematically in fig. 4. There are in fact many ways (correspondences) to generate a bcc product from an fcc parent by means of a homogeneous distortion, and another possible correspondence is shown in fig. 5, but analysis shows that the Bain deformation involves the smallest principal strains and should therefore be favoured. The lattice correspondence, a unique relationship between any lattice point in the initial lattice and the point it becomes in the final lattice, implied by the Bain distortion has also been verified experimentally using an ordered Fe_3Pt alloy which undergoes nominally a fcc to bcc transformation. By observing corresponding superlattice reflections in the parent and martensitic phases by means of transmission electron diffraction it was deduced that the Bain correspondence actually applies (TADAKI and SHIMIZU [1970]).

From the lattice correspondence shown in fig. 4 one would expect for example, $[0\,0\,1]_M \parallel [0\,0\,1]_P$, $[0\,1\,0]_M \parallel [1\,1\,0]_P$, $(1\,1\,2)_M \parallel (1\,0\,1)_P$, $(0\,1\,1)_M \parallel (1\,1\,1)_P$, etc. However, such exact parallelisms are not observed. For the Fe–Pt alloy shown in fig. 1 the following orientation relationship was observed:

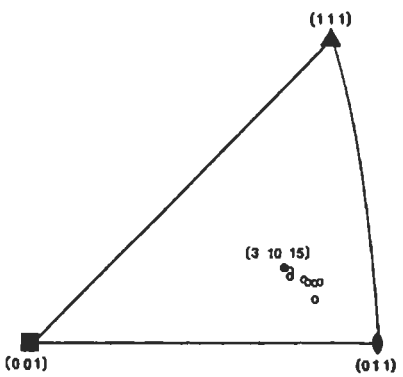


Fig. 3. Unit triangle of stereographic projection showing experimental habit planes (open circles) for martensite formed in an Fe–7Al–1.5C wt.% alloy. The average habit plane cannot be expressed in terms of simple Miller indices, and is near the plane $\{3\ 10\ 15\}_P$ of the parent phase.

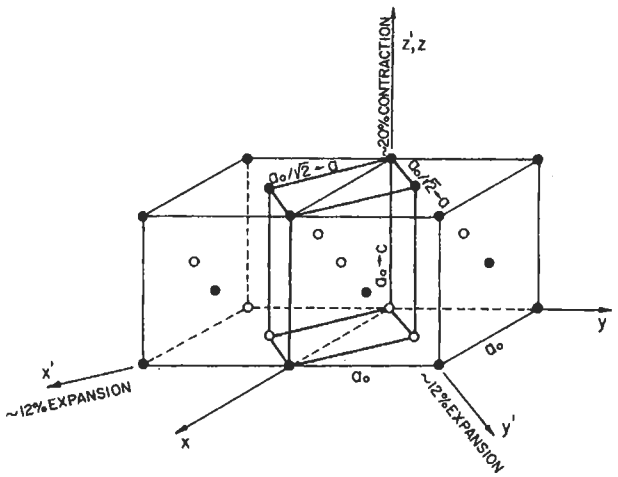


Fig. 4. Lattice distortion and correspondence proposed by BAIN [1924] for the fcc–bcc (bct) martensitic transformation in iron alloys. The correspondence related cell in the parent phase (indicated by bold lines) becomes a unit cell in the martensite as a consequence of a homogeneous "upsetting" with respect to the z' axis.

$[0\,0\,1]_P - [0\,0\,1]_M$: 9.10° apart,
$[\bar{1}\,0\,1]_P - [\bar{1}\,\bar{1}\,1]_M$: 4.42° apart,
$(1\,1\,1)_P - (0\,1\,1)_M$: 0.86° apart,

which is typical of the orientation relationships found in iron alloys and steels. Further consideration of the above orientation relationship shows that the correspondence cell is not only distorted according to the Bain strain, but also is rotated (about 10° from $[0\,0\,1]_P$ towards $[1\,1\,0]_P$). This is termed a *rigid body rotation*. The combined distortion–rotation

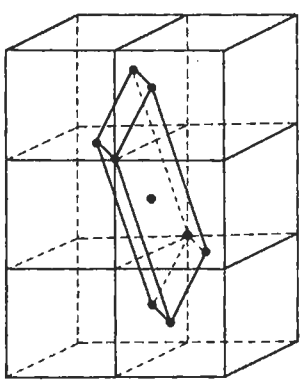


Fig. 5. Alternative lattice correspondence for the fcc–bcc (bct) martensitic transformation which involves larger principal distortions than the Bain correspondence.

References: p. 1552.

is known as the *lattice deformation*. Additional examination of the above orientation relationship shows that the close-packed planes $(1\,1\,1)_P$ and $(0\,1\,1)_M$ are nearly parallel to each other and the close-packed directions $[\bar{1}\,0\,1]_P$ and $[\bar{1}\,\bar{1}\,1]_M$ are a few degrees apart. This is an example of the *Greninger–Troiano* orientation relationship. When the close-packed planes and directions in two coexisting structures are almost parallel, the *Kurdjumov–Sachs* orientation relationship is obtained.

In summary, the Bain strain correctly transforms the crystal structure of the austenite into that of martensite, and when combined with an appropriate rigid body rotation leads to the correct orientation relationship. In fact, the combined lattice deformation (Bain + rotation) is an *invariant-line strain* which leaves a single line unrotated and undistorted. This line lies in the interface between the austenite and martensite, and permits the interfacial structure to be semi-coherent and glissile. Hence, two crystals can only transform into one another by a martensitic mechanism if they can be related by a lattice deformation which is at the very least an invariant-line strain.

2.3. The phenomenological crystallographic theory of martensitic transformations

We have seen that the Bain strain (**B**) can transform the parent into the product crystal structure with minimal atomic displacements. In itself the Bain strain alters every vector, but it can in combination with a rigid body rotation (**R**) be converted into a *lattice deformation* which is an invariant-line strain (**RB**) which also yields the experimentally observed orientation relationship. However, this invariant-line strain is inconsistent with the experimentally established invariant-plane strain shape deformation ($\mathbf{P}_1$).

The reason for this is as follows. Referring to fig. 4 again, it is noted that the correspondence cell is contracted (~20%) along the z' axis and expanded (~12%) along the x' and y' axes. Such a homogeneous distortion will leave no plane invariant (i.e., undistorted and unrotated). But suppose that the distortion along y' vanished. One of the principal distortions would then be less than unity (along z'), one greater than unity (along x') and the remaining one (along y') exactly unity. These conditions imply the distortion of an initial sphere (parent phase) into a triaxial ellipsoid (martensite) following which the sphere and ellipsoid can fit together along an undistorted plane of contact (habit plane) as shown in fig. 6. But this special set of conditions is not generally found in practice; the principal distortions are determined by the lattice correspondence and observed lattice parameters of the two phases.

The apparent inconsistency described above is resolved by envisioning a *lattice-invariant deformation* $\bar{\mathbf{P}}$ involving no structural change, such as slip or twinning to occur in conjunction with the lattice deformation **RB**. The additional deformation must be lattice-invariant because the necessary structural change is effected by the Bain distortion alone. The role of the additional deformation is essentially to shear (distort) the ellipsoid resulting from the Bain distortion into tangency with the initial sphere; then one of the principal distortions (OX in fig. 6) becomes unity and an undistorted contact plane exists. This additional deformation (slip, twinning, or faulting) being a shear is known as the *inhomogeneous shear* or *complementary shear* of the crystallographic theory.

These are the three phenomenological steps **B**, **R**, $\bar{\mathbf{P}}$ describing the total transform-

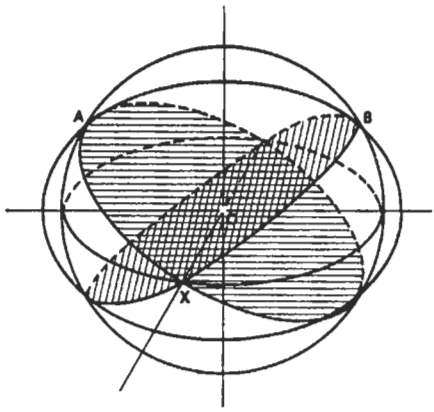


Fig. 6. The Bain distortion has a mathematical analog wherein an initial unit sphere is distorted into an ellipsoid. If one of the principal distortions is unity and the other two are of opposite sign, then the sphere and ellipsoid can fit together along the undistorted planes AOX or BOX. See text for discussion.

ation crystallography. There is no time sequence implied as to which step occurs when. Of course the combined effect of these three operations must be equivalent to the shape deformation.

Working with the theory just described, different crystallographic features such as the habit plane and orientation relationship can be predicted by supposing the inhomogeneous shear to occur on different crystallographic planes and directions. For example, in most iron alloys an inhomogeneous shear on the $(1\,1\,2)$–$[\bar{1}\,\bar{1}\,1]_M$ twinning system will predict a habit plane near $(3\,15\,10)_P$, but assuming the inhomogeneous shear system to be $(0\,1\,1)$–$[\bar{1}\,\bar{1}\,1]_M$ predicts a habit plane near $(1\,1\,1)_P$. Since the lattice parameters and correspondence are usually known, the Bain distortion for a given transformation is specified, and the flexibility in the theory comes through different suppositions concerning the plane and direction of the inhomogeneous shear. Once the inhomogeneous shear system is assumed, a shear of a certain magnitude will produce an undistorted plane, which when rigidly rotated to its original position becomes the invariant habit plane.

The previous discussion and examples cited above have centered around iron alloys. Because of the importance of steels there has been substantial work on them and thus an abundance of experimental data exists. However, the principles presented are quite general and apply to all martensitic transformations.

2.3.1. Summary of crystallographic theory

The Bain strain converts the structure of the parent phase into that of the product phase. When combined with an appropriate rigid body rotation, the net homogeneous lattice deformation **RB** is an invariant-line strain (step *a* to *c* in fig. 7). However, the observed shape deformation is an invariant-plane strain $\mathbf{P}_1$ (step *a* to *b* in fig. 7), but this gives the wrong crystal structure. If, however, a second homogeneous shear $\mathbf{P}_2$ is combined with $\mathbf{P}_1$ (step *b* to *c*), then the correct structure is obtained but the wrong shape since

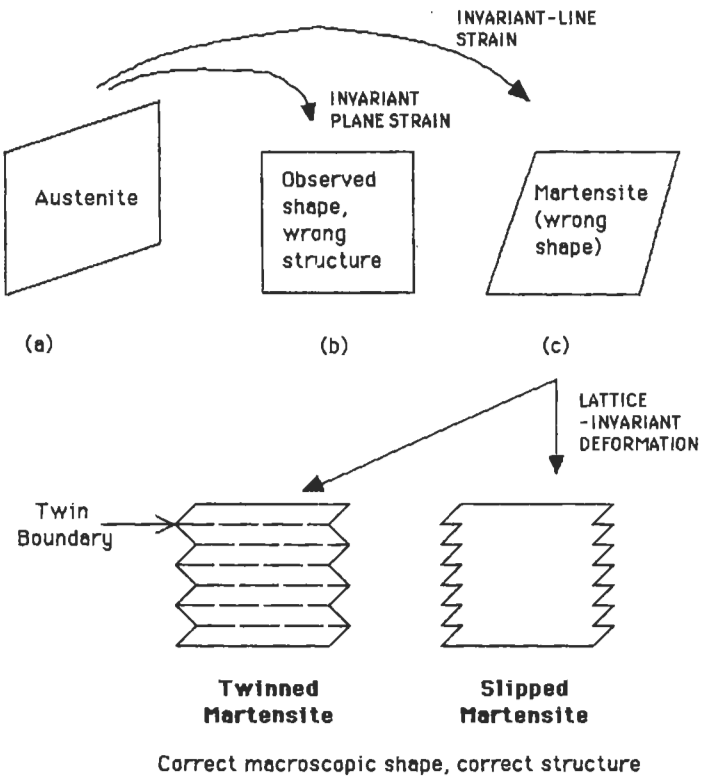


Fig. 7. Schematic illustration of the phenomenological theory of martensite crystallography.

$$\mathbf{P}_1\mathbf{P}_2 = \mathbf{RB}$$

These discrepancies are all resolved if the shape changing effect of $\mathbf{P}_2$ is cancelled macroscopically by an inhomogeneous lattice-invariant deformation, which may be slip or twinning as illustrated in fig. 7.

The theory illustrated elegantly explains all the observed features of the martensite crystallography. It is easy to predict the orientation relationship, by deducing the Bain strain and adding a rigid body rotation which makes the net lattice deformation an invariant-line strain. The habit plane does not have rational indices because the amount of lattice-invariant deformation needed to recover the correct the macroscopic shape is not usually rational. The theory predicts a substructure in plates of martensite (either twins or slip steps) as is observed experimentally. The transformation goes to all the trouble of ensuring that the shape deformation is macroscopically an invariant-plane strain because this reduces the strain energy when compared with the case where the shape deformation might be an invariant-line strain.

We now proceed to discuss the details of this theory.

2.3.2. The inhomogeneous shear and martensite substructure

The inhomogeneous shear was introduced in the crystallographic description of martensite transformations to ensure that the habit plane is macroscopically undistorted. Figure 8 is a schematic representation of the appearance of internally twinned and internally slipped martensite plates. Although there are localized distortions at the interface, the "saw tooth" effect because of alternating twins (called transformation twins) or slip lamellae prevents accumulation of any strain at the interface over large distances.

For internally twinned martensite the Bain distortion is envisioned to occur along different contraction axes in the two twin-related regions, and the twinning plane in the martensite is derived from a mirror plane in the parent. In the case of internally slipped martensite the Bain distortion is the same in all regions of a plate. Figure 9 shows that the same effective shear angle γ can be accomplished by slip or twinning. Note that the relative thickness of the two twin components determines the angle γ.

Because of the inhomogeneous shear mentioned above, one would expect to observe some kind of substructure in the martensite. Such observations have been made since the introduction of the theory. Figure 10 is a transmission electron micrograph of a martensite plate in an Fe–Ni–C alloy. The fine striations crossing the plate are transformation twins ($\{1\,1\,2\}_M$ twinning plane). Regions adjacent to the martensite are retained austenite. If the twin plane is indexed specifically as $(1\,1\,2)_M$, the habit plane trace becomes specifically $(3\,15\,10)_P$. If $(1\,1\,2)$–$[\bar{1}\,\bar{1}\,1]_M$ twinning is used in the habit plane calculations, the predicted habit plane is in fact $(3\,15\,10)_P$. Thus the particular variant of the twin plane is found to be consistent with the particular variant of the habit plane, and for the Fe–Ni–C alloy the experimental observations are in excellent agreement with those features which are predicted using the phenomenological crystallographic theory. This is also the case for many other martensitic transformations which have been studied in some detail using transmission electron microscopy and diffraction. Notable exceptions,

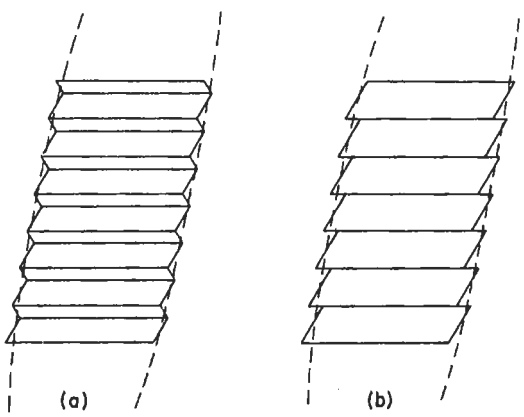


Fig. 8. Schematic representations of the inhomogeneous shear in a martensitic transformation involving (a) internally twinned. and (b) internally slipped plates of martensite. The serrated effect at the interface (habit plane) in each case prevents the long-range accumulation of strain and consequently the interface remains macroscopically undistorted.

References: p. 1552.

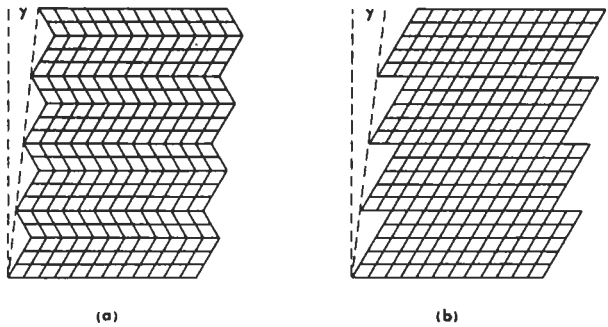


Fig. 9. As shown for twinning (a) and slip (b) the same magnitude of the inhomogeneous shear, as given by the angle γ, can be accomplished by either.

however, are certain steels which transform to martensite with a $\{2\,2\,5\}_p$ habit plane. The martensite in these materials has a complex substructure, as seen in the electron microscope, and is not very well explained using the theory described above. In addition, there is no adequate crystallographic description of the lath martensites found in ferrous alloys.

2.3.3. Mathematical description of the phenomenological crystallographic theory

The basic equation of the crystallography theory just described is:

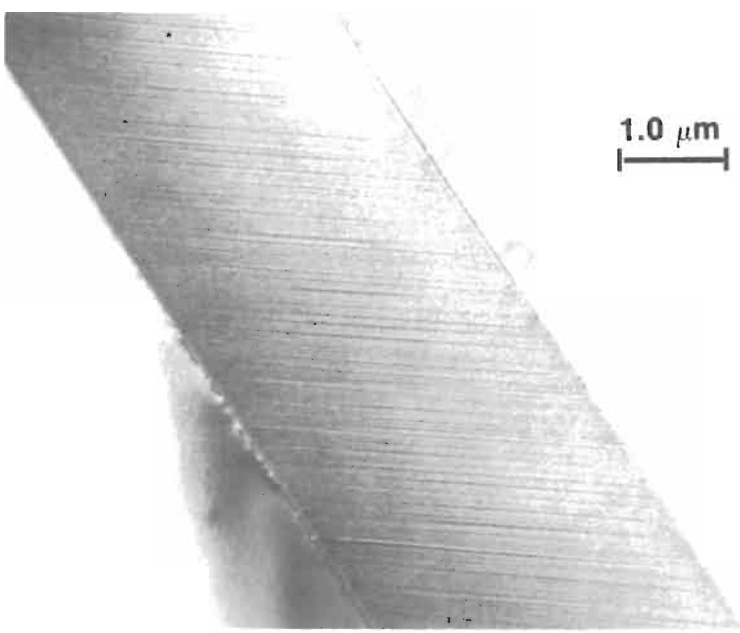


Fig. 10. Transmission electron micrograph showing a martensite plate and adjacent retained austensite in an Fe–30% Ni–0.4% C alloy. The striations within the plate running from left to right are transformation twins.

$$\mathbf{P}_1 = \mathbf{R}\bar{\mathbf{P}}\mathbf{B} \tag{1}$$

where **B** represents the Bain distortion, $\bar{\mathbf{P}}$ is a simple shear (following in a mathematical sense the Bain distortion), **R** is the rigid body rotation previously mentioned, and $\mathbf{P}_1$ is the invariant-plane strain shape deformation. **B**, $\bar{\mathbf{P}}$, **R** and $\mathbf{P}_1$ can all be represented as (3×3) deformation matrices. A vector, on the other hand, can be represented by a 1×3 single-row matrix or a 3×1 single-column matrix. The product of a deformation matrix with a vector gives a new vector which may in general be of a different length and direction.

The matrix product $\mathbf{R}\bar{\mathbf{P}}\mathbf{B}$ is equivalent to the shape deformation $\mathbf{P}_1$ and the rotation matrix **R** rotates the plane left undistorted by $\bar{\mathbf{P}}\mathbf{B}$ to its original position. That is, $\mathbf{P}_1 = \mathbf{R}\bar{\mathbf{P}}\mathbf{B}$ is an invariant-plane strain.

Although eq. 1 indicates that the inhomogeneous shear $\bar{\mathbf{P}}$ follows the Bain distortion, the same end result can be obtained by "allowing" the shear to occur in the parent phase prior to the Bain distortion. This alternative procedure simplifies computations. In this case the basic equation becomes:

$$\mathbf{P}_1 = \mathbf{RBP} \tag{2}$$

where **P** as before is a simple shear.

The matrix representation of the invariant-plane strain shape deformation is:

$$\begin{aligned}\mathbf{P}_1 &= \mathbf{I} + m\mathbf{dp}' \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + m\begin{bmatrix} d_1 & d_2 & d_3 \end{bmatrix}\begin{pmatrix} p_1 & p_2 & p_3 \end{pmatrix} \\ &= \begin{pmatrix} 1 + md_1p_1 & md_1p_2 & md_1p_3 \\ md_2p_1 & 1 + md_2p_2 & md_2p_3 \\ md_3p_1 & md_3p_2 & 1 + md_3p_3 \end{pmatrix}\end{aligned} \tag{3}$$

where $\mathbf{p}'$ (prime meaning transpose) being a plane normal is written as a (1×3) single-row matrix, in contrast to **d**, which is a vector written as a (3×1) single-column matrix.

The Bain distortion illustrated in fig. 4, for the fcc $\rightarrow$ bcc (or bct) transformation can be represented as:

$$\mathbf{B} = \begin{pmatrix} \sqrt{2}a/a_0 & 0 & 0 \\ 0 & \sqrt{2}a/a_0 & 0 \\ 0 & 0 & c/a_0 \end{pmatrix} \tag{4}$$

For typical values of the lattice parameters of the fcc lattice, a, and of the bct lattice, a_0 and c, the matrix **B** becomes

References: p. 1552.

$$\mathbf{B} = \begin{pmatrix} 1.12 & 0 & 0 \\ 0 & 1.12 & 0 \\ 0 & 0 & 0.8 \end{pmatrix} \tag{5}$$

It is apparent from eq. 5 that two of the principal distortions are greater than unity (i.e., 1.12) and the third is less than unity (i.e., 0.8). Therefore, the Bain distortion cannot leave a plane invariant; as stated earlier, for a strain to be an invariant-plane strain, one of the principal distortions must be zero, one greater than unity and the other less than unity.

Referring back to eq. 2, it is noted that **P** is a simple shear, and therefore can be represented as $(\mathbf{I}+m\mathbf{dp}')$. Also, the inverse of **P**, $\mathbf{P}^{-1}=(\mathbf{I}-m\mathbf{dp}')$ is a simple shear of the same magnitude, on the same plane, but in the opposite direction. That is, both **P** and $\mathbf{P}^{-1}$ are invariant-plane strains. It is now convenient to rewrite eq. 2 as:

$$\mathbf{P}_1\mathbf{P}_2 = \mathbf{RB}, \tag{6}$$

where $\mathbf{P}_2=\mathbf{P}^{-1}$. Since both $\mathbf{P}_1$ and $\mathbf{P}_2$ are invariant-plane strains, their product **RB** is an invariant-*line* strain **S**. The invariant-line is defined as the line of intersection of the *planes* which are invariant to $\mathbf{P}_1$ and $\mathbf{P}_2$. If the invariant-line strain **S** is known, all the crystallographic features of a given martensitic transformation can be predicted. It is not within the scope of the present account to go into the details of the invariant-line strain analysis (BOWLES and MACKENZIE [1954]), but certain highlights should be mentioned. The Bain correspondence and distortion **B** are known from the lattice parameters of the parent and martensite phases. **R** is determined once the plane $\mathbf{p}_2'$ and the direction $\mathbf{d}_2$ of $\mathbf{P}_2$ are assumed. It should be noted that $\mathbf{RB}=\mathbf{S}$ constitutes the lattice deformation of the transformation.

The important results of the invariant-line strain analysis, noting that the shape strain is $\mathbf{P}_1=\mathbf{I}+m_1\mathbf{d}_1\mathbf{p}_1'$ and that the simple shear (mathematically preceding the Bain distortion) is $\mathbf{P}_2=\mathbf{I}+m_2\mathbf{d}_2\mathbf{p}_2'$ are as follows (where the magnitudes, directions and planes of the component invariant-plane strains are given respectively by m, **d** and **p**′):

$$\mathbf{d}_1 = \left[\mathbf{S}\mathbf{y}_2 - \mathbf{y}_2\right]/\mathbf{p}_1'\mathbf{y}_2, \tag{7}$$

$$\mathbf{p}_1' = \left(\mathbf{q}_2' - \mathbf{q}_2'\mathbf{S}^{-1}\right)/\mathbf{q}_2'\mathbf{S}^{-1}\mathbf{d}_1, \tag{8}$$

where $\mathbf{y}_2$ is any vector lying in $\mathbf{p}_2'$ (except the invariant line **x**) and $\mathbf{q}_2'$ is any normal (other than **n**′, the row eigenvector of $\mathbf{S}^{-1}$, i.e., $\mathbf{n}'\mathbf{S}^{-1}=\mathbf{n}'$) to a plane containing $\mathbf{d}_2$. The normalization factor for $\mathbf{d}_1$ in eq. 7 is $1/m_1$ and therefore $\mathbf{P}_1$, m_1, $\mathbf{d}_1$ and $\mathbf{p}_1'$ are all determinable. The matrix **R** is determined from the requirement that **x** and **n**′ which are displaced by the Bain distortion must be totally invariant. **R** defines the orientation relationship within any small region of the martensite plate not involving $\mathbf{P}_2$. Thus, the assumed correspondence and lattice parameters determine **B**, the assumption of $\mathbf{p}_2'$ and $\mathbf{d}_2$ allows **R** to be determined, and $\mathbf{RB}=\mathbf{S}$ defines the elements of $\mathbf{P}_1$.

The above description of the crystallographic theory parallels the analysis given by BOWLES and MACKENZIE [1954] but the treatments of WECHSLER *et al.* [1953] and

BULLOGH and BILBY [1956] are equivalent.

Some variations in the basic theory just presented include the introduction of a dilatation parameter, which relaxes slightly the requirement for the habit plane to be undistorted, and the incorporation of two inhomogeneous shear systems such that

$$\mathbf{P}_1 = \mathbf{RBS}_2\mathbf{S}_1 \tag{9}$$

where $\mathbf{S}_2$ and $\mathbf{S}_1$ are the two inhomogeneous shear systems involved. Neither of these modified approaches is without criticism, and it has been argued that the double shear approach loses generality.

2.3.4. Some other crystallographic observations

As pointed out earlier in this discussion, the martensite habit plane is near $\{3\,10\,15\}_P$ for an Fe–24.5Pt at.% alloy. A similar habit plane is also found for martensite formed in Fe–Ni alloys containing approximately 30 wt.% Ni. However, in some steels, e.g., Fe–8Cr-1.1C wt.%, the habit plane is near $\{2\,2\,5\}_P$, and the habit plane of lath martensites typical of low-carbon steels and dilute iron alloys is near $\{1\,1\,1\}_P$. However, the martensite-parent orientation relationship is almost the same in the {3 10 15}, {2 2 5} and {1 1 1} transformations.

It is unclear why different steels transform to martensite with a different habit plane, but it is generally thought that there are different operative modes of the inhomogeneous shear, i.e., different planes and directions, or even the operation of two or more modes within a given martensite plate or lath as suggested by the multiple shear theories. Figure 11 shows some experimental habit plane determinations for a variety of iron alloys.

Techniques are improving for measuring the shape strain associated with martensitic transformations and the information so obtained is leading to a clearer understanding of

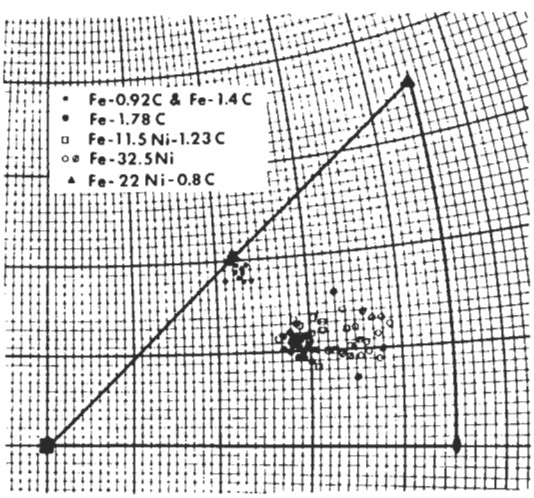


Fig. 11. Unit triangle of a stereographic projection showing the results of habit plane measurements on martensite formed in five different Fe-alloys.

References: p. 1552.

the transformation crystallography. In addition, the use of transmission electron microscopy has greatly enhanced the understanding of the crystallography of martensitic transformations.

2.3.5. Further observations on the martensite morphology and substructure

In general, martensite in ferrous alloys takes the form of lenticular plates as seen in fig. 1 for an Fe–Pt alloy. This morphology is readily understood because a lenticular plate, as a mechanical twin, is a low strain energy shape in a "shear" transformation. However, other morphologies are not uncommonly found in ferrous alloys, and these are briefly mentioned here.

In low-carbon steels (up to 0.4 wt.% C) and in some other ferrous alloys the martensite takes the form of laths rather than plates. A typical lath has dimensions $0.3 \times 4 \times 200$ μm. As mentioned above, the habit plane of such laths is near $\{1\,1\,1\}_P$. In Fe–Ni–C alloys, three distinct morphologies are found and are classified as the lenticular, thin plate, and butterfly morphologies. In one well-studied case it has been shown that butterfly martensite (fig. 16, below) forms at the highest temperatures, lenticular martensite forms at intermediate temperatures, and thin-plate martensite forms at the lowest temperatures. In some materials, particularly ferrous alloys, a needle-like morphology known as surface martensite is formed under some conditions. Apparently this morphology is easily nucleated at surfaces and can develop into a "needle" form because matrix constraints are relaxed near a free surface.

Whereas the plate types of ferrous martensite frequently form in a self-accommodating manner involving several habit plane variants, lath martensites do not appear to do so. Instead, the laths tend to cluster together into packets, each adjacent lath having the same habit plane, orientation relationship, and shape deformation. The intrinsic lath morphology (as opposed to plate morphology) is thought to be related to the larger shape strain associated with lath martensite. Unlike the plate morphology which typically contains a high density of transformation twins as shown in fig. 10, martensite laths are typically untwinned and contain a high density of internal dislocations as shown in fig. 12. Since the inhomogeneous shear in ferrous lath martensite is not twinning, it is expected that slip is the mode involved and that the lattice-invariant deformation would be effected by the movement of interface dislocations. An example of such interface dislocations in an Fe–Ni–Mn martensite lath is shown in fig. 13. The high density of dislocations usually observed within lath martensites may not be directly related to the transformation crystallography. It may result from the constraints a lath experiences during growth, i.e., the internal dislocations probably arise from accommodation distortion.

Another martensite morphology found in certain steels, notably austenitic stainless steels and Fe–Mn–C alloys, is the banded morphology (fig. 18, below). However in these alloys, the transformation is not fcc $\rightarrow$ bcc, but fcc $\rightarrow$ hcp instead. This transformation is accomplished by distorting the fcc parent phase on $\{1\,1\,1\}_P$ planes by the passage of partial dislocations which generate the hcp martensite. As would be expected, the bands delineate $\{1\,1\,1\}_P$ planes. This type of fcc $\rightarrow$ hcp martensite transformation is similar to that which occurs in cobalt and cobalt alloys, and, as mentioned earlier, involves a fully coherent interface and no lattice-invariant deformation in the usual sense.

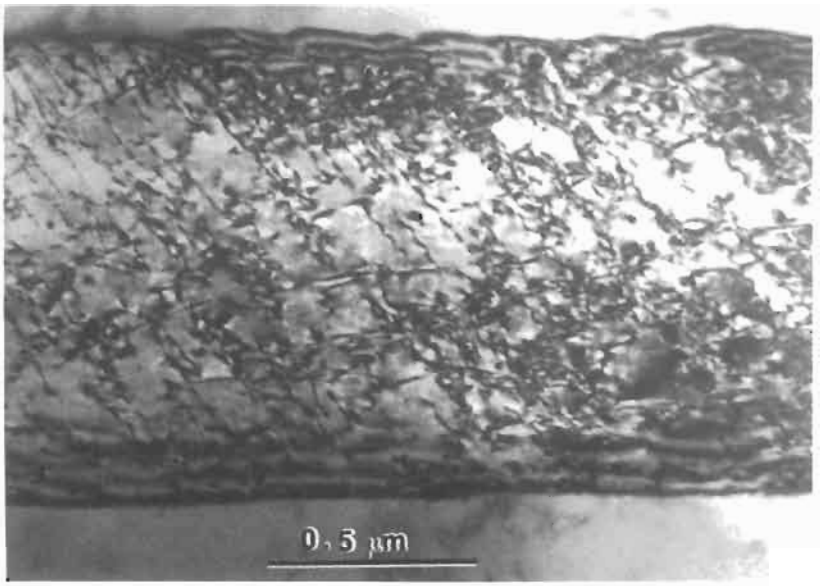


Fig. 12. Transmission electron micrograph showing dislocations (rather than twins) in a martensite lath formed in an Fe–20% Ni–5% Mn alloy.

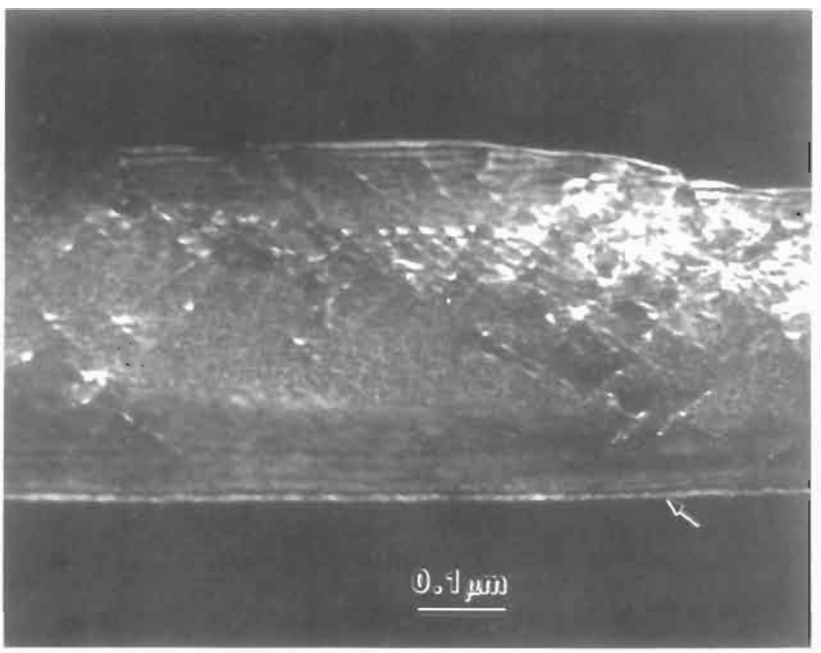


Fig. 13. Transmission electron mmicrograph showing interface dislocations (parallel to the arrow) in a lath of martensite in an Fe–20% Ni–5% Mn alloy.

References: p. 1552.

Figures 14–18 are optical micrographs showing typical martensite morphologies in ferrous alloys.

2.4. Martensite–parent interfaces

The phenomenological theory of martensitic transformations permits all the crystallographic features of the parent and product phases to be mathematically related, but does not yield detailed information about the mechanism of transformation. The latter depends critically on the structure of the transformation interface.

The fact that martensite can grow at very low temperatures and very high velocities means that the transformation interface must be very mobile and be able to move without any need for diffusion. The interface must be *glissile*. A fully coherent interface is glissile, but since the parent and martensite crystals cannot in many cases be converted into one another by a lattice deformation which is an invariant-plane strain, the interfacial structure is more likely to be semi-coherent.

A semi-coherent interface consists of coherent regions separated by discontinuities which prevent the misfit in the interface plane from accumulating over large distances. There are two kinds of semi-coherency [CHRISTIAN and CROCKER, 1980]; if the discontinuities mentioned above are intrinsic dislocations with Burgers vectors in the interface plane, not parallel to the dislocation line, then the interface is said to be epitaxially semi-coherent. The normal displacement of such an interface requires the thermally activated climb of the intrinsic dislocations. A martensite interface cannot therefore be epitaxially semi-coherent.

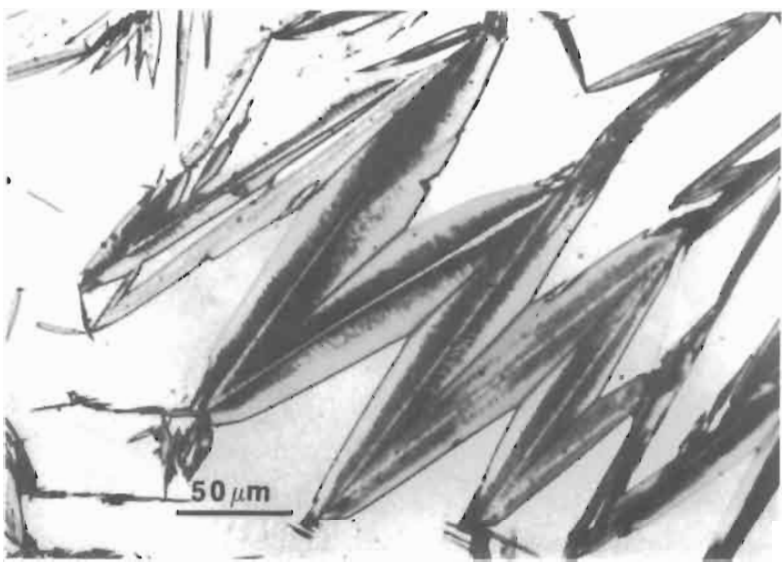


Fig. 14. Optical micrograph showing lenticular martensite formed at –79°C in an Fe–31Ni–0.23C wt.% alloy. The more heavily etched linear central region of the plates is known as a midrib.

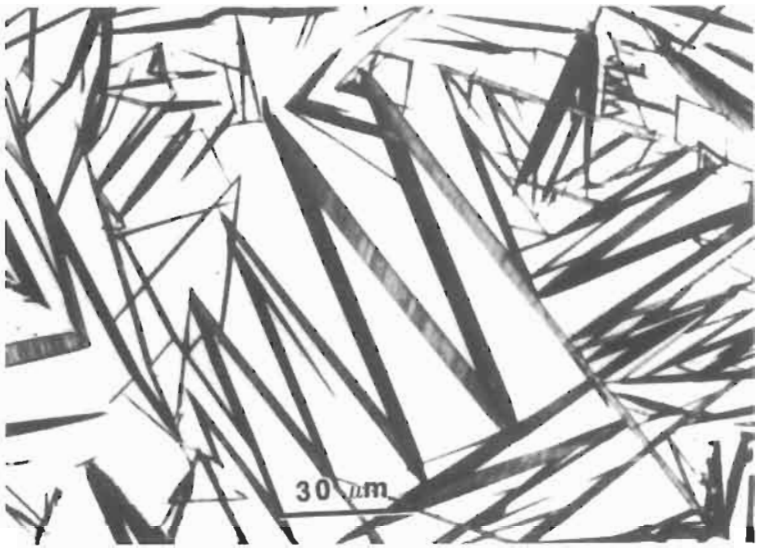


Fig. 15. Optical micrograph showing the "thin-plate" morphology of martensite, observed in an Fe–31Ni–0.29C wt.% alloy transformed at –160°C.

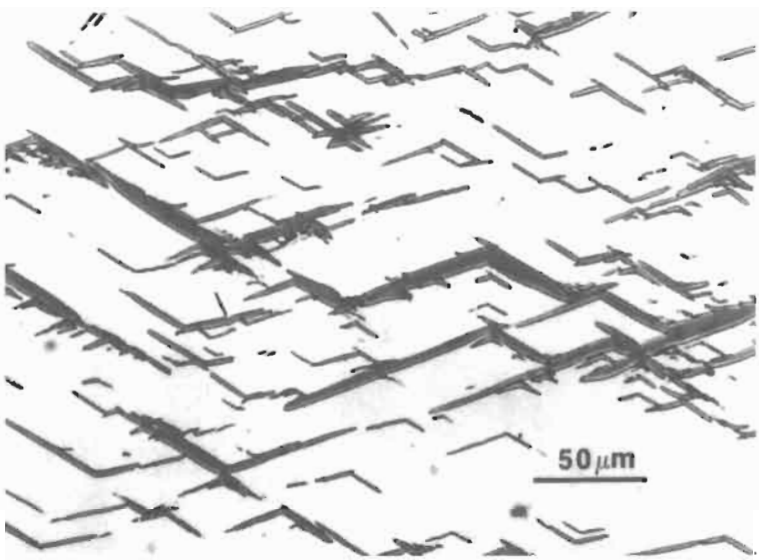


Fig. 16. Optical micrograph showing butterfly martensite morphology formed at 0°C in an Fe–20Ni–0.7C wt.% alloy.

References: p. 1552.

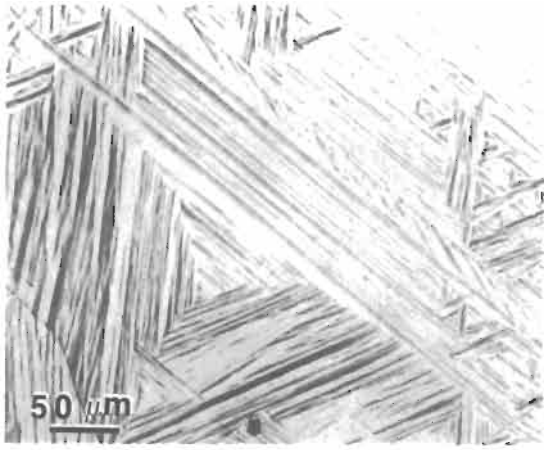


Fig. 17. Optical micrograph showing surface relief following the formation of lath martensite in an Fe–0.2C wt.% alloy.

In the second type of semi-coherency the discontinuities are screw dislocations, or dislocations whose Burgers vectors do not lie in the interface. This is the kind of semi-coherency associated with glissile martensite interfaces, whose motion is conservative (i.e., it does not lead to the creation or destruction of lattice sites). Two further conditions must be satisfied before even this interface can be said to be glissile [CHRISTIAN and CROCKER, 1980]:

(i) A glissile interface requires that the glide planes of the intrinsic dislocations associated with the product lattice must meet the corresponding glide planes of the parent lattice edge to edge in the interface, along the dislocation lines.

(ii) If more than one set of intrinsic dislocation exist, then these should either have the same line vector in the interface, or their respective Burgers vectors must be parallel.

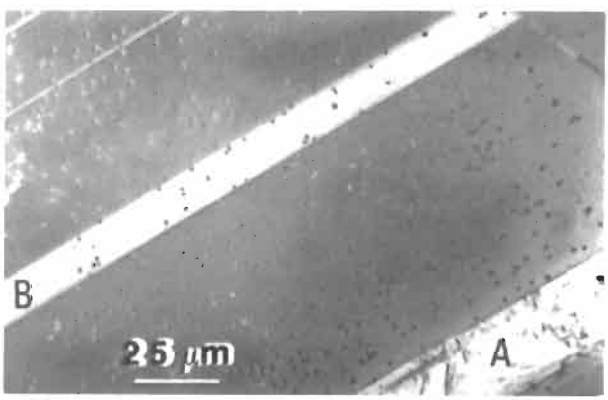


Fig. 18. Optical micrograph showing bands of hcp martensite (A and B) formed in the fcc austenite matrix in an Fe–12Mn–0.48C wt.% alloy.

This ensures that the interface can move as an integral unit, and places severe restrictions on any theory involving more than one lattice-invariant deformation.

The intrinsic dislocations accomplish the lattice-invariant deformation as the interface glides. The interface also contains small, atomic height steps, which are coherent and whose motion leads to the transformation of the parent into the product phase. These have the character of transformation dislocations, and their Burgers vector can be defined in terms of the lattice deformation **S** (CHRISTIAN [1982]). The density of such steps depends on the local curvature of the interface, but it must be emphasized that their motion is glissile and that it is their motion which leads to transformation, whereas the intrinsic dislocations cause the lattice-invariant deformation.

When the transformation dislocations (the coherent steps) generate the final lattice in a deformation which is an invariant-plain strain, as in the fcc → hcp cobalt transformation, no other defects are required in the interface.

2.5. Energetics of martensitic transformations

2.5.1. Transformation hysteresis and the reverse transformation

Two representative but widely different cases are considered here: the non-thermoelastic martensitic transformation in Fe–Ni (≈ 30 wt.%) alloys, and the thermoelastic martensite in typical β-phase alloys such as β-brass, often known as the $\frac{3}{2}$ electrons per atom Hume–Rothery phases. In both cases the shear component of the macroscopic shape strain is about 0.2 (shear angle about 12°) but the way in which this strain is accommodated by the parent phase is quite different for the two cases, as is the transformation hysteresis.

In the Fe–Ni alloys the martensite-parent interface appears to become immobilized after a plate of martensite has thickened to a certain extent, and when the martensite is heated, the interface will not move backwards, apparently being pinned by its damaged environment. Instead, the reverse martensite → parent (austenite) transformation takes place by the nucleation of small platelets of the parent phase within the martensite plates (KESSLER and PITSCH [1967]). And in ferrous alloys containing carbon, i.e., steels, the usual stages of martensite tempering occur (KRAUSS [1980]) within which the martensite decomposes through diffusional processes.

On the other hand, in alloys of the β-phase type, the shape deformation of the martensite is elastically accommodated by the matrix with no dislocation or "debris" generation; the interface therefore remains glissile, capable of "backwards" motion leading to the shrinkage of martensite plates during heating. In addition, in such *thermoelastic transformations* (KURDJUMOV and KHANDROS [1949]), stored elastic energy apparently predisposes the reverse martensite-parent transformation, allowing it to occur "prematurely" (KAUFMAN and COHEN [1958]). In such cases, the stored elastic energy contributes to the driving force for the reverse transformation.

Whether martensitic transformation in any particular alloy will be thermoelastic or not can be predicted by an analysis of the stress field around a plate of martensite. If the stress induced by the shape deformation causes the matrix to yield plastically, then the transformation is not thermoelastic [OLSON and OWEN, 1976; LING and OWEN, 1981]. Austenite in iron–platinum alloys undergoes ordering which greatly affects the elastic

properties. The shear modulus decreases and the yield strength increases, so that it becomes possible to elastically accommodate the martensite. The ordered alloy therefore has thermoelastic martensite whereas the disordered alloy has non-thermoelastic martensite. The plastic deformation for the two cases is illustrated in fig. 19.

2.6. Thermoelastic and non-thermoelastic martensitic transformations

The differences between thermoelastic and non-thermoelastic martensitic transformations may be highlighted by comparing Au–Cd and Fe–Ni alloys, respectively (OTSUKA and WAYMAN [1977]). Figure 20 shows that there is a substantial difference in the transformation hysteresis for the two alloys, which is the difference between the A_F (completion temperature for the reverse transformation) and M_S temperatures (KAUFMAN and COHEN [1958]). The thermoelastic transformation (Au–Cd) is characterized by a small hysteresis. Another difference is in the manner of the martensite-parent reverse transformation. In the thermoelastic case, the martensitic transformation on cooling proceeds by the continuous growth of martensite plates and the nucleation of new plates. When cooling is stopped, growth ceases, but resumes during further cooling, until the martensite plates impinge with each other or with grain boundaries. The reverse transformation upon heating occurs by the backwards movement of the martensite-parent interface, and the plates of martensite shrink and revert completely to the initial parent phase orientation.

On the other hand, a plate of non-thermoelastic martensite generated at a given temperature does not grow further upon subsequent cooling because the interface apparently becomes immobilized. Nor does the immobilized interface reverse its direction of motion during heating. Instead, the parent phase has to nucleate within the immobilized

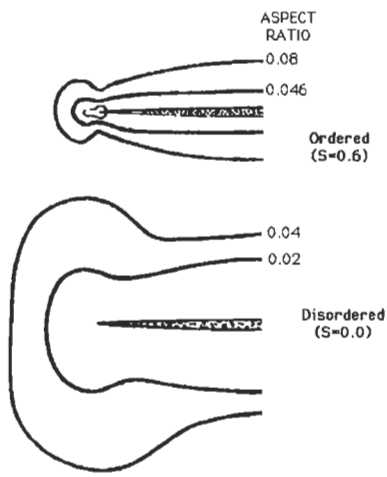


Fig. 19. Ling and Owens' calculations showing the contours where the matrix satisfies a yield criterion, as a function of the aspect ratio of the martensite plate and the degree of ordering (S). The region between each contour and the martensite plate represents plastically deformed austenite. Only half the martensite plate is illustrated in each case.

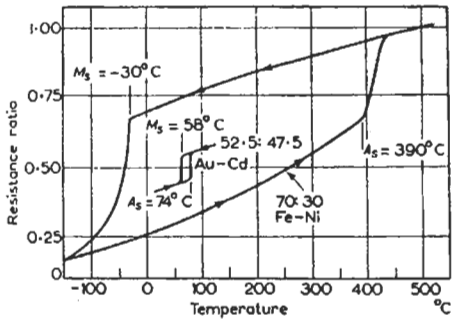


Fig. 20. Graph showing electrical resistance change during heating and cooling for Fe–Ni and Au–Cd alloys, indicating the hysteresis for the martensitic transformations on cooling and the reverse transformation on heating for non-thermoelastic and thermoelastic transformation, respectively (after KAUFMAN and COHEN [1957]).

martensite plates; since the parent phase can nucleate in many orientations, the original parent phase orientation is not recovered on heating (KESSLER and PITSCH [1967]).

2.6.1. Free energy change due to martensitic transformation

Figure 21 shows schematically the change in chemical free energies of martensite and austenite (parent phase) with temperature. T_0 is the temperature at which the austenite and martensite are in thermodynamic equilibrium and M_S is the temperature at which the transformation starts upon cooling. The difference in free energies between austenite (γ) and martensite (α'), $\Delta G_{M_S}^{\gamma\to\alpha'}$ at the M_S temperature is the critical chemical driving force for the onset of the martensitic transformation (other features of fig. 21 will be described later).

In general, the free energy change associated with a martensitic transformation is given by:

$$\Delta G^{P\to M} = \Delta G_C^{P\to M} + \Delta G_{NC}^{P\to M} \tag{10}$$

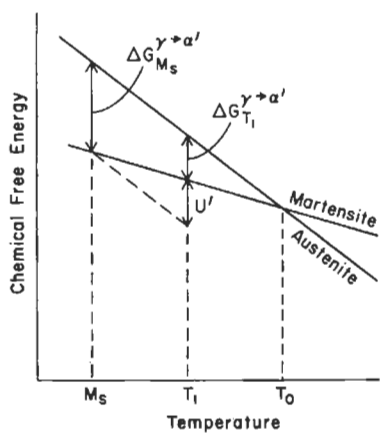


Fig. 21. Schematic diagram showing the free energy change for a martensitic transformation. See text for details.

References: p. 1552.

where $\Delta G_C^{P\to M}$ is the chemical free energy change (per unit volume of transformation) associated with the transformation from the parent to martensite, and $\Delta G_{NC}^{P\to M}$ is the non-chemical energy opposing the transformation (consisting of elastic strain and surface energy). Since the martensite-parent interface is semi-coherent the surface energy term should be small, and the elastic energy term will dominate. As will be seen later, the elastic strain energy per unit volume of martensite scales with the aspect ratio c/r, where $2c$ and $2r$ represent the plate thickness and length respectively.

During thermoelastic equilibrium, the plate of martensite continues to increase its aspect ratio unit the chemical and non-chemical terms balance. Consequently, a plate of constant length can be thickened or thinned when the transformation temperature is lowered or raised, respectively. This is the origin of the terminology *thermoelastic transformation.*

2.6.2. Nucleation of martensite

If a martensite transformation were able to nucleate (by a random fluctuation of structure) in an ideally perfect crystal of the parent phase, the attendant free energy change could be written (COHEN and WAYMAN [1981]) as:

$$\Delta G = (4\pi/3)r^2 c\Delta G_C^{P\to M} + (4\pi/3)r^2 c(Ac/r) + 2\pi r^2\sigma \tag{11}$$

assuming that a martensite plate can be represented as an oblate spheroid of radius r and semi-thickness c. The plate volume is then $(4\pi/3)r^2c$ and its surface area $\simeq 2\pi r^2$. The elastic strain energy per unit volume, due to the shape deformation, is given by Ac/r, where

$$A = \frac{\pi(2-\nu)}{8(1-\nu)}\mu s^2 + \frac{\pi}{4}\mu\delta^2, \tag{12}$$

where s and δ are the shear and dilatational components of the invariant-plane strain shape deformation, the dilatational strain being directed normal to the habit plane of the martensite. The Poisson ratio (ν) and the shear modulus (μ) are each taken to be the same for the parent and product phases. With typical values for the martensitic reaction in ferrous alloys, A is about 2.4×10^3 MJ m^{-3}.

Most martensitic transformations, and certainly the fcc $\rightarrow$ bcc (or bct) structural change in iron-base alloys, involve semi-coherent rather than fully coherent interfaces, and so the interfacial energy is estimated to be about 100–200 mJ m^{-2} (100–200 erg cm^{-2}) rather than to lie in the coherent range of 10–20 mJ m^{-2} (10–20 erg cm^{-2}). $\Delta G_C^{P\to M}$ is approximately –170 MJ m^{-3} for a typical ferrous alloy at its martensite-start temperature (M_s), and when this quantity is substituted into eq. 11 along with the above values of A and σ, the resulting function of ΔG versus r and c is a saddle-shaped surface with a barrier on the order of 8×10^{-16} J per nucleation event ($\simeq 5\times 10^3$ eV). This is about $10^5\ kT$ at temperatures where the transformation typically occurs. Clearly, the available thermal energy is much too small for homogeneous nucleation by heterophase fluctuations, and so special nucleation sites must be considered. It is frequently thought that dislocation arrays serve as martensite nuclei.

Martensitic reactions in ferrous alloys can occur either isothermally or athermally (i.e., on continuous cooling). Most transformations in steels have been observed athermally as a function of dropping temperature, but the best quantitative insights into the reaction kinetics are obtained from isothermal transformations in which case nucleation and transformation rates can be determined experimentally. In any event (i.e., athermal or isothermal reaction kinetics) the transformation kinetics are controlled by the nucleation rate and not the growth rate. It is the relatively rapid growth rate which leads to this situation.

The important problem of martensite nucleation was previously addressed by CHRISTIAN [1970] who concluded that the critical stage in martensite nucleation involves the attainment of a condition from which non-thermally activated growth can begin, which implies the formation of a semi-coherent interface. He further suggested a framework within which to consider the martensite nucleation process involving (a) classical nucleation utilizing thermal fluctuations (either homogeneously or at randomly distributed preferred sites), (b) the quenching-in of a pattern of embryos or compositional fluctuations which are subcritical at high temperatures but become supercritical at lower temperatures, and (c) the non-thermally activated rearrangement of the structure of a defect, or the interaction of different defects to produce a configuration resembling the martensitic structure. Recent thinking (CLAPP [1973], OLSON and COHEN [1976] would tend to favor the last of these categories.

Whatever the details of the mechanism of martensitic nucleation, one very important result is that martensite is triggered when the temperature T is such that free energy change for the diffusionless transformation of the parent phase γ to the product phase α', i.e., $\Delta G_T^{\gamma\to\alpha'}$ becomes less than a critical value $\Delta G_{M_s}^{\gamma\to\alpha'}$ (KAUFMAN and COHEN [1958], MAGEE [1970]). This critical value is about -1100 J mole^{-1} for low-alloy steels. However, it depends on the strength of the parent phase; dissolved elements which strengthen the austenite make martensitic transformation more difficult and this effect alone can depress the M_S temperature. Thus, the value of $\Delta G_{M_s}^{\gamma\to\alpha'}$ may be much smaller (say -3000 J mole^{-1}) for steels which have large concentrations of strengthening elements (GHOSH and OLSON [1992]).

2.7. Mechanical effects in martensitic transformations

2.7.1. Introductory comments

The most familiar mechanisms of plastic deformation are slip, mechanical twinning and diffusion-induced creep. For very small plastic strains, the first two of these deformation modes are conservative — i.e., they preserve an atomic correspondence between the deformed and undeformed parts of the crystal so that the crystal contains a memory of its original shape. All of these deformation modes are *lattice-invariant* because although they cause a change in the shape, the crystal structure remains as it was prior to deformation.

Martensitic transformation can also be regarded as a mode of deformation, but a mode which at the same time causes a change in the crystal structure. The shape change is of course, an invariant-plane strain with a large shear component and a relatively small

dilatational strain. The distinguishing features of a variety of deformation modes are listed in table 2.

Given that the growth of martensite causes a change in shape, it is natural to expect the transformation to be influenced by an externally applied stress. The permanent strain caused by martensitic transformation is called *transformation plasticity*. A phase change in a stress-free material is usually triggered by heat treatment, when the parent phase passes through an equilibrium transformation temperature. Alternatively, the application of a stress in isothermal conditions can trigger transformation in circumstances where it would not otherwise occur. Unusual effects can occur when stress and temperature work together. The transformation may occur at remarkably low stresses or at very small deviations from the equilibrium temperature. This is why even minute stresses can have a large influence on the development of microstructure. It is not surprising that transformation plasticity can be obtained at stresses which are much smaller than the conventional yield stress of the parent phase.

2.7.2. Chemical and mechanical driving forces

The interaction of an externally applied stress (which is below the yield strength of the parent phase) can manifest in two ways:

(i) The stress can alter the driving force for the transformation.
(ii) It can change the appearance of the microstructure by favouring the formation of those variants which best comply with the applied stress.

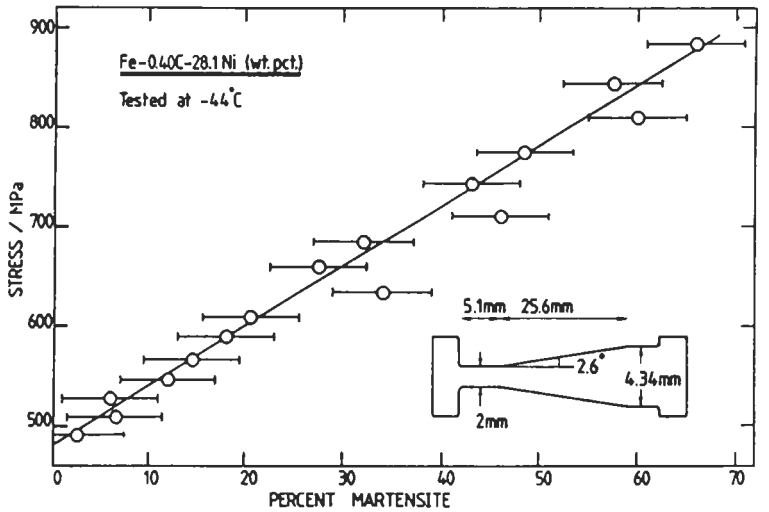
Both of these factors are illustrated in fig. 22, where a fine grained polycrystalline sample of austenite was stressed at a temperature *above* its normal martensite-start temperature. The amount of martensite obtained is seen to vary directly with the magnitude of the applied stress. The stress manifests as a *mechanical driving force* whose contribution assists the chemical driving force, which on its own is inadequate to trigger transformation. Not only does the stress induce the formation of martensite, but only those variants which comply with the tensile stress grow in profusion (fig. 22b). Thus, most of the plates are tempted to grow on those planes which are close to the plane of maximum shear stress (45° to the tensile axis). The microstructure would have been much more chaotic in the absence of the stress. The stress caused the alignment of

Table 2
Characteristics of different modes of deformation/transformation

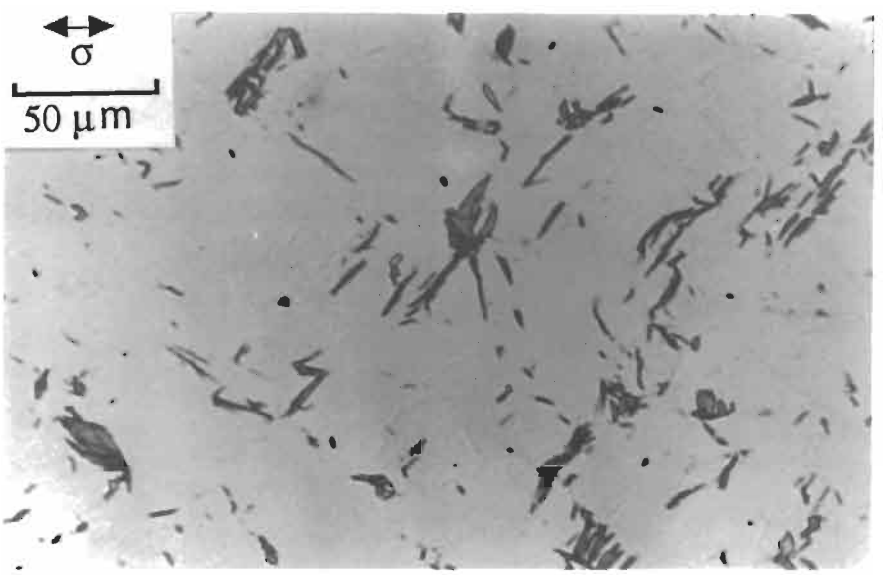
	Slip deformation	Mechanical twinning	Displacive transformation	Reconstructive transformation
Causes permanent change in shape	Yes	Yes	Yes	Yes
Invariant-plane strain shape change with a large shear component	Yes	Yes	Yes	No
Changes crystallographic orientation	No	Yes	Yes	Yes
Changes lattice type	No	No	Yes	Yes
Can lead to a density change	No	No	Yes	Yes

plates, and a more ordered, organised microstructure.

The chemical free energy change associated with a martensitic transformation was depicted in fig. 20. Referring to fig. 20 again, when a stress is applied to the austenite at T_1 (between M_S and T_0), the mechanical driving force, U, due to the stress is added to the chemical driving force, $\Delta G_{T_1}^{\gamma\to\alpha'}$, and the martensitic transformation starts at the critical stress where the total driving force is equal to $\Delta G_{M_s}^{\gamma\to\alpha'}$. U' ($= \Delta G_{M_s}^{\gamma\to\alpha'} - \Delta G_{T_1}^{\gamma\to\alpha'}$) in fig. 20 is the critical mechanical driving force necessary for the stress-induced martensitic



(a)



References: p. 1552.

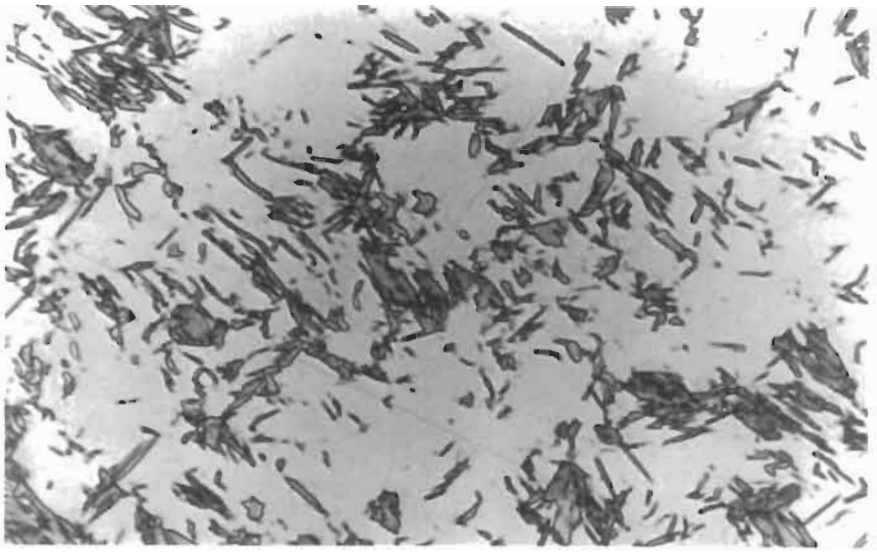


Fig. 22. Stress-affected martensitic transformation in a Fe–28Ni–0.4C wt.% alloy tested at a temperature above the martensite-start temperature. (a) Volume fraction of martensite as a function of stress. (b) Optical micrograph from a low stress region. (c) Optical micrograph from a high-stress region. The arrow indicates the direction of the tensile stress.

transformation at T_1. The mechanical driving force is a function of stress and the orientation of a transforming martensite plate, and can be expressed (PATEL and COHEN [1953]) as:

$$U = \tau s + \sigma\delta, \tag{13}$$

where τ is the resolved shear stress along the transformation shear direction in the martensite habit plane, σ is the normal stress resolved perpendicular to the habit plane, and s and δ are the shear and dilatational components of the shape deformation due to martensitic transformation. For a small applied stress σ_1, τ and σ (fig. 23) may be expressed as follows for any given orientation of a martensite plate (ONODERA *et al.* [1976]);

$$\tau = \tfrac{1}{2}\sigma_1 \sin 2\theta \cos\alpha, \tag{14}$$

$$\sigma = \pm\tfrac{1}{2}\sigma_1[1 + \cos 2\theta], \tag{15}$$

where σ_1 is the magnitude of the applied stress, θ is the angle between the applied stress and the normal to the habit plane, and α is the angle between the transformation shear direction and the maximum shear stress in the habit plane. The plus and minus signs in eq. 15 correspond to uniaxial tension and compression, respectively. From eqs. 13–15 the mechanical driving force is

$$U = \tfrac{1}{2}\sigma_1[s \sin 2\theta \cos\alpha \pm \delta(1 + \cos 2\theta)]. \tag{16}$$

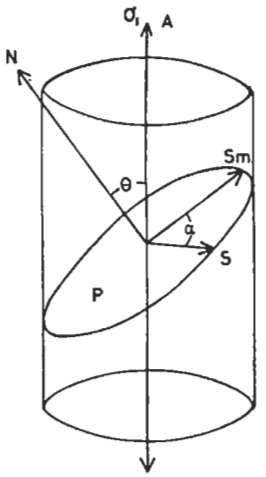


Fig. 23. Schmid-factor diagram for the influence of an applied stress σ_1 along stress-axis A in inducing a martensitic transformation on habit plane P with normal N. S is the direction of the shape strain for the martensite and S_m is the maximum shape strain elongation parallel to the habit plane.

2.7.3. Critical stress to induce martensitic transformation

When martensite is induced in polycrystalline austenite (containing randomly oriented grains) by an external stress, the martensite plate whose orientation yields a maximum value of U, eq. 16, will be formed first. If it is assumed that the angle α is zero, then the maximum value of U is obtained by setting $dU/d\theta = 0$. For Fe–Ni alloys ($s = 0.20$, $\delta = 0.04$), the value of θ which would maximise U is therefore found to be as 39.5° for tension and 50.5° for compression. These angles would become exactly 45° if $\delta = 0$.

If the chemical driving force ($\Delta G^{\gamma \to \alpha'}$) decreases linearly with an increase in temperature above M_S, then a proportionally greater mechanical driving force is required before martensite formation is triggered. The necessary stress might then be expected to increase linearly with temperature, and this is in fact observed experimentally (OLSON and COHEN [1972], ONODERA *et al.* [1976]) over the range $M_S \to M_S^\sigma$ (fig. 24). However, above M_S^σ (e.g., at T_2), martensite is induced at a stress σ_b after plastic deformation of the austenite occurs.

The microstructure also changes when martensite is generated under the influence of an applied stress. It has been observed (OTSUKA and WAYMAN [1977]) that only a single variant, out of (up to) twenty four that might be possible, of martensite is induced when an elastic stress is applied. This is the variant which gives rise to maximum elongation in the direction of the tensile axis.

When the austenite is deformed at temperatures above M_S^σ (e.g., at T_2 in fig. 24), it begins to deform plastically at a stress σ_a, and is strain-hardened up to σ_b. Then the martensitic transformation starts to take place. σ_b is considerably lower than σ_c, which is obtained by extrapolating the critical stress–temperature line between M_S and M_S^σ. This decrease (i.e., $\sigma_c - \sigma_b$) in the critical applied stress for martensite formation is due to plastic deformation of austenite. As to the role of plastic deformation of austenite on the deformation-induced martensitic transformation, two different views have been expressed.

References: p. 1552.

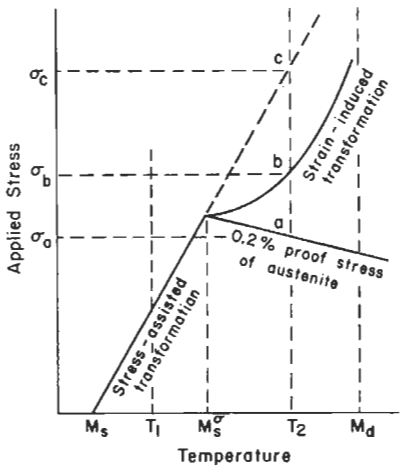


Fig. 24. Schematic stress–temperature diagram showing the critical stress to initiate the formation of martensite as a function of temperature. The various regimes of behavior are described in the text.

One is the strain-induced martensite nucleation hypothesis proposed by OLSON and COHEN [1972, 1976, 1979], and the other holds that the stress is locally concentrated at obstacles (e.g., grain boundaries, twin boundaries, etc.) by plastic deformation of austenite, and thus the concentrated stress becomes equivalent to σ_c in fig. 24, the latter proposed by ONODERA and TAMURA [1979] and others.

In contrast to the M_S temperature where a martensitic transformation will begin as a consequence of the chemical driving force, reference is frequently made to the M_d temperature. M_d is the temperature above which the chemical driving force becomes so small that nucleation of martensite cannot be mechanically induced, even in the plastic strain regime.

2.7.4. Transformation-induced plasticity (TRIP)

As an example of TRIP behavior (ZACKAY *et al.* [1967]), we consider an austenitic Fe–29Ni–0.26C wt.% alloy (M_S = –60°C, M_d = 25°C) deformed in tension at various temperatures. The relationship between tensile properties and test temperature for this alloy is shown in the upper part of fig. 25 (TAMURA *et al.* [1970]). Note that the elongation is nearly 100%. An inverse temperature dependence of the yield stress is observed between M_S^σ and M_S. This results from the shape strain of martensite formed before yielding of the austenite occurs. The tensile strength is increased with decreasing test temperature just above M_S^σ. Such an elongation enhancement is attributed mainly to the suppression of necking, due to the increase in work-hardening rate by the formation of martensite. Furthermore, it is thought that the initiation and propagation of cracks is suppressed by the formation of martensite during deformation, because the stress concentrations may be released by the formation of preferential variants of martensite at regions of stress concentration. Similar results are seen in fig. 25 for two other alloys.

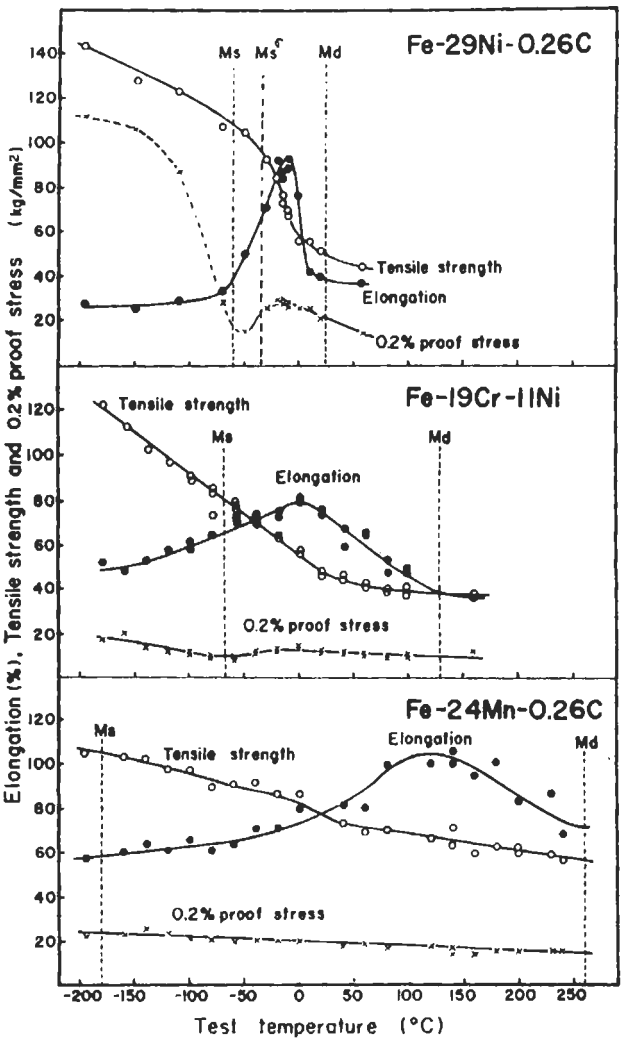


Fig. 25. Experimental data showing the effect of test temperature on the tensile properties of three metastable austenitic alloys strained at the rate $\varepsilon = 5.5 \times 10^{-4}/s$ (after TAMURA *et al.* [1970]).

With an increase in carbon content, the martensite hardness is increased substantially, but the austenite hardness only slightly increases. Accordingly, in the case of high-carbon steels, since a larger work-hardening rate can be obtained even with a small amount of martensite formation, the uniform elongation by TRIP becomes much larger (TAMURA *et al.* [1972, 1973]).

References: p. 1552.

2.8. Mechanical effects specific to thermoelastic martensitic transformations

2.8.1. General

Thermoelastic martensitic transformations are accompanied by very unusual mechanical effects in both the parent and martensite phases. In addition to the now popular shape memory effect (SME) and the mechanical behavior of alloys exhibiting this type of behavior, closer inspection shows that SME alloys also show a variety of other kinds of interesting mechanical behavior. Examples are the superelasticity associated with the formation of a reversible stress-induced martensite, the rubberlike effect, "training" and the two-way shape memory effect, extensive deformation resulting from stress-induced martensite-to-martensite transformations, unusual damping behavior, and finally, high stresses generated during the reverse martensite-to-parent transformation.

The number of materials exhibiting the shape memory effect is now extensive, including many Cu-based alloys, those of the noble metals based on Ag and Au, the classic Ni–Ti alloys (NITINOL), ternary variations of the same such as Ni–Ti–Cu and Ni–Ti–Fe, Ni–Al alloys, and Fe–Pt alloys, to mention a few. The martensite that forms in these alloys is thermoelastic and crystallographically reversible, and both the parent and product phases are in general atomically ordered [SHIMIZU and WAYMAN, 1972]. The SME martensites are either internally twinned or internally faulted due to the lattice-invariant deformation. In addition, the parent phase is usually of the B2 or DO3 type and the martensite crystal structure is given by the stacking sequence of planes originating from the parent phase {1 1 0} planes, e.g., 2H, 3R, 9R and 18R.

2.8.2. The shape memory effect (SME)

A brief description of the SME is as follows. An object in the low temperature martensitic condition* when deformed and then unstressed regains its original shape on heating through the $A_S \rightarrow A_F$ temperature interval, as the parent phase grows at the expense of the martensite. The strains (6–8%) describing the original deformation of the martensite are completely recovered. The recovery of the original shape is entirely due to the reverse transformation of the deformed martensite.

The SME effect can be explained as follows (SCHROEDER and WAYMAN [1977], SABURI *et al.* [1979]). A single crystal of the parent phase will usually transform into 24 orientations of martensite (variants of {*hkl*}). But when this mixture of many variants of martensite is deformed, the microstructure changes into a single orientation of martensite. This change occurs by the motion of martensite–martensite boundaries, during which the variant most compatible with the applied deformation grows at the expense of the others. The motion of the martensite–martensite boundaries involves some twinning/detwinning deformation, because the twins are in fact simply other orientations of martensite. It has been shown that twinning can convert one orientation (variant) of martensite to another (SCHROEDER and WAYMAN [1977a]). Consequently, when a fully martensitic mixture of many orientations of martensite is deformed, one plate orientation grows at the expense

* Usually below M_F, the temperature at which the martensitic transformation is completed.

of the others. The favoured martensite is that whose shear component of the shape deformation permits the maximum elongation of the specimen in the direction of the tensile axis (SABURI *et al.* [1979]). The entire deformation is therefore due solely to martensite variant reorientation. Any other form of deformation (e.g., ordinary slip) tends to destroy the shape memory effect.

Although a single crystal of the parent phase transforms into many orientations of martensite, the reverse situation does not occur. Rather, the single crystal of martensite resulting from deformation below the M_F temperature, transforms during heating into a single orientation of parent phase. This is a consequence of the martensite and parent lattice symmetries involved and the necessity to maintain atomic ordering during the reverse transformation. The highly symmetric parent phase (usually cubic) has many crystallographically equivalent ways in which the Bain distortion can occur, giving numerous variants of martensite. On the other hand, the relatively unsymmetric martensite (e.g., monoclinic in Cu–Zn–Al) does not have such a multiplicity of choices, so that only the original orientation of parent phase can form. The single crystal of martensite "unshears" to form a single crystal of the parent phase, thereby restoring the specimen to its original shape. A schematic representation of the shape memory process is shown in fig. 26 (SABURI *et al.* [1979]).

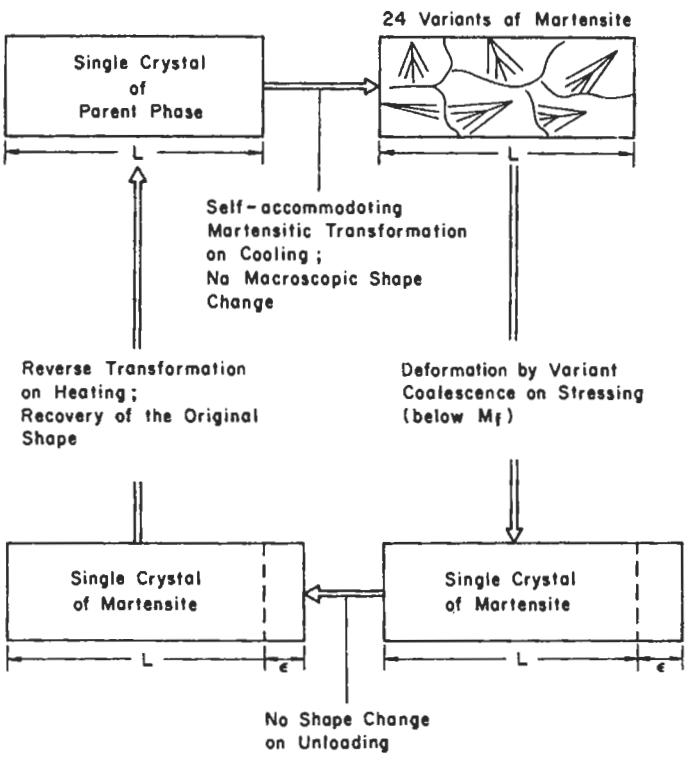


Fig. 26. Schematic representation of deformation processes associated with the shape memory effect.

References: p. 1552.

This description of the shape memory effect is quite general, irrespective of the particular alloy system or crystal structure of the martensite.

Figure 27 (SCHROEDER and WAYMAN [1977a] shows a stress–strain curve for a Cu–Zn SME alloy single-crystal specimen deformed below the M_F temperature. It is seen that the martensite begins to deform (a') at a relatively low stress, 35 MN m^{-2}. The residual strain at point c was completely recovered during heating.

2.8.3. The two-way shape memory effect

With the shape memory effect, as previously described, a specimen deformed by martensite variant conversion will undeform and regain its original shape during heating from A_S to A_F. In contrast the two-way shape memory (TWSM) involves a reversible deformation. A specimen will spontaneously deform during cooling from M_S to M_F and then undeform during heating from A_S to A_F. Such behavior occurs as a result of deformation in either the parent phase or in the martensite (WASILEWSKI [1975]). There are in fact, two types of two-way shape memory depending on the way in which the alloy is *trained* (SCHROEDER and WAYMAN [1977b]):

(i) *SME cycling* involves cooling a specimen below the M_F temperature, deforming it to produce a preferential martensite variant, as described earlier, and heating it to above the A_F temperature. This procedure is repeated several times, the manner of martensite deformation (e.g., tension, compression, bending) remaining unchanged.

(ii) *SIM cycling* involves the deformation of a specimen above the M_S temperature in order to produce stress-induced martensite (SIM), followed by reversal of the SIM when the applied load is released. This process is also repeated several times, using the same means of stressing each time.

The two-way shape memory is observed after both SME and SIM cycling, and the terms *SME training* and *SIM training* have been suggested to describe the process of cycling in the desired manner (SCHROEDER and WAYMAN [1977b]). In either case, the two-way behavior comes from the preferential formation and reversal of a trained variant of martensite formed after either SME or SIM cycling and cooling the specimen from M_S

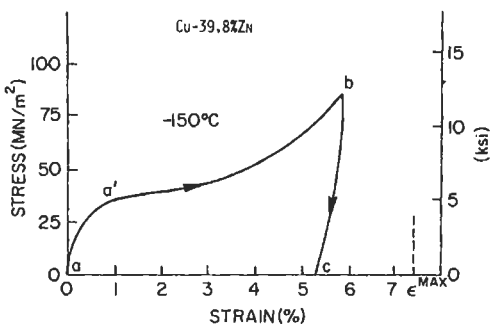


Fig. 27. Stress–strain curve for a Cu–Zn shape-memory alloy deformed below its M_f temperature. The low flow stress a' associated with the martensite deformation is to be noted.

to M_F. That is, the training of a specimen to form a preferential variant of martensite upon cooling to M_F can result from prior deformation of the martensite formed by cooling followed by heating, or SIM cycling above the M_S temperature. In both cases the first part of the TWSM occurs upon cooling. In either case, the result is that a major portion of the martensite in a given specimen corresponds to a preferred variant, which produces a spontaneous strain during cooling, the amount of the strain corresponding to the shape strain of the preferred variant.

The effectiveness of SME training is found to be inferior to that of SIM training. Nevertheless, in both cases, the preferential formation of a selected variant of martensite is explained by the progressive introduction of a pattern of stresses or defects into the material during the training processes. These stresses (or defects) apparently become better defined and established with increased cycling (SCHROEDER and WAYMAN [1976]).

2.8.4. The engine effect in shape memory alloys

Although SME materials generally deform in the martensitic condition at comparatively low stresses, as seen in fig. 27, surprisingly large stresses are generated when the deformed martensite is heated from A_S to A_F. As a case in point (JACKSON *et al.* [1972]), a nearly equiatomic Ni–Ti alloy when martensitic will undergo deformation (as at point a' of fig. 27) at a stress of about 70 MN m^{-2}. Yet when it is constrained and heated to A_F it is found that a *thermomechanical recovery stress* of about 700 MN m^{-2} is generated. In other words, heat can be used to create a mechanical force which can do work. This is the basis for the design of engines based on shape memory alloys (GOLDSTEIN and MCNAMARA [1979]).

2.8.5. Pseudoelastic effects

Pseudoelasticity refers to a situation where large strains, in excess of the elastic limit, are completely recovered upon unloading at a constant temperature. There are two categories, *superelasticity* and the *rubber-like* effect (OTSUKA and WAYMAN [1977]). In superelasticity, the martensite is stress-induced and reverts to the parent phase on the removal of stress; the stress therefore contributes a mechanical driving force for transformation. By contrast, rubber-like behavior involves the deformation of existing martensite with no phase transformation.

2.8.6. Pseudoelastic effects: superelasticity

The application of a suitable stress above the M_S temperature (but below M_d), induces the formation of martensite. In thermoelastic alloys, this stress-induced martensite disappears when the stress is released, giving rise to superelasticity. A stress–strain curve showing typical superelastic behavior for a single-crystal sample is shown in fig. 28. Two plateau regions are seen. The upper plateau corresponds to the formation of a preferred variant of stress-induced martensite plates. The preferred variant is that whose shape strain best complies with the applied stress (SCHROEDER and WAYMAN [1979]). The parallel plates of martensite nucleate, lengthen and coalesce, leading eventually to a sample which is a single crystal of martensite, at the end of the upper plateau region in fig. 28. Upon releasing the load, the stress–strain curve follows the lower plateau

References: p. 1552.

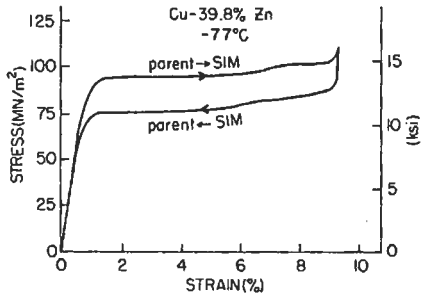


Fig. 28. Stress–strain curve for a thermoelastic Cu–Zn shape-memory alloy deformed above the M_s temperature, showing superelastic behavior as a consequence of the formation and reversion of a reversible, stress-induced martensite.

region, which involves the formation of parallel plates of only one variant of the parent phase. The levels of the plateau stresses shown in fig. 28 depend on the test temperature. The upper plateau stress is naturally zero at the M_S temperature. The temperature dependence of the stress to produce SIM is shown in fig. 29. The slope of the M_S variation with temperature may be used to estimate the latent heat of transformation (OTSUKA and WAYMAN [1977]). Stress–strain curves of the type shown in fig. 29 are frequently referred to as *superelastic loops.*

2.8.7. Pseudoelastic effects: rubber-like behavior

Rubber-like behavior occurs in the martensite phase of some alloys which undergo thermoelastic transformation. It does not involve a phase transformation, but is related to the reversible movement of transformation twin boundaries or martensite boundaries.

It is useful to compare the usual shape-memory effect (fig. 26) with rubber-like behaviour, since the latter often occurs in SME alloys which are aged in their martensitic

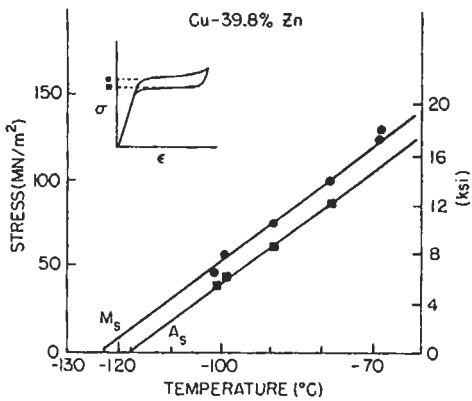


Fig. 29. Stress–temperature plot of experimental data showing the temperature dependence of the applied stress required to produce stress-induced martensite in a Cu–39.8% Zn alloy.

state. In the single-crystal shape-memory effect, applied deformation can convert a microstructure which consists of many variants of martensite into a single crystal of martensite. Heating then displacively transforms the martensite into the parent phase single-crystal, thereby removing the applied deformation.

However, if the microstructure consisting of many variants of martensite is first allowed to age, and then deformed to a single crystal of martensite, the removal of the applied stress causes the single crystal of martensite to revert to the many variants, thereby reversing the applied deformation. This restoring force only arises after ageing, and its origins lie in the changes that occur in the state of order on ageing (AHLERS *et al.* [1978]; MARUKAWA and TSUCHIYA [1995]; TSUCHIYA *et al.* [1995]).

The long-range order in the parent structure is usually imperfect because the parent phase occurs at relatively high temperatures. On quenching, the martensite inherits this imperfect order. Ageing at ambient temperature can, with the help of quenched-in excess vacancies, allow the state of order in the martensite to become more perfect. The martensite also has a lower crystallographic symmetry than the parent phase, so the *orientation* of the more perfect state of order is different for each martensite variant. This makes the conversion of one martensite variant into another more difficult, since the conversion would lead to wrongly oriented order in the growing variant. The resulting increase in energy provides the restoring force which tends to make the single-crystal of martensite revert to the original many variants when the applied stress is removed.

2.8.8. Martensite-to-martensite transformations

Figure 30 shows the deformation behavior of a Cu–39.8Zn wt.% single crystal strained at –88°C, some 35°C above the M_S temperature (SCHROEDER and WAYMAN [1978]) The first upper plateau corresponds to the formation of SIM as described previously. However, the second plateau which starts at about 9% strain, corresponds to a second martensitic *transformation* which is stress-induced from the first martensite "mother". The two lower plateaux are a result of the reverse transformations, occurring in an inverse sequence. By means of these successive martensite-to-martensite transformations, completely recoverable strains as high at 17% can be realized. Figure 30 shows a double superelastic loop. It should be noted that the second stress-induced martensite

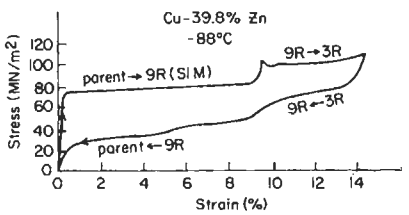


Fig. 30. Stress–strain diagram showing the deformation behavior of a Cu–39.8% Zn alloy single crystal specimen deformed above the M_s temperature. A double superelastic loop is observed which is the result of two successive stress-induced martensitic transformations. The first stress-induced transformation is that depicted by fig. 25, while the second stress-induced transformation originates from the first-formed (under stress) martensite.

References: p. 1552.

in the Cu–Zn alloy can be formed from the first SIM, or from the normal thermally formed (upon cooling) martensite.

In the case of the Cu–Zn alloy discussed above, the initial martensite has a 9R structure, and additional deformation changes this into a 3R structure. This occurs by shearing on the basal (close-packed) plane of the 9R such that the structural stacking sequence is changed from ...*ABCBCACAB*... to ...*ABCABCABC*...

3. *Crystallographically similar transformations*

3.1. The bainite transformation in steels

It is widely agreed that the bainite reaction in steels has certain martensitic characteristics but there may be some diffusion of carbon at some stage of the reaction. The surface relief accompanying the growth of bainite is an invariant-plane strain with a large shear component. There are also other martensitic-like features such as an irrational habit plane and transformation substructure, and various studies of the austenite-bainite orientation relationship imply that the Bain correspondence between the austenite and bainitic ferrite exists. (See also ch. 15, § 2.6).

Whereas the detailed role of carbon in the mechanism of the bainite transformation requires further work, it is clear that the iron and substitutional atoms are transferred in a coordinated manner across the bainite–austenite interface. Consequently, it is reasonable to expect the phenomenological crystallographic theory to apply to the bainite transformation, and this is indeed found to be the case in practice.

3.2. Oxides and hydrides

Various metallic oxides and hydrides form with a plate morphology and exhibit an invariant-plane strain shape deformation with a large shear component. Such plates are also internally twinned or internally slipped. An example of an internally twinned plate in a tantalum oxide is shown in fig. 31. In this case the observed crystallographic features of the oxide plates are accurately predicted by the phenomenological crystallographic theory developed for martensitic transformations (Van Landuyt and Wayman [1968]). A lattice correspondence therefore exists between the metallic atoms in the two phases and remains intact despite the occurrence of interstitial diffusion (of oxygen) during the phase change.

3.3. The CuAu II ordering reaction

The formation of the ordered orthorhombic CuAu II phase from the disordered cubic phase parent at about 390°C generates an invariant-plane strain relief with a large shear component, and conforms well with the phenomenological crystallographic theory of martensitic transformations. This led Smith and Bowles [1960] to conclude that the mechanism by which the internally twinned plates of the CuAu II form, is similar to that of a martensite, even though place changes between neighboring atoms are necessarily involved during the ordering reaction. But Aaronson and Kinsman [1977] dispute this,

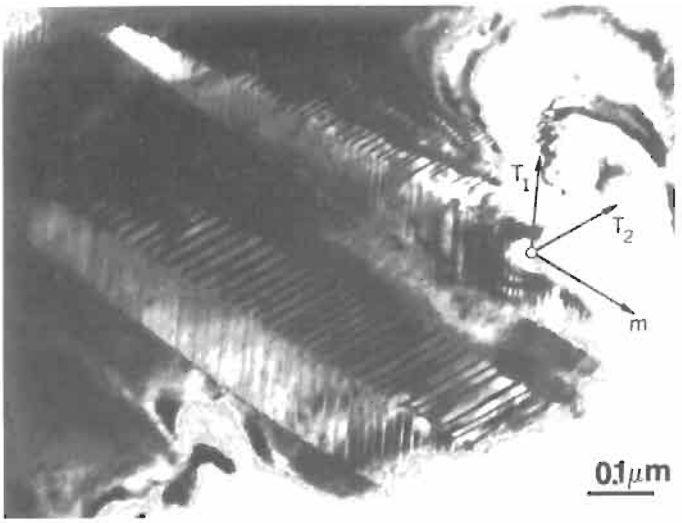


Fig. 31. Transmission electron micrograph showing plates of the sub-oxide TaO_y following the oxidation of tantalum. The striations within the plates are twins, and the general crystallygraphy of the process is well described by the phenomenological crystallograpic theory developed for martensitic transformations.

claiming that the atomic jumps required for ordering are inconsistent with a shear transformation and that conformity with the crystallographic theory is not a sufficient criterion to identify a shear transformation.

The CuAu II ordering reaction was later considered by BOWLES and WAYMAN [1979] who suggested that the place exchanges between neighboring atoms, required to achieve the ordering, are redundant in the sense that they do not contribute to the transformation shape change so that the theory still prescribes the displacements involved in producing the lattice change. They proposed that the growth of CuAu II plates occurs by the glide of transformation dislocations, i.e., the structure of the interface is glissile as required for displacive transformation.

CHRISTIAN [1994] has analysed these difficulties in detail. Solid-state phase transformations are often classified as either reconstructive or displacive depending on whether or not individual atoms are mobile. In an entirely displacive transformation, each atom moves, relative to its neighbours, by less than an interatomic distance. Thus, the number of atoms must be conserved in any incremental volume transformed from the parent to the product phase by an entirely displacive (i.e., martensitic) mechanism. This is not necessarily true for a reconstructive transformation where the atoms are highly mobile, and material transport can occur in a way which reduces the strain energy so that the shape change does not have a shear component. The ordered products or precipitates (such as CuAu II) whose growth involves very short-range diffusion, but nevertheless exhibit shape changes with appreciable shear components, are not simply reconstructive, but *diffusional-displacive*. The interfacial structure of this kind of a transformation product must be the same as that of martensite, and the transformation will exhibit all the

crystallographic features of martensite, with the plate-shape determined by a need to minimise the strain energy associated with the shape deformation.

4. *Omega phase formation*

The $\beta \rightarrow \omega$ phase change (cubic-hexagonal) involves correlated atomic displacements and occurs in Ti, Zr and Hf alloys and in β-phase alloys of the noble metals. It has been extensively investigated in recent years. The reader is referred to two reviews: SASS [1972] and WILLIAMS *et al.* [1973]. The $\beta \rightarrow \omega$ phase change occurs both athermally and isothermally, with the same mechanism for both paths (WILLIAMS *et al.* [1973]).

Omega phase formation can occur reversibly and without diffusion at quite low temperatures, and the ω-phase cannot be suppressed by rapid quenching. These characteristics parallel those found in martensitic transformations, but unlike martensite, the ω-phase "particles" exhibit a cuboidal or ellipsoidal morphology, depending upon the relative parent and product lattice parameters. At least during the early stages of transformation, the ω particles are quite small ($\simeq$ 1.5 nm), and they are aligned along <111> directions. Their observed number-density is very high (10^{18}–10^{19} cm^{-3}). The smallness of the particles and their high number-density suggest that their nucleation is not as serious a barrier as their growth, in contrast to martensitic transformations where the opposite is the case. Figure 32 shows an example of omega particles.

A number of alloys exhibiting the $\beta \rightarrow \omega$ transformation feature the following orientation relationship:

$$(1\,1\,1)_\beta \parallel (0\,0\,0\,1)_\omega$$

$$[1\,\bar{1}\,0]_\beta \parallel [2\,\bar{1}\,1\,0]_\omega$$

Interestingly, these same alloys (although at different compositions) also undergo a $\beta \rightarrow \alpha'$ martensitic transformation in which case the orientation relationship is of the Burger's type:

$$(1\,1\,0)_\beta \parallel (0\,0\,0\,1)_{\alpha'}$$

$$[1\,\bar{1}\,1]_\beta \parallel [1\,1\,\bar{2}\,0]_{\alpha'}$$

It should be noted that the $\beta \rightarrow \omega$ orientation relationship shows a multiplicity of only four, which, when compared to that of martensitic transformations (where the multiplicity is much higher) suggests that the respective lattice correspondences are different.

Another comparison can be made between the diffusionless $\beta \rightarrow \omega$ and $\beta \rightarrow \alpha'$ (martensitic) transformations. When martensite forms, the atomic displacements are mostly homogeneous, as given by a Bain-type distortion. Each atom (apart from additional shuffles in some cases) is homogeneously transported to its final position according to the Bain deformation and correspondence. But in the $\beta \rightarrow \omega$ transformation,

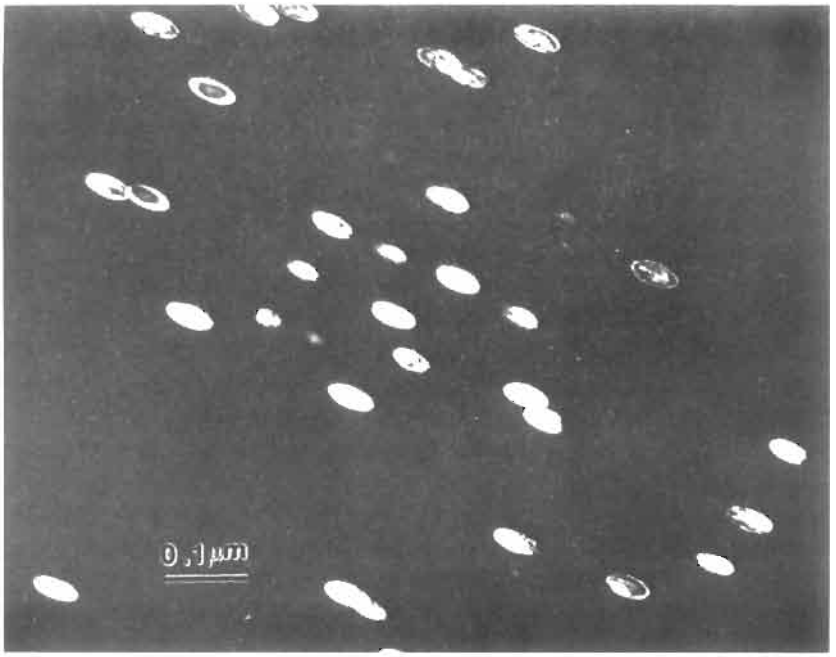


Fig. 32. Transmission electron micrograph (dark field) showing ω-particles formed at 480°C in a Ti–Mo–Sn–Zr alloy. Only one of four variants of the ω-phase is in contrast.

some of the corresponding atoms undergo movements whereas others do not. Figure 33 (SASS [1972]) shows a $(1\,0\,\bar{1})$ section through a β unit cell. Atoms A, B, E and F remain in place during the $\beta \rightarrow \omega$ transformation, whereas atoms C and D undergo shuffle

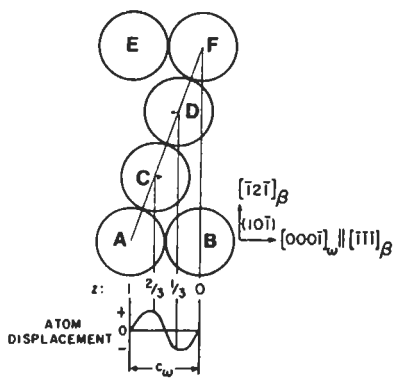


Fig. 33. Schematic diagram showing atomic movements involved during ω-phase formation. This is a $(10\bar{1})$ section through the parent bcc unit cell. The 2/3 <111> longitudinal displacement waves cause atomic movements (in directions indicated by the two arrows) which are those needed to form the ω-structure (after SASS [1972]).

References: p. 1552.

movements in opposite sense along $[\bar{1}\,\bar{1}\,\bar{1}]_\beta$, in the end reaching the position $\frac{1}{2}[\bar{1}\,\bar{1}\,\bar{1}]$, initially having been situated at the $\frac{2}{3}$ and $\frac{1}{3}$ positions. Accordingly, the 1st, 4th, 7th, 10th, etc. planes remain unchanged while 2 and 3, 5 and 6, 8 and 9, etc., shuffle past each other in opposite senses. The $\beta \rightarrow \omega$ atomic movements are formally consistent with atom movements given by a $\frac{2}{3}\langle 111\rangle$ displacement wave having nodes at $n\langle 111\rangle_\beta$ positions (n = any integer).

Advances in understanding the $\beta \rightarrow \omega$ transformation have come from the application of transmission electron microscopy and diffraction. Diffraction patterns frequently exhibit networks of diffuse intensity (sheets of intensity on $\langle 111\rangle_\beta$ planes) which obscure the identification of diffraction maxima from the ω-phase. These diffuse diffraction effects are observed at temperatures above which those from the ω-phase are clearly identified, and the former have been ascribed to pretransformation "linear defects".

Impurities have a marked influence on the $\beta \rightarrow \omega$ transformation. For example, the presence of 1200 ppm of oxygen can lower the ω-start temperature by about 600 K, presumably because oxygen stiffens the matrix and depresses the transformation start temperature. Apparently the oxygen atoms somehow interact with the pretransformation linear defects along $\langle 111\rangle_\beta$ and impede their ordering.

There are some outstanding problems with ω-phase formation, including the possible role of lattice vacancies in promoting linear defects and hence nucleation, the high nucleation frequency and low particle growth rate, and the extreme embrittlement of the matrix β-phase after the ω-phase forms.

5. *Phase changes and charge density waves*

A charge density wave (CDW) is a static modulation of conduction electrons and is a Fermi-surface driven phenomenon usually accompanied by a periodic distortion of the lattice. In essence, the electronic energy of the solid is lowered as a consequence of the lattice distortion, the attendant strain energy of which is more than compensated by the reduction in electronic energy. The present understanding of CDWs in conductors follows from the pioneering work of PEIERLS [1955], KOHN [1959] and OVERHAUSER [1968, 1971]. A detailed description of CDW phenomena has been published by WILSON *et al.* [1975].

PEIERLS [1955] suggested that a simple one-dimensional solid is susceptible to a periodic lattice distortion which can lower the total energy of the system by means of a distortion which changes the lattice periodicity. The basic concept is illustrated in fig. 34. fig. 34a shows the familiar $E-k$ relationship for a one-dimensional monovalent metal. The first Brillouin zone with boundary at $\pm\pi/a$ is half occupied. The upper part of the figure shows that the electronic charge density is of periodic form, being maximum in the vicinity of the ion cores. In fig. 34b the one-dimensional lattice has been periodically distorted as shown in the upper part of the figure and the resulting period $2a$ causes the formation of a charge density wave of period $2a$ associated with the new superlattice. Since the lattice has been doubled in real space, a new band gap will appear at $\pm\pi/2a$ in k-space which permits some electrons to "spill down", thus lowering the Fermi energy. Thus the (periodic) lattice distortion has a rather simple driving force which is purely

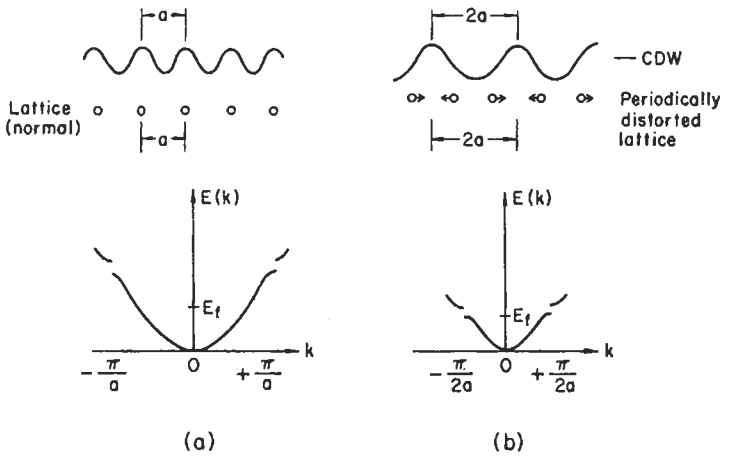


Fig. 34. Schematic drawing showing a periodic lattice distortion, charge density waves, and energy band gaps before and after a periodic lattice distortion in a hypothetical one-dimensional metal. Note the overall lowering of the Fermi level following the periodic lattice distortion. See text.

electronic in origin. Numerous examples of a CDW phase change have recently been found in quasi-one-dimensional organic conductors and quasi-two-dimensional layered compounds (e.g., transition-metal dichalcogenides).

A favorable Fermi surface geometry is necessary for the formation of a CDW, which will most likely occur when the shape of the Fermi surface permits a connection by the same wavevector **Q**, i.e., $\mathbf{Q}=2\mathbf{k}_f$. This modulation with wavevector **Q** will modify the Fermi surface by creating gaps at these nested positions. If the nested (i.e., dimpled) portion of the Fermi surface is significant, the energy "gain" by creating energy gaps may overcome the energy "cost" arising from the strain associated with the periodic lattice distortion, thus allowing the formation of a CDW. In other words, a structural change will occur when the CDW formation is accompanied by ion displacements which stabilize the charge perturbation. Another requirement for forming a CDW is a strong electron–phonon coupling, required to permit ionic displacements to reduce the otherwise prohibitive Coulomb energy, and precursor phenomena such as a soft phonon mode (CHAN and HEINE [1973]) might occur above the transition temperature to assist the CDW instability. The normal crystalline periodicity is altered in the presence of a CDW. However, it should be noted that the wavevector of a CDW is determined by the Fermi surface and is therefore not necessarily an integral fraction of a reciprocal lattice vector of the undistorted parent phase. Consequently an "incommensurate" phase may result, which is considered to have lost its translational symmetry (AXE [1977]) and is thus a quasi-crystalline phase.

The incommensurate state described above may not actually correspond to the lowest possible energy state and, accordingly, the CDW or the lattice may undergo a further distortion which makes the two "commensurate" in which case the CDW wavevector is an integral fraction of the underlying lattice. The commensurate state is usually referred

to as a "locked-in" state. Thus, there can be two phase changes associated with a CDW formation: the normal (parent)-to-incommensurate transition (usually second order) and the incommensurate-to-commensurate transformation (first order). Since a CDW is accompanied by a lattice distortion, diffraction techniques (electron, neutron, X-ray) can be used to reveal satellite reflections appearing near the Bragg reflections of the parent phase as a consequence of the new superlattice associated with the formation of the CDW. These satellites are separated from the associated Bragg reflections by a reciprocal lattice vector determined by the CDW wavevector.

In lower-dimensional systems (1-d and 2-d materials) the possibility of CDW formation is enhanced because the simple structures involved lead to a high probability for favorable Fermi surface nesting. But in three-dimensional materials CDW phenomena should, in principle, be rather rare because of the unlikelihood of favorable Fermi surface nesting.

Besides the incommensurate state mentioned above, McMillan [1976] introduced the concept of *discommensurations*, which are narrow out-of-phase regions (defects) between large in-phase (commensurate) regions. Accordingly, the incommensurate-to-commensurate change is viewed as a defect (discommensuration) "melting" transition.

A recent study (Hwang *et al.* [1983]) of the so-called "premartensitic effects" in Ti–Ni alloys has been carried out using Ti–Ni–Fe alloys containing a few percent of Fe in order to suppress the martensitic transformation (lower M_S) while at the same time leaving the temperature regime of the "premartensite" phenomena essentially unchanged. Space does not permit a detailed presentation of this work, but some of the highlights are interesting and these are briefly considered.

The "pre-martensitic" resistivity anomaly (increase with decreasing temperature) in Ti–Ni type alloys has been ascribed (Hwang *et al.* [1983]) to the formation of a three-dimensional charge density wave which appears to evolve in two stages. In the first, the B2 (CsCl) parent is gradually (second order) distorted into an incommensurate phase with decreasing temperature. The incommensurate phase is most simply described as one with "distorted cubic" symmetry. Its gradual transition involves the appearance of new $\frac{1}{3}\langle 110\rangle_{B2}$ superlattice reflections which intensify with decreasing temperature. At this stage, the $\frac{1}{3}\langle 110\rangle$ superlattice reflections are in slightly irrational (incommensurate) positions. As the temperature is further decreased these superlattice reflections "lock in" to commensurate positions which are precise multiples of $\frac{1}{3}\langle 110\rangle_{B2}$, and the resulting structure is rhombohedral (designated the R-phase) as a consequence of a homogeneous distortion which involves an expansion along <111> cube diagonals (and a small contraction in all directions in the plane normal to <111>). This incommensurate-to-commensurate transformation is first-order in nature. Thus, the events observed during the so-called *premartensitic state* of Ti–Ni–X alloys identify with the sequence: B2(CsCl) → distorted cubic (incommensurate) → R, rhombohedral (commensurate). In addition, microdomains typical of anti-phase domains in ordered alloys have been identified using dark field electron microscope images formed by using the new superlattice reflections.

Finally, it should be mentioned that the incommensurate-to-commensurate transformation in Ti–Ni–Fe alloys is also associated with the appearance of twin-like domains, as shown in fig. 35. The twin plane has been identified as a {1 1 0} plane. It would

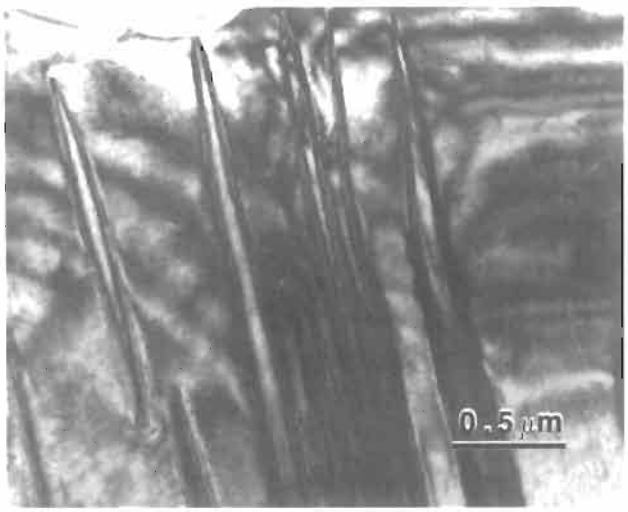


Fig. 35. Twin-related domains as a consequence of an incommensurate-to-commensurate phase change in a $Ti_{50}Ni_{47}Fe_3$(at%) alloy. See text for explanation.

appear that the only difference between the "twins" and the "matrix" within which they exist is that the sense of the <111> expansion (Bain distortion) occurs along different *c* axes in each. In this respect, the formation of twins during the incommensurate-to-commensurate change appears to be a result of overall strain compensation, not unlike that occurring in a martensitic transformation. But in the case of the Ti–Ni–Fe alloys, when the actual martensite is formed at a lower temperature, the {1 1 0} twins and the commensurate R-phase are completely destroyed by the advancing martensite interface, and, seemingly, the "premartensitic" commensurate twinned structure presents no obstacle to the growth of the martensite plates.

In the above description of charge density waves and related phenomena, a measure of consistency has been attempted by referring to the *second order* parent → incommensurate (gradual) change as a transition, as opposed to the *first-order* incommensurate → commensurate (abrupt) change which is viewed as a *transformation*.

Acknowledgment

CMW would like to acknowledge many helpful comments from numerous colleagues and members of his research group, and to thank particularly Professor G. Krauss for providing fig. 17, Professor K. Shimizu for fig. 18, Dr. M. Umemoto for figs. 14–16, and Professor J. C. Williams for fig. 32. HKDHB would like to express his gratitude to Professor Colin Humphreys for his help and encouragement, and to Professor Robert Cahn for providing several important articles.

References: p. 1552.

References

AARONSON, H. I., and K. R. KINSMAN, 1977, Acta Metall. **25**, 367.

AHLERS, M., BARCELÓ, G. and RAPACIOLI, R., 1978, Scripta Metallurgica, **12**, 1075–1078.

AXE, J. D., 1977, in: Electron–Phonon Interactions and Phase Transitions, ed. T. Riste (Plenum, New York) p. 50.

BAIN, E. C., 1924, Trans. AIME **70**, 25.

BOWLES, J. S., and J. K. MACKENZIE, 1954, Acta Metall. **2**, 129, 138 and 224.

BOWLES, J. S., and C. M. WAYMAN, 1979, Acta Metall. **27**, 833.

BULLOUGH, R., and B. A. BILBY, 1956, Proc. Phys. Soc. **B69**, 1276.

CHAN, S. K., and V. HEINE, 1973, J. Phys. **F3**, 795.

CHRISTIAN, J. W., 1970, Phase Transformations, in: Physical Metallurgy, 2nd Ed., ed. R.W. Cahn (North-Holland, Amsterdam) ch. 10.

CHRISTIAN, J. W., 1975, Theory of Transformations in Metals and Alloys, Part I, 2nd edition, Pergamon Press, Oxford.

CHRISTIAN, J. W. and CROCKER, A. G., 1980, Dislocations in Solids, ed. F.R.N. Nabarro, North-Holland, Amsterdam, vol. 3, p. 369.

CHRISTIAN, J. W., 1982, Metallurg. Trans. **13A**, 509.

CHRISTIAN, J. W., 1994, Metallurgical and Materials Transactions A, **25A**, 1821–1839.

CLAPP, P. C., 1973, Phys. Stat. Sol. (b) **57**, 561.

COHEN, M., and C. M. WAYMAN, 1981, Fundamentals of Martensitic Reactions, in: Metallurgical Treatises, eds. J.K. Tien and J.F. Elliott (Metallurgical Society AIME, Warrendale, PA) p. 445.

GHOSH, G. and OLSON, G. B., 1992 Proc. Int. Conf. on Martensitic Transformations — ICOMAT 92, published by the Monteray Inst. Adv. Studies, California, USA, eds. C.M. Wayman and J.F. Perkins, pp. 353–358.

GOLDSTEIN, D. M., and L. J. MCNAMARA, 1979, Report NSWC MP 79–441, Proc. of Heat Engine Conference (U.S. Naval Surface Weapons Center, Silver Springs, MD).

HWANG, C. M., M. MEICHLE, M. B. SALAMON and C. M. WAYMAN, 1983, Phil. Mag. **A47**, 9 and 31.

JACKSON, C. M., H. J. WAGNER and R. J. WASILEWSKI, 1972, National Aeronautics and Space Agency Report SP-5110.

KAUFMAN, L. and COHEN, M., 1958, Progress in Metal Physics, **7**, 165.

KESSLER, H., and W. PITSCH, 1967, Acta Metall. **15**, 401.

KOHN, W., 1959, Phys. Rev. Lett. **2**, 393.

KRAUSS, G., 1980, Principles of the Heat Treatment of Steels (ASM, Metals Park, OH) ch. 8.

KURDJUMOV, G., and L. G. KHANDROS, 1949, Dokl. Akad. Nauk SSSR **66**, 211.

LING, H. C. and OWEN, W. S., 1981, Acta Metallurgica, **29**, 1721–1736.

MAGEE, C. L., 1970, Phase Transformations, ASM, Metals Park, Ohio, U.S.A., pp. 115–156.

MARUKAWA, K. and TSUCHIYA, K., 1995, Scripta Metallurgica et Materialia, **32**, 77–82.

MCMILLAN, W., 1976, Phys. Rev. **B14**, 1496.

OLSON, G. B., and M. COHEN, 1972, J. Less-Common Met. **28**, 107.

OLSON, G. B., and M. COHEN, 1976, Metallurg. Trans. **7A**, 1897, 1905 and 1915.

OLSON, G. B. and OWEN, W. S., 1976, New Aspects of Martensitic Transformations, Japan Institute of Metals, Tokyo, p. 105–110.

OLSON, G. B., and M. COHEN, 1979, Proc. US–Japan Symp. on Mechanical Behavior of Metals and Alloys Associated with Displacive Transformations (Rensselaer Polytechnic Inst., Troy, NY) p. 7.

ONODERA, H., and I. TAMURA, 1979, Proc. US–Japan Symp. on Mechanical Behavior of Metals and Alloys Associated with Displacive Transformations (Rensselaer Polytechnic Inst., Troy, NY) p. 24.

ONODERA, H., H. GOTO and I. TAMURA, 1976, Suppl. Trans. Japan Inst. Met. **17**, 327.

OTSUKA, K., and C. M. WAYMAN, 1977, Reviews on the Deformation Behavior of Materials* **2**, 81.

OVERHAUSER, A. W., 1968, Phys. Rev. **167**, 691.

* A quarterly, published by Freund, Tel Aviv.

OVERHAUSER, A. W., 1971, Phys. Rev. **B3**, 3173.
PATEL, J. R., and M. COHEN, 1953, Acta Metall. **1**, 531.
PEIERLS, R. E., 1955, Quantum Theory of Solids (Oxford Univ. Press, Oxford).
SABURI, T., S. NENNO and C. M. WAYMAN, 1979, Proc. ICOMAT-79, Int. Conf. on Martensitic Transformations, Cambridge, Massachusetts (Alpine Press, Boston, MA) p. 619.
SASS, S. L., 1972, J. Less-Common Met. **28**, 157.
SCHROEDER, T. A., and C. M. WAYMAN, 1977a, Acta Metall. **25**, 1375.
SCHROEDER, T. A., and C. M. WAYMAN, 1977b, Scripta Metall. **11**, 225.
SCHROEDER, T. A., and C. M. WAYMAN, 1978, Acta Metall. **26**, 1745.
SCHROEDER, T. A., and C. M. WAYMAN, 1979, Acta Metall. **27**, 405.
SHIMIZU, K., and C. M. WAYMAN, 1972, Met. Sci. J. **6**, 175.
SMITH, R., and J. S. BOWLES, 1960, Acta Metall. **8**, 405.
TADAKI, T., and K. SHIMIZU, 1970, Trans. Japan Inst. Met., **11**, 44.
TAMURA, I., T. MAKI, H. HATO, Y. TOMOTA and M. OKADA, 1970, Proc. 2nd Int. Conf. on Strength of Metals and Alloys, Asilomar, CA, vol. 3, p. 894.
TAMURA, I., T. MAKI, S. SHIMOOKA, M. OKADA and Y. TOMOTA, 1972, Proc. 3rd Conf. on High Strength Martensitic Steels, Havirov, Czechoslovakia, p. 118.
TAMURA, I., Y. TOMOTA and M. OZAWA, 1973, Proc. 3rd Int. Conf. on Strength of Metals and Alloys, Cambridge, UK, vol. 1, p. 611 and vol. 2, p. 540.
TSUCHIYA, K., TATEYAMA, K., SUGINO, K. and MARUKAWA, K., 1995, Scripta Metallurgica et Materialia, **32**, 259–264.
VAN LANDUYT. J., and C. M. WAYMAN, 1968, Acta Metall. **16**, 803 and 815.
WASILEWSKI, R., 1975, in: Shape Memory Effects in Alloys, ed. J. Perkins (Plenum, New York) p. 245.
WECHSLER, M. S., D. S. LIEBERMAN and T. A. READ, 1953, Trans. AIME **197**, 1503.
WILLIAMS, J. C., D. DE FONTAINE and N. E. PATON, 1973, Metallurg. Trans. **4**, 2701.
WILSON, J. A., F. J. DISALVO and S. MAHAJAN, 1975, Adv. Phys. **24**, 117.
ZACKAY, V. F., E. R. PARKER, D. FARH and R. BUSCH, 1967, Trans. ASM **60**, 252.

General bibliography

Williams, J. C., D. de Fontaine and N.E Paton, 1973, Metallurg. Trans. 4, 2701. The ω-phase as an Example of an Unusual Shear Transformation.
Christian, J. W., 1964, The Theory of Transformation in Metals and Alloys (Pergamon Press, Oxford).
Wayman, C. M., 1964, Introduction to the Crystallography of Martensitic Transformations (Macmillan, New York).
Wilson, J. A., F. J. DiSalvo and S. Mahajan, 1975, Charge-density Waves and Superlattices in the Metallic Layered Transition Metal Dichalcogenides, Adv. Phys. 24, 117.
J. Perkins, ed., 1975, The Shape Memory Effect in Alloys (Plenum, New York).
Proceedings JIMIS-1, 1976, Japan Inst. Met. Int. Symp. on Martensitic Transformations, Suppl. Trans. Japan Inst. Met., vol. 17.
Proceedings ICOMAT-77, 1977, International Conference on Martensitic Transformations, Kiev USSR.
Phase Transformations, 1979, Series 3, No. 11, vols. I and 2 (The Institution of Metallurgists, London).
Proceedings ICOMAT-79, 1979, International Conference on Martensitic Transformations, Cambridge, MA (Alpine Press, Boston, MA).
Nishiyama, Z., 1979, Martensitic Transformation, eds. M. E. Fine, M. Meshii and C. M. Wayman (Academic, New York).
Barrett, C. S., and T. B. Massalski, 1980, Structure of Metals, 3rd Ed. (Pergamon Press, Oxford) ch. 18.
Proceedings of International Conference on Solid–Solid Phase Transformations, 1981. Pittsburgh. PA (Metallurgical Society AIME, Warrendale, PA, 1983).
Cohen, M., and C. M. Wayman, 1981, Fundamentals of Martensitic Reactions, in: Metallurgical Treatises (Metallurgical Society AIME, Warrendale, PA) p. 445.

Bhadeshia, H. K. D. H., 1987, Worked Examples in the Geometry of Crystals, Institute of Materials, London, pp. 1–96.

Bhadeshia, H. K. D. H., 1992, Bainite in Steels, Institute of Materials, London, pp. 1–458.

Christian, J. W., 1982, Metallurgical Transactions A, "Deformation by moving interfaces", 13A, 509–538.

Wayman, C. M., 1989, Materials Science Forum, **56–58**, 1–32.

Wayman, C. M., and H. R. P. Inoue, 1994, Crystallographic Transformations, in: Intermetallic Compounds, eds. J. H. Westbrook and R. L. Fleischer, John Wiley and Sons, New York, U.S.A., Vol. 1, pp. 827–847.

Planes, A., J. Ortín and L. Mañose, 1995, Proc. 3rd European Symposium: a Martensitic Transformation, *J. de Physique,* **5**, suppl. 2, pp. CZ-1 to 551.

CHAPTER 17

PHYSICAL METALLURGY OF STEELS

W. C. LESLIE

Professor Emeritus
Department of Materials Science and Engineering
University of Michigan
Ann Arbor, MI 48109 USA

E. HORNBOGEN

Institut für Werkstoffe
Ruhr-Universität Bochum
D-44780 Bochum, Germany

R. W. Cahn and P. Haasen†, eds.
Physical Metallurgy; fourth, revised and enhanced edition

1. Iron and steel

1.1. Introduction

Steels are alloys of the element iron with carbon and usually several other elements, but the term is also applied to some newer alloys that contain virtually no carbon. There are three major groups of alloying elements, those that form interstitial and substitutional solutions, and those that are immiscible with the crystal lattices of iron (table 1). We shall not deal with the last category in the following but we do not imply that they are of no importance. Commercial steels always contain both types of alloying elements, some added intentionally, others unintentionally during production and processing of the steel. Unwanted elements are, for example, N (from air), S (from coke), P, As (from some ores), and H (introduced during pickling). Examples of the compositions of some commercial steels are given in table 2.

Iron and steels have been our most important metallic materials for several millennia and will continue to be so for the foreseeable future. There are several reasons for this preeminent position:

a) The crust of the earth contains about 4.2 wt% Fe, which can be rather easily reduced to the metallic state.

b) Iron has a melting temperature (1540°C) that allows thermally activated processes at moderate temperatures (T > 400°C) while these processes (except those involving

Table 1
Solubility of the elements in the crystal lattices of iron.

H 1																	He 3
Li 3	Be 2											B 1	C 1	N 1	O 1	F 3	Ne 3
Na 3	Mg 3											Al 2	Si 2	P 2	S 3	Cl 3	Ar 3
K 3	Ca 3	Sc 3?	Ti 2	V 2[b]	Cr 2[b]	Mn 2[b]	Fe	Co 2[b]	Ni 2[b]	Cu 2	Zn 2	Ga 2	Ge 2	As 2	Se 3	Br 3	Kr 3
Rb 3	Sr 3	Y 3?	Zr 2	Nb 2	Mo 2	Tc 2	Ru 2	Rh 2[b]	Pd 2[b]	Ag 3	Cd 3	In 3	Sn 2	Sb 2	Te 3	I 3	Xe 3
Cs 3	Ba 3	La 3?	Hf 2?	Ta 2	W 2	Re 2	Os 2	Ir 2[b]	Pt 2[b]	Au 2	Hg 3	Tl 3	Pb 3	Bi 3	Po 3?	At 3	Rn 3
Fr 3	Ra 3	Ac 3?	Th 3?	Pa 3?	U 3	Np 3?	Pu 3?	Am 3?	Cm 3?	Bk 3?	Cf 3?						

1. interstitial elements
2. substitutional elements
3. practically no solubility

P may be interstitial in the fcc structure.
[b] complete solubility at certain temperatures.

Table 2
Important alloying elements in some commercial steels.

no.	Composition (wt%)								Structure	Application
	C	Mn	Cr	Ni	Mo	W	Si	V		
1	0.12	0.45	–	–	–	–	–	–	ferrite plus pearlite	construction steel
2	0.65	0.70	–	–	–	–	–	–	tempered martensite	tool steel
3	0.50	< 1.0	13.0	–	–	–	–	–	tempered martensite	knife blade steel (stainless)
4	0.10	0.40	5.0	–	0.5	–	–	–	bainite	high temperature steel
5	< 0.15	0.40	–	6.0	–	–	–	–	ferrite	low temperature steel
6	0.80	0.30	4.0	–	–	18.0	–	2.0	tempered martensite and tungsten-carbide	high speed steel
7	0.10	< 1.0	18.0	8.0	–	–	–	–	austenite	stainless steel
8	< 0.07	–	–	–	–	–	4.0	–	ferrite	transformer steel

interstitial diffusion) are very slow at room temperature.

c) There are two phase transformations in solid iron that allow the formation of a wide variety of microstructures. As a consequence a large range of physical and mechanical properties can be obtained.

d) Iron is ferromagnetic at temperatures below 768°C the Curie temperature, T_c, a highly important characteristic which leads to several anomalies of other properties.

Despite the long history of steel there is still room for innovations, even in mass production — witness the advent, since about 1975, of micro-alloyed steels, dual-phase steels and interstitial-free (IF) steels. Steady progress continues to be made in the understanding of the finer details of microstructures in steels, the development of crystallographic orientations, the distribution and interactions of solute elements and the control of mechanical properties by precipitation of solutes from solid solution. These advances have been made possible through the use of new tools that enable us to observe internal structure on an atomic scale.

1.2. Some properties of pure iron

The iron atom is situated towards the end of the first transition series with an electron configuration $Ar3d^64s^2$. The neighboring elements with larger atomic number than iron (Co, Ni, Cu) crystallize in close-packed structures until the 3d states are filled in Cu. The

References: p. 1618.

transition elements with less-filled 3d states (β-Ti, V, Cr, δ-Mn) have a bcc structure. The maximum of the binding energy for Cr with its half-filled 3d band indicates that a contribution of covalent bonding by 3d-states may stabilize this structure (fig. 1).

Iron is found between these two groups of elements. It crystallizes in both the fcc ($912° < T_\gamma < 1394°C$) and the bcc lattices ($1394° > T_\alpha < 1538°C$) and again at $T_\alpha < 912°C$. The re-occurrence of the bcc structure at high temperatures can be rationalized by dividing the difference in free energy of bcc and fcc iron, $\Delta F_{\alpha-\gamma}$, into a magnetic and a nonmagnetic term (ZENER [1955]). The magnetic term is negligible up to the temperatures at which magnetization of α-Fe is nearly constant. At higher temperatures (depending on T_c) the bcc structure is increasingly stabilized due to the entropy of demagnetization. Without this effect, α-Fe should transform to γ-Fe at about 700°C without another transformation at 1394°C. In fig. 2, $\Delta F_{\alpha-\gamma} = F_\alpha - F_\gamma$ is plotted as a function of temperature. The contribution due to magnetic uncoupling in the α-phase is shown as a dashed line. This term produces the curvature of $\Delta F_{\alpha-\gamma}(T)$ as determined by JOHANNSON [1937] and the re-occurrence of the bcc structure above 1394°C.

The increasing value of $\Delta F_{\alpha-\gamma}$ with decreasing temperature indicates that an increasing pressure is needed to produce the closest packed iron at temperatures below 912°C. At 20°C this pressure is 11.3 GPa, determined by both static and dynamic measurements. Iron can exist in at least three crystal structures, fcc γ, bcc α and hcp ε. The closest packed phase, ε, exists at pressures in excess of 11.3 GPa and at temperatures from ambient to at least 2000K. A triple point at which α, γ and ε coexist occurs at about 775K and 10 GPa (fig. 3). There is some evidence that a fourth or even a fifth phase may exist (BOEHLER [1986]) but confirmation is required.

There is no experimental evidence for the existence of thermal vacancies in quenched α-Fe, perhaps because of the high ratio of energy of formation, u_F, to energy of motion, u_M: $u_F/u_M \approx 4$, as compared to about 1 for the fcc lattice. Vacancies probably anneal out during quenching of α-Fe (JOHNSON [1960]). It is likely that these results are strongly influenced by interstitial elements.

The Burgers vector of the minimum energy dislocation in α-Fe is $(a/2)\langle 111\rangle$. Segments of $a\langle 100\rangle$ can form by the reaction $(a/2)[1\bar{1}1] + (a/2)[11\bar{1}] \rightarrow a[100]$. The

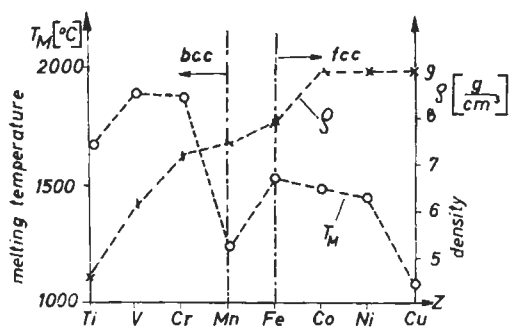


Fig. 1. Melting temperature T_M, and density, ρ, versus atomic number Z of the elements of the first transition period.

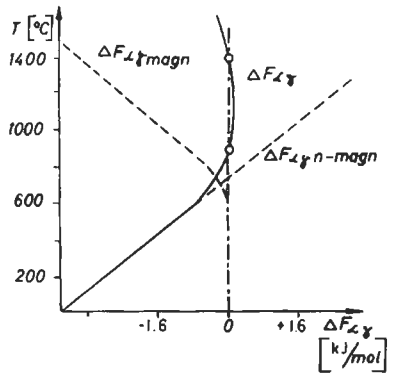


Fig. 2. Free energy difference $\Delta F_{\alpha\gamma}$ resolved into magnetic and nonmagnetic components (after ZENER [1955]).

stacking-fault energy of α-Fe (as of the other bcc transition metals) is very high. Stacking faults have therefore not been observed in α-Fe even by subtle weak beam electron microscopy. Details of the core structure of dislocations in the bcc lattice and their consequences for crystal plasticity are discussed in ch. 20, § 3.2. The probability for the occurrence of annealing twins is low but twins generated during cooling through the

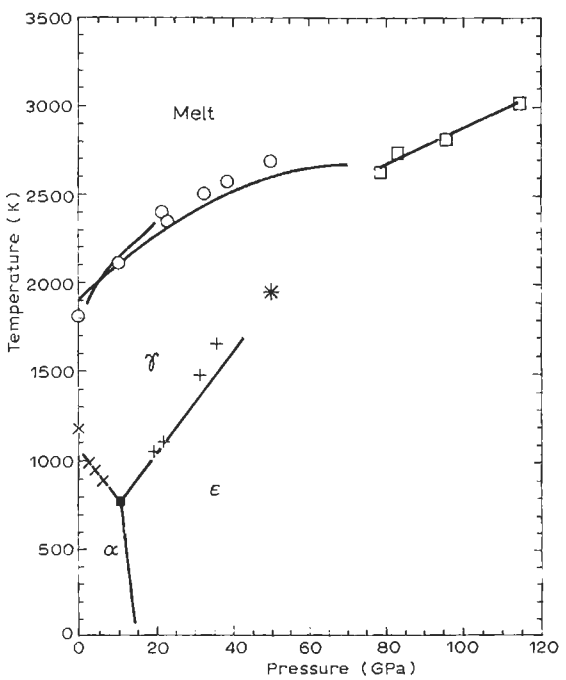


Fig. 3. Iron phase relations (SAXENA *et al.* [1993]).

$\gamma \rightarrow \alpha$ transformation are common. In γ-Fe the Burgers vector of the complete dislocations is $(a/2)<110>$. The stacking-fault energy (at 1000°C), determined from the frequency of annealing twins, is 75 mJ/m^2 (NUTTING and CHARNOCK [1967]). (See ch. 28, fig. 55). Stacking faults and annealing twins are frequently found in γ-Fe and its alloys.

Many discontinuously during the phase transformations of iron. They are often used to determine the transformation temperatures in steels. There are discontinuities of the self-diffusion coefficient of iron at the transformation temperatures. The diffusivity of iron in α is about 10^2 times higher than in γ at the same temperatures (fig. 4). This can be explained qualitatively by the closer packing of γ-Fe (BIRCHENALL [1951]). Below T_c there is an anomalous decrease of the diffusion coefficient of about 60% during the transition to complete magnetic order. A change in equilibrium vacancy concentration, not magnetostriction, was used as an explanation for this behavior (BIRCHENALL and BORG [1960]). Since then, much work has been done on this anomaly (reviewed by JÖNSSON [1992]); it is now attributed to a change in the activation energy for atomic jumps.

An amorphous structure can be obtained for pure iron neither by vapor deposition nor by splat cooling down to temperatures close to 0 K. Such a structure is exclusively found in certain alloys of iron. (See also sonochemical approach, ch. 19, § 3.6).

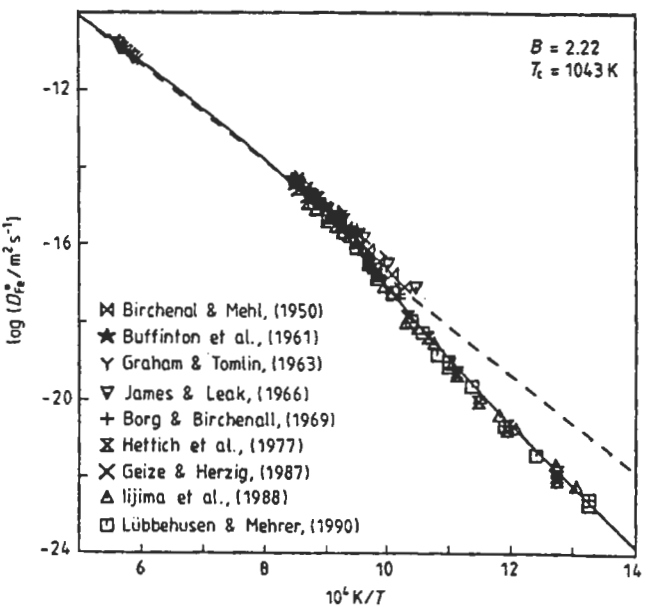


Fig. 4. Self-diffusion coefficient of paramagnetic ($D_{\alpha p}$) and ferromagnetic ($D_{\alpha f}$) α-iron; Q in kJ. (After JÖNSSON [1992]).

2. *Alloys of iron*

2.1. Interstitial alloys

The elements H, C, N and O can occupy interstitial sites in the lattices of iron. All these elements fulfill the condition $r_I \ll r_{Fe}$ (table 3), The interstitial voids in fcc iron are larger than those in the bcc lattice (fig. 5). The largest void in the fcc lattice, 0.052 nm in radius, is at an octahedral site; the next largest is 0.028 nm in radius. The radius of a carbon atom is 0.08 nm and for nitrogen is 0.07 nm. These atoms in solution reside in the octahedral sites, thereby expanding the lattice of austenite.

In bcc iron the interstitial voids are smaller than those in austenite. Although the octahedral site is only 0.019 nm in radius as compared to 0.036 nm for the tetrahedral site, C and N atoms reside in the octahedral sites, presumably because there they have only two iron atoms as nearest neighbors rather than four in the tetrahedral sites. It is this difference in size of the interstitial sites that permits the higher solubility of C in austenite than in ferrite which in turn allows the formation of a great variety of microstructures as C precipitates from supersaturated solid solution in α iron. As shown in fig. 6, diffusivities in bcc iron depend upon the size of solute atoms.

The solubilities depend on the phase with which the α-iron interstitial solid solution is in equilibrium: stable equilibria would be, for example, those with nitrogen gas or with graphite. There is a high activation energy of nucleation of these two elements in iron because of large surface energy terms. Metastable phases form instead, with lower activation energy and higher solubilities in the iron-rich solid solution. Phases that form in Fe–C and Fe–N alloys are listed in table 4. In fig. 7 the solubilities of some of these phases are shown. Which phase forms depends on temperature and time of heat treatment and on the imperfection structure of the solid solution. The probability of formation of nitrogen gas or graphite is so low that these phases rarely precipitate in steels. An

Table 3
Maximum solubility of interstitials in α-Fe.

element	atomic radius (for coord. no. 12) (nm)	max. solubility (at%)	at temperature (°C)	in equil. with
H	< 0.04	$1\text{-}2\times10^{-3}$	905	H_2
B	0.098	0.005[a]	912	Fe_2B
C	0.091	0.10	727	Fe_3C
N	0.092	0.38	585	Fe_4N
O	~0.092	$0.7\text{–}13\times10^{-4}$ [b]	906	FeO

[a] Ohtani [1988]
[b] Swisher [1967]

References: p. 1618.

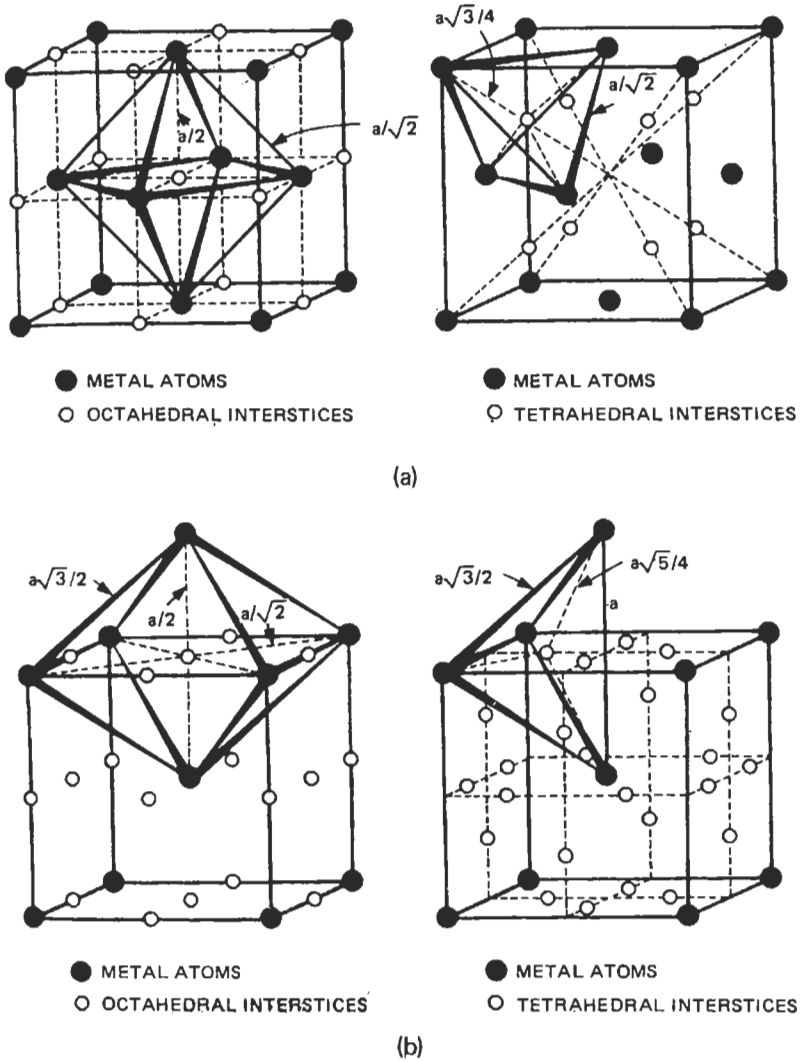


Fig. 5. Interstitial voids in iron, (a) Interstitial voids in the fcc structure, octahedral (1) and tetrahedral (2). (b) Interstitial voids in the bcc structure; octahedral (1) and tetrahedral (2).

important exception occurs when carbon or low-alloy steels are held in the range 450°-650°C for hundreds or thousands of hours. The pearlite in the steels decomposes to ferrite and graphite. Normally Fe_3C (cementite,) or less stable carbides form instead of graphite. Recently, COTTRELL [1993] has published an electronic interpretation of the high stability and curious crystal structure of cementite. In fig. 8 the metastable Fe–Fe_3C diagram is shown in addition to the stable Fe-graphite diagram.

CHIPMAN [1972] gives the following equations for the solubility of carbon in α-iron in equilibrium with graphite, cementite or ε carbide ($Fe_{2,4}C$):

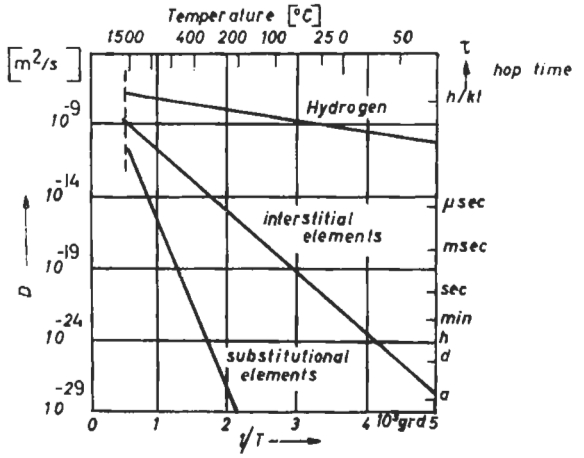


Fig. 6. Average diffusion coefficients D of substitutional (s) and interstitial elements in iron. $Q_s = 250 \pm 40$ kJ/mol; in α-Fe: $Q_C = 83$ kJ/mol; $Q_N = 74$ kJ/mol; $Q_O = 95$ kJ/mol.

$$\text{graphite: } \log [\text{C}]_{\text{ppm}} = 7.81 - \frac{5550}{T}$$

$$\text{cementite: } \log [\text{C}]_{\text{ppm}} = 6.38 - \frac{4040}{T}$$

$$\varepsilon\text{-carbide: } \log [\text{C}]_{\text{ppm}} = 4.06 - \frac{1335}{T} \text{ (approximate)}$$

HASSEBE *et al.* [1985] agree with these solubilities for cementite and graphite if the slight deviation from the log [C] and $1/T$ linear relation caused by the magnetically induced anomaly below the Curie temperature is ignored.

There are several important composition ranges in the Fe–Fe_3C diagram (fig. 8):

a. The maximum solubility of carbon in γ-iron (austenite) 2.11 wt% (9.12 at%) is the limiting composition separating steel and cast iron.
b. Cast iron has usually a hypo-eutectic or eutectic composition: $C_E = 4.3$ wt% = 17.3 at%.

Table 4
Metastable compounds that precipitate in Fe–C and Fe–N alloys

	Composition	Crystal structure	Nucleation sites	Formation temperatures
cementite	Fe_3C	orthorhombic	dislocations, grain b.	> 200°C
ε-carbide	$Fe_{2.4}C$	hcp	matrix, dislocations	< 250°C
γ'-nitride	Fe_4N	fcc	dislocations, grain b.	> 250°C
α'-nitride	Fe_8N	fig. 14	matrix	< 300°C

References: p. 1618.

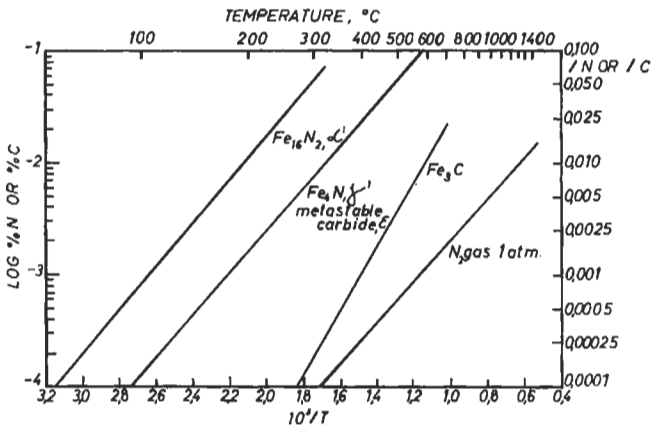


Fig. 7. Solubilities of C and N in α-Fe in equilibrium with different phases (average values from several investigations, after LESLIE and KEH [1965]).

$$L_E \rightarrow \gamma\text{-Fe(C)} + Fe_3C.$$

c. Steels can be subdivided in hypo-eutectoid and hypereutectoid depending on whether their carbon content is smaller or larger than 0.77 wt% (3.48 at%). The microstructure that originates from the eutectoid reaction

$$\gamma\text{-Fe(C)} \rightarrow \alpha\text{-Fe(C)} + Fe_3C$$

is known as *pearlite*, if it is formed by lamellar growth.

d. The maximum solubility of carbon in α-Fe is 0.0218 wt% (0.10 at%). At lower carbon contents, carbides can only form by precipitation from ferrite.

The lower the carbon content, the lower the temperature below which precipitation can start. Below 250°C the less stable ε-carbide can form. The smallest carbon contents that can be obtained currently $(1\text{–}2)\times 10^{-4}$ wt%, correspond to the solubility at about 50°C.

Nitrogen has a relatively large solubility in α-Fe as compared to carbon. The solubility of oxygen is very small (SWISHER and TURKDOGAN [1967]). The difference between the solubilities of N, C and O in α-iron cannot be explained on the basis of atomic-size ratios alone (table 3).

Solid solubilities are determined on annealed specimens of the highest obtainable purity. Interstitial solute atoms tend to segregate to lattice defects because they offer sites that are energetically more favorable than those in the undisturbed lattice. These can be point, linear and planar defects such as vacancies, some substitutional solute atoms, dislocations, sub-boundaries, grain boundaries and free surfaces. However, in annealed specimens the concentration of such defects is insufficient to cause significant variation in the measured solubility, except perhaps of B and 0 whose solubilities in α-iron is extremely low. The important effects of segregation of solutes in steels are treated in detail later.

The precipitation of interstitial solutes from solution in α-iron at low temperatures is

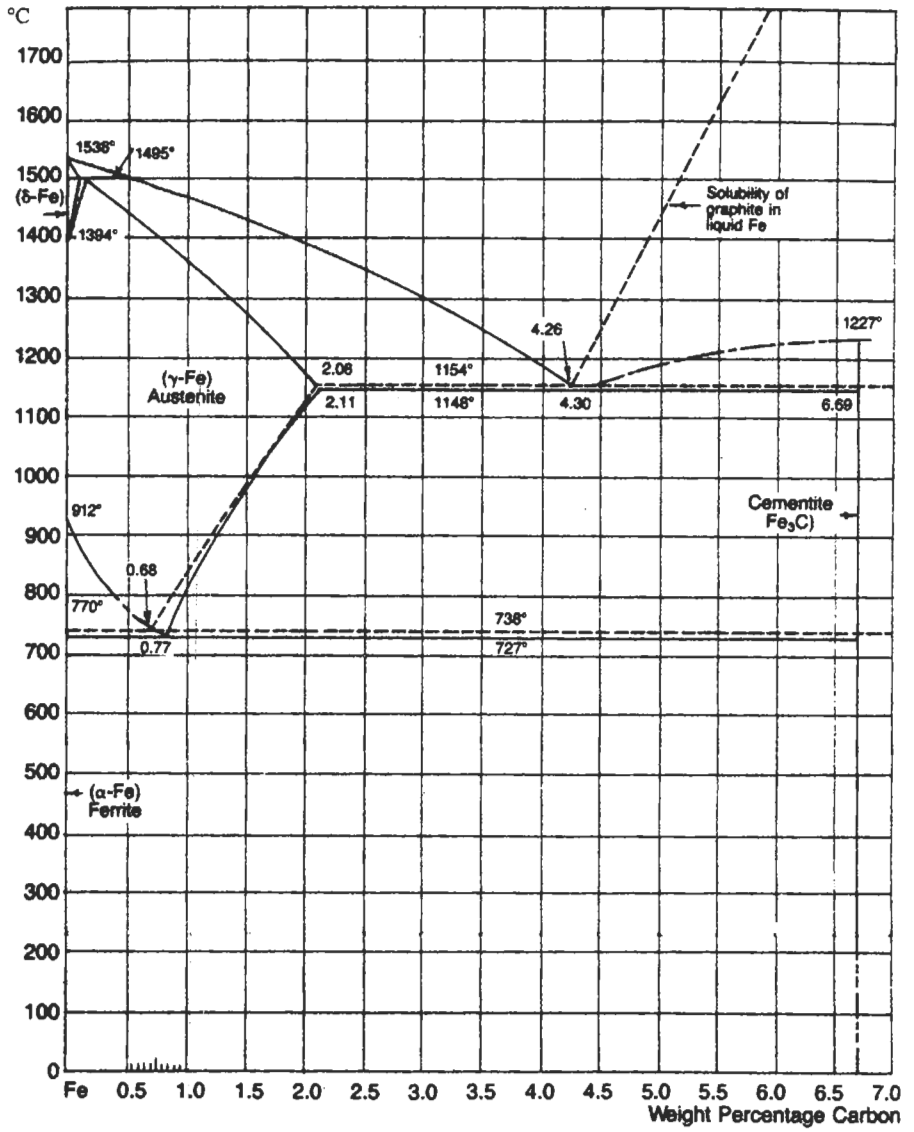


Fig. 8. The Fe–C equilibrium diagram up to 7% carbon. Solid lines indicate Fe–Fe_3C diagram; dashed lines indicate Fe-graphite diagram.

strongly affected by lattice defects, especially dislocations (table 4). The presence of the defect serves to minimize the strain energy term in the activation energy for nucleation of the precipitates, so that the distribution of precipitates is determined by the distribution of defects.

References: p. 1618.

2.2. Substitutional alloys

The elements marked 2 (table 1) are completely or partially soluble as substitutional elements in the iron lattices. The Hume–Rothery rules for solubility are well fulfilled for alloys with transition elements: large solubility of the elements in the neighborhood of Fe (Ni, Co, Mn, Cr and V), limited solubility of Ti, and practically no solubility of Ca and K. The alloying elements have different energies of solution in α- and γ-iron. The $\Delta F_{\alpha\gamma}(T)$ curve (fig. 2) is shifted in different directions depending on whether $\Delta F_{\alpha\gamma}$ is increased or decreased by the alloying element (fig. 9). This leads to an increase or decrease of the temperature range in which austenite is stable. For simplicity the curves are drawn for α- and γ-phase of the same composition (i.e., T_0-temperatures are shown for $\Delta F_{\alpha\gamma}=0$). We can divide the alloying elements into two groups, depending on whether they close or open the γ-field. Typical phase diagrams are shown in figs. 8, 10 and 11. The following rules tell how the alloying elements behave in this respect:

a) the γ-field is opened by all interstitial elements, because the fcc lattice provides sites of smaller strain energy than does the bcc lattice;
b) the γ-field is opened by elements with a fcc or hcp structure (Cu, Au, Ni, Pt, γ-Mn);
c) the γ-field is closed by elements that form Hume-Rothery phases with Cu, Ag and Au, even if they form a fcc lattice (Al, Si, P, Zn, Ga, Ge, As etc.);
d) the γ-field is closed by the bcc transition metals (V, Ti, Mo, W, Cr etc.); The phase diagram of Fe–Cr alloys shows that small additions of chromium lower the $\gamma \rightarrow \alpha$ transformation temperature while the loop is closed by larger Cr-additions. ZENER [1955] has pointed out that this behavior can be explained by the magnetic term of the free energy of α-Fe (fig. 2). The addition of Cr to Fe has two effects: It shifts the $\Delta F_{\alpha\gamma}(T)$ curve as shown in fig. 9 and it lowers the Curie temperature T_c. Consequently the term $\Delta F_{\alpha\gamma\,magn}$ (fig. 2) appears at a lower temperature and changes the shape of the curve to that shown in fig. 11b. This T_c-effect should affect transform-

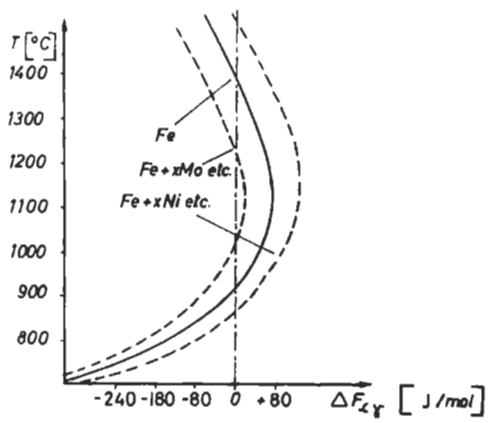


Fig. 9. The effect of solute elements with larger heat of solution in γ-Fe than in α-Fe (e.g. Mo) and vice versa (e.g. Ni) on the transformation temperatures ($\Delta F_{\alpha\gamma}=0$), assuming constant concentrations of α and γ-solid solution and no change of T_c in α-Fe (schematic).

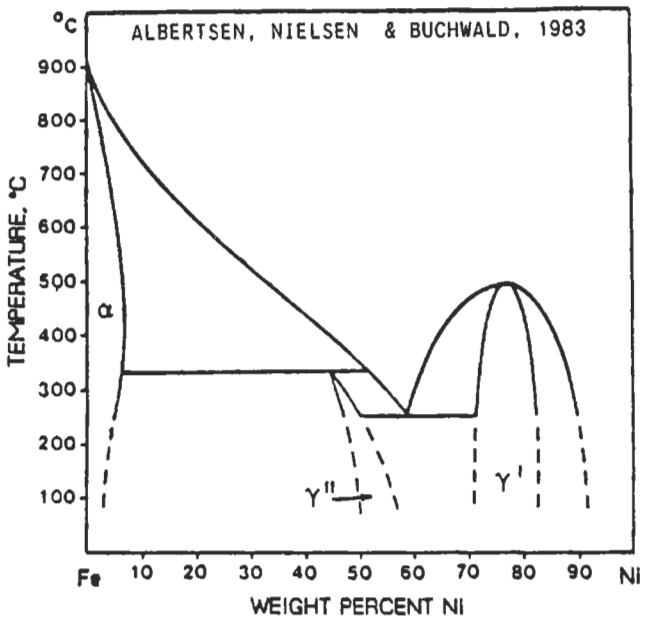


Fig. 10. Proposed Fe–Ni phase diagram based on electron microscopy, electron microprobe studies, and Mossbauer spectroscopy of iron meteorites (ALBERTSEN *et al.* [1983]).

ation temperatures of all alloys of iron. In Fe–Cr alloys it is evident because the difference in heat of solution of Cr in α-Fe and γ-Fe is relatively small.

In many alloys the α- and γ-field is bounded by a *miscibility gap* (fig. 11c). The α- or γ-solid solutions are then in equilibrium with an intermetallic compound or a solid solution. Special attention is drawn to the ordered bcc structures Fe_3Al and Fe_3Si that can form as stable coherent precipitates in α-iron. In addition metastable equilibria occur

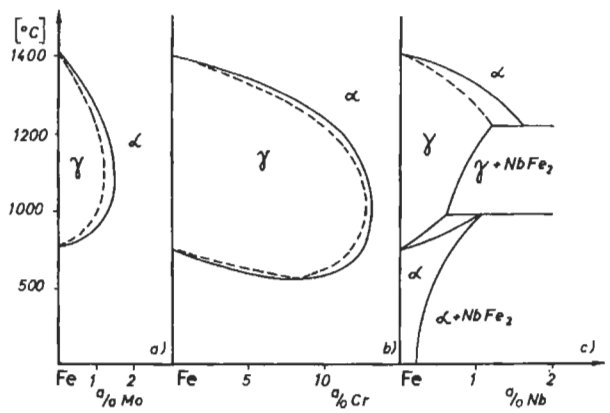


Fig. 11. Closed γ-fields in Fe–Mo, Fe–Cr, and Fe–Nb phase diagrams.

frequently in substitutional alloys of Fe which form the matrix of microstructures. Such coherent phases have been found in binary bcc solid solutions with Cu, Au, Al, Mo, W, Be (HORNBOGEN *et al.* [1966]). They play an important role in maraging steels.

The crystal structure of the solid solution of Fe at room temperatures, as it follows from the phase diagrams (figs. 10, 11), is used to classify steels. If they consist predominantly of α-Fe solid solutions they are known as *ferritic* steels; solid solutions of γ-Fe are the base of *austenitic* steels. A metastable fcc structure can be obtained, for example, from the ternary alloy indicated in fig. 12. This metastable phase is the well-known austenitic stainless steel. Precipitation of stable or metastable coherent, ordered phases (γ') is very common in the fcc iron alloys. Such ($\gamma + \gamma'$) microstructures provide the base for creep resistance of precipitation-hardened stainless steels and Ni-base superalloys.

2.3. Interstitial plus substitutional alloys

Even plain carbon steels contain substitutional elements, especially Mn, present either intentionally or inadvertently. Therefore interstitials are always in interaction with atoms substituted in the iron lattice. The interaction energy as compared to that with iron, Δu, can be positive or negative. $\Delta u = (u_{Fe-i}) - (u_{s-i})$. The lattice of iron with interstitial (i) and substitutional (s) atom in solution can be expected to show a distribution of the i-atoms that depends on that of the s-atoms (fig. 13). A qualitative idea of the sign of Δu can be obtained from the stability of the carbides of the s-atoms but the situation is complicated by the effects of substitutional solutes on carbon activity in cementite and in ε carbide. For example, aluminum decreases the activity coefficient of carbon in ferrite and increases solubility; silicon increases the activity coefficient of carbon in ferrite and decreases solubility (LESLIE and RAUCH [1978]). Manganese decreases the activity coefficient of carbon in ferrite but also decreases carbon activity in cementite, the result being no change in the solubility of carbon in ferrite, at least up to 1% Mn (SAITOH and USHIODA [1989]).

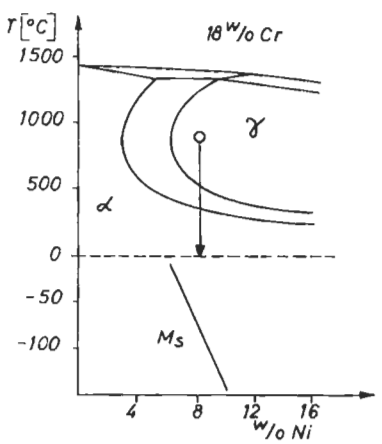


Fig. 12. Section of the Fe–Cr–Ni diagram showing the composition and heat treatment of austenitic stainless steel. Below M_s the metastable fcc alloys become unstable and transform to bcc martensite.

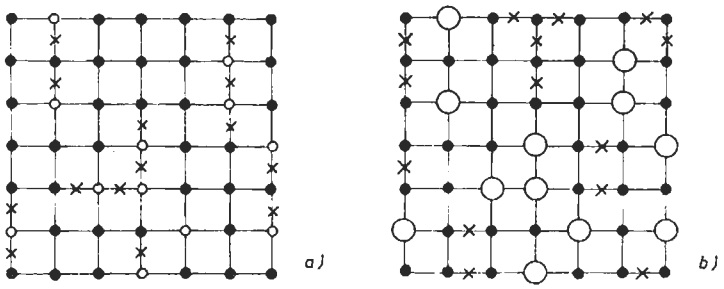


Fig. 13. Schematic drawing of the structure of interstitial (×) plus substitutional (○) solid solutions, in iron (●): (a) attraction, (b) repulsion between interstitial and substitutional solute.

A very large number of carbides can form in Fe–s–C alloys in stable or metastable equilibria. In table 5 the crystal structures of carbides of the elements of the first transition series are listed. There are the following possibilities for the composition of carbides that form in Fe–s–C alloys.

a) Fe_xC_y solute-free carbide;
b) $Fe(s)_xC_y$ carbide with substitutional composition of the alloy;
c) $[Fe_xs_z]C_y$ carbide with Fe replaced by a definite amount of s;
d) s_xC_y iron-free carbide.

Many $[Fe_xs_z]C_y$-type carbides occur in alloy steels, still more in rapidly solidified alloys of iron, in addition to those listed in tables 4 and 5.

The M_6C η-carbide is cubic with 96 metal and 16 carbon atoms in the elementary cell. The metal atoms are iron and group V, VI transition metals. A typical representative, $[Fe_4W_2]C$, forms in high-speed steels. W, Mo and Cr can be substituted in small, Mn in large amounts in Fe_3C. If a higher concentration of group V or VI elements is present a mixed carbide of the type $Cr_{23}C_6$ forms (table 5). Fe, Mn, V, Nb, Mo and W can replace Cr in this carbide in a wide range of concentrations. In tungsten steels it has a composition $[Fe_{21}W_2]C_6$. Another important group of mixed carbides are the χ-phases. Their structure can be derived from substitutional compounds with the β-Mn structure. They form in Cr–Mo and Cr–W steels. C dissolves in the β-Mn structure and stabilizes it. A pure χ-phase has, for example, a composition $Fe_{28}Cr_{23}W_7$ which is shifted to higher W-contents by C additions.

Metastable carbides that form in ferrite are characterized by greater similarity to α-iron as compared to the more stable carbides. This allows them to form semicoherent or coherent interfaces with α-Fe. Figure 14a shows the structure which fulfills the requirements for full coherency if the particles are small. An ordered arrangement of interstitials that only distorts the bcc lattice (α') has been found by JACK [1951] in Fe–N alloys. This structure forms during ageing of Fe–N alloys at a low temperature. The iron-free carbides TiC, VC and NbC have face-centered-cubic structures with lattice parameters 20–25% greater than those of the austenite from which they precipitate (fig. 14b). They can form as ultra-fine dispersions, which raise the yield stress of micro-alloyed steel (HSLA-steel, i.e., high-strength low-alloy steel).

References: p. 1618.

Table 5
Composition and melting temperatures [°C] of carbides that form in equilibrium with transition metals (after GOLDSCHMIDT [1967]).

TiC	V_2C	$Cr_{23}C_6$	$Mn_{23}C_6$	(Fe_3C)	(Co_3C)	(Ni_3C)
3140	2200	1580	1010	*g	*g	*g
ZrC	Nb_2C	Mo_2C	Tc	Ru	Rh	Pd
3550	3100	2410	g?	g	g	g
HfC	Ta_2C	W_2C	Re	Os	Ir	Pt
3890	3400	2800	g	g	g	g

TC cubic (B1)
T_2C hcp (L'3)
$T_{23}C$ cubic ($D8_4$)
T_3C orthorhombic (fig. 4.8, DO_{11})

* carbide not stable in liquid state

3. *Transformation reactions*

3.1. Pearlite

The iron-rich side of the Fe–Fe_3C phase diagram (fig. 8) provides the background for a large number of solid-state reactions. In alloy steels, reactions can be controlled by substitutional or interstitial diffusion. The microstructure of a steel can only be understood if all the individual reactions and their mutual effects are known.

If an alloy of eutectoid composition is cooled below the eutectoid temperature T_p the pearlite reaction can occur. Nucleation of two new types of crystals inside the austenite crystal is unlikely because of the high surface- and strain-energy required. Instead, transformation begins at austenite grain boundaries and moves into grains with a reaction front that leaves an aggregate of lamellar particles behind. Austenite grain corners are favored nucleation sites, followed by grain edges and planar grain boundaries. In commercial steels nucleation can occur at inclusions. In hypereutectoid steels cementite forms first during transformation of austenite, then pearlite can nucleate at the austenite–cementite boundaries. Similarly, in hypoeutectoid steels ferrite forms first and nucleation can occur at ferrite–austenite boundaries.

Each of these so-called pearlite nodules consists of two interlocking single crystals, one of cementite, the other of ferrite (HILLERT [1962]). Each nodule continues to grow (at velocity G) within an austenite grain until it impinges on another in the same grain. The pearlite reaction is completed at a time t_f that depends upon the austenite grain diameter, d:

$$t_f = 0.5\frac{d}{G}$$

where d/G is the time required for one nodule to occupy one grain. It is now generally considered that the rate-controlling step in the growth of pearlite is the rate of diffusion of carbon in the pearlite–austenite boundary modified by some volume diffusion in the austenite ahead of the

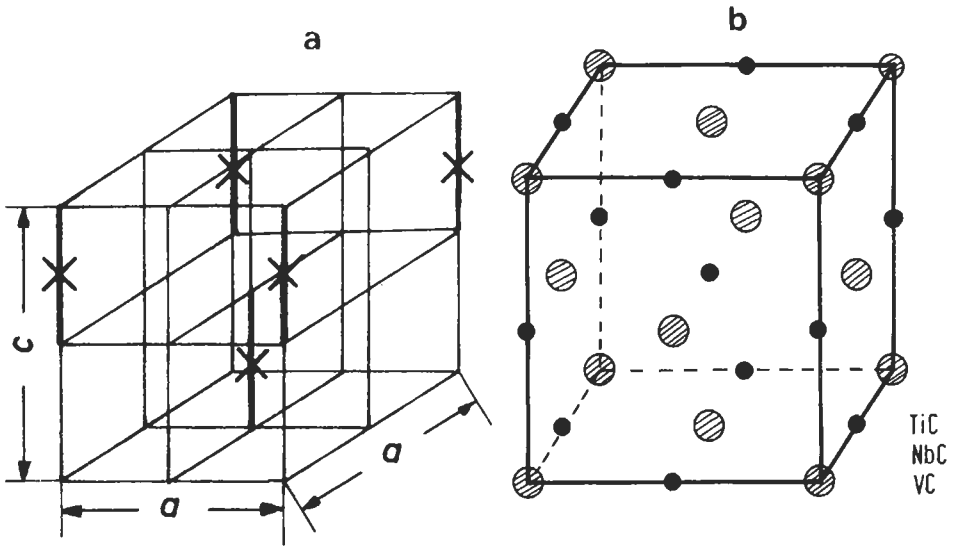


Fig. 14. (a) Structure of the metastable compound $Fe_{16}N_2(\alpha')$, crosses indicate N atoms. $a = 5.72$ Å, $c = 6.29$ Å (after JACK [1951]). (b) Structure of the carbides TiC, VC, NbC and carbonitride as Ti(C,N).

moving interface to provide the variation in carbon content of austenite at the interface. Recent studies (ZHOU and SHIFLET [1991]) have shown that pearlite growth is accomplished by the lateral migration of steps across the growth front of cementite and ferrite.

The addition of a third element to the Fe–C system inevitably decreases the rate of formation of pearlite. These effects occur in two ways:

a) elements that segregate to boundaries can reduce the frequency of nucleation of ferrite. Boron is spectacularly effective in this regard.
b) Substitutional solutes can influence both the rate of nucleation and the rate of growth of pearlite. They can do this by changing the activities of C in the three phases involved or by partitioning between ferrite and cementite at the pearlite–austenite interface. Since the latter process requires diffusion of substitutional atoms it is most effective at elevated transformation temperatures. Molybdenum is particularly effective in decreasing the rate of growth of pearlite nodules whereas cobalt substantially increases the rate.

When a hypoeutectoid steel is slowly cooled from the austenite region, ferrite forms below the upper critical temperature, which is strongly dependent upon alloy content, and continues to form until the eutectoid temperature, A_{e1}, is reached. With more rapid cooling ferrite can form well below A_{e1}. The ferrite takes on four distinct morphologies (AARONSON [1962]). Appearing roughly in order of decreasing temperatures, these are termed *grain-boundary allotriomorphs*, *Widmannstätten side-plates*, *intragranular idiomorphs* and *intragranular plates*. The Kurdjumov–Sachs relationship between austenite and ferrite:

$$\{111\}\gamma \;//\; \{110\}\alpha$$
$$\langle 110\rangle\gamma \;//\; \langle 111\rangle\alpha$$

is observed. The Widmannstätten ferrite is often called *acicular ferrite*.

The formation of cementite plates from hypereutectoid steels follows a similar pattern but with the Pitsch orientation:

$(100)_{cem}$ // $(554)\gamma$
$(010)_{cem}$ // $(110)\gamma$
$(001)_{cem}$ // $(225)\gamma$

Figure 15 shows some of the typical microstructures which are observed in Fe–C steels of various compositions.

3.2. Martensite

If austenite is cooled fast enough so that no diffusion-controlled transformation reaction can begin until an extent of under-cooling can be reached at which the fcc structure becomes highly unstable, it transforms by shear to a distorted bcc lattice, supersaturated with carbon, called martensite. The minimum cooling rate required for the formation of martensite varies widely depending upon the composition of the steel, which in turn controls the temperatures and rates at which non-martensitic transformation products form. The temperature at which diffusionless transformation starts is designated M_s. The M_s-temperature is always lower than the temperature T_0 at which α- and γ-solid solutions of the same composition are in equilibrium (T_0, $\Delta F_{\alpha\gamma} = 0$, fig. 9). The M_s-temperature can be above or below room temperature, and steels can be classified accordingly. For austenitic steels M_s << room temperature, martensitic steels M_s > room temperature; both must of course be cooled faster than the critical rate to avoid formation of other transformation products. Between T_0 and M_s metastable austenite can transform partially into martensite under external stress. During the cooling below M_s the martensitic transformation continues until it stops at the martensite finish temperature, M_f. Frequently, some small amount of austenite is retained.

Unalloyed iron or substitutional solutions of γ-iron can transform on cooling to a bcc structure.

As noted before, interstitial atoms occupy octahedral sites in both α and γ iron. Only one set of such sites, Z, is common to both lattices. Since the carbon atoms have no time to move during the shear (displacive) transformation, the number of carbon atoms at these Z sites in bcc iron is increased, causing a tetragonal distortion of the bcc lattice. The tetragonality is minimal below about 0.20% carbon because of the association of carbon atoms with lattice defects produced by the displacive transformation (fig. 16).

In addition to the shear (Bain strain) that leads to the change in crystal structure, a second deformation, a lattice invariant strain, is required to allow the martensite to correspond to the original crystal. The amount of shear for the $\gamma \rightarrow \alpha$ transformation is about 0.12 (WAYMAN [1964]). The internal plastic deformation takes place by twinning or slip (KELLY and NUTTING [1961]. It leads to martensite crystals that contain dislocations and thin twins. The M_s of a steel depends upon a) the equilibrium temperature, T_0 and b) the strength of the austenite. Alloying elements affect both factors, as shown by one of several empirical equations used to calculate M_s:

$$M_s(°C) = 539 - 423\ C - 30.4\ Mn - 12.1\ Cr - 17.7\ Ni - 7.5\ Mo$$

(ANDREWS [1965]) where the chemical symbols represent wt% of the alloying elements.

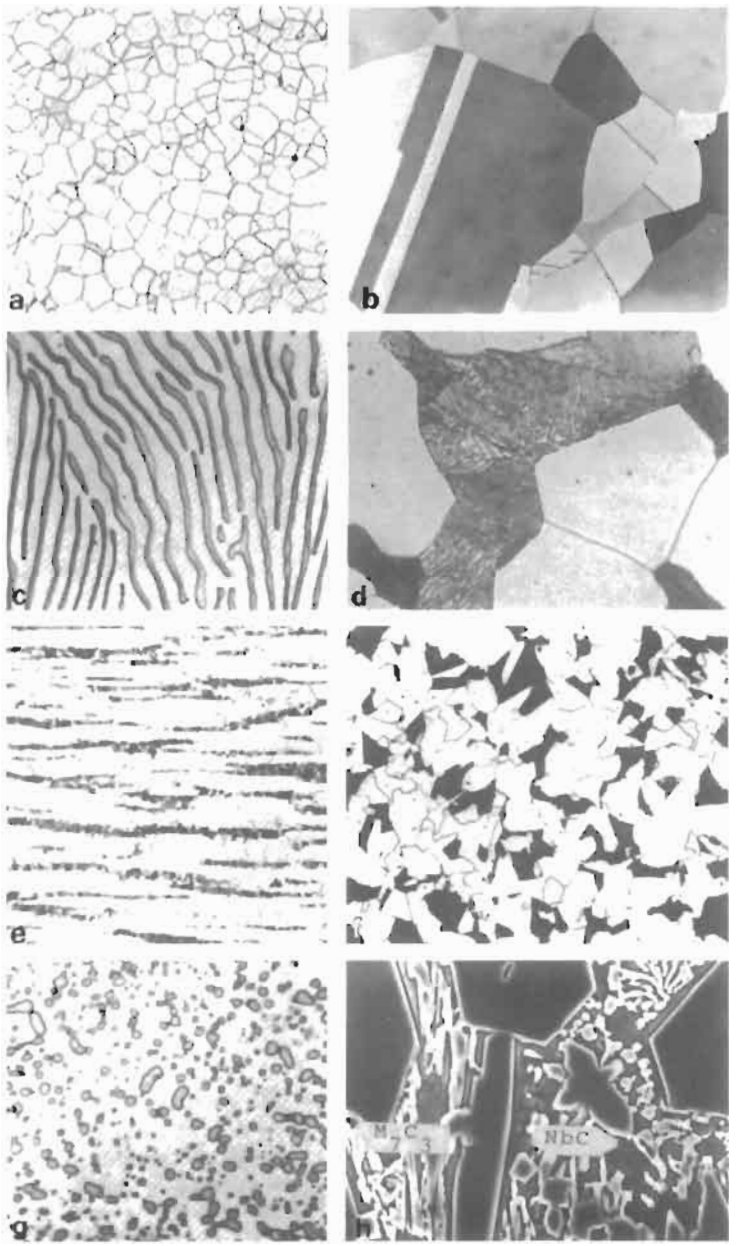


Fig. 15. (a) Single-phase microstructures; ferrite. 100 ×. (b) Austenite, 1.2 wt% C, 12 wt% Mn, 1100°C water quenched. 200 ×. (c) Lamellar structure: pearlite, 0,8 wt% C. 2500 ×. (d) Ferrite–pearlite structure, 0.4 wt% C. 500 ×. (e) Anisotropic ferrite–pearlite structure (banding in rolled steel, dark areas: pearlite). 100 ×. (f) ferrite–martensite structure, dual-phase, 0.11 C, 0.15 Mo, 0.24 P wt%, intercritical annealing: 60 min, 725°C. 150 ×. (g) Dispersed Fe_3C-particles, 1 wt% C. 1000 ×. (h) Hard-facing Fe-base alloy, high volume fraction of NbC and M_7C_3, M = Cr, Fe, 5.5 C, 30 Cr, 7 Nb wt%. 1000 ×. Figs. 15a, c, e, g. J.R. VILELLA, U.S. Steel Corp. Fig. 15b, h. H. BERNS, Ruhr-University.

References: p. 1618.

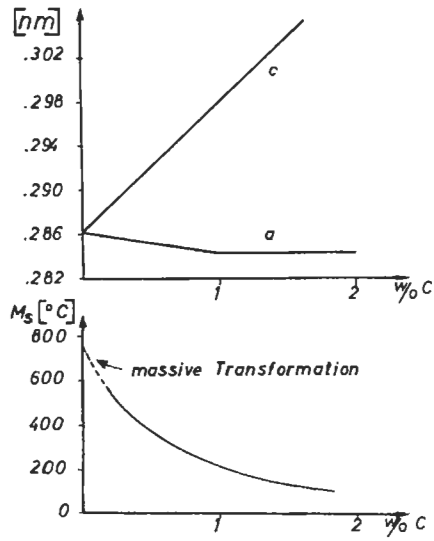


Fig. 16. c/a ratio and M_s for body-centered tetragonal Fe–C martensite.

All elements save Co lower M_s. All the others in the equation produce solid solution strengthening of austenite and lower M_s, regardless of whether they raise or lower T_0. Also, austenite can be stablilized against transformation by work hardening, by grain refinement, by deformation twins, and by prior precipitation of coherent second-phase particles, in brief, by anything that increases strength. Both nucleation and growth of martensite are inhibited by increased strength (KAJIWARA [1986], LESLIE and MILLER [1964], MEYER and HORNBOGEN [1968]).

The very large effect on M_s of the prior precipitation of coherent particles in the austenite of an Fe–Ni–Al alloy is shown in fig. 17. The resistance to displacive transformation is maximized with maximum precipitation hardening. With over-aging the M_s rises as Al and Ni are removed from solution. Figure 18 illustrates the microstructure after partial and complete transformation to martensite.

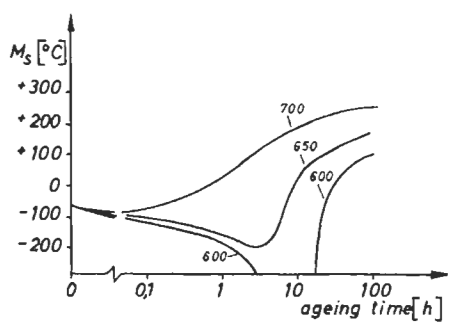


Fig. 17. Variation of the M_s-temperature of an Fe, 27.4 at% Ni, 12.5 at% Al, 0.06 at% C alloy by aging the austenite at different temperatures, to form coherent precipitates.

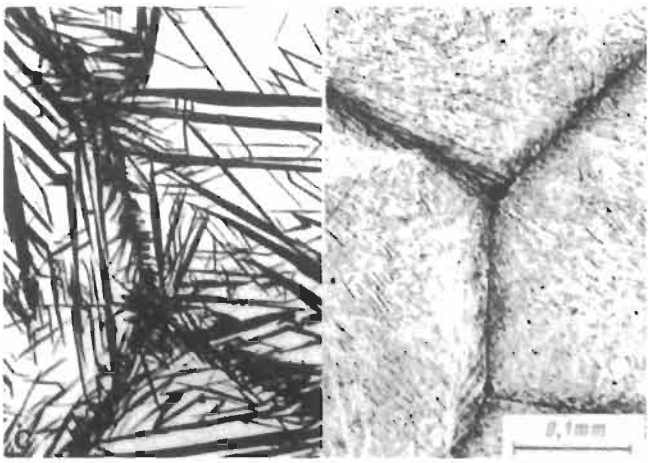


Fig. 18. Incomplete (40%) and complete transformation of austenite to martensite of an 0.06 at% C, 27.4 at% Ni, 12.5 at% Al steel ($M_s = -69°C$). The transformation starts at austenite grain boundaries. Austenite grain size determines the maximum dimension of martensite crystals. Former austenite grains can be revealed even in the completely transformed alloy. Carbide precipitates are visible at the grain-boundary nodes.

In order to obtain complete transformation to martensite a steel has to be cooled considerably below M_s (fig. 19). Retained austenite is found in the spaces between the martensite. This austenite becomes work-hardened by the martensite crystals as they form, and is thus stabilized against transformation. If cooling is interrupted at a temperature below M_s, then resumed after a period of time, the transformation does not continue until an additional undercooling is reached (*stabilization* of austenite). This can be due to strain-aging of deformed austenite and the thermally activated reduction of transformation stresses that would have aided further transformation (e.g., GLOVER [1956]).

The size of martensite crystals is limited by the austenite grain size. In steels, small martensite crystals are wanted because isotropic properties and small internal stresses are wanted. Heat-treatments that lead to grain growth of austenite before cooling below M_s

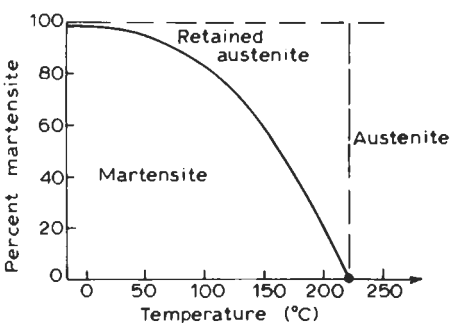


Fig. 19. Martensite that has formed below M_s in a steel with 1.1 at% C and 2.8 wt% Cr after austenitizing at 1040°C. $M_s = 230°C$, the transformation is almost complete at 20°C. (After COHEN [1949].)

are therefore avoided.

It has become conventional to classify the morphology of martensite as being either the lath type or the plate type. The former is typical of low to medium carbon steels. When the carbon content is below about 0.20% the structure is often body-centered cubic. The lattice invariant shear is mainly accomplished by dislocations. The units of martensite are in the shape of laths containing a high density of dislocations. The laths are grouped into larger packets. Plate-type martensite is found in alloys with M_s temperatures below ambient. This structure is composed of individual lenticular plates with internal very fine twins. In many instances the structure consists of mixtures of lath and plate martensite. The conventional distinction between the two types has been questioned by TAN *et al.* [1992], who found twins in lath martensite in 0.20% C steel by careful transmission electron microscopy and demonstrated that it is easy to miss these.

3.3. Bainite

Although bainite was detected in a study of the isothermal transformation of austenite in 1929 (DAVENPORT and BAIN [1930]) the modes of formation of the various admixtures of ferrite and carbides comprising bainite are still matters of controversy; see the proceedings of the International Conference on Bainite, 1990. The controversy is hardly surprising, for bainite forms in a wide variety of steels, over a wide range of temperatures during the transformation of austenite isothermally or during continuous cooling and the transformation products are finely divided and only resolvable by transmission electron microscopy. The disagreements center on the relative extent of transformation by shear or by diffusional mechanisms. The answer may lie in the observation by JUNG *et al.* [1994] that bainite can nucleate by diffusion and grow by a displacive process. Regardless of the outcome, the importance of bainite lies in the desirable combination of mechanical properties it imparts to commercial steels transformed isothermally or by continuous cooling.

Conventionally, the morphology of bainite is divided into two types, upper bainite and lower bainite. These terms arose from structural differences noted after isothermal transformation and may have little relevance to structures developed during continuous cooling. Bainite forms isothermally at temperatures above the M_s and below that required for formation of ferrite and pearlite, or at cooling rates too rapid for nucleation of these other transformation products but not rapid enough to form martensite.

The structure of upper bainite is generally agreed to consist of aggregates of ferrite platelets called *sheaves* with Fe_3C particles between the sheaves. In lower bainite the carbides, probably ε, $Fe_{2.4}C$, are contained within the ferrite platelets. In hypoeutectoid steels the ferrite platelets form first. The controversy arises over whether the iron atoms cross the ferrite-austenite interface by shear or by diffusive jumps. The structure of lower bainite is shown in figs. 20a and b.

There is a fourth type of reaction in iron alloys with M_s temperatures above 400°C. Even at extremely high cooling rates transformation occurs above M_s by a mechanism in which the rapid motion of the γ–α interface produces ferrite with the same composition as the austenite. The morphology of the transformation product differs from that of

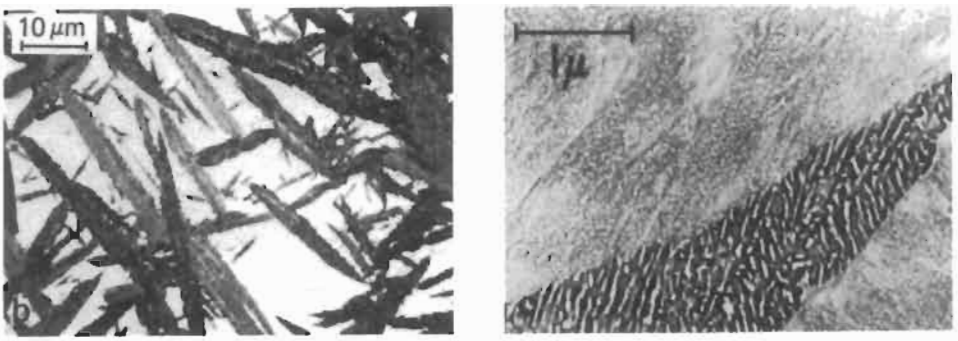


Fig. 20. (a) and (b) Microstructure of lower bainite. 3.23 wt% Cr; 30 min at 350°C. Partially transformed 0.66 wt% C (SPEICH [1962]). (a) Optical micrograph. (b) Electron micrograph (replica).

martensite because thermally activated motion of the interface occurs by individual jumps across the reaction front in a direction opposite to that of its motion. The transformed structure is almost free of the defects present in martensite. This reaction is termed a *massive transformation* (GILBERT and OWEN [1962]) (see also chap. 15).

3.4. Transformation diagrams and hardenability

The evolution of microstructures in steels can be understood best by reference to *isothermal* (IT) or to *continuous cooling transformation diagrams*. These diagrams indicate the beginning and end of each reaction in the transformation of austenite at constant sub-critical temperatures or the cooling rates at which various reactions begin and end. In the making of an IT diagram, small thin specimens of a selected steel are heated into the austenite range, quickly transferred to a liquid medium held at a constant subcritical temperature, held for a pre-determined period and then quenched into iced brine and examined metallographically to determine the extent of phase transformation. The process is illustrated schematically in fig. 21. The result of many such determinations is a diagram such as fig. 22 for a typical high-strength alloy steel, AISI 4340. Strictly speaking, each diagram is representative only of reactions in one sample of steel having one austenite grain size. Even within one grade of steel variations in chemical composition and austenite grain size will result in small changes in the times for beginning and ending of transformation. Nevertheless, such diagrams are indispensable for the planning of certain heat treatments to obtain desired microstructures and properties, as shown in figs. 23 a, b, c, and d.

The IT diagram of fig. 22 shows two "C-curves", the upper one for the formation of ferrite and pearlite and the lower for formation of bainite, with a deep "bay" between them (bainite is not labelled in the figure). Bainite in this steel forms at constant temperatures from just above 500°C down to below M_s at about 280°C. At temperatures between 500° and about 425°C the reaction does not go to completion; much austenite remains. The upper and lower critical temperatures and M_s are shown.

In production and heat treating of steels it is often desirable to have some knowledge

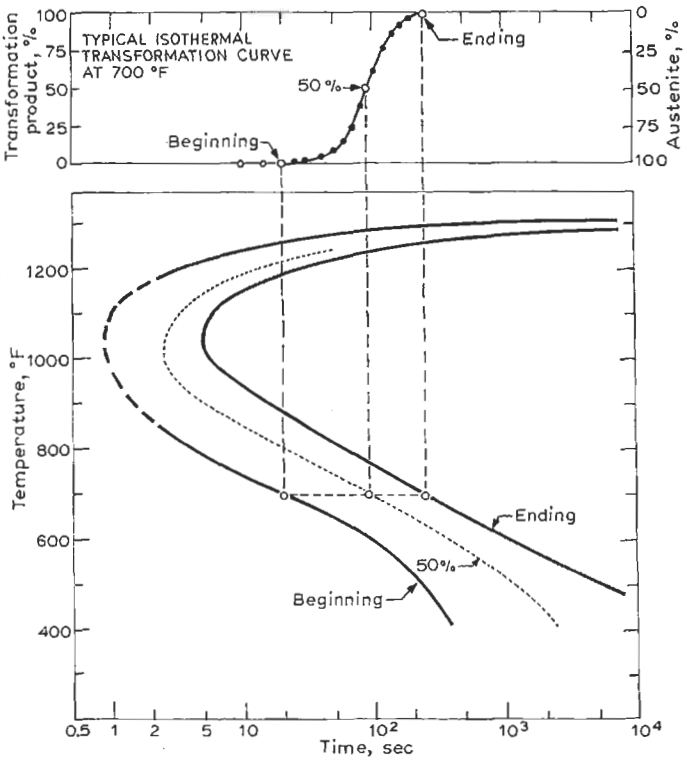


Fig. 21. Diagram showing how measurements of isothermal transformation are summarized by the isothermal transformation diagram.

of microstructures following *continuous* cooling from austenite and this information is not readily obtained from IT diagrams. For these estimations we turn to CCT diagrams. These are developed through a combination of dilatometry and metallography. In dilatometry, temperature, time and specimen length are recorded for a variety of cooling rates. The initiation and termination of transformation are marked by an increase in specimen length and the cessation of growth, respectively. Microstructures are determined after quenching. From these data a CCT diagram can be constructed, fig. 24 (ELDIS [1978]), with the linear temperature-log time scale as used in IT diagrams. Alternatively, instead of plotting cooling curves, individual bar diameters are plotted on the abscissa instead of transformation times. The diagrams indicate the structures that can be obtained at the centers of the bars of the appropriate size, fig. 25 (ATKINS [1977]). Thus a bar 20 mm in diameter, air cooled, will have a structure of approximately 70% bainite, 30% martensite. A bar of 50 mm diameter, water cooled, will have a desirable structure of about 100% martensite at the center.

Hardenability is commonly defined as the depth to which steel of a given composition can be quenched to yield a structure of 50% martensite — 50% other transfor-

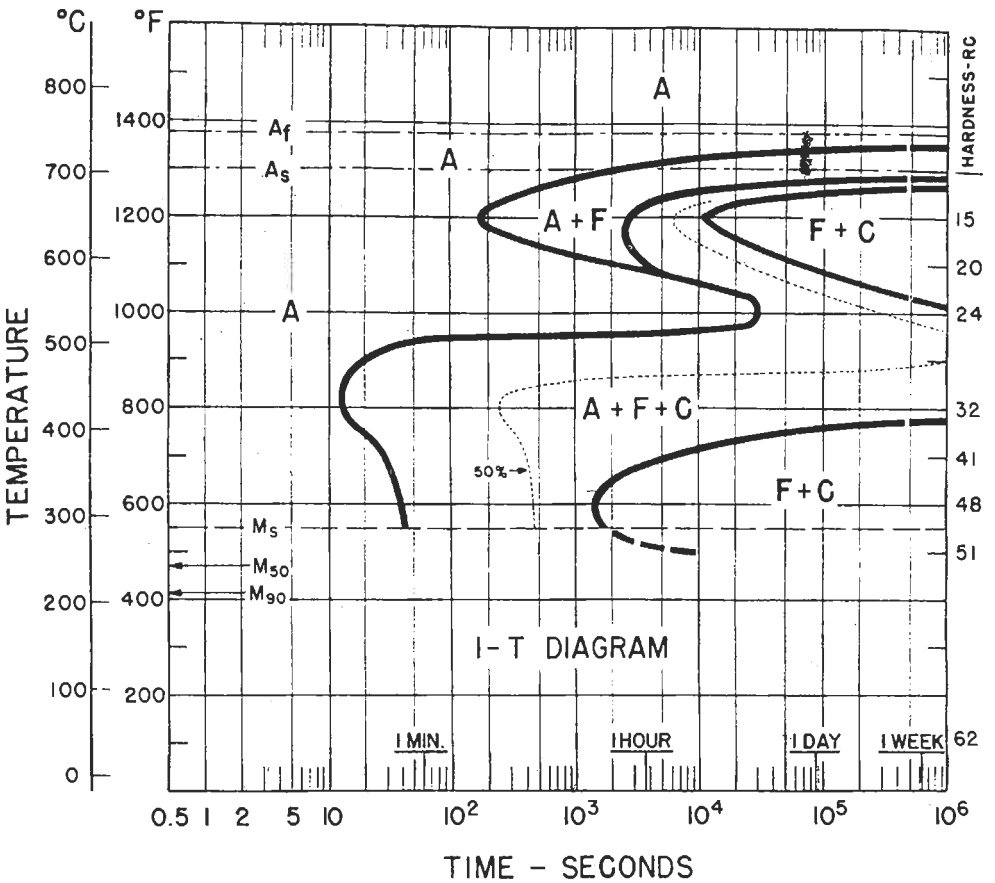


Fig. 22. IT diagram for AISI 4340 steel. 0.42 C, 0.78 Mn, 1.79 Ni, 0.80 Cr, 0.33 Mo. Austenitized at 845°C. Grain size ASTM 7–8. A = austenite, C = cementite, F = ferrite.

mation products. A criterion of 90% martensite also may be used. Hardenability is affected by austenite grain size, carbon content and alloy content. It is measured by the Jominy end-quench test (JOMINY and BOEGEHOLD [1938]) or by a more recent test better suited for steels of lower hardenability (GRANGE [1973]). There are various methods of predicting hardenability of steels from austenite grain size, carbon content and alloy content, reviewed by DOANE [1978]. The goal of steel producers is to attain the maximum hardenability with the minimum use of expensive alloying elements. Steels can be ordered within specified limits of hardenability (hardenability bands) and the producer can use whatever composition will yield that result.

3.5. Tempering of martensite

Ferrous martensites are highly supersaturated solid solutions of C in iron, obtained by

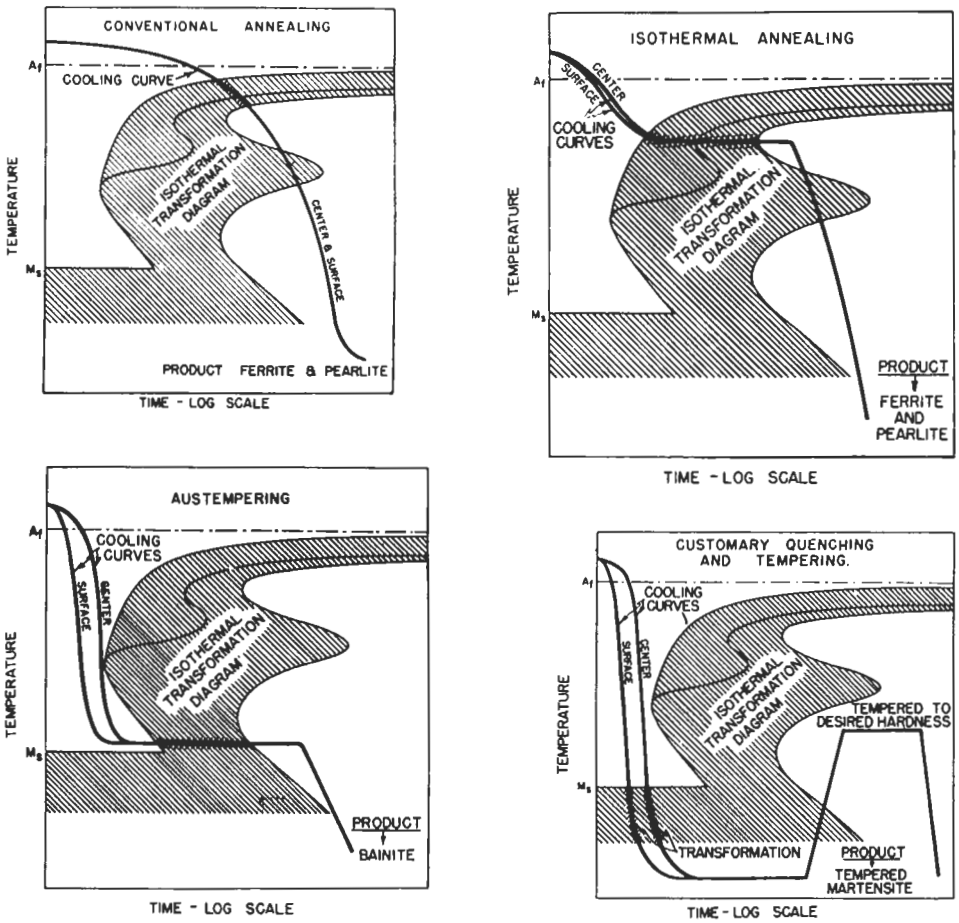


Fig. 23. Four illustrations of the use of IT diagrams in planning practical heat-treatments.

rapid cooling from temperatures at which the alloy is single-phase fcc austenite. The transformation to bcc or bct iron produces a very high density of defects — dislocations, twins, former austenite grain boundaries, vacant lattice sites and any substitutional solute atoms initially present. Generally speaking, such structures impart high strength but poor toughness. To improve toughness with some reduction in strength the steel is reheated to temperatures above ambient. The result is the removal of C from solid solution and the precipitation of carbides. The operation is simple in practice but the processes occurring are highly complicated. Despite decades of study there is still no complete agreement on all the details. These vary depending upon the carbon content and alloy content, which determine the M_s temperature. If this is above ambient the tempering process (auto-tempering) can begin during quenching. In steels with very low M_s temperatures, subsequently held at temperatures below 100°C, processes termed "aging" can occur.

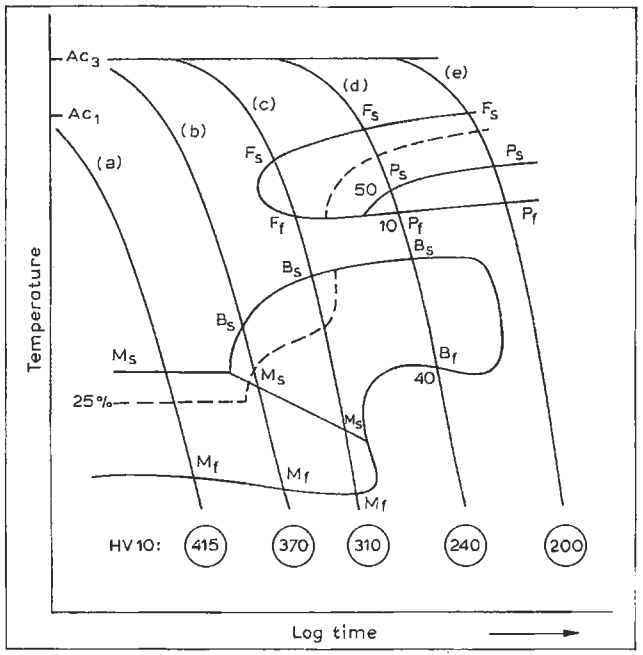


Fig. 24. Schematic illustration of the construction of a CCT diagram. F-ferrite, P-pearlite, B-bainite, M-martensite, HV10-hardness.

UWAKWEH *et al.* [1991] state that there are two steps of aging, first clustering of C atoms, then the ordering of these clusters into a carbide of composition Fe_6C. In contrast, SPEICH and TAYLOR [1992], consider reactions during aging to be one process of spinodal decomposition.

There is a continuing tendency towards lower carbon contents in steels, for elimination of quench-cracking, improved weldability and greater toughness. There is general agreement on the processes occurring when these lower-carbon steels with high M_s temperatures are quenched and tempered:

1. Segregation of C to dislocations, former austenite boundaries and retained austenite
2. Precipitation of ε-carbide, $Fe_{2.4}C$, in the temperature range 50–200°C
3. Between about 200 to 300°C, transformation of retained austenite to ferrite and carbide
4. Solution of ε-carbide and precipitation of χ carbide and/or Fe_3C.
5. Between 400 and 700°C, recovery of the dislocation structure, spheroidization of cementite and recrystallization of the ferrite.
6. If carbide-forming elements are present in substantial amounts alloy carbides such as VC, Cr_7C_3 or Mo_2C can replace Fe_3C during tempering at above about 500°C.

The hardness of a steel after tempering at any given temperature for the customary time

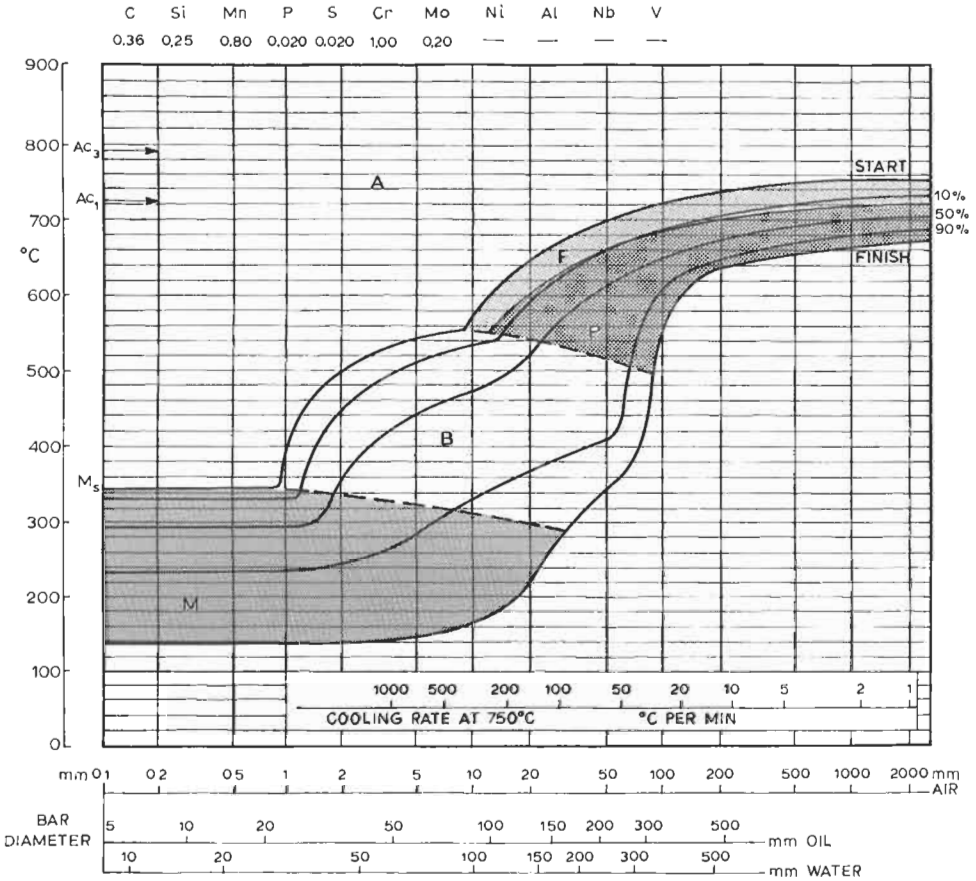


Fig. 25. CCT diagram of a Si–Mn–Cr–Mo steel. A = austenite, B = bainite, F = ferrite, M = martensite, P = pearlite.

of one hour can be calculated from a knowledge of the chemical composition (GRANGE *et al.* [1977]). Alloying elements increase resistance to tempering, the most effective being V, P, Mo and Mn.

When high-strength medium-C steels are tempered between 250 and 400°C an annoying phenomenon called *tempered martensite embrittlement* appears. Although the hardness after tempering decreases continuously throughout this range the toughness, as measured by Charpy V-notch tests, reaches a minimum, fig. 26 (BRIANT and BANERJI [1978]). During austenitizing, S and P atoms segregate to austenite grain boundaries (see chapter 13). The segregration remains during the transformation to martensite and during tempering, reducing cohesion at the prior austenite boundaries. Cementite precipitates at these same boundaries and serves as nucleation sites for cracks during impact loading. These cracks propagate along the weakened boundaries, leading to intergranular fracture.

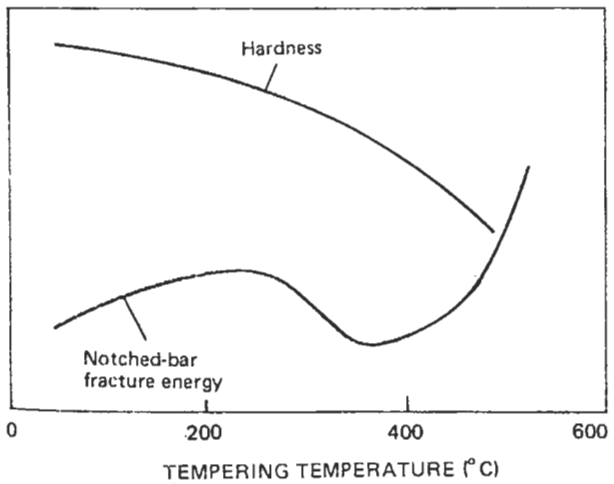


Fig. 26. Schematic illustration of tempered martensite embrittlement of a quenched and tempered steel.

Tempered martensite embrittlement can be prevented in steels of very high hardenability and very high strength by the addition of 1 to 2% Si. Silicon inhibits the growth of ε carbide and cementite (OWEN [1954], KEH and LESLIE [1963]), reducing their effectiveness as crack starters and allowing such steels to be tempered in the range 260-315°C, where both yield strength and notch toughness reach a maximum.

4. *Deformation and recrystallization*

4.1. Microstructure of deformed steel

Iron and ferritic steels, in common with other bcc metals, are characterized by a pronounced dependence of the yield and flow stresses upon temperature and strain rate (CHRISTIAN [1983]). As shown in fig. 27 the yield stress of polycrystalline iron increases sharply at temperatures below about 325 K, ~0.18 T_m, where T_m is the melting point in kelvins. Above that temperature a long plateau appears and the yield stress does not begin to drop until a temperature of about 800 K, ~0.45 T_m, is reached. Although it is not clearly shown in fig. 27, a concave-downward region appears in the yield stress-temperature plot between about 180 and 280 K (MATSUI *et al.* [1978]). Similarly, the yield stress increases markedly increasing with strain rate, fig. 28 (LESLIE *et al.* [1969]). These are intrinsic characteristics of the bcc metal lattice and are intimately connected with the decreasing toughness of ferritic steels with decreasing temperature and/or increasing strain rate. The cause of these phenomena is the slight dissociation of screw dislocations in bcc metals which renders them relatively immobile (sessile) at low temperatures. The dissociation must be constricted by thermal activation before it can glide. In contrast, edge dislocations can glide at any temperature.

In bcc metals slip can occur on any plane containing a <111> slip direction, so-called

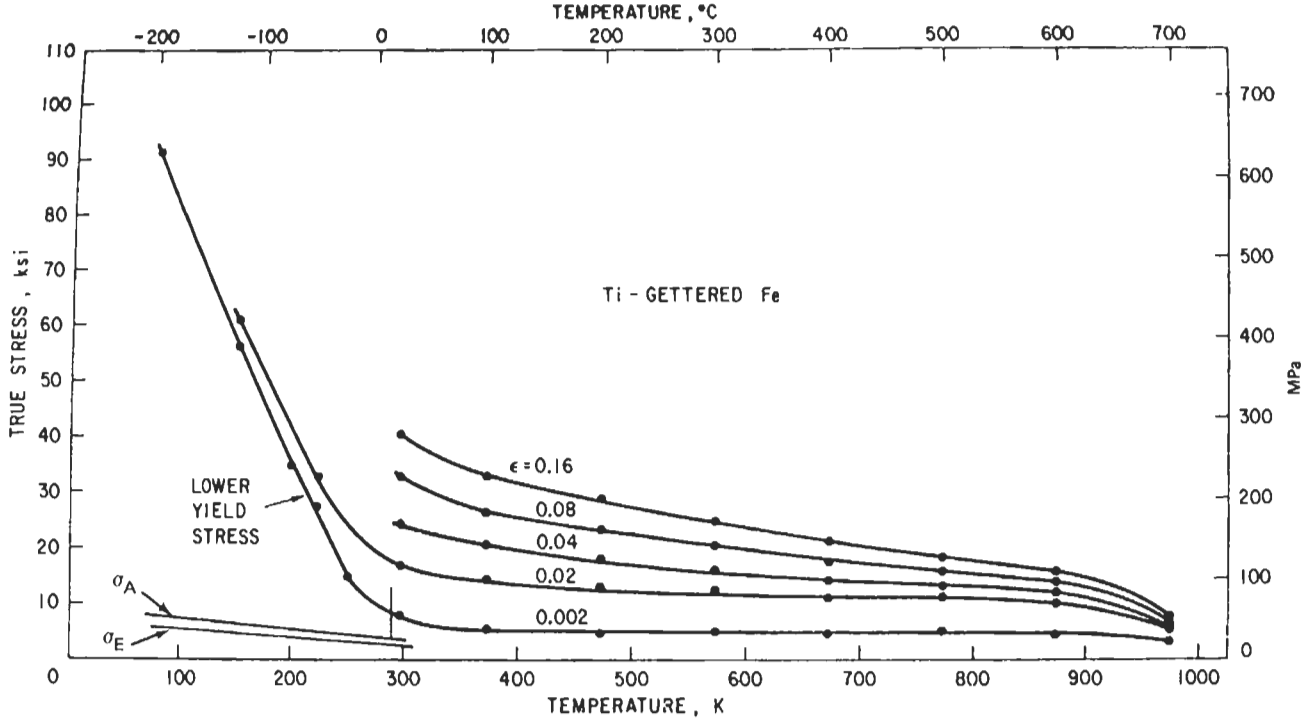


Fig. 27. Temperature dependence of yield and flow stresses in Fe–0.15%Ti. $\dot{\varepsilon} \cong 2.5 \times 10^{-4}$ s^{-1}, G.S. ASTM 5–6. σ_E is the elastic limit, σ_A the anelastic limit.

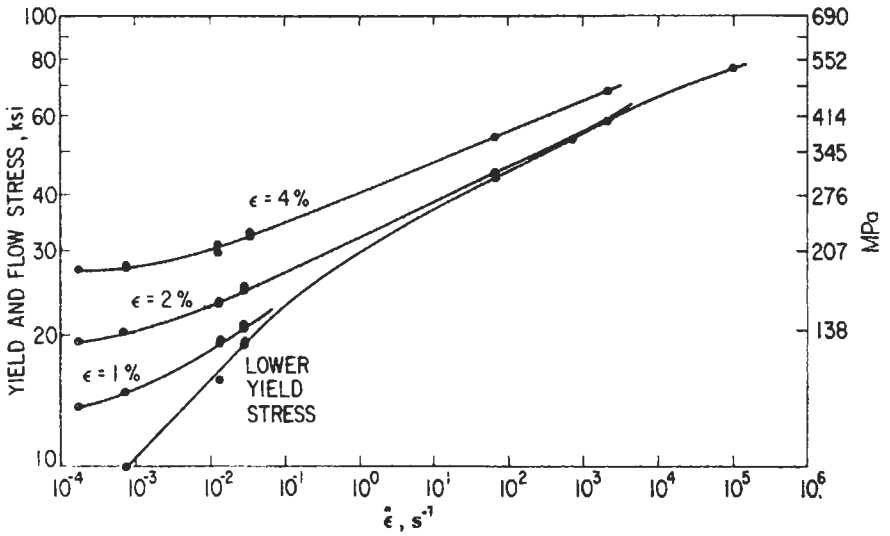


Fig. 28. The effect of strain rate on the strength of polycrystalline bcc iron at room temperature.

pencil glide. The result is wavy slip lines on a polished surface of ferrite deformed at ambient temperature (fig. 29a). Deformation at temperatures well below ambient leaves a uniform distribution of the sessile screw dislocations, the mobile edge components having moved out of the crystal. During deformation of ferrite at room temperature, dense tangles of dislocations form a cell structure. The spacing between cell walls decreases with increasing strain (KEH [1961]).

In fcc metals (austenitic steels) slip occurs by glide of $a/2<110>$ dislocations on {111} planes, leading to a total of 12 possible slip systems. The dislocations are confined to their slip planes because of low stacking-fault energy, the result being well-defined, straight slip lines (fig. 29b). In the presence of a fine dispersion of shearable coherent particles slip steps are well defined and widely spaced (fig. 29c).

Dislocations can be generated at grain boundaries and at the interfaces of inclusions and carbide particles. The density of dislocations increases with deformation (fig. 29d), KEH and WEISSMANN [1962]. The curve shown there does not apply when a dispersion of any strong second phases is present; in those instances the density of dislocations for a given strain, i.e., the rate of work hardening, is greater than indicated by fig. 29d.

Cementite, whether present in pearlite or as discrete particles, resists deformation, with the important exception of wiredrawing. Beginning with a eutectoid carbon content in the form of fine pearlite and upper bainite, a true strain of 4 reduces the thickness of Fe_3C lamellae to about 2 nm, or just a few unit cells.

When stressed, annealed low-carbon steels usually yield discontinuously, i.e., with an abrupt drop in stress, followed by yielding at constant stress and then continuous work-

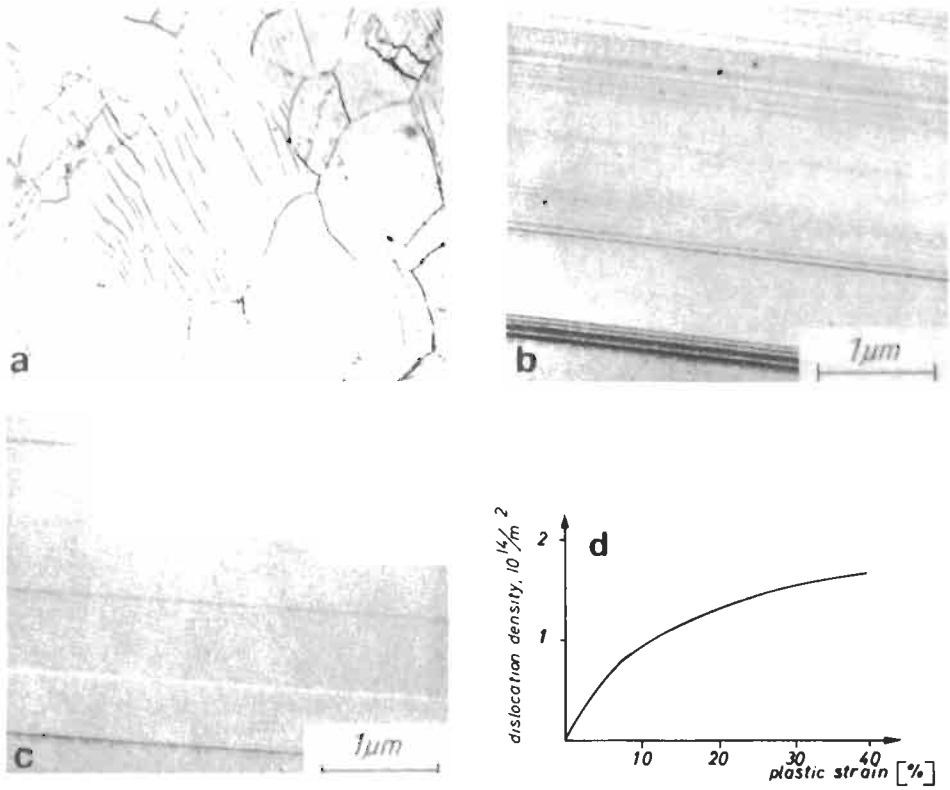


Fig. 29. (a) Wavy slip bands in ferrite after 3% strain at 20°C, 200 ×. (b) Traces of {111} slip planes in austenitic steel after 2% strain, replica. (c) High slip steps in austenitic steels containing γ' precipitates after 2% strain. (d) Relation between plastic strain and dislocation density in polycrystalline α-Fe.

hardening. The strain at constant stress is called the yield point elongation or *Lüders strain*. For this to occur, three requirements must be met:

1. A low density of mobile dislocations present before yielding. (Most or all of the dislocations present are normally pinned by interstitial "atmospheres", see ch. 20).
2. Rapid generation and multiplication of dislocations.
3. A low to moderate dependence of dislocation velocity on stress.

Under these conditions, when dislocations are generated heterogeneously at some spot of stress concentration they multiply rapidly by a multiple cross-slip mechanism. Edge dislocations in bcc iron move about 20 times faster than screws at a given stress, but the motion of screws is required for macroscopic yielding. The flow initiated at the upper yield stress coincides with the motion, cross slip and multiplication of screw dislocations. From the point of generation of dislocations a zone of high dislocation density (localized strain) sweeps through the specimen. This is called a *Lüders band*. The cross-section of

the specimen is reduced, raising the stress and inducing other such bands until uniform work hardening begins.

Deformation of ferritic steels also can proceed by twinning. This is favored by low temperatures, high strain rates, large grain size, lack of mobile dislocations and solid solution hardening, especially by Si. As compared to deformation by dislocation generation and motion it is of minor importance. Mechanical twins can be produced in austenitic steels but only at very high strain rates and/or low temperatures (HORNBOGEN [1964b]).

4.2. Recovery and recrystallization

Recrystallization can be defined as crystallographic reorientation by passage of a high-angle boundary. Recovery is the name given to all other processes that reduce the energy stored by deformation (see ch. 26). Both are thermally activated and both can be driven by stored energy of deformation but recrystallization can also be driven by the reduction of surface energy, whether the surfaces are external or internal, i.e., grain boundaries. Both can be static or dynamic. We will consider static processes first, those that occur at constant temperature after cold work.

There can be several measures of recovery, such as decreases in electrical resistivity, decreases in X-ray line broadening, decreases in hardness, decreases in dislocation density and rearrangement of dislocations. These all proceed rapidly initially, then reach a plateau. Unless recrystallization intervenes they will go to completion in a range of temperature from below room temperature to about 500°C, fig. 30 (MICHALAK and PAXTON [1961]). The motion of lattice defects can be hindered by factors such as solute atoms — especially interstitials — pre-precipitation clusters and nucleation of precipitates on such defects, thereby retarding or stopping recovery.

Recrystallization at constant temperatures after cold working can begin by migration of existing high-angle grain boundaries, migration of sub-boundaries and by sub-grain

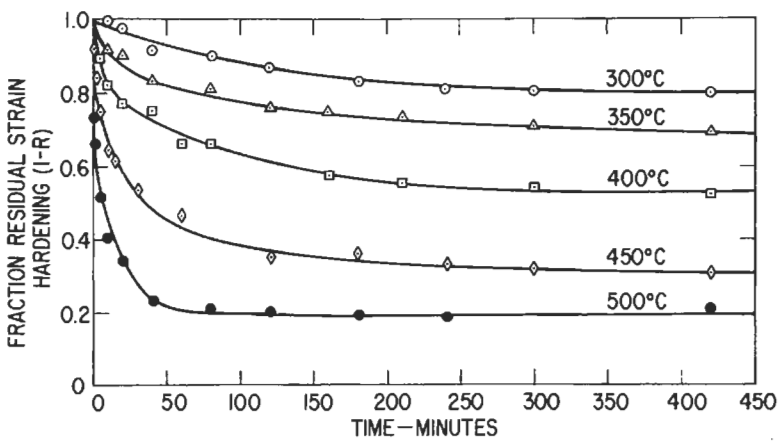


Fig. 30. Recovery of α-Fe after 5% strain at 0°C (MICHALAK and PAXTON [1961]).

coalescence (HU [1962]). These processes will begin where energy density is greatest, as at grain-boundary edges, inclusions and intersections of mechanical twins. As stated previously, recovery and recrystallization are driven by the stored energy of cold work so they can be concomitant, competing processes. After slight cold working recovery can be complete before recrystallization begins, and the rate of growth of recrystallized grains will be constant until the growing grains impinge. After heavy cold work the reduction of stored energy by recovery causes the rate of growth of new, recrystallized grains to slow markedly with increasing time at constant temperature (fig. 31). This iron was highly purified by zone-refining, which greatly enhances the rate of recovery.

The rates of formation and growth of recrystallized grains can be changed by several orders of magnitude by even small additions of second elements in solid solution (VENTURELLO *et al.* [1963], LESLIE *et al.* [1963], fig. 32). However, the inhibition by fine precipitates or pre-precipitation clusters is even greater. If ferritic steels are deformed in the austenite range, dynamic recrystallization can occur immediately after working, before or during cooling. Control of the process by temperature, degree of reduction and the presence of fine precipitates is important in the production of high-strength low-alloy (HSLA) steels. The fine precipitates of AlN, Nb(C,N), V(C,N) or Ti(C,N) are used to refine austenite grain size and hence ferrite grain size; AlN precipitated during annealing after cold work provides a desirable texture in sheets for cold forming. Copper precipitated *before* cold rolling and annealing can produce a similar effect (ERA *et al.* [1988]). MnS, MnSe, AlN or BN can be employed in grain-oriented silicon steels to provide a very pronounced cube-on-edge, {110}<001> texture.

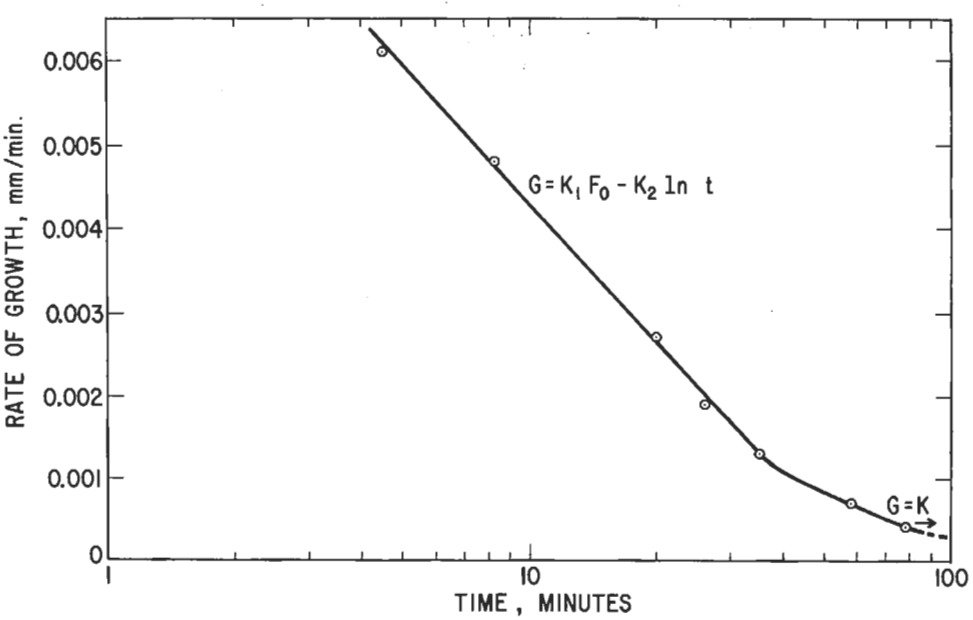


Fig. 31. Time dependence of rate of grain growth in zone-refined iron, cold rolled 60%, annealed at 480°C.

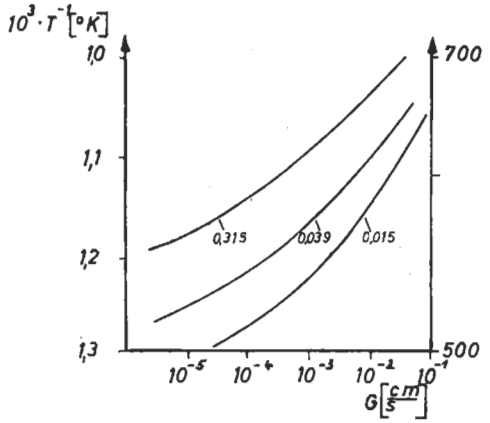


Fig. 32. The effect of small Mo additions on the velocity of the recrystallization front of iron at 5% recrystallization, 60% deformation, at% Mo as indicated (After LESLIE *et al.* [1963]).

Second-phase particles have two different effects on the kinetics of recrystallization. Particles larger than about 1 μm, such as inclusions or pearlite nodules promote nucleation of recrystallization. If particles are fine, <0.1 μm and closely spaced, dislocations after cold working are uniformly distributed, without regions of high lattice curvature required for nucleation and recrystallization is inhibited (fig. 33). The effect of second-phase particles of different sizes is further treated in ch. 26.

5. *Mechanical properties*

5.1. Strength of ferrite

The ideal strength of iron crystals can be measured by straining iron whiskers. Depending on crystallographic orientation, the tensile strength lies between 5 and 13 GPa and the shear strength between 3.5 and 8 GPa (BRENNER [1956]). In contrast, the critical resolved shear stress of high-purity iron single crystals at room temperature is about 10 MPa and of polycrystals about 30 to 40 MPa. The approximate ranges of strength of iron-base materials is shown in fig. 34.

Single-phase ferrite can be strengthened by

1. grain refinement;
2. cold work;
3. solid solution additions;
4. shear-type transformation;
5. radiation.

The relation between the lower yield stress and the grain size in ferrite is customarily expressed by the Hall–Petch equation.

$$\sigma_{LYS} = \sigma_o + kd^{-\frac{1}{2}}$$

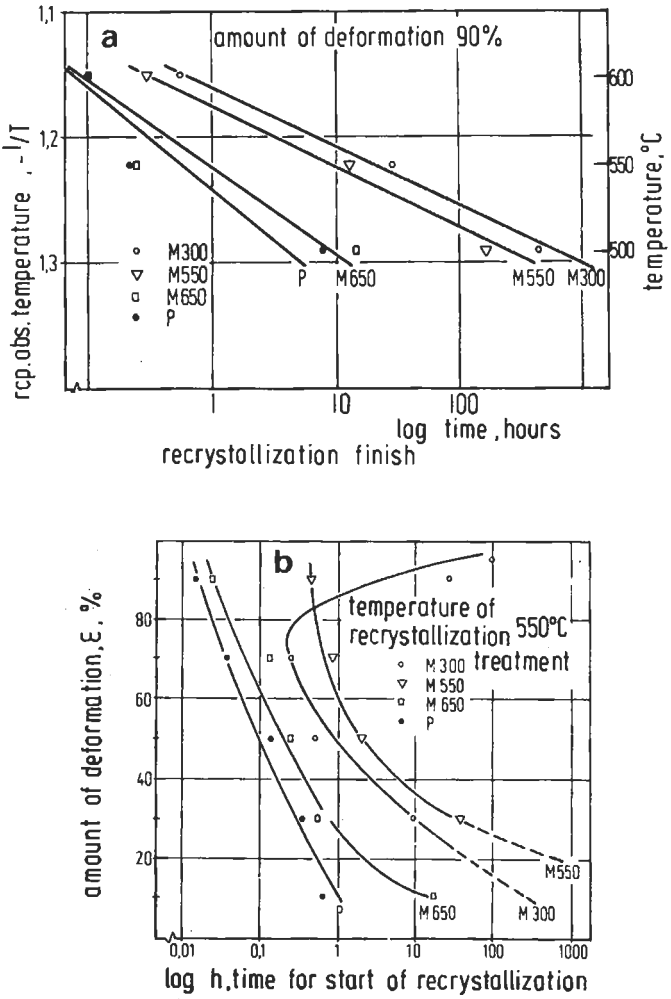


Fig. 33. Recrystallization of a steel with 0.22 wt% C and different distributions of Fe_3C particles: P: Ferrite–Pearlite; M: Martensite tempered at 300, 550, 650°C. (a) T–t-diagram for complete recrystallization. (b) T–ε-diagram for start of recrystallization.

where d is the average intercept grain size. The value of the exponent is now generally accepted as $\frac{1}{2}$. Although empirical, this equation is highly useful; fig. 35 is a plot covering a wide range of grain sizes.

After small plastic strains the dislocation density in bcc iron is related to the flow stress, as shown in fig. 36 (KEH and WEISMANN [1963]). The relation can be expressed by

$$\sigma = \sigma_o + \alpha G b \sqrt{N}$$

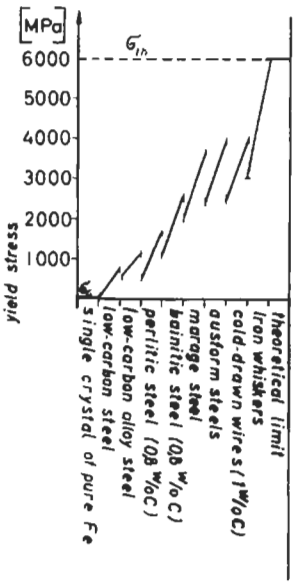


Fig. 34. Strength ranges of iron-base materials.

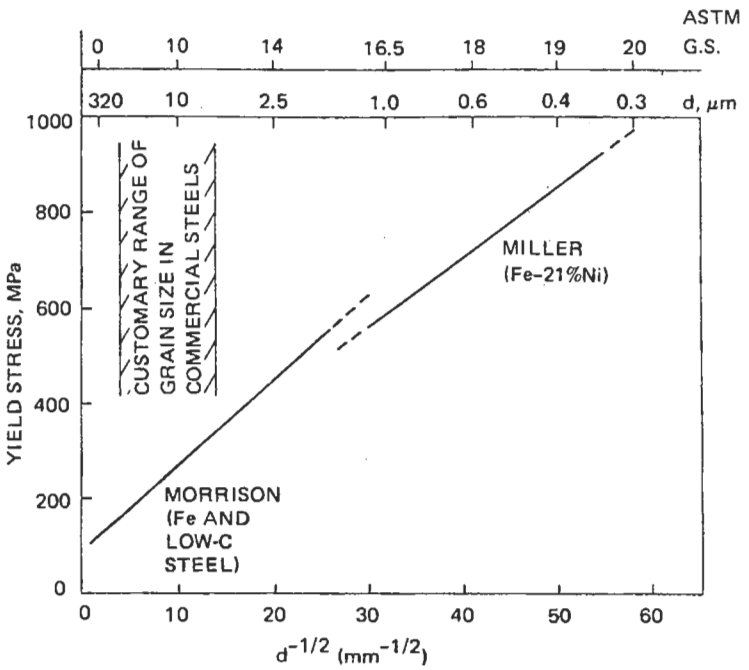


Fig. 35. The effect of grain size on the lower yield stress of alpha iron.

References: p. 1618.

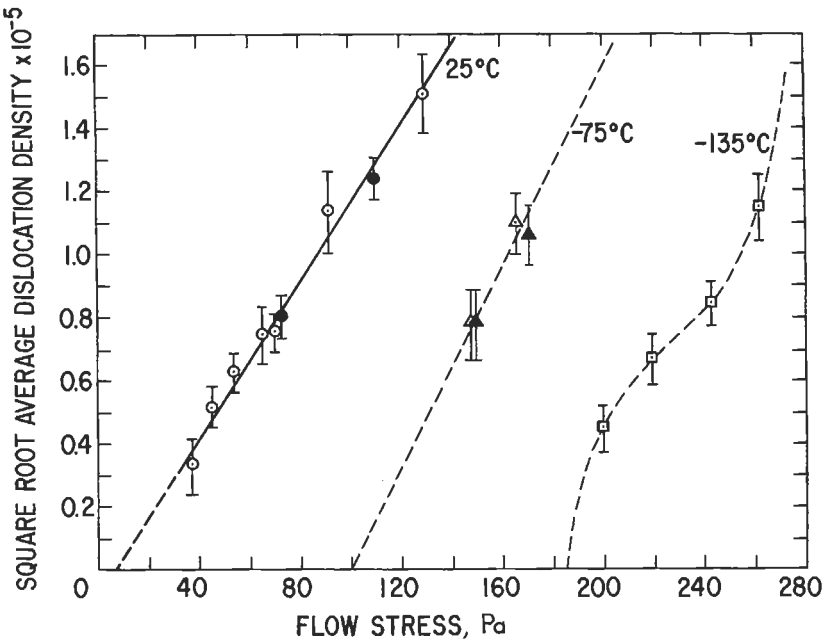


Fig. 36. Relationship between the square root of average dislocation density and flow stress of α-iron (KEH and WEISSMANN [1963]).

where α is a proportionality constant, G is the shear modulus, b is the Burgers vector, and N is the dislocation density.

Dislocation density at a given strain is greater in fine-grained than in coarse-grained iron.

For uniform plastic strain the true stress-true strain relation is expressed by the empirical equation

$$\sigma = K\varepsilon^n$$

where σ is true flow stress, ε is true strain and K and n are constants (LUDWIK [1909]). K is the true stress at a strain of unity and n, the strain hardening exponent, is the slope of a log true stress-log true strain plot. The strain hardening exponent is numerically equal to the limit of uniform plastic strain and is an indication of the stretchability of low-carbon steel sheets.

The work-hardening of iron at large strains depends upon the mode of deformation. Titanium-gettered iron containing essentially no interstitial solutes can be cold drawn to a true strain of at least 10, to a strength of 1650 MPa (LANGFORD *et al.* [1972]). In torsion, work hardening saturates at a true strain of just under 4.8 because of dynamic recovery (YOUNG *et al.* [1974]). High strains are more difficult to obtain by cold rolling because of less severe geometrical restraints, but at a true strain of 2.8, work hardening was still increasing (LAKE [1977]).

True solid solution hardening of ferrite refers to the interaction of the strain fields of solute atoms with dislocations when the solute atoms are immobile. It is difficult to separate dynamic and static effects when interstitial solutes are involved, but tests done at 0°C and at a fairly high strain rate indicate that the change in shear stress with change in carbon content, $d\tau/dC$ is about $G/2$.

When substitutional solutes are added to bcc iron, the athermal behavior of the yield stress, shown in fig. 27, is eliminated. The yield stress is always a function of temperature. Even more interesting is the phenomenon of alloy softening. In restricted regimes of temperature and composition, solutes, whether substitutional or interstitial, can *reduce* the yield stress of iron, fig. 37. The magnitude of the softening, the yield stress of the alloy less the yield stress of the iron, is comparable to the maximum strengthening. Softening can be produced in unalloyed iron by knocking iron atoms into interstitial positions by high-energy radiation (SATOH *et al.* [1970]).

Outside the composition limits in which alloy softening occurs, solutes strengthen bcc iron. The atom size misfit is a rough guide to the magnitude of strengthening but it fails when the solute atoms are smaller than iron atoms, fig. 38.

Elements of limited solubility in bcc iron tend to segregate in grain boundaries, fig. 39, (HONDROS and SEAH [1977]) with significant effects on mechanical properties, as described later (see also ch. 13). Grain boundaries are sources of dislocations and the stress required to activate these sources can vary with solute type and concentration at the boundaries. As a consequence, k in the Hall–Petch equation, $\sigma_{LYS}=\sigma_o+kd^{-1/2}$, also

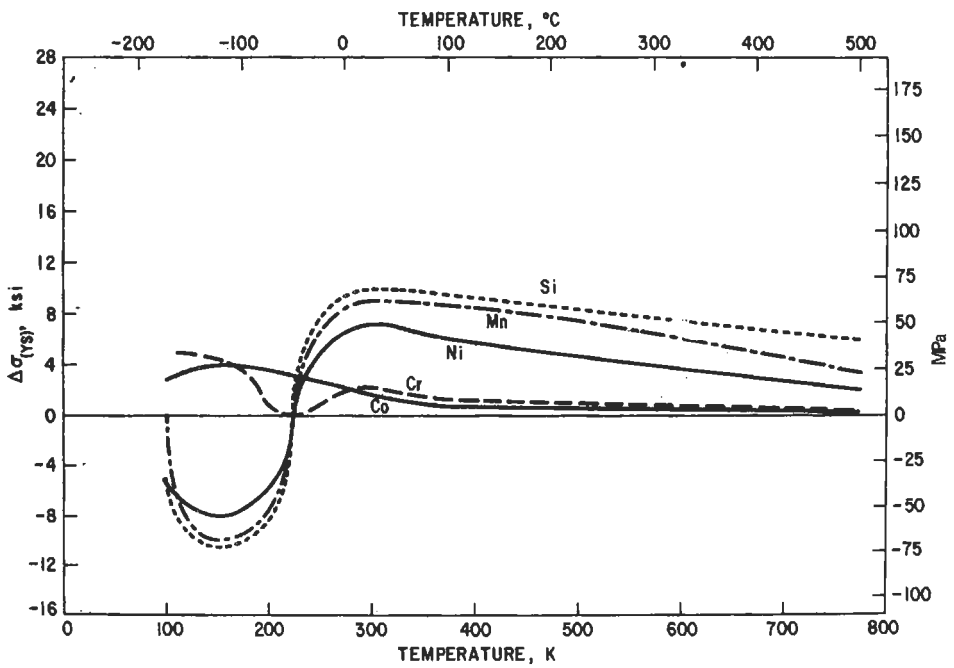


Fig. 37. Solid solution hardening and softening in iron-base alloys (1.5 at%), $\dot{\varepsilon} \cong 2.5 \times 10^{-4}$ s^{-1}.

References: p. 1618.

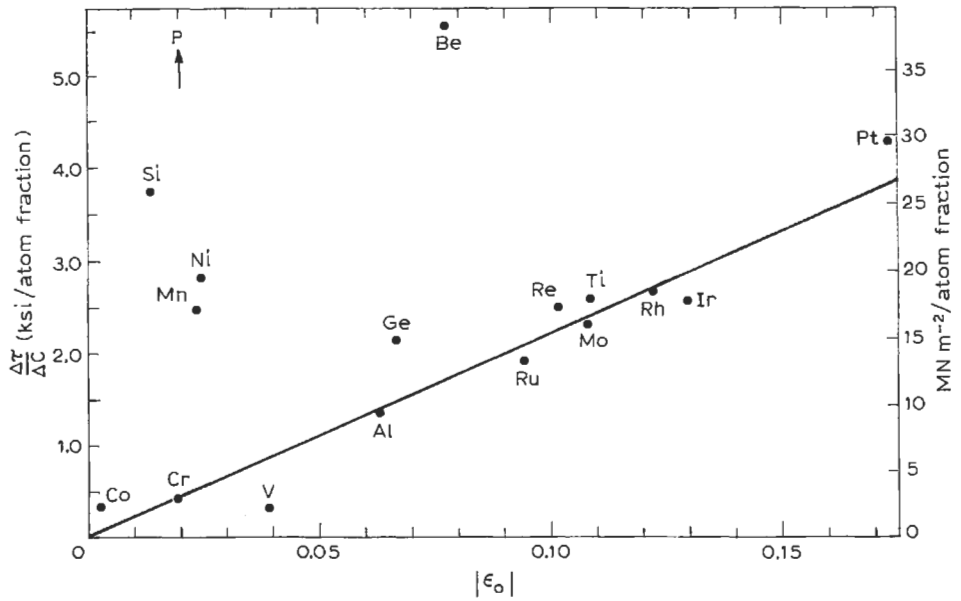


Fig. 38. Correlation of solid solution strengthening of iron-base binary alloys at 298 K with size misfit parameter, ε_0.

varies, C, N and P producing the largest increases. These segregation effects, equilibrium and non-equilibrium, further complicate the studies of solid solution hardening.

Although the process is not technologically important, high-purity bcc iron can be hardened by shear transformations induced at ambient temperatures by passage of a shock wave generating a pressure greater than about 13 GPa. The reaction involves the shear transformation of bcc iron to ε at high pressure then the reverse after the shock wave passes. The defect structure remaining resembles that of lath type martensite, without the presence of carbon. The yield strength is increased by about 540 MPa (LESLIE and SOBER [1967]). A martensite-like structure of high hardness can also be generated in pure iron by ultrarapid quenching from the melt. The hardness is increased more than sevenfold compared to metal quenched in the solid state (DUFLOS and CANTOR [1978]).

Iron can be hardened and embrittled by bombardment with high-energy neutrons or electrons which knock atoms from their normal lattice positions into interstitial sites (KLUEH [1990]).

5.2. Properties of structural steels

We are concerned here with steels that are not heat treated after manufacture. The newest of these are interstitial-free (IF) steels (TAKECHI [1994]). They are made possible by commercially practicable steel-making techniques that can reduce C, N, O and S to very low concentrations; C contents can be as low as 0.002%. Sufficient Ti or Ti+Nb

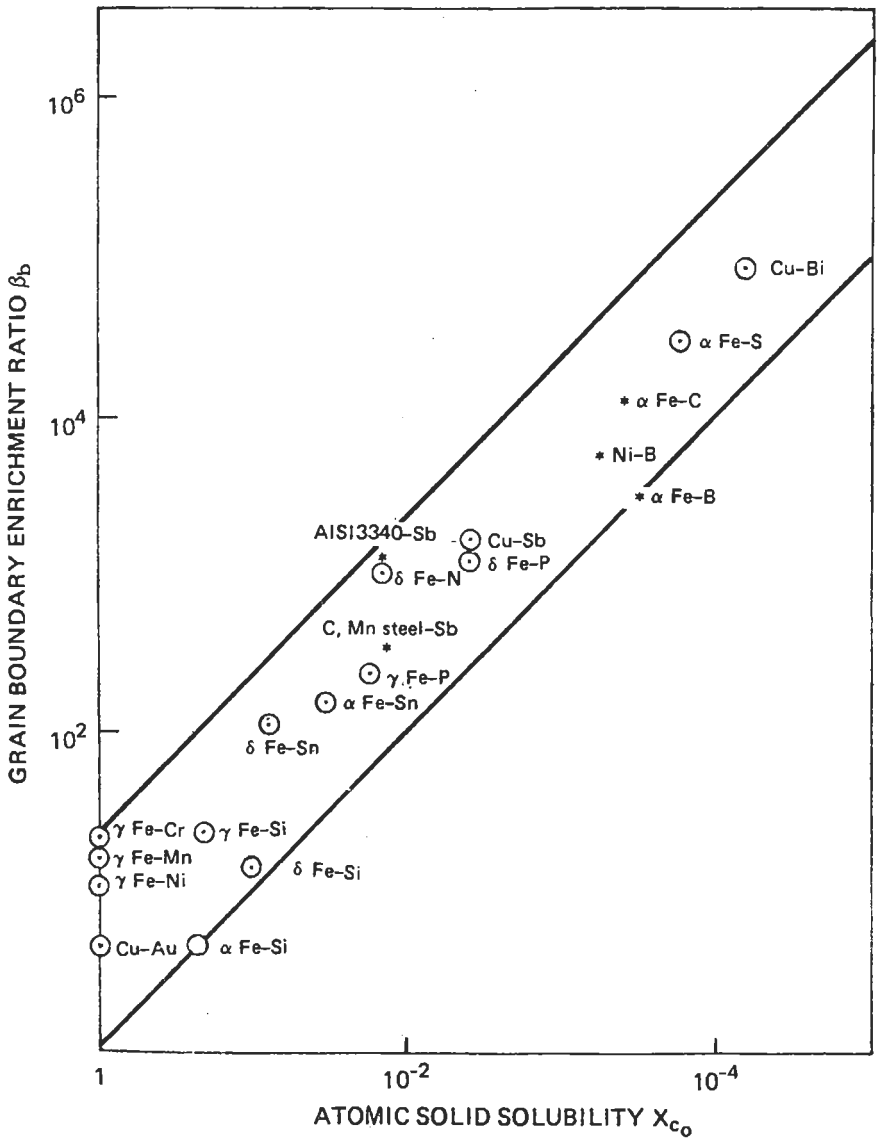


Fig. 39. Correlation of grain-boundary enrichment ratio β_b with the inverse of solid solubility for several systems.

is added to combine with C, N and S, reducing C in solution to less than 0.0001% after annealing. The steels are continuously cast, hot rolled, cold-rolled and continuously annealed.

The impetus for the rapid growth in production of these steels came from the

References: p. 1618.

realization that continuously cast Al-killed sheet steel did not have satisfactory deep drawing and non-aging properties, especially after hot-dip galvanizing which is now in great demand for automobile construction. For good forming properties, high n (in the stress/strain relationship $\sigma = K\varepsilon^n$) values and a well-developed {111}<110> texture are required. The extent of the latter governs the r value, which is defined as the ratio of width true strain to thickness true strain in tension tests of sheet specimens. With the assumption of constant volume:

$$r = \frac{\ln\left(W_i/W_f\right)}{\ln\left(L_f W_f / L_i W_i\right)}$$

where W_i = initial width
W_f = final width
L_i = initial length
L_f = final length

For deep-drawing steels r should be at least 1.5; some IF steels reach values of 2 to 3.

Interstitial-free steels were developed initially to have a low yield strength and high elongation in a tension test and excellent ductility, but it is now obvious that the IF technique can be expanded to higher strength steels. P and Mn can be added for solid solution strengthening, B can be added to improve grain boundary cohesion and Cr can be added to make IF ferritic stainless steel. Many current grades of cold-rolled sheet steels may be superseded by these C-free steels.

Formerly, many low-C sheet steels were not deoxidized in the furnace. As a result, C and O in the liquid steel reacted during cooling of the ingot, leaving an outer rim of solid steel lower in carbon than the interior. After cold-rolling and annealing the steel surface tended to be free of defects. Now the great majority of low-carbon steels are continuously cast and the rimming action cannot be tolerated, so the steel is "killed" with an addition of 0.03 to 0.04% Al. Any aluminum that does not combine with oxygen is free to form AlN, thereby minimizing subsequent strain-aging. *Strain-aging* is the term applied to various changes in properties of a metal that occur by the interaction of interstitial solute atoms with dislocations during or after plastic deformation. Such interactions appear in many different alloy systems. In steels the significant interstitials are C and N atoms. When property changes occur after plastic deformation the process is called *static strain-aging*; when they are concurrent with plastic strain it is called *dynamic strain-aging*. The effects of static strain-aging on the tensile properties of a low-carbon steel are shown schematically in fig. 40. When a specimen is strained, allowed to stand for a period of time at temperatures as low as ambient then re-strained the following effects are noted:

1. an increase in yield stress, ΔY
2. a return of the Lüders strain, ε_L
3. an increase in ultimate tensile stress, ΔU
4. a decrease in total elongation, $\Delta\varepsilon$
5. embrittlement.

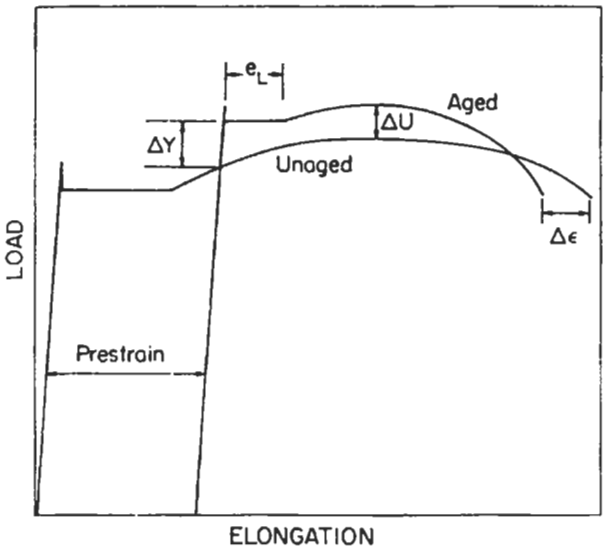


Fig. 40. Effects of static aging on the load-elongation curve of iron containing interstitial solutes: ΔY = change in yield stress due to strain-aging, e_L = Lüders strain after strain-aging, ΔU = increase in ultimate tensile strength due to strain-aging, $\Delta\varepsilon$ = decrease in elongation due to strain-aging.

There is an unfortunate consequence of the return of the Lüders strain. If cold-rolled and annealed sheets are temper rolled to flatten, then allowed to age, during subsequent forming "stretcher strains" appear. These are streaks in which the thickness of the sheet is very slightly reduced, a consequence of inhomogeneous yielding, and the appearance of the sheet is spoiled. The principal cause of strain aging in C steel is N in solid solution in ferrite. It is more effective than C because of its greater solubility. However, if aging after plastic strain is done at temperatures of about 100°C or above, C becomes effective. The conventional method of eliminating strain aging is to use aluminum-killed steels, but this is not completely effective because some C remains in solid solution. To bring the interstitial solutes to the necessary concentration of about one part per million requires addition of elements such as titanium, as in IF steels.

Figure 41 shows the effects of dynamic strain aging on the tensile properties of a common carbon steel. When the steel is strained at temperatures between 100 and 250°C, serrations appear on the stress-strain curves and the rate of work-hardening is greatly increased. Dislocations generated during plastic straining are quickly pinned by mobile interstitial solutes and more dislocations must be generated before the strain can continue. The results are an increased dislocation density for a given strain and a high rate of work-hardening.

Most effects of strain-aging are detrimental but in at least one instance, bake-hardening, the process proves useful. As automobile body sheets have been made thinner to reduce weight they become more susceptible to denting. To counteract that, strength must be increased, but that causes difficulties in forming. An answer is found with steels

References: p. 1618.

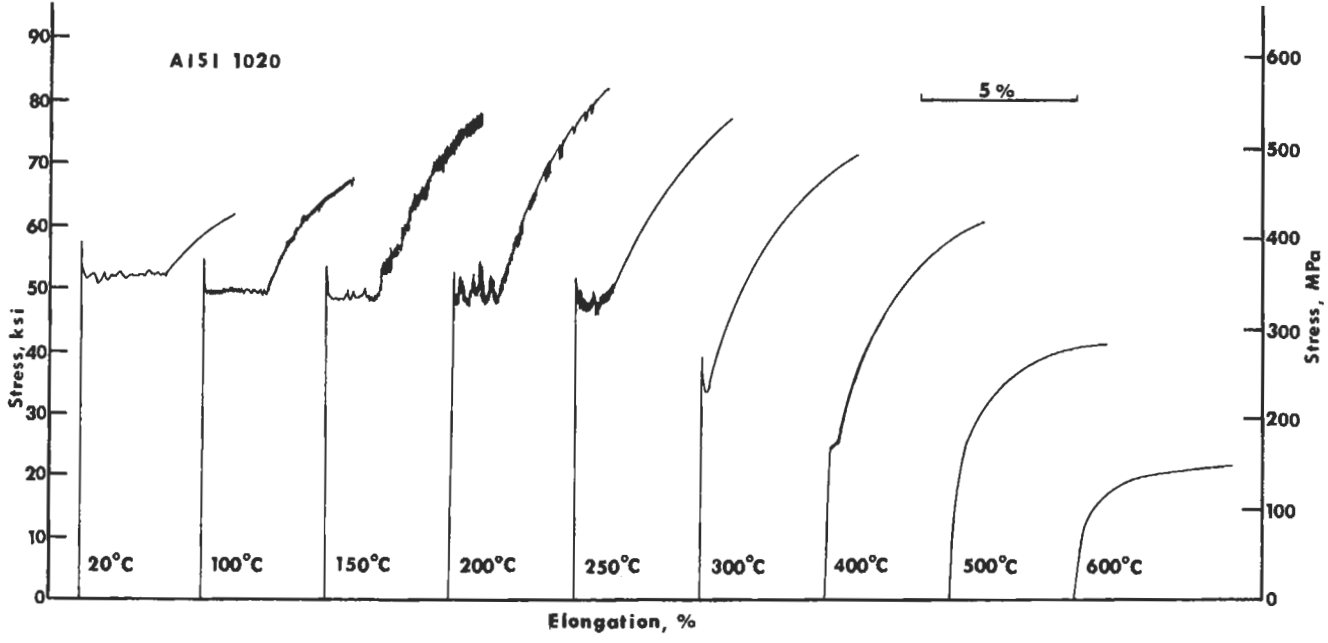


Fig. 41. Stress-elongation curves of a carbon steel strained in tension, $\dot{\varepsilon} = 1.75 \times 10^{-4}\ s^{-1}$.

containing 0.005 to 0.01% C with Al plus Nb or Ti, cold-rolled, continuous-annealed at temperatures near 900°C, then cooled rapidly, leaving about 0.001% C in solid solution, fig. 42 (SATOH *et al.* [1992]). When the sheet is formed into a fender, painted, then baked at 170 to 200°C, strain aging adds an increment of strength of about 50 MPa.

The latest innovation in the production of low-carbon sheet steels is the continuous casting of slabs 50 mm thick. Without being cooled the slab is reheated to rolling temperature, then passed through a five-stand hot rolling mill, rolled to desired thickness and coiled.

A process somewhat further removed from commercial production is the continuous casting between water-cooled copper rolls of sheet about 2 mm or less in thickness. These require only subsequent cold-rolling and annealing (SHANG and WRAY [1989]). Questions arise as to whether current techniques of obtaining desirable microstructures, textures and surface quality will be applicable to these new products.

The properties of low-C steels are determined primarily by ferrite grain size and the type and distribution of carbide particles (RICKETT and KRISTUFEK [1949]). Ferrite grain size is determined by prior austenite grain size which in turn is controlled by the deoxidation of liquid steel, or lack thereof. As discussed before, grain growth in austenite is inhibited by precipitation of AlN, or less commonly by Nb, V or Ti carbonitrides. Grain size is also affected by the degree of cold work and the annealing temperature and time. Grain size can be reduced or made more uniform by re-austenitizing at a low

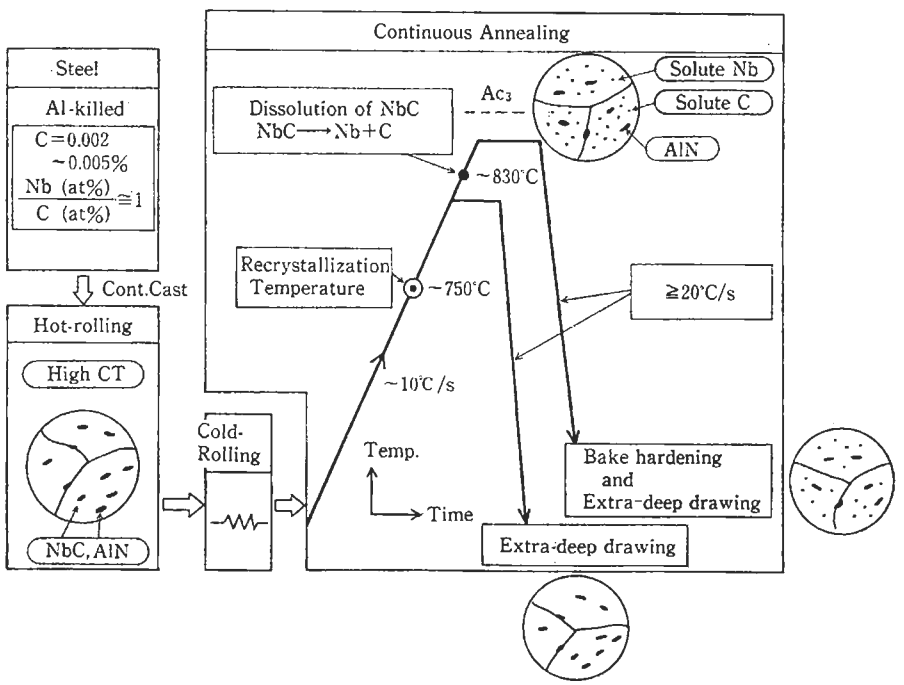


Fig. 42. Processing of bake-hardening and extra-deep drawing sheet steel (SATOH *et al.* [1992]).

References: p. 1618.

temperature and air cooling (normalizing). Carbides can take the form of massive particles of cementite or volumes of pearlite, depending upon cooling rate from above A_1 temperature.

As long as ferrite is the continuous phase, the yield strength of the steel is determined by the grain size and solute content of the ferrite. Banded ferrite–pearlite structures are common in hot-rolled steels, resulting in anisotropy of mechanical properties. The banding results from the segregation of Mn to austenite grain boundaries during heating. Hot rolling spreads the Mn-rich regions into sheets. During cooling, pearlite nucleates preferentially in the Mn-rich portions, giving rise to sheets of pearlite islands.

As C content increases, pearlite becomes the continuous phase and we must consider the properties of pearlite particularly when the structure is entirely pearlite, as in eutectoid carbon steels used for rails. The yield strength of pearlite is dependent upon the interlamellar spacing; GENSAMER *et al.* [1942] determined a proportionality between yield strength and the logarithm of interlamellar spacing. Pearlite lamellae do not have a constant true spacing; there is a distribution of spacings about a mean value, $\bar{\lambda}_o$. A mean intercept spacing $\bar{l}$ is measured by drawing random straight lines on a micrograph. The total length of line divided by the number of lamellae intersected is defined as $\bar{l}$. It can be shown that

$$\bar{l} = 2\bar{\lambda}_o$$

(RIDLEY [1984]).

The yielding of pearlitic is presumed to occur by motion of dislocations in the ferrite lamellae (ALEXANDER and BERNSTEIN [1989]). The ductility and impact toughness of fully pearlitic steels are more dependent upon prior austenite grain size than upon interlamellar spacing, large grains being detrimental (LEWANDOWSKI and THOMPSON [1986]).

5.2.1. High-strength low-alloy steels

The users of steels are greatly interested in reducing costs by eliminating heat treatments and purchasing steels that have properties equal to those of heat-treated alloy steels. These newer steels should have minimal carbon content to improve toughness and weldability and minimal expensive alloy additions. To answer this demand steel producers have developed a wide variety of *high-strength low-alloy* (HSLA) steels. These are intended to meet certain mechanical properties rather than specified chemical compositions. They have certain features in common, one being refinement of ferrite grains. Grain refinement is the only strengthening mechanism that simultaneously increases strength and toughness. Strengthening by precipitation of very fine particles reduces toughness so it must be offset by grain refinement.

Precipitation and grain refinement are managed by control of hot-rolling temperatures and deformations. Inclusions of MnS are elongated during hot rolling thereby reducing toughness in the transverse direction so elements are added to make the inclusions resistant to hot deformation. This is termed inclusion shape control. To control the structure formed during cooling from hot-forming processes, (equiaxed ferrite–pearlite, or acicular ferrite, sometimes called low-carbon bainite), alloying elements are added to

lower the transformation temperature and/or the cooling rate from the hot-rolling temperature can be accelerated.

The fine precipitates used to strengthen HSLA steels and to control grain size are carbo-nitrides of Ti, Nb and V. Titanium has the advantage of also serving to change elongated sulfide inclusions into discrete particles. Additions of calcium to the liquid steel before casting also serve to control inclusion shape. The carbo-nitrides noted are chosen because of their very low solubility in austenite, Ti(C,N) being the least soluble and V(C,N) the most. By precipitating as fine dispersions in austenite they inhibit austenite recrystallization and grain growth. They are much less soluble in ferrite than in austenite; in consequence, any C, Nb, or V remaining in solution in austenite tend to precipitate at successive positions of the intermittently moving γ–α boundaries leaving behind in the ferrite planar arrays of fine, closely spaced particles only resolvable by transmission electron microscopy. This is termed *interphase precipitation.* These microalloy additions are often combined, Ti or Nb for grain refinement, V for precipitation in ferrite.

Copper occupies a special place in HSLA steels. When added in concentrations between 0.5 and 1.0% it can remain in solution even during slow cooling from the austenite range. An object can be formed when strength is low, then aged by holding in the range 500 to 600°C. Coherent bcc particles form, containing iron as well as copper. The maximum strengthening, fig. 43, is attained just before the particles transform in situ to fcc copper (HORNBOGEN [1964], HORNBOGEN *et al.* [1966], GOODMAN *et al.* [1973]). Copper, in combination with phosphorus, is used in the so-called weathering steels that are resistant to atmospheric corrosion.

The most common types of HSLA steel have a ferrite–pearlite structure after hot rolling. They can have yield strengths up to 700 MPa, as compared to 200 MPa in plain carbon steels with the same type of microstructure. The C content of most HSLA steels is less than 0.20%.

To obtain an acicular ferrite structure after hot-rolling additions of B, Mo or Mn are used to decrease the transformation temperature. For improved toughness and weldability at high-strength levels, C must be kept below about 0.08%.

Not all HSLA steels are shaped in the mills. With microalloying additions, medium carbon steels can be hot forged and with controlled cooling can eliminate the need for quenching, tempering, straightening and stress relieving.

Dual-phase steels are a special category of HSLA steels. Their microstructure consists of a ferrite matrix with 10 to 20% martensite (or lower bainite) islands. The virtue of this structure lies in the fact that the formation of martensite produces stresses around each island. Yielding then occurs at low applied stress and the yielding is continuous, but because of the presence of the martensite islands, work-hardening is rapid. These features, together with the soft, low-carbon, $<0.10\%$, ferrite provide good formability and strength near 600 MPa after 10% strain. Dual-phase steels generally contain about 1.5% Mn and 0.50% Si. After hot rolling or cold rolling the steels are *intercritically annealed* (heated into the $\alpha+\gamma$ region). The amount of austenite formed varies with the initial carbon content and the annealing temperature. During rapid cooling the austenite pools transform to martensite, with perhaps some retained austenite. The intercritical annealing

References: p. 1618.

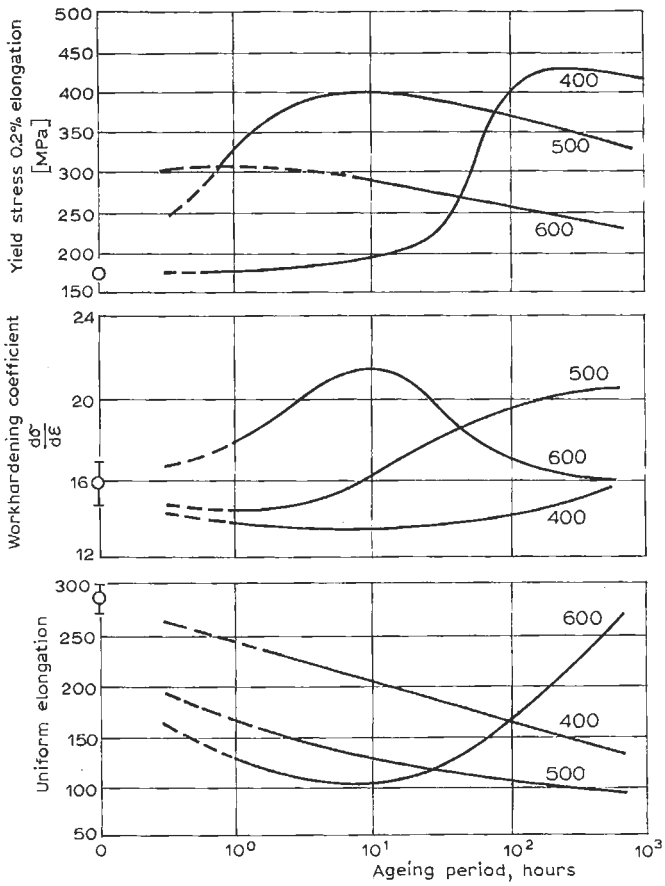


Fig. 43. Precipitation hardening of Fe–0.91% Cu alloy, quenched from 840°C and aged at temperatures indicated (HORNBOGEN *et al.* [1966]).

is done in continuous annealing lines, if they are designed for sufficiently high temperatures (see, for instance, OHASHI [1988]). During forming operations the ductility can be enhanced by strain-induced transformation of retained austenite. Plastic deformation takes place almost exclusively in the ferrite component of the microstructure. Additions of strength can be gained by strain aging during paint baking. The principal market is the automobile industry.

5.3. Strength of martensite

The effect of carbon content on the hardness of untempered martensite is shown in fig. 44 (KRAUSS [1978]). Yield strength is a linear function of carbon content up to about 0.40% (fig. 45). The yield strength of shock-transformed Ti-gettered iron containing virtually no carbon in solid solution establishes the zero point. The departure from

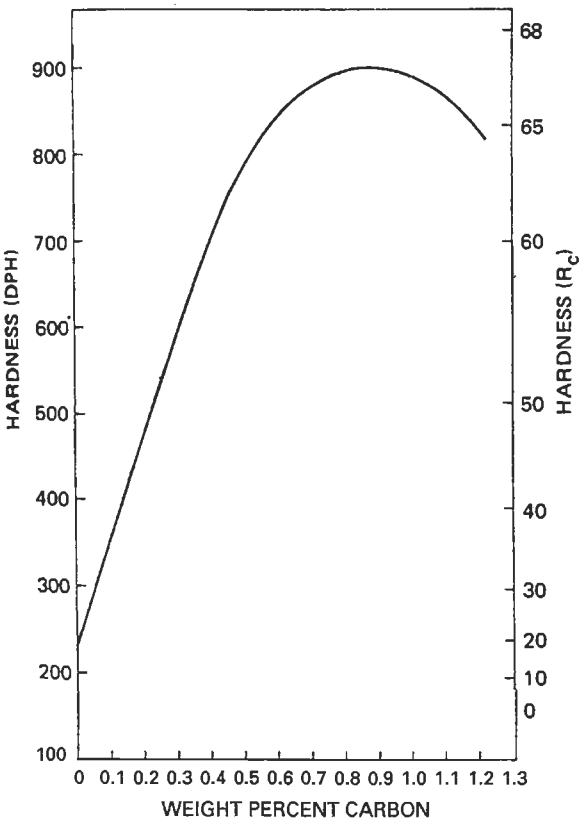


Fig. 44. Hardness of martensite as a function of C content in Fe–C alloys and steels (KRAUSS [1978]).

linearity above 0.40% C shown in fig. 44 is caused by the retention of austenite. As this phase increases in amount the hardness reaches a maximum at about 0.80% C, then decreases rapidly. The simple relation shown in fig. 45 results from a combination of several factors. The relative importance of these factors remains a matter of dispute (LESLIE and SOBER [1967], OWEN [1992]).

Current interest centers on low-carbon martensites because of concern with toughness and weldability. In martensite with carbon contents of about 0.15% or less the defect structure, including prior austenite grain boundaries, provides the largest contribution to strength. At or above carbon contents of 0.30% rearrangement of carbon atoms and dislocations during the quench can dominate the strength, unless the M_s temperature is below about –60°C, when true solid-solution hardening by carbon dominates (WINCHELL and COHEN [1962]). As a further complication it should be noted that iron-carbon martensites do not have a unique strength, the yield and flow stresses are greater in compression than in tension, by about 15% (RAUCH and LESLIE [1972]).

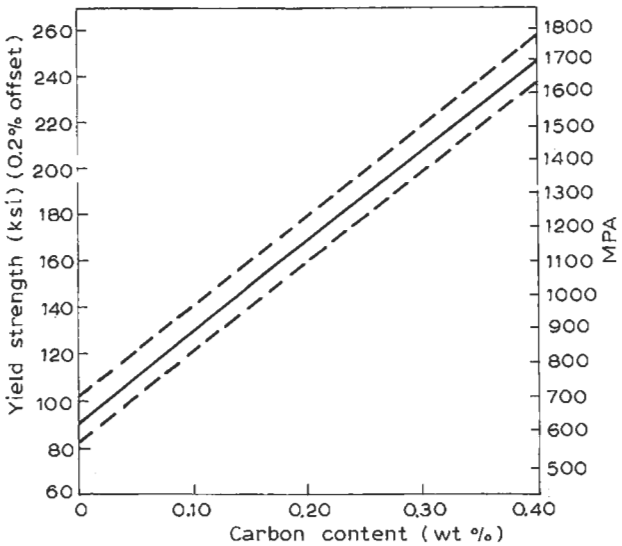


Fig. 45. The effect of carbon content on the yield strength of untempered martensites.

5.4. Strength and ductility of tempered martensite

Tempering is the process of heating martensitic steels to obtain a desired combination of strength and ductility (§ 3.5). High strength is easily obtained; to achieve a combination of strength, sufficient ductility to resist fracture at points of stress concentration or during impact loading, retention of ductility at low temperatures, weldability and low cost is a major challenge for designers of steels. For a given carbon or low-alloy steel the hardness (strength) can be calculated by the method of GRANGE *et al.* [1977]. Figure 46 is a plot of the hardness of iron-carbon alloys, quenched to martensite, then tempered for one hour at the temperatures indicated. Additions of Mn, P, Si, Ni, Cr, Mo or V do not affect the hardness of as-quenched martensite but each of them retards softening during tempering. To estimate the hardness of a given steel the hardness associated with a given carbon content and tempering temperature is read from fig. 46. To this is added an increment of hardness attributed to each of the alloying elements present, taken from plots provided. The temperature range in which tempered martensite embrittlement occurs must be kept in mind.

Within the range of austenite grain sizes normally present in commercial steels, i.e., ASTM grain sizes 5 to 10, there is little effect of grain size on yield strength after tempering. However, if the grain size is reduced beyond that range there is a significant gain in both yield strength and toughness. GRANGE [1966] attained ultra-fine austenite grains by four cycles of brief austenitizing at low temperature, quenching, then finally tempering at 205°C. Results are shown in fig. 47 for three kinds of steel. They can be represented by a Hall–Petch type equation

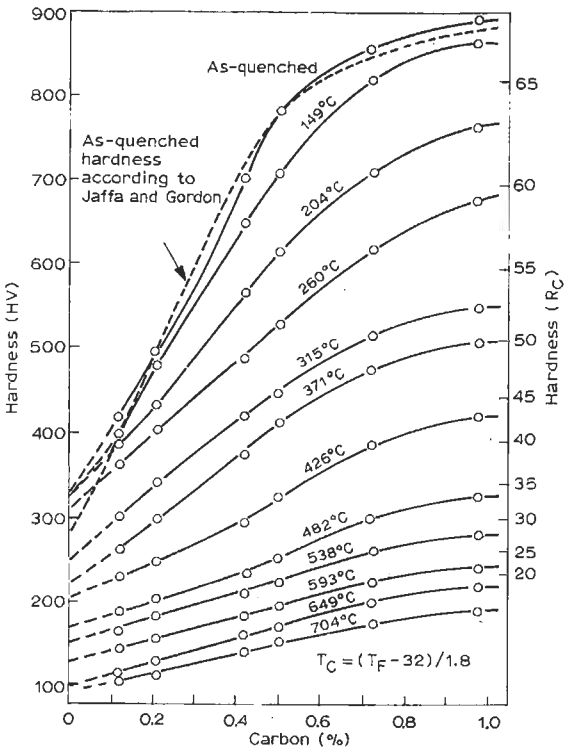


Fig. 46. Hardness of tempered martensite in Fe–C alloys (GRANGE *et al.* [1977]).

$$\sigma_{ys} = \sigma_o + Kd^{-\frac{1}{2}} \tag{7}$$

where σ_o is the intercept at zero on the abscissa and d is grain diameter. The increase in yield strength was not accompanied by a decrease in ductility. Later, TOMITA and OKABAYASHI [1985] followed this procedure but after the last austenitizing step they transformed isothermally at 320°C to form 15 to 25% bainite, then quenched to martensite and tempered at 200°C. The result was an increase in the yield strength found by GRANGE in a similar steel, with no loss in ductility. The bainite is constrained by the enveloping martensite and presents another set of boundaries to stop the growth of martensite crystals.

In instances when the steel contains elements such as V, Mo, and W, the results of tempering can be quite different. When present in concentrations between 1% and 18%, together with 0.80 to 1.50% C, these elements form the basis of high-speed tool steels. These must be austenitized at a high temperature, sufficient to dissolve most of the stable carbides of V, Mo and W. The undissolved carbides provide wear resistance to the final product. After quenching, the steels are tempered between 400 and 500°C to precipitate fine carbides. As shown in fig. 48, strength is preserved at these temperatures, which can be encountered during high-speed machining of steels.

References: p. 1618.

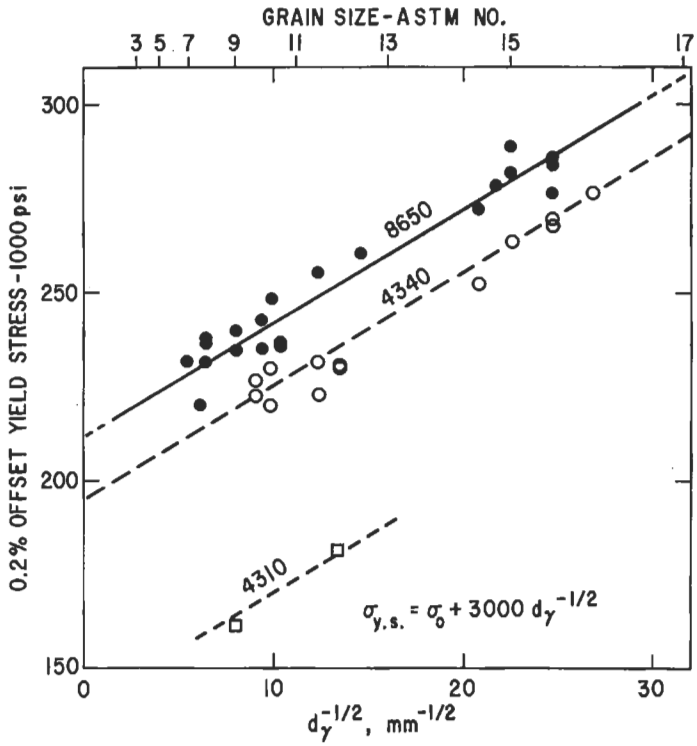


Fig. 47. Effect of prior austenite grain size on the yield strength of martensite in three commercial steels tempered at 205°C (GRANGE [1966]).

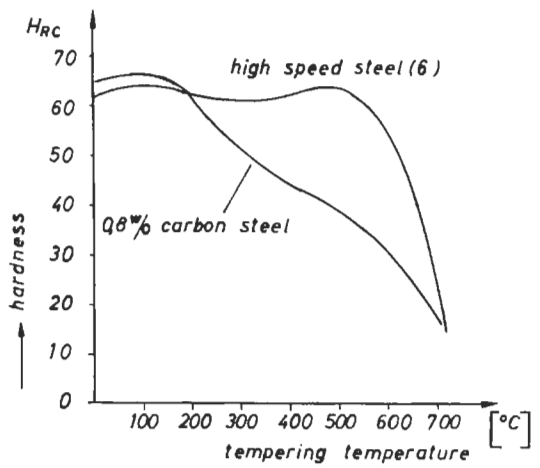


Fig. 48. Strengths of a high-carbon steel and a high-speed tool steel (table 2, no. 6) after tempering, illustrating secondary hardening of the tool steel.

5.5. Ultra-high-strength steels

5.5.1. Maraging steels

According to the definition of steel in the first sentence of this chapter, maraging steels are not steels at all, for they contain no deliberately added carbon; carbon is an undesirable impurity. They are iron-nickel alloys, commonly containing 18 wt% Ni. They also contain 3 to 5 wt% Mo, 0.2 to 1.8 wt% Ti and 0.1 to 0.2 wt% Al. Some contain Co, others have none. The heat-treatments to develop strength are simple. The alloys are hot-worked at above about 920°C, austenitized at about 850°C and cooled. The M_s temperature is between 200 and 300°C and the alloy is completely martensitic before reaching room temperature. The structure is iron–nickel bcc lath-type martensite, which unlike Fe–C martensite is soft, ductile and machinable. However, the rate of work hardening is low since there are no interstitial atoms available to pin dislocations. The uniform elongation in tension is only about 1 to 3%. The function of the martensite is to provide a high density of nucleation sites for the precipitation of intermetallic compounds and probably to provide paths for accelerated diffusion of substitutional solutes. The alloys are aged at about 480°C for three hours.

The result is the precipitation of intermetallics of the type Ni_3X, where X may be Ti, Mo or V. These particles are of the order of 10 nm in size with a number density of 10^{23} m^{-3} (SHA *et al.* [1993]). Titanium is the element that precipitates most rapidly and is the one used to control strength, fig. 49 (SPEICH and FLOREEN [1963]). When present, Co raises the activity of Mo, thus hastening its precipitation as Fe_7Mo_6 or Fe_2Mo. Overaging results in reversion to the stable austenite + ferrite, agglomeration of particles and softening (MEYER and HORNBOGEN [1967]).

Maraging steels can be provided with yield strengths ranging from 1400 to 2400 MPa. The most remarkable feature of these alloys is their combination of high yield stress with high fracture toughness over a wide range of temperatures, fig. 50 (BRUCH

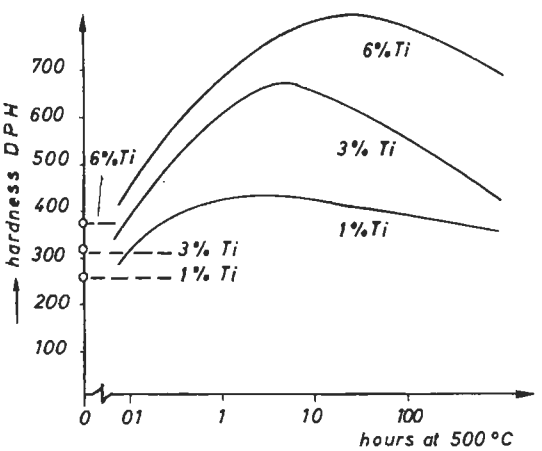


Fig. 49. Effect of Ti content on hardness of an Fe–20% Ni martensite after aging at 500°C (SPEICH and FLOREEN [1963]).

References: p. 1618.

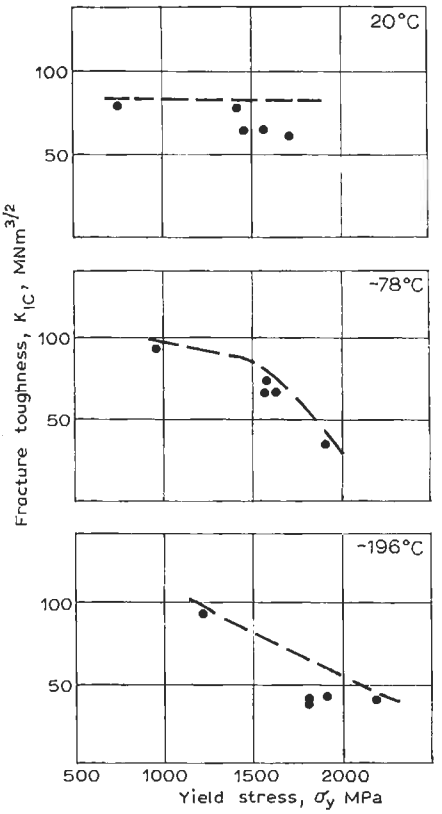


Fig. 50. Fracture toughness of an 18 Ni, 8 Co, 5 Mo, 0.1 C maraging steel as a function of test temperature and yield stress (BRUCH and HORNBOGEN [1978]).

and HORNBOGEN [1978]). Fracture toughness is not greatly sensitive to the aging treatment so at maximum yield stress the alloy is still tough.

In a drastic modification of maraging steels, the nickel content is reduced from 18% to 11%, molybdenum is reduced to 1.2%, 3% Cr is added and Co is increased to 13.4% and about 0.25% C is added. The presence of Co increases the activity of C and raises the M_s temperature. The alloy is austenitized at 885°C, leaving some undissolved carbides. After air cooling the alloy is refrigerated at −73°C to transform retained austenite and finally aged at 480°C. Precipitates formed during aging appear to be M_2C carbides. Some martensite reverts to austenite in the form of thin films. The yield strength can be 1650 MPa and the plane strain fracture toughness 130 MPa$\sqrt{m}$. Charpy V-notch energy at −100°C can be about 32 J (AYER and MACHMEIER [1993]).

5.5.2. Modified standard steels

An arbitrary definition of an ultra-high strength steel is any steel having a yield strength greater than 1380 MPa. The standard against which modifications are judged is

AISI 4340 and equivalent medium-carbon Ni, Cr, Mo steels produced in other nations. These can be oil quenched and tempered to any desired hardness. Properties can be improved by vacuum arc or electroslag remelting to lower the H, N and O contents and to reduce the number of non-metallic inclusions.

An improvement is obtained by increasing the Si content of 4340 from 0.15–0.30% to 1.45–1.80% to take advantage of the well-known effect of Si in inhibiting the growth of carbide particles. This steel, called 300 M, can be tempered at a higher temperature than 4340 to develop the same hardness, thereby reducing quenching stresses. The addition of Si also moves the temper embrittlement range to higher temperatures. Tempering at 315°C produces a maximum in notch impact toughness and near a maximum yield strength. The additional Si also increases hardenability and adds a component of solid solution strengthening. A small addition of vanadium may provide fine V(C,N) particles during tempering.

Crucible Hy–Tuf is a modification of AISI 4130 with increased Mn and Si and an addition of Ni. In Ladish D6AC the Cr and Mo contents are increased over those in 4340.

5.5.3. Thermomechanically-treated steels

Ultra-high-strength steels can be produced by thermomechanical treatments. For our purposes we can divide such treatments into three groups:

1. Supercritical TMT (above Ae_3)
2. Intercritical TMT (between Ae_1 and Ae_3)
3. Subcritical TMT (below Ae_1)

Considering the first group, if austenite is heavily deformed without recrystallizing it takes up a very sharp texture, {124} <112>. During transformation the martensite also takes up a sharp texture, (112)+(111) which is retained during tempering (HU *et al.* [1976]). The texture provides a high modulus of elasticity in a direction normal to the surface and an increased yield strength in compression in the through-thickness direction, thereby improving the ballistic properties (GHOSH and PATON [1978]).

Forming in the ferrite-plus-austenite range, followed by quenching before recrystallization, yields a structure of martensite fibers in ferrite, with very high toughness in a direction normal to the fibers (GRANGE [1970]).

Deformation of austenite below A_{e1} is the process called ausforming. It requires a steel having a large bay between pearlite and bainite in an isothermal transformation diagram. The steel is cooled — avoiding pearlite — heavily worked, then cooled below M_s. Yield strengths of about 2600 MPa can be developed. The strengthening is attributed to refinement of martensite platelets formed in flattened austenite grains, strain-induced precipitation of carbides in austenite and development of a strong texture in the martensite.

In the third category of thermomechanical treatments, a steel can be deformed as it is transforming to pearlite or bainite.

As an example of the strength that can be developed in thermomechanically-treated steels, consider the following: a steel containing 0.6% C and 5% Ni was ausformed, quenched, tempered, strained 3% then aged at 120°C, a combination of ausforming and strain aging. Yield strength was 3590 MPa, tensile strength 3630 MPa (MATAS *et al.*

[1964]) This compares with attainable tensile strengths in cold-drawn eutectoid carbon steel wire of 4800 MPa.

With the exception of super-critical treatment (controlled hot-rolling) these thermo-mechanical processes have not been applied in industry for the following reasons:

1. Large reductions in cross section must be made, so large sections cannot be treated.
2. At the temperatures involved, loads on equipment are severe.
3. The steel must be used in the simple shape in which it is formed.
4. Joining is difficult; welding is impossible.

5.6. Tool steels

Tool steels are used for the shaping of other metals or for the cutting of other hard or abrasive materials. There are seven categories of tool steels but lack of space forbids discussion of more than one, the most complex and important, high-speed steels. These are of two types, containing Mo or W. A typical composition of a W type is given in table 2.

The heat treatment of high-speed steels consists of austenitizing at a very high temperature, 1250 to 1300°C to dissolve $M_{23}C_6$ carbides but leaving about 10% of the volume consisting of MC and M_6C carbides. Hardenability is very high with a deep bay in the IT diagram. Transformation is accomplished by quenching to about 500°C to relieve thermal stresses, then air cooling. The structure is composed of martensite, about 15 to 25% retained austenite and about 10 to 20% undissolved carbides. The heat treatment is distinguished by the necessity for multiple tempering to transform the retained austenite. The first temper at 550 to 595°C "conditions" the austenite (presumably austenite loses carbon to martensite in which carbides are precipitating). The second (or successive) temper at the same temperature transforms the austenite and completes the secondary hardening process in which the precipitation of carbides raises the hardness to 63 to 65 Rockwell C, equivalent to the hardness after quenching, with greater ductility. The volume of carbides in the structure can be as great as 40%. These high-speed steels can be used at temperatures of 500°C or more.

5.7. Austenitic steels

Some alloying elements, principally Ni and Mn, expand the γ loop in the Fe–C system and other alloying elements can retard the γ–α transformation so that the fcc structure can be metastable to ambient temperatures and below. Therefore, we must consider the differences between the mechanical properties of fcc and bcc iron. These differences are inherent in the crystal structures. Diffusivities of solutes are lower in fcc than in bcc at the same temperatures; the solubility of interstitial solutes is greater in fcc than in bcc. The stacking fault energy in bcc is very high, in fcc it is much lower but varies with solute additions.

In fcc iron alloys dislocations tend to dissociate readily and cross slip is difficult. Dislocations are largely restricted to {111} slip planes, leading to planar arrays, narrow slip lines (fig. 29), a high rate of work hardening and large uniform elongations in tensile tests. There is only a small dependence of strength on temperature or strain rate. In

contrast, in bcc iron cross-slip is easy; slip occurs on any plane that contains the <111> slip direction, leading to cellular arrays of dislocations. Macroscopic deformation depends upon motion of screw dislocations which are slightly dissociated and immobile and must be constricted by thermal activation. The result is a pronounced dependence of strength on temperature and strain rate and a tendency toward brittleness at low temperatures and/or high strain rates.

Because of the relatively high solubility of interstitial elements in fcc iron, strain aging is infrequent. A notable exception occurs in Hadfield manganese steel, our oldest alloy steel. This contains 10 to 14 wt% Mn and 1 to 1.4 wt% C. As a result, the fcc structure is stable at ambient temperatures. The rate of work-hardening in gouging wear is exceptionally high. Early on this was attributed to a strain-induced transformation to martensite but when the carbon content is maintained at the proper level this does not occur. Later, the high work-hardening and wear resistance were attributed to profuse mechanical twinning but when twinning is eliminated by a small addition of an element that increases the stacking fault energy, the high work-hardening remains (ZUIDEMA *et al.* [1987]). Because of the high concentrations of Mn and C and their mutual attraction in solution, it was concluded that dynamic strain-aging brought about by the reorientation of C members of Mn–C couples in the cores of dislocations is the principal cause of the rapid work-hardening (DASTUR and LESLIE [1981]).

The most important class of austenitic steels is that of the stainless steels containing 17 to 25% Cr and 8–35% Ni. The Cr content is the basis of all stainless steels, whether ferritic, martensitic, austenitic or duplex. It provides corrosion resistance. Nickel, as stated previously, expands the γ loop so that the steels remain austenitic at low temperatures. They have excellent cryogenic properties as a consequence of their fcc structure. They can be hardened by cold work, by elements in solid solution, especially N, or by precipitation strengthening. MURATA *et al.* [1993] have reviewed these mechanisms. The solubility of C in the standard 18–8 stainless steel declines sharply with decreasing temperature; at 600°C it is less than 0.01%. Since the diffusion coefficient of interstitial atoms in austenite is much lower than in ferrite precipitation of carbides is delayed. However, if heated as during welding, $Cr_{23}C_6$ carbides can form at grain boundaries, thereby causing a local depletion of Cr and providing paths for corrosive attack. This problem can be alleviated by reducing the C content to less than 0.03% or by adding Ti or Nb in sufficient quantity to combine with C and precipitate as stable carbides.

5.8. Steels for low-temperature applications

Austenitic Ni–Cr steels would be ideal for low-temperature service, except that cost is prohibitive and the strength is often inadequate. Therefore, ferritic steels must be modified in composition and microstructure to counteract their inherent tendency to fracture in a brittle manner at low temperature. In brittle fracture plastic strain and energy absorbed are negligible. Cracks propagate along grain boundaries or along {100} planes of ferrite. A steel to be used at low temperature must fulfill the requirement that its ductile-to-brittle transition must be below the temperature of use. Questions still remain regarding the relation between methods of determining this transition temperature and behavior in service.

References: p. 1618.

In designing ferritic steels for use at low temperatures certain rules must be followed; grain size must be minimized, C content must be kept below 0.2% and the lower the better, Ni must be added and strength should be obtained by a structure of acicular ferrite or low-carbon bainite.

When Ni is added in amounts of 5 or 9% the notch toughness of ferritic steel at low temperatures can rival that of austenitic steels, at a much lower cost, fig. 51. It has been known for some time that the elements Pt, Ni, Ru, Rh, Ir and Re in solid solution both strengthen and toughen bcc Fe (FLOREEN and HAYDEN [1967]). All other additions decrease toughness. The toughening effect is on the bcc lattice, not on interactions with interstitials. There is no theory pertaining to this toughening effect, but we can speculate that these metals increase cross slip of screw dislocations by reducing their dissociation and regret that Pt is not more abundant.

5.9. Segregation of solutes and steel purity

Segregation of solutes to grain boundaries in steel is important because it is closely connected to decreasing ductility. *Tempered martensite embrittlement* is described in section 3.5. It occurs when as-quenched martensite is tempered between 250 and 400°C. In contrast, *temper embrittlement* occurs in steels that are quenched, tempered at a high temperature, 600 to 700°C, then placed in service at temperatures from 350 to 575°C. After slow cooling, as in the center of heavy sections, or after long exposure to those temperatures in service, the steel becomes embrittled. This is of particular concern in the generation of electricity because of embrittlement of steam turbine rotors.

As indicated in fig. 39, the grain boundary segregation of a solute is inversely proportional to its bulk solubility. However, the phenomenon of co-segregation also must be considered (GUTTMANN *et al.* [1982]). When two solutes in Fe are mutually attracted, as are Mn and S or Ni and Sb, they will tend to segregate together at equilibrium. A third factor is site competition at boundaries. Carbon and P, or C and S, or B and P compete for these sites. Carbon and B have the desirable characteristic of enhancing grain boundary cohesion. Because of co-segregation between (Mn + Si) and (P + Sn +

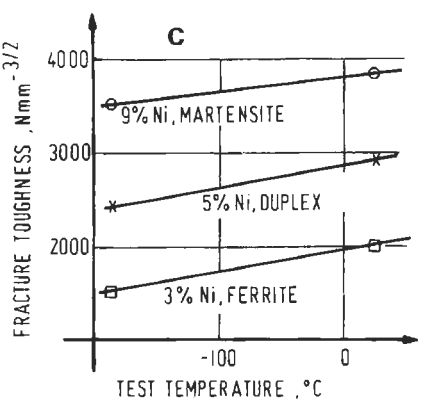


Fig. 51. Low-temperature toughness of three Ni steels, with varied microstructures.

As+Sb) it is desirable to reduce the concentrations of these elements to the lowest possible level.

The culmination of 70 years of reserach was the casting of 120 ton ingots of 3.5% Ni Cr Mo V steel for turbine rotor forgings (JAFFEE *et al.* [1989]). The chemical composition in weight % was

C	Mn	Si	S	Ni	Cr	Mo	V	P	Al	Sn	As	Sb
0.25	0.02	0.03	0.001	3.65	1.75	0.42	0.13	0.002	<0.005	0.003	0.003	0.0012

This super-clean steel was made by electric-arc furnace melting under oxidizing conditions, ladle furnace refining, vacuum carbon deoxidation with argon bubbling and casting with vacuum stream degassing. It is free of temper embrittlement over the entire temperature range and safe operation at 480 to 500°C is possible, thereby increasing the efficiency of power generation.

Removal of P and S from austenitic stainless steels can improve their resistance to intergranular corrosion and intergranular stress corrosion cracking (BRIANT and ANDRESEN [1988]). This topic is also treated in ch. 13.

6. *Other physical properties*

6.1. Steels for nuclear applications

Steels used in the construction of nuclear reactors, for example as canning materials, must not contain elements with isotopes that have a high absorption cross-section for neutrons. B, Co, Ta, Nb, Ti, Al, N are such elements that are normally contained in steels.

The opposite is true for steels that are used as absorber materials or for shielding. Among alloying elements B has the best ability to absorb neutrons over a wide energy range. The macroscopic neutron-absorption coefficient depends on the B concentration in the steel. Because B is hardly soluble in α-Fe it is either present segregated or as particles (Fe_2B). Above 2 wt% B ferritic steels become so brittle that they cannot be worked (grain-boundary brittleness). Therefore austenitic steels are often used that contain up to 5 wt% B as a dispersion of $(Fe,Cr)_2B$. The absorption coefficient of a steel is structure-sensitive, because self-shielding leads to a decrease of the macroscopic absorption coefficient above a certain particle size.

Irradiation with neutrons produces a variety of effects on a very fine scale. In the C–Mn–Mo low-alloy plate steels used in the containment vessels of nuclear fission reactors the flux of neutrons can produce embrittlement (GELLES [1990]). The cause is nanoscale precipitation processes that would not be expected to occur at the temperature of use. Copper is present as a residual element in low-alloy steels, normally in concentrations of about 0.15% or less. Although solubility at temperatures below 500°C is very low the diffusivity of copper at those temperatures is so slight that precipitation does not occur. However, under irradiation the diffusivity is enhanced and the irradiation kinetics at 300°C is about equal to thermal kinetics at 500°C. As a result, ε Cu-enriched

precipitate particles about 5 nm or less in size have formed at 288°C. These particles strengthen the steel and raise the ductile-brittle transformation temperature (SIMONEN [1993]). Steels for containment vessels are now required to have extra low Cu contents.

Another unexpected result of radiation is Cr depletion at grain boundaries in austenitic stainless steels, not connected with precipitation of Cr carbides. The depletion amounts of about 5% Cr over a width of about 5 nm. The depletion correlates with increased intergranular stress-corrosion cracking.

A major problem with austenitic stainless steels under neutron irradiation is development of cavities and volumetric expansion, called swelling. Ferritic/martensitic steels are less susceptible and can replace the stainless steels.

Steels used in fusion reactors have problems in addition to swelling. They are exposed to very high energy neutrons. The alloying elements Cu, Ni, Mo, Nb and N are converted to radioactive isotopes with long half lives, intensifying the problem of waste disposal. New alloys, called low activation materials have been developed to replace the standard stainless steels with Mn substituting for Ni. Martensitic Cr steels with W added for strengthening appear to have promise. This issues are treated in great detail in a recent monograph (FROST [1994]).

6.2. Steels for electrical applications

Steels for electrical and magnetic applications are sold on the basis of core loss, which is the amount of electrical energy dissipated as heat when the steel is magnetized by an alternating current. The reduction of core loss leads to greater efficiency in the generation, transmission and use of electrical power. Silicon is the most effective element for increasing the electrical resistivity of Fe which in turn reduces eddy current losses in alternating current magnetization.

There are two types of Si steels, non-oriented and oriented. Non-oriented grades are low-C alloys containing 0.8 to 3.5% Si plus Al. They are used in less critical applications such as small motors and radio or television transformers. They are hot-rolled, cold-rolled, annealed and decarburized to develop large, clean ferrite grains with no particular effort made to develop a favorable texture except perhaps by the addition of a solute such as Sb which can inhibit growth of certain grains during recrystallization.

The easy direction of magnetization in bcc metals is <100>; to take advantage of this, grain oriented steels are made with <100> in the rolling direction of the sheet and {110} planes in the plane of the sheet. When transformer cores are made, the rolling direction is put in the main flux path. Grain oriented silicon steels are the most sophisticated mass products of the steel industry. They are essentially high-purity Fe–Si alloys with a constant Si content of about 3.25%. Higher Si contents are not used because of difficulties in cold rolling. Production is begun by reheating slabs to the unusually high temperatures of 1315–1375°C, the purposes being to dissolve various compounds (MnS, AlN, BN, TiC, MnSe, etc.) that are subsequently precipitated to serve as grain growth inhibitors and also to ensure that hot rolling is completed in the ferrite region. After cold rolling the steel is annealed to recrystallize and to remove C then annealed at about 1150°C in dry H_2. The purposes of the second anneal are to develop the desired <100>{110} texture by secondary recrystallization in which most primary recrystallized

grains are pinned by the inhibitors so that the favorably oriented grains grow at their expense and to remove the S, N, C or Se in the inhibitors by reaction with H_2. Finally, the magnetic flux density and core loss can be improved by coatings that provide a smooth surface and a residual tensile stress to the very thin steel sheet (INOKUTI *et al.* [1993]). Si steels for transformer laminations are also treated in ch. 28, § 4.4.2 and ch. 29, § 4.2.

7. *Solidification*

7.1. Rimming steel, killed steel

The solubility of oxygen in liquid steel decreases rapidly with decreasing temperature (fig. 52): $Fe_l + \underline{O} \rightarrow FeO$. The product FeO reacts with C to form CO which evolves as bubbles in the liquid. When poured into an ingot mold this evolution of gas causes the steel to rise in the mold. The boiling action can be stopped by putting a steel cap on top of the mold — thus the term *capped steel.* This procedure is often used for steels containing 0.15% C or more. If the evolution of gas, the "rimming action" is allowed to continue, the formation of honeycomb blowholes in the solidifying steel is sufficient to offset the shrinkage that occurs during solidification. The amount of gas evolved must be controlled so that a skin of solid metal forms close to the mold wall. This metal that solidifies first has about the same composition as the steel being poured but as solidification continues the remaining liquid becomes enriched in solute elements, especially S, P, C, Si and Mn. *Rimmed steel* has certain advantages for the manufacture of sheet products; the rim zone contains few non-metallic inclusions leading to high surface quality and ingot-to-product yield is higher than for killed steel.

The term *killed steel* is applied to steels that have been deoxidized by addition of Al or Al + Si to the ladle of liquid steel before the ingots are cast. Because there is little or no oxygen free to combine with C to form bubbles a shrinkage cavity called "pipe" forms at top of the ingot which must be cropped before hot rolling, thus the ingot-to-product yield is reduced. Also, there is a greater number of non-metallic inclusions present in killed steel than in rimmed steel but less segregation of solutes. In rimmed steel the turbulence caused by gas evolution promotes segregation, so alloy steels are killed to obtain greater homogeneity.

With the increased dominance of continuous casting in the production of steels and especially for sheet steels, rimmed and capped steels may soon fade into history. Gas evolution cannot be tolerated in continuous casting, so vacuum carbon deoxidation or deoxidation with a minimal addition of Al are employed.

Vacuum carbon deoxidation is necessary for the production of ultra-low-carbon steels and for the removal of H_2. Carbon is removed as CO. The H_2 content of the liquid steel is proportional to the square root of the partial pressure of H_2 in the vacuum chamber,

$$\%[H] = k\sqrt{P_{H_2}}$$

The removal of hydrogen prevents the formation of internal cracks, called flakes, that occur during cooling from the forging or hot rolling temperature.

References: p. 1618.

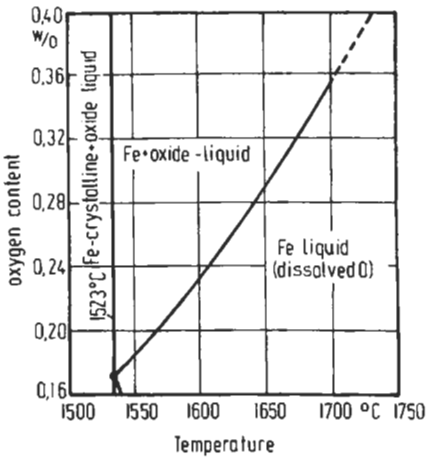


Fig. 52. Partial Fe–O phase diagram.

7.2. Cast irons

Cast irons are alloys of Fe with more than 2 wt% and usually 2.6 to 3.6 wt% C. By varying the composition, cooling rate and minor additions to the melt, the solidification can follow either the thermodynamically stable eutectic reaction liquid $\rightarrow \gamma$ + graphite or the metastable reaction, liquid $\rightarrow \gamma + Fe_3C$. During cooling through the eutectoid interval the structure can be the stable one, γ + graphite $\rightarrow \alpha$ + graphite, or the metastable γ + graphite $\rightarrow$ pearlite + graphite. These yield the common *gray iron*, so called because of the appearance of its fracture. If the eutectoid reaction follows the metastable path $\gamma + Fe_3C \rightarrow$ pearlite + Fe_3C, the result is *white cast iron.* Microstructures are shown in fig. 53.

Which of these paths is followed depends upon the graphitization potential of the melt which in turn is determined by the C and alloy content of the melt, principally Si. Gray irons usually contain 1 to 3% Si. Silicon increases the activity of C, as commonly expressed by the carbon equivalent, CE = %C + 0.3(%Si) + 0.33(%P) – 0.027(%Mn) + 0.4(%S). Gray iron is widely used because of its low cost, strength in compression and high damping capacity.

White irons contain from 1.0 to 2.0% Si. They can be annealed after casting at temperatures between 900–970°C to produce irregularly shaped particles of graphite called *temper carbon,* resulting in malleable iron. If slowly cooled the matrix is ferrite. More rapid cooling or subsequent heat treatments yield matrix structures of pearlite or martensite. Thus, a wide range of mechanical properties is available in these malleable irons.

Ductile cast iron, also called *nodular cast iron,* differs from malleable iron in that the graphite particles form as spheroids during cooling from the liquid (fig. 53). The matrix in the as-cast condition can be ferrite or pearlite or can be altered by subsequent heat-treatment. Ductile iron has advantages over other cast irons in that the graphite spheroids form directly from the melt and no annealing is required as in the production of malleable iron. Strength and toughness are greater than in malleable iron and can be

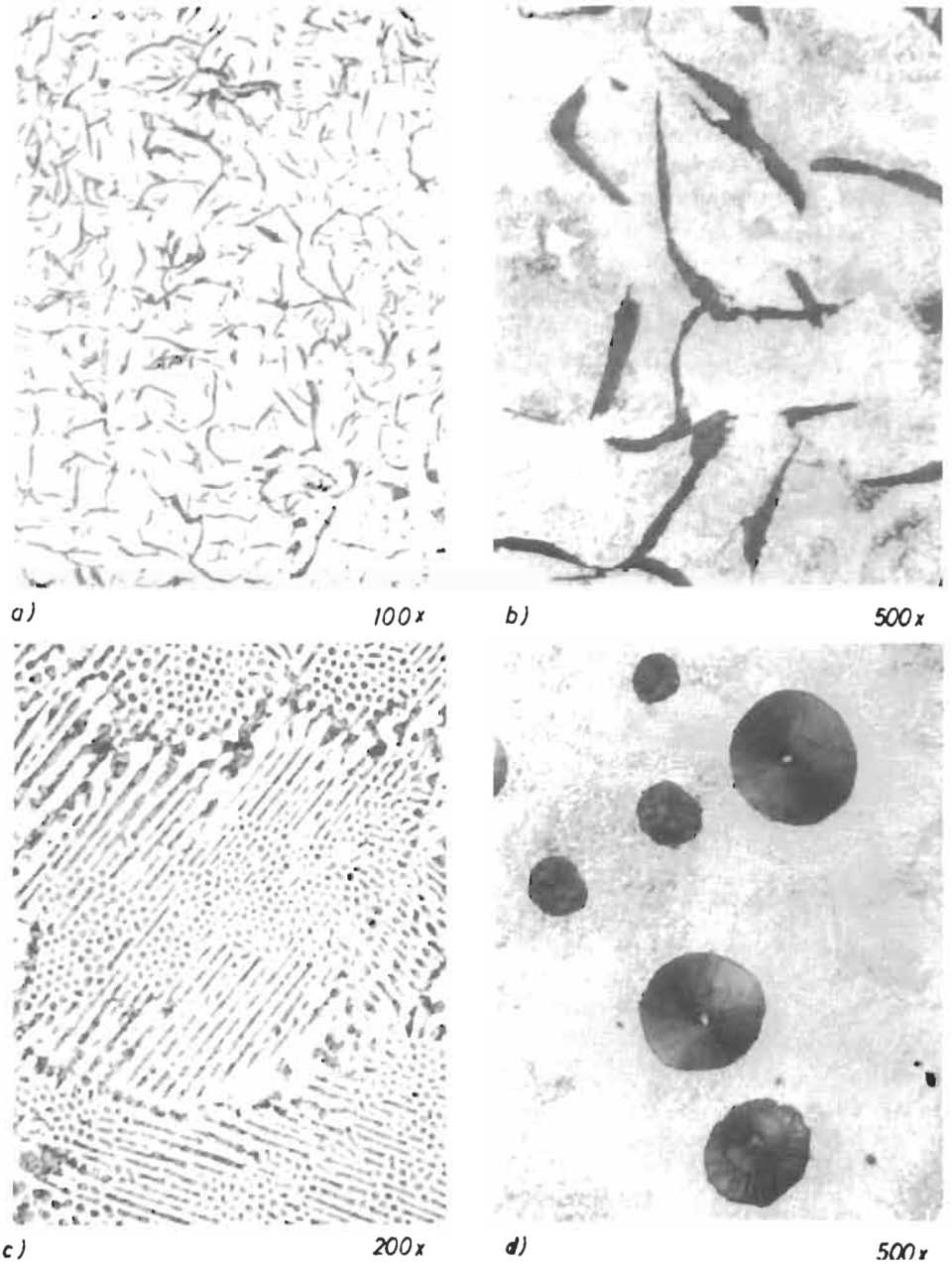


Fig. 53. Microstructures of grey, white, and nodular cast iron: (a) graphite and ferrite; (b) graphite and pearlite; (c) cementite and pearlite; (d) nodular graphite and pearlite. (Courtesy J. MOTZ).

varied over wide ranges by heat treatment. The section size limitation on malleable iron does not apply to ductile iron.

References: p. 1618.

In the production of ductile iron, two additions are made to the melt (SKALAND *et al.* [1993]). The first is the spheroidizing agent, Mg, added as ferrosilicon with 3 to 12% Mg. The Mg vaporizes and rises through the liquid iron, removing S from solution. The second is the inoculant, which provides nucleation sites for graphite spheroids. The inoculant is Ca, Sr or Ba, added in ferrosilicon containing Al. These elements form hexagonal silicates whose (001) basal planes provide interfaces with graphite that are coherent or semi-coherent with low interfacial energy.

There are special grades of cast iron for particular applications; 15% Si provides resistance to corrosion, 11 to 28% Cr gives oxidation resistance, Ni provides a stable austenite matrix.

References

AARONSON, H. I., 1962, in Decomposition of Austenite by Diffusional Processes, Interscience, NY, 387.
ALBERTSEN, J. F., H. P. NIELSEN and V. F. BUCHWALD, 1983, Phys. Scripta, **27**, 314.
ALEXANDER, D. J. and I. M. BERNSTEIN, 1989, Metallurg. Trans. **20a**, 2321.
ANDREWS, K. W., 1965, J. Iron Steel Inst. **203**, 721.
ATKINS, M., 1977, Atlas of Continuous Cooling Diagrams, British Steel Corp, 107.
AYER, R., and P. M. MACHMEIER, 1993, Metallurg. Trans. **24A**, 1943.
BIRCHENALL, C. E., 1951, Atom Movements, ASM, Metals Park OH, 112.
BIRCHENALL, C. E. and R. J. BORG, 1960, Trans. AIME, **218**, 980.
BOEHLER, R., 1986, Geophys. Res. Lett. **13**, 1153.
BRENNER, S. S., 1956, J. App. Phys. **27**, 854.
BRIANT, C. L. and S. K. BANERJI, 1978, Int. Metallurg. Rev. **23**, 164.
BRIANT, C. L. and R. L. ANDRESEN, 1988, Metallurg. Trans. **19A**, 495.
BRUCH, U. and E. HORNBOGEN, 1978, Arch. Eisenhüttenw. **49**, 357 and 409.
CHIPMAN, J., 1973, Metallurg. Trans. **3**, 55.
COHEN, M., 1949, Trans. ASM, **41**, 35.
CHRISTIAN, J. W., 1983, Metallurg. Trans. **14A**, 1237.
COTTRELL, A. H., 1993, Mat. Sci. Tech. **9**, 277.
DASTUR, Y. N. and W. C. LESLIE, 1981, Metallurg. Trans. **12A**, 749.
DAVENPORT, E. S. and E. C. BAIN, 1930, Trans. AIME, **90**, 117.
DOANE, D. V., 1978, Hardenability Concepts With Applications to Steels, AIME, 351.
DUFLOS, F. and B. CANTOR, 1978, in 'Rapidly Quenched Metals III' (ed. B. Cantor), London: Institute of Metals, p. 110.
ELDIS, G. T., 1978, Hardenability Concepts With Applications to Steels, AIME, 126.
ERA, H., M. SHIMIZU and T. HIRUTA, 1986, Metallurg. Trans. **17A**, 621.
FLOREEN, S. and H. W. HAYDEN, 1967, Trans. AIME, **239**, 1405.
FROST, B. R. T., 1994, Nuclear Materials (Vols. 10A and 10B of 'Materials Science and Engineering').
GELLES, D. S., 1990, ISIJ Intern. **30**, 905.
GENSAMER, M.., E. B. PEARSALL, W. S. PELLINI and J. R. LOW, 1942, Trans. ASM, **30**, 983.
GHOSH, A. K. and N. E. PATON, 1978, AMMRC Report 7R 78-40.
GILBERT, A. and W. S. OWEN, 1962, Acta Metall. **10**, 45.
GLOVER, S. G., 1956, in Mechanisms of Phase Transformation in Metals, Inst. Metals, London.
GOLDSCHMIDT, H. J., 1967, Interstitial Alloys, Butterworths, London.
GOODMAN, S. R., S. S. BRENNER and J. R. LOW, Jr., 1973, Metallurg. Trans. **4**, 2363.
GRANGE, R. A., 1966, Trans. ASM, **59**, 26.
GRANGE, R. A., 1970, Proc. 2nd ICSMA, ASM, 861.
GRANGE, R. A., 1973, Metallurg. Trans. **4**, 2231.

GRANGE, R. A., C. R. HRIBAL and L. F. PORTER, 1977, Metallurg. Trans. **8A**, 1775.
GUTTMANN, M., Ph. DUMOULIN and M. WAYMAN, 1982, Metallurg. Trans. **13A**, 1693.
HASEBE, M., H. OHTANI and T. NISHIZAWA, 1985, Metallurg. Trans. **16A**, 913.
HILLERT, M., 1962, Decomposition of Austenite by Diffusional Processes, Interscience, NY, 197.
HONDROS, E. D. and M. P. SEAH, 1977, Int. Met. Rev. **22**, 262.
HORNBOGEN, E., 1964a, Trans. ASM, **57**, 120.
HORNBOGEN, E., 1964b, in High Energy Rate Working of Metals, Central Inst. for Industrial Research, Oslo, 345.
HORNBOGEN, E., G. LUTGERING and M. ROTH, 1966, Arch. Eisenhüttenw. **37**, 523.
HU, H., 1962, Trans. AIME, **224**, 75.
HU, H., G. R. SPEICH and R. L. MILLER, 1976, AMMRC Report CTR 76–22.
HU, H., 1983, Metallurg. Trans. **14A**, 85.
INOKUTI, Y., K. SUSUKI and Y. KOBAYASHI, 1993, ISIJ Intern. **33**, 957.
INTERNATIONAL CONFERENCE ON BAINITE, 1990, Met. Trans. **21A**, 767–889, 1343–1547.
JACK, K. H., 1951, Proc. Roy. Soc. **A208**, 200.
JAFFEE, R. I., T. OHHASI, O. WATANABE and P. MACHNER, 1989, Trans. Iron Steel Soc., Feb., 45.
JOHANNSON, H., 1937, Arch. Eisenhüttenw. **11**, 241.
JOHNSON, J. W., 1960, J. Less-Common Met. **2**, 241.
JOMINY, W. E. and A. L. BOEGEHOLD, 1938, Trans ASM, **26**, 574.
JÖNSSON, B., 1992, Z. Metallkde. **83**, 349.
JUNG, Y. C., K. NAKAI, H. OHTSUBO and Y. OHMORI, 1994, ISIJ International, **34**, 43.
KAJIWARA, S., 1986, Metallurg. Trans. **17A**, 1693.
KEH, A. S. and W. C. PORR, 1960, Trans. ASM, **53**, 81.
KEH, A. S., 1961, in Imperfections in Crystals, Interscience, NY, 213.
KEH, A. S. and S. WEISSMANN, 1962A, in Electron Microscopy and the Strength of Crystals, Interscience, NY, 231.
KEH, A. S., 1962b, Acta. Metall. **11**, 1101.
KEH, A. S. and W. C. LESLIE, 1963, Materials Science Res. 1, Plenum Publishing, NY, 208.
KELLY, P. M. and J. NUTTING, 1961, J. Iron Steel Inst. **197**, 199.
KLEUH, R. L., 1990, Metals Handbook, 10th ed., ASM Intern, **1**, 653.
KRAUSS, G., 1978, Hardenability Concepts With Applications to Steel, AIME, 235.
KRAUSS, G., 1992 and C. J. MCMAHON, Jr., 1992, Martensite, ASM Intern, 295.
LAKE, J. S. H., 1977, AIM Conference, paper 4A-2.
LANGFORD, G., P. K. NAGATA, R. J. SOBER and W. C. LESLIE, 1972, Metallurg. Trans. **3**, 1843.
LESLIE, W. C., E. HORNBOGEN and G. E. DIETER, 1962, J. Iron Steel Inst. **200**, 622.
LESLIE, W. C., J. T. MICHALAK and F. W. AUL, 1963, in Iron and its Dilute Solid Solutions, Interscience, NY 119.
LESLIE, W. C. and R. L. MILLER, 1964, Trans AIM, **57**, 972.
LESLIE, W. C. and A. S. KEH, 1965, in Mech. Working of Steel 2, Gordon and Breach, NY, 337.
LESLIE, W. C. and J. R. SOBER, 1967, Trans. ASM, **60**, 459.
LESLIE, W. C., R. J. SOBER, S. G. BABCOCK and S. J. GREEN, 1969, Trans ASM, **62**, 690.
LESLIE, W. C. and G. C. RAUCH, 1978, Metallurg. Trans. **9A**, 343.
LEWANDOWSKI, J. J. and A. W. THOMPSON, 1986, Metallurg. Trans. **17A**, 461.
LUDWIK, P., 1909, Elemente der technologischen Mechanik, Springer-Verlag, 32.
MATAS, S. J., M. HILL and H. P. MUNGER, 1964, Mech. Working of Steel 1, Gordon and Breach, NY, 143.
MATSUI, H., S. MORIYA, S. TAKAKI and H. KIMURA, 1978, Trans. Jpn. Inst. Met. **19**, 163.
MEYER, W., and E. HORNBOGEN, 1968, Arch. Eisenhüttenw, **39**, 73.
MICHALAK, J. T. and H. W. PAXTON, 1961, Trans AIME, **221**, 850.
MILLER, R. L., 1972, Metallurg Trans. **3**, 905.
MORRISON, W. B., 1966, Trans ASM, **59**, 824.
MURATA, Y., S. OHASHI and Y. UEMATSU, 1993, ISIJ Intern. **33**, 711.
NUTTING, J. W. and W. CHARNOCK, 1967, Met. Sci. J. **1**, 77 and 123.
OHASHI, N., 1988, in Supplementary Volume 1 of the Encyclopedia of Materials Science and Engineering, ed. R.W. Cahn (Pergamon Press, Oxford), p. 185.

OHTANI, H., M. HASEBE, K. ISHIDA and T. NISHIZAWA, 1988, Trans. Iron Steel Inst. Jpn. **28**, 1043.
OWEN, W. S., 1954, Trans. ASM, **46**, 812.
OWEN, W. S., 1992, Martensite, ASM International, 227.
RAUCH, G. C. and W. C. LESLIE, 1972, Metallurg. Trans. **3**, 373.
RICKETT, R. L. and F. C. KRISTUFEK, 1949, Trans. ASM **41**, 1113.
RIDLEY, N., 1984, Metallurg. Trans. **15A**, 1019.
SAITOH, H. and K. USHIODA, 1989, ISIJ Intern. **29**, 960.
SATOH, A., J. MIFUNE and M. MESHI, 1970, Proc. 2nd ICSMA, ASM, 2, 747.
SATOH, S., S. OKADA, T. KATO, O. HASHIMOTO, T. HANAZAWA and H. TSUNEKAWA, 1992, Kawasaki Steel Tech. Rept. No. 27, 31.
SAXENA, S. K., G. SHEN and P. LAZOR, 1993, Science, **260**, 1312.
SHA, W., A. CEREZO and G. D. W. SMITH, 1993, Metallurg. Trans. **24A**, 1221.
SHANG, L-T., and P. J. WRAY, 1989, Metallurg. Trans. **20A**, 1191.
SIMONEN, E. P., 1993, J. Metals, Dec., 45.
SKALAND, T., O. GRONG and T. GRONG, 1993, Metallurg. Trans. **24A**, 2321.
SPEICH, G. R., 1962, in Decomposition of Austenite by Diffusional Processes, Interscience, NY, 353.
SPEICH, G. R. and S. FLOREEN, 1963, Trans. ASM, **57**, 714.
SPEICH, G. R., 1990, Metals Handbook, 10th ed., ASM Intern. **1**, 423.
SPEICH, G. R. and K. A. TAYLOR, 1992, in *Martensite* (ed. G.B. Olsen and W.S. Owen), ASM Internat., 243.
SWISHER, J. H. and E. T. TURKDOGAN, 1967, Trans. AIME, **239**, 426.
TAKECHI, H., 1994, ISIJ International, 34, 1. Special Issue, Physical Metallurgy of Ultra-Low Carbon IF Steels.
TAN, Y. H., D. C. ZENG, K. C. DONG, Y. H. HE and S. A. HU, 1992, Metallurg. Trans. **23A**, 1413.
TOMITA, Y. and K. OKABAYASHI, 1985, Metallurg. Trans. **16A**, 83.
UWAKWEH, O. N. C., J-M. GÉNIN and J-F. SILVAIN, 1991, Metallurg. Trans. **22A**, 797.
VENTURELLO, G. C., C. ANTONIONE and F. BONACCORSO, 1963, Trans. AIME, **227**, 433.
WAYMAN, C. M., 1964, Introduction to the Crystallography of Martensitic Transformations, McMillan, London, 138.
WINCHELL, P. G. and M. COHEN, 1962, Trans. ASM, **55**, 347.
YOUNG, C. M., L. J. ANDERSEN and O. D. SHERBY, 1974, Metallurg. Trans. **5**, 519.
ZENER, C., 1955, Trans. AIME, **203**, 619.
ZHOU, D. S. and G. J. SHIFLET, 1991, Metallurg. Trans. **22A**, 1349.
ZUIDEMA, B. K., D. K. SUBRAMANYAM and W. C. LESLIE, 1987, Metallurg. Trans. **18A**, 1629.

Further reading

Bhadeshia, H. K. D. H., Bainite in steels, Cambridge Univ. Press, 1992.
Decker, R. F. and S. Floreen, Maraging Steels; Recent Developments and Applications, TMS-AIME, Warrendale, PA, 1988.
Pickering, F. B. ed. Constitution and Properties of Steels, VCH, Cambridge, U. K., 1992.
Honeycombe, R. W. K. and H. K. D. H. Bhadeshia, Steels, Microstructure and Properties. 2nd Edn. Edward Arnold, London, 1995.
Olson, G. B. and W. S. Owen, eds, Martensite, ASM International, Materials Park, OH, 1992.
Leslie, W. C. Physical Metallurgy of Steels, Tech Books, Herndon, VA. 1991.
Krauss, G. Principles of Heat Treatment of Steels, ASM International, Metals Park,. OH, 1980.
Kot, R. A. and B. L. Bramfitt, eds., Fundamentals of Dual-Phase Steels, TMS-AIME, Warrendale, PA, 1981.
Doane, D. V. and J. S. Kirkaldy, eds., Hardenability Concepts with Applications to Steel, TMS, AIME, Warrendale, PA, 1978.
Cramb, A. W. New Steel Casting Processes for Thin Slabs and Strip, Iron and Steelmaker, 15, July, 1988, 45.
Marder, A. R. and J. I. Goldstein, eds. Phase Transformations in Ferrous Alloys, TMS-AIME, Warrendale, PA, 1984.
Metals Handbook, ASM International, 10th ed. Vol. 1, 1990.
Progress of the Iron and Steel Technologies in Japan in the Past Decade, Trans. ISIJ. Vol. 25, No. 8, 1985.

CHAPTER 18

POINT DEFECTS

H. J. WOLLENBERGER

Hahn–Meitner-Institut Berlin GmbH,
D-14109 Berlin, FRG

R. W. Cahn and P. Haasen†, eds.
Physical Metallurgy; fourth, revised and enhanced edition

1. Introduction

Point defects are lattice defects of zero dimensionality, i.e., they do not possess lattice structure in any dimension. Typical point defects are *impurity atoms* in a pure metal, *vacancies* and *self-interstitials*. This chapter covers the properties of vacancies and self-interstitials, their interaction with other lattice defects, the production mechanisms and their importance for radiation damage of materials. Properties of atomic solutes per se are of interest mainly with respect to thermodynamics of alloys which are treated in chs. 5 and 6.

Vacancies are produced simply by heating, in concentrations sufficiently high for quantitative investigations. To obtain similar concentrations of interstitials, they must be produced by doing external work on the crystal. Such work is done on an atomic scale by energetic particle irradiation. Collisions between the projectiles and lattice atoms cause displacements of atoms from substitutional sites to interstitial sites. Thus, vacancies and interstitials are produced in equal numbers. As one vacancy and one interstitial together form a *Frenkel* defect, irradiation is essentially a Frenkel defect production process. This is disadvantageous with respect to experimental research on interstitial properties, because radiation induced property changes of crystals always involve vacancy contributions. Vacancy formation is treated in § 2.2.2 and Frenkel defect production in § 3.1.

Plastic deformation also produces vacancies and interstitials. Although its application is much less expensive than particle irradiation, it has not become a common procedure for point defect production because it does not allow a controlled defect production independent of complex dislocation networks. The significance of point defects in plastic deformation is treated in chs. 21 and 22, and their recovery behavior after plastic deformation is touched upon in ch. 28.

Anomalously high point defect concentrations occur in some non-stoichiometric intermetallic compounds. There, vacancies and interstitials clearly play the role of additional alloying elements and are of thermodynamic significance in this sense. A brief insight is given here in § 2.3.

Other methods such as rapid quenching, evaporation on cold substrates or laser annealing depend on thermally activated production, which is treated in the present chapter for controlled conditions which allow successful vacancy investigation.

In pure metals and in the majority of alloys, vacancies provide thermally activated atom transport and, hence, vacancy properties directly influence the atomic transport (ch. 7). Vacancy properties give information on interatomic forces by means of specific perturbations due to the vacant lattice site.

The interstitial is a very interesting defect because of the large lattice perturbation it causes. This large lattice perturbation and the related dynamic and static properties play a key role in the evolution of typical damage structures. These involve not only the microstructure but also the microchemistry of alloys.

The properties of vacancies and interstitials are described in §§ 2 and 3, respectively. In the late 1960s, a controversy arose on the interpretation of the generally observed radiation damage recovery. The question was whether a satisfactory explanation requires

the existence of two interstitial species annihilating at quite different temperatures or needs only one interstitial species. Accordingly, quite different migration energies were assumed for vacancies and interstitials, respectively, by the advocates of the two different models. The reader is referred to the report on the panel discussion held at the Jülich Conference on Vacancies and Interstitials in Metals (see further reading: SEEGER *et al.* [1970]) in order to get the flavour of the atmosphere of that period. Many papers in the literature give a lively impression. In the meantime, a large number of experimental results were reported and even completely new methods were applied. They certainly narrowed the scope of the controversy considerably. It would be impossible to outline within this article the controversy in terms of *all* experimental results which were claimed at the time to be of relevance. Instead of this, the scientific content and today's state of the controversy is briefly outlined in a dedicated section (§ 3.3.2.5).

The application of alloy materials in nuclear reactors, particularly in the core region with its high neutron flux density, causes radiation damage to be a technological problem. This is, in particular, true for the fast breeder reactor and the future fusion reactor. Therefore, § 4 is dedicated to the basic mechanisms of point defect production and reactions under reactor irradiation conditions. As reactor irradiations are frequently simulated by heavy ion irradiation, the characteristics of this are treated as well. The interaction of swift heavy ions with metals has been included (§ 4.3) in order to complete the description of the interaction of heavy ions with metallic materials although this topic is not directly related to reactor technology. Important property changes such as swelling, radiation-induced creep and radiation-induced atom redistribution are outlined in §§ 4.4, 4.5 and 4.7. Irradiations under reactor conditions also demonstrated the pattern formation of vacancy clusters. This topic is treated in § 4.6.

2. *Vacancy properties*

2.1. Theoretical background

The entropy of a defect-containing crystal is larger than that of a perfect crystal. Therefore the Gibbs free energy change resulting from changing the atomic-defect concentration c by δc is given by:

$$\delta G = \left(\Delta H^{\mathrm{f}} - T\Delta S^{\mathrm{f}} + k_B T \ln c\right)\delta c, \tag{1}$$

where ΔH^{f} is the activation enthalpy of formation, $k_{\mathrm{B}} \ln c$ is the ideal entropy of mixing and ΔS^{f} is that entropy change which arises in addition to $k_{\mathrm{B}} \ln c$ (excess entropy of mixing), and k_{B} is the Boltzmann constant. In the close-packed metals, formation of vacancies requires the smallest amount of energy when compared with other lattice defects. Hence, they are dominant in thermal equilibrium.

The entropy change ΔS^{f} is mainly due to the change of the phonon spectrum of the crystal by the introduced defects. For the high-temperature harmonic approximation one obtains:

$$\Delta S_v^f = k_B \sum_i \ln(\omega_{0i}/\omega_i) \tag{2}$$

where ω_{0i} and ω_i are the eigenfrequencies of the crystal without and with vacancies, respectively. The phonon spectrum is changed by the change of the atomic coupling for the nearest neighbors of the vacancy and by the overall lattice volume change. The entropy change ΔS^f is positive if on average the ω_i are lowered and negative if they are enhanced.

The simplest model assumes nearest-neighbor interaction by a "spiral spring" with force constant f and calculates the change of the Einstein frequencies of the nearest neighbors of the vacancy by removing one of its twelve coupling springs (in the fcc lattice). For the vibration towards the vacancy, $\omega_{0E}^2 = 4f/M$ (M = atomic mass) is changed to $\omega_{0E}^2 = 3f/M$. Vibrations perpendicular to the removed spring are unchanged. For all other atoms all eigenfrequencies are unchanged. Hence we obtain $\Delta S_v^f = \frac{1}{2}(12\ k_B \ln 4/3) = 1.73\ k_B$.

Refinement of the calculation must take into account the static atomic relaxations around the vacancy (DEDERICHS and ZELLER [1980]). These change the force constants for many more atoms than the nearest neighbors. Correct consideration of this effect seems to be difficult. A negative sign of the relaxation contribution to the formation entropy change was obtained by SCHOTTKY *et al.* [1964] and a positive one by BURTON [1971]. In total, ΔS_v^f was obtained by SCHOTTKY *et al.* [1964] to be 0.49 k_B for Cu, 0.92 k_B for Ag and 1.22 k_B for Au, whereas BURTON [1971] obtained 1.8 k_B–2.0 k_B for fcc and 2.2 k_B - 2.4 k_B for bcc metals.

The sensitivity of the result to the number of atomic shells considered as contributing to the frequency redistribution has been investigated systematically for different pair-interaction potentials by HATCHER *et al.* [1979]. In the model crystal about 5×10^3 atoms around the vacancy were allowed to relax statically and for 102 to 5×10^2 atoms in the central region dynamical displacements (vibrations) were allowed. The pair interaction was simulated by the *Born–Mayer potential* introduced by GIBSON *et al.* [1960] for computer simulation of Cu (see also § 3.3.1) and by the *Morse potential* introduced by COTTERILL and DOYAMA [1966]. This last potential was smoothly cut-off at $r = 1.2a$, where a is the lattice-constant by HATCHER *et al.* [1979]. It then fits the observed lattice constant, bulk modulus and enthalpy of vacancy formation for Cu. The entropy of formation was found to depend approximately linearly on the inverse of the number of atoms allowed to vibrate. The relationship was followed up to the 19th atomic shell around the vacancy. By extrapolation, $\Delta S_{1v}^f = 2.3\ k_B$ was obtained for the Morse potential and 1.6 k_B for the Born–Mayer potential. The influence of the static lattice relaxation on the magnitude of ΔS_{1v}^f is demonstrated by comparing the above with 2.28 k_B for the Morse potential and 2.52 k_B for the Born–Mayer potential; both values are for the unrelaxed atoms (perfect lattice around the vacancy). HATCHER *et al.* [1979] have also found $\Delta S_{1v}^f = 2.1\ k_B$ for a potential simulating α-Fe. The cut-off Morse potential describes the formation and migration properties of mono- and di-vacancies in Cu reasonably well. The results are unsatisfactory for the Born–Mayer and the cited long-range Morse potential (COTTERILL and DOYAMA [1966]). For the hcp Mg, $\Delta S_{1v}^f = 1.5\ k_B$–2 k_B was found by using an empirical potential (MONTI and SAVINO [1981]).

Discrepancies between the results from different potentials are particularly large for the *relaxation volume* ΔV_{1v}^{rel}, i.e., the volume change of a crystal caused by removing one atom from the interior and withdrawing it from the crystal. The quantity is measured by the lattice-parameter change. The Born–Mayer potential for Cu yields –0.47 atomic volume, whereas the Morse potential yields –0.02 atomic volume for both the extended and the cut-off version (DEDERICHS *et al.* [1978]). With $\Delta H_{1v}^{f} = -0.41$ eV, the Born–Mayer potential fails to give a reasonable enthalpy of formation. With the corresponding values +1.17 eV and +1.29 eV for the extended and the cut-off Morse potential, respectively, satisfactory agreement with experimental values is obtained.

Ab-initio calculations of vacancy formation enthalpies are obviously problematic. Reviews are by FRIEDEL [1970], EVANS [1977], HEALD [1977], and STOTT [1978]. The problem consists of calculating the energetics of the electron system for a vacancy containing crystal. Only so-called simple metals have been treated. It had to be assumed that the core electrons are rigidly confined to the nucleus and any charge redistribution caused by the missing atom core only affects the conduction electrons. For transition metals the vacancy effect on the outermost *d* electrons has not been described successfully as yet. For the simple metals two entirely different approaches have been applied.

The first consists in replacing the vacancy by a repulsive impurity potential acting in a free electron gas (jellium). The second approach is based on pseudo-potential theory and assumes the vacancy effect to be weak enough to follow a linear response formalism. The results of both approaches are unsatisfactory in terms of agreement with experimental data. Reasons are the improper treatment of the effect of the local charge redistribution as well as of that of the lattice relaxation around a vacancy on the electron system. More promising seems to be the combination of selfconsistent cluster calculations (quantum-chemical approach) with lattice defect calculations (see ADAMS and FOILES [1990]).

Vacancy migration has been treated by analytical methods (FLYNN [1968]) as well as by computer simulation (BEELER [1983]). The model developed by FLYNN [1968] predicts activation enthalpies for the vacancy migration ΔH_{v}^{m} in remarkably good agreement with experimental data for quite a number of metals. The model relates ΔH_{v}^{m} to the elastic moduli. According to the author, the jumping atom passes the saddle point by means of a fluctuation of the kinetic energy which was picked up near the atom's equilibrium position where the kinetic energy takes maximum values. There, however, the atomic movement can be described by the harmonic approximation.

2.2. Experimental methods and results

2.2.1. Introductory remarks

Vacancies are of atomic size and therefore difficult to observe directly. However, agglomerates of numerous vacancies are visible even with a naked eye. Such agglomerates are formed in suitable diffusion samples owing to the *Kirkendall effect* (SMIGELSKAS and KIRKENDALL [1947]). Two metals A and B are brazed to each other and then thermally annealed at high temperatures. After sufficient atom transport, one observes that the two elements diffuse into each other, but with different velocities. As a conse-

quence, more atoms leave the fast diffuser crystal than slow diffusing atoms arrive there. The crystal responses to this migrational loss of matter by forming voids as shown in fig. 1. The quantitative evaluation of such experiments moreover prove that the migration of single vacancies is the origin of the atom diffusion (ch. 7, § 5.3.1.1).

Field ion microscopy does resolve the atomic structure of crystals and single vacancies have indeed been imaged (SEIDMAN [1973]). But these are vacancies at the surface or just beneath a surface and might not prove the existence of this defect within the bulk of a crystal.

The volume change of a crystal due to vacancy formation (§ 2.2.2.2) seems to be the most direct measurable change caused by single vacancies. Other properties are treated in §§ 2.2.2.3 and 2.2.2.4.

2.2.2. Enthalpy and entropy of formation

2.2.2.1. Single vacancies and di-vacancies. For thermal equilibrium one obtains from eq. (1):

$$c_v^0 = \exp\left(\Delta S_v^f / k_B\right) \exp\left(-\Delta H_v^f / k_B T\right). \tag{3}$$

The determination of enthalpy and entropy of formation according to this equation requires the determination of equilibrium vacancy concentrations c_v^0. Two different procedures are applied: either measurement of c_v at high temperatures such that c_v is indeed in thermal equilibrium during the measurement, or quenching of the sample from high temperatures to low temperatures such that the defects are immobile. The latter method avoids difficult measurements at high temperatures but requires considerable efforts to avoid vacancy losses during quenching. Review articles are by HOCH [1970], BALLUFFI *et al.* [1970] and SIEGEL [1978].

Even the most careful measurement of c_v^0 at high temperatures cannot yield the wanted formation enthalpy ΔH_{1v}^f and entropy ΔS_{1v}^f of the monovacancy because c_v^0 does

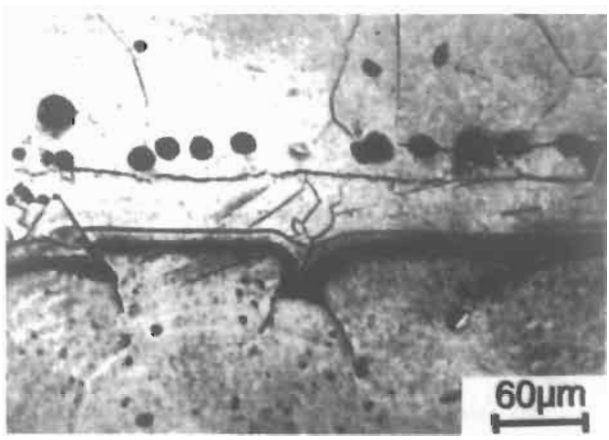


Fig. 1. Cross section of a diffusion couple Cu (above) against Ni (below). Thin black line represents the Cu–Ni interface before the diffusion annealing. Black dots near the interface line are cavity images.

not necessarily involve only monovacancies. Actually it is given by $c_v^0 = c_{1v}^0 + 2c_{2v}^0 + \ldots = \sum_{n=1}^{N} n c_{nv}^0$, where c_{nv}^0 is the equilibrium concentration of clusters consisting of n vacancies. For equilibrium concentration measurements, higher aggregates than di-vacancies need not be considered owing to their negligibly small concentration in the normal cubic metals. We then have in the dilute solution approximation:

$$c_v^0 = c_{1v}^0 + 2c_{2v}^0 = \exp\left(\Delta S_{1v}^f / k_B\right) \exp\left(-\Delta H_{1v}^f / k_B T\right) + 2g_{2v} \exp\left[\left(2\Delta S_{1v}^f - \Delta S_{2v}^b\right)/k_B\right] \exp\left[\left(\Delta H_{2v}^b - 2\Delta H_{1v}^f\right)/k_B T\right], \tag{4}$$

where ΔS_{2v}^b and ΔH_{2v}^b are entropy and enthalpy of binding of the di-vacancy, respectively. For the di-vacancy formed by two vacancies at nearest-neighbor sites in the fcc structure, the ageometry factor becomes $g_{2v} = 6$. In bcc and hcp structures different di-vacancy configurations may exist simultaneously with non-degenerate values of $g_{2v}, \Delta S_{2v}^b$ and ΔH_{2v}^b. Up to now, there is no reliable method of measurement which separates c_{2v}^0 and c_{1v}^0. Hence, the formation enthalpy as derived from the temperature dependence of c_v^0 is an effective one:

$$\Delta H_v^f = -\frac{\partial \ln c_v^0(T)}{\partial\left(1/k_B T\right)} = \Delta H_{1v}^f + 2g_{2v} \frac{\left(\Delta H_{1v}^f - \Delta H_{2v}^b\right) \exp\left[\left(G_{2v}^b - G_{1v}^f\right)/k_B T\right]}{1 + g_{2v} \exp\left[\left(G_{2v}^b - G_{1v}^f\right)/k_B T\right]} \tag{5}$$

where G denotes the Gibbs free energy. For bcc and hcp metals, eq. (5) becomes much more complicated due to the possible occurrence of different di-vacancy configurations (SEEGER [1973a]). Equation (5) introduces a temperature dependence of ΔH_v^f which must be considered whenever ΔH_v^f is determined over a large temperature range. Binding enthalpies of the di-vacancy have been calculated for Cu, Ni, Au, and Pt to amount to 0.08, 0.07, 0.08, and 0.11 eV, respectively (KLEMRADT *et al.* [1991], § 2.2.5). Such values would not lead to measurable deviations from the Arrhenius behavior.

2.2.2.2. Differential dilatometry. The classical method of direct c_v^0 determination is *differential dilatometry* (WAGNER and BEYER [1936], SIMMONS and BALLUFFI [1960a, b, 1962, 1963]). It is based on a theory by ESHELBY [1956], which relates the microscopic volume change ΔV^{rel} of randomly distributed dilatation centers (point defects) to the total volume change of the defectcontaining crystal as compared to the perfect crystal. The total crystal volume change consists of the Eshelby contribution plus that caused by adding to the total crystal volume the atomic volumes of those atoms which are taken out of the interior of the crystal in order to produce vacancies. The crystal volume change can be measured by the length change, $\Delta l/l$, of the sample for cubic crystals and the average microscopic (lattice cell) change, $\Delta a/a$, by the lattice parameter change of the same sample. One obtains:

$$c_v = 3\left(\Delta l/l - \Delta a/a\right), \tag{6}$$

which is correct for cubic crystals and the small vacancy concentrations, usual in such metals.

Precise measurements of this type are extremely difficult. For illustration of the orders of magnitude to be measured, the Al data by SIMMONS and BALLUFFI [1960a] are shown in fig. 2. From these data the authors derived $\Delta H_v^f = 0.76$ eV. Since then ΔH_v^f has frequently been determined also by means of the methods of measurement described below with the result of values ranging from 0.60 eV to 0.77 eV with 0.67 ± 0.03 eV as recommended by BALLUFFI [1978]. Recommended values for the other metals are given in table 1. A complete collection of data was given by EHRHART [1991] and SCHULTZ [199la].

The relaxation volume is given by $\Delta V_v^{rel}/\Omega = 3\Delta a/c_v^0$ (where Ω is the atomic volume). The diffuse X-ray scattering cross section (DXS, § 3.3.2.2.) is given by the product of c_v and the square of the scattering amplitude caused per vacancy (DEDERICHS [1973]) for low vacancy concentrations and random spatial distribution. The scattering amplitude is essentially given by the elastic moduli and the dipole force tensor which describes the atomic displacements resulting from the insertion of the vacancy. As ΔV_v^{rel} is determined by the same quantities and can be derived from $\Delta a/a$ measurements as shown above, the combination of DXS and $\Delta a/a$ measurements allows the evaluation of both, c_v and ΔV_v^{rel} (EHRHART *et al.* [1979]). The DXS — $\Delta a/a$ measurements yielded $\Delta V_v^{rel} = -0.15\ \Omega$ for gold when measured at 4 K. The DD method yielded $-0.5\ \Omega$ instead (SIMMONS and BALLUFFI [1962]). EHRHART *et al.* [1979] trace this difference back to a potential temperature dependence of ΔV_v^{rel}. For other metals, ΔV_v^{rel} data are given in table 1. (See also ch. 12, § 3.3).

The third method determines the volume of vacancy formation $\Delta V_v^f = \Delta V_v^{rel} + \Omega$ by measuring the quenched resistivity increment (as a measure for c_v^0 at the quenching temperature T_q) as a function of temperature and hydrostatic pressure. The order of magnitude of the pressure dependence is such that for gold, a 6 kbar pressure increase corresponds to a temperature decrease of about 30 K around 900 K for constant c_v^0.

The way of determining ΔV_v^{rel} from $(\Delta a/a)_q$ and $\Delta\rho_q$ measurements on quenched

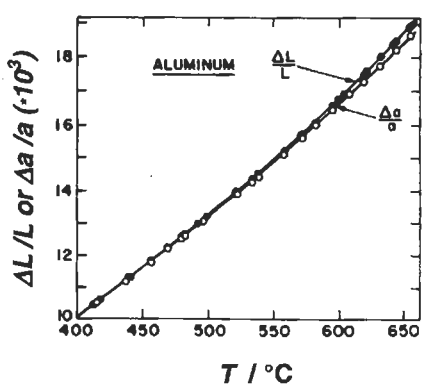


Fig. 2. Differential dilatometry in vacancy equilibrium conditions for aluminium. $\Delta L/L$ is the relative length change of a sample, $\Delta a/a$ is the relative lattice-constant change. (After SIMMONS and BALLUFFI [1960a].)

Table 1
Properties of vacancies in pure metal.

Glossary: Property symbol	Definition and units
$c_v(T_M)$	Vacancy concentration at the melting temperature (x 10^4)
ΔS_v^f	Activation entropy of formation (units of k_B), referring either to single vacancy values or to effective values depending on data evaluation
ΔH_v^f	Activation enthalpy of formation (in eV) referring to single-vacancy or effective values, depending on data evaluation
ΔV_v^{rel}	Relaxation volume of the vacancy (in atomic-volume units)
ρ_v	Electric resistivity contribution per unit concentration of vacancies (in units of 10^{-4} Ωcm)
ΔH_v^m	Activation enthalpy of migration (in eV)
Measurement symbol	**Method of measurement for data evaluation**
DD	Differential dilatometry
$\Delta\rho_q$, $\Delta\rho_{irr}$, $\Delta\rho_{cw}$	Electrical resistivity of quenched, irradiated and cold worked samples respectively
TEM	Transmission electron microscopy
PAS	Positron-annihilation spectroscopy
DXS	Diffuse X-ray scattering
NMR	Nuclear magnetic resonance
μSR	Muon spin rotation experiments
THDS	Thermal helium desorption spectroscopy

Property	Value	Measurement	Reference
Ag (silver, fcc)			
$c_v(T_M)$	1.7	DD	SIMMONS and BALLUFFI [1960b]
ΔS_v^f	1.5	DD	SIMMONS and BALLUFFI [1960b]
ΔH_v^f	1.11 ± 0.05		SIEGEL [1982a]
ρ_v	1.3	$\Delta\rho_q$	DOYAMA and KOEHLER [1962]
ΔH_v^M	0.66 ± 0.05		BALLUFFI [1978]
Al (aluminium, fcc)			
$c_v(T_M)$	9.4	DD, $\Delta\rho_q$	SIEGEL [1982b]
ΔS_v^f	0.7	DD, $\Delta\rho_q$	SIEGEL [1978]
ΔH_v^f	0.67 ± 0.03		BALLUFFI [1978] SIEGEL [1978] SCHAEFER *et al.* [1987a]
ΔV_v^{rel}	–0.05	DXS	EHRHART *et al.* [1986]
ρ_v	1.0		SIEGEL [1978]
ΔH_v^M	0.61 ± 0.03		SIEGEL [1982b]
Au (gold, fcc)			
$c_v(T_M)$	7.2	DD	SIMMONS and BALLUFFI [1962]
ΔS_v^f	0.72		SIEGEL [1978]
ΔH_v^f	0.93 ± 0.04		SIEGEL [1982a]
ΔV_v^{rel}	–0.15	$\Delta\rho_q$, HDS, $\Delta a/a$, valid at 4 K	EHRHART *et al.* [1979]
ρ_v	1.70	$\Delta\rho_q$	EHRHART *et al.* [1979]
ΔH_v^M	0.71 ± 0.05		SONNENBERG and DEDEK [1982]

References: p. 1710.

Property	Value	Measurement	Reference
Be (beryllium, hcp)			
ΔH_v^M	0.8	$\Delta\rho_{irr}$	SCHUMACHER [1970]
Bi (bismuth, rhombohedral)			
$c_v(T_M)$	6.17	DD	MATSUNO [1977]
ΔS_v^f	0.3	DD	MATSUNO [1977]
ΔH_v^f	0.35	DD	MATSUNO [1977]
ΔH_v^M	0.5	PAS, $\Delta\rho_{irr,q}$	CORBEL *et al.* [1987]
Cd (cadmium, hcp)			
$c_v(T_M)$	4.5	DD	JANOT and GEORGE [1975]
ΔS_v^f	0.5	DD	JANOT and GEORGE [1975]
ΔH_v^f	0.46 ± 0.05	PAS	EHRHART [1991]
ΔV_v^{rel}	–0.54	PAS, pressure dependence	DICKMAN *et al.* [1978]
ΔH_v^M	0.4 ± 0.05		EHRHART [1991]
Co (cobalt, hcp)			
ΔV_v^{rel}	–0.05	DXS, valid at 4 K	EHRHART and SCHÖNFELD [1982]
ΔH_v^M	0.72	$\Delta\rho_{irr}$	DANDER and SCHAEFER [1977]
Cr (chromium, bcc)			
ΔH_v^f	2.0 ± 0.2	PAS	LOPER *et al.* [1985]
ΔH_v^M	–0.95		SCHULTZ [1991]
Cu (copper, fcc)			
$c_v(T_M)$	2.1 ± 0.1	DD	DIFFERT *et al.* [1987]
ΔS_v^f	2.8	$\Delta\rho_q$	BERGER *et al.* [1979]
	1.6	DD	DIFFERT *et al.* [1987]
ΔH_v^f	1.28 ± 0.05		SCHAEFER *et al.* [1987b]
ΔV_v^{rel}	–0.25	DXS, low temperature value	EHRHART *et al.* [1986]
ρ_v	0.62	$\Delta\rho_q$, DD	BERGER *et al.* [1979]
ΔH_v^M	0.70 ± 0.02		BALLUFFI [1978]
Fe (iron, bcc)			
ΔV_v^{rel}	–0.05	$\Delta a/a$, $\Delta\pi$ DXS	EHRHART [1985]
ΔH_v^f (para)	1.79 ± 0.1	PAS	DE SCHEPPER *et al.* [1983]
	1.85 ± 0.1	PAS	SHIRAI *et al.* [1989]
ΔH_v^f (ferro)	1.59 – 1.73;	μSR, ferromagnetic state	FURDERER *et al.* [1987]
ΔH_v^M (ferro)	0.55		TAKAKI *et al.* [1983]
	1.11 ± 0.2	Q^{SD}–H_v^f	SHIRARI *et al.* [1989]
In (indium, tetragonal)			
ΔS_v^f	5	PAS	WEILER and SCHAEFER [1985]
ΔH_v^f	0.52 ± 0.04		EHRHART [1991]
ΔV_v^{rel}	–0.6	PAS	DICKMAN *et al.* [1977]
ΔH_v^M	(0.31)	Q_V^{SD}–H_V^F	WEILER and SCHAEFER [1985]
K (potassium, bcc)			
ΔH_v^f	0.34	Q_v^{SD}–ΔH_v^M	WALLNER *et al.* [1986]
ΔH_v^M	0.038	$\Delta\rho_{irr}$	WALLNER *et al.* [1986]
Li (lithium, bcc)			
ΔH_v^f	0.48	Q_v^{SD}–ΔH_v^M	SCHULTZ [1991]
ΔH_v^M	0.038		WALLNER [1983]

Property	Value	Measurement	Reference
Mg (magnesium, hcp)			
$c_v(T_M)$	7.2	DD	JANOT *et al.* [1970]
ΔS_v^f	0 ± 0.3	DD	JANOT *et al.* [1970]
ΔH_v^f	0.79	$\Delta\rho_q$	TZANETAKIS *et al.* [1976]
ΔH_v^M	0.5–0.6	$\Delta\rho_q$	TZANETAKIS *et al.* [1976]
Mo (molybdenum, bcc)			
$c_v(T_M)$	0.13		SCHULTZ [1991]
ΔS_v^f	1.6		SCHULTZ [1991]
ΔH_v^f	3.0 ± 0.2	PAS	SCHAEFER [1987]
ΔV_v^{rel}	–0.1	DXS	EHRHART [1978]
ρ_v	4.3		SCHWIRTLICH and SCHULTZ [1980]
ΔH_v^M	1.35	$\Delta\rho$	SCHWIRTLICH and SCHULTZ [1980]
Na (sodium, bcc)			
$c_v(T_M)$	7.8 ± 0.3		ADLHART *et al.* [1975]
ΔS_v^f	3.9 ± 0.3		ADLHART *et al.* [1975]
ΔH_v^f	0.354 ± 0.035	$\Delta a a$, $\Delta l/l$	ADLHART *et al.* [1975]
ΔH_v^M	0.03		WALLNER [1983]
Nb (niobium, bcc)			
ΔH_v^f	3.07	Q_v^{SD}–ΔH_v^M	SCHULTZ [1991]
ΔH_v^M	0.55		FABER and SCHULTZ [1977]
Ni (nickel, fcc)			
ΔH_v^f	1.79 ± 0.05		SCHAEFER [1982]
ΔV_v^{rel}	–0.2	DXS	BENDER and EHRHART [1983]
ρ_v	2.3–3.6	$\Delta\rho_q$	WYCISK and FELLER–KNIEPMEIER [1978]
ΔH_v^M	1.04 ± 0.04		SEEGER and KRONMÜLLER [1987]
Pb (lead, fcc)			
$c_v(T_M)$	1.7	DD	FEDER and NOWICK [1967]
Δsvf	2.6	$\Delta\rho_q$, DD	LEADBETTER *et al.* [1966]
	0.7 ± 2.0	DD	FEDER and NOWICK [1967]
ΔH_v^f	0.58 ± 0.04	PAS	ITO *et al.* [1982], HARA and NANAO [1982]
ρ_v	2.8	$\Delta\rho_q$	KNODLE and KOEHLER [1978]
ΔH_v^M	0.43 ± 0.02	$\Delta\rho_{irr}$	SCHROEDER and SCHILLING [1976]
Pd (palladium, fcc)			
$c_v(T_M)$	0.14	PAS	SCHAEFER [1982]
ΔH_v^f	1.85	PAS	SCHAEFER [1982]
ΔH_v^M	1.03 ± 0.3	$\Delta\rho_{cw}$	KÖSTER and KEHRER [1965]
Pt (platinum, fcc)			
$c_v(T_M)$	9.4	$\Delta\rho_q$	KHELLAF *et al.* [1988]
ΔS_v^f	0.4	$\Delta\rho_q$	KHELLAF *et al.* [1988]
ΔH_v^f	1.35 ± 0.05		EHRHART [1991]
ΔV_v^{rel}	–0.28	$\Delta\rho_q$	EMRICK [1982]
ρ_v	5.75	FIM, $\Delta\rho_q$, upper limit	BERGER *et al.* [1973]
ΔH_v^M	1.43 ± 0.05		BALLUFFI [1978]
Re (rhenium, hcp)			
ΔH_v^M	2.2	$\Delta\rho_{cw}$	VANDENBORRE *et al.* [1968]
	0.5–0.6	$\Delta\rho_{irr}$	VANDENBORRE *et al.* [1972]

References: p. 1710.

Property	Value	Measurement	Reference
Sb (antimon, rhombohedral)			
$c_v(T_M)$	10		DE DIEGO *et al.* [1987]
ΔH_v^f	1.4	Q^{SD}–$H^M{}_v$	EHRHART [1991]
ΔH_v^M	~0.3	PAS	DE DIEGO *et al.* [1987]
Sn (tin; β-Sn (for T ≥ 13°C): tetragonal, bct (A5))			
ΔS_v^f	1.1	DD, PAS	BALZER and SIGVALDASON [1979]
$c_v(T_M)$	≤ 0.3	DD	BALZER and SIGVALDASON [1979]
ΔH_v^f	0.51	PAS	SHAH and CATZ [1984]
ΔH_v^M	0.3–0.4		MCILWAIN *et al.* [1975]
Ta (tantalum, bcc)			
ΔH_v^f	3.1	Q_v^{SD}–ΔH_v^M	SCHULTZ [1991]
ΔH_v^M	0.7		FABER *et al.* [1974]
Th (thorium, fcc)			
ΔH_v^f	1.28 ± 0.2	PAS	KIM *et al.* [1984]
ΔH_v^M	2.04	Q^{SD}–H_v^f	KIM *et al.* [1984]
V (vanadium, bcc)			
ΔH_v^f	2.2 ± 0.4	PAS, life time	JANOT *et al.* [1982]
ΔH_v^M	0.7		SCHULTZ [1991]
	0.5	Q_v^{SD}–ΔH_v^M, PAS	SCHULTZ [1991]
W (tungsten, bcc)			
ρ_v	7		PARK *et al.* [1983]
$c_v(T_M)$	3		PARK *et al.* [1983]
ΔS_v^f	3.2		PARK *et al.* [1983]
ΔH_v^f	3.6 ± 0.2		PARK *et al.* [1983]
ΔH_v^M	1.70 ± 0.1	valid for low T	MUNDY *et al.* [1987]
Zn (zinc, hcp)			
$c_v(T_M)$	5 ± 0.3	DD	BALZER and SIGVALDASON [1979]
ΔS_v^f	1 ± 1	DD	BALZER and SIGVALDASON [1979]
ΔH_v^f	0.54 ± 0.03		EHRHART [1991]
ΔV_v^{rel}	–0.6	DXS	EHRHART and SCHONFELD [1979]
ΔH_v^M	0.42 ± 0.02		EHRHART [1991]
Zr (zirconium, hcp)			
ΔH_v^f	> 1.5	PAS	HOOD [1986]
ΔV_v^{rel}	~0.05	DXS, valid at 4 K	EHRHART and SCHONFELD [1982], EHRHART *et al.* [1986]
ΔH_v^M	0.54–0.62	$\Delta\rho_{irr}$	NEELY 11970]
	0.6–0.7	PAS_{irr}	HOOD *et al.* [1984], HOOD and SCHULTZ [1988]

samples must be based upon well known ρ_v values. As these are obtained by use of absolute c_v^0 data from DD measurements this method goes back essentially to DD data but does involve all quenching problems.

For the bcc metals ΔV_v^{rel} values do not exist because of difficulties in determining absolute c_v^0 values (SCHULTZ [1991]). For hcp metals, similar values have been reported as for the above-quoted fcc metals.

2.2.2.3. Positron-annihilation spectroscopy. The second common method of measuring vacancy concentrations in thermal equilibrium is *positron-annihilation spectroscopy* (PAS). Review articles on PAS application for studies of vacancy properties are by SEEGER [1973b], DOYAMA and HASIGUTI [1973], TRIFTSHÄUSER [1975], SIEGEL [1978], MIJNARENDS [1979], WEST [1979] and HAUTOJÄRVI [1987]. High-energy positrons injected into metal crystals are rapidly thermalized by electron-hole excitations and interactions with phonons. The thermalized positron diffuses through the lattice and ends its life by annihilation with an electron. The lifetime depends on the total electron density occurring along the diffusional path of the positron. Vacancies obviously trap positrons in a bound state, and because of the missing core electrons at the vacant lattice site, the local electron density is significantly reduced. This condition causes the lifetime of trapped positrons to be enhanced by 20–80% as compared to that of free positrons in the perfect lattice. Consequently, positrons in a vacancy-containing crystal end their lifes by annihilation either as free positrons or as trapped positrons. Critical discussion on non-thermal trapping was given by KLUIN and HEHENKAMP [1991]. The lifetimes for both fates are different and the probability of trapping is proportional to the vacancy concentration. Lifetime measurements are possible as γ-quanta are emitted at the birth of a positron as well as at its decay. Fortunately, thermalization happens within about one picosecond whereas the average lifetime in the metal crystal is in the order of 200 ps.

With PAS, information is extracted not only from the positron lifetimes but also from the angular correlation between the directions of the emitted annihilation γ-rays and from the Doppler broadening of these γ-rays. These quantities yield information on the net momenta of the annihilating electron-positron pairs. They allow a distinction to be made between annihilations with the higher-momentum core electrons and the lower-momentum valence or conduction electrons.

Vacancy concentrations are determined by means of lifetime spectrum measurements as well as momentum techniques (angular correlation and Doppler broadening). The former avoids additional assumptions on positron-annihilation parameters but requires high-resolution measuring techniques and expanded data deconvolution. Although the deconvolution of momentum-distribution data requires critical assumptions on the temperature dependence of positron-annihilation parameters, these methods have become more popular for ΔH_v^f determinations. In fig. 3, vacancy concentrations derived from angular correlation measurements are shown for Cu and Au. The range of measurement by differential dilatometry is indicated for comparison. The PAS data extend to about two orders of magnitude lower vacancy concentration than the differential dilatometry data. As monovacancies certainly predominate in this concentration range, PAS studies are of great importance as a complement to differential dilatometry on the one side, and to resistivity measurements in quenched samples on the other side.

References: p. 1710.

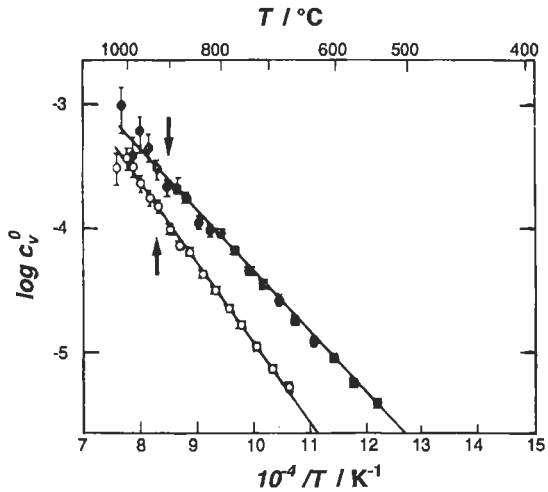


Fig. 3. Arrhenius plot of the vacancy concentration as derived from positron annihilation spectroscopy for gold (full symbols) and copper (open symbols) according to TRIFTSHÄUSER and MCGERVEY [1975]. The left hand side of the arrows represents the range covered by the differential dilatometry.

2.2.2.4. Resistivity measurements after quenching. All problems of measurements at temperatures near the melting point can be avoided if equilibrium vacancy concentrations can quantitatively be quenched from such temperatures T_q to temperatures which immobilize the vacancies. Properties which are proportional to the vacancy concentration could then be measured on quenched and unquenched samples for direct comparison. This idea appeared so challenging that large efforts were put into the development of quenching methods following the pioneering work by KAUFFMAN and KOEHLER [1952]. Methods and problems are reviewed by BALLUFFI *et al.* [1970] and SIEGEL [1978]. There are indeed problems inherent to the quenching process. During quenching, the vacancies are still highly mobile in a significant part of the total temperature interval being passed through. The migrating vacancies are able to react with other defects or with one another, with the following consequences: (i) Vacancy losses to sinks such as dislocations, grain boundaries and surfaces. The quenched concentration c_v will therefore be smaller than $c_v^0(T_q)$. The inhomogeneous distributions of sinks will cause locally varying concentrations. (ii) Vacancy clustering, which causes repartitioning of the cluster size distribution existing at T_q. High-order clusters are favored in comparison to the equilibrium distribution at T_q.

By modelling the vacancy reaction scheme for the conditions of quenching, vacancy loss and repartitioning of cluster sizes were studied in great detail (BALLUFFI *et al.* [1970]). With the aid of such calculations, quenching results obtained for systematically varied quenching rates could be corrected and extrapolated to infinite quenching rates in the case of Au (FLYNN *et al.* [1965]). Systematic studies of this kind were performed also on Al (BASS [1967] and BERGER *et al.* [1978]).

An entirely different approach was successfully followed by LENGELER [1976] and

LENGELER and BOURASSA [1976]. They were able to rapidly quench single crystals of Au and Cu with dislocation densities so low that vacancy losses are negligible even for high T_q values. The effect is demonstrated in fig. 4. Quenching data were reviewed by BALLUFFI [1978].

2.2.3. Activation enthalpy of migration

2.2.3.1. Problems of methods of determination. The common way of investigating the migration properties of vacancies consists of quenching a sample from high temperatures and subsequent annealing at increasing temperatures in order to induce the diffusion-controlled annihilation of the excess vacancies at the annealing temperature (further reading: Seeger *et al.* [1970]). The residual resistivity has mostly been taken as a measure of the vacancy concentration in the lattice. This property is certainly preferable to all others in view of its convenient application and high sensitivity. The isochronal recovery behavior of a pure metal containing point defects is sketched in fig. 5. The resistivity increase recovers in distinct steps which are labelled according to VAN BUEREN [1955]. Stage I is observed in irradiated samples only. Stage II occurs in irradiated and plastically deformed samples. Quenched samples show resistivity recovery at temperatures varying from stage III to stage IV, depending on the individual metal. Stage V is caused by recrystallization (ch. 28) and occurs in all samples which contain larger defect agglomerates. As the first recovery stage of quenched samples does not

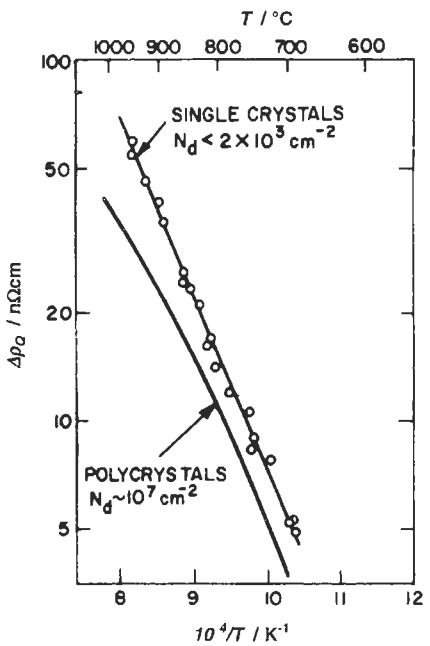


Fig. 4. Arrhenius plot of the quenched-in residual resistivity of gold. The influence of the dislocation density N_d is shown by the curve for quenched polycrystals. (After LENGELER [1976].)

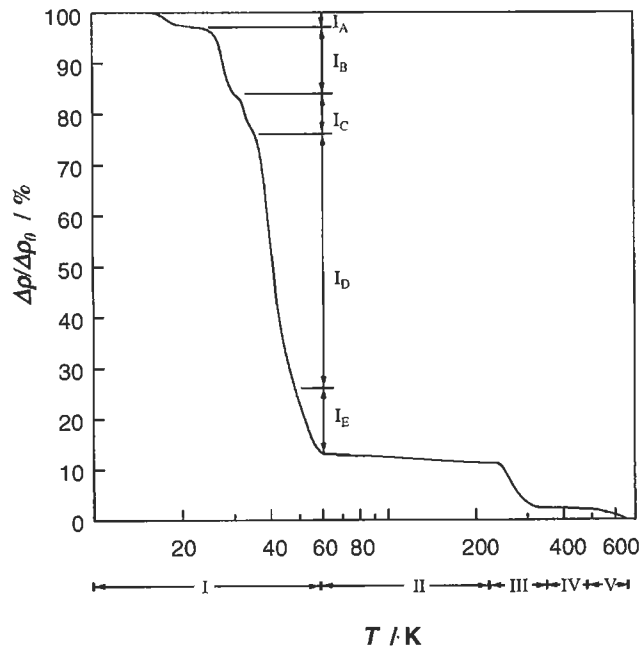


Fig. 5. Isochronal recovery curve of the electrical resistivity of Cu electron irradiated at 4 K.

generally coincide with stage III of irradiated samples, a straightforward classification of this stage was not possible. This situation created the *stage III controversy* (§ 3.3.2.5).

The resistivity method is not appropriate for deducing information on details of the underlying defect reactions. Its inability to indicate secondary vacancy reactions like clustering and trapping at impurities beside annihilation is one of the origins of the stage III controversy. It was not before the mid seventies that other methods such as PAS or *perturbed* γ–γ *angular correlation* (PAC) came up, which allowed a quantitative separation of different vacancy reactions.

A breakthrough was achieved by applying PAS in order to determine the temperature at which vacancies become mobile. Another significant step was taken when the ability of nuclear probe atom methods such as Mössbauer spectroscopy and PAC to distinguish between interstitials and vacancies was discovered.

As many as 21 experimental results are listed by BALLUFFI [1978] which favor stage III interpretation by vacancy migration. Since then methods such as the perturbed γ-γ angular correlation measurements have been applied fruitfully. Their results also favor the vacancy interpretation (see, for example, PLEITER and HOHENEMSER [1982]).

2.2.3.2. Two selected pieces of evidence for vacancy migration in stage III. Among the earlier results, the positron-annihilation behavior in stage III seems especially conclusive with respect to vacancy migration in that stage. The important quantity is the so-called *lineshape parameter* $R = |(I_v^t - I_v^f)/(I_c^t - I_c^f)|$ The energy of the positron-annihilation γ quanta is Doppler–broadened because of the non-vanishing momenta of the

annihilating electron-positron pairs. The center of the curve (intensity versus energy) is caused by low-momentum pairs, i.e., involves conduction and valence (v) electrons, whereas the branches arise from high-momentum pairs, i.e., involve core (c) electrons. The quantity I is simply a suitably broad section of the Doppler line integral, one below the maximum (I_v) and the other below the branches (I_c). The upper subscripts t and f refer to trapped and free positrons, respectively, as obtained from the irradiated and unirradiated state.

The lineshape parameter R was found to be considerably larger for positrons trapped by voids than for those trapped by single vacancies (MANTL and TRIFTSHÄUSER [1978]). For trapping by dislocation loops its magnitude is close to that for single vacancy trapping. The result from measurements in electron-irradiated Cu through stage III recovery is shown in fig. 6. The corresponding resistivity-decrease between 200 K and 300 K amounted to about 85% of its value at 200 K. The increase of R indicates a further fractional decrease of core electron density as compared to the presence of merely single vacancies below 200 K. This can only be visualized by vacancy agglomeration in three-dimensional arrangements. The effect of interstitial-type dislocation loops which could be formed in stage III according to the two-interstitial model was found to produce $R = 0.62 \pm 0.03$ in neutron-irradiated aluminium (GAUSTER *et al.* [1975]). Interstitial agglomeration can therefore be excluded as an origin of the observed R enhancement. MANTL and TRIFTSHÄUSER [1978] assume the three-dimensional agglomerates to be the nuclei of the vacancy loops observed at higher temperatures by electron microscopy. The same behavior of R as in Cu has been observed in the bcc Mo whereas in Al the observed increase in R did not significantly exceed the error limit (MANTL and TRIFTSHÄUSER [1978]).

Fortunately, the PAC technique indicated the arrival of a vacancy type defect in stage III at the probe atom Cd in Al (RINNEBERG *et al.* [1978], RINNEBERG and HAAS [1978]). The PAC technique essentially measures the hyperfine interaction of the probe-atom nucleus with the electric crystal field. Interstitials or vacancies situated at the nearest-neighbor sites of such probe atoms obviously cause sufficiently large field gradients to produce measurable signals. Review articles on this type of application of PAC are by PLEITER and HOHENEMSER [1982], WICHERT [1982] and NIESEN [1981]. The PAC signals are determined by the electric field gradient at the site of the probe-atom nucleus and are

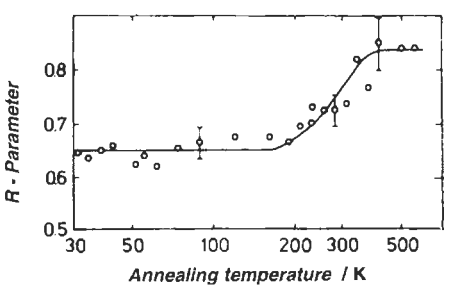


Fig. 6. Lineshape parameter R of low-temperature electron-irradiated Cu upon isochronal annealing (after MANTL and TRIFTSHÄUSER [1978]).

References: p. 1710.

in that sense specific for the type of defect which is placed at a nearest-neighbor site. If it happens that the probe atom in a given host metal traps the vacancy, a quenched sample containing probe atoms will exhibit the PAC signal caused by the vacancy. Similarly one obtains the signal caused by the interstitial when one irradiates the sample at low temperatures and anneals through stage I or irradiates in stage II. In this way one obtains the "fingerprints" of vacancy and interstitial. Clear evidence for the arrival of vacancies at the probe atoms in stage III has been obtained for Al, Ag (BUTT *et al.* [1979], DEICHER *et al.* [1981a]), Au (DEICHER *et al.* [1981b]), Cu (WICHERT *et al.* [1978]), Pt (MÜLLER [1979]), Cd and Zn (SEEBOECK *et al.* [1982]).

The PAC signal also gives information on the deviation of the electric field-gradient tensor from axial symmetry. When measurements are performed in single crystals, the tensor orientation with respect to the crystal orientation can be obtained. Knowledge of orientation and deviation from axial symmetry provides valuable criteria for answering the question whether mono-, di-, or tri-vacancies decorate the probe atom and cause the observed signal. Indeed, the PAC technique has resolved a number of different configurations for the metals quoted above (review by RECKNAGEL *et al.* [1983]). The variety of defect arrangements trapped at In in Au after different treatments is illustrated in fig. 7. From the field-gradient tensor properties, it has been concluded that defect 3 is a monovacancy at a nearest-neighbor site, defect 4 is a planar <111> vacancy loop and defects 1 and 2 are multiple vacancies in different arrangements (DEICHER *et al.* [1981b]). It should be kept in mind that the spectrum of visible vacancy arrangements does depend on the nature of the probe atom because of the required trap property. Nevertheless, the variety of arrangements to be seen in fig. 7 gives a good impression of the complexity of the stage III processes which, in the earlier controversy, were often identified with either monovacancy or monointerstitial reactions.

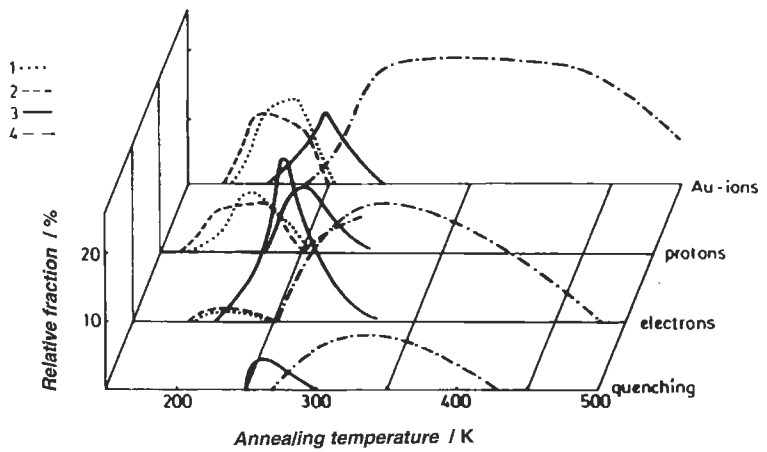


Fig. 7. Fraction of In atoms in gold which emit perturbed γ–γ angular correlation signals upon isochronal annealing after quenching and low-temperature irradiation with the particles indicated. The different types of lines indicate different types of signals and hence, different defects trapped by the probe atoms (after DEICHER *et al.* [1981b]).

2.2.3.3. Experimental determination of ΔH_v^m. The primary problem of experiments involving quenching and subsequent annealing consists in the large variety of defect reactions which are likely to occur simultaneously. In addition, the boundary conditions for these diffusion-controlled reactions are significantly influenced by the quenching and annealing procedure. The following conditions are important: (i) locally inhomogeneous distribution of vacancies, due to losses at sinks or cluster formation (§ 2.2.4), (ii) time-dependent sink concentration and population of different configurations during the annealing process, e.g., by nucleation and growth of vacancy clusters, (iii) sink efficiencies which depend on both, vacancy concentration and temperature, (iv) vacancy-impurity interactions, (v) drift diffusion effects caused by elastic interaction of vacancies and sinks. Computer simulation by means of the rate equation approach shows that the main defect parameters can hardly be inferred from a standard analysis of isothermal and isochronal data if, for example, the actual clustering kinetics are unknown (JOHNSON [1968]).

The experimentally observed influence of clustering is illustrated in fig. 8 which shows for Au and Al the observed temperatures T_a of annealing stages as function of the quenching temperature T_q. Within the temperature range shown for T_a, vacancies either annihilate at fixed sinks or form immobile clusters which dissociate at considerably higher temperatures than 450 K. The behavior of T_a for Al is illustrated in more detail in fig. 9 which shows the temperature-differentiated isochronal recovery curve for the T_q's indicated. Peak *B* was quantitatively interpreted by a single diffusion-controlled annihilation process with an activation enthalpy of 0.65 eV, which was ascribed to monovacancy migration. Peak *A* could not be interpreted by a single process and is characterized by an effective activation enthalpy of 0.44–0.5 eV. Obviously, at low vacancy concentrations as quenched from low temperatures, monovacancies are the dominant defects. For higher vacancy concentrations as quenched from temperatures above 500°C, multiple vacancies are formed which migrate faster than monovacancies and accelerate the annihilation process. This feature describes, in general, the annihilation behavior of quenched-in vacancies in most metals, although the separation between the processes involving

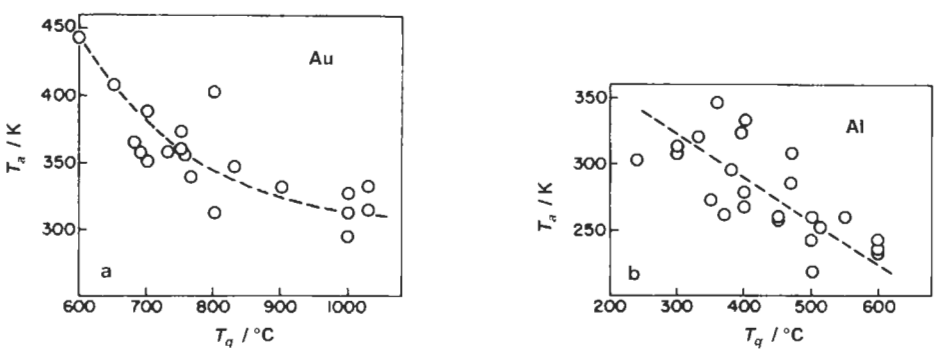


Fig. 8. Dependence of the temperature T_a, characterizing the temperatures of vacancy-annihilation stages in isochronal annealing experiments on quenching temperature T_q for Au and Al (after BALLUFFI [1978]).

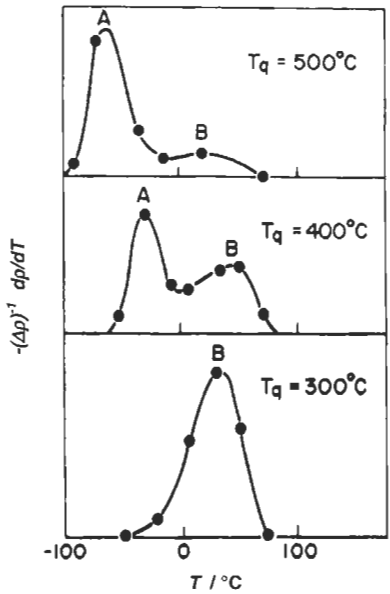


Fig. 9. Temperature-differentiated isochronal resistivity recovery curve of Al for different quenching temperatures T_q (after LEVY *et al.* [1973]).

monovacancies and those involving mobile multivacancies is not so obvious as in Al.

In order to obtain reliable data, the boundary conditions must carefully be controlled in the experiments (LENGELER and BOURASSA [1976], fig. 4) and/or systematically varied solutions of clustering models must be fitted to the data for evaluation (SAHU *et al.* [1978]). The results of such fits are not single sets of enthalpy and entropy changes for migration of single vacancies but "fields of existence" for possible parameter sets which include also the vacancy-binding parameters up to higher clusters. Enthalpies of monovacancy migration are listed in table 1.

The stage III controversy arose when the recovery stage III of irradiated samples always occurred at lower temperatures than the main recovery stage in quenched samples. A systematic comparison is by BALLUFFI [1978]. As an example, the temperatures of recovery T_a measured by more than 10 groups of authors in Pt are compared in fig. 10. The widespread data-scatter for both types of treatment as well as the systematic difference between irradiation and quenching is obvious. It is quite suggestive to take the discrepancy of 200 K between the most often observed recovery temperatures in irradiated and quenched samples as overwhelming evidence for the migration of two entirely different defects, as was done by the advocates of the two-interstitial model. According to this model, stage III in irradiated metals is caused by the migration and subsequent recombination of the so-called stable interstitial which arises from the metastable one by conversion (SEEGER [1958, 1975]). Its enthalpy of migration is supposed to equal the effective enthalpy of migration found for stage III, whereas the

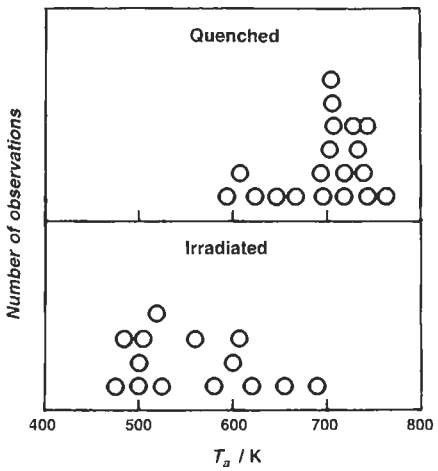


Fig. 10. Temperatures T_a (definition see fig. 7) measured in Pt between 400 K and 800 K after quenching or irradiation, by different authors. Each circle represents one published isochronal recovery stage. The ordinate gives the observation frequency of a given T_a. (after BALLUFFI [1978].)

vacancy migration enthalpy follows from the respective quantity found in quenched samples. The behavior of prequenched and irradiated samples has been studied for Al (BAUER [1969]), Au (LEE and KOEHLER [1968]) and Pt (DUESING and SCHILLING [1969], SONNENBERG *et al.* [1972]). Typical curves for Pt are shown in fig. 11. The splitting of the single stage III in the unquenched state into two stages in the prequenched state has been interpreted by SCHILLING *et al.* [1970] in terms of a mono-di-vacancy model which takes into account the different concentrations and local distributions of sinks in the quenched and unquenched samples. Very careful measurements and data evaluation revealed a monotonic variation of ΔH^m_{eff} when proceeding through the stage III–IV region. The data were quantitatively interpreted by a mono-divacancy model with $\Delta H^m_{1v} = 1.45$ eV, $\Delta H^m_{2v} = 1.00$ eV and $\Delta H^b_{2v} = 0.15$ eV (SONNENBERG *et al.* [1972]).

BALLUFFI (1978) compared the reaction processes in stage III for the irradiated and quenched state by plotting ΔH^m_{eff} against $\bar{T}_a$, the T_a values averaged over all reported measurements, as shown in fig. 12. The data for the different metals lie well on straight lines through the origin, but quenched and irradiated states are characterized by different lines. The equality $\Delta H^m_{eff}/k_B T_a$ for the different metals indicates the close similarity of the parameters determining migration and annihilation, such as frequency factor, sink geometry and density, and initial defect concentration. The different slopes for quenched and irradiated states indicate deviation of at least one of these parameters. The difference can be explained by a number of jumps-per-vacancy until annihilation which is larger by a factor of 10^3 in the quenched state than in the irradiated state. This order of magnitude had already been derived earlier (SONNENBERG *et al.* [1972], JOHNSON [1970], SCHILLING *et al.* [1970], LEE and KOEHLER [1968], LWIN *et al.* [1968]).

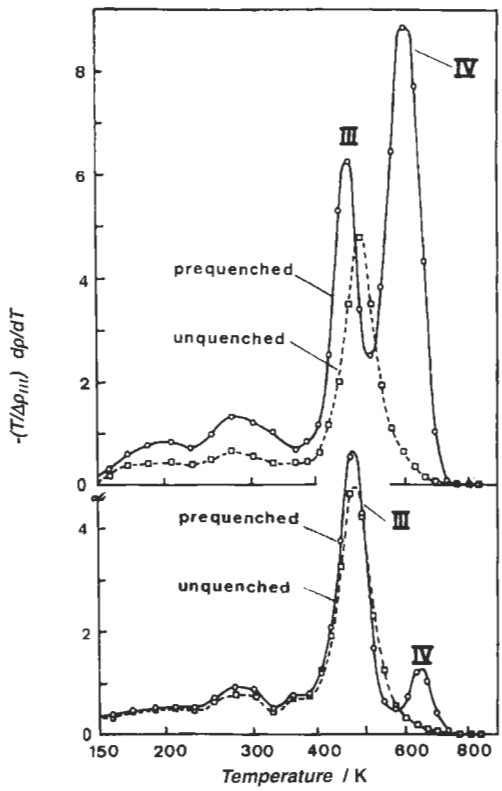


Fig. 11. Temperature-differentiated isochronal recovery curves of Pt after low-temperature electron-irradiation. Shown is the influence of pre-quenching and that of initial defect concentration (the ratio of quenched-in resistivity to irradiation-induced resistivity is large for the upper curves, small for the lower curves). The curves are normalized by different values of $\Delta\rho_{III}$ (see ordinate). Merely for that reason a deviation between dashed and solid curves occurs below 350 K. (after SCHILLING *et al.* [1970].)

2.2.4. Agglomeration

Vacancy agglomerates which are large enough to be observable in the electron microscope have been studied in great detail (ch. 11). The topology of configurations involves dislocation loops, stacking-fault tetrahedra and voids. Review articles are by EYRE *et al.* [1977], KIRITANI [1982], and SHIMOMURA *et al.* [1982]. Since the resolution of common electron microscopy is limited to 1–2 nm, the observable agglomerates certainly contain more than ten vacancies (see also § 4.2.2). About the same resolution limit holds for diffuse X-ray scattering (EHRHART *et al.* [1982] and LARSON and YOUNG [1982]). Field ion microscopy does allow imaging of agglomerates consisting of less than ten vacancies (WAGNER [1982]). It has, however, not been applied yet to questions like shape and size distribution of small vacancy agglomerates as formed by the encounter of migrating vacancies.

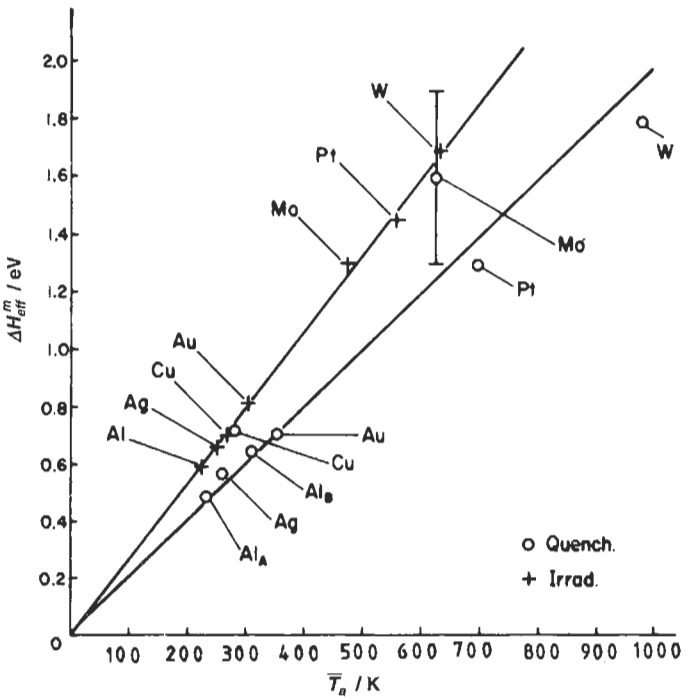


Fig. 12. Effective enthalpies of defect migration versus T_a (definition see fig. 7) averaged over the results by different authors for quenched and irradiated samples (after BALLUFFI [1978]).

While experimental information on configuration and stability of di-, tri- and tetra-vacancies is lacking, computer simulations give some information by using a special nearest-neighbor interaction potential constructed to similate the bcc α-Fe (JOHNSON [1964]). This potential was also used for simulating a fcc model crystal which behaves similar to Ni with respect to the elastic properties. This model yielded for the single vacancy: $\Delta H^f_{1v} = 1.49$ eV, $\Delta H^m_{1v} = 1.32$ eV and $\Delta V^{rel}_{1v} = -0.15$ atomic volume. The stable di-vacancy consisting of the vacancies at nearest-neighbor sites is characterized by $\Delta H^b_{1v} = 0.25$ eV and $\Delta H^m_{2v} = 0.9$ eV. The most stable tri-vacancy, shown in fig. 13a, yielded $\Delta H^b_{3v} = 0.75$ eV and $\Delta H^m_{3v} = 1.02$ eV. The migration includes an intermediate dissociation step. Reorientation of the trivacancy with a lower activation enthalpy than ΔH^m_{3v} occurs by jump of the atom which forms a tetrahedron together with the tri-vacancy (fig. 13b). The stable tetravacancy is that shown in fig. 13d with $\Delta H^b_{4v} = 1.51$ eV. The stability of the three-dimensional tetrahedron as compared to the two-dimensional rhombic configuration (fig. 13c) might be a direct consequence of the nearest-neighbor interaction potential (DEDERICHS *et al.* [1978]).

For the α-Fe model crystal $\Delta H^m_{1v} = 0.68$ eV was found for the nearest-neighbor jump. The most stable di-vacancy configuration places the vacancies at second-neighbor distances $\Delta H^b_{2v} = 0.2$ eV. Migration happens either via the metastable nearest-neighbor

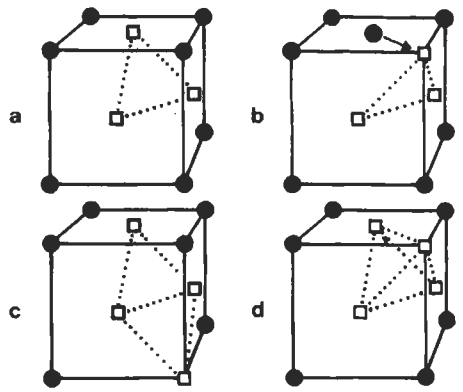


Fig. 13. Tri- and tetra-vacancy configurations in the fcc structure (see text).

configuration or via a fourth-neighbor configuration, both with an activation enthalpy being about the same as for the mono-vacancy migration. The most stable tri-vacancy has two nearest-neighbor and one second-neighbor spacings with =0.49 eV (JOHNSON and BEELER [1977]). The first immobile aggregate was found to be the tetravacancy. The local vibration densities of states of the neigbouring atoms of vacancy clusters were calculated for Cu, α-Fe and α-Ti by YAMAMOTO [1982]. Formation and migration enthalpies of di-vacancies have frequently been derived from experiments as described in §§ 2.2.2 and 2.2.3 (EHRHART [1991]). But the derived numbers result from fitting models to measured data by assuming that eq. (4) holds. However, direct evidence for divacancies causing the observed deviation of $c_v^0(1/T)$ from linearity has not been reported. Another reason would be a temperature dependence of the activation enthalpies of vacancy formation (AUDIT [1982], GANNE and QUÉRÉ [1982], VAROTSOS and ALEXOPOULOS [1982]).

2.2.5. Interaction with solutes

The interaction of vacancies with solutes controls solute diffusion, also known as impurity diffusion, and the self-diffusion of the solvent atoms. This topic is treated in ch. 7. Here, we briefly look at the interaction energies and their influence on c_v^0. In theory, a number of different approaches have been followed in the past in order to obtain quantitative information about the vacancy solute interaction (review by DOYAMA [1978]). Nowadays ab initio calculations are performed by means of the KKR Green's function method based upon the density functional theory in the local spin density approximation (KLEMRADT *et al.* [1991]). The method allows one to calculate the nearest neighbor interaction of vacancies with 3 d and 4 sp solute atoms in Cu and Ni as well as with 4 d and 5 sp solute atoms in Ag and Pd. As a result the data shown in fig. 14 were obtained. A positive sign of the interaction energy means repulsion and a negative sign, attraction between vacancy and solute. For the sp solute atoms in Ag and Pd, the attraction is approximately proportional to the valence difference. Such proportionality was discussed for a long time in the literature for experimental data and was often

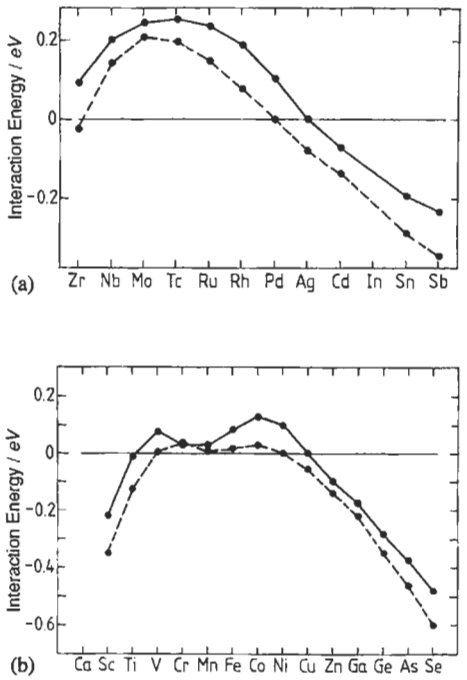


Fig. 14. Calculated interaction energies of a vacancy with a solute atom of the element given by the abscissa for the solvents silver (—) and palladium (---) in (a) and copper (—) and nickel (---) in (b) according to KLEMRADT *et al.* [1991].

compared with a proportionality to the lattice parameter according to a different approach of modelling (review by BENEDEK [1978]). The same behavior is seen for the sp solute atoms in Cu and Ni. The significantly different behavior of the 3 d solute atoms in Cu and Ni as compared to that of the 4 d solute atoms in Ag and Pd is due to their magnetic moment. The magnetic exchange energy reduces the repulsive energy to very small values. KLEMRADT *et al.* [1991] find remarkably good agreement of their data with experimental ones as obtained from solute diffusion measurements. The authors also report on first results for binding energies of a second solute atom with a vacancy solute pair. This configuration and high order clusters are of great interest for the understanding of the dependence of solvent and solute diffusion on solute concentration.

The direct experimental method of determining ΔH^b_{vsa} for vacancy solute atom pairs is the comparison of DD measurements for dilute alloys and the pure solvent metal. In the dilute alloy c_v^0 is enhanced by a term being proportiorial to $c_{sa} \exp[-(\Delta H^f_v - \Delta H^b_{vsa})/k_B T]$. Further methods are equilibrium PAC measurements and resistivity measurements after quenching, both comparing dilute alloys and pure solvent metals (DOYAMA [1978]). A careful comparison of DD measurements and PAS measurements in Cu and dilute Cu-Ge alloys performed by KLUIN and HEHENKAMP [1991] yields good agreement for the

results from the two methods after a reassessment of the PAS models applied for the evaluation of c_v^0.

2.3. Vacancies in ordered alloys

Long-range atomic order is a common phenomenon in intermetallic phases, (ch. 4). Examples are Cu_3Au and CoGa. For many intermetallic compounds the critical temperatures for the order–disorder transformation exceed the melting temperature, which indicates high ordering energies. For such alloys the point defects show additional features to those known from pure metals or disordered solid solutions. Let us inspect the already quoted CoGa which belongs to the group of β-brass electron compounds and crystallizes in the CsCl structure. This structure is made up of two cubic primitive sublattices α and β occupied by A (Co) and B (Ga) atoms, respectively. The latter atoms are positioned at the body-centered sites of the first lattice. Vacant lattice sites in the two sublattices are not likely to be energetically equivalent, owing to their different nearest neighbor atomic shells. The vacancy fractions in the two sublattices may be quite different, which can be inferred from the following naive picture. The vacancy in the α–sublattice is surrounded by only B atoms. This arrangement might suggest for the enthalpies of formation $\Delta H_v^\alpha = \Delta H_v^B$ where the right hand side describes the pure B metal case. This assumption is indeed confirmed for a number of group VIII-group IIIA compounds. For Ga we have $\Delta H_v^B \approx 0.5$ eV leading to vacancy concentration in the Co sublattice of about 10% at 900°C. The more detailed treatment by MIEDEMA [1979] yields $\Delta H_v^\alpha = 0.48$ eV while equilibrium measurements yielded 0.23 ± 0.06 eV. The same reasoning leads to $\Delta H_v^\beta = 1.4$ eV and, hence, vacancy fractions in the β-sublattice being many orders of magnitudes smaller than that in the α-sublattice. Certainly, such large total amounts of vacancies will significantly determine macroscopic properties of the material.

How are the excess vacancies in the α-sublattices being formed? An excess vacancy (excess with respect to the vacancies in the β-sublattices) can be formed only by transferring an A atom into the β-sublattice. Such an atom forms an anti-site defect (also called antistructure atom), i.e., a point defect specific for ordered alloys. Because of the requirement of equal numbers of sites per sublattice the one anti-site defect must be accompanied by two vacancies in the α-sublattice. We have a triple defect.

The creation of anti-site defects introduces the ordering energy as a controlling parameter for concentration and mobility of the vacancies in the α-sublattice. The usual nearest neighbor jump distance (<111>/2) must be replaced by the next nearest neighbor distance (<100>). In particular, the vacancies are retained from annihilation during quenching or even slow cooling. By this means, vacancy concentrations of the 1% order of magnitude can easily be observed at room temperature for a number of intermetallics. On the other hand, the degree of order does not only depend on the ordering energy but also on the formation enthalpy of the α-vacancies. This condition leads to a curved Arrhenius plot for the degree of order as being observed.

For some of the intermetallics significant repulsive interaction of the vacancies has been observed and explained by the electrical charge which may be as high as one

electron per vacancy (for FeAl, KOCH and KOENIG [1986]). Repulsive interaction leads to dispersed vacancy distribution as observed for FeAl whereas formation of large voids (50–100 nm in diameter) is observed for NiAl. For NiAl and for CoAl the vacancy concentration is significantly smaller than for FeAl and CoGa. The latter compounds show a higher degree of intrinsic disorder than the former ones. Long-range ordering of vacancies has been observed for a number of more complex intermetallics (LIU PING and DUNLOP [1988]).

An important feature of many intermetallic compounds is their stability with respect to deviation from stoichiometry. The compound CoGa is found to be stable from 45 to 65 at.% Co for slowly cooled specimens. The vacancy concentration in the Co sublattice amounts to about 10% for 45 at.% Co and falls below 0.1% at 65 at.% Co. At the stoichiometric composition it amounts to about 2.5%. At first sight, these vacancies might be taken as structural ones, i.e., formed to assure the lattice stability for the respective composition. On the other hand, the above-mentioned sluggish approach towards the thermal equilibrium concentration of vacancies does substantially aggravate the discrimination between thermal and structural vacancies. For CsCl structure compounds the existence of structural vacancies is still discussed controversially in the literature (KIM [1986], WEVER [1992]).

For the intermetallic compounds of more complex structures the above-treated questions are studied to even lesser extent by systematic means which is mainly due to the great difficulties with the preparation of reliable sample material. Nevertheless, for a few cases, most interesting properties have been studied in great detail. One example is the Zintl phase β-LiAl which is of great interest as an anode material for lithium sulfur batteries. At room temperature this material contains 0.4% and 7% Li vacancies for 47 at.% Al and 52 at.% Al, respectively. The activation energy for Li self-diffusion is of the order of 0.1 eV. The Li vacancies are long-range ordered at 80 K (KIM [1986]).

The intermetallic compounds Nb_3Sn, V_3Ga, and Nb_3Ge (A15 structure) are of great interest because of their excellent superconducting properties. These properties are closely related to a specific feature of the A15 structure which contains linear chains of the transition metal atoms. The transition temperature for the superconducting to the normal state was found to depend strongly on the quenching temperature for V_3Ga. This dependence has successfully been interpreted in terms of the anti-site defects (Ga atoms within the V atom chains) produced by thermal disordering at the quenching temperatures (BAKKER [1987]). Diffusion properties of intermetallic compounds are reviewed by WEVER [1992]. (See also ch. 7, § 5.2). A good recent review is by DE NOVION [1995].

3. *Self-interstitials*

3.1. Production of interstitial atoms

3.1.1. Introduction

The interstitial formation energy of a few eV is easily provided by irradiating a crystal with energetic particles. For example, an electron of 400 keV energy transfers

19 eV recoil energy to a Cu nucleus via head-on collision. The maximum energy transferred to Cu by fission neutrons of 2 MeV amounts to 125 keV. The fundamental problems of radiation damage in materials intensively stimulated research regarding particle–lattice-atom interaction and the Frenkel defect production which results. In the present section we deal with this matter with regard to the atomic defect concentration c_d produced by a *given fluence* Φ (time-integrated flux density $\emptyset$) of particles penetrating a crystal. Among the numerous questions on the Frenkel defect production this point is particularly important for the evaluation of specific (microscopic) defect properties from radiation-induced changes of macroscopic materials properties.

3.1.2. Atomic displacement cross-section for electron irradiation and the production of stable Frenkel defects

The atomic fraction of collisions between projectiles and target atoms for a fluence Φ is given by $\sigma\Phi$. The term *collision* ought to be specified in order to render the *cross-section* σ a well-defined quantity. Electrons penetrating a crystal interact with the electrons as well as with the nuclei of the target atoms. In metals, the electron-electron interaction produces heat and, hence, requires cooling of irradiated samples. At the same time it causes an angular spread of the electron beam around the direction of incidence. This spread is of relevance for the angular resolution of threshold-energy determinations (see below) and for the flux-density determination in any kind of electron-irradiation experiment except for very thin samples.

The electron–nucleus interaction causes the displacement of atoms from regular lattice sites to interstitial positions. The necessary specification of this collision is given by the minimum recoil energy required for permanent displacement of an atom. Only collisions with recoil energies T larger than the *displacement threshold* T_d are of relevance for the Frenkel defect production. The total cross-section can be written as the recoil-energy integral of the differential cross-section $d\sigma$,

$$\sigma(E) = \int_{T_d}^{T_{max}} \frac{d\sigma(E,T)}{dT}\, dT, \tag{7}$$

taken from the displacement threshold energy to the maximum transferred energy (head-on collision), E is the electron energy. The differential cross-section for the scattering of a relativistic (Dirac) electron by a point nucleus was calculated by MOTT [1932]. Relativistic electron scattering favors mean recoil energies and reduces high ones when compared to Rutherford scattering, for which $d\sigma/dT \propto T^{-2} = T_{max}^{-2} \cos^{-4}\theta$ holds, where θ is the starting angle of the recoiling atom with respect to the incident electron-beam direction (further reading: CORBETT [1966]).

The lattice structure causes T_d to depend on the recoil impact direction with respect to the lattice orientation. This dependence is shown in fig. 15 for Cu (method of measurement see further below).

Irradiation of textureless polycrystalline samples means randomness of the recoil impact directions with respect to the lattice orientations. If for T_d in eq. (7) the absolute minimum $T_{d,min}$ of the threshold-energy surface were taken, σ would be calculated much larger than the experimental total displacement cross-section. All those collisions with

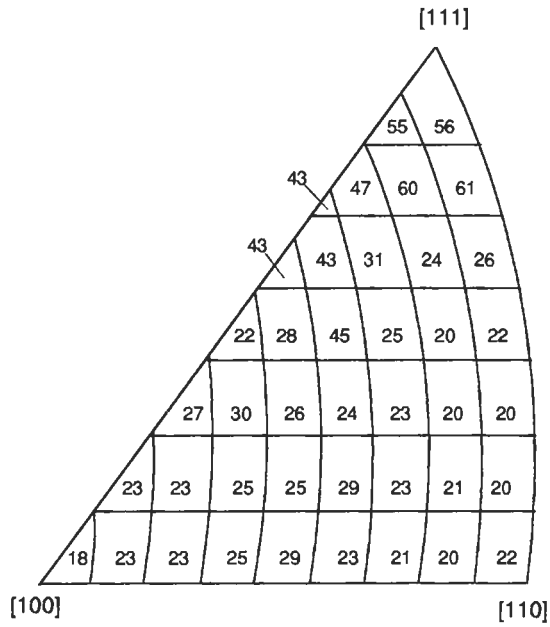


Fig. 15. Displacement threshold energy T_d as a function of the atomic recoil impact direction within the fundamental orientation triangle in Cu (after KING *et al.* [1981]).

$T > T_{d,min}$ but with impact directions for which $T < T_d$ holds do not produce Frenkel defects. This effect is often considered phenomenologically by introducing the *displacement probability, p(T)* into eq. (7), leading to the *displacement cross section*

$$\sigma_d(E) = \int_{T_{d,min}}^{T_{max}(E)} p(T) \frac{d\sigma(E,T)}{dT} dT. \tag{8}$$

When the angular dependence of the threshold energy is known, $p(T)$ and hence σ_d can be calculated when $d\sigma/dT$ is known. For the determination of the angular dependent threshold energy, particle irradiation of a single crystal is desired which causes atomic displacements with just one recoil impact direction. By variation of the crystal orientation with respect to the recoil impact direction and by variation of the particle energy, the angular dependence would thus be scanned. Unfortunately, the statistical nature of particle interaction causes a distribution of recoil impact directions which fills the solid angle 2π (maximum angle between incident projectile direction and target-atom recoil impact is $\pi/2$) for any given incident particle direction. The distribution is governed by the angular dependent differential cross-section of the particular type of particle–atom collision. Hence, any scanning of the angular dependent threshold energy requires deconvolution of the measured data with respect to the angular dependent differential cross-section. An additional problem arises for electron irradiation from the considerable beam spread upon passing a sample. It is due to the multiple scattering by target

electrons. For electrons of 0.5 MeV, the angular distribution at the reverse side of a 25 μm thick Cu sample is such that about 40% of the electrons are scattered out of their incident direction by more than 40°. Review articles on earlier threshold-energy surface determinations are by SOSIN and BAUER [1968], VAJDA [1977] and JUNG [1981a].

The angular resolution for T_d has been improved substantially by applying high-voltage electron microscopy (HVEM) for in-situ defect production. Two measuring methods for defect production rates have been developed: (i) residual resistivity measurements in electron-microscope samples, i.e., samples of about 400 nm thickness and 0.1×0.1 mm^2 irradiated area (KING *et al.* [1981]) and (ii) rate measurements for nucleation and growth of interstitial-type dislocation loops (URBAN and YOSHIDA [1981]). The first method is applied at irradiation temperatures below 10 K, whereas the second one requires mobile interstitials, i.e., irradiation temperatures above 50 K (see § 3.3.2.4). Both methods benefit by the small sample thickness, by the ease of sample tilting with respect to the beam direction and by the electron flux density which is some orders of magnitude larger than at common accelerator irradiations. The orientational dependence of threshold energy shown in fig. 15 was obtained by resistivity-change rate measurements at six different electron energies and about 35 different crystal orientations, yielding about 200 data points altogether.

The displacement probability $p(T)$ derived from the threshold energies in fig. 15 is shown in fig. 16 with error bars. Calculations of the total displacement cross-section according to eq. (8) with such $p(T)$ data leads to uncertainties of 10–15%. A similarly exhaustive determination of the angular dependence of T_d as in Cu does not exist for any other metal yet. As a consequence, defect production rates can generally not be calculated with satisfactory accuracy for given irradiation experiments by applying eq. (8). Another method often applied is the measurement of a standard quantity as the electrical resistivity under the given irradiation condition and relating all measured property

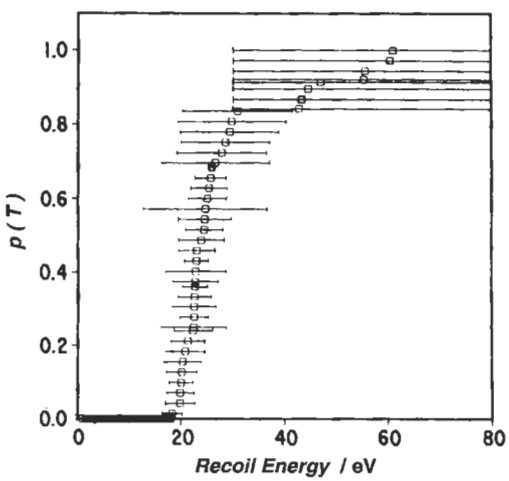


Fig. 16. Displacement probability versus recoil energy T_d as resulting from the angular dependence in fig. 15. The error bars indicate the sensitivity of the threshold-energy surface fit to the measured damage rate data.

changes to that of the standard quantity. This procedure is described in § 3.2. The resistivity contribution per unit concentration of Frenkel defects, ρ_F enters into the determination of the angular dependent T_d as the displacement cross-section $\sigma_d(E, \theta)$ for the electron energy E and the incident beam direction θ is measured by the resistivity damage rate $d\rho/d\Phi = \rho_F \sigma_d$. If certain conditions regarding (i) the angular dependence of the displacement probability (not discussed here) and (ii) the magnitude of the available recoil energies T relative to T_d are fulfilled, the value of ρ_F can be derived from such measurements as well as the angular dependent T_d (ABROMEIT [1983]). The results in figs. 14 and 15 have been obtained with $\rho_F = 2 \times 10^{-4}\,\Omega$ cm. But the optimal fit assuming the above-mentioned conditions to be fulfilled yielded $\rho_F = 2.85 \times 10^{-4}\,\Omega$ cm (KING and BENEDEK [1981], see also table 2).

The data in fig. 15 tell us that the production of stable Frenkel defects is energetically more favorable for recoil impacts near <100> and <110> than for those near <111> and its environment. Analytical calculations (LEIBFRIED [1965]) and computer simulations (GIBSON *et al.* [1960]) have shown that impacts along these favorable directions initiate *replacement collision sequences* (fig. 17) which produce the interstitial separated from the vacancy by several nearest neighbor distances (nnd). The energy release per nnd along the replacement chain amounts to about 0.7 eV. A recoil impact along <111>, for comparison, leads to simultaneous collisions of the knock-on atom with three nearest neighbor atoms which form a equilateral triangular window around the travelling path of the primary knock-on atom. As a result, the recoil energy of this 4-body collision is divided by three and accordingly the kinetic energy of the recoiling atoms which remains available for defect production. Around 60 eV primary recoil energy is required to produce one Frenkel defect while about 20 eV are sufficient for recoil impacts along <100> and <110>.

It is obvious that an interstitial cannot be stable at or close to the nearest neighbor position of a vacancy. *Spontaneous recombination*, occurring even at $T = 0$, will annihilate

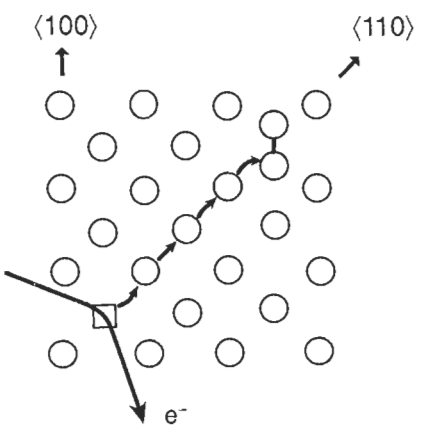


Fig. 17. Frenkel defect production via replacement collision sequence by an incident electron in the MeV range (schematic).

References: p. 1710.

the displacement process. Hence, a minimum distance between vacancy and interstitial is required to render the Frenkel defect mechanically stable. This fact and the above-discussed energetically unfavorable many-body collisions along higher-indexed lattice directions suggest that the direct transfer of the primary knock-on atom into a stable interstitial position is unlikely to occur. The displacement of the primary knock-on atom to the nearest neighbor position has indeed been proven experimentally by means of PAC analysis after neutrino recoil-induced Frenkel pair production (METZNER *et al.* [1987]). The occurrence of replacement collisions is proven by the radiation induced disordering of long-range ordered alloys (KIRK *et al.* [1978]).

The minimum distance between vacancy and interstitial assuring their stability at $T=0$ K has been studied by computer simulation (DRITTLER *et al.* [1969]). The result is shown in fig. 18. The volume given by the unstable lattice sites around the interstitial at which the vacancy is unstable at $T=0$ K is called the *spontaneous recombination volume* v_0 (fig. 18a). An irradiated crystal allows stable Frenkel defect production only in that volume fraction which is not occupied by the recombination volumes of the preceedingly produced Frenkel defects. Due to the spontaneous recombination the defect production is reduced according to the factor $f_{sr}=(1-2v_0c_d)$ to a first order approximation (review by

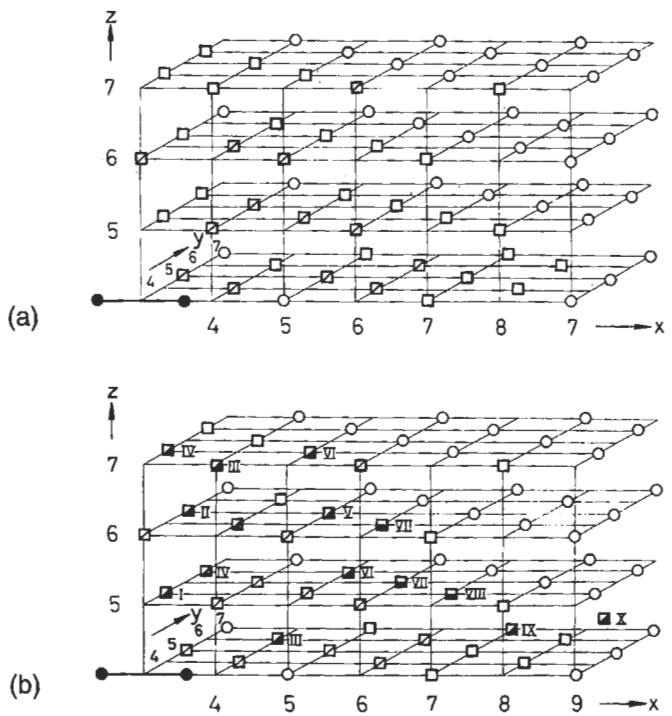


Fig. 18. Unstable vacancy positions around the <100> dumbbell in Cu for (a) T = 0 K, ● interstitial, ◪ unstable, □ stable, ○ not investigated, (b) 0.1 eV recoil energy given to one of the dumbbell atoms, ◪ unstable at $T=0$, ◪ dynamically unstable (0.1 eV), ◒ probably unstable, □ stable.

WOLLENBERGER [1970]). Under electron irradiation the spontaneous recombination volume is apparently enlarged by the so-called subthreshold collisions. The effect is demonstrated in fig. 18b showing those lattice sites which are unstable for the vacancy when one of the interstitial atoms is recoiling with 0.1 eV. The cross section of the lattice atoms for recoil energies larger than 0.1 eV are following from eq. (7) when the lower boundary T_d of the integral is replaced by 0.1 eV. The cross-section for such subthreshold collisions is by two orders of magnitude larger than the displacement cross section and is therefore of similar magnitude as the term $2v_0$. The computer simulations yield $v_0 \approx 62\,\Omega$.

From measurements of the resistivity damage rate as a function of the radiation-induced resistivity increase (defect concentration) for electron irradiation, f_{sr} has been derived (review by WOLLENBERGER [1970]) with the result $v_0 = 125\,\Omega$ for Cu. The subthreshold annihilation rate was indeed found to approximately equal $2v_0 c_d$. Similar values were found for other metals. The difference for v_0 between experiment and computer simulation is most probably due to the fact that in reality the Frenkel defect is produced with about 50 eV recoil energy in the average accompanied by strong lattice exitation while in the model the configuration was initially set without any lattice excitation. The cross-section for subthreshold Frenkel defect annihilation was found experimentally to be 4×10^{-25} m^2, which agrees well with 0.1 eV for the lower limit of effective subthreshold recoil energies as assumed in the computer simulations.

At finite temperatures, the displacement process is influenced by the lattice vibration affecting the energy release (phonon production) during the replacement sequence and the minimum separation of vacancy and interstitial required for the Frenkel defect stability. The minimum displacement threshold energy for Cu was found to decrease with increasing irradiation temperature from 19 eV at 4 K to 11 eV at 450 K (URBAN *et al.* [1982]) indicating a substantial assistance of the energetically most favorable Frenkel defect production by lattice vibrations.

The minimum vacancy-interstitial separation naturally increases when the interstitials become mobile. The *close pairs* (see § 3.3.2.4) recombine at temperatures from 15 K upwards to about 30 K for Cu. As a result, the recombination volume is larger than v_0 by the volume shell given by the close pair interstitial sites (when the vacancy is assumed to be at the center of v_0). At the same time the effective defect production rate for $c_d = 0$ is reduced by nearly $f_{cp} = 25\%$ (fractional recovery due to close pair annihilation).

At temperatures above 30 K, the correlated recombination of *freely migrating* interstitials (§ 3.3.2.4) further reduces the production rate. The fraction of freely migrating interstitials recombining correlatedly is given by $1 - \langle r_v / r_p \rangle$, where r_p is the average vacancy interstitial separation attained by the replacement collision sequence and r_v the radius of the spherically assumed recombination volume for thermally activated migrating interstitials. Both quantities have been determined by means of resistivity damage rate measurements. The ratio $\langle r_p / r_v \rangle = 2$ was found for incident electron energies above 1 MeV (BECKER *et al.* [1973]) and 80 K irradiation temperature. Near the displacement threshold, $\langle r_p / r_v \rangle$ rises steeply up to about ten. The radius r_v was found to depend on temperature according to $T^{-1/3}$ between 50 K and 105 K (LENNARTZ *et al.*

[1977]). This temperature dependence is expected when the elastic interaction of vacancy and interstitial controls the migration path of the interstitial. In effect, the initial displacement rate at temperatures at which interstitials migrate freely (above 32 K for Cu) is given by $\rho_d(1-f_{cp})\ (1-\langle r_v/r_p\rangle)$. The dependence of the defect production rate on c_d is strongly influenced by the microstructure of the sample and thus provides a useful method of analysis for the microstructure (review by WOLLENBERGER [1978]).

3.2. Determination of Frenkel defect concentrations

Specific properties of Frenkel defects are determined by measuring a certain property change of a macroscopic sample upon irradiation and relating the measured change to the underlying atomic defect concentration, c_d. Determination of c_d from the irradiation parameters as treated in the foregoing section is subject to substantial uncertainty. In order to avoid such uncertainties it has become usual to measure the electrical residual resistivity increase $\Delta\rho$ of a sample made of a proper standard material for which the constant of proportionality ρ_F is known and to calculate $c_d = \Delta\rho/\rho_F$. Resistivity samples can generally be irradiated together with the sample under specific investigation. Thus ρ_F plays the role of a defect-concentration standard. As has been shown in the foregoing section, ρ_F can be derived from resistivity-damage rate measurements in single crystals under electron irradiation at low temperatures and under suitable variation of crystal orientation and electron energy.

For the most extensively investigated Cu the uncertainty ρ_F amounts to about ±15 %. For other metals the uncertainty is larger. Values are given in table 2.

An entirely different method of determining c_d is the proper evaluation of the X-ray scattering intensity in the Huang range and in the range between the Bragg reflections (see § 3.3.2.2, also ch. 12). The Huang scattering intensity is proportional to $c_d[(\Delta V_{1i}^{rel})^2+(\Delta V_{1v}^{rel})^2]$ (HAUBOLD [1975], DEDERICHS [1973]). By determining the lattice constant change, $\Delta a/a$, one obtains $c_d(\Delta V_{1v}^{rel}+\Delta V_{1I}^{rel})$. From both results, ΔV_{1i}^{rel} as well as c_d can be derived once ΔV_{1v}^{rel} is known from other measurements. Resistivity measurements on the same sample or under the same irradiation conditions then give ρ_F. The error limits amount to about ±20% (SCHILLING [1978], EHRHART [1991] and SCHULTZ [1991]).

3.3. Interstitial properties

3.3.1. Results of model calculations

Methods for calculation of point-defect properties are provided by the theories of lattice statics and dynamics (reviews are by DEDERICHS and ZELLER [1980], LEIBFRIED and BREUER [1978] MARADUDIN *et al.* [1971]) and by computer simulation (reviews are by BEELER [1983], JOHNSON [1973], GEHLEN *et al.* [1972]). Computer simulation was introduced into the point-defect research field by GIBSON *et al.* [1960] and ERGINSOY *et al.* [1964]. They studied the fundamental displacement processes initiated by recoiling atoms and with that gave the first hints for an understanding of the shape of the angular dependence of the threshold-energy (§ 3.1.2) on the basis of replacement-collision

Table 2
Electrical resistivity contribution per unit concentration of Frenkel defects $\rho_F[10^{-4}\ \Omega\ \text{cm}]$ for various metals.

Metal	Contribution	Method[a]	Reference
Ag	2.1	$\Delta\rho_{irr}$	LUCASSON and WALKER [1962]
			ROBERTS *et al.* [1966]
Al	4.0	$\Delta\rho_{irr}$	JUNG [1981b]
	3.9	HS & $\Delta a/a$	BENDER and EHRHART [1983]
Au	3.0	$\Delta\rho_{irr}$, HS & $\Delta a/a$	SEGURA and EHRHART [1979]
Co	15	$\Delta\rho_{irr}$	VAJDA [1977]
	16	HS & $\Delta a/a$	EHRHART and SCHONFELD [1982]
Cr	37	$\Delta\rho_{irr}$	BIGET *et al.* [1979]
Cu	2.5	DXS & $\Delta a/a$	EHRHART *et al.* [1986]
	2.85	$\Delta\rho_{irr}$	KING and BENEDEK [1981]
	2.8	HS & $\Delta a/a$	BENDER and EHRHART [1983]
α-Fe	30	$\Delta\rho_{irr}$	VAJDA [1977]
	20	DXS & $\Delta\rho_{irr}$	EHRHART [1985]
Ga	5.4	$\Delta\rho_{irr}$	MYHRA and GARDINER [1975]
K	3	$\Delta\rho_{irr}$	WALLNER *et al.* [1986]
Mo	15	HS & $\Delta a/a$	EHRHART and SEGURA [1975]
	13	$\Delta\rho_{irr}$	VAJDA [1977]
Nb	14.6	DXS	EHRHART [1978]
Ni	6	$\Delta\rho_{irr}$	LUCASSON and WALKER [1962]
	7.1	HS & $\Delta a/a$	BENDER and EHRHART [1983]
Pt	9.5	$\Delta\rho_{irr}$	JUNG *et al.* [1973]
Re	20	$\Delta\rho_{irr}$	LUCASSON [1975]
Rh	15	$\Delta\rho_{irr}$	GUINAN *et al.* [1973]
	19	$\Delta\rho_{irr}$	LUCASSON [1975]
Sc	50	$\Delta\rho_{irr}$	DAOU *et al.* [1985]
Sn	1.1	$\Delta\rho_{irr}$	DIMITROV and DIMITROV [1985]
Ta	16	$\Delta\rho_{irr}$	JUNG *et al.* [1973]
			BIGET *et al.* [1979]
α-Ti	14..18	$\Delta\rho_{irr}$	SHIRLEY and CHAPLIN [1972]
Zn	15	$\Delta\rho_{irr}$	VAJDA [1977]
	15	HS & $\Delta a/a$	EHRHART and SCHONFELD [1982]
Zr	35	$\Delta\rho_{irr}$	NEELY [1970]
	40	$\Delta\rho_{irr}$	LUCASSON [1975]

[a] $\Delta\rho_{irr}$: electrical resistivity of irradiated samples; HS: Huang scattering; $\Delta a/a$: lattice parameter, DXS: diffuse X-ray scattering.

sequences (JUNG [1981b], URBAN and YOSHIDA [1981], TENENBAUM and DOAN [1977], VAJDA [1977], ROTH *et al.* [1975], BECKER *et al.* [1973], WOLLENBERGER and WURM [1965]).

Such computer simulations are based on the numerical solution of the equations of motion for a set of atoms arranged as in the crystal and coupled by spring forces which simulate the actual lattice–atom interaction. Critical aspects of the method are the choice of the the interaction potentials and the adjustment of the boundary conditions for the necessarily size-limited model crystal in a way which represents the correct embedding in an elastic continuum. The potentials have generally been chosen so as to generate two-

body central forces and to match known physical properties, such as the elastic constants. Mostly, a particular functional form of the potential, such as the Morse, Lennard–Jones or Born–Mayer type, is taken to be valid and the respective coefficients are determined by matching the model's properties to various physical properties of the modelled metal. In other cases, purely empirical potentials have been developed by the matching procedure. Potentials which result from pseudopotential theory have also been used. Nowadays, realistic manybody potentials are prepared by means of the *embedded atom method* (FOILES *et al.* [1986]). Computer simulations have been applied to study not only defect production mechanisms but also static and dynamic properties of defects (§ 2.1).

For analytical calculations of the interstitial properties, the real space relaxation model is often replaced by the normal coordinate expansion model. The latter was introduced by KANZAKI [1957] and treats the defect-lattice interaction as a source function (*Kanzaki forces*) for the displacement in a harmonic lattice. The force equations for equilibrium are Fourier-transformed, yielding variables which are Fourier inverses of real space displacements. All energy changes are related to the first and second derivatives of the interactions, and these derivatives can directly be matched to the force constants as following from Born-v. Karman fits to the phonon dispersion curves. The properties of primary interest are the activation enthalpies of formation and of migration of the interstitial. The latter is determined as the difference between the interstitial formation enthalpy in the saddle point configuration and that in the equilibrium configuration, in accordance with common use in rate theory.

3.3.1.1. Activation enthalpies of formation in equilibrium and saddle-point configurations. Specific feature of the self-interstitial is the strong lattice distortion with its large displacements in the neighboring atom shells (0.14 nnd for the nearest neighbors of the two atoms forming a split interstitial configuration in Al (HAUBOLD [1976])). Consequently the lattice relaxation significantly influences the repulsive energy contribution. One easily realizes that the large capacity of computers is very helpful for handling just this contribution as accurately as possible.

The electronic contribution consists of the electron–energy change due to the volume change of the crystal upon insertion of the interstitial atom. Those results given in table 3 which were not obtained by computer simulation essentially deviate by different treatment of the electronic contribution.

In order to find the stable interstitial configuration, the formation enthalpy has been calculated by a number of authors for all configurations shown in fig. 19. In table 3 only minimum and maximum values are given. They indicate that with one exception the difference between minimum and maximum of the formation energy is less than 15% of the minimum value, and detailed comparison would show that there are always two or three different configurations for which the formation energy deviates by less than 5%. The origin of these small differences lies in the large relaxation effect (see § 3.3.2.1). Without the lattice relaxation the formation enthalpy would be larger by nearly one order of magnitude (DEDERICHS *et al.* [1978]). Rearrangement of the interstitial atom or atom pair does not cause a significant change of the relaxation volume (see table 3) and, hence, changes the formation enthalpy only by a small amount. The data reported by DEDERICHS *et al.* [1978] are obtained by computer simulation with a Morse potential

Table 3
Calculated properties of self-interstitials.

Metal	Configuration	ΔH_i^f (eV)	ΔV_i^{rl} (at. volume)	ΔH_i^m (eV)	Reference
Cu	<100>-split	5.07–5.82		< 0.24	HUNTINGTON [1942, 1953]
	Octahedral	5.14–6.09			
	Octahedral	2.5–2.6	1.67–2.01		TEWORDT [1958]
	<100>-split	3.2	1.10–1.25		
	Octahedral	2.73–3.42	1.21–1.78		SEEGER and MANN [1960]
	<100>-split	2.187	1.126	0.103	BENNEMANN and TEWORDT [1960]
	Octahedral	2.43–2.44	1.219–1.441		
	<100>-split	4.351		0.090	BENNEMANN [1961]
	<100>-split	4.139	2.20	0.05	JOHNSON and BROWN [1962] [a]
	Crowdion	4.840	2.57		
	<100>-split	2.47–2.84	1.37–1.44	0.46–0.57	SEEGER *et al.* [1962] [a]
	<111>-split	3.83–4.27	1.43–1.65		
	<100>-split	3.39	1.34	< 0.06	DOYAMA and COTTERILL [1967] [a]
	Tetrahedral	3.70	1.38		
	<100>-split	3.42	1.5	0.13	DEDERICHS *et al.* [1978]
	Tetrahedral	3.S9	1.48		
Ni	<100>-split	4.0S	1.7	0.15	JOHNSON [1966]
	Crowdion	4.10		0.04	
	<111>-split	4.24		0.13	
α-Fe	<100>-split	4.G	2.34	0.21	JOHNSON [1965]
	Octahedral	5.73			DEDERICHS *et al.* [1978] [a]
Al	<100>-split	2.S9		0.15	LAM *et al.* [1980]
Mg	A-split	2.36			IMAFUKU *et al.* [1982]
	Hexahedral	2.66			

[a] Gives data on more configurations than listed in this table.

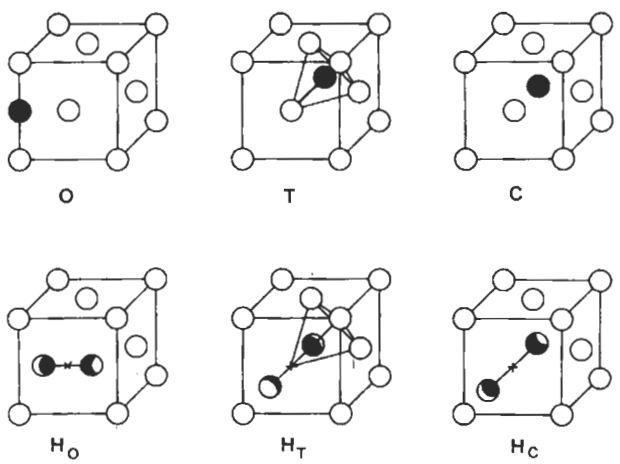


Fig. 19. Self-interstitial configurations in the free lattice. Symmetries: O — octahedral; T – tetrahedral, C – crowdion; H_O, H_T, H_C — dumbbell having axis along <100>, <111>, and <110>, respectively.

References: p. 1710.

modified in order to soften the strong repulsive core of this potential and to fit (in addition to other properties) lattice constant, compression modulus, vacancy formation enthalpy) the relaxation volume of the interstitial to the experimental value (see table 4). For Al an inter-ionic potential as derived from pseudo-potential theory was used.

The activation enthalpy for migration of the Cu interstitial was accordingly obtained to be around 0.1 eV, with the exception of 0.5 eV by SEEGER *et al.* [1962]. This difference was one of the sources which energized the stage III controversy (§ 3.3.2.5). The small difference between the formation enthalpies of the saddle point configuration and that of the equilibrium position shows again that the activation enthalpy of formation depends rather weakly on the interstitial configuration.

Details of possible migration steps can conveniently be investigated by computer simulation. The most probable migrational step obtained for the fcc structure (DEDERICHS *et al.* [1978], LAM *et al.* [1980]) is shown in fig. 20a. It consists of a translational motion of the center of gravity of the dumbbell by one atomic distance and a rotational motion of the dumbbell axis by 90°. It should be noted that any other form of motion of the <100>-split interstitial requires a considerably higher activation enthalpy. Especially the 90° axis rotation with fixed center of gravity requires a four times larger activation enthalpy.

For the bcc α-Fe potential, the elementary jump is found again to be a translational step of the center of gravity of one atomic distance and rotation of the axis by 60° as shown in fig. 20b. The activation enthalpy is 0.21 eV. Contrary to the situation in the fcc structure, the pure 90° axis rotation requires only a little larger activation enthalpy, namely 0.25 eV.

3.3.1.2.Dynamic properties. Computer simulation furthermore revealed the occurrence of low-frequency resonant modes of the <100> dumbbell in the fcc structure besides high-frequency localized modes (SCHOLZ and LEHMANN [1972], IMAFUKU *et al.* [1982]). This unusual occurrence of both types of vibrational modes at one defect

Table 4
Relaxation volume of self-interstitials[a].

Metal	ΔV_i^{rel} (at. volume)	Measurement	Reference
Al	1.9	HS & $\Delta a/a$	BENDER and EHRHART [1983]
Cd	11–19	HS & $\Delta a/a$	EHRHART and SCHONFELD [1982]
Co	1.5	HS & $\Delta a/a$	EHRHART and SCHONNFELD [1982]
Cu	1.3	$\Delta a/a$	DWORSCHAK *et al.* [1972]
	1.45	DXS & $\Delta a/a$	HAUBOLD and MARTINSEN [1978]
	1.7	HS & $\Delta a/a$	BENDER and EHRHART [1983]
Fe	1.1	DXS & $\Delta a/a$, $\Delta\rho$	EHRHART [1985]
Mo	1.1	HS & $\Delta a/a$	EHRHART [1978]
Nb	1.11	DXS & $\Delta a/a$	DÖNITZ *et al.* [1974]
Ni	1.8	HS & $\Delta a/a$	BENDER and EHRHART [1983]
Pt	2.0	$\Delta a/a$	HERTZ and PEISL [1975]
Zn	3.5	HS & $\Delta a/a$	EHRHART and SCHONFELD [1982]
Zr	0.6	HS & $\Delta a/a$	EHRHART and SCHONFELD [1982]

[a]Glossary see tables 1–3.

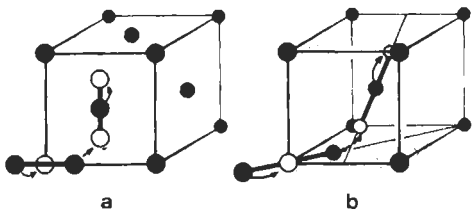


Fig. 20. Migrational steps (a) of the <100>-split interstitial in the fcc lattice and (b) the <110>-split in the bcc lattice, according to computer simulation results.

originates from the highly compressed lattice around the interstitial and its special configuration (DEDERICHS *et al.* [1973]). For the oppositely directed vibrations of the two dumbbell atoms along <100> as shown in fig. 21a, the small equilibrium separation between the two atoms (0.77 nnd) leads to a very strong force-constant which couples the two atoms. This constant leads to a localized mode and for the modified Morse potential one obtains the mode A_{1g} shown in fig. 22, which lies well above the maximum lattice frequency, ω_{max}.

For the displacement directions of the two dumbbell atoms as shown in fig. 21c the strongly compressed spiral spring between the two atoms exhibits a negative bending spring component which acts perpendicular to the spiral spring axes. Quantitative evaluation shows that the force-constant of the negative bending spring becomes comparable to that of the restoring force of the perfect lattice. The resulting force-constant is small so that the librational mode in fig. 21c is a resonant mode with a very low frequency, as shown in fig. 22 for E_g. Another resonant mode (A_{2u}) is excited with the direction of atomic motion as shown in fig. 21b. As the motions in figs. 21b and c also strain the compressed springs with the nearest neighbors of the dumbbell, localized modes are excited as well (see A_{2u} and E_g beyond ω_{max} in fig. 22). The low-frequency resonances of the dumbbell lead to comparatively large thermal displacements of the interstitial atoms. In fig. 23 the mean squares of the atomic displacements are compared for dumbbell atoms and those in perfect lattice positions, as calculated for a Morse

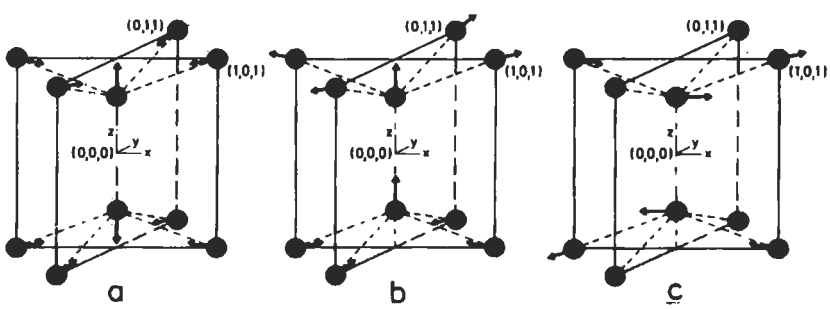


Fig. 21. Localized and resonance modes of the <100>-split interstitial: (a) localized mode A_{1g} (see fig. 22); (b) resonance and localized mode A_{2u}; (c) resonance and localized mode E_g. (From DEDERICHS *et al.* [1978].)

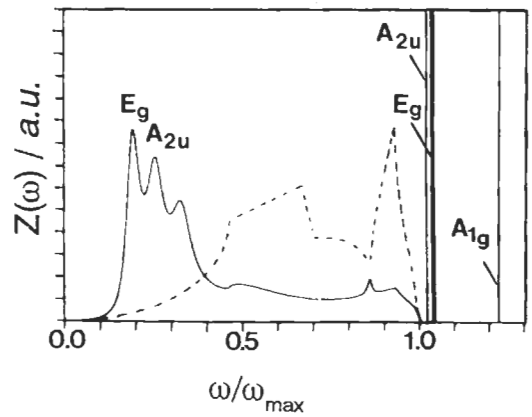


Fig. 22. Local frequency spectrum of the <100>-split interstitial (averaged over all directions) for a modified Morse potential (from DEDERICHS *et al.* [1978]). Dashed curve is for perfect lattice.

potential. It becomes obvious that the resonant modes are thermally populated already around 30 K, the beginning of stage I_D recovery (§ 3.3.2.4) in Cu. Indeed, large amplitudes of the resonant modes lead directly to the saddle point configuration for the migration jump. The flatness of the energy contour along the migration jump path is a direct consequence of the negative bending spring effect.

The presence of resonant modes should cause a low temperature maximum in the temperature dependence of the specific heat (ZELLER and DEDERICHS [1976]), and some experimental indication for its occurrence has been found in electron-irradiated Cu (MONSAU and WOLLENBERGER [1980]). Experimental difficulties, however, prevented a reliable error analysis of the result.

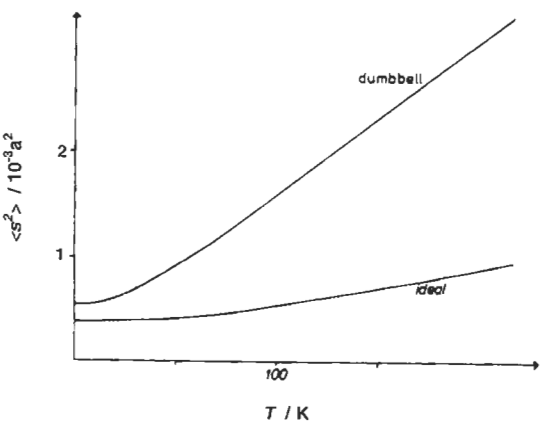


Fig. 23. Averaged square of the thermal displacements of one of the <100>-split interstitial atoms compared with that of a regular lattice atom (from DEDERICHS *et al.* [1978]).

The influence of the resonance modes on the temperature dependence of elastic moduli was observed by HOLDER *et al.* [1974]. They derived $\omega_R = \omega_{max}/8$ which is in good agreement with the calculated modes in fig. 22. The most effective consequence of the resonant mode is the considerable elastic polarizability. In fig. 24 the displacement directions of the dumbbell atoms and their nearest neighbors are shown for two different applied shear stresses and for uniform compression. Obviously the <100>-shear stress excites the resonant mode E_g, and computer experiments show that the rotation angle of the dumbbell axes is by a factor of twenty larger than the shear angle because of the bending spring action. The presence of dumbbell interstitials in a crystal therefore must decrease the elastic constant c_{44} with respect to a perfect crystal. Large decreases of the shear moduli of irradiated polycrystals have indeed been observed many years ago (KÖNIG *et al.* [1964], WENZL *et al.* [1971]). This negative sign of the modulus change was a long-standing problem in theory and had not been solved earlier than the detection of the resonant mode in the computer experiments.

The existence of a para-elastic polarizability of interstitials was discussed by DEDERICHS *et al.* [1978]. For anisotropic defects it could be even one order of magnitude larger than the dia-elastic polarizability. It can be observed directly, however, only for such interstitials as are able to reorient in an external field without performing migration steps towards annihilation. On the other hand, migration in an external field will be influenced by the para-elastic polarizability. This effect might be of great importance for irradiation-induced creep (BULLOUGH and WILLIS [1975], see also § 4.5).

3.3.1.3. Arrhenius behavior of diffusion. The low activation enthalpy for interstitial migration and the low temperature at which it takes place (stage I_D recovery in Cu around 35 K, § 3.3.2.4) raised the question whether Arrhenius behavior, as it has always been assumed for stage I recovery interpretation, actually exists (FLYNN [1975]). Indeed, the majority of vibrational lattice modes occupy their ground states of zero-point motion at temperatures much lower than the Debye temperature ($T_{Deb} = 310$ K for Cu). Hence little justification can be given for a classical description of the migration process from this point of view. On the other hand, computer experiments clearly show that the migrational step is always a consequence of sufficient excitation of both resonant modes of the split interstitial (§§ 3.3.1.2 and 3.3.2.6). Excitation of these modes is possible only by phonons of matching frequency. Since $\Delta H_i^m = 30\ \hbar\omega_R$, extremely high excitations must be provided by incoming phonons. The fluctuation possibility for this state is

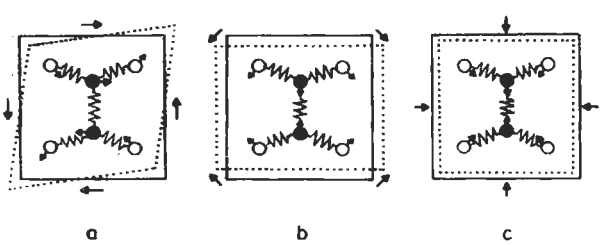


Fig. 24. Homogeneous deformation of a crystal containing a <100>-split interstitial: (a) <100> shear (modulus c_{44}); (b) <110> shear (modulus $(c_{11}-c_{12})/2$); (c) compression (modulus $(c_{11}+2c_{12})/3$).

References: p. 1710.

essentially classical, i.e., the migration velocity depends on $\exp-(\Delta H_i^m/kT)$. This is indeed the classical Arrhenius behavior because the phonon states corresponding to ω_R are classically populated already at low temperatures. The motion of the atoms neighboring the jumping dumbbell observed in computer simulations suggests a description in terms of the classical activated state.

3.3.1.4. Multiple interstitials. In the foregoing sections it has been shown that interstitials are dilatation centers with large relaxation volumes. Two interstitials approaching one another interact via these atomic displacement fields. The interaction energy can be calculated by means of Kanzaki forces simulating the defect-induced displacement field. For distances large compared to the range of the Kanzaki forces (a few atomic distances), the energy can be calculated by means of the Green's function of the elastic continuum. From a multipole expansion of the forces one obtains a leading term of the direct interaction, the dipole–dipole interaction or the so-called first order size interaction. Here, the interaction energy decreases with increasing distance, r, as r^{-3}. Its sign and magnitude depends on the direction of the defect-connecting line with respect to the axes of the elastic anisotropy. For cubic crystals and isotropic defects one obtains ESHELBY's formula [1956] as first-order term of a perturbation expansion with respect to the anisotropy parameter $d = c_{11} - c_{12} - 2c_{44}$:

$$E_{\text{int}} = -\frac{15}{8\pi} d \left(\frac{5}{3c_{11} + 2c_{12} + 4c_{44}} \right) \frac{\Delta V_1^{\text{rel}} \Delta V_2^{\text{rel}}}{r^3} \left(\frac{3}{5} - \sum_i \frac{r_i^4}{r^4} \right). \tag{9}$$

For $d > 0$ and interstitial-interstitial interaction, i.e., $\Delta V_1^{rel} = \Delta V_2^{rel} > 0$ we have an attractive interaction along <100> and repulsive ones along <110> and <111>. For $d = 0$ the dipole interaction vanishes but the multipole interaction does not. The influence of higher-order terms of the perturbation expansion was studied by DEDERICHS and POLLMANN [1972]. A numerical solution free of approximations was given by MASAMURA and SINES [1970].

Higher multipole interactions decrease with r^{-5} or even faster (SIEMS [1968], HARDY and BULLOUGH [1967]). Of special interest is the induced interaction based upon the polarizability of the defects. It depends on r as r^{-6} (SIEMS [1968]). For the interaction of two interstitials in Cu or Al with their high polarizabilities, the induced-interaction energy is comparable to the dipole interaction only for $r \leq 3$ nnd (TRINKAUS [1975]).

The above described relationships are certainly not applicable to interstitials approaching one another as close as is necessary to form multiple interstitials. This range of interaction has been studied by computer simulation (INGLE *et al.* [1981]). Quite detailed information on static and dynamic properties of di- and tri-interstitials for different Cu potentials were obtained. The stability of configurations was studied up to clusters consisting of 37 interstitials. The activation enthalpies of binding were found to be larger than the activation enthalpies of migration for all cases. Dissociation of multiple interstitials must therefore not be expected according to this fcc model. The relaxation volume change per clustering interstitial amounts to –5% to –15% of ΔV_i^{rel}. The formation enthalpy per interstitial decreases by about 30% when going from two to ten interstitials.

The stable di-interstitial consists of parallel dumbbells at nearest-neighbor sites, tilted

by a small angle (< 10°) in the <110> plane. Stable tri-interstitials are formed by mutually orthogonal dumbbells at nearest neighbor sites. Adding one atom in the octahedral lattice position to the tri-interstitial configuration yields the stable four-interstitial. Larger clusters arise from the four-interstitial cluster by adding further dumbbells such that each is equidistant from the central octahedral interstitial and is aligned orthogonally to its nearest-neighbor dumbbell. Such three-dimensional structures are more stable than two-dimensional ones for fewer than nine interstitials and are less stable than these for more than 13 interstitials. The stable two-dimensional clusters are platelets of octahedral interstitials on <111> planes.

The activation enthalpy of di-interstitial migration was found to lie below or above that of mono-interstitial migration, depending on the details of the potential. For the higher-order clusters, the migration enthalpies increase rapidly with the number of interstitials in the cluster.

The dynamic behavior of the di-interstitial is essentially similar to that for the mono-interstitial. A variety of resonant and localized modes occur. According to the resonant modes the thermal displacements of the dumbbell atom are nearly as large as for the single dumbbell. The change of the elastic constant C_{44} per interstitial is even larger than that of the single dumbbell. Although the quantitative results depend significantly on the details of the pair-interaction potential used for the simulation, the general dynamic behavior makes at least di-interstitials a very interesting object for experimental studies.

For information on interstitial loop formation from MDC studies see § 4.2.2.

3.3.2. Experimental methods and results

3.3.2.1. Relaxation volume. Introduction of an additional atom into a perfect finite crystal enhances the volume of this crystal by the relaxation volume of the self-interstitial, ΔV_i^{rel}. Because of possible relaxation anisotropies, interstitials of concentration c_i must be introduced randomly with respect to the anisotropy axes in order to produce the correct total volume change $c_i \Delta V_i^{rel}$. The method of determining ΔV_i^{rel} is given by the appropriate evaluation of the Huang scattering (EHRHART [1978]) or the diffuse scattering between the Bragg peaks (HAUBOLD and MARTINSEN [1978], §§ 3.2 and 3.3.2.2, also ch. 12). As the relaxation volumes of interstitial and vacancy enter quadratically into the Huang scattering intensity and ΔV_v^{rel} is considerably smaller than ΔV_i^{rel}, the former needs to be known only approximately. The experimental results are collected in table 4. The deviation of ΔV_i^{rel} from one atomic volume can be imaged as that additional volume which must be spent because of the perturbation of the regular packing of spheres by the inserted interstitial atom. Perturbation of the fcc structure obviously requires more additional volume than the less dense bcc structure.

3.3.2.2. Configuration. Experimental information on interstitial configurations comes from diffuse X-ray scattering and mechanical (§ 3.3.2.6) or magnetic relaxation experiments. The open question of the potential existence of two interstitial configurations in one and the same metal as postulated by the two-interstitial model generated a strong impact on the development of both the theory of the point-defect-induced diffuse X-ray scattering (DEDERICHS [1973]) and the measuring technique (EHRHART *et al.* [1974]).

References: p. 1710.

As the result of these efforts, the configuration of the interstitial in Al, Cu, Ni, Mo, Fe and Zn has been determined. For the fcc metals the <100> split interstitial, for the bcc the <110> split and for the hexagonal Zn the <0001> split interstitial has been found.

The analysis of diffuse X-ray scattering makes use of the information which is included in the intensity profile scattered outside the Bragg reflections due to the imperfect crystal. Here, the relevant imperfections are the atomic displacements around interstitial atoms and vacancies. An order-of-magnitude comparison of the different scattering contributions to the total diffuse scattering cross-section is shown in fig. 25, according to EHRHART *et al.* [1974]. The interesting intensity profile occurs superposed to a background which is one to two orders of magnitude larger than this. Information can be obtained only by a careful subtraction of the intensity scattered by an interstitial-free crystal from that of an interstitial-containing one.

Owing to the minor lattice distortions around a vacancy (§§ 2.1 and 2.2.2.2) the diffuse intensity caused by this defect is only a small correction to the interstitial contribution.

The diffuse intensity profile contains the symmetry of the atomic distortions as well

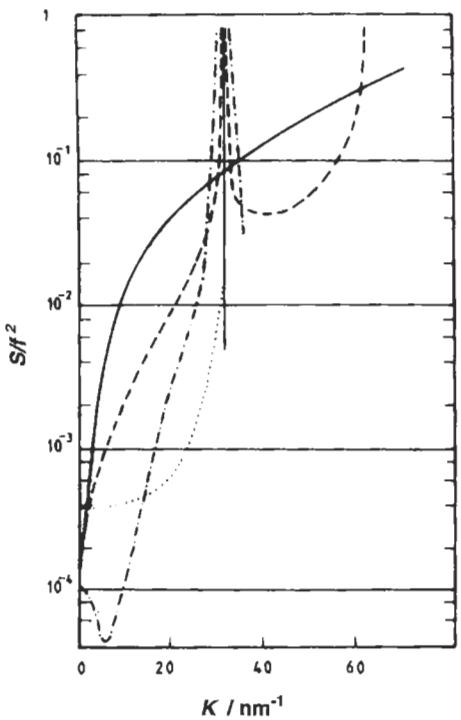


Fig. 25. Cross-sections for Compton scattering, thermal diffuse scattering (TDS) at 4 K and point-defect X-ray scattering by Al as a function of the scattering wave vector ($\lambda = 15.4$ nm) (after EHRHART *et al.* [1974]). Line presentations: — Compton scattering, --- thermal diffuse scattering at 4.2 K, -·-· contribution of <100> split interstitial and ... by the vacancy, both of concentration 5×10^{-4}.

as the distortion strength, i.e., the spatial extension of the distortion field. The unknown configuration is determined by comparing calculated intensity profiles with the measured ones. Figure 26 shows calculated scattering profiles for three different interstitial configurations in Al and the comparison with measured data. The agreement between measured and calculated data for the <100> split interstitial is convincing. The profiles for this configuration have been calculated with a distance between the two dumbbell atoms of 0.85 nnd. The four nearest neighbors of the dumbbell were found to be displaced outwards by about 0.07 nnd. The long-range atomic displacement field around lattice defects determines the scattering at small values of the scattering vector, q. Measurements close to the Bragg peaks (Huang scattering) can therefore resolve the symmetry of interstitials (EHRHART [1978]). Although high-angular-resolution techniques must be applied, the overall experimental expenditure and the requirements for sample quality are less exceptional than in the case of diffuse X-ray measurements between the Bragg peaks. This advantage originates from the much higher defect-induced scattering intensity. Huang scattering measurements are particularly powerful for point-defect clustering studies.

3.3.2.3. Formation enthalpy. The formation enthalpy of a Frenkel defect can be determined by measuring the heat release caused by annihilating Frenkel defects. In practice, Frenkel defects are generated in a sample by low-temperature irradiation. In order to obtain a reliable measure for the defect concentration generated, a resistivity sample is irradiated simultaneously. After sufficient irradiation, the calorimetric sample is then heated at a constant heating rate. The power input required for this heating condition varies as the stored energy releases. Different methods of measurement for this entity have been applied with varying success. One of the most serious problems in measuring stored energies in irradiated samples is the required decoupling of the sample from good thermal contact with a cooling stage during the irradiation to optimal thermal

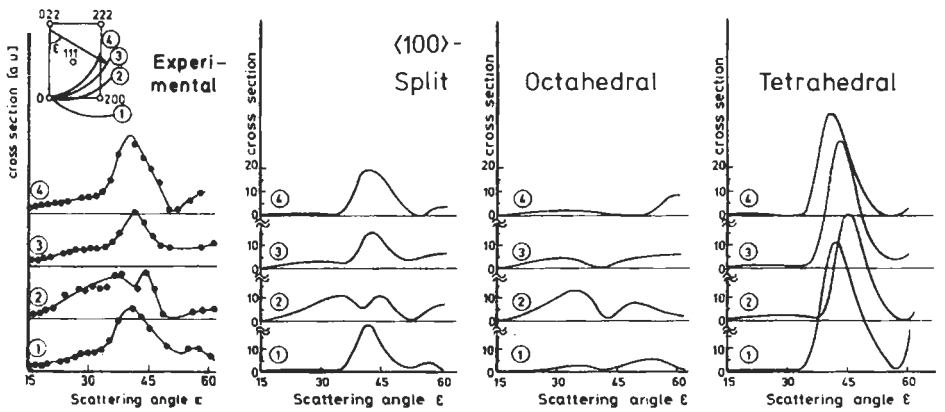


Fig. 26. Comparison of the diffuse X-ray-scattering intensity measured in Al electron-irradiated at 4 K with curves calculated for different interstitial configurations. The numbers 1 to 4 indicate the circles of the Ewald sphere shown in the insert (left upper corner) along which the intensity was measured and calculated. (after EHRHART *et al.* [1974].)

References: p. 1710.

isolation during the calorimetric measurements. This procedure must be performed at temperatures below 10 K (stage I recovery). Another very inconvenient condition is the continuous nonlinear increase of the heat capacity of the sample with increasing temperature. The total enhancement of the heat capacity exceeds one order of magnitude across the stage I recovery interval. The stored-energy release must be measured on this strongly varying background (review by WENZL [1970]).

Results of stored-energy measurements have generally been given in terms of the ratio stored-energy release per resistivity recovery $\Delta Q/\Delta\rho$. The ΔH_F^f values in table 5 were calculated from the reported ratios by multiplying them by the ρ_F values given in table 2. The enthalpy of interstitial formation has been calculated according to $\Delta H_F^f - \Delta H_v^f = \Delta H_v^f = \Delta H_i^f$, with the vacancy data from table 1.

3.3.2.4.Activation enthalpy of migration. Most experimental information on the migration enthalpy of interstitials comes from electrical resistivity recovery measurements performed after low-temperature irradiation. The recovery curve as found for Cu after electron irradiation at 4.2 K is shown schematically in fig. 5. The curve can be taken as a prototype for many metals, although the detailed structure slightly varies from metal to metal. Substantially different curves are observed for Au, Cd and V (EHRHART [1991] and SCHULTZ [1991]).

The generally agreed interpretation of the stage I recovery consists of the so-called close pair recombination in stages I_A–I_C and the free migration of the interstitial in stages I_D and I_E with interstitial-vacancy recombination upon random diffusion and eventual encounter with a vacancy. In close pairs vacancies and interstitials have such a strong elastic interaction that the thermally activated jumping interstitial has no chance to escape recombination with its correlated vacancy at temperatures not higher than stage I_C. For an extensive review on earlier experimental work in fcc metals and its interpretation, see SCHILLING *et al.* [1970]. The local correlation of those vacancies and interstitials which

Table 5
Activation enthalpy of Frenkel defect and self-interstitial formation[a].

Metal	ΔH_F^f (eV)	ΔH_i^f (eV)	Measurement[b]	Reference
Al	4.3	3.6	n	SCHILLING and TISCHER [1967]
	3.7	3.0	e^-	WOLLENBERGER [1965]
Cu	2.9	1.6	n	BLEWITT [1962]
	4.1	2.8	n	LOSEHAND *et al.* [1969]
	4.2	2.9	e^-	MEECHAN and SOSIN [1959]
	5.5	4.2	d^{++}	GRANATO and NILAN [1965]
	5.4	4.1	e^-	WOLLENBERGER [1965]
α-Fe	6.3	4.7	e^-	MOSER [1966]
	6.6	5.0	e^-	BILGER *et al.* [1968]
	13.6	12.0	n	BILGER *et al.* [1968]
Pt	2.9	1.5	e^-	FEESE *et al.* [1970]
	2.4	1.1	d^{++}	JACKSON [1979]

[a]Denoted by ΔH_F^f (Frenkel defect) and ΔH_i^f (self interstitial). For evaluation procedure see text.
[b]Stored-energy release and resistivity recovery measured after irradiation with electrons (e^-), neutrons (n) and deuterons (d^{++}), respectively.

were produced by one and the same displacement process (displacement distance r_p of a correlated Frenkel defect) must cause a recovery stage for which the significant average number of migrational jumps per interstitial is determined by r_p. For recombination with non-correlated vacancies the specific number of interstitial jumps depends on the concentration of the Frenkel defects. Hence, we expect two recovery stages which are well separated when the average distance of the Frenkel defect is large compared to r_p and overlap when both quantities are similar in size. These two stages are identified as I_D and I_E.

The curve in fig. 5 does not fall to zero in stage I, which means incomplete Frenkel defect recombination in stage I_E. Among many possible reasons, two are mainly effective in those cases represented by fig. 5: (i) formation of immobile di-interstitials, tri-interstitials and so forth, (ii) trapping of interstitials by solutes. A complete description of the stage I_E recovery must consider these reactions. A review on the theory of diffusion-controlled defect reactions has been given by SCHROEDER [1980].

The first systematic quantitative study of stage I_D–I_E recovery was performed by CORBETT *et al.* [1959] for electron-irradiated Cu by applying the numerical description of a diffusion-controlled model reaction scheme (WAITE [1957]) to suitably measured electrical resistivity recovery curves. Excellent agreement between measured data and theoretical prediction was obtained. The obtained ΔH_i^m values are collected in table 6.

Since then a large number of stage I resistivity recovery investigations have been performed. A critical review with respect to the experimental verification of interstitial configuration and migration models was given by YOUNG [1978]. The corresponding

Table 6
Activation enthalpy of interstitial migration[a].

Metal	ΔH_i^m	Measurement[b] (eV)	Reference
Ag	0.088	$\Delta\rho_{irr}$	RIZK *et al.* [1977]
Al	0.115	MR	SPIRIC *et al.* [1977]
	0.112	$\Delta\rho_{irr}$	RIZK *et al.* [1976]
Co	0.1	MA	DANDER and SCHAEFER [1977]
			SULPICE *et al.* [1968]
Cu	0.117	$\Delta\rho_{irr}$	CORBETT *et al.* [1959]
Ga	0.073	$\Delta\rho_{irr}$	MYHRA and GARDINER [1975]
α-Fe	0.3	MA	SCHAEFER *et al.* [1975]
Mo	0.083	$\Delta\rho_{irr}$	KUGLER *et al.* [1982]
Ni	0.15	$\Delta\rho_{irr}$, MA	KNOLL *et al.* [1974]
Pb	0.01	$\Delta\rho_{irr}$	BIRTCHER *et al.* [1974]
			SCHROEDER and SCHILLING [1976]
Pt	0.063	$\Delta\rho_{irr}$	SONNENBERG *et al.* [1972b]
			DIBBERT *et al.* [1972]
W	0.054	$\Delta\rho_{irr}$	DAUSINGER [1978]
Zn	0.015	$\Delta\rho_{irr}$	MYHRA and GARDINER [1973]
Zr	0.26	$\Delta\rho_{irr}$	NEELY [1970]
	0.30	MR	PICHON *et al.* [1973]

[a] Denoted by ΔH_i^m.

[b] See glossary of foregoing tables; in addition MR and MA indicate mechanical relaxation and magnetic after-effect, respectively.

References: p. 1710.

behavior of bcc metals has been reviewed by SCHULTZ [1987]. Complete data collection see EHRHART [1991] and SCHULTZ [1991].

As an example of data obtained by means of a rather elaborate measuring technique, fig. 27 shows differential isochronal resistivity-recovery curves for electron-irradiated Pt. The influence of the initial defect concentration (given by the initial residual resistivity $\Delta\rho_0$ and corresponding to about 10 at-ppm to 1000 at-ppm) becomes quite obvious. The quality of the fit of appropriate rate equation solutions to the stage I_D-I_E recovery is shown in fig. 28. The series of figures clearly shows the increasing weight of both uncorrelated recovery and suppression of recovery by di-interstitial formation with increasing initial defect concentration. The influence of interstitial-trapping solutes is shown by $i+f$ in the upper two figures.

Table 6 gives the migration enthalpies obtained by this type of investigations. The table also includes data obtained by elastic after-effect measurements in Al (SPIRIC *et al.* [1977]) and magnetic after-effect measurements in Ni (KNOLL *et al.* [1974]) and in Fe (SCHAEFER *et al.* [1975]). These properties, measured in single crystal samples at a fixed temperature after isochronal annealing periods at increasing temperatures, give information not only on the recombination rate but also on the symmetry and a possible

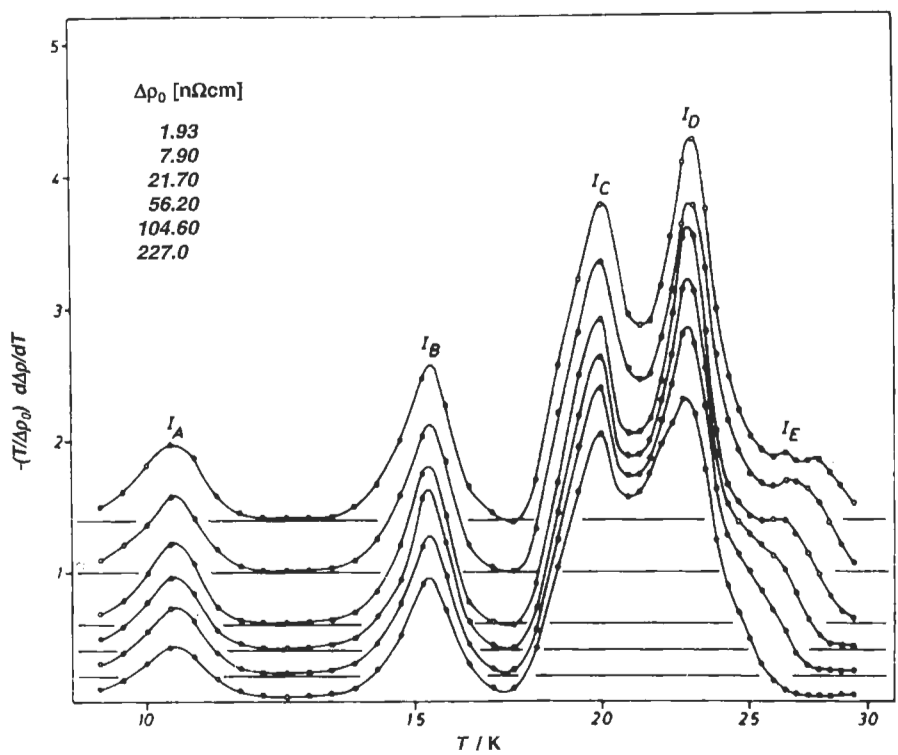


Fig. 27. Temperature-differentiated isochronal resistivity recovery curves of electron-irradiated Pt for different initial irradiation-induced resistivity increments $\Delta\rho$ (after DIBBERT *et al.* [1972]).

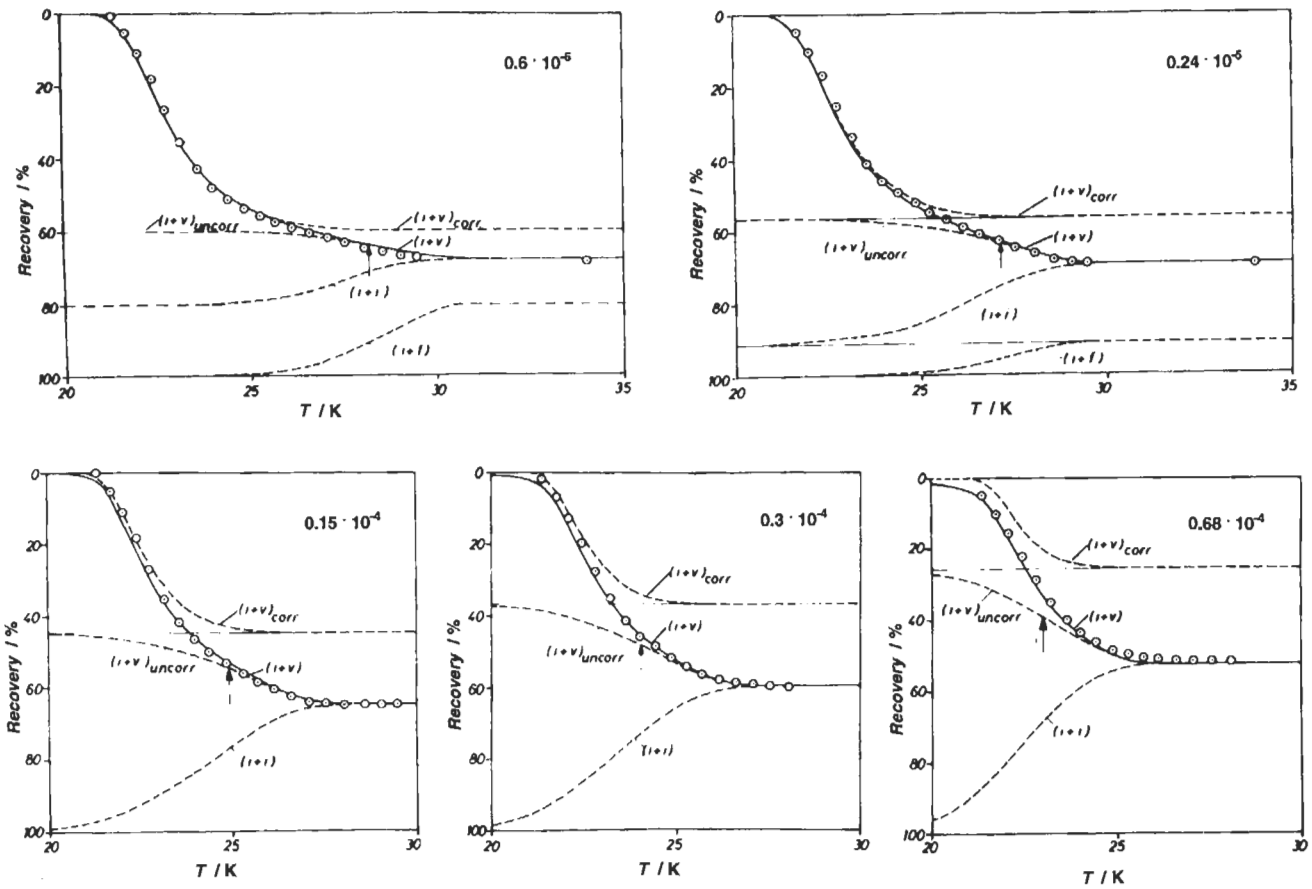


Fig. 28. Experimental resistivity-recovery data of electron-irradiated Pt for different initial Frenkel defect concentrations c_0, compared with calculated curves. The dashed curves give the fractional influence of the different point-defect reactions. (after SONNENBERG *et al.* [1972b].)

orientation relaxation of the migrating defect. In Al the results show that a defect with tetragonal symmetry reorients while migrating, until it recombines with a vacancy. This behavior is exactly predicted by the computer experiments with the combined translational and rotational jump of the <100>-dumbbell (§ 3.3.1). The annihilation kinetics of the after-effect follows closely that of the resistivity. At temperatures below stage I_D no reorientation relaxation has been observed which could be ascribed to free interstitials. Therefore, a pure dumbbell rotation without migration can be excluded, which is again in agreement with the theoretical prediction.

For Ni, reorientation relaxation of a defect with <100>-symmetry in stage I_D has been derived from magnetic after-effect measurements.

For α-Fe, magnetic alignment experiments have identified the <110>-symmetry of the interstitial (CHAMBRON *et al.* [1975]) and magnetic after-effect measurements have shown the reorientation relaxation during annealing in stage I_D (SCHAEFER *et al.* [1975]).

For other metals, including Au (DWORSCHAK *et al.* [1981]), Cd (COLTMAN *et al.* [1971], O'NEAL and CHAPLIN [1972]), Nb and Ta (SCHWIRTLICH *et al.* [1987]) and V (SCHULTZ [1987]), no recovery stages which could unambiguously be ascribed to correlated and uncorrelated recombination of freely migrating interstitials have been detected. On the other hand, long-range mobility of the interstitials below 5 K has been concluded from Huang scattering experiments which do show interstitial cluster formation during electron irradiation at 5 K in Au, Cd, Mg and Nb (EHRHART [1978]). Many other observations give more indirect hints of long-range interstitial mobility. In particular, Au has been studied with respect to this problem in great detail (DWORSCHAK *et al.* [1981] and SEEGER [1970]). A stage I_E, if it exists, must occur below 0.3 K (BIRTCHER *et al.* [1975], SCHROEDER and STRITZKER [1977]. The unexpectedly large displacement distances (DWORSCHAK *et al.* [1981]) at low temperatures could be due to the special coupling of interstitials to the vibrational lattice modes (§ 3.3.1.2).

As for Mo and W, recovery stages observed around 35 K for Mo (MAURY et al. [1978]) and 27 K for W (DAUSINGER [1978]) do show a temperature shift as function of the initial defect concentration. But the kinetics are completely different from what is observed in Cu, for example. According to SCHULTZ [1987], the observed kinetics would be in accordance with a model that allows not only for vacancy-interstitial recombination but also for interstitial trapping at certain neighboring sites of a vacancy. By this effect the amount of recovery in stage I_E will be reduced. At some higher temperature the close pair formed by the trapping event is allowed to recombine leading to the typical close-pair recovery stage, as indeed observed a few degrees above the quasi-I_E stage. Direct evidence for this interpretation is, still lacking.

3.3.2.5. The controversy on the interpretation of the recovery-stage III. The controversy resulted from the fact that the main recovery stage of *quenched* samples did not exactly coincide with stage III in *irradiated* samples (§ 2.2.3.3). Unfortunately, initially the difference was particularly large for Cu, the standard research material of the metal physicists. The lack of total recovery at the end of stage III also played a significant role in the dispute. The latter problem, however, is of less significance than the former and is omitted here.

The primary source of disagreement is the interpretation of the stage III recovery.

The advocates of the "one-interstitial-" or (a better expression) *vacancy model* explain stage III by diffusion-controlled annihilation of vacancies. The temperature difference of the recovery-rate maxima of isochronal annealing curves were ascribed to the different number of jumps the vacancies had to perform in order to reach the sinks which are of different density and configuration after quenching and irradiation, respectively (§ 2.2.3.3).

The advocates of the two-interstitial model solved the problem by ascribing stage III after irradiation to migration and annihilation of the stable interstitial, and by calling the first stage after quenching stage IV and ascribing it to migrating vacancies. The interstitial migrating in stage I (fig. 5) was thought to be a metastable configuration. Throughout a period of 20 years, the open question of the stage III-stage IV origin energized research in the point defect field. Considerable improvement of experimental methods was achieved in three directions: (i) determination of interstitial properties like configuration, dynamic behavior, migration, recombination and clustering behavior, (ii) improvement and understanding of the quenching procedure, especially with regard to impurity effects, (iii) detection of vacancy migration and clustering (PAC and PAS, see § 2.2.3.2). As an introduction to the controversy the reader is referred to SEEGER [1975], YOUNG [1978] and BALLUFFI [1978]. The two latter authors conclude that in many metals the experimental observations can satisfactorily be explained by the one-interstitial model, i.e., by the vacancy migration in stage III. The experimental basis of the papers by SCHINDLER *et al.* [1978] and FRANK *et al.* [1979], which was taken as a strong support of the two-interstitial model by the authors just at the time of YOUNG's [1978] and BALLUFFI's [1978] reports could not be confirmed later (LAUPHEIMER *et al.* [1981]).

In recent years, the intensity of the debate has diminished although the advocates of the two-interstitial model report on new deficiencies of the one-interstitial model and successful applications of the two-interstitial model (KIENLE *et al.* [1983], SEEGER [1982], SEEGER and FRANK [1983]). The general argumentation regarding interstitial properties is, as before, based on observations which cannot be explained by the well-known stage I interstitial properties directly. The second set of parameters introduced by the second interstitial of the two-interstitial model generally allows easy explanation of such observations.

The key question of the controversy is the consistent explanation of stage III by vacancy migration. There are many experimental results reported in the literature which favor vacancy migration in stage III (§ 2.2.3). As an example we take results for Au for which in specially pre-treated electron-irradiated samples $\Delta H_{eff}^{m} = 0.71$ eV was found in stage III (SONNENBERG and DEDEK [1982]). The treatment consisted of high-dose α–particle irradiations which resulted in a very high density of dislocation loops. This high loop density acting as a high fixed sink density for the migrating stage III defect suppresses all other possible reactions of this defect. The model fitted to equilibrium, quenching and self-diffusion data by SAHU *et al.* [1978] yielded $\Delta H_{1v}^{m} = 0.83$ eV. According to SEEGER and FRANK [1983] the differing migration enthalpies are the evidence for the migration of an interstitial in stage III with $\Delta H_{i}^{m} = 0.71$ eV. To the present author this conclusion, although suggestive, might be premature. The quenching data yield ΔH_{eff}^{m} values between 0.55 eV and 0.63 eV, depending on the initial vacancy concentration. The result $\Delta H_{1v}^{m} = 0.83$ eV is derived by fitting a model which assumes a

References: p. 1710.

substantial di-vacancy contribution. Di-vacancies are made responsible (i) for the curvature of the $c_v^0(1/T)$ curve which determines ΔH_{2v}^b and (ii) for the experimental $\Delta H_v^f = 0.62$ eV. This model is certainly not unambiguous with regard to its basic assumptions. The question whether the above curvature could be due to temperature dependence of ΔH_{1v}^f (VAROTSOS and ALEXOPOULOS [1982]), for example, must carefully be reinvestigated as well as the atomic processes during annealing of quenched resistivity samples. The electron-microscopic work on quenched gold clearly demonstrates the strong tendency for vacancy agglomeration. At the same time, there are reliable indications that the electrical resistivity does not change during the formation of small clusters containing not more than about five vacancies. Hence, resistivity recovery experiments can show only a later state of a complex vacancy-annihilation regime with certainly time-dependent sink concentrations.

3.3.2.6. Dynamic properties. The dynamic behavior of the self-interstitial is perhaps the property most sensitive to the configuration. As has been shown in § 3.3.1.2, the decrease of the elastic modulus with increasing interstitial concentration played a key role in the understanding of the nature of interstitial structure. After a number of modulus measurements in polycrystals which demonstrated the negative sign and the large absolute value of the modulus change (KÖNIG *et al.* [1964], WENZL *et al.* [1971]), the first measurements in single crystals containing randomly distributed single Frenkel defects have been reported by HOLDER *et al.* [1974] on thermal-neutron-irradiated Cu and by REHN and ROBROCK [1977] in electron-irradiated Cu. The data obtained by the former authors are shown in fig. 29, indicating the largest effect, for c_{44}. An interstitial concentration of 300 at-ppm softens the crystal by $\Delta c_{44}/c_{44} \approx 1\%$. If the concentration of single interstitials were to be enhanced up to 3 % (limiting factor: spontaneous recombination, see § 3.1.2) and the softening proceeded linearly, the crystal would be unstable against a <100> shear stress ($c_{44} = 0$). Quenched-in vacancies were found to produce much smaller changes of the elastic constants (FOLWEILER and BROTZEN

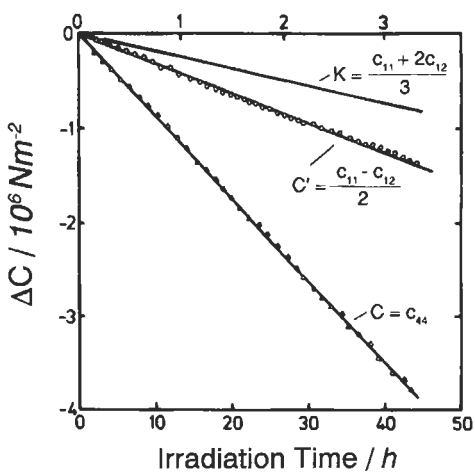


Fig. 29. Elastic moduli of Cu during electron irradiation at 4 K (after HOLDER *et al.* [1974].)

[1959]). The sequence of decreasing softening for C' and K can be understood in terms of the dumbbell interstitial. In the case of <110> shear stress which measures $C' = (c_{11}-c_{12})/2$, the bending spring of the dumbbell coupling is not activated, as can be seen in fig. 24, but the bending springs of the interaction with the nearest neighbors of the dumbbell are operating. As the equilibrium distance is larger than that between the dumbbell atoms the force-constants are smaller and hence the softening is less pronounced. This effect is even smaller for hydrostatic compression (bulk modulus κ). Data for Al are given by ROBROCK and SCHILLING [1976]. In the bcc Mo the largest softening has been observed for C' (OKUDA [1975]), which shows the <110>-symmetry of the split interstitial in this metal.

The large change of the elastic constants should significantly influence the phonon dispersion curves (WOOD and MOSTOLLER [1975], SCHOBER *et al.* [1975]). Such an effect has indeed been observed by NICKLOW *et al.* [1979]. SCHOBER [1989] has pointed out that irradiated crystals with their self-interstitials show many features which are known as low temperature glass anomalies (low-frequency vibrations, relaxation processes and general two-level-system behavior). GRANATO [1993] proposes the self-interstitial to be the crucial defect which determines all thermodynamic properties of each of the crystalline, liquid or amorphous states, as well as their changes upon melting.

3.3.2.7. Interstitial agglomeration. The formation of small multiple interstitials, even di- or tri-interstitials, has been concluded from measurements of the diffuse X-ray scattering, Huang scattering (§ 3.3.2.2) and mechanical relaxation measurements (§ 3.3.2.6). In fig. 30 diffuse X-ray scattering measurements on electron-irradiated Al subsequently annealed to 40 K are compared with calculations for two di-interstitial models. Obviously the parallel dumbbell di-interstitial is formed in Al at the end of

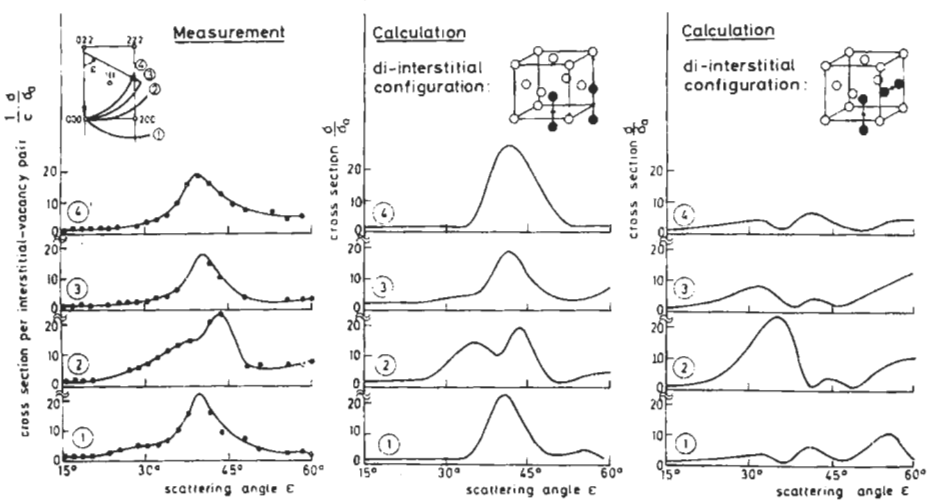


Fig. 30. Comparison of the diffuse X-ray-scattering intensity measured in Al electron-irradiated at 4 K and subsequently heated to 40 K, with calculated curves for two different di-interstitial configurations. Numbers 1 to 4 as in fig. 25. (after EHRHART *et al.* [1974].)

recovery stage I. This configuration was predicted by computer simulation for a Cu potential (§ 3.3.1.4).

For the Huang scattering intensity the formation of clusters consisting of n interstitials leads to $(c_d/n)(N\Delta V_i^{rel})^2 = nc_d(N\Delta V_i^{rel})^2$ instead of $c_d(N\Delta V_i^{rel})^2$ for single interstitials of concentration c_d. This enhancement is predicted for a linear superposition of the atomic displacements in the cluster. The effect of the actual relaxation upon clustering seems, however, to remain smaller than the gain by the factor n (EHRHART and SCHILLING [1974]). The cluster formation can easily be monitored, as the wavevector dependence of the intensity according to q^{-2} is substituted by a q^{-4} dependence in the upper q range. From the boundary of the q^{-2} validity the cluster radius can be estimated. For Al, with the same experimental conditions as in fig. 26, $n=2$ was found in agreement with the above result. Between 40 K and 45 K, n was found to increase from 2 to 4 (see fig. 31). For Cu it has been found that already at the end of recovery stage I, small loop-shaped clusters with $n=5$–10 exist. The lack of clusters smaller than with $n=5$ leads to the conclusion that in Cu, di- and tri-interstitials are mobile in the same temperature region as the mono-interstitial is. This behavior was also found for di-interstitials by the computer simulations (§ 3.3.1.4). More information on small interstitial clusters in Au, Cd, Co, Nb, Mg, Zn and Zr from Huang scattering is given by EHRHART and SCHÖNFELD [1982]. The strong interstitial clustering found for Au already at 10 K (fig. 31) is in accordance with the interstitial mobility already found at 0.3 K (BIRTCHER *et al.* [1975]. In fig. 32 we see the strong reduction of the interstitial clustering in dilute alloys when compared with the pure solvant.

In pure metals, the interstitial clustering produces the characteristic feature of the stage II recovery as seen in fig. 33. Similar behavior is found for material cold-worked at low temperatures (fig. 34). This continuous resistivity decrease with approximately

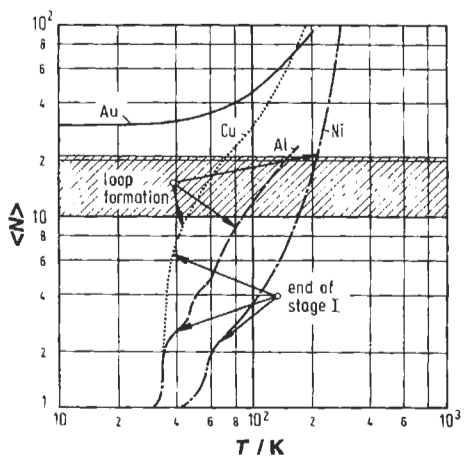


Fig. 31. Average interstitial cluster sizes as function of the annealing temperature for Au, Cu and Ni and initial defect concentration of about 300 ppm (after EHRHART *et al.* [1986]).

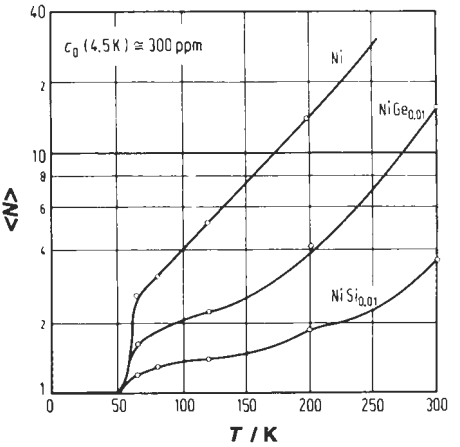


Fig. 32. Average interstitial cluster sizes for dilute Ni alloys as function of the annealing temperature after 3 MeV electron irradiation to 300 ppm defect concentration (after EHRHART [1991]).

constant rates through relatively large temperature ranges (nearly two orders of magnitude for Au, DWORSCHAK *et al.* [1981]) must be due to interstitial rearrangements rather than due to recombination. Constant recovery rates mean "white" spectra of activation enthalpies for the rate controlling processes, in the case of Au ranging through nearly two orders of magnitude. Although this observation involves substantial information

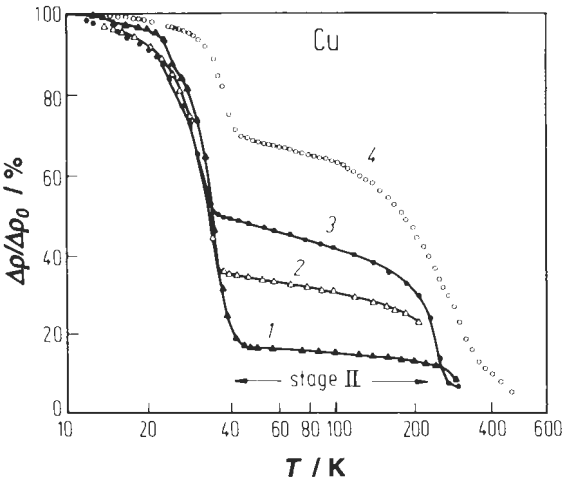


Fig. 33. Isochronal recovery (10 min annealing period) of the electrical resistivity of electron (1–3) and neutron (4) irradiated Cu. Radiation-induced resistivity increase: 1–0.157 nΩcm, 2–3.03 nΩcm, 3–5.9 nΩcm, 4–0.0186 nΩcm (after EHRHART [1991]).

References: p. 1710.

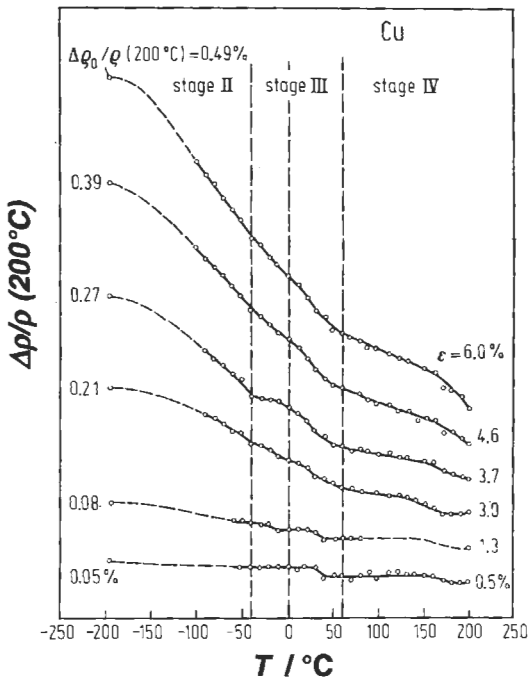


Fig. 34. Isochronal recovery of the electrical resistivity for Cu plastically deformed at –195°C to different strains ε. The given data are normalized by ρ at 200°C (after VAN DEN BEUKEL [1970]).

about cluster mobility and reactions (rearrangements), this urgently needed information (see § 4.2.2) could not be utilized yet because of the unknown resistivity change per interstitial in the various cluster arrangements.

3.3.2.8. Interstitial-solute interaction. Comparison of the stage I recovery of pure metals with that of impure ones showed that solutes trap migrating self-interstitials within certain temperature ranges (SCHILLING *et al.* [1970]). If this trapping was due to a long-range interaction between solutes and interstitials this interaction should be of elastic nature, because of the effective screening of charges by the conduction electrons. The interaction potential would then be described to a first approximation by eq. (9). The overall effect of solutes on migrating interstitials is an attractive one, despite the existence of repulsive directions (PROFANT and WOLLENBERGER [1976]). The drift-diffusion of interstitials in the interaction potential has been treated in detail by SCHROEDER [1980].

The long-range interaction causing drift diffusion cannot reliably be applied at distances between solute and interstitial of three atomic distances or less. Answers to questions such as the configuration of an interstitial placed close to a solute or the binding energy of this pair must be derived from microscopic calculations. A first attempt was made by DEDERICHS *et al.* [1978], who simulated a solute in Cu in computer

experiments by simply shifting the Cu–Cu interaction potential by a certain distance R_0. Enlarging the equilibrium distance of the neighbor simulates oversized solutes, diminishing it simulates undersized solutes. The model predicts a number of interesting results. For undersized solutes it shows the formation of a mixed dumbbell. Solute and one solvent atom share one lattice position in the <100> split configuration. By oversized solutes the self-interstitial dumbbell is trapped at different neighboring lattice positions of the solute atom, with different binding energies, of the order of magnitude of 0.1 eV.

The interaction energies in saddle-point configuration of the self-interstitials have not been reported. The saddle-point energies for different types of jumps of the mixed dumbbell atoms have been investigated with the result shown in fig. 35. The three different kinds of jumps are illustrated in fig. 36. As $\Delta H_i^m = 0.1$ eV, activation enthalpies for cage motion are very small according to fig. 35. For volume size misfits larger than about 10%, the mixed dumbbell should be able to rotate before dissociation occurs. Since subsequent rotational and caging jumps allow long-range migration of the mixed dumbbell (see fig. 36) activation enthalpies of migration smaller than that of the vacancy are to be expected. Such conditions are indeed found experimentally with considerable practical implications for solute transport (§ 4.7).

Similar calculations have been performed for different solutes in Al by using interionic potentials derived from first principles (LAM *et al.* [1980, 1981]). The pair potentials for the Al–Al and Al-solute interaction extended up to the nineth neighbor. For the solute Zn, mixed dumbbell formation was found with a binding energy of 0.38 eV.

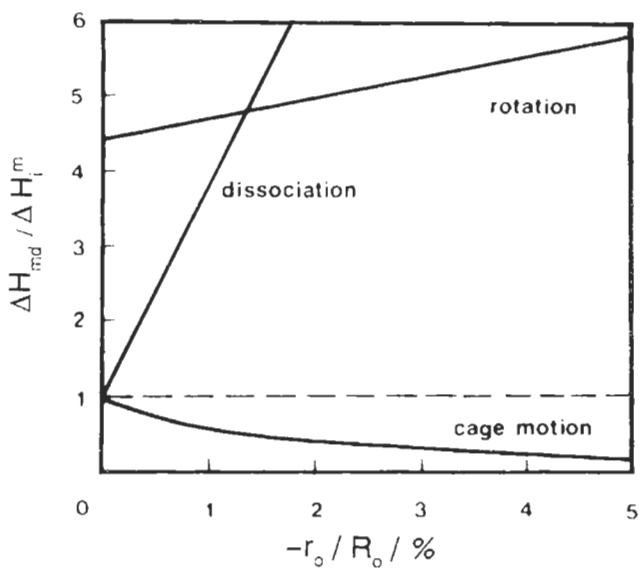


Fig. 35. Normalized enthalpies of mixed dumbbell dissociation, rotation and cage motion according to computer simulation of Cu (DEDERICHS *et al.* [1978] versus size-misfit. The ratio r_0/R_0, of the potential shift r_0) divided by the equilibrium atom distance R_0 of Cu must be multiplied by about six in order to obtain the volume size-misfit comparable with experiments.

References: p. 1710.

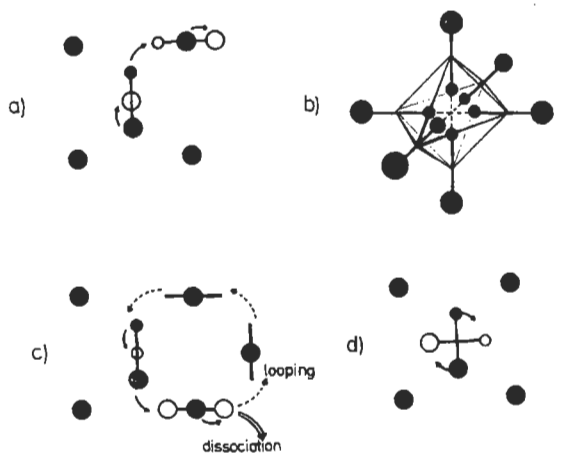


Fig. 36. Elementary jumps of a mixed dumbbell: (a) Jump of the solute; (b) by cage motion of the solute around the octahedral site the six mixed dumbbells shown are formed consecutively; (c) jump sequence of solvent atoms and solute which in case of looping leads to migration steps of the solute; (d) rotation (from DEDERICHS *et al.* [1978]).

For Be, Ca, K and Li, maximum binding energies between 0.62 eV (K) and 1.75 eV (Li) were found for different configurations but in no case a split configuration with solvent atom and solute sharing a lattice site was found. For Mg no binding was found. The saddle-point energy for mixed-dumbbell migration was found to be much larger than this dissociation energy.

The experimental situation of the solute-interstitial interaction has been reviewed by WOLLENBERGER [1978]. Since then, a number of observations have been reported which indicate the existence of quite a number of different solute-interstitial complex configurations with different physical properties in one and the same irradiated dilute alloy. These results are reviewed by ROBROCK [1983].

Trapping of migrating interstitials by solutes prevents the interstitials from recombining with vacancies. This can clearly be seen in the isochronal resistivity recovery curve, as examplified in fig. 37. The presence of 112 at-ppm Au in Cu suppressed recovery in stage I_E. The recovery is retarded and occurs stepwise around 60 K and between 90 K and 120 K. Beyond this the resistivities of dilute alloy and pure solvent are equal, indicating that the Au atoms no longer trap interstitials. By fitting diffusion models to the recovery curves in the stage I_D–I_E range the reaction, capture or trapping radius of the solute for migrating interstitials is obtained. Appropriate evaluation of the final recovery stage yields the activation enthalpy for dissociation of the complex. Activation enthalpies obtained this way for different solutes in Cu are shown in table 7.

The final recovery stage could also be caused by the solute-interstitial pairs becoming mobile as a whole and migrating until the interstitial recombines with a vacancy. Then the activation enthalpy derived from this recovery stage would characterize the saddle point for the complex migration. The resistivity recovery experiments do not discriminate

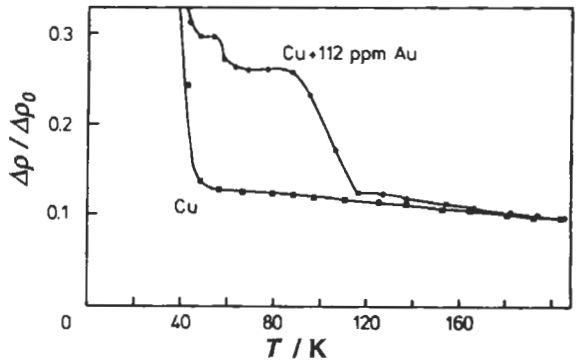


Fig. 37. Isochronal resistivity recovery of Cu and dilute Cu–Au electron-irradiated at 4 K (after CANNON and SOSIN [1975]).

between these two types of reactions. This question is particularly interesting for the solutes Co, Fe, Ge, Ni and Zn in Cu for which only an upper limit for the binding energy can be deduced. In these cases trapping was not observed at all. This might be due to migration of the complex already at the lowest temperature investigated (55 K).

Solutes in Al seem to cause even stronger binding since the recovery curves of dilute alloys do not coincide with that of the pure Al before stage III recovery, i.e., before vacancies start moving (WOLLENBERGER [1975], DWORSCHAK *et al.* [1976]). The quantitative evaluation of resistivity measurements with respect to reaction radii is somewhat handicapped by the fact that the resistivity does not give any hint on defect reactions occurring if they are not accompanied by resistivity change. But whenever the resistivity decreases, we simply interpret the change as a result of recombination. Reactions which must carefully be considered in this respect are formation of mobile or immobile di-interstitials which seem not to be accompanied by a measurable resistivity change but do change the defect reaction scheme significantly. Configurational changes of interstitial-solute pairs, on the other hand, could be accompanied by resistivity changes. Careful choice of the experimental conditions can to some extent help to exclude undesired reactions. The significance of di-interstitial formation in this respect has recently been discussed in some detail by MAURY *et al.* [1980].

The simplest reaction scheme in view of the solute-interstitial trapping is obtained by electron irradiation at temperatures corresponding to the recovery stages I_E and II. A large number of capture radii have been determined by resistivity-change rate measure-

Table 7
Activation enthalpies of self-interstitial solute pair dissociation [a].

Metal	Ag	Au	Be	Co	Cr	Fe	Ge	In
ΔH^{diss} (eV)	0.40	0.33	>1	<0.15	0.36	<0.20	<0.20	0.44
Metal	Mg	Mn	Ni	Pd	Sb	Si	Ti	Zn
ΔH^{diss} (eV)	0.43	0.32	<0.15	0.30	0.32	0.44	0.40	<0.20

[a] Denoted by ΔH^{diss}, values according to WOLLENBERGER [1978].

ments in the stage II recovery range. For Cu and Al as solvents, the so determined capture radii of nearly all solutes are found to depend on temperature as T^{-2}. If the capture radii were determined by drift diffusion of the interstitial in the long-range interaction field, the potential energy should depend on distance R as $R^{-1/2}$ in order to explain the T^{-2} dependence (SCHROEDER [1980]). No physical reason can be given for such a long-range potential. As for the recombination radius, a $T^{-1/3}$ dependence was found (LENNARTZ *et al.* [1977]), as expected for the elastic dipole interaction. The curious T^{-2} dependence might arise from the fact that the trapped interstitials are detrapped from varying well depths during the period of measurement. By assuming the appropriate number of trapping sites and interstitial transition rates between them, the T^{-2} dependence was quantitatively explained for Au in Cu as an example (ABROMEIT and WOLLENBERGER [1983]).

The absolute size of the effective capture radii is derived in the order of a few atomic distances for Cu, and for Al as solvent about twice as large. The sizes found in Cu disagree with the elastic interaction theory by a factor of about two. For Al, however, which is nearly isotropic, there is a disagreement of approximately one order of magnitude. The disagreement in the case of Al indicates that this type of theoretical description is not correct. It should be noted that diffusion jumps of the dumbbell interstitial in the vicinity of a solute have not been investigated by computer experiments until now.

While resistivity measurements have yielded substantial information on the trapping behavior of solutes in terms of capture rate constants, methods such as ion-channeling measurements, Mössbauer effect analysis, perturbed γ–γ angular correlation and mechanical relaxation measurements have extended our knowledge of configurations and motion characteristics of solute-interstitial pairs.

The *ion-channeling method* has been reviewed by SWANSON *et al.* [1978] and HOWE and SWANSON [1982]. It is the most direct method for localizing impurity atoms in a host lattice. Whenever solutes are displaced from regular lattice sites by trapped interstitials, ion channeling, in principle, can detect them and measure the displacement distance. Most of the solute–interstitial pair investigations have been performed by means of Rutherford backscattering of channeled ions. This method has the limitation that it has sufficient sensitivity only for solute masses large compared to the solvent mass. Consequently this method has been applied mainly to dilute Al alloys. The existence of <100> mixed dumbbells has been concluded for the investigated alloys with undersized solutes and no solute displacement was found for oversized ones. By using proton-induced X-ray emission as an indicator for the presence of impurities in interstitial positions (ECKER [1982]) the mass difference problem is avoided.

It should be kept in mind that channeling measurements can unambiguously identify the interstitial position of a solute only when the solute does not occupy more than one type of position. Data evaluation in case of simultaneous presence of more than one position becomes ambiguous and in practice very uncertain for even one of the assumed positions. Besides the different configurations possible for the solute-one-interstitial complex, the trapping of more than one interstitial per solute could lead to different interstitial positions of the solute. The latter problem can be solved, in principle, by

careful investigations of the defect-concentration dependence of the solute position (MÜLLER [1986]).

The *mechanical relaxation methods* measure the damping of applied elastic vibrations caused by the stress-induced reorientation of solute–interstitial complexes (see reviews by KRONMÜLLER [1970], NOWICK and BERRY [1972]). By investigating single crystals the symmetry of the complex can be derived. Complexes of different configurations are expected to show different activation enthalpies for reorientation and migration. For In in Cu, for example, not less than six different damping maxima are obtained (KOLLERS *et al.* [1981]), pointing to six different complex configurations. Four of them annihilate close to the final resistivity recovery stage of this alloy (DWORSCHAK *et al.* [1978]) with small differences in annihilation temperature among each other. The remaining two annihilate at considerably lower temperatures. A similar multiplicity of reorientation-damping peaks and annihilation temperatures has been found in dilute Al alloys (ROBROCK [1982], HULTMAN *et al.* [1981]). It has also been shown that the activation enthalpy for re-orientation significantly depends on the solute species (GRANATO *et al.* [1982]).

For Fe in Al a re-orientation process occurs at 8 K with an activation enthalpy of about 13 MeV (REHN *et al.* [1978]). It has been identified to arise from the cage motion detected earlier by Mössbauer studies (see below). Some disagreement regarding details of the cage geometry as concluded from internal friction (ROBROCK and SCHOBER [1981]) and Mössbauer spectroscopy (PETRY *et al.* [1982]) shows progress and limits reached with the application of these methods.

By *Mössbauer effect studies*, the *cage motion* of a solute which has trapped a self-interstitial has been detected. In a dilute alloy of ^{57}Co in Al or Ag irradiated at low temperatures and subsequently annealed through recovery stage I or irradiated at temperatures within stage II, a second Mössbauer line of the ^{57}Co (^{57}Fe) transition was observed besides the well-known one arising from the solute dissolved substitutionally (VOGL and MANSEL [1975]). From the isomer shift of the new line it must be concluded that the Mössbauer isotope is placed closer to the next solvent atom than a nearest-neighbor distance. Hence, it must be situated in the close vicinity of an interstitial atom. The Debye–Waller factor f of this new line shows an unusual temperature dependence, as shown in fig. 38. It decreases between 10 K and 20 K (for Al) by about 80%. This decrease can be measured reversibly so long as the sample is not heated into recovery stage III (Al) or to 100 K (Ag), i.e., until the trapped interstitials are annihilated. It has been quantitatively explained by a cage motion of the Mössbauer isotope in the cage shown in fig. 36b (VOGL *et al.* [1976]). The essential origin of the reversible Debye–Waller-factor drop is a locally restricted motion of the Mössbauer isotope within the lifetime of the Mössbauer quanta (10^{-7} s). The low-temperature range of its occurrence fits well to the theoretical result in fig. 35.

Mössbauer studies have also demonstrated the existence of more than one solute-interstitial-complex configuration. In irradiated Mo-^{57}Co five different new Mössbauer lines were resolved (MANSEL *et al.* [1982]).

The PAC technique introduced in § 2.2.3.2 has also been applied to questions of interstitial-solute interaction (§ 2.2.3.2). Again, a variety of different configurations of

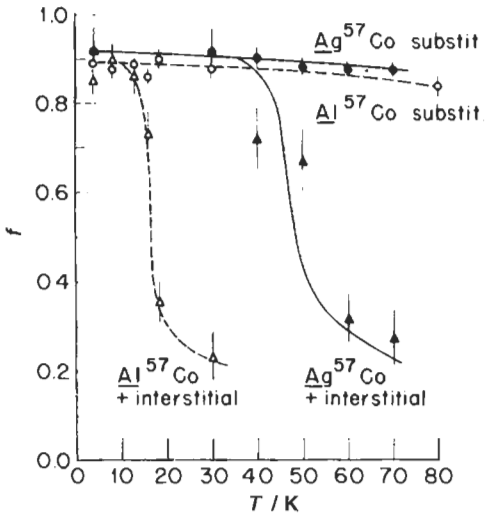


Fig. 38. Debye–Waller factor as derived from Mössbauer effect measurements in ^{57}Co in electron-irradiated Al and Ag. Irradiated and unirradiated states are indicated with "+ interstitial" and "substit.", respectively. (after VOGL *et al.* [1976].)

interstitial probe atom pairs must be concluded from the results for certain dilute alloys.

The application of probe atom methods still seems to possess considerable potential. A systematic application to single crystal samples could give more information on the symmetry of the complexes and systematic studies of the kinetics could give valuable information on the interstitial population of the complexes.

4. *Miscellaneous radiation effects*

4.1. Introductory remarks

The information collected in § 3 results from research which was to substantial degree carried out in order to solve problems of nuclear technology. Therefore, it seems to be reasonable to extend the considerations in § 3 beyond the limits of "simple" conditions for the defect production and the properties of single defects or small clusters towards more complex conditions common for technical processes. In § 4.2 the defect production by neutron and heavy ion irradiation is treated with special emphasis to the cascade phenomenon and its consequences for the formation of defect clusters and single defects. The role of the electronic interaction of energetic ions with metallic materials is described in § 4.3 by looking at the influence of incident swift heavy ions on the stability of Frenkel defects. In §§ 4.4 and 4.5 the *dimensional instability* of materials under high fluence irradiations at elevated temperatures is treated. This phenomenon is one of the most serious problems in the technology of fast fission and a prospective fusion reactor. The multiple reaction regime built-up by the defect production, agglomer-

ation, trapping, and annihilation allows for self-organization of vacancy clusters. The boundary conditions for this reaction mode are treated in § 4.6. The coupling of defect fluxes with atom fluxes leads to substantial locally varying compositional changes and eventually to phase transformations. This topic is treated in § 4.7.

4.2. Frenkel defect production by neutrons and ions

4.2.1. Atomic displacement cross-section and cascade formation

Electron irradiation is the basic method of Frenkel defect production when properties of single Frenkel defects are to be studied. Nuclear reactor technology introduced neutron-irradiated materials and raised the question about the underlying damage structure. Neutrons are "heavy" projectiles as compared with electrons. The energy transfer to target atoms for head-on collisions is given by the classical formula $T_{max} = 4MmE/(M+m)^2$, where E is the projectile energy, and m and M are the masses of projectile and target atom, respectively. Reviews on radiation-solid interaction are by LEHMANN [1977], LEIBFRIED [1965] and SEITZ and KOEHLER [1956]. For Cu we have $T_{ma} \approx E/16$ and the average recoil energy $<T> = 5 \times 10^4$ eV for 1.6 MeV fission neutrons. This energy is three orders of magnitude larger than the displacement threshold energy (§ 3.1.2), which leads to damage features completely different from those of electron irradiation. Before looking more closely at details of this type of damage, we realize that bombarding a sample with ions of the sample material (called self-ion irradiation) having 5×10^4 eV incident energy must have a similar effect as the fission neutron irradiation. An incident ion behaves like a lattice atom recoiling from a fission neutron collision. The only difference being: the lattice atom starts from a lattice site while the incident ion moves interstitially. This difference is of importance only when the primary collision is to be considered. For the secondary and all subsequent collisions the nature of the projectile is irrelevant.

Self-ion irradiation is highly welcome as a form of simulation for neutron irradiation because any measurement on ion irradiated samples is much easier than on often highly radioactive neutron irradiated ones (ISHINO *et al.* [1977]; GABRIEL *et al.* [1976]; further reading: STIEGLER [1976], eds. W. V. GREEN *et al.* [1990], SINGH *et al.* [1990]). Moreover, ion irradiations allow one to achieve given fluences (damage levels) in much shorter times than neutron irradiations do. This acceleration of the defect production shifts (among others) the dominance of the defect annihilation at sinks towards that by recombination and may thus impair correct simulation of more complex processes like swelling (§ 4.4) and phase transformations (§ 4.7). In practice, a self-ion energy of 5×10^4 eV would not be appropriate for damage production in the bulk of samples as the penetration depth of such ions measures only a few atomic distances. With self-ions of 0.3–1 MeV energy, samples for electron microscopy (100–200 nm thickness) or electrical resistivity measurements can be homogeneously damaged. If even thicker samples are required, as for mechanical property measurements for example, neutron irradiation is simulated by high-energy light-ion irradiation. Light ions are protons, deuterons and α-particles and applied energies are 20 MeV, for example. For such projectiles the penetration depth is of the order of 10–100 μm. The induced recoil energy spectra of the

target atoms are then different from those caused by reactor neutrons. Simulation is no longer straightforward.

Lattice atoms which obtained recoil energies as large as 100 keV act themselves as projectiles in the lattice and release their primary recoil energy by collisions with other lattice atoms. In this way a primary knock-on-atom initiates a *collision cascade* which, according to the transferred energies, leads to a displacement cascade which, in turn, ends up in a defect cascade. The ranges of Cu^+ ions with energies 5×10^4 eV, 2×10^4 eV, and 5×10^3 eV in Cu are 65 nnd, 31 nnd, 13 nnd, respectively. Hence, the density of collisions and defects strongly increases towards the end of the cascade.

As long as the collisions can be treated as binary collisions the process can analytically be described by a linear Boltzmann-type transport equation, as was originally done by KINCHIN and PEASE [1955]. These authors obtained the ratio $n=T/2T_d$, where T is the energy of the primary knock-on atom displacing n atoms in a cascade with recoil energies larger than the displacement threshold T_d. This treatment can only be a first-order approximation as the binary collision model breaks down as soon as the recoil energies become of the order of 10^4 eV. The multiple-collision events in this energy region form so-called *displacement spikes*, illustrated in figs. 39a, b. The corresponding result for n as obtained from computer simulation reads $n=0.8\ \varepsilon\ (T)/2T_{d,eff}$ for $T\geq 2.5T_{d,eff}$, $n=1$ for $T_{d,eff}\leq T\leq 2.5T_{d,eff}$ and $n=0$ for $T<T_{d,eff}$, where $\varepsilon(T)$ is the damage energy, which is the total energy T diminished by the losses due to electron excitations (NORGETT *et al.* [1974], § 4.3). These excitations result from the Coulomb interaction of the moving ion with the electrons of the solid. By electron–phonon interaction the energy is lost as heat. In Ni, for example, the excitation losses become appreciable around 10^4 eV and measure about 60% of the total energy at 10^6 eV. By taking into account the number of Frenkel defects produced per primary recoil atom, we obtain for the total displacement cross-section:

$$\sigma_d(E) = \int_{T_{d,eff}}^{T_{max}(E)} \frac{d\sigma(E,T)}{dT}\, n(T)dT, \tag{10}$$

where E is the energy of the incident particles (neutrons or ions) and $d\sigma/dT$ is the differential cross-section of a target atom to obtain energies T by collisions with the incident particles, and $T_{d,eff}$ an effective threshold energy suitably chosen to replace the anisotropic threshold-energy surface by a spherical surface (§ 3.1.2). For neutrons and ions the differential cross-sections are entirely different functions of T.

The validity of eq. (10) for ion irradiation on Cu has extensively been studied by means of electrical resistivity measurements for a wide range of recoil energies by AVERBACK *et al.* [1978]. The important finding is that for 4.2 K irradiation temperature the quantity $\sigma_{d,meas}/\sigma_{d,calc}=\xi$, called the *efficiency*, falls from unity for recoil energies around 10^2 eV to 0.5 around 10^3 eV and to 0.35 above 10^3 eV. As the effective threshold $T_{d,eff}$ for Cu amounts to about 50 eV (JUNG [1991]) n values around 1 and 10 correspond to the above quoted recoil energies of 10^2 and 10^3 eV. The decrease of ξ below unity already from $n=1$ on indicates a substantial defect loss in reality. The sequence of images in fig. 39 shows that also in the computer model the number of Frenkel defects is strongly reduced by recombination as the collision cascade decays.

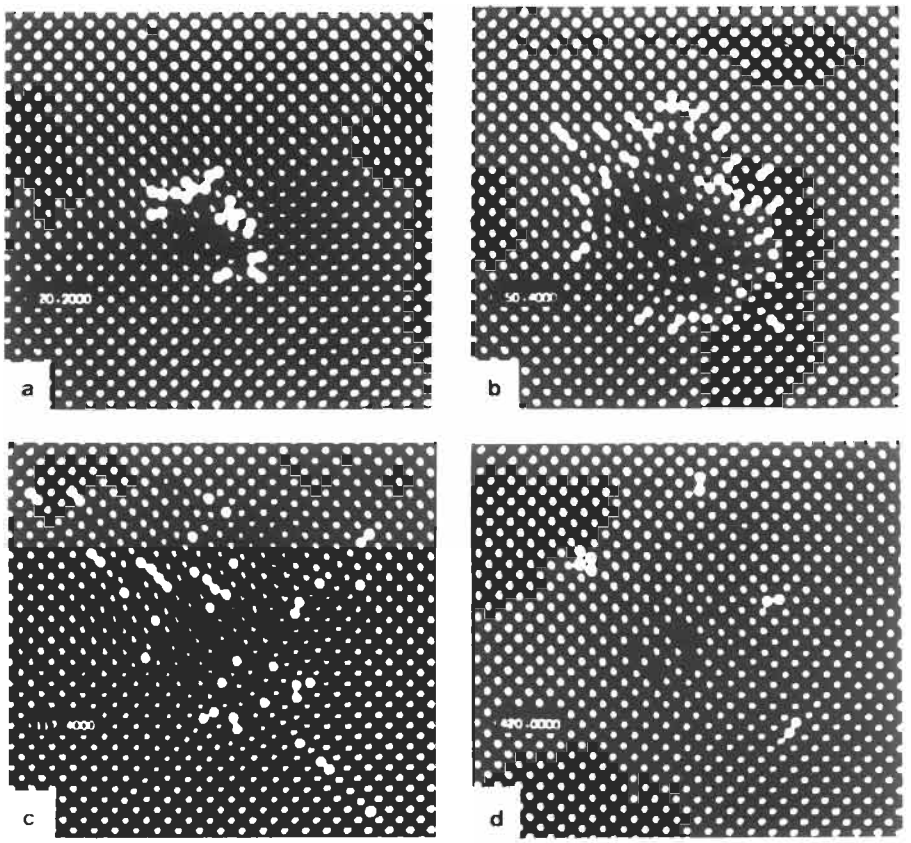


Fig. 39. Computer simulation of a displacement cascade with $T = 25$ eV in Cu. Shown is the {100} plane. The primary recoil impact direction was <110> given to the atom at that now vacant site in the cascade area which lies closest to the left lower corner of fig. 39a at $t = 0$. Atoms imaged by bright dots have more than 0.5 eV. In (b): $t = 6 \times 10^{-14}$ s (real time); in (c): $t = 3.5 \times 10^{-13}$ s; in (d), when $t = 1O^{-12}$ s a stable configuration is reached, consisting of five Frenkel defects. (from LEHMANN *et al.* [1981].)

This process has been studied in some detail by *molecular dynamics calculations* (MDC), (reviews: ENGLISH *et al.* [1992] and DIAZ DE LA RUBIA and GUINAN [1992]). Since the first application of MDC to point-defect production by GIBSON *et al.* [1960] and LEHMANN *et al.* [1981] (fig. 39), the method has considerably been improved, (i) by implementing more realistic isotropic manybody potentials as calculated by the embedded atom method (FOILES *et al.* [1986]) and (ii) by the use of modern supercomputers with vector architecture and (iii) by including the *phonon–electron coupling* (CARO and VICTORIA [1989], FINNIS *et al.* [1991]). As a result, calculations became possible up to primary recoil energies of 25 keV applied to one of the atoms of a crystal block containing up to 5×10^5 atoms. These calculations generally confirm what is seen in fig. 39 but uncovered a large amount of details. In regard to our question about efficiency ξ,

it is of fundamental importance that the cascade core assumes liquid structure as shown in fig. 40 for an intermediate time period. The atom density in fig. 40a was derived from a snap-shot of the cascade core at the given time. The density distribution in fig. 40b represents that of liquid Cu. The atom density in the core reaches only 80% of that of liquid Cu at 2200 K for which the radial distribution function matches the one observed in the cascade core. The lacking atoms are distributed around the liquid drop volume. The studies have further shown how vacancies and interstitials recombine mutually and agglomerate within each of both populations. As a result, the production efficiency ξ was obtained in good agreement with the measurements. The strong decrease between 10^2 and 10^3 eV coincides with the increasing importance of the thermal spike within the cascade.

The MDC studies further show that enhancing the temperature of the cascade surrounding matrix enlarges the cascade cooling period with significant effects on the defect agglomeration (§ 4.2.2), i.e, on the escape from recombination. For Cu by counting the vacancies agglomerated in loops an increase of ξ by a factor of two was observed when going from 4 K irradiation temperature to room temperature (JENKINS *et al.* [1993]).

Cascades for recoil energies in the order of magnitude 10^5 eV are computer-simulated by means of binary collision codes. Figure 41 gives a three-dimensional impression of a computer simulation result. Two features are typical and important: (i) The vacancies are frequently clustered and are surrounded by single interstitials. The picture was postulated earlier by SEEGER [1958]. (ii) The defect density strongly fluctuates and the structure could be described as a complex of subcascades. This feature is typical for high-energy cascades.

The defect cascades can be imaged by electron microscopy (reviews: JÄGER [1981], KIRITANI [1982] and SHIMOMURA *et al.* [1982]) and field ion microscopy (reviews: WAGNER [1982] and SEIDMAN [1978]). A particular enlightening example of electron microscopy results is shown in fig. 42. By producing collision cascades in long-range ordered Cu_3Au, disordered volume regions are created which are the images of the collision cascades. By using the superlattice reflections and dark-field imaging, the disordered zones appear as dark contrast within the brightly imaged ordered matrix.

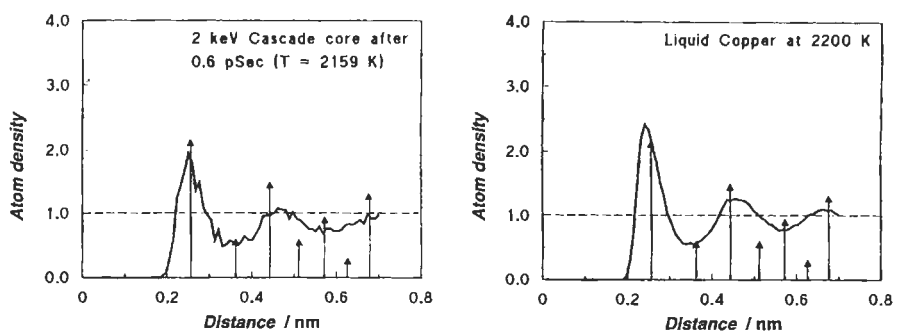


Fig. 40. Radial atomic distribution functions for Cu (a) the core of a 2 keV cascade after 0.6 psec, (b) liquid at 2200 K after ENGLISH *et al.* [1992].

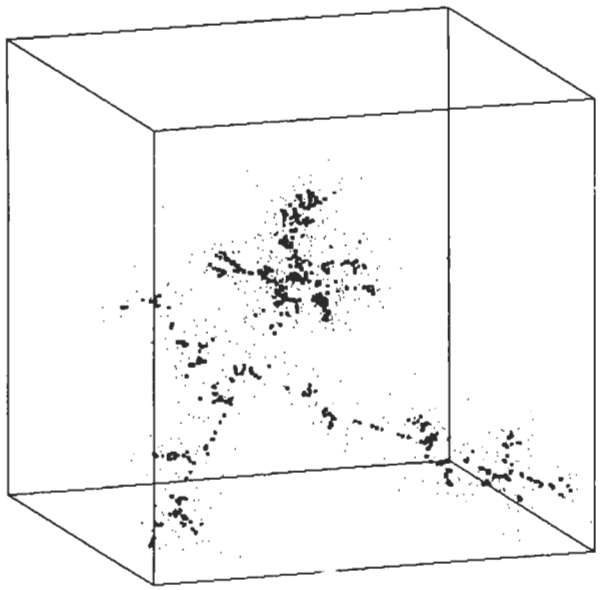


Fig. 41. Computer simulation of a displacement cascade with $T = 2 \times 10^5$ eV in a Cu model. The primary knock-on occurred in the lower left. The squares are vacancies and the (smaller) dots are interstitials. Cube edge measures 125 nearest-neighbor distances. (from HEINISCH [1981].)

Subcascade formation can clearly be seen in fig. 42b. The average cascade diameter was found to lie between 3 nm at $T = 10$ keV and 12 nm at $T = 100$ keV, in satisfactory agreement with calculated values (SIGMUND [1972]) for $T \leq 50$ keV. This example shows

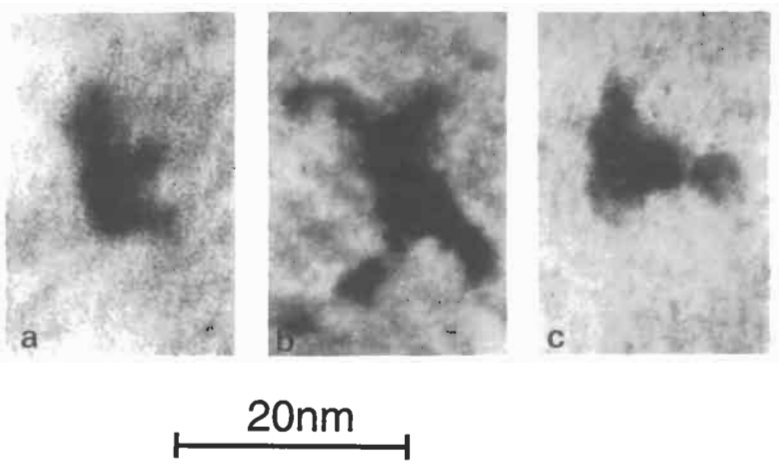


Fig. 42. Disordered zones in long-range ordered Cu_3Au irradiated with Cu^+ ions of different energy: (a) $T = 5 \times 10^4$ eV; (b) $T = 10^5$ eV; (c) separated subcascade at $T = 10^5$ eV (from JENKINS and WILKENS [1976]).

References: p. 1710.

the direct phase transformation by cascade-type damage. Another important aspect of the resulting defect cluster structure is their nature as point-defect sink. Both aspects are treated in more detail in §§ 4.2.2 and 4.6.

4.2.2. Intra-cascade defect reactions

Since about 1990, a predominant fraction of the fundamental radiation damage studies refer to intra-cascade defect reactions. This interest arises from a considerable amount of data, which could no longer be explained satisfactory by the chemical rate equation models with the simplifying assumptions regarding homogeneous single defect production and reactions. Challenging problems were the damage correlation for ion and neutron irradiation (see further reading: Singh *et al.* [1990]), extreme high swelling rates (Woo *et al.* [1993]), fluctuating cluster populations in time (Ishino [1993]), and the evolution of loop and void concentrations with increasing fluence (Kiritani [1993], Singh and Zinkle [1993], Abromeit [1994]).

The detailed MDC studies have first led to quantitative understanding of the atomic mixing by the thermal spike process and the molten zone (Diaz De La Rubia and Guinan [1992]). Secondly, they revealed valuable information on the *cascade collapse*. The solidification process of the molten zone determines the configuration of the vacancy agglomerate collected in the central part of the cascade. The number of vacancies collected depends sensitively on the occurrence of collision chains, which transport interstitials out of the molten zone. The replacement chains, however, were found to be comparatively short. They are more important at lower temperatures and for lower cascade energies. For higher cascade energies, *ballistic vacancy-interstitial separation* was found to be more important (Jenkins *et al.* [1993]). Frequently, the ballistic ejection of interstitials from the cascade core was followed by interstitial clustering at the cascade periphery. These small interstitial clusters in general had Burgers vectors tangential to the cascade periphery allowing them to escape from the cascade region by thermally-assisted glide. This mechanism might be critical in preventing recombination and thus enhancing vacancy survival. It initiated the formation of new swelling models (§ 4.4).

The fraction of single vacancies and interstitials, which may escape the intra-cascade reactions cannot yet be derived from MDC simulations with satisfactory accuracy because of poor statistical relevance. The binary collision codes with slightly better statistical relevance yield substantially more free interstitials (14% of total defect production) than free vacancies (3.5%) (Heinisch [1990]).

It should be noted that the MDC simulations give very little evidence for vacancy loop production during the cascade event. However, loop formation was observed by MDC when areas of high vacancy concentration are subjected to intense thermal motion (Kapinos *et al.* [1990]).

There is still a substantial gap between the few cascades studied by MDC and the wealth of experimental data on cascade collaps for ion and neutron irradiation. The quantities frequently evaluated are the defect yield Y and cascade efficiency ε. The defect yield relates the number of visible vacancy loops per unit area to the ion flux and the efficiency relates the number of vacancies ascribed to the observable loops to the number of vacancies calculated per cascade according to eq. (10). The number of vacancies per loop is difficult to determine

because of problems with the image conditions for the loops. In addition, the sample surface as defect sink strongly influences the yield. Figure 43 shows the defect cluster distribution in a sample with variable thickness. The lack of vacancy clusters in the foil center is certainly due to interstitials annihilating clusters, while the interstitials are lost to the surface from the adjacent layer. For ion irradiated samples, which are generally studied in the thin geometry, the defect yields amount to a few percent for Cu and to around 10% for Ni when the incident ion energies range from 10 to 50 keV.

For any kind of damage model the loop production rate is one of the key quantities as it is essentially a production rate of sinks and of dislocations which may contribute to the microstructure controlling the mechanical properties of the material. At temperatures beyond 0.5 T_m the loops act as vacancy sources and thus affect the effective defect production.

The problems of reactor technology are relevant for fluences in the dpa range from 1 to 100. Therefore, the question of the interaction of cascades with remnants of formerly produced cascades becomes very important. In fig. 44 we have the fluence dependence of the cluster density for neutron irradiated Cu at different irradiation temperatures. The upper solid curve reflects the behavior of fission, spallation, and fusion neutrons, i.e., for rather different neutron spectra. The cluster production rate vanishes at fluence levels between 10^2 and 10^{-1} dpa. Obviously, cascades overlapping the existing clusters induce their annihilation. This effect was indeed observed by MDC simulations (JENKINS *et al.* [1993]). Deviations from the linear dependence of the cluster density on the fluence are explained by cluster annihilation due to freely migrating interstitials. The temperature dependence of the cluster density shows Arrhenius behavior caused by the vacancy evaporation.

Experimental information on the fraction of single vacancies and interstitials which escaped reactions with their parent defect cascade comes from measurements of damage rates (THEIS and WOLLENBERGER [1980]), mass transport (REHN and BIRTCHER [1993]) and swelling rates. Evaluation of the mass transport data yield fractions of less than 5% for Ni, for example, when equal fractions for vacancies and interstitials are assumed. For the explanation of the observed maximum swelling rates fractions beyond 7% are needed when the conventional single defect dislocation bias is presumed. This discrepancy lead to the detailed consideration of direct glissile interstitial loop production as part of the cascade event (TRINKAUS *et al.* [1993]). If this type of loops were indeed produced to a substantial degree the mass transport data would have to be re-interpreted, as discussed by TRINKAUS *et al.* [1994].

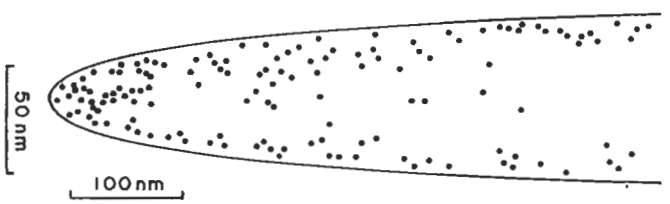


Fig. 43. Distribution of defect clusters in Ni irradiated as thin foil at 573 K with 14 MeV neutrons. The measured positions are projected to the paper plane (after KIRITANI [1993]).

References: p. 1710.

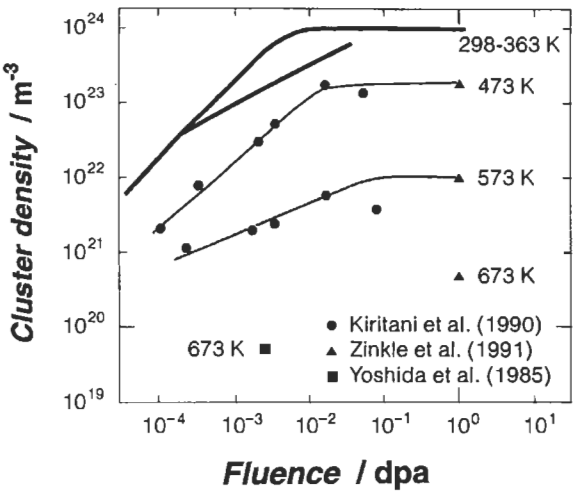


Fig. 44. Fluence dependence of defect cluster accumulation in neutron-irradiated copper at different irradiation temperatures. The upper solid line represents data from fission, fusion, spallation neutrons, and 800 MeV protons (SINGH and ZINKLE [1993]).

4.3. Fast heavy-ion irradiation

The electronic energy loss of charged projectiles in the metallic crystal lattice has been treated in § 4.2.1 as a quantity which must be reduced from the total recoil energy in order to obtain the available damage energy. The electronic energy loss was taken as a heat source and other effects have been neglected. This approximation holds up to a certain threshold $(dE/dx)_e$ of the electronic stopping power. In fig. 45 the nuclear and electronic stopping powers are compared with each other as function of the incident energy of a Xe ion penetrating Cu. One realizes that the electronic stopping power predominates above a certain threshold. This threshold amounts to about 40 keV/nm for metals. By using oxygen to uranium ions with energies of 10 MeV/nucleon one covers the (dE/dx)e range between about 1 and 10 keV/nm. Heavy ions with energies above 1 MeV/nucleon are available from a number of accelerators, which have formerly been built for the purpose of nuclear physics research. The penetration depth of such ions in matter is of the order of 10^{-4} m, which leads to a sample thickness well suited for microscopic and macroscopic investigations.

The differential cross-section of a target electron to receive the recoil energy T varies as $1/T^2$ with, for example, $T_{max} = 10$ keV and $\langle T \rangle = 100$ eV. These recoil energies are very small as compared to the projectile energy because of the unfavorable mass ratio of the colliding particles. The excited electrons are not able to initiate atomic displacements by collision with target atoms, again because of the unfavorable mass ratio. The energy released to the target electrons is shared between the kinetic energy of them and the potential energy which is essentially the potential energy of the ionized atoms. It amounts to about 10% of the total deposited energy for the ions located in the ion wake. The

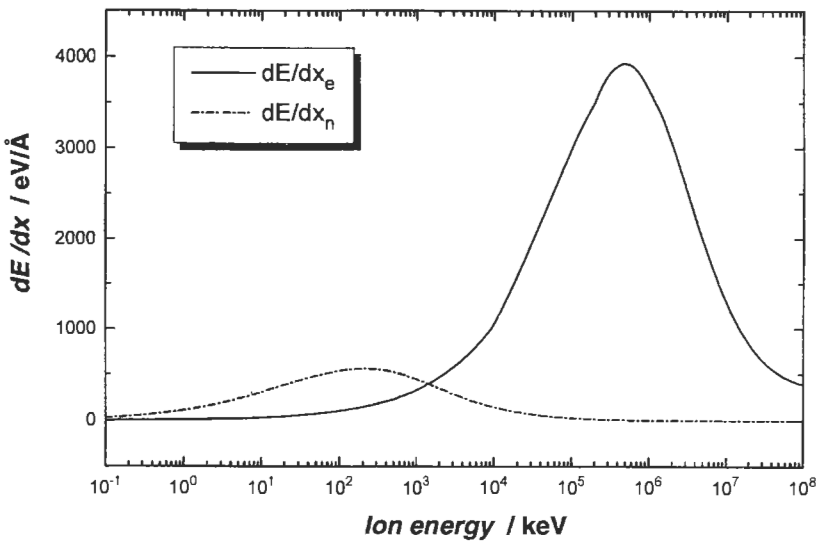


Fig. 45. Electronic and nuclear stopping power for Xe in Cu.

corresponding recoil energy per atom around the projectile trajectory amounts from 0.1 to a few eV. Quite different from the nuclear energy loss process, where the recoil energy transfer is a spatially and timely stochastic process the potential energy transfer to the trajectory surrounding atoms is a highly collective phenomenon in space and time. The response time of the free electron gas is characterized by the plasma frequency, which corresponds to a relaxation time of 10^{-16} s. The whole process is termed a Coulomb explosion and leads to a shockwave emanating from the projectile trajectory.

Coulomb explosion has been studied by means of MDC (LEGRAND *et al.* [1993]) with the goal of illuminating the result of such collective lattice deformations. The calculations have been performed for a Fe crystal cell containing 5,000 to 6,000 atoms with cell dimensions of about $6 \times 6 \times 2$ nm^3. The authors investigated the shockwave effect for projectile trajectories parallel to <100>, <110>, <111>, and <120>. The excitation energies per atom were chosen around 1 eV/atom and the radial impacts were given simultaneously to about 600 atoms forming the cylinder around the assumed projectile trajectory. The excitation energy was dissipated at the crystal boundaries by properly arranged *stochastic walls*. As a result, the formation of one to two Frenkel defects was observed per event. The results indicate that Frenkel defect production might be possible by electronic excitation via a Coulomb explosion.

Experimental work proved a significant influence of the electronic excitation on the defect production above a certain threshold of $(dE/dx)_e$. Figure 46 shows the thermal recovery of the radiation-induced resistivity increase after irradiation with the given ions and energies. Attention is focussed to the stage I annealing. While the 0.5 MeV proton irradiation produces the wellknown stage I annealing, the amount of recovery becomes smaller when the electronic stopping power of the projectiles is increased. The effect is

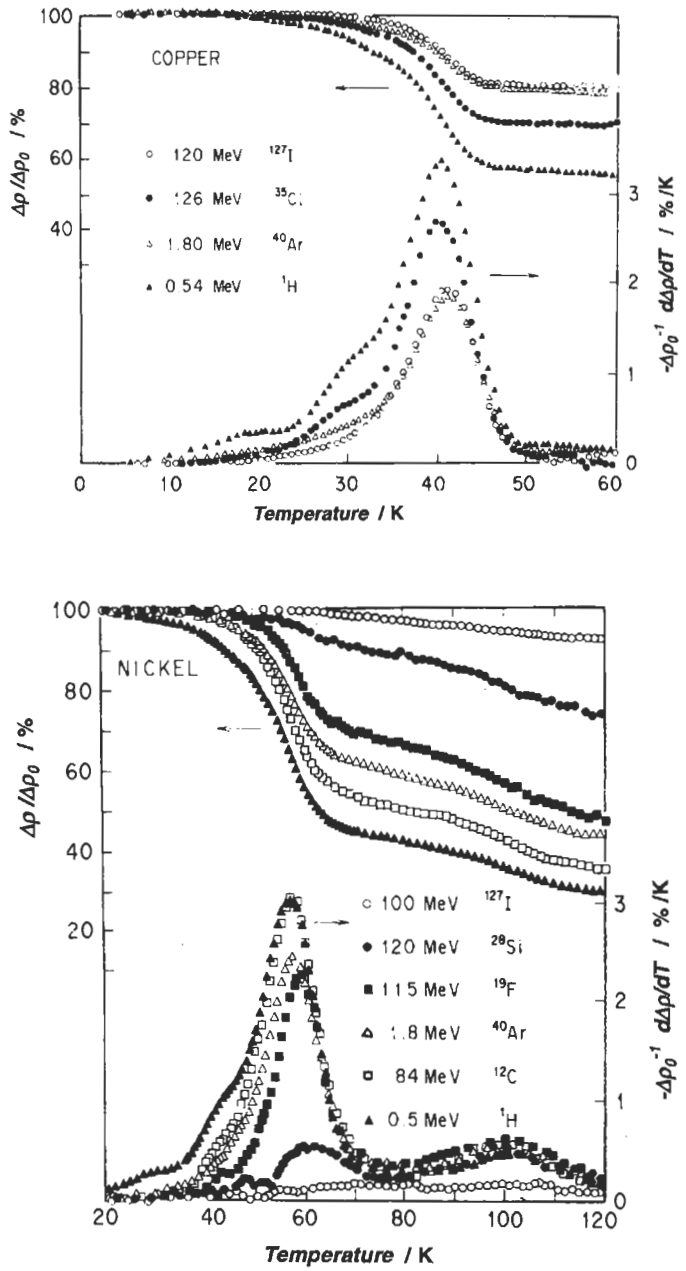


Fig. 46. Electrical resistivity recovery upon isochronal annealing and recovery rates for heavy-ion irradiated Cu and Ni (after IWATA and IWASE [1990]).

stronger for Ni than for Cu. The recovery reduction occurs at the low temperature side of stage I indicating that the fractional population of close pairs and correlatedly recombining freely migrating interstitials (§ 3.3.2.4) is reduced. By increasing mass and incident energy of the projectiles not only $(dE/dx)_e$ is enhanced but also the recoil spectrum due to elastic nuclear collisions is changed. In order to separate potential effects arising from both of the mechanisms it is reasonable to plot the stage I recovery one time vs. the weighted mean elastic recoil energy $T_{1/2}$ from nuclear collisions (AVERBACK *et al.* [1978] and the other time vs. $(dE/dx)_e$. Increasing $T_{1/2}$ means essentially fractional decrease of single Frenkel defect production in favor of defect cluster production. Both are done in fig. 47 for Ni. The two branches (open and closed symbols) in fig. 47a tell us that equal values of $T_{1/2}$ lead to completely different stage I populations, depending on the incident ion energy. Fig. 47b shows all datapoints describing just one curve. Obviously, the electronic stopping power is responsible for the stage I reduction. For Cu the datapoints describe just one curve already in the $T_{1/2}$-plot indicating negligible influence of $(dE/dx)_e$ on the stage I population. The metals Al and Pt were found to behave like Ni while Ag behaves like Cu. The two groups of metals deviate by their electron-phonon coupling (§ 4.2.1 and § 4.2.2) with a strong coupling for Ni and Pt and a weak one for Cu and Ag. The authors explain these observations by Frenkel defect production via elastic nuclear collisions and the reduction of the recovery in stage I by subthreshold annihilation (§ 3.1.2) of formerly produced Frenkel defects as a result of the electronic excitation. The behavior of the defect production rate as function of the defect concentration confirms this interpretation. The rate is constant initially and begins to decrease with increasing defect concentration when the accumulated defect concentration allows for a measurable annihilation rate (IWATA and IWASE [1990]).

For the metals Ti, Co, and Cr, DUNLOP *et al.* [1993] observed positive electrical resistivity damage rates which increase linearly with $(dE/dx)_e$. The effect is largest for Ti. For Ti irradiated by GeV Pb ions electron microscopy resolved alignments of small defects along the incident ion beam direction. The defects appear by the same contrast as known for dislocation loops. The defect size appears to stay below 10 nm. By irradiating with 4 GeV U ions at 20 K the phase transformation from the hcp equilibrium α-phase to the hexagonal ω-phase is induced. Without irradiation, this phase transformation is initiated under hydrostatic pressure of the order of 10 kbar. For Fe the authors find the Ni type behavior for irradiation with ions up to $(dE/dx)_e \approx 40$ keV/nm. Beyond this value Fe behaves like Ti, Co, and Cr. DUNLOP *et al.* [1993] point out that each of the metals Ti, Fe, Co, and Cr does form various allotropic phases and exhibits soft modes in the phonon spectra, which are associated with the displacive phase transformation. This correlation might suggest that the damage detected by means of the electrical resistivity increase consists of second phase precipitates also for Fe, Co, and Cr.

Up to now there is no experimental evidence for direct Frenkel defect production by electronic excitation in pure metal crystals. But, the deposited energy is obviously sufficient to induce close Frenkel pair recombination (subthreshold collisions) with strong electron–phonon coupling. The specific collective and radial impact generation emanating from the swift ion path may well trigger displacive phase transformations. This has been demonstrated, in particular, for the crystal-amorphous transition in intermetallic com-

References: p. 1710.

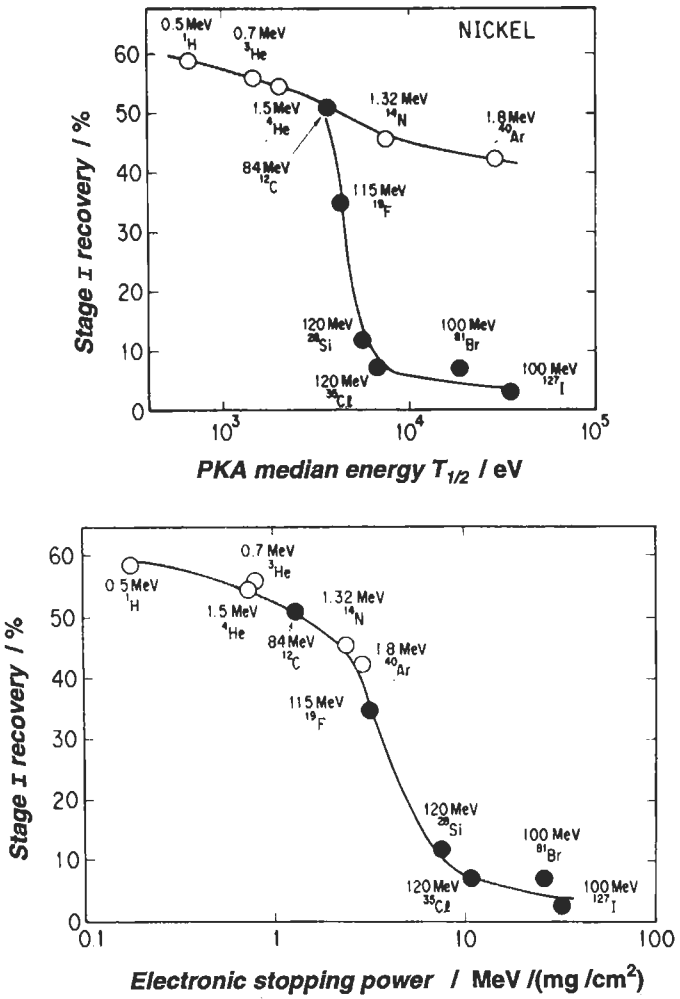


Fig. 47. Fraction of stage I recovery for heavy-ion-irradiated Ni (a) vs. weighted average recoil energy of the irradiating ions indicated, (b) vs. electronic stopping power of the irradiating ions indicated (after IWATA and IWASE [1990]).

pounds. In compounds like $NiZr_2$ and NiTi which can easily be amorphized by other kinds of irradiation (elastic collisions) or plastic deformation, ion tracks of amorphous material of about 8 nm diameter are produced for $(dE/dx)_e > 50$ keV/nm (800 MeV Pb or U ions). For Cu_3Au no phase transformation by electronic excitation has been observed.

In amorphous materials, fast-ion irradiation causes strong plastic flow perpendicular to the ion beam direction (KLAUMÜNZER [1992]). This phenomenon is again caused by

the collective cylindrical impact generation which, in amorphous materials, can develop significant internal stress levels as the atoms must not necessarily return to their initial positions before the Coulomb explosion. In crystals they do return because of the topological stability. The amorphous material responds to the internal stress by intense plastic flow. This flow requires the presence of shear sites of sufficiently large concentrations. These sites are likely to be identical with the *free volume* (ch. 7, § 9.1.2).

4.4. Swelling

One of the serious radiation damage problems in nuclear engineering is the *dimensional instability* of metallic structural materials. It shows up as *swelling*, i.e., decrease of the materials density with increasing fluence, and as volume-conserving creep, which is observed at significantly lower temperatures than is the thermally activated creep (§ 4.5). Swelling and the formation of voids in the bulk of the material was detected in structural core parts of the Dounreay Fast Reactor (CAWTHORNE and FULTON [1967]). Electron micrographs of voids in Ni-ion-irradiated Ni and neutron irradiated type 316 stainless steel are shown in fig. 48. The observed macroscopic volume increase $\Delta V/V$, i.e., the swelling equals the total void volume. The morphology of voids as three-dimensional vacancy arrangements is multiform (further reading: Corbett and Ianiello [1972], Bleiberg and Bennett [1977] and Carpenter *et al.* [1980], Gavillet *et al.* [1993]). Voids have also been studied by small-angle X-ray scattering, which can fruitfully be combined with other techniques (ch. 12). The phenomenon is widespread in view of the materials as well as the irradiating particle species. Microstructure and chemistry

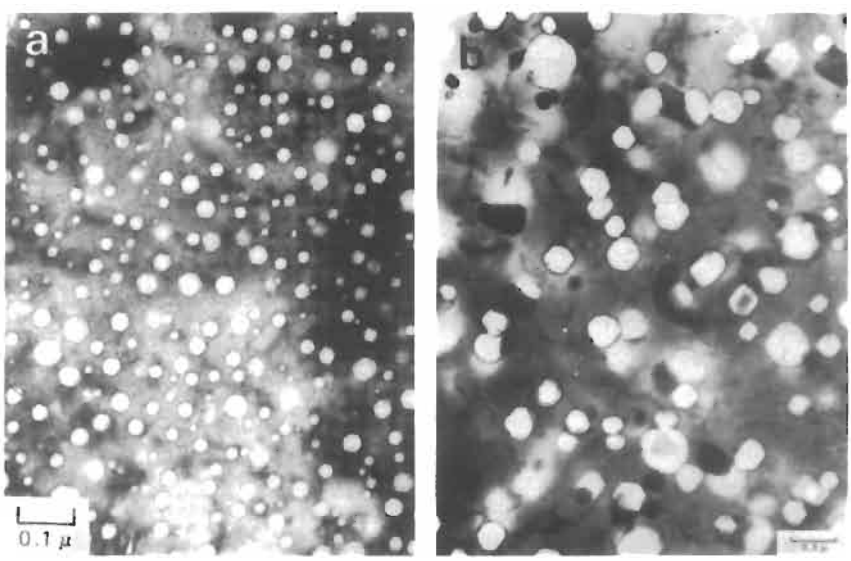


Fig. 48. Voids in (a) Ni irradiated by 200 keV Ni^+ ions at 500°C (after ADDA [1972]) and (b) type 316 stainless steel irradiated by 7.1×10^{22} neutrons/cm^2 at 525°C (after APPLEBY *et al.* [1972]).

(composition) of the materials and irradiation temperature are of considerable influence on the swelling rate. The general feature of the swelling behavior of relevant structural materials is sketched in fig. 49. We note a number of important characteristics: (i) significant swelling starts after accumulation of an incubation fluence, (ii) saturation of swelling appears not to exist within tolerable swelling levels (except a few void lattice forming materials), (iii) ferritic steels show considerably less swelling than austenitic ones, and, (iv) cold-working enhances the incubation fluence. For comparison, the behavior of Mo is shown: this develops a void lattice (§ 4.6).

In the present section we focus our attention on the basic processes controlling void nucleation and growth. This view will enable us to understand the above characteristics in terms of general defect reaction schemes. Although voids have been known for a long time from the Kirkendall effect (§ 2.2.1) (see also ch. 7, 5.3), they were not expected in irradiated materials as all earlier investigations showed only planar aggregates of vacancies (dislocation loops, stacking fault tetrahedra). Moreover, such aggregates were found to dissolve at temperatures at which voids are still being formed. Research of the nucleation problem was always impeded by the fact that impurities, in particular, gaseous ones will certainly stabilize three-dimensional vacancy agglomerates. In addition, Fast Reactor irradiations not only produce Frenkel defects but also relatively large amounts of He resulting from the large cross-sections for (n,α) transmutations of frequently present metal isotopes. On the other hand, it proved to be difficult to prepare sufficient gas-free or generally impurity-free samples to study homogeneous void nucleation under charged particle irradiation. MDC studies for Cu and H as impurity (SHIMOMURA *et al.* [1993]) revealed that the tri-vacancy relaxation towards the stacking fault tetrahedric structure may be suppressed by involving H atoms. However, the trapping energy for H is small for vacancy agglomerates up to 5 vacancies. From 6 vacancies onwards, trapping

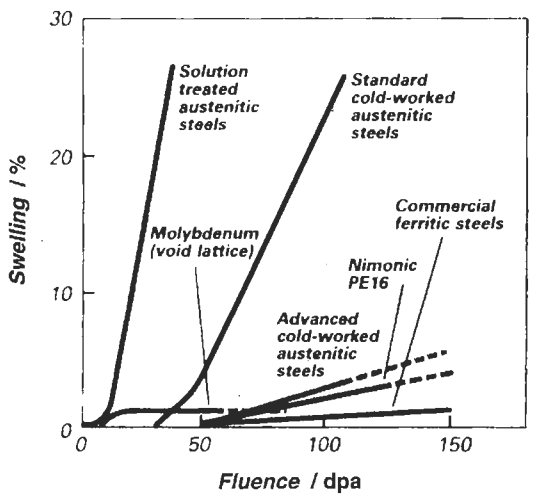


Fig. 49. Swelling vs. fluence for different materials and conditions (schematic) after EYRE and MATTHEWS [1993].

of H stabilizes the void structure. An H atom which is trapped in the hexa-vacancy pushes out four atoms which are positioned in the tetradedral 10-vacancy of 6-vacancy stacking fault tetrahedra. This stabilization is equivalent to the suppression of the relaxation of a 3-vacancy to a stacking fault tetrahedra. The void structure becomes more and more stable between 10 and 20 vacancies, but needs an increasing number of H atoms for the stabilization. The study also shows that vacancy agglomerates with less than 15 vacancies become thermally unstable at the *peak swelling temperature* (see further below).

Gaseous impurities are now generally accepted as necessary prerequisites for the stabilization of the void structure. The fission product He has received most attention with respect to its role in void nucleation (GLASGOW *et al.* [1981], STOLLER and ODETTE [1985], PARKER and RUSSELL [1982]). Void formation by electron irradiation demonstrates the nucleation for pure single vacancy production. Here, void densities are often enhanced at and near grain boundaries and dislocations (HORSEWELL and SINGH [1988]). This inhomogeneity corresponds to the well-known impurity segregation.

For cascade-producing irradiation, intra-cascade clustering and the considerable fluctuation of the defect concentration of freely migrating defects (which escape from cascades) on void nucleation, has been studied in some detail (KITAJIMA *et al.* [1985], TRINKAUS [1990]). It turns out that these effects should have only minor influence on the effective void nucleation rates, mainly because of the small probability for the direct agglomeration of large enough numbers of vacancies.

Gas-containing three-dimensional vacancy agglomerates (cavities) are termed *bubbles*, the size of which is controlled by the gas pressure, i.e., the ratio of the numbers of gas atoms to vacancies and the metal surface energy (TRINKAUS [1983], MANSUR [1993]). The transition to apparently unlimited void growth as seen in fig. 49 is correlated with the critical bubble radius, which is given by the deviation of the effective gas conditions from those of the ideal gas, the surface energy, and the excess vacancy flux to the cavity. Hence, extension of the incubation fluence requires materials design such that the gas and/or vacancy fluxes towards bubble embryos are reduced.

The excess vacancy flux feeding the void growth was generally assumed to arise from biased vacancy and interstitial annihilation at sinks (reviews: BRAILSFORD and BULLOUGH [1978], BULLOUGH and WOOD [1980], MANSUR [1993] and further reading: Wiffen *et al.* [1979], Nygren *et al.* [1981]). The common explanation of such biased vacancy and interstitial annihilation at sinks considers the elastic interaction between the migrating point defects and dislocations as causing a drift diffusion towards the dislocations. The interaction energy for a dilatation centre A with the volume change ΔV_A^m in the strain field of an edge dislocation is proportional to $\Delta V_A^m / r$, where r is the distance of A from the dislocation core. The proportionality to ΔV_A^m makes the interaction strength quite different for vacancy and interstitial (Cu, for example: $\Delta V_v^m = -0.2$, $\Delta V_i^m = 1.6$). In a given material, the topologies and concentrations of saturable sinks (e.g., sessile dislocations), unsaturable sinks and saturable traps composes a complex scheme of spatially varying defect reactions.

The swelling theories were based upon the chemical rate equation approach which assumes spatially homogeneous defect reactions. Spatially inhomogeneous reactions like

References: p. 1710.

point-defect annihilation at an extended localized inexhaustable sink (e.g., sessile edge dislocation) were incorporated into the rate equations by appropriate divergence terms for the defect concentrations. For a long time, freely migrating single interstitials and vacancies were assumed to be the only reactants for traps and sinks and in addition the displacement rates according to eq. (10) were taken as production rates for these freely migrating defects. Intra-cascade recombination and clustering (§ 4.2.2) were widely neglected. The wealth of experimental data on the microstructural evolution under all kinds of irradiation (further reading: Klueh *et al.* [1992], Ishino *et al.* [1991]) as well as the progress in modelling suggest that satisfying agreement between the early theories and experiments whenever achieved may be attributed to the large number of less well-known parameters which were involved in the theories and adjustable within relatively wide ranges. However, taking the various defect reactions into account allowed a satisfying qualitative explanation of the swelling response to the various materials variables and irradiation conditions. For a brief illustration, we look at the influence of cold-working and alloy composition on the irradiation temperature dependence of the swelling (fig. 50). The cold work controls the vacancy supersaturation via dislocation sink strength and bias. The temperature controls the defect mobility and, by this, the stationary defect concentration which in turn determines the relative influence of the recombination. At low temperatures the recombination rate is large because of the slowly migrating vacancies. The biased sink annihilation is suppressed. At higher temperatures the latter is favored at the expense of recombination. At the same time, vacancy loops start decaying which enhances the vacancy supersaturation. At even higher temperatures, the stationary vacancy concentration approaches thermal equilibrium values and, hence, does not provide the supersaturation required for void formation. Slowing down of one of the migrating defect species can be achieved by adding solutes which temporarily trap interstitials (§ 3.3.2.8). In this way swelling is reduced (see, for example, POTTER *et al.* [1977]) as shown in fig. 50b. The peak swelling temperature depends on the displacement rate because of the increasing dominance of recombination with increasing displacement rate (steady-state defect concentration).

Quantitative interpretation of the vacancy supersaturation by the annihilation bias alone became more and more difficult when new experimental work revealed swelling rates up to 5%/dpa in neutron-irradiated Ni (KIRITANI *et al.* [1990]) and in Cu containing 0 (YAMAKAWA *et al.* [1992]). Moreover, detailed investigation of local variations of the swelling rates showed maxima up to 7%/dpa for Al in the neighborhood of grain boundaries and in subgrains surrounded by dislocation walls (HORSEWELL and SINGH [1988]). The width of these maximum swelling regions is of the order of 10^{-6} m. A swelling rate of 5%/dpa requires an excess vacancy flux to the voids by the same fraction of the total flux of the freely migrating single defects. Hence, for the less probable case that all interstitials annihilate at sinks different from the voids, the fractional freely migrating defect escaping from the cascades must be 5% or significantly larger for the probable case of partial interstitial annihilation at voids. Such values do not fit theoretical and experimental findings (§ 4.2.2). In addition, the width of 10^{-6} m is considerably larger than the average diffusion distance of single vacancies and interstitials for the sink strength observed for respective irradiation conditions. These

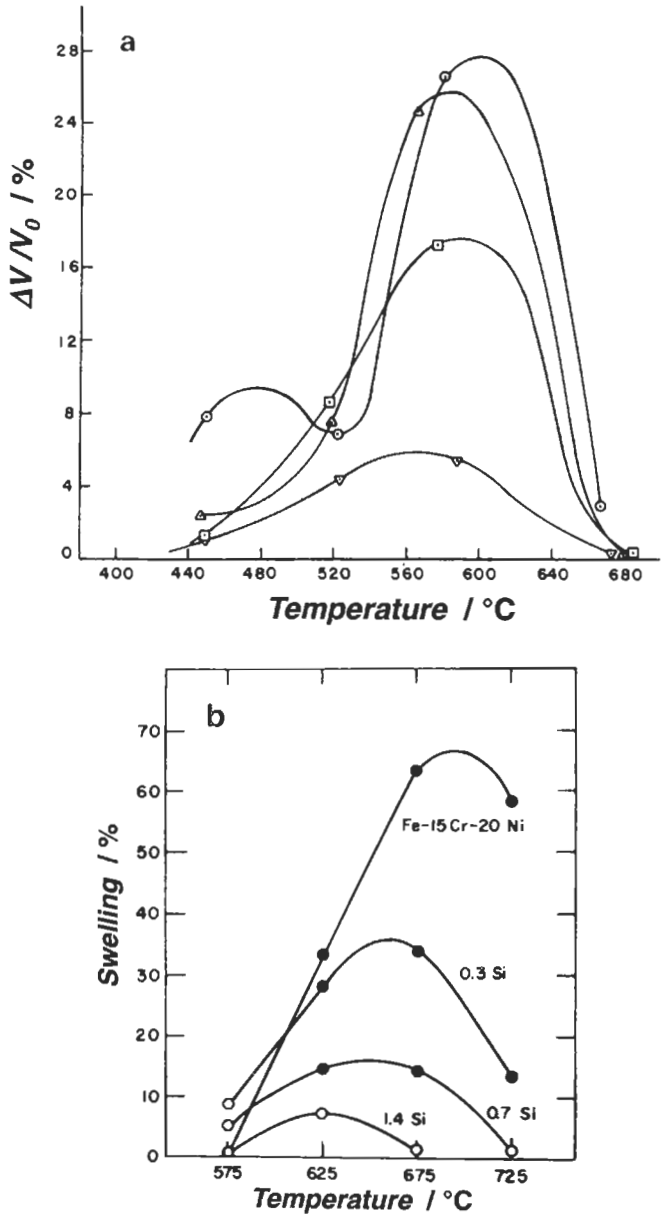


Fig. 50. Influence of pretreatment and solute additions on void-swelling: (a) 316 type stainless steel after 14.1×10^{22} neutrons/cm^2, ⊙ - annealed, △ - 10% cold worked (cw), □ - 20% cw, ▽ - 30% cw (from FLINN and KENFIELD [1976]); (b) Fe-15 at% Cr-20 at% Ni with quoted Si additions after a fluence corresponding to 140 displacements per atom (from JOHNSTON *et al.* [1976]).

References: p. 1710.

difficulties can be overcome by taking into account the production of glissile interstitial loops during the cascade event (TRINKAUS *et al.* [1992]). Such loops are known to move by climb and glide. Small interstitial loops are believed to be faulted. However, MDC studies for Cu have shown that small interstitial loops do transform spontaneously from the faulted to the highly glissile unfaulted configuration (FOREMAN *et al.* [1991]). Such loops would perform a one-dimensional glide with an estimated activation enthalpy of about 0.1 eV. For small loops of four interstitials, the direction of the Burgers vector may change to equivalent lattice orientations with an activation enthalpy of about 0.5 eV. Such loops would migrate by describing a zig-zag path until they are trapped or annihilated by proper components of the microstructure. The significant result of the one-dimensional glide motion is the much larger free travel distance than the diffusion distance of the three-dimensionally migrating defect for the same microstructure. Under these circumstances, spontaneously produced small interstitial loops may be trapped or annihilated having a very small chance to get involved in recombination or to encounter voids. Numerical calculations showed that the free travel distance may indeed be of the order of 10^{-6} m and thus explain the required vacancy supersaturation in the above mentioned width of maximum swelling regions near grain boundaries. By this explanation, the annihilation bias is supplemented by a *production bias* in the sense that the numbers of freely migrating vacancies and interstitials escaping the cascade are biased by the gliding interstitial loops. More specific experimental work is necessary to confirm this defect reaction path.

The remarkable swelling resistance of the ferritic steels (fig. 49), and of other bcc metals and alloys, was earlier ascribed to the small elastic interstitial dislocation interaction due to the small relaxation volume of the interstitial in these structures (table 4). The small elastic interaction would result in a small drift diffusion effect towards the dislocations and, hence, a small annihilation bias (SNIEGOWSKI and WOLFER [1984]). MDC studies of the interstitial dislocation interaction, however, showed that interstitials are likely to be trapped near dislocations (KURAMOTO [1992]) and thus kept available for recombination. More detailed experimental information leads to the conclusion that impurity trapping is the leading effect among a number of other defect reactions which are of greater influence on the vacancy supersaturation than corresponding reactions are in fcc metals and alloys (LITTLE [1993]). The interstitials C and N and the substitutional Si are particularly seen as being effective. The C-vacancy binding enthalpy seems to be as high as 0.85 eV and would therefore act as an effective recombination catalyzer even at high temperatures. Also, Cottrell atmospheres around dislocations may screen the sinks and reduce climbing rates.

4.5. Radiation-induced creep

The thermally activated creep of materials is treated in chs. 20–25. An important feature of this phenomenon is the exponential temperature dependence of the creep rate, $\dot{\varepsilon}$, for a fixed stress, σ. If the creep experiment is done under irradiation, $\dot{\varepsilon}$ becomes nearly independent of temperature below about $T_M/2$. This behavior originates from the high radiation-induced vacancy- and interstitial-concentration which allows sufficient disloca-

tion climb and deformation, despite the small diffusivities of these defects (reviews: LUCAS [1993], MANSUR [1992] and *further reading:* Carpenter *et al.* [1980], Gittus [1978], Woo and McElroy [1988]). For not too large applied stresses σ, the creep rate $\dot{\varepsilon}$ is found to be proportional to σ and to the defect production rate. One of the successful models uses the *stress-induced preferential absorption* (SIPA) of defects at dislocations capable of climbing (HEALD and SPEIGHT [1974]), as illustrated in fig. 51. If again the elastic-energy-controlled drift diffusion is assumed to determine the defect current towards dislocation sinks, the orientation of the dislocations with respect to the external stress will control the rate constants for defect annihilation. According to the model calculation, dislocations of type I (fig. 51) absorb more interstitials than vacancies, and vice versa for type II dislocations. The resulting climb along the directions indicated cause the volume-conserving elongation of the sample parallel to σ. Within the void-swelling regime, the excess interstitial flux to the dislocations determines the creep rate, which then obtains a contribution proportional to the swelling rate.

The SIPA mechanism is certainly of fundamental importance for the radiation-induced creep but the really observed behavior is more complex due to the complex microstructure of the structural materials (dislocation interactions, grain-boundary diffusion, chemical heterogeneities). At high stress levels the creep rate becomes proportional to σ^2 which has been explained by a climb-enabled dislocation glide mechanism (MANSUR [1979]).

4.6. Self-organization of point defect agglomerates

Many metals and alloys develop three-dimensionally ordered arrangements of voids (void lattices) when irradiated with heavy ions or neutrons at temperatures around 0.45 T_m. Figure 52 shows the void lattice in Nb, for example. The phenomenon is observed for fcc, bcc and hcp structures. In the cubic metals, the void lattices are isomorphous to the atomic lattice. For the hcp metals, only planar ordering parallel to the basal planes is observed. The void radii amount to about 5 nm while the void lattice parameters are about ten times larger. This ratio varies between 5 and 15 for the various metals with the exception of Al, for which it amounts to about 40. The ratio is nearly independent of the irradiation temperature. The minimum fluence required for the lattice formation ranges

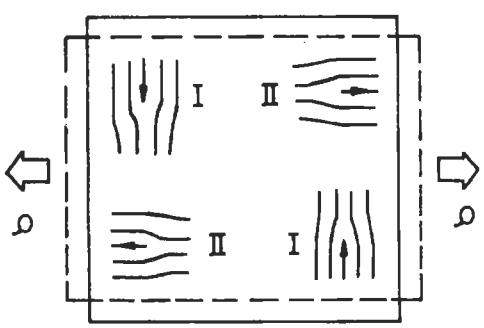
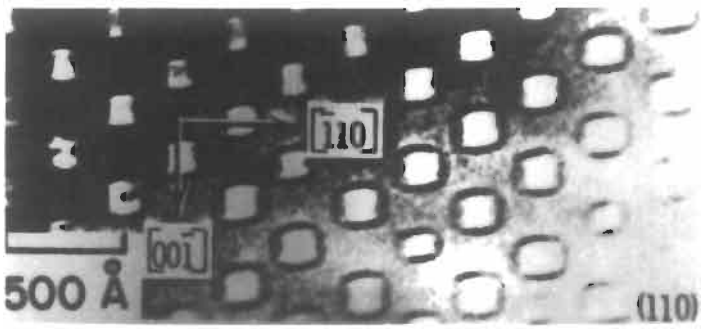
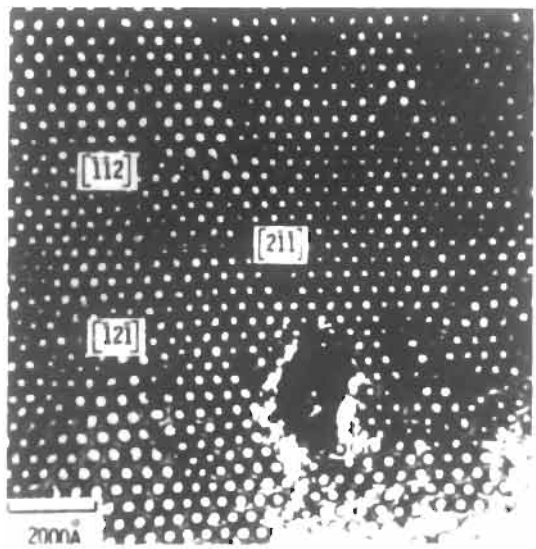


Fig. 51. Dislocation climb directions by preferential vacancy and interstitial absorption under applied stress σ.

References: p. 1710.



(a)



(b)

Fig. 52. Void lattice in Nb irradiated at 1073 K with 7.5 MeV Ta ions to fluences of (a) 140 dpa and (b) 300 dpa.

from 5 dpa to 400 dpa, depending on the metal species. This fluence also depends on the irradiation temperature and on the gas and impurity level of the material. In pure metals void lattices are not observed for electron irradiation. Electron irradiation of disordered void arrangements as produced by ion irradiation did not induce the formation of the lattice. As the total swelling caused by the void lattice generally amounts to a few percent, the total number of displacements is by two to four orders of magnitude larger than that of the vacancies accumulated in the void lattice. Correspondingly, the energy temporarily stored in the total number of defects produced (given by the total fluence level) is large compared to the energy stored in the void lattice. The ordered structure is the result of the highly energy dissipative process.

Sufficiently long irradiation with ions or neutrons at temperatures around 0.3 T_m leads to periodic arrangements of dislocation walls (see fig. 53, for example). The existence ranges for disordered and ordered vacancy agglomerates are well distinguished in the defect production rate vs. temperature diagram as shown in fig. 54. At still lower temperatures, simple cubic lattices of stacking fault tetrahedra are observed in electron irradiated fcc metals and alloys (ZAISER *et al.* [1992]).

Earlier theoretical treatments of void lattice formation were based upon equilibrium thermodynamics (minimization of the total energy) and the elastic interaction of the microstructural components (WILLIS [1975]). These attempts failed as they can neither explain the observed dependence on temperature and irradiating particle species nor the lattice formation in isotropic metals like W. The now generally accepted explanation considers the void lattice as the result of a self-organizing open system far from thermodynamic equilibrium (MARTIN [1975], KRISHAN [1982]). The various theoretical approaches reported since follow this concept but deviate by the various defect reactions taken into account and considered to be dominant (ABROMEIT and TRINKAUS [1992], HÄHNER and FRANK [1992], review: ABROMEIT [1989]). The simplest defect reaction system describing a heavy ion or neutron irradiated crystal already fulfils the conditions for self-organization (HAKEN [1977]). The irradiation produces vacancy loops and freely migrating vacancies and interstitials by constant rates. The loops grow by vacancy

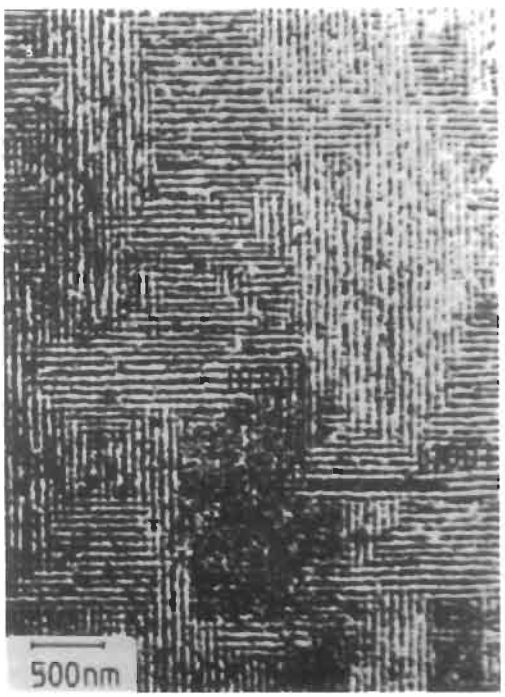


Fig. 53. Periodic arrays of planar {001} walls of dislocations and stacking fault tetrahedra in copper irradiated with 3.4 MeV protons around 100° C to 2 dpa fluence. Imaged in <100> projection (after JÄGER *et al.* [1990]).

References: p. 1710.

annihilation and shrink by interstitial annihilation. The annihilation rates are biased as described in § 4.4. For simplicity we neglect recombination, which is allowed for small defect production rates and/or high temperatures. The important characteristics for the self-organization are (i) the diffusion length of the reacting vacancy and interstitial commensurates with the characteristic length of the evolving ordered structure, (ii) the reaction rates controlling the loop arrangement are non-linear as they are given by the concentrations of the loops times the concentrations of the migrating defects. Mathematically) the system is described by a proper chemical rate equation system which is analysed by means of the linear stability analysis with respect to the stability of spatially homogeneous defect distributions. Our diffusion-controlled reaction system does not contain any information about lattice orientation. Hence, it will not give predictions on the relationship between the ordered structure and the lattice orientation. Nevertheless, we stay with this model in order to understand the physical origin of the self-organization. The stability analysis considers the behavior of concentration fluctuations superposed to the spatial averages. The fluctuations are described by a Fourier series. One component is shown in fig. 55. The index s stands for sink and c_s is proportional to the vacancy loop concentration. Because of the annihilation bias the fluctuation amplitudes of the interstitial concentration c_i and of the vacancy concentration c_v are different in the way as shown in the figure. This difference causes amplification of the fluctuation of the loop concentration $\delta c_s/\langle c_s\rangle$ as the vacancy concentration is larger than the interstitial concentration at the maximum of the loop concentration and higher vacancy arrival rates than interstitial arrival rates induce loop growth. The growing fluctuations induce vacancy and interstitial fluxes from the regions of positive fluctuation amplitudes to those of negative ones (fig. 55). These fluxes are initially larger for the interstitials than for the vacancies due to the difference in the fluctuation amplitude. However, they reduce this difference until total equilibration. The linear stability analysis shows that this amplification of $\delta c_s/\langle c_s\rangle$ vanishes when the fluctuation wavelength λ goes to zero and to infinity. The fluctuation of the loop concentration must therefore have a maximum at a certain

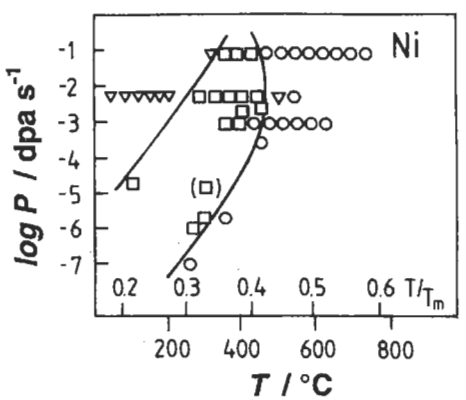


Fig. 54. Map of periodic cluster arrays for Cu irradiated with neutrons, heavy ions, and protons, ▽ - random, □ - periodic cluster array, ○ - periodic voids and dislocations (after JÄGER and TRINKAUS [1993]).

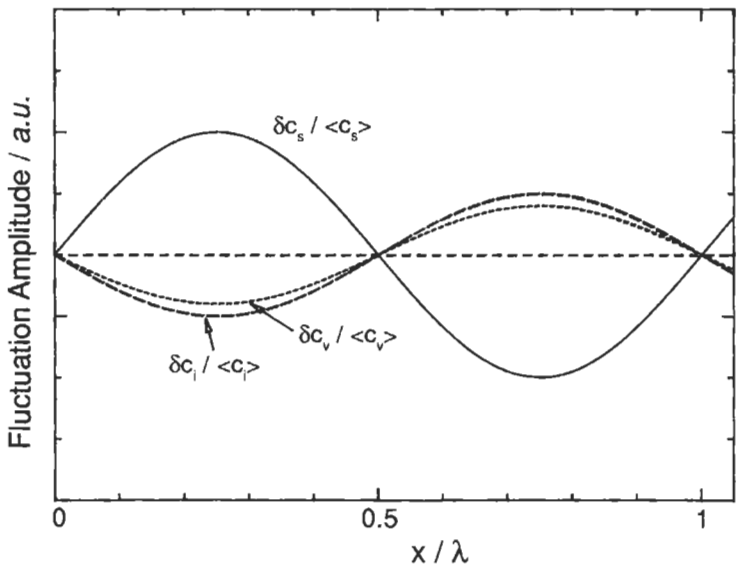


Fig. 55. Sketch of the spatial fluctuation of the concentrations of sinks (*s*), vacancies (*v*) and interstitials (*i*).

finite value of λ.

It becomes obvious that the driving force for the self-organization is the annihilation bias for the two diffusing reactants and the fact that one reactant increases the sink strength of the loops (by growing) and the other (interstitial) decreases it. When the bias had the opposite sign the system would not be self-organizing. Adding recombination or secondary sinks (grain boundaries) or traps to the reaction system would reduce the driving force for self-organization to the measure at which the loops contribute to the total sink and trapping strength. Introducing a production bias (§ 4.4) shifts the annihilation bias threshold accordingly: the difference of the sink strength for interstitials and vacancies must be larger than the respective difference of the production rates. At low temperatures where the recombination dominates the sink annihilation, loop ordering will be impossible in accordance with fig. 54. At high temperatures the vacancy loops decay but void formation becomes possible if nucleation is favored and self-organization may be possible when they dominate as sinks and the respective annihilation rates are biased.

The linear stability analysis yields the initial conditions for which homogeneous cluster distributions may become unstable. It does not give information on the shape of the evolving structure. Such information is obtained by analysis of the stationary

solutions in the instability regime. As a result one obtains the concentration fluctuations for vacancy loops as shown in fig. 56. The regions of high-loop concentrations are much smaller than those of low concentrations and with proceeding fluence the separation between the empty and the highly populated regions becomes even sharper and the wavelength slightly increases.

The stability analysis for defect reaction systems controlled by isotropic diffusion cannot give information about the symmetry relationship between the ordered defect arrangement and the crystal structure. Anisotropic diffusion controlled by the crystal lattice symmetry must influence the process of pattern formation. Two different attempts have been made to introduce anisotropic diffusion: (i) the diffusion of the vacancies and interstitials is assumed to be asymmetric in the sense that the vacancy migrate three-dimensionally while the interstitials migrate one-dimensionally (dynamic crowdion, gliding interstitial loops) or two-dimensionally (EVANS [1987], HÄHNER and FRANK [1992], DUBINKO [1991], JÄGER and TRINKAUS [1993]), (ii) the elastic interaction between (predominantly) the interstitials and the ordering vacancy cluster introduces the crystal lattice symmetry, leads to respective anisotropic drift diffusion and thus controls the shrinkage of the clusters (BORODIN *et al.* [1993]).

The effect of one-dimensionally migrating interstitials can easily be seen in fig. 57 which shows how a mispositioned void is shifted. Arriving interstitials whose path is not blocked by neighboring voids induce shrinkage at the one side and three-dimensionally migrating excess vacancies arriving within the screened regions induce growth on the other side. In a three-dimensional crystal lattice the evolution of the void lattice would therefore start with linear ordering along the migration paths of the interstitials or with planar ordering within one of the planes containing the maximum number of the migration lines. The fundamental question with this explanation of the pattern formation concerns the structure of the one-dimensionally migrating interstitial. Stability of the dynamic crowdion has not been confirmed by the MDC studies for fcc and bcc metals

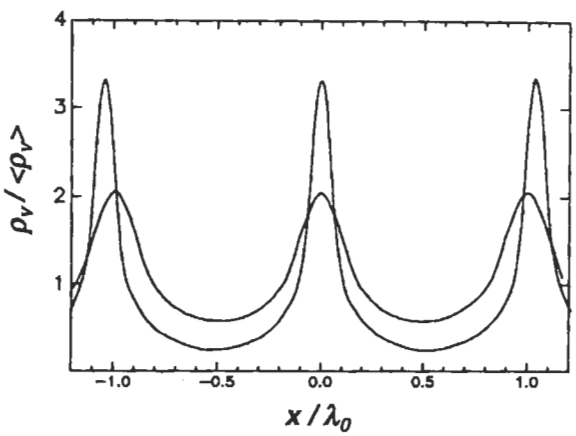


Fig. 56. Spatial variation of the loop concentration for two different maximum minimum ratios (after ABROMEIT and TRINKAUS [1992]).

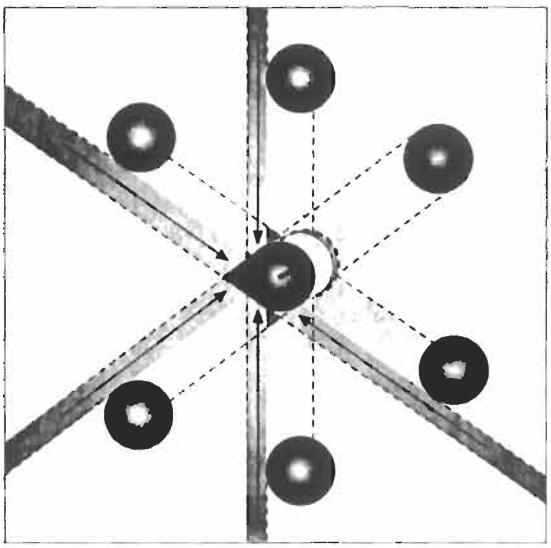


Fig. 57. Rearrangement of voids by one-dimensionally migrating interstitials (after FOREMAN [1972]).

up to now (see § 3.3). In addition, the difference of the formation enthalpies for the dumbbell interstitial and the crowdion were found to only measure a few tens of an eV. Such values are at variance to the values required for the explanation of void lattices observed at the upper temperature limit and the observed maximum void lattice parameters. The mean free path of the crowdion between its creation and conversion to a stable dumbbell interstitial must be at least of the order of the void lattice constant. The opposite tendency of the mean free path to decrease and the void lattice distance to increase with increasing temperature defines an upper temperature limit for the void lattice formation. The observed maximum void lattice distances and the upper temperature limit requires a difference between the enthalpies for interstitial conversion and migration of the crowdion of values between 0.5 and 1 eV in contrast to the findings from MDC studies. Of special interest in this regard is the behavior of Au for which the absence of a pronounced annealing stage I has been interpreted by the conversion of the crowdion to the stable dumbbell configuration at very low temperatures (FRANK and SEEGER [1983]). Therefore, the observed bubble lattice (JOHNSON *et al.* [1983] cannot be explained by a one-dimensionally migrating interstitial. Similar arguments hold for Nb. The gliding interstitial loops discussed in § 4.4 might explain the symmetry relationship in a consistent manner. For hcp metals two-dimensionally migrating interstitials are more likely to be stable and would explain the observed void lattice geometries.

The various attempts to explain the symmetry relationship by anisotropic drift diffusion due to the elastic interaction have been successful in rationalizing the symmetry relationship (MARTIN [1977]) but quantitative explanation seems to be successful for sinks with significant elastic asymmetry (as dislocation loops) only.

References: p. 1710.

4.7. Radiation-induced solute segregation

Irradiation of a crystal causes the supersaturation of vacancies and interstitials in the lattice. At irradiation temperatures such that both defect species are mobile the defects will annihilate by recombination and annihilation at sinks. For very low sink densities the annihilation rate will be controlled by the recombination reaction. As the defects migrate randomly recombinations will occur homogeneously in space. For descriptions in terms of the rate equation approach the defect annihilation is considered by the "lossy medium". This treatment neglects an important sink effect. Unsaturable sinks like surfaces or non-sessile dislocations cause persistent defect fluxes under stationary circumstances. The situation in front of such a sink is sketched in fig. 58 for a solid solution consisting of A and B atoms. If the partial diffusion coefficients D_A^v and D_B^v of the constituents A and B of a binary alloy due to vacancy migration are different, the vacancy flux J_v to the sink will cause different atom fluxes J^A and J^B with the result of

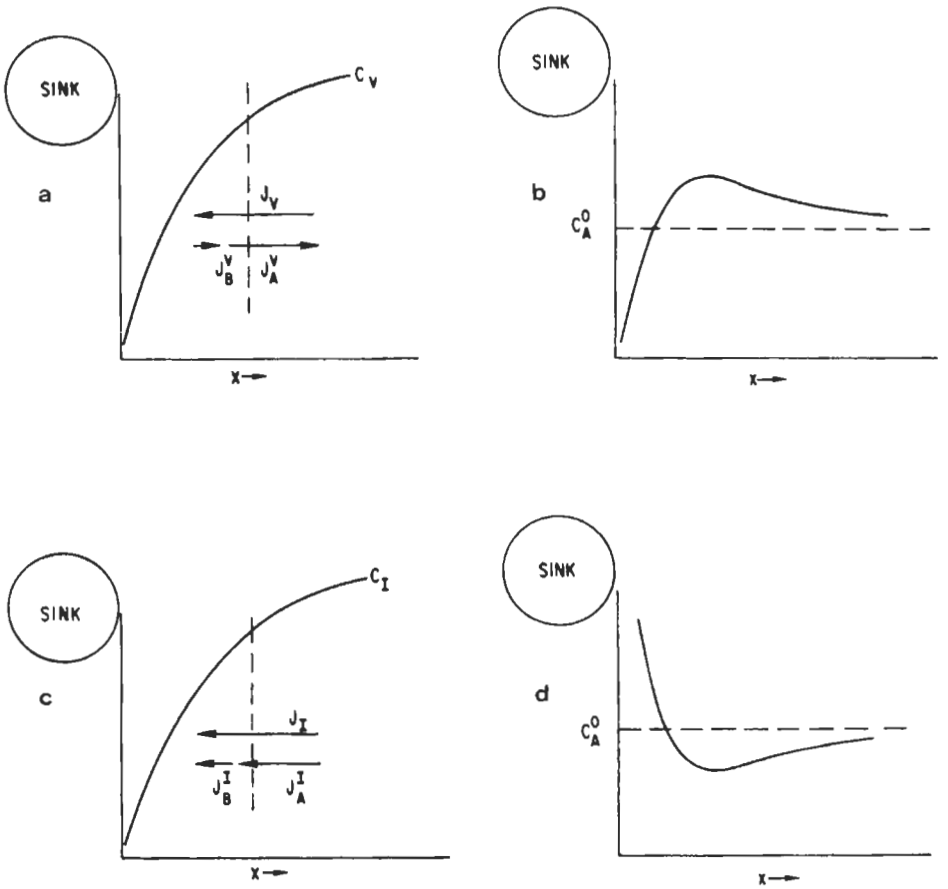


Fig. 58. Inverse Kirkendall effect produced by vacancy (a, b) and interstitial (c, d) fluxes to sinks.

a depletion of the faster transported constituent near the sink as shown in fig. 58b. The process is the inverse Kirkendall effect (SMIGELSKAS and KIRKENDALL [1947]). If the interstitial flux J_i provides atom fluxes as shown in fig. 58c constituent A will be enriched as shown in fig. 58d. For steady state conditions and $r_{vs} = r_{is}$ the concentration gradient ∇c^A was derived by WIEDERSICH *et al.* [1979] to be proportional to $[(D_A^v/D_B^v) - (D_A^i/D_B^i)]\nabla c_v$.

Atomic redistribution of this kind also occurs if one of the defect species forms a mobile tightly bound complex with one of the constituent atoms and annihilates at the sink. In this case, the complex forming constituent will be enriched at the sink. One realizes that such atomic redistributions may lead to the formation of new phases. The most simple case is the irradiation of an undersaturated solid solution of B in A. When the redistribution amplitude is large enough, the solubility limit will be exceeded near or at the sink and an intermediate phase will be formed whenever nucleation conditions allow it. At sufficiently high irradiation temperatures the process may be reversed by the action of thermal vacancies after stopping the irradiation. Theoretical description of the radiation-induced segregation has been reviewed by WIEDERSICH and LAM [1983].

For studies of precipitation near internal sinks high voltage electron microscopy has proved to be a very valuable tool. Electron energies from about 0.5 MeV upwards are sufficient to produce Frenkel defects during the imaging. The beam intensity is extremely high such that the number of atomic displacements produced by irradiation in nuclear reactors can be exceeded by orders of magnitude within minutes. Radiation-induced formation of ordered phases like Ni_3Si can thus be imaged easily by means of the dark field technique. For Cu–Be the formation of radiation-induced Guinier–Preston zones (WOLLENBERGER and KELL [1989]) has been studied by means of field ion microscopy and that of the long range ordered γ-phase CuBe by transmission electron microscopy (KOCH *et al.* [1981]). The experimental material on radiation-induced segregation was reviewed by REHN and OKAMOTO [1983].

Radiation-driven atomic redistribution is not necessarily bound to sink annihilation of the defects. If the rate of Frenkel pair recombination depends on the alloy composition, which might be the case for a strong influence of the chemical potential, the recombination rate would become inhomogeneous in space. This again would create persistent defect fluxes which could amplify existing concentration fluctuations (MARTIN [1980]). The field of phase transformations in irradiated materials has been reviewed by RUSSELL [1993].

Acknowledgement

The author is deeply indebted to Drs. C. Abromeit and V. Naundorf for many enlightening discussions and for critical reading of the manuscript.

References: p. 1710.

References

ABROMEIT, C., 1983, J. Phys. F: Met. Phys. **13**, L169.

ABROMEIT, C., 1989, Int. J. Modern Phys. **B 3**, 1301.

ABROMEIT, C., 1994, J. Nucl. Mater. **216**, 78.

ABROMEIT, C., and H. TRINKAUS, 1992, Solid State Phenomena **23** & **24**, 181.

ABROMEIT, C., and H. WOLLENBERGER, 1983, Phil. Mag. **A 6**, 951.

ADAMS, J. B., and S. M. FOILES, 1990, Phys. Rev. **B 41**, 3316.

ADLHART, W., G. FRITSCH and E. LÜSCHER, 1975, J. Phys. Chem. Solids **36**, 1405.

APPLEBY, W. K., D. W. SANDUSKY and U. E. WOLFF, 1972, Fluence and Temperature Dependence of Void Formation in Highly Irradiated Stainless Steels, in: Radiation-Induced Voids in Metals, eds. J. W. Corbett and L. C. Ianniello (Conf-710601, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 156.

AUDIT, P., 1982, Vacancy-Thermal Expansion in Aluminium, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 291.

AVERBACK, R. S., R. BENEDEK and K. L. MERKLE, 1978, Phys. Rev. **B 18**, 4156.

BAKKER, H., 1987, Mater. Sci. Forum **15-18**, 1155.

BALLUFFI, R. W., 1978, J. Nucl. Mater. **69/70**, 240.

BALLUFFI, R. W., K. H. LIE, D. N. SEIDMAN and R. W. SIEGEL, 1970, Determination of Concentrations and Formation Energies and Entropies of Vacancy Defects from Quenching Experiments, in: Vacancies and Interstitials in Metals, eds. A. Seeger, D. Schumacher, W. Schilling and J. Diehl (NorthHolland, Amsterdam) p. 125.

BALZER, R., and H. SIGVALDASON, 1979, Phys. Stat. Sol. (b) **92**, 143.

BASS, J., 1967, Phil. Mag. **15**, 717.

BAUER, W., 1969, Radiation Effects **1**, 23.

BECKER, D. E., F. DWORSCHAK and H. WOLLENBERGER, 1973, Radiation Effects **17**, 25.

BEELER, J. R., Jr., 1983, in: Radiation Effects Computer Experiments (North-Holland Publ. Co., Amsterdam).

BENDER, O., and P. EHRHART, 1983, J. Phys. **F.13**, 911.

BENEDEK, R., 1978, J. Phys. F: Met. Phys. **17**, 569.

BENNEMANN, K. H., and L. TEWORDT, 1960, Z. Naturf. **15a**, 772.

BERGER, A. S., D. N. SEIDMAN and R. W. BALLUFFI, 1973, Acta Metall. **21**, 123.

BERGER, A. S., S. T. OCKERS and R. W. SIEGEL, 1979, J. Phys. **F 9**, 1023.

BERGER, A. S., S. T. OCKERS, M. K. CHASON and R. W. SIEGEL, 1978, J. Nucl. Mater. **69/70**, 734.

BIGET, M., F. MAURY, P. VAJDA, A. LUCASSON and P. LUCASSON, 1979, Phys. Rev. **B 19**, 820.

BILGER, H., V. HIVERT, J. VERDONE, J. L. LEVEQUE and J. C. SOULIE, 1968, Point Defects in Iron, in: International Conference on Vacancies and Interstitials in Metals, Jülich (Jül-Conf-2, p. 751), unpublished.

BIRTCHER, R. C., W. HERTZ, G. FRITSCH and J. F. WATSON, 1975, Very Low Temperature Electron-Irradiation and Annealing of Gold and Lead, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 405.

BLEWITT, T. H., 1962, in: Radiation Damage in Solids, ed. D. S. Billington (Academic, New York) p. 630.

BORODIN, V. A., A. I. RYAZANOV, and C. ABROMEIT, 1993, J. Nucl. Mater. **207**, 242.

BRAILSFORD, A. D., and R. BULLOUGH, 1978, J. Nucl. Mater. **69/70**, 434.

BRANDT, W., 1974, Appl. Phys. **4**, 1.

BRANDT, W., and H. F. WAUNG, 1968, Phys. Lett. **27A**, 100.

BULLOUGH, R., and J. R. WILLIS, 1975, Phil. Mag. **31**, 855.

BULLOUGH, R., and M. H. WOOD, 1980, J. Nucl. Mater. **90**, 1.

BURTON, J. J., 1971, Phys. Rev. **B 5**, 2948.

BUTT, R., R. KEITEL and G. VOGL, 1979, HMI-Report, unpublished.

CANNON, C. P., and A. SOSIN, 1975, Radiation Effects **25**, 253.

CARO, A. and M. VICTORIA, 1989, Phys. Rev. **A 40**, 2287.

CAWTHORNE, C., and E. J. FULTON, 1967, Nature **216**, 515.

CHAMBRON, W., J. VERDONE and P. MOSER, 1975, Determination of Point Defects Symmetry in a Cubic Lattice by Magnetic Relaxation, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 261.

COLTMAN, R. R., C. E. KLABUNDE, J. K. REDMAN and A. L. SOUTHERN, 1971, Radiation Effects **7**, 235.

CORBEL, C., P. BOIS, P. MOSER and I. LEMAHIEU, 1987, Mater. Sci. Forum **15-18**, 721.

CORBETT, J. W., R. B. SMITH and R. M. WALKER, 1959, Phys. Rev. **114**, 1442 and 1460.

COTTERILL, R. M. J. and M. DOYAMA, 1966, Energies and Atomic Configurations of Line Defects and Plane Defects in fcc Metals, in: Lattice Defects and Their Interaction, ed. R. R. Hasiguti (Gordon and Breach, New York) p. 1.

DANDER, W., and H. E. SCHAEFER, 1977, Phys. Stat. Sol. (b) **80**, 173.

DAOU, J. N., P. VAJDA, A. LUCASSON, P. LUCASSON and J. P. BURGER, 1985, Phys. Lett. **107A**, 142.

DAUSINGER, F., 1978, Phil. Mag. **A 37**, 819.

DE DIEGO, N., C. HIDALGO and P. MOSER, 1987, Mater. Sci. Forum **15-18**, 193.

DE SCHEPPER, L., D. SEGERS, L. DORIKENS-VANPRAET, M. DORIKENS, G. KNUYT, L. M. STALS and P. MOSER, 1983, Phys. Rev. **B 27**, 5257.

DEDERICHS, P. H., 1973, J. Phys. **F. 3**, 471.

DEDERICHS, P. H., and J. POLLMANN 1972, Z. Phys. **255**, 315.

DEDERICHS, P. H., and R. ZELLER 1980. Dynamical Properties of Point Defects in Metals, in: Springer Tracts in Modern Physics, Vol. 87, eds. G. Höhler and E. A. Niekisch (Springer. Berlin) ch. 6.4.

DEDERICHS, P. H., C. LEHMANN and A. SCHOLZ, 1973, Phys. Rev. Lett. **31**, 1130.

DEDERICHS, P. H., C. LEHMANN, H. R. SCHOBER, A. SCHOLZ and R. ZELLER, 1978, J. Nucl. Mater. **69/70**, 176.

DEICHER, M., E. RECKNAGEL and TH. WICHERT, 1981a, Radiation Effects **54**, 155.

DEICHER, M., O. ECHT, E. RECKNAGEL and TH. WICHERT, 1981b, Geometrical Structure of Lattice Defect-Impurity Configurations Determined by TDPAC, in: Nuclear and Electron Resonance Spectroscopies Applied to Materials Science, eds. E. N. Kaufmann and G. K. Shenoy (Elsevier-North-Holland, New York) p. 435.

DE NOVION, C., 1995, in: Intermetallic Compounds — Principles and Practice, eds. J. H. WESTBROOK and R. L. FLEISCHER (Wiley, Chichester), Vol. 1, p. 559.

DIAZ DE LA RUBIA, T. and M. W. GUINAN, 1992, Mater. Sci. Forum **97-99**, 23.

DIBBERT, H. J., K. SONNENBERG, W. SCHILLING and U. DEDEK, 1972, Radiation Effects **15**, 115.

DICKMAN, J. E., R. N. JEFFERY and D. R. GUSTAFSON, 1977, Phys. Rev. **B 16**, 3334.

DICKMAN, J. E., R. N. JEFFERY and D. R. GUSTAFSON, 1978, J. Nucl. Mater. **69/70**, 604.

DIFFERT, K., A. SEEGER and W. TROST, 1987, Mater. Sci. Forum **15-18**, 99.

DIMITROV, O., and C. DIMITROV, 1985, Radiation Effects **84**, 117.

DOYAMA, M., 1978, J. Nucl. Mater. **69/70**, 350.

DOYAMA, M., and J. S. KOEHLER, 1962, Phys. Rev. **127**, 21.

DOYAMA, M., and R. M. J. COTTERILL, 1967, Energies and Atomic Configurations of Point Defects in fcc Metals, in: Lattice Defects and Their Interactions, ed. R. R. Hasiguti (Gordon and Breach, New York) p. 79.

DOYAMA, M., and R. R. HASIGUTI, 1973, Cryst. Lattice Defects **4**, 139.

DRITTLER, K., H. J. LAHANN and H. WOLLENBERGER, 1969, Radiation Effects **2**, 51.

DUBINKO, V. I., 1991, J. Nucl. Mater. **178**, 108.

DUESING, G., and W. SCHILLING, 1969, Radiation Effects **1**, 65.

DUNLOP, A., D. LESUEUR and A. BARBU, 1993, J. Nucl. Mater. **205**, 426.

DWORSCHAK, F., G. HOLFELDER and H. WOLLENBERGER, 1981, Radiation Effects **59**, 35.

DWORSCHAK, F., H. WAGNER and P. WOMBACHER, 1972, Phys. Stat. Sol. (b) **52**, 103.

DWORSCHAK, F., R. LENNARTZ, J. SELKE and H. WOLLENBERGER, 1978, J. Nucl. Mater. **69/70**, 748.

DWORSCHAK, F., TH. MONSAU and H. WOLLENBERGER, 1976, J. Phys. **F 6**, 2207.

ECKER, K. H., 1982, Verhandl. DPG (VI) **17**, 892.

EHRHART, P., 1978, J. Nucl. Mater. **69/70**, 200.

EHRHART, P., 1985, Mater. Res. Soc. Symp. **41**, 13.

EHRHART, P., 1991, in: Atomic Defects in Metals, ed. H. Ullmaier, Landoldt-Börnstein, ed. O. Madelung, Group III, Vol. 25 (Springer Berlin) p. 88.

EHRHART, P., and B. SCHÖNFELD, 1979, Phys. Rev. **B 19**, 3896.

EHRHART, P., and B. SCHÖNFELD, 1982, Self-Interstitial Atoms and their Agglomerates in hcp Metals, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 47.

EHRHART, P., and E. SEGURA, 1975, X-ray Investigation of Interstitials and Interstitial Clusters after Low Temperature Electron-Irradiation and Thermal Annealing of Gold, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 295.

EHRHART, P., and W. SCHILLING, 1974, Phys. Rev. **B 8**, 2604.

EHRHART, P., B. SCHÖNFELD and K. SONNENBERG, 1982, Agglomerates of Interstitial Atoms and Vacancies in Electron-Irradiated Copper, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 687.

EHRHART, P., H. D. CARSTANJEN, A. M. FATTAH and J. B. ROBERTO, 1979, Phil. Mag. **A 40**, 843.

EHRHART, P., H. G. HAUBOLD and W. SCHILLING, 1974, Investigation of Point Defects and Their Agglomerates in Irradiated Metals by Diffuse X-ray Scattering, in: Festkörperprobleme XIV/ Advances in Solid State Physics, ed. H. J. Queisser (Vieweg, Braunschweig) p. 87.

EHRHART, P., K.-H. ROBROCK and H. R. SCHOBER, 1986, in: Physics of Radiation Effects in Crystals, eds. R. A. Johnson and A. N. Orlov, Elsevier Science Publishers, p. 3.

EMRICK, R. M., 1982, J. Phys. **F 12**, 1327.

ENGLISH, C. A., A. J. E. FOREMAN, W. J. PHYTHIAN, D. J. BACON and M. L. JENKINS, 1992, Mater. Sci. Forum **97-99**, 1.

ERGINSOY, C., G. H. VINEYARD and A. ENGLERT, 1964, Phys. Rev. **133**, A595.

ESHELBY, J. D., 1956, The Continuum Theory of Lattice Defects, in: Solid State Physics, 3, eds. F. Seitz and D. Turnbull (Academic, New York) p. 79.

EVANS, J. H., 1987, Materials Science Forum **15-18**, 869.

EVANS, R., 1977, Calculation of Point Defect Formation Energies in Metals, in: Vacancies '76, eds. R. E. Smallman and J. E. Harris (The Metals Society, London) p. 30.

EYRE, B. L., and J. R. MATTHEWS, 1993, J. Nucl. Mater. **205**, 1.

EYRE, B. L., M. H. LORETTO and R. E. SMALLMAN, 1977, Electron Microscopy Studies of Point Defect Clusters in Metals, in: Vacancies '76, eds. R. E. Smallman and J. E. Harris (The Metals Society, London) p. 63.

FABER, K., and H. SCHULTZ, 1977, Radiation Effects **31**, 157.

FABER, K., J. SCHWEIKHARDT and H. SCHULTZ, 1974, Scripta Metall. **8**, 713.

FEDER, R., and A. S. NOWICK, 1967, Phil. Mag. **15**, 805.

FEESE, K., D. HOFFMANN and H. WOLLENBERGER, 1970, Cryst. Lattice Defects **1**, 245.

FINNIS, M., P. AGNEW and A. J. E. FOREMAN, 1991, Phys. Rev. **B 44**, 567.

FLINN, J. E., and T. A. KENFIELD, 1976, Neutron Swelling Observations on Austenitic Stainless Steels Irradiated in EBR-II, in: Correlation of Neutron and Charged Particle Damage, Proc. of the workshop in Oak Ridge, ed. J. O. Stiegler (CONF-760673, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 253.

FLYNN, C. P., 1968, Phys. Rev. **171**, 682.

FLYNN, C. P., 1975, Thin Solid Films **25**, 37.

FLYNN, C. P., J. BASS and D. LAZARUS, 1965, Phil. Mag. **11**, 521.

FOILES, S., M. I. BASKES and M. S. DAW, 1986, Phys. Rev. **B33**, 7983.

FOLWEILER, R. C., and F. R. BROTZEN, 1959, Acta Metall. **7**, 716.

FOREMAN, A. J. E., 1972, Harwell Report, AERE-R 7135.

FOREMAN, A. J. E., C. A. ENGLISH and W. J. PHYTHIAN, 1991, AEA Technology Harwell Report, AEA-TRS-2028 and 2031.

FRANK, W., A. SEEGER and R. SCHINDLER, 1979, Radiation Effects **40**, 239.

FRANK, W., and A. SEEGER, 1983, Radiation Effects **71**, 157.

FRIEDEL, J., 1970, Theory of Point Defects in Metals, in: Vacancies and Interstitials in Metals, eds. A. Seeger, D. Schumacher, W. Schilling and J. Diehl (North-Holland, Amsterdam) p. 787.

FÜRDERER, K., K.-P. DÖRING, M. GLADISCH, N. HAAS, D. HERLACH, J. MAJOR, H.-J. MUNDINGER, J. ROSENKRANZ, W. SCHAFER, L. SCHIMMELE, M. SCHMOLZ, W. SCHWARZ and A. SEEGER, 1987, Mater. Sci. Forum **15-18**, 125.

GABRIEL, T. A., J. D. AMBURGEY and N. M. GREEWE, 1976, Nucl. Sci. and Eng. **61**, 21.

GANNE, J. P., and Y. QUÉRÉ, 1982, Intrinsic Thermal Expansion of Point Defects in Metals, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, 1982) p. 232.

GAUSTER, W. B., S. MANTL, T. SCHOBER and W. TRIFTSHÄUSER, 1975, Annealing of Dislocation Loops in Neutron-Irradiated Copper Investigated by Positron Annihilation, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 1143.

GEHLEN, P. C., I. R. BEELER and R. L. JAFFEE, eds., 1972, Interatomic Potentials and Simulation of Lattice Defects (Plenum, New York).

GIBSON, J. B., A. N. GOLAND, M. MILGRAM and G. H. VINEYARD, 1960, Phys. Rev. **120**, 1229.

GLASGOW, B. B., A. SI-AHMED, W. G. WOLFER and F. A. GARNER, 1981, J. Nucl. Mater. **103/104**, 981.

GRANATO, A. V., and T. G. NILAN, 1965, Phys. Rev. **137**, A1250.

GRANATO, A. V., J. HOLDER, K. L. HULTMAN, D. L. JOHNSON, G. G. SETSER, P. WALLACE and H. WONG, 1982, in: Point Defects and Defect Interaction in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (University of Tokyo Press, Tokyo) p. 360.

GRANATO, A. V., 1993, J. Non-Crystalline Solids **156-158**, 402.

GUINAN, M. W., C. L. SNEAD and A. N. GOLAND, 1973, Radiation Effects **20**, 33.

HÄHNER, P., and W. FRANK, 1992, Solid State Phenomena **23 & 24**, 203.

HAKEN, H., 1977, Synergetics (Springer Verlag, Berlin).

HARA, Y., and S. NANAO, 1982, in: Point Defects and Interactions in Metals, Proc. of Yamada Conf. V, eds. J. Takamura, M. Doyama, M. Kiritani (University of Tokyo Press, Tokyo) p. 595.

HARDY, J. R., and R. BULLOUGH, 1967, Phil. Mag. **15**, 237.

HATCHER, R., D. R. ZELLER and P. H. DEDERICHS, 1979, Phys. Rev. **B 19**, 5083.

HAUBOLD, H.-G., 1975, Study of Irradiation-Induced Point Defects by Diffuse Scattering, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 268.

HAUBOLD, H.-G., 1976, Rev. Physique Appl. **11**, 73.

HAUBOLD, H.-G., and D. MARTINSEN, 1978, J. Nucl. Mater. **69/70**, 644.

HAUTOJÄRVI, P., 1987, Mater. Sci. Forum **15-18**, 81.

HEALD, P. T., 1977, Discrete Lattice Models of Point Defects, in: Vacancies '76, eds. R. E. Smallman and J. E. Harris (The Metals Society, London) p. 11.

HEALD, P. T., and M. V. SPEIGHT, 1974, Phil. Mag. **29**, 1075.

HEINISCH, H., 1981, J. Nucl. Mater. **103/104**, 1325.

HEINISCH, H., 1990, Radiation Effects and Defects in Solids **113**, 53.

HERTZ, W., and H. PEISL, 1975, J. Phys. **F 5**, 2241.

HETTICH, G., H. MEHRER and K. MAIER, 1977, Scripta Metall. **11**, 795.

HOCH, M., 1970, Equilibrium Measurements in High-Melting-Point Materials, in: Vacancies and Interstitials in Metals, eds. A. Seeger, D. Schumacher, W. Schilling and J. Diehl (North-Holland, Amsterdam) p. 81.

HODGES, C. H., 1970, Phys. Rev. Lett. **25**, 285.

HOLDER, J., A. V. GRANATO and L. E. REHN, 1974, Phys. Rev. **B10**, 363.

HOOD, G. M., 1986, J. Nucl. Mater. **139**, 179.

HOOD, G. M., and R. J. SCHULTZ, 1988, J. Nucl. Mater. **151**, 172.

HOOD, G. M., R. J. SCHULTZ and J. A. JACKMAN, 1984, J. Nucl. Mater. **126**, 79.

HORSEWELL, A. and B. N. SINGH, 1988, ASTM STP **955**, 220.

HOWE, M., and M. L. SWANSON, 1982, Ion Channeling Investigations of the Interactions between Irradiation-produced defects and solute atoms in metals, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 53.

HULTMAN, K. L., J. HOLDER and A. V. GRANATO, 1981, J. Physique, Colloque C5-42, Suppl. **10**, 753.

HUNTINGTON, H. B., 1942, Phys. Rev. **61**, 325.

HUNTINGTON, H. B., 1953, Phys. Rev. **91**, 1092.
IMAFUKU, M., R. YAMAMOTO and M. DOYAMA, 1982, Computer Studies of Self-Interstitials in Magnesium, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 145.
INGLE, K. W., R. C. PERRIN and H. R. SCHOBER, 1981, J. Phys. **F 11**, 1161.
ISHINO, S., 1993, J. Nucl. Mater. **206**, 139.
ISHINO, S., S. IWATA, Y. MATSUTANI and S. TANAKA, 1977, Computer Simulation and Neutron, Heavy-Ion and Electron Irradiation Correlation, in: Radiation Effects in Breeder Reactor Structural Materials, Scottsdale, eds. M. L. Bleiberg and J. W. Bennett (Met. Soc. AIME, Warrendale, PA) p. 879.
ITO, K., Y. OHTSU and S. TANIGAWA, 1982, in: Positron Annihilation, Proc. 6th Int. Conf. on Positron Ann., eds. P. G. Coleman, S. C. Sharma, L. M. Diana (North-Holland Publ. Co., Amsterdam) p. 514.
IWATA, T., and A. IWASE, 1990, Radiation Effects and Defects in Solids **113**, 135.
JACKSON, J. J., 1979, Phys. Rev. **B 20**, 534.
JÄGER, W., 1981, J. Microsc. Spectrosc. Electron. **6**, 437.
JÄGER, W., and H. TRINKAUS, 1993, J. Nucl. Mater. 205, 394.
JÄGER, W., T. P. EHRHART and W. SCHILLING, 1990, Radiation Effects and Defects in Solids **113**, 201.
JANOT, C., and B. GEORGE, 1975, Phys. Rev. **B 12**, 2212.
JANOT, C., B. GEORGE and P. DELCROIX, 1982, J. Phys. **F 12**, 47.
JANOT, C., D. MALLEJAC and B. GEORGE, 1970, Phys. Rev. **B 2**, 3088.
JENKINS, M. L., and M. WILKENS, 1976, Phil. Mag. **34**, 1155.
JENKINS, M. L., M. A. KIRK and W. J. PHYTHIAN, 1993, J. Nucl. Mater. **205**, 16.
JOHNSON, P. B., D. J. MAZEI, and J. H. EVANS, 1983, Radiation Effects **78**, 147.
JOHNSON, R. A., 1964, Phys. Rev. **134**, A1329.
JOHNSON, R. A., 1965, J. Phys. Chem. Solids **26**, 75.
JOHNSON, R. A., 1966, Phys. Rev. **145**, 423.
JOHNSON, R. A., 1968, Phys. Rev. **174**, 684.
JOHNSON, R. A., 1970, Phys. Rev. **B 1**, 3956.
JOHNSON, R. A., 1973, J. Phys. **F 3**, 295.
JOHNSON, R. A., and E. BROWN, 1962, Phys. Rev. **127**, 446.
JOHNSON, R. A., and J. R. BEELER, 1977, Phys. Rev. **156**, 677.
JOHNSTON, W. G., T. LAURITZEN, J. W. ROSOLOWSKI and A. M. TURKALO, 1976, J. Metals **28**, 19.
JUNG, P., 1981a, Phys. Rev. **B 23**, 664.
JUNG, P., 1981b, Radiation Effects **59**, 103.
JUNG, P., 1991, in: Atomic Defects in Metals, ed. H. Ullmaier, Landoldt-Börnstein, ed. O. Madelung, Group III, Vol. 25 (Springer Verlag, Berlin) p. 1.
JUNG, P., R. L. CHAPLIN, H. J. FENZL, K. REICHELT and P. WOMBACHER, 1973, Phys. Rev. **B 8**, 553.
KANZAKI, H. J., 1957, J. Phys. Chem. Solids **2**, 24.
KAPINOS, V. G., Y. U. N. OSETSKII and P. A. PLATANOV, 1990, J. Nucl. Mater. **173**, 229.
KAUFFMAN, J. W., and J. S. KOEHLER, 1952, Phys. Rev. **88**, 149.
KHELLAF, A., R. M. EMRICK and J. J. VUILLEMIN, 1988, Phys. Rev. **B 37**, 6717.
KIENLE, W., W. FRANK and A. SEEGER, 1983, Radiation Effects **71**, 163.
KIM, S. M., 1986, Phys. Rev. **B 33**, 1509.
KIM, S. M., and A. T. STEWART, 1975, Phys. Rev. **B 11**, 4012.
KIM, S. M., J. A. JACKMAN, W. J. L. BUYERS and D. T. PETERSON, 1984, J. Phys. **F 14**, 2323.
KINCHIN, G. H., and R. S. PEASE, 1955, Rep. Prog. Phys. **18**, 1.
KING, W. E., and R. BENEDEK, 1981, Phys. Rev. **B 23**, 6335.
KING, W. E., K. L. MERKLE and M. MESHII, 1981, Phys. Rev. **B 23**, 6319.
KIRITANI, M., 1982, Nature of Point Defects and Their Interactions Revealed by Electron Microscope Observation of Their Clusters, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 59.
KIRITANI, M., T. YOSHIIE, S. KOJIMA, Y. SATOH and K. HAMADA, 1990, J. Nucl. Mater. **174**, 327.
KIRITANI, M., 1993, J. Nucl. Mater. **206**, 156.
KIRK, M. A., T. H. BLEWITT and T. L. SCOTT, 1978, J. Nucl. Mater. **69/70**, 780.

KITAJIMA, K., Y. AKASHI and E. KURAMOTO, 1985, J. Nucl. Mater. **133&134**, 486.
KLAUMÜNZER, S., 1992, Mater. Sci. Forum **97-99**, 623.
KLEMRADT, U., B. DRITTLER, T. HOSHINO, R. ZELLER, P. H. DEDERICHS and N. STEFANOU, 1991, Phys. Rev. **B 43**, 9487.
KLUIN, J.-E., and T. HEHENKAMP, 1991, Phys. Rev. **B 44**, 11597.
KNODLE, W. C., and J. S. KOEHLER, 1978, J. Nucl. Mater. **69/70**, 620.
KNÖLL, H., U. DEDEK and W. SCHILLING, 1974, J. Phys. **F 4**, 1095.
KOCH, J. M., and C. KOENIG, 1986, Phil. Mag. **B 54**, 177.
KOCH, R., R. P. WAHI and H. WOLLENBERGER, 1981, J. Nucl. Mater. **103/104**, 1211.
KOLLERS, G., H. JACQUES, L. E. REHN and K. H. ROBROCK, 1981, J. Physique, Colloque **C5-42**, Suppl. 10, 729.
KÖNIG, D., J. VÖLKL and W. SCHILLING, 1964, Phys. Stat. Sol. **7**, 591.
KÖSTER, W., and H. P. KEHRER, 1965, Z. Metallkunde **56**, 760.
KRISHAN, K., 1982, Radiation Effects **66**, 121.
KRONMÜLLER, H., 1970, Studies of Point Defects in Metals by Means of Mechanical and Magnetic Relaxation, in: Vacancies and Interstitials in Metals, eds. A. Seeger, D. Schumacher, W. Schilling and J. Diehl (North-Holland, Amsterdam) p. 667.
KUGLER, H., I. A. SCHWIRTLICH, S. TAKAKI, K. YAMAKAWA, U. ZIEBART, J. PETZOLD and H. SCHULTZ, 1982, Stage III Recovery in Electron-Irradiated bcc Transition Metals, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 520.
KURAMOTO, E., 1992, J. Nucl. Mater. **191-194**, 1297.
KUSMISS, J. H., and A. T. STEWART, 1967, Adv. Phys. **16**, 471.
LAM, N. Q., N. V. DOAN and Y. ADDA, 1980, J. Phys. **F 10**, 2359.
LAM, N. Q., N. V. DOAN, L. DAGENS and Y. ADDA, 1981, J. Phys. **F 11**, 2231.
LARSON, B. C., and F. W. YOUNG, Jr., 1982, Vacancy and Interstitial Loops in Irradiated Metals, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 679.
LAUPHEIMER, A., W. FRANK, M. RÜHLE, A. SEEGER and M. WILKENS, 1981, Radiation Effects Lett. **67**, 95.
LEADBETTER, A. J., D. M. T. NEWSHAM and N. H. PICTON, 1966, Phil. Mag. **13**, 371.
LEE, C., and J. S. KOEHLER, 1968, Phys. Rev. **176**, 813.
LEGRAND, P., J. MORILLO and V. PONTIKIS, 1993, Radiation Effects and Defects in Solids **126**, 151.
LEHMANN, C., 1977, Interaction of Radiation with Solids, in: Defects in Crystalline Solids, vol. 10, eds. S. Amelinckx, R. Gevers and J. Nihoul (North-Holland, Amsterdam).
LEHMANN, C., A. SCHOLZ and W. SCHILLING, 1981, Computer Simulation of Radiation-Induced Defects, film (16 mm, colour, 32 min) produced and distributed by EMW. Huschert Filmstudio (Weststr. 34-36, D-40597, Düsseldorf-Benrath, Germany).
LEIBFRIED, G., 1965, Bestrahlungseffekte in Festkörpern (Teubner, Stuttgart).
LEIBFRIED, G., and N. BREUER, 1978, Point Defects in Metals I, in: Springer Tracts in Modern Physics, vol. 81, eds. G. Höhler and E. A. Niekisch (Springer, Berlin).
LENGELER, B., 1976, Phil. Mag. **34**, 259.
LENGELER, B., and R. R. BOURASSA, 1976, J. Phys. **F 6**, 1405.
LENNARTZ, R., F. DWORSCHAK and H. WOLLENBERGER, 1977, J. Phys. **F 7**, 2011.
LEVY, V., J. M. LANORE and J. HILLAIRET, 1973, Phil. Mag. **28**, 373.
LITTLE, E. A., 1993, J. Nucl. Mater., **206**, 324.
LIU PING and G. L. DUNLOP, 1988, J. Mater. Sci. **23**, 1419.
LOPER, G. D., L. C. SMEDSKJAER, M. K. CHASON and R. W. SIEGEL, 1985, in: Positron Annihilation, eds. P. C. Jain, R. M. Singru, K. P. Gopinathan (World Scientific Publ. Co., Singapore) p. 461.
LOSEHAND, R., F. RAU and H. WENZL, 1969, Radiation Effects **2**, 69.
LUCAS, G. E., 1993, J. Nucl. Mater. **206**, 287.
LUCASSON, P. G., 1975, in: Fundamental Aspects of Radiation Damage in Metals, Proc. Int. Conf. Gatlinburg, Tenn., USA, eds. M. T. Robinson and F. W. Young (USERDA Conf. 751006P1/P2) p. 42.
LUCASSON, P. G., and R. M. WALKER, 1962, Phys. Rev. **127**, 485.
LWIN, Y. N., M. DOYAMA and J. S. KOEHLER, 1968, Phys. Rev. **165**, 787.
MACKENZIE, J. K., T. W. CRAIG and B. T. A. MCKEE, 1971, Phys. Lett. **33A**, 227.

MAIER, K., H. MEHRER and G. REIN, 1979c, Z. Metallkunde **70**, 271.
MAIER, K., H. METZ, D. HERLACH, H. E. SCHAEFER and A. SEEGER, 1978, J. Nucl. Mater. **69/70**, 589.
MANNINEN, M., R. NIEMINEN, P. HAUTOJÄRVI and J. ARPONEN, 1975, Phys. Rev. **B 12**, 4012.
MANSEL, W., J. MARANGOS and D. WAHL, 1982, J. Nucl. Mater. **108/109**, 137.
MANSUR, L. K., 1979, Phil. Mag. **A 39**, 497.
MANSUR, L. K., 1992, Materials Science Forum **97-99**, 489.
MANSUR, L. K., 1993, J. Nucl. Mater. **206**, 306.
MANTL, S. and W. TRIFTSHÄUSER, 1978, Phys. Rev. **B 17**, 1645.
MARADUDIN, A. A., E. W. MONTROLL, G. H. WEISS and I. P. IPATOVA, 1971, Theory of Lattice Dynamics in the Harmonic Approximation, Solid State Phys., Suppl. 111, eds. H. Ehrenreich, F. Seitz and D. Turnbull (Academic, New York).
MARTIN, G., 1975, Phil. Mag. **32**, 615.
MARTIN, G., 1977, J. de Physique, Colloque C-7, suppl. 4, **38**, 419.
MARTIN, G., 1980, Phys. Rev. **B 21**, 2122.
MASAMURA, R. A., and G. SINES, 1970, J. Appl. Phys. **41**, 3930.
MATSUNO, N., 1977, J. Phys. Soc. Jpn. **42**, 1675.
MAURY, F., A. LUCASSON, P. LUCASSON, J. LE HERICY, P. VAJDA, C. DIMITROV and O. DIMITROV, 1980, Radiation Effects **51**, 57.
MAURY, F., M. BIGET, P. VAJDA, A. LUCASSON and P. LUCASSON, 1978, Radiation Effects **38**, 53.
MCILWAIN, J., R. GARDINER, A. SOSIN and S. MYHRA, 1975, Radiation Effects **24**, 19.
MEECHAN, C. J., and A. SOSIN, 1959, Phys. Rev. **113**, 422.
METZNER, H., R. SIELEMANN, S: KLAUMÜNZER and E. HUNGER, 1987, Mater. Sci. Forum **15-18**, 1063.
MIEDEMA, A. R., 1979, Z. Metallkunde **70**, 345.
MIJNARENDS, P. E., 1979, Electron Momentum Densities in Metals and Alloys, in: Topics in Current Physics 12, ed. P. Hautojärvi (Springer, Berlin) ch. 2.
MISEK, K., 1979, Czech. J. Phys. **B 29**, 1243.
MONSAU, TH., and H. WOLLENBERGER, 1980, unpublished results.
MONTI, A. M., and E. J. SAVINO, 1981, Phys. Rev. **B 23**, 6494.
MOSER, P., 1966, Mém. Sci. Rev. Métallurg. **63**, 431.
MOTT, N. F., 1932, Proc. Roy. Soc. **A 135**, 429.
MÜLLER, H.-G., 1979, thesis, Univ. of Bonn, unpublished.
MÜLLER, M., 1986, Reports of the Hahn-Meitner-Institut Berlin, B 433.
MUNDY, J. N., S. J. ROTHMAN, N. Q. LAM, H. A. HOFF and L. J. NOWICKI, 1978, Phys. Rev. **B 18**, 6566.
MUNDY, J. N., S. T. OCKERS and L. C. SMEDSKYAER, 1987, Mater. Sci. Forum **15-18**, 199.
MYHRA, S., and R. B. GARDINER, 1973, Radiation Effects **18**, 39.
MYHRA, S., and R. B. GARDINER, 1975, Radiation Effects **27**, 35.
NEELY, H. H., 1970, Radiation Effects **3**, 189.
NICKLOW, R. M., W. P. CRUMMETT and J. M. WILLIAMS, 1979, Phys. Rev. **B 20**, 5034.
NIESEN, L., 1981, Hyperfine Interactions **10**, 619.
NORGETT, M. J., M. T. ROBINSON and I. M. TORRENS, 1974, Nucl. Eng. Design **33**, 50.
NOWICK, A. S., and B. S. BERRY, 1972, Anelastic Relaxation in Crystalline Solids, in: Mater. Science Series, eds. A. M. Alper, J. L. Margrave and A. S. Nowick (Academic, New York).
O'NEAL, T. N., and R. C. CHAPLIN, 1972, Phys. Rev. **B 5**, 3810.
OKUDA, S., 1975, Experimental Studies on Self-Interstitials in bcc Metals, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 361.
PARK, J. Y., H.-C. HUANG, R. W. SIEGEL and R. W. BALLUFFI, 1983, Philos. Mag. **A 48**, 397.
PARKER, C. A., and K. C. RUSSELL, 1982, in: Effects of Radiation on Materials, eds. H. R. Brager and J. S. Perrin (ASTM Publ. 04-78 2000-35, Philadelphia, PA.) p. 1042.
PETRY, W., G. VOGL and W. MANSEL, 1982, Z. Phys. **B 46**, 319.
PICHON, R., E. BISOGNI and P. MOSER, 1973, Radiation Effects **20**, 159.
PLEITER, F., and C. HOHENEMSER, 1982, Phys. Rev. **B 25**, 106.

POTTER, D. I., L. E. REHN, P. R. OKAMOTO and H. WIEDERSICH, 1977, Void Swelling and Segregation in Dilute Nickel Alloys, in: Radiation Effects In Breeder Structural Materials (AIME, New York) p. 377.

PROFANT, M., and H. WOLLENBERGER, 1976, Phys. Stat. Sol. (b) **71**, 515.

RECKNAGEL, E., G. SCHATZ and TH. WIECHERT, 1983, in: Hyperfine Interactions of Radioactive Nuclei, Topics in Current Physics 31, ed. J. Christiansen (Springer Berlin) p. 133.

REHN, L. E. and P. R. OKAMOTO, 1983, in: Phase Transformations during Irradiation, ed. F. V. Nolfi, Jr. (Appl. Science Publ., London) p. 247.

REHN, L. E. and R. C. BIRTCHER, 1993, J. Nucl. Mater. **205**, 31.

REHN, L. E., and K. H. ROBROCK, 1977, J. Phys. **F 7**, 1107.

REHN, L. E., K. H. ROBROCK and H. JACQUES, 1978, J. Phys. **F 8**, 1835.

RINNEBERG, H., and H. HAAS, 1978, Hyperfine Interactions **4**, 678.

RINNEBERG, H., W. SEMMLER and G. ANTESBERGER, 1978, Phys. Lett. **A 66**, 57.

RIZK, R., P. VAJDA, F. MAURY, A. LUCASSON and P. LUCASSON, 1977, J. Appl. Phys. **48**, 481.

RIZK, R., P. VAJDA, F. MAURY, A. LUCASSON, P. LUCASSON, C. DIMITROV and O. DIMITROV, 1976, J. Appl. Phys. **47**, 4740.

ROBERTS, C. G., W. P. RICKEY and P. E. SHEARIN, 1966, J. Appl. Phys. **37**, 4517.

ROBROCK, K. H., 1982, Study of Self-Interstitial-Atom-Solute-Atom-Complexes by Mechanical Relaxation, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 353.

ROBROCK, K. H., 1983, The Interaction of Self-Interstitial Atoms with Solute Atoms in Metals, in: Phase Stability and Solute Redistribution, eds. F. Nolfi and J. Gittus (Appl. Sci. Publ., London) p. 115.

ROBROCK, K. H., and H. R. SCHOBER, 1981, J. Physique, Colloque **C5-42**, Suppl. 10, 735.

ROBROCK, K. H., and W. SCHILLING, 1976, J. Phys. **F 6**, 303.

ROTH, G., H. WOLLENBERGER, C. ZECKAU and K. LÜCKE, 1975, Radiation Effects **26**, 141.

RUSSELL, K. C., 1993, J. Nucl. Mater. **206**, 129.

SAHU, R. P., K. C. JAIN and R. W. SIEGEL, 1978, J. Nucl. Mater. **69/70**, 264.

SCHAEFER, H.-E., 1987, Phys. Stat. Sol. (a) **102**, 47.

SCHAEFER, H.-E., D. BUTTEWEG and W. DANDER, 1975, Defects in High Purity Iron after 27 K Electron Irradiation, in: Proc. Int. Conf: on Fundamerltal Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA, 22161) p. 463.

SCHAEFER, H. E., 1982, in: Positron Annihilation, Proc. 6th Int. Conf. on Positron Ann., eds. P. G. Coleman, S. C. Sharma, L. M. Diana (North-Holland Publ. Co., Amsterdam) p. 369.

SCHAEFER, H. E., R. GUGELMEIER, R. SCHMOLZ and A. SEEGER, 1987a, Mater. Sci. Forum **15-18**, 111.

SCHAEFER, H. E., W. STUCK, F. BANHART and W. BAUER, 1987b, Mater. Sci. Forum **15-18**, 117.

SCHILLING, W., 1978, J. Nucl. Mater. **69/70**, 465.

SCHILLING, W., and P. TISCHER, 1967, Z. Angew. Phys. **22**, 56.

SCHILLING, W., G. BURGER, K. ISEBECK and H. WENZL, 1970, Annealing Stages in the Electrical Resistivity of Irradiated fcc Metals, in: Vacancies and Interstitials in Metals, eds. A. Seeger, D. Schumacher, W. Schilling and J. Diehl (North-Holland, Amsterdam) p. 255.

SCHINDLER, R., W. FRANK, M. RÜHLE, A. SEEGER and M. WILKENS, 1978, J. Nucl. Mater. **69/70**, 331.

SCHOBER, H. R., V. K. TEWARY and P. H. DEDERICHS, 1975, Z. Phys. **B 21**, 255.

SCHOBER, H. R., 1989, in: Phonon '89, eds. S. Hunklinger, W. Ludwig and G. Weiss, Vol. 1 (World Scientific, Singapore) p. 444.

SCHOLZ, A., and C. LEHMANN, 1972, Phys. Rev. **B 6**, 813.

SCHOTTKY, G., A. SEEGER and G. SCHMID, 1964, Phys. Stat. Sol. **4**, 439.

SCHROEDER, H., and B. STRITZKER, 1977, Radiation Effects **33**, 125.

SCHROEDER, H., and W. SCHILLING, 1976, Radiation Effects **30**, 243.

SCHROEDER, H., R. LENNARTZ and U. DEDEK, 1975, Recovery of Pure Lead after Electron Irradiation at 4.7 K and below 3 K, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA, 22161) p. 411.

SCHROEDER, K., 1980, Theory of Diffusion-Controlled Reactions of Point Defects in Metals, in: Springer Tracts in Modern Physics, vol. 87, eds. G. Höhler and E. A. Niekisch (Springer, Berlin) p. 171.
SCHULTZ, H., 1987, Mater. Sci. Forum **15-18**, 727.
SCHULTZ, H., 1991, Mater. Sci. Eng. **A 14**, 149.
SCHULTZ, H., 1991a, in: Atomic Defects in Metals, ed. H. Ullmaier, Landoldt-Börnstein, ed. O. Madelung, Group III, Vol. 25 (Springer Berlin) p. 115.
SCHUMACHER, D., 1970, in: Vacancies and Interstitials in Metals, eds. J. Diehl, W. Schilling, D. Schumacher, A. Seeger (North-Holland Publ. Co., Amsterdam) p. 889.
SCHWIRTLICH, I. A., J. PETZOLD and H. SCHULTZ, 1987, Phys. Stat. Sol. (a) **104**, 297.
SEEBOECK, R., W. ENGEL, S. HOTH, R. KEITEL and W. WITTHUHN, 1982, Vacancies in Zinc and Cadmium Produced by Proton- and Electron-Irradiation, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 271.
SEEGER, A. and E. MANN, 1960, J. Phys. Chem. Solids **12**, 326.
SEEGER, A., 1958, On the Theory of Radiation Damage and Radiation Hardening, in: Proc. 2nd U.N. Int. Conf. on the Peaceful Uses of Atomic Energy, vol. 6 (United Nations, Geneva) p. 250.
SEEGER, A., 1970, Radiation Effects **2**, 165.
SEEGER, A., 1973a, Cryst. Lattice Defects **4**, 221.
SEEGER, A., 1973b, J. Phys. **F 3**, 248.
SEEGER, A., 1975, The Interpretation of Radiation Damage in Metals, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 493.
SEEGER, A., 1982, Phys. Lett. **89A**, 241.
SEEGER, A., and H. KRONMÜLLER, 1987, Mater. Sci. Forum **15–18**, 65.
SEEGER, A., and W. FRANK, 1983, Radiation Effects **71**, 157.
SEEGER, A., E. MANN and R. V. JAN, 1962, J. Phys. Chem. Solids **23**, 639.
SEGURA, E., and P. EHRHART, 1979, Radiation Effects **42**, 233.
SEIDMAN, D.N., 1973, J. Phys. F: Met. Phys. **3**, 393.
SEIDMAN, D.N., 1978, Surf. Sci. **70**, 532.
SEITZ, F., and J.S. KOEHLER, 1956, Displacement of Atoms during Irradiation, in: Solid State Physics. vol. 2, eds. F. Seitz and D. Turnbull (Academic, New York) p. 305.
SHAH, N., and A.L. CATZ, 1984, Phys. Rev. **B 30**, 2498.
SHIMOMURA, Y., K. YAMAHAWA, K. KITAGAWA and H. ODA, 1982, Studies of Point Defect Clusters with an Electron Microscope, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 712.
SHIMOMURA, Y., M.W. GUINAN and T. DIAZ DE LA RUBIA, 1993, J. Nucl. Mater. **205**, 374.
SHIRAI, Y., H.-E. SCHAEFER and A. SEEGER, 1989, in: Positron Annihilation, Proc. 8th Int. Conf. on Positron Annihilation, Gent, Belgium, 1988, eds. L. Dorikens-Vanpraet, M. Dorikens, D. Segers (World Scientific Publ. Co., Singapore) p.... (?)
SHIRLEY, C.G., and R.L. CHAPLIN, 1972, Phys. Rev. **B 5**, 2027.
SIEGEL, R.W., 1978, J. Nucl. Mater. **69/70**, 117.
SIEGEL, R.W., 1982a, in: Positron Annihilation, Proc. 6th Int. Conf. on Positron Ann., eds. P.G. Coleman, S. C. Sharma, L. M. Diana (North-Holland Publ. Co., Amsterdam) p. 351.
SIEGEL, R.W., 1982b, in: Point Defects and Interactions in Metals, Proc. of Yamada Conf. V, eds. J. Takamura, M. Doyama, M. Kiritani, University of Tokyo Press, Tokyo, p. 533.
SIEMS, R., 1968, Phys. Stat. Sol. **30**, 645.
SIGMUND, P., 1972, Rev. Roum. Phys. **17**, 969.
SIMMONS, R.O., and R.W. BALLUFFI, 1960a, Phys. Rev. **117**, 52.
SIMMONS, R.O., and R.W. BALLUFFI, 1960b, Phys. Rev. **119**, 600.
SIMMONS, R.O., and R.W. BALLUFFI, 1962, Phys. Rev. **125**, 862.
SIMMONS, R.O., and R.W. BALLUFFI, 1963, Phys. Rev. **129**, 1533.
SINGH, B.N. and S.J. ZINKLE, 1993, J. Nucl. Mater. **206**, 212.
SMIGELSKAS, A.D. and E.O. KIRKENDALL, 1947, Trans. AIME **171**, 130.

SNIEGOWSKI, J. J. and W. G. WOLFER, 1984, Proc. Topical Conf. on Ferritic Alloys for Use in Nuclear Energy Technology, AIME, 1984, 579.
SONNENBERG, K., and U. DEDEK, 1982, Radiation Effects **61**, 175.
SONNENBERG, K., W. SCHILLING, H. J. DIBBERT, K. MIKA and K. SCHROEDER, 1972b, Radiation Effects **15**, 129.
SONNENBERG, K., W. SCHILLING, K. MIKA and K. DETTMANN, 1972a, Radiation Effects **16**, 65.
SOSIN, A., and W. BAUER, 1968, Atomic Displacement Mechanism in Metals and Semi-Conductors, in: Studies in Radiation Effects in Solids, ed. G. J. Dienes, vol. 3 (Gordon and Breach, New York).
SPIRIC, V., L. E. REHN, K. H. ROBROCK and W. SCHILLING, 1977, Phys. Rev. **B 15**, 672.
STOLLER, R. E., and G. R. ODETTE, 1985, J. Nucl. Mater. **131**, 118.
STOTT, M. J., 1978, J. Nucl. Mater. **69/70**, 157.
SULPICE, G., C. MINIER, P. MOSER and H. BILGER, 1968, J. Physique **29**, 253.
TAKAKI, S., J. FUSS, H. KUGLER, U. DEDEK and H. SCHULTZ, 1983, Radiation Effects **79**, 87.
TENENBAUM, A., and N. V. DOAN, 1977, Phil. Mag. **35**, 379.
TEWORDT, L., 1958, Phys. Rev. **109**, 61.
THEIS, U. and H. WOLLENBERGER, 1980, J. Nucl. Mater. **88**, 121.
TRIFTSHÄUSER, W., 1975, Festkörperprobleme XV/Adv. in Solid State Phys., ed. H. J. Queisser (Vieweg, Braunschweig) p. 381.
TRIFTSHÄUSER, W., and J. D. MCGERVEY, 1975, Appl. Phys. **6**, 177.
TRINKAUS, H., 1975, Theory of Polarization-Induced Elastic Interaction of Point Defects, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006. Natl. Techn. Inf. Service, Springfield, VA 22161) p. 254.
TRINKAUS, H., 1983, Radiation Effects **78**, 189.
TRINKAUS, H., 1990, J. Nucl. Mater. **174**, 178.
TRINKAUS, H., B. N. SINGH and A. J. E. FOREMAN, 1992, J. Nucl. Mater. **199**, 1.
TRINKAUS, H., B. N. SINGH and A. J. E. FOREMAN, 1993, J. Nucl. Mater. **206**, 200.
TRINKAUS, H., V. NAUNDORF, B. N. SINGH and W. H. WOO, 1994, J. Nucl. Mater. **210**, 244.
TZANETAKIS, P., J. HILLAIRET and G. REVEL, 1976, Phys. Stat. Sol. (b) **75**, 433.
URBAN K., and N. YOSHIDA, 1981, Phil. Mag. **A 44**, 1193.
URBAN, K., B. SAILE, N. YOSHIDA and W. ZAG, 1982, in: Point Defects and Defect Interaction in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 783.
VAJDA, P., 1977, Rev. Mod. Phys. **49**, 481.
VAN BUEREN, H. G., 1955, Z. Metallk. **46**, 272.
VAN DEN BEUKEL, A., 1970, in: Vacancies and Interstitials in Metals, eds. J. Diehl, W. Schilling, D. Schumacher, A. Seeger (North-Holland Publ. Co., Amsterdam) p. 427.
VANDENBORRE, H., L. STALS and J. NIHOUL, 1968, Int. Conf. Vacancies and Interstitials in Metals (Jülich-Conf. 2) p. 802.
VANDENBORRE, H., L. STALS, K. SONNENBERG and J. NIHOUL, 1972, in: Defects in Refractory Metals, eds. R. De Batist, J. Nihoul, L. Stals (Centre d'Etudes de l'Energie Nucléaire, Mol, Belgien) p. 49.
VAROTSOS, P., and K. ALEXOPOULOS, 1982, Phys. Stat. Sol. (b) **110**, 9.
VOGL, G., and W. MANSEL, 1975, Mössbauer Studies of Interstitials in fcc Metals, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 349.
VOGL, G., W. MANSEL and P. H. DEDERICHS, 1976, Phys. Rev. Lett. **36**, 1497.
WAGNER, C., and J. BEYER, 1936, Z. Phys. Chem. **B 32**, 113.
WAGNER, R., 1982, Field Ion Microscopy, in: Crystals, Growth, Properties and Applications, ed. H. C. Freyhardt (Springer, Berlin).
WAITE, T. R., 1957, Phys. Rev. **107**, 463 and 471.
WALLNER, G., 1983, Dissertation, Techn. Universität München.
WALLNER, G., K. BÖNING and U. DEDEK, 1986, J. Phys. **F 16**, 257.
WEILER, W., and H. E. SCHAEFER, 1985, J. Phys. **F 15**, 1651.
WENZL, H., 1970, Physical Properties of Point Defects in Cubic Metals, in: Vacancies and Interstitials in Metals, eds. A. Seeger, D. Schumacher, W. Schilling and J. Diehl (North-Holland, Amsterdam) p. 363.

WENZL, H., F. KERSCHER, V. FISCHER, K. EHRENSPERGER and K. PAPATHANASSOPOULOS, 1971, Z. Naturf. **26**, 489.
WEST, R. N., 1973, Adv. Phys. **22**, 263.
WEST, R. N., 1979, Positron Studies of Lattice Defects in Metals, in: Topics in Current Physics 12, ed. P. Hautojärvi (Springer, Berlin) ch. 3.
WEVER, H., 1992, in: Diffusion in Solids - Unsolved Problems, ed. G. E. Murch (Trans. Tech. Publ., Aedermannsdorf, Switzerland) p. 55.
WICHERT, TH., 1982, PAC Study of Point Defects in Metals, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 19.
WICHERT, TH., M. DEICHER, O. ECHT and E. RECKNAGEL, 1978, Phys. Rev. Lett. **41**, 1659.
WIEDERSICH, H., and N. Q. LAM, 1983, in: Phase Transformations during Irradiation, ed. F. V. Nolfi, Jr. (Appl. Science Publ., London) p. 1.
WIEDERSICH, H., P. R. OKAMOTO and N. Q. LAM, 1979, J. Nucl. Mater. **83**, 98.
WILLIS, J. R., 1975, J. Mech. Phys. Sol. **23**, 129.
WOLLENBERGER, H., 1965, Fruhjahrstagung DPG Freudenstadt, unpublished .
WOLLENBERGER, H., 1970, Production of Frenkel Defects during Low-Temperature Irradiations, in: Vacancies and Interstitials in Metals, eds. A. Seeger, D. Schumacher, W. Schilling and J. Diehl (North-Holland, Amsterdam) p. 215.
WOLLENBERGER, H., 1975, Defect interactions above stage I, in: Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, 1975, eds. M. T. Robinson and F. W. Young, Jr. (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA 22161) p. 582.
WOLLENBERGER, H., 1978, J. Nucl. Mater. **69/70**, 362.
WOLLENBERGER, H., 1983, in: Physical Metallurgy, Vol. 2, eds. R. W. Cahn, P. Haasen (Elsevier Science Publishers, Amsterdam) p. 1140.
WOLLENBERGER, H., and B. KELL, 1989, J. Nucl. Mater. **169**, 126.
WOLLENBERGER, H., and J. WURM, 1965, Phys. Stat. Sol. **9**, 601.
WOO, C. H., A. A. SEMENOV and B. N. SINGH, 1993, J. Nucl. Mater. **206**, 170.
WOOD, R. F., and M. MOSTOLLER, 1975, Phys. Rev. Lett. **35**, 45.
YAMAKAWA, K., I. MUKOUDA and Y. SHIMOMURA, 1992, J. Nucl. Mater. **191-194**, 396.
YAMAMOTO, R., 1982, Lattice Vibrations around a Vacancy and Vacancy Clusters in Metals, in: Point Defects and Defect Interactions in Metals, eds. J. Takamura, M. Doyama and M. Kiritani (Univ. of Tokyo Press, Tokyo) p. 120.
YOUNG, F. W., Jr., 1978, J. Nucl. Mater. **69/70**, 310.
ZAISER, M., W. FRANK and A. SEEGER, 1992, Solid State Phenomena **23 & 24**, 221.
ZELLER, R., and P. H. DEDERICHS, 1976, Z. Phys. **B 25**, 139.

Further reading

Bleiberg, M. L., and J. W. Bennett, eds., 1977, Radiation Effects in Breeder Reactor Structural Materials, Proc. Int. Conf. Scottsdale 1977 (AIME, New York).
Carpenter, G. J. C., C. E. Coleman and S. R. MacEwen, eds., 1980, Fundamental Mechanisms of Radiation-Induced Creep and Growth, Proc. Int. Conf. Chalk River 1979, J. Nucl. Mater. 90.
Corbett, J. W., 1966, Electron Radiation Damage in Semiconductors and Metals, in: Solid State Physics Suppl. 7, eds. F. Seitz and D. Turnbull (Academic, New York).
Corbett, J. W., and L. C. Ianniello, eds., 1972, Radiation-Induced Voids in Metals, Proc. Int. Conf. Albany 1971 (CONF-710601, Natl. Techn. Inf. Service, Springfield, VA 22161).
Flynn, C. P., 1972, Point Defects and Diffusion (Clarendon Press, Oxford).
Gavillet, D., M. Victoria, B. N. Singh and A. Horsewell, eds., 1993, Proc. Workshop on Time Dependence of Radiation Damage Accumulation and its Impact on Materials Properties, J. Nucl. Mater. 206.
Gehlen, P. C., J. R. Beeler, Jr., and R. I. Jaffee, eds., 1971, Interatomic Potentials and Simulation of Lattice Defects, Battelle Inst. Materials Science Colloquia, Seattle and Harrison, Hotsprings (Plenum, New York).
Gittus, J., 1978, Irradiation Effects in Crystalline Solids (Appl. Science Publ., London).

Green, W. V., M. Victoria, T. Leffers and B. N. Singh, eds., 1990, Radiation Effects and Defects in Solids 113, 1.

Hautojärvi, P., ed., 1979, Positrons in Solids, Topics in Current Physics, vol. 12 (Springer, Berlin).

Höhler, G., and E. A. Niekisch, eds., 1978, Point Defects in Metals I, Springer Tracts in Modern Physics, vol. 81 (Springer, Berlin).

Höhler, G., and E. A. Niekisch, eds., 1980, Point Defects in Metals II, Springer Tracts in Modern Physics, vol. 87 (Springer, Berlin).

Holland, J. R., L. K. Mansur and D. I. Potter, eds., 1981, Proc. Symp. On Phase Stability during Irradiation (Met. Soc. AIME, Warrendale, PA).

Ishino, S., M. Kiritani, T. Kondo and J. L. Scott, eds., 1991, Fusion Reactor Materials, J. Nucl. Mater. 179-181.

Klueh, R. L., R. E. Stoller and D. S. Gelles, eds., 1992, Fusion Reactor Materials, J. Nucl. Mater. 191-194.

Krippner, M., ed., 1969, Radiation Damage in Reactor Materials, Proc. Symp. Vienna 1969 (IAEA, Vienna).

Nelson, R.S, 1968, The Observation of Atomic Collisions in Crystalline Solids, in: Defects in Crystalline Solids, vol. 1, eds. S. Amelinckx, R. Gevers and J. Nihoul (North-Holland, Amsterdam).

Nygren, R. E., R. E. Gold and R. H. Jones, eds., 1981, Fusion Reactor Materials, Proc. 2nd Topical Meeting on Fusion reactor Materials, Seattle, J. Nucl. Mater. 103/104.

Peterson, N. L., and R. W. Siegel, eds., 1976, Properties of Atomic Defects in Metals, Proc. Int. Conf. Argonne, J. Nucl. Mater. 69/70.

Proc. Workshop "Time Dependence of Radiation Damage Accumulation and its Impact on Materials Properties", 1993, J. Nucl. Mater. 206.

Robinson, M. T., and F. W. Young, Jr., eds., 1975, Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals, Gatlinburg 1975 (CONF-751006, Natl. Techn. Inf. Service, Springfield, VA 22161).

Schilling, W. and H. Ullmaier, 1994, Physics of Radiation Damage in Metals, in: Nuclear Materials, Part II, 179, ed. B. R. T. Frost, (Volume 10B of Materials Science and Technology, ed. R. W. Cahn, P. Haasen and E. J. Kramer, VCH, Weinheim)

Seeger, A., D. Schumacher, W. Schilling and J. Diehl, eds., 1970, Vacancies and Interstitials in Metals (North-Holland, Amsterdam).

Singh, B. N., A. Horsewell, M. Victoria and W. V. Green, 1990, eds., J. Nucl. Mater. 174, 125.

Stiegler, J. O., ed., 1976, Correlation of Neutron and Charged Particle Damage, Proc. Workshop Oak Ridge 1976 (CONF-760673, Natl. Techn. Inf. Service, Springfield, VA 22161).

Stiegler, J. O., ed., 1979, Workshop on Solute Segregation and Phase Stability During Irradiation, Gatlinburg 1978, J. Nucl. Mater. 83.

Wiffen, F. W., J. H. DeVan and J. O. Stiegler, eds., 1979, Fusion Reactor Materials, Proc. 1st Topical Meeting Miami, J. Nucl. Mater. 85/86.

Woo, C. H., and R. J. McElroy, eds., 1988, J. Nucl. Mater. 159, 1.

CHAPTER 19

METASTABLE STATES OF ALLOYS

R. W. CAHN and A. L. GREER

Department of Materials Science and Metallurgy
University of Cambridge
Cambridge CB2 3QZ, England

R. W. Cahn and P. Haasen†, eds.
Physical Metallurgy; fourth, revised and enhanced edition

1. Introduction

1.1. General features

Much of the development of practical metallic materials has involved dealing with metastable states. Just what is meant by "metastable" we will consider more carefully in § 1.3; for the moment we use it to describe any state with a Gibbs free energy higher than the lowest value corresponding to stable equilibrium. Conventionally the access to metastable states is achieved by solid-state quenching. Examples are the quenching of steels from the austenite phase field to form martensite (ch. 17, § 3.2) and the quenching of solution-treated Al–Cu alloys to obtain a supersaturated solid solution suitable for age-hardening (§ 5.1, below). These examples illustrate the point of obtaining metastable states: the relaxation of the constraint of being in equilibrium gives access to a wider range of structures, some of which can be expected to show superior properties. They also illustrate that the metastable state may be of interest as the final product or as an intermediate on the way to the final product. At least since the 1920s there has been interest in developing new materials by increasing the departure from equilibrium, mainly by rapid quenching of the liquid. This grew into a large research effort following the work of Duwez and colleagues in 1959–60 (reviewed by DUWEZ [1967]), in which gun splat-quenching was used to demonstrate complete, metastable, solid solubility in Ag–Cu and metallic glass formation in Au–Si and Au–Ge. Rapid liquid quenching has since been used in a very wide range of fundamental studies as well as being adopted as a significant commercial production process, particularly for magnetic materials. In the early 1980s interest turned also to solid-state processing as a way to obtain alloys in states far from equilibrium. This built on a great deal of earlier work, including the development in the 1960s of mechanical alloying to produce high-temperature alloys (BENJAMIN [1970]).

The microstructural manifestations of the departures from equilibrium achieved by novel processing routes can be classified under five headings:

(i) *Increased defect concentrations.* These include increased concentrations of vacancies (and to a lesser extent interstitials), dislocations, stacking faults, twin boundaries and grain boundaries. Another characteristic to be put in this category is an increased level of chemical disorder (that is, reduced order parameter) in ordered solid solutions and compounds.

(ii) *Microstructural refinement.* This involves finer scale distributions of different phases and of solute. Examples include finer dendrite arm spacings, eutectic spacings and precipitate diameters.

(iii) *Extended solid solubility.* A stable crystalline phase may be found with solute levels beyond the solubility limit at the ambient temperature, or indeed beyond the equilibrium limit at any temperature (see account of solute trapping in ch. 8, §§ 5.2, 8.1.7).

(iv) *Metastable phases.* A metastable phase may form during processing. It may be one that is not found *in equilibrium* under any conditions in the system of interest, or it may be so found but appear metastably under different conditions of composition, temperature or pressure. This heading includes crystalline and quasicrystalline intermetallic compounds.

(v) *Metallic glasses.* These are a special case of a metastable phase. When the processing is by cooling a liquid they form not by a phase transformation, but by a continuous congealing of the liquid.

Some of the kinetic aspects of rapid solidification have been dealt with in ch. 8. The present chapter is concerned more generally with the wide variety of methods for producing alloys in states which are far from equilibrium, that is, with novel methods which lead to larger departures from equilibrium than can be achieved by conventional methods. It is also concerned with the metastable products themselves, and in particular with the metallic glassy state, which is the most distinctive product of the novel methods. Few surveys have sought to cover metastable states from such a broad perspective; another recent example is a book chapter (MARTIN, CANTOR and DOHERTY [1996]).

1.2. Methods for achieving metastability

Here, an outline is given of some of the very diverse methods which have been applied, and links between them are discussed. Details of the methods are presented in § 3. Most of the effort on developing new production methods has focused on cooling rate. This depends on the dimensions of the sample, the heat conduction in the material, and the rate of heat transfer to the surrounding quenching medium. If the sample to be quenched is initially solid, its dimensions are fixed and often not ideally suited to rapid heat extraction. In addition, efficient heat transfer may be difficult to achieve. On the other hand if the sample is initially liquid, it can be flattened into a thin sheet in intimate contact with a good solid heat sink. This is the basis of the Duwez method, subsequently developed into diverse continuous production techniques described in § 3.1. In addition to solid heat sinks, the heat from the liquid alloy can be extracted into another liquid or a gas. The most common rapid liquid quenching methods give products in droplet, wire, ribbon or sheet form, with at least one dimension as small as 10 to 100 μm. The cooling rate is commonly $10^5 - 10^6$ K/s. Rapid liquid quenching is also readily achieved when the surface of a bulk sample is melted; the bulk itself then acts as the heat sink (§ 3.5). The melting can be by a scanned continuous heating beam or by a stationary pulsed beam. For laser beams, pulse durations can be ultra short — in the nanosecond or even picosecond range — giving very thin melted layers and quenching rates as high as 10^{14} K/s; this is the highest liquid quenching rate which can be achieved (LIN and SPAEPEN [1984]).

The objective of rapid heat extraction is to obtain rapid movement of the solid–liquid interface. The latent heat of freezing released at the interface is extracted through the growing solid into the quenching medium. However, the latent heat can alternatively be extracted into the liquid ahead of the interface if that liquid is sufficiently undercooled. The methods for obtaining highly undercooled liquids are described in § 3.2; they have attracted much interest because of their suitability for fundamental studies of the mechanisms of metastable phase formation.

A number of other techniques are associated with the concept of rapid quenching, though less obviously than with the liquid case. When a solid is grown from atoms in the vapor phase (generated by evaporation or sputtering), the effective quench rate in the

process of condensation is as high as 10^{12} K/s (§ 3.3). Other methods in which the mobility of the atoms joining the solid is rapidly reduced are those based on chemical or electrochemical precipitation or deposition from solution (§ 3.6). Irradiation of a solid leads to locally disturbed configurations which regain thermal equilibrium very quickly; again the effective local quench rate is of the order of 10^{12} K/s. Amorphization by irradiation is considered in § 3.4.

Chemical reactions can proceed towards equilibrium in stages, and the intermediate stages can yield a metastable phase. Just this is found in the solid-state amorphization reaction (§ 3.7) in which an amorphous alloy can be produced by the reaction of two solid metallic elements.

Severe mechanical deformation can lead to metastable states (§ 3.8). The deformation forces the production of disturbed configurations, and in this way can be considered akin to techniques such as irradiation. On the other hand it can be used to bring different phases into intimate contact, promoting solid-state reactions which can yield metastable phases. Finally, high pressure can induce metastability.

1.3. The nature of metastability

A classification of thermodynamic states is illustrated in fig. 1. In stable equilibrium the free energy (the Gibbs free energy for a system under constant pressure) is at the lowest possible minimum. The minimum shown schematically in the figure may represent the stable state of a single phase or of a phase mixture. A metastable state is one in internal equilibrium, that is, within the range of configurations to which there is access by continuous change, the system has the lowest possible free energy. However, if there were a large fluctuation — the nucleation of a more stable phase — transformation to the new phase would occur. A good example of a thermodynamically metastable state is an undercooled liquid; this has well defined equilibrium properties at each

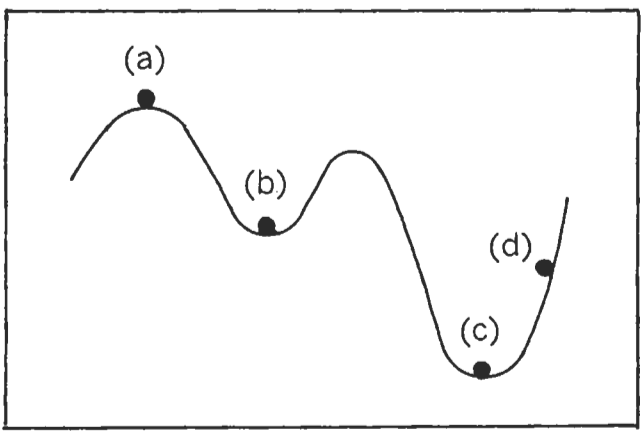


Fig. 1. A schematic illustration of possible thermodynamic states of alloys: (a) unstable equilibrium, (b) metastable equilibrium, (c) stable equilibrium, and (d) unstable non-equilibrium state. The vertical axis shows Gibbs free energy, the horizontal, a parameter representing changes in structure.

temperature, but contact with a nucleant will result in rapid transformation to the stable crystalline phase. A state of unstable equilibrium is destroyed by any fluctuation, however small; an example would be a precipitate dispersion (subject to Ostwald ripening) in which all the precipitates have exactly the same size. And there are unstable states, not in equilibrium at all.

The term "metastable" was invented by the German physical chemist, Wilhelm Ostwald, in 1893. In his textbook (OSTWALD [1893]) he defined the newly named concept as follows: "Of a physical system, persisting (in its existing state) when undisturbed or subject to disturbances smaller than some small or infinitesimal amount, but passing to a more stable state when subject to greater disturbances."

As elucidated by TURNBULL [1981], all solids of practical interest are configurationally frozen, that is, under the conditions of observation, the atomic configuration does not undergo significant changes. In this way microstructures which are thermodynamically unstable are readily found, and are indeed much more common than microstructures in true equilibrium, whether stable, metastable or unstable. Examples are a supersaturated solid solution in the spinodal range, a dispersion of precipitates of varied size, and a glass; all of these can relax toward lower free energy states continuously, without any nucleation barrier. That these states persist at all is of course because of kinetic barriers to atomic motion. As suggested by Turnbull, the best practice may be to describe a microstructure in terms of its thermodynamic state before configurational freezing set in. In this way, a glass would be considered metastable and a supersaturated solid solution in the spinodal range would be considered unstable. A common practice, however, is to term all such configurationally frozen states "metastable". Although such a usage is approximate at best, we shall adopt it in this chapter. By considering the local potential wells between which atoms must make diffusional jumps, even states which are *thermodynamically* unstable may be thought of as *kinetically* metastable. Such kinetic metastability can exist only where thermal activation of atomic jumps is required. Transformations which do not require thermal activation cannot be suppressed; martensitic transformations come in this category, and Turnbull suggests that crystallization of amorphous pure metals does also.

Taking the definition of a metastable state to be any state which has a free energy in excess of that of stable equilibrium, it is then natural to quantify the degree of metastability by the value of the excess. The novel production techniques and microstructures associated with large excess free energies are the focus of this chapter. By novel methods such as those outlined in § 1.2, excess free energies as high as $R\bar{T}_m$ (where $\bar{T}_m$ represents the average melting point of the elements in the system) can be retained in configurationally frozen states. Such values are up to twice the highest achieved by conventional processing. In the various microstructural effects of novel processing methods (§ 1.1), the excess free energy can be seen to arise from three basic kinds of metastability (TURNBULL [1981]): morphological, compositional and structural. In *morphological metastability*, the excess free energy comes from defects of all kinds including interphase boundaries; it is associated with (i) increased defect concentrations and (ii) microstructural refinement. *Compositional metastability* is found in (iii) extended solid solubility. *Structural metastability* is associated with (iv) metastable crystalline

References: p. 1818.

phases and (v) metallic glasses. Typical excess free energies associated with these types of metastability are:

morphological $\leq 0.1\ R\overline{T}_m$
structural $\leq 0.5\ R\overline{T}_m$
compositional $\leq 1.0\ R\overline{T}_m$

As described by Turnbull, such excess values can be considered to be attained in general by a process of energizing the system followed by quenching to effect configurational freezing. Considering the processing routes outlined in § 1.2, examples of energization can be seen to include melting, generation of gaseous species by evaporation or sputtering, dissolution, irradiation or mechanical deformation. In the solid-state amorphization reaction, the energization is the creation (by deposition or mechanical working) of interphase interfaces. The quenching, typically characterized by a cooling rate, is needed to reach a configurationally frozen state while retaining a large excess energy. It is useful to consider what constitutes configurational freezing. This does not occur at a single well defined temperature, but reflects the range of types of atomic motion. Roughly, however, it can be considered that diffusive motion of atoms within bulk phases (that is, lattice diffusion in a crystal, or diffusion in a glass) is frozen out below a reduced temperature (T/T_m) of ≈ 0.4. For grain-boundary and free-surface diffusion, the corresponding reduced temperatures are ≈ 0.1 and ≈ 0.2. Though a highly metastable material may have a lower temperature for effective configurational freezing because of for example a larger grain-boundary area, the fundamental temperatures for freezing the various atomic processes stay the same as in materials near equilibrium. Thus the retention of large excess free energies in practical materials is not necessarily thwarted under normal service conditions by any increased susceptibility to structural change.

Configurationally frozen normal materials, even when processed by novel methods, could be in thermodynamically stable states, and such states by definition have the largest driving forces for their formation. It is reasonable to ask, therefore, why metastable materials are so commonly obtained. They must be favored kinetically. The nucleation of one phase in another is easier with lower interfacial energy between the phases, and that energy tends to scale with the entropy difference between the phases. Growth of one phase in another is favored if substantial solute partitioning is not required. Thus a transformation is in general more likely if the new phase is more similar to the old, either in entropy or composition. This has been described as a "step rule" (TURNBULL [1981]), and it will often lead to metastable products.

2. *Formation of metallic glasses (amorphous alloys)*

2.1. Formation and thermodynamics

The metastable alloy structure most associated with novel processing methods and with large departures from equilibrium is the metallic glass. Strictly, a *glass* is formed when a liquid which is cooled fails to crystallize and instead continuously and uniformly congeals into a solid. The glass has essentially the structure of the liquid, combined with

the resistance to shear deformation of a solid. Non-crystalline, *amorphous*, alloys can be obtained by any of the methods outlined in § 1.2, and not just those involving quenching of a liquid. In this section we show that amorphous alloys made by liquid quenching can be considered as true glasses and that those prepared by other methods are not significantly different. The thermodynamics of the glassy/amorphous state are considered in relation to the formation of amorphous alloys.

We start by pointing out that amorphous and glassy states are not constrained to be the same, and illustrate this with the case of silicon. Liquid silicon has a high density and is metallic. Crystalline silicon has tetrahedral covalent bonding. Amorphous silicon, which can be made by vapor deposition as well as by solidification from the liquid, also has local tetrahedral coordination with covalent bonding (POATE *et al.* [1987]). In its bonding type, solid amorphous silicon is therefore quite distinct from the liquid and cannot be regarded as a glass, which if formed would be dense and metallic. For example, this has the result that crystallization of amorphous silicon is a quite different transformation from crystallization of liquid silicon, as illustrated by the different temperature dependences of the two processes (STIFFLER, EVANS and GREER [1992]).

Figure 2 shows schematically the changes in properties which occur when a liquid is cooled into the crystalline or glassy states. At the higher temperatures, range A, the melt is in equilibrium. It is extremely difficult (though possible under certain circumstances, PEREPEZKO and GLEITER [1983]) to superheat a crystal above the thermodynamic melting point, so that metastable states are very rare in range A. In defining the boundary between ranges A and B, we take T_f (the equilibrium freezing temperature); this is not necessarily the same as T_m, the melting temperature, for alloys with a solidus–liquidus gap, but we ignore the complexities arising from an interposed solid + liquid range. In effect, the diagram is for a system in which all the transformations are without solute partitioning.

In range B, the melt is thermodynamically undercooled and will crystallize rapidly if a critical nucleus is provided. At least in the higher temperatures in this range the liquid has a low viscosity, and crystallization can be avoided only because fast cooling gives no time for nucleus formation. At the lowest temperatures in this range the liquid viscosity increases rapidly with falling temperature, and a glass can still be formed even in the presence of nuclei because the low atomic mobility stifles crystal growth. Devitrification studies (§ 4.3) provide evidence for copious nucleation occurring in the later stages of a quench without significant overall transformation of the liquid. In range B, the viscosity varies over some fifteen orders of magnitude, and though most of the range is not accessible to direct measurement (because crystallization would intervene during the measurement), the curve can be estimated by interpolation between experimentally accessible points in the neighborhood of P and Q. At all temperatures in range B the atomic mobility is adequate for the liquid to remain in configurational equilibrium (that is, in an internal equilibrium in which it has the lowest free energy of any state which it can reach by continuous change).

Crystallization, if it occurs, involves a discontinuous change in properties (as shown, for example, by volume in fig. 2). Crystallization can occur only in range B, and if it is avoided on cooling, range C is reached. Range C is defined by the *glass transition*

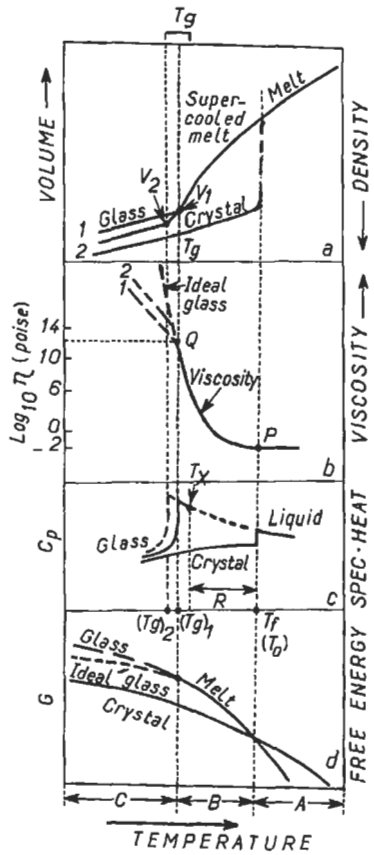


Fig. 2. The glass transition: variation of several properties with temperature. See text.

temperature T_g, below which the liquid configuration is congealed in a pattern which corresponds to equilibrium at T_g. Experimentally, T_g is close to the temperature at which the viscosity reaches 10^{12} Pa.s. The nature of the transition at T_g can be seen by considering the temperature dependence of liquid/glass properties, for example volume and viscosity as shown in fig. 2. Above T_g the properties show a strong temperature dependence, reflecting the configurational changes in the liquid. Below T_g, however, atomic motions in the melt are so slow that during a quench (as presupposed in drawing fig. 2) there is no longer time for configurational changes; the mobile melt has become a rigid glass. Corresponding to the lack of configurational change in the glass, its properties have a rather weak temperature dependence, roughly matching the behavior of the crystal.

It is important to recognize that T_g is *not* a thermodynamically defined temperature; its location is determined wholly by kinetic considerations. If the quenching rate is reduced (but not so much as to permit crystallization in range B), then the liquid

configurations can stay in equilibrium to a lower temperature, that is $(T_g)_1$ is lowered to, say $(T_g)_2$. This slower quench entails a smaller volume (higher density) and higher viscosity for the glass. Thus different glassy states can exist; § 4.2 deals with the structural relaxation by which glassy states can change. On simple annealing glasses densify and evolve towards an ideal glassy state represented by the extrapolation of the equilibrium properties, shown on fig. 2 for the viscosity. How dense could an ideal glass be? This intriguing question was first addressed by KAUZMANN [1948]. He pointed out that extrapolation of liquid properties to lower temperatures would soon (that is, at temperatures not far below experimental T_gs) lead to what he termed an *apparent paradox* in which the characteristic difference between liquid and crystal properties would be inverted. For example, the liquid would become *denser* than the crystal and it would have *lower* entropy. While not thermodynamically forbidden, structurally such an inversion seems very unlikely. Kauzmann proposed that some transformation, either an ideal (non-kinetic) glass transition, or some type of spontaneous crystallization, must intervene to prevent the paradox being reached. He suggested that the isentropic point (at which the extrapolated liquid entropy matches that of the crystal) could be taken to represent the ideal glass transition temperature and that the observed glass transition (found at finite rather than infinitely slow cooling rates) occurs somewhat before the ideal transition. There has been a steady flow of papers attempting to make sense of Kauzmann's *entropy catastrophes* (as they are termed nowadays), not only in Kauzmann's original version at low temperatures, but also other catastrophes in the neighborhood of the liquid/vapor transition, which the most recent study (KISHORE and SHOBHA [1994]) has sought to relate to van der Waals's well-known critical temperature. The Kauzmann paradox remains very much on the research agenda.

In connection with fig. 2 we have described the classical features of glass formation, well known for oxides and polymers. Metallic systems are not so well known as glass-formers, yet experiments have consistently shown that they fit into the same picture. Around T_g, metallic glass-forming systems show all the characteristics (volumetric, rheological, thermal) of the conventional glass transition illustrated in the figure. On annealing, the properties of metallic glasses change in a way characteristic of the structural relaxation of conventional glasses. Most importantly, some easy glass-forming compositions are sufficiently stable that measurements are possible in range B without crystallization. In this way the link between the glass and liquid is directly established (KUI and TURNBULL [1985]). Measurements of properties and structures (and, as we will see later in this section, thermodynamic measurements) suggest that amorphous metallic phases produced by other methods are (within the range of structure expected) identical to glasses produced by liquid quenching.

Of central importance in considering glass formation and the properties of glassy or amorphous phases is the atomic mobility. Low mobilities permit the glass to be formed and a glassy or amorphous phase to persist. Much of the discussion of atomic mobility in dense non-crystalline systems has centred on the concept of free volume, which we now examine. As seen in fig. 2, a faster quenched glass (1) has a lower density than a more slowly quenched one (2) or, in other terms, "1" has a larger *free volume* than "2". Free volume was originally introduced and analysed by polymer physicists (FOX and

FLORY [1951]); RAMACHANDRARAO *et al.* [1977a, b] give a full discussion of the concept. In the simplest terms, free volume is the unfilled space between the atoms or molecules (which at once poses the problem of defining the contours of these entities); the central concern is with the space in excess of that which would exist even in ideal close-packing. The free volume can be defined in terms of thermal expansion starting from an ideal glassy state, or in terms of the hole theory of liquids in which free volume is postulated to consist of identifiable holes with a rather narrow size distribution.

The crucial point is that, the smaller the free volume, the more sluggish is any structural rearrangement. The viscosity, η, varies with free volume approximately according to the *Doolittle equation*, originally derived for polymers: $\eta = A \exp(B/f_T)$, where f_T is the *relative free volume* defined by $(V_T - V_0)/V_0$, where V_T, V_0 are the volumes of material at temperatures T and 0 K respectively. The equation is valid throughout range B in fig. 2, but cannot be applied in the glassy state itself. The contentious question whether the free volume is more or less uniformly distributed, or else concentrated in holes which are probably at least the size of the smallest atoms in the glass, is still open, as reviewed by CAHN [1979]. Glass stability against annealing-induced crystallization (which will be considered further in § 4.3) is directly related to atomic mobility. Figure 3 shows that the crystallization temperature has a good correlation with hole formation enthalpy; a stabler glass has a higher formation energy for holes, therefore fewer holes form, the total free volume is smaller and thus crystallization more sluggish. In range B (fig. 2) the relative free volume remains in thermal equilibrium. On cooling it decreases particularly rapidly as T_g is approached, leading to a very sharp rise in viscosity. Below T_g free volume is frozen-in. In the glassy state, free volume is a very useful parameter for describing and analysing the property changes which occur on annealing; this is exemplified in the work of SPAEPEN and TAUB [1983] on viscosity and of CHASON and MIZOGUCHI [1987] on atomic diffusion (see § 4.2).

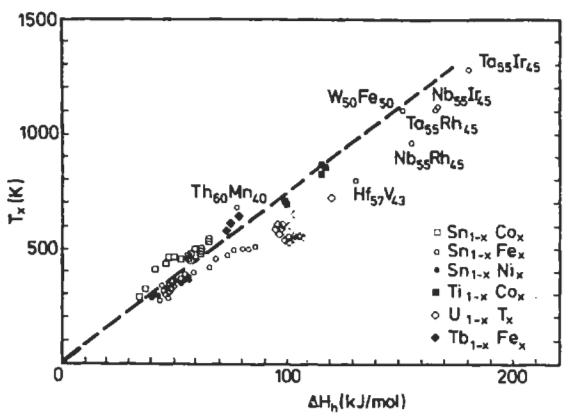


Fig. 3. Dependence of the measured crystallization temperature, T_x, of various binary metallic glasses on the formation enthalpy of a hole the size of the smaller of the constituent atoms: the hole formation enthalpy is calculated from thermochemical data. (After BUSCHOW [1982]).

Different types of glass-forming system, network formers like SiO_2, ionic liquids and those showing van der Waals bonding, do show different types of behavior, reflecting the different sensitivities of the liquid structure to temperature changes. The so-called *strong* liquids like SiO_2 have a near-Arrhenius dependence of atomic mobility (characterized by the viscosity). On the other hand an organic liquid like toluene is a *fragile* liquid with a strongly non-Arrhenius temperature dependence following the Doolittle equation. The distinction is fully discussed by ANGELL [1995]. Metallic liquids fall between these extremes, nearer to the fragile limit. It is characteristic of such systems that there is a large step in specific heat at T_g. As shown in fig. 2c, the specific heat of a glass-forming liquid alloy is significantly greater (perhaps as much as two times greater just above T_g) than that of the corresponding crystalline solid. Integration of liquid and solid specific heats over the temperature interval between T_f and T_g shows that the heat of crystallization of the glass at T_g should be significantly less than the latent heat of melting at T_f; this is confirmed by direct measurement and provides a way of estimating the liquid specific heat in the undercooled region when values are not otherwise available (GARRONE and BATTEZZATI [1985]). The large specific heats of glass-forming liquids relate directly to the loss of entropy on cooling. Glass-forming alloys typically show a strongly negative heat of mixing of their components. This heat is greater at lower temperatures, again reflecting the increased order in lower-temperature liquids. Figure 4 shows the entropy difference between liquid and crystalline states in an easy glass-forming system for which reliable specific heat data are available. The ideal glass transition temperature (zero entropy difference) is just a little below the measured glass transition (determined as in fig. 35, below). The proximity means that thermodynamic parameters can be used to estimate glass transition temperatures, as shown by BATTEZZATI and GREER [1987].

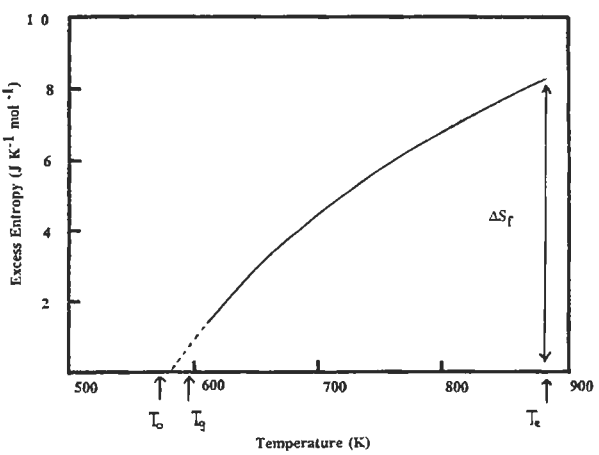


Fig. 4. Excess entropy of liquid $Pd_{40}Ni_{40}P_{20}$ as a function of temperature, calculated from specific heat data of EVANS *et al.* [1987]. The excess entropy has the value of the entropy of fusion, ΔS_f, at the eutectic temperature, T_e, and extrapolates to zero at the ideal glass transition temperature. This is seen to be very close to both the observed kinetic glass transition, T_g, and the ideal transition temperature, T_0, calculated from the hole theory of DUBEY and RAMACHANDRARAO [1984].

References: p. 1818.

The thermodynamics of the glassy state itself can be estimated using specific heat measurements as above. In addition, electrochemical (E.M.F.) measurements on the glass can be used to determine its thermodynamic parameters directly; the procedures have been discussed by BORMANN and ZÖLTZER [1992]. CALPHAD (CALculation of PHAse Diagrams) methods involve the determination of a consistent set of thermodynamic parameters for an alloy system by fitting the phase diagram. When this takes account of specific heat and other data (such as E.M.F. measurements) reliable estimates are possible, even in the undercooled liquid/glass regime. SAUNDERS and MIODOWNIK [1986] first successfully applied this approach to interpret amorphous phase formation by solid-state reactions, and it has since been developed further (BORMANN *et al.* [1988]).

A striking example of the link between the liquid and an amorphous alloy formed in the solid state, and of the use of thermodynamics in interpreting experimental results, is provided by *inverse melting*. This phenomenon occurs in the metastable body-centred cubic solid solution in the Cr–Ti system. When a solid solution of $Cr_{55}Ti_{45}$ is annealed below ~750°C it transforms polymorphically to an amorphous phase of the same composition. Yet the bcc phase itself forms from the liquid. Thus if the amorphous phase is thermodynamically continuous with the liquid, its Gibbs free energy must equal that of the bcc at *two* temperatures (fig. 5). The upper temperature is the normal melting/freezing point. At the lower temperature there is an equilibrium between the amorphous and crystalline phases, but the amorphous phase is the more stable phase below the temperature; this is an inverse melting point, because in equilibrium at this point a crystal would melt *on cooling*. The Cr–Ti system has been subjected to a CALPHAD analysis and this predicts behavior of the form shown in fig. 5. The corresponding metastable phase diagram is shown in fig. 6; in this diagram only three phases are included — the

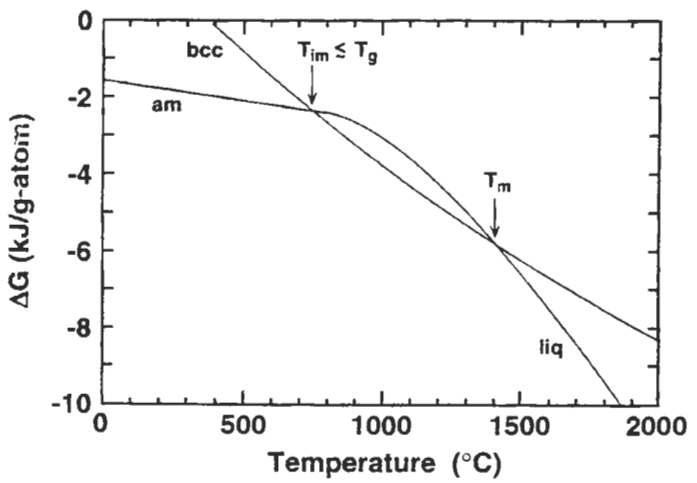


Fig. 5. Free energy curve of the liquid/amorphous phase with respect to the bcc solid solution for Ti–55 at% Cr. Hcp titanium and bcc chromium are taken as reference states at each temperature. T_m and T_{im} denote the melting and the inverse melting temperatures, respectively (After BORMANN [1994]).

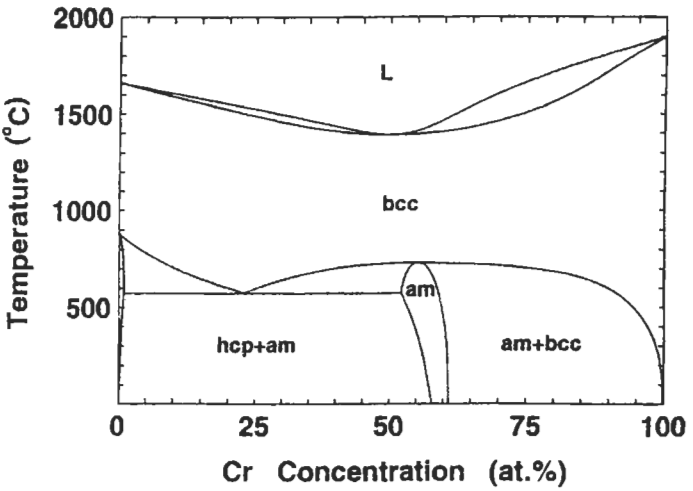


Fig. 6. Metastable phase diagram of the Ti–Cr system as calculated by the CALPHAD method, considering only equilibria between the hcp and bcc solid solutions and the amorphous or the liquid phase (After BORMANN [1994]).

hcp and bcc solid solutions and the liquid. Intermetallic compounds are excluded. The metastable diagram is relevant to experiments on solid-state amorphization because at the temperatures used the compounds have difficulty in nucleating. As shown in fig. 6, the amorphous phase appears at low temperatures. The thermodynamic analysis shows that this is really the same phase as the liquid. Consideration of the temperature dependence of the free energies in fig. 5 shows that the liquid or amorphous phase at the inverse melting point must have a lower entropy than the bcc solid solution. This remarkable result is possible because the amorphous phase is strongly chemically ordered, whereas the solid solution is disordered. The greater chemical order in the amorphous phase outweighs its greater positional entropy. Given the chemical order in the amorphous phase, it is possible to regard it as a type of intermetallic compound, as its phase field in fig. 6 would imply.

The analysis of inverse melting was concerned with a transformation occurring near thermodynamic equilibrium (albeit a constrained equilibrium in which intermetallic compounds do not appear). Other types of solid-state amorphization have prompted consideration of what the ultimate limit to crystalline stability might be — if a liquid or amorphous phase did not nucleate, how far could the crystal be superheated? FECHT and JOHNSON [1988] have considered this, and show that the proliferation of vacancies in the crystal would lead to its entropy becoming greater than that of the liquid at a finite temperature (roughly twice the thermodynamic melting temperature T_m). In a further application of Kauzmann's concept of a structural paradox, they suggest that this point would represent the absolute stability limit of the crystal. More generally, how far could the crystal exist beyond its equilibrium phase field? Its thermodynamic stability could be reduced not only by heating, but also for example by changing its composition or by

injecting defects (by irradiation or deformation). These points have been considered by FECHT, DESRÉ AND JOHNSON [1989], who suggest that melting or amorphization could become a continuous second-order transition at high enough solute contents. It is as yet unclear whether any observed amorphization fits this pattern, or occurs closer to thermodynamic equilibrium.

2.2. Compositions of amorphous alloys

Amorphous alloys are now known for many compositions. Within a given alloy system, the range of compositions which can be made amorphous depends on the production method and conditions. For example, with rapid liquid quenching, the range of glass formation is wider for higher cooling rates. However, even at extreme cooling rates (for example in pulsed laser quenching, § 3.5), the glass-forming composition ranges remain limited. Table 1 gives some examples of compositions which can be made amorphous by 'standard' techniques such as melt-spinning (§ 3.1). More complete listings may be found in TAKAYAMA [1976], SURYANARAYANA [1980] and JOHNSON [1986].

The main categories are:

(i) *Late transition metal–metalloid.* This is probably the most studied category of amorphous alloys, as it includes the Fe-, Co- and Ni-based alloys which are commercially exploited because of their soft magnetic properties (SMITH [1993]). The alloys consist of one, or a combination of, the late transition metals and one, or a combination, of the metalloids. The late transition metals are here considered to be in groups VIII and IB. The metalloids are in groups IIIB, IVB or VB, including Al and Sb. Amorphous alloys are mostly formed by rapid liquid quenching, and the glass-forming compositions are centered around deep eutectics, typically in the range 15 to 25 at.% metalloid. Binary amorphous alloys in this category are common, but the glass-forming ability is generally found to be higher when there are more components, probably because the eutectics are deeper.

(ii) *Early transition metal–metalloid.* Early transition metals from groups IVA, VA and VIA also form amorphous alloys in combination with metalloids, with compositions again centered on deep eutectics in the range around 20 at% metalloid. Partial substitution of an early transition metal for the late transition metal in category (i) can improve the glass-forming ability. When the alloys contain a refractory metal, the crystallization temperatures can be particularly high, and this has led to interest in early transition metal-based glasses as diffusion barriers (DÖRNER *et al.* [1991]).

(iii) *Early transition metal - late transition metal.* These alloys of group VIII or IB elements with group IVA, VA or VIA elements form the second most-studied category. When they are made by rapid liquid quenching, the glass-forming composition ranges are typically much wider than in the metalloid-containing alloys (categories i and ii), covering several eutectics and intermetallic compounds. Although the glass-forming ability is often greatest around the centre of the composition range, Fe, Ni and Co alloys with as little as 9 at% Zr have been made amorphous. These alloys have attracted much attention because they are the main category which can be made amorphous by solid-state reaction (§ 3.7).

Table 1
Examples of glass-forming alloys[a], classified into the categories discussed in the text.

Type	Alloy	Glass-forming range	References
(i) Late transition metal–metalloid	$Fe_{100-x}B_x$	x=12–25	HASEGAWA and RAY [1978]
	$Pd_{100-x}Si_x$	x=14–22	LEWIS and DAVIES [1976]
	$Ni_{100-x}B_x$	x=17–18.5, 31–41	DONALD and DAVIES [1980]
	$Pt_{100-x}Sb_x$	x=34–36.5	SRIVASTAVA *et al.* [1972]
	$Fe_{80}P_{13}C_7$		DUWEZ [1967]
	$Fe_{40}Ni_{40}B_{20}$		LUBORSKY [1977]
	$Fe_{40}Ni_{40}P_{14}B_6$		POLK and CHEN [1974]
	$Co_{70.5}Fe_{4.5}Si_{15}B_{10}$		MAKINO *et al.* [1990]
	$Pd_{40}Ni_{40}P_{20}$		MAITREPIERRE [1970]
(ii) Early transition metal–metalloid	$Ti_{100-x}Si_x$	x=15–20	INOUE and MASUMOTO [1980]
	$Re_{65}Si_{35}$		FISCHER *et al.* [1978]
	$W_{80}N_{20}$		DÖRNER *et al.* [1991]
	$W_{60}Ir_{20}B_{20}$		FISCHER *et al.* [1978]
(iii) Early transition metal–late transition metal	$Nb_{100-x}Ni_x$	x=40–70	RUHL *et al.* [1967]
	$Cu_{100-x}Zr_x$	x=25–60	RAY *et al.* [1968]
	$Ni_{100-x}Zr_x$	x=10–12, 33–80	CLEMENS *et al.* [1984], GREGAN *et al.* [1981]
	$Fe_{100-x}Zr_x$	x=9, 72, 76	HORVÁTH *et al.* [1988]
	$Ta_{100-x}Ni_x$	x=40–70	GIESSEN *et al.* [1976]
	$Ta_{55}Rh_{45}$		FISCHER *et al.* [1978]
	$Zr_{60}Ni_{25}Al_{15}$		INOUE *et al.* [1990]
(iv) Aluminum-based	$Al_{100-x}La_x$	x=10, 50–80	INOUE *et al.* [1988b]
	$Al_{70}Ni_{15}Si_{15}$		DUBOIS *et al.* [1992]
	$Al_{75}Cu_{15}V_{10}$		HOLZER and KELTON [1991]
	$Al_{85}Ni_{10}Zr_5$		DUBOIS *et al.* [1992]
	$Al_{80}Ni_{10}Y_{10}$		INOUE *et al.* [1988a]
	$Al_{85}Fe_{10}La_5$		INOUE *et al.* [1988a]
	$Al_{85}Ni_5Y_8Co_2$		INOUE *et al.* [1990]
(v) Lanthanide-based	$La_{100-x}Al_x$	x=18–34	AGYEMAN *et al.* [1979]
	$La_{100-x}Ge_x$	x=17–22	AGYEMAN *et al.* [1979]
	$La_{100-x}Au_x$	x=18–26	JOHNSON *et al.* [1975]
	$Gd_{100-x}Fe_x$	x=32–50	BUSCHOW [1979]
	$La_{55}Al_{25}Ni_{20}$		INOUE *et al.* [1989]
(vi) Alkaline-earth	$Mg_{100-x}Zn_x$	x=25–32	CALKA *et al.* [1977]
	$Ca_{100-x}Al_x$	x=12.5–47.5	SOMMER *et al.* [1978]
	$Mg_{65}Cu_{25}Y_{10}$		KIM *et al.* [1990]
	$Be_{40}Zr_{10}Ti_{50}$		HASEGAWA and TANNER [1977]
(vii) Actinide-based	$U_{100-x}Co_x$	x=24–40	RAY and MUSSO [1976]

[a] All compositions are quoted in at.%. The amorphous alloys have been formed by a variety of techniques, mostly rapid liquid quenching. The compositions given are those of representative alloys and do not imply that amorphous phase formation at other compositions in the same system would not be possible. For binary alloys, however, some indication is given of glass-forming ranges.

References: p. 1818.

(iv) *Aluminum-based alloys.* Aluminum-based amorphous alloys have attracted attention only quite recently, when ductile compositions with >75 at.% Al were discovered; these are of interest because of their high specific strength (§ 4.4.1). The alloys most studied have 5 to 15 at% of a lanthanide and 5 to 15 at% of a late transition metal. Amorphous phase formation is also found in the binary aluminum–lanthanide systems, at approximately 10 and 50–80 at% lanthanide. Other amorphous aluminum alloys are based on additions of an early transition metal and a late transition metal, or of a late transition metal and metalloid.

(v) *Lanthanide-based.* Alloys of a lanthanide with 10 to 35 at.% metalloid can form amorphous alloys. Also, binary alloys of the lanthanides, or Y, with late transition metals (groups VIII and IB) have deep eutectics at lanthanide-rich compositions, and some glass-forming ranges have been found around these. An alloy of this type (La–Au) was the first to demonstrate solid-state amorphization by interfacial reaction (SCHWARZ and JOHNSON [1983]). Of particular interest are alloys from the same Al–lanthanide-late-transition metal systems as yield ductile Al-rich amorphous alloys. When lanthanide-based, e.g., $La_{55}Al_{25}Ni_{20}$ (INOUE *et al.* [1989], these alloys exhibit unusually high glass-forming abilities. The critical cooling rate for glass formation is exceptionally low < 10 K/s) and the amorphous phases are remarkably resistant to crystallization, crystallizing only at temperatures ~ 100 K above T_g. (Most amorphous alloys crystallize rapidly at or slightly above T_g.) The very good glass-forming ability has the consequence that bulk amorphous samples (for example, rods of diameter up to ~ 1 cm) can be cast (INOUE *et al.* [1993]), in marked contrast with the thin sections associated with the rapid quenching required for more normal glass-forming alloys.

(vi) *Alkaline-earth based.* The group IIA metals can have quite wide composition ranges for amorphous alloy formation when alloyed with a metal from groups IB, IIB, IIIB, IVA, VIIA or VIII. A ternary addition of a lanthanide can greatly improve the glass-forming ability (KIM *et al.* [1990]). For example, compositions such as $Mg_{65}Cu_{25}Y_{10}$ show glass-forming ability similar to that of the lanthanide–aluminum-transition metal alloys discussed in (v), and can similarly be cast in bulk (LI *et al.* [1992], INOUE *et al.* [1993a]).

(vii) *Actinide-based.* A number of amorphous binary alloys based on actinides are known, which do not fit into a clear pattern.

The examples in table 1 illustrate the main types of amorphous alloys, but there are many others. The table emphasizes binary compositions, but many of the compositions which form amorphous phases most readily have three or more components, and some of these may be difficult to categorize. As amorphous alloys are optimized for various applications, multicomponent compositions are developed. An example is the composition $Fe_{73.5}Cu_1Nb_3Si_{13.5}B_9$, designed to devitrify to a fine-grained structure with good soft magnetic properties (§ 4.3). Like most complex compositions, however, this is based on a simple amorphous alloy type, in this case type (i). Other compositions, such as $Zr_{41.2}Ti_{13.8}Cu_{12.5}Ni_{10.0}Be_{22.5}$, have been developed to obtain exceptionally good glass-forming ability (PEKER and JOHNSON [1993]. These recent findings have been put in their context in a reportage by GREER [1993]. Subsequently, INOUE [1995] has found a whole series of easy-forming glasses, quaternary and quinary compositions of the type

$Zr_{65}Al_{7.5}(Co,Ni,Cu)_{27.5}$, e.g., $Zr_{65}Al_{7.5}Ni_{10}Cu_{17.5}$ which, like Johnson's alloy, can be prepared as an amorphous cylinder as much as 16 mm in diameter. In an alloy with many components crystallization may be strongly inhibited; this is sometimes referred to as the *confusion principle*. The confusion is increased if the atoms are of differing sizes. In the first example given, zirconium and titanium atoms are large, copper and nickel of intermediate size, and beryllium has the smallest atoms of any metallic element. Alloys of this type also have an exceptionally wide temperature gap between the glass transition and the temperature at which crystallization begins. Other than these broad generalizations, other criteria for selecting compositions to form amorphous phases readily are considered in the next section.

We show here (fig. 7) seven phase diagrams, one from each of the categories in table 1, each with the glass-forming ranges (for cooling rates such as are achieved in melt-spinning). It is very clear that glasses form preferentially near deep eutectics, where viscosity is high before the thermodynamic freezing temperature is reached on cooling, although the glass-forming ranges are usually not symmetrical about the eutectic composition. These phase diagrams link with several of the criteria discussed in the next section.

2.3. Criteria for formation of amorphous phases

In this section, we concentrate initially on glasses made by rapid solidification; those made by irradiation, mechanical milling or interdiffusion will be discussed afterwards. The generic term most commonly used to represent what we are trying to interpret here is *glass-forming ability*, or GFA. This is usually measured by the critical cooling rate, R_c, which is necessary to turn a melt into a glass, avoiding crystallization: when this is in the experimentally accessible range for mainline rapid solidification processes such as melt-spinning, approximately up to 10^7 K/s, then the alloy can be vitrified by such a process. If theory tells us that R_c is higher than this, then at best tiny amounts of glass can be made by some much faster quenching process such as in-situ quenching with nanosecond or picosecond laser pulses (§ 3.5).

MOTORIN [1983] calculated from first principles the expected homogeneous nucleation rates of crystals in supercooled pure metal melts, using known physical parameters as input. For silver, copper nickel and lead, minimum cooling rates of 10^{12}–10^{13} K/s were derived, which demonstrates clearly enough why no pure metals have been amorphized by melt-quenching. Not even picosecond laser-quenching has been successful in this respect. Thus practically interesting GFAs refer only to alloys — as mentioned in § 1.3, TURNBULL [1981] did propose that crystallization of pure amorphous metals would come into the category of transformations which do not require thermal activation and thus cannot be suppressed.

The range of ideas and approaches which have been proposed to attain an understanding of why some alloy systems can be vitrified and others not, and of what determines the composition ranges over which metallic glasses can be made in practice, is disconcertingly broad. There are almost as many criteria for glass formation as there are for the good life, and almost as little agreement among the protagonists as to which is preferable.

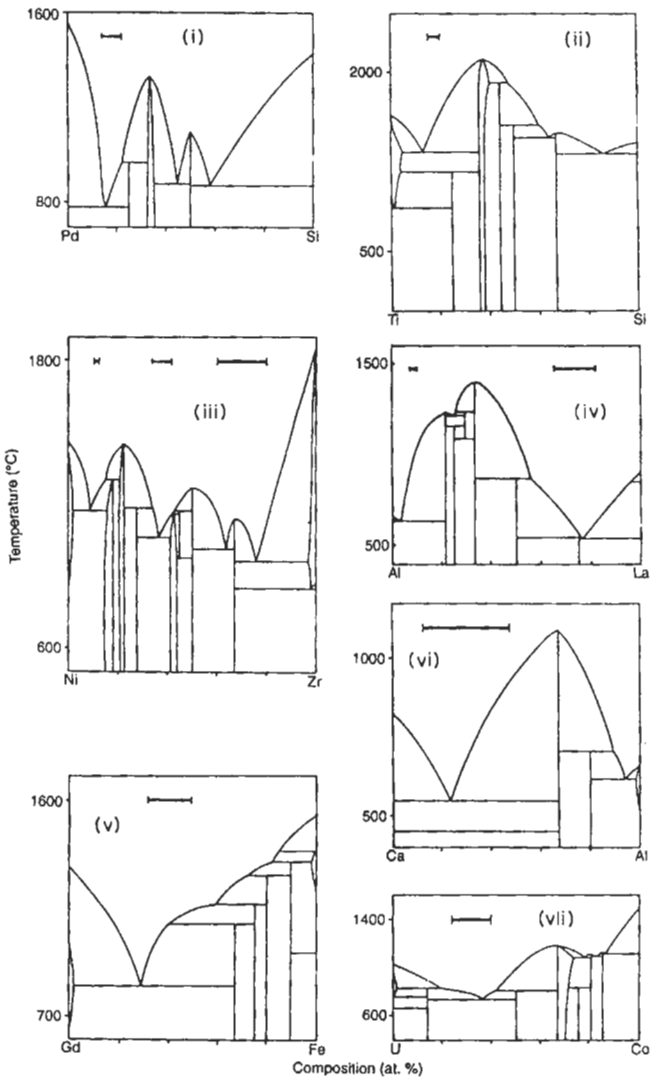


Fig. 7. Equilibrium phase diagrams for examples of each of the categories (i) to (vii) listed in table 1. Glass-forming ranges are indicated.

An early attempt to come to grips with this difficult issue was by POLK and GIESSEN [1978]. The best comprehensive overview of the development of these criteria is by RAMACHANDRARAO [1984]. CAHN [1986a], HAFNER [1986] and MASSALSKI [1986] have published more specialised treatments. Here, only an abbreviated review is feasible.

In spite of a tendency by some theorists to regard this claim as culpably naive, there can be no doubt that atomic size is the single factor that plays the major role in

determining GFA. The recognition that the constituent atoms in a melt must have sufficiently different Goldschmidt radii to permit glass formation goes back to very early experiments on vapor-quenching, in a classic study by MADER *et al.* (1967). A critical radius mismatch of ~ 15% fitted the results, and this value has been accepted ever since. This criterion might be termed an *anti-Hume Rothery criterion,* since Hume–Rothery's celebrated law laid it down that a radius mismatch not exceeding 15% was necessary for extensive solid solution formation in terminal alloy phases. — Mader himself had simulated glass formation by jiggling populations of wooden balls of different sizes on a tray, and found that the radii had to differ sufficiently for "crystallization" to be prevented; subsequently, SIMPSON and HODKINSON (1972) performed the first of several simulations of this kind with rafts of soap bubbles of ~ 1 mm diameter which better simulate real interatomic forces, and again found a critical radius mismatch of ~ 15%.

The next stage was a series of attempts to create GFA *maps,* in which radius mismatch was plotted along one axis of a graph, and some measure of bond strength such as the negative heat of mixing or heat of evaporation along the other — the idea being that strong interatomic bonds render crystallization more difficult. One of several such attempts in this period was by GIESSEN (1981). Such graphs suggest that the bond strength does indeed play a part, but it is a "weak" variable compared to the size mismatch. It should be noted, however, that strong bonding can itself modify an effective atomic size: this was demonstrated by BECHET and REGAZZONI [1991] who showed that the effective size of late-transition-metal atoms in amorphous aluminum-rich alloys is significantly reduced by the strong bonding between the elements.

A very important development of the radius mismatch approach is due to EGAMI and WASEDA [1984], followed by a further development of the same ideas by EGAMI and AUR [1987]. These authors were interested in calculating the microscopic stress levels at the scale of individual atom groups in a glass or a crystalline solid solution. Briefly, they concluded that in a glass, neither the local stress fluctuations nor the total strain energy vary much with solute concentration when these quantities are normalized with respect to the elastic moduli, whereas in a crystalline phase the strain energy rises steadily with solute content. Thus, beyond a critical solute content, glass becomes favored over crystal formation, and in this way, not only GFA but also the glass-forming composition range can be calculated.

The conclusion is that $c_B^{\min}(v_B - v_A) \approx 0.1$, where $c_B^{\min}$ is the minimum solute concentration of B in A required for a glass to form and the v's are atomic volumes. The agreement with observation is good for many systems, though melt-quenching by normal melt-spinning is not necessarily fast enough to give agreement with Egami and Waseda's criterion. Figure 8, for the Ni–Nb system, shows at the bottom the results of applying Egami and Waseda's criterion, and also three different experimentally determined glass-forming ranges for different quenching methods. The "faster" methods agree very well with the theoretical criterion.

Several other theories, of a thermodynamic nature, based on atomic volumes have appeared. One version is due to RAMACHANDRARAO [1980]: he estimated the departure from ideality of melts, in terms of a kind of deviation from a liquid Vegard's Law, and concluded, both theoretically and by comparison with experiment, that melts with a large

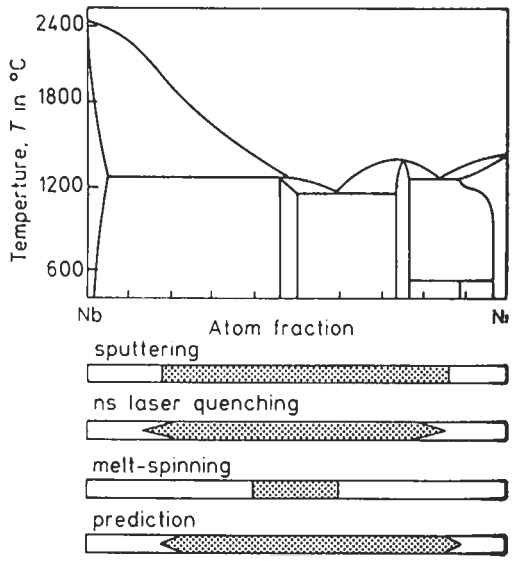


Fig. 8. The Ni–Nb phase diagram, with experimentally determined glass-forming ranges for three methods of progressively diminishing effective quenching speed, together with the range predicted by Egami and Waseda's theory.

deviation, in the sense of having an anomalously small atomic volume, were most likely to form glasses. This finding was interpreted in terms of the enhanced viscosity associated with a small atomic volume (which implies little free volume available to support atomic motion in the melt). — YAVARI *et al.* [1983] independently developed a closely related idea, suggesting that a zero (or negative) change of specific volume on melting of a crystalline species favors glass formation on subsequent rapid solidification. In effect, their idea is that if a crystal is denser than the melt from which it grows, then in growing it rejects free volume into the melt and thereby reduces its viscosity; thus crystal growth becomes self-catalytic. Contrariwise, a less dense crystal raises viscosity and favors glass formation.

Theoretical physicists have developed a number of models to rationalize GFA. The best known is that due to NAGEL and TAUC [1977]. They proposed that a glass is most likely to form if its electronic energy lies in a local metastable minimum with respect to composition change. They showed that if the structure factor corresponding to the first, strong peak of the diffuse X-ray scattering curve, K_p, satisfies the relationship $K_p = 2\,k_F$, where k_F, is the wave vector at the Fermi energy, then the electronic energy does indeed occupy a local minimum. A number of familiar glasses, in the Au–Si, Au–Ge and Co–P systems for example, accurately obey this criterion, but there are also exceptions, e.g., some obscure glasses in the Cs–O and Rb–O systems. This kind of approach, using modern approaches such as pseudopotential theory, has been taken much further in recent years, but a discussion would take us too far here. For further details, the reader is referred to reviews by HAFNER [1981, 1986]. However, as we shall see below, the

ability, demonstrated by Hafner, to make theoretical estimates of the glass transition temperatures of unknown glasses offers a novel way to estimate GFA.

Theories based more explicitly on the need to prevent nucleation of a crystalline phase take two main forms. First, there are models based on an explicit calculation of homogeneous nucleation rates (this begs the question whether heterogeneous nucleation plays a role). The standard approach here is due to DAVIES (1976). He adapts a theory of isothermal crystallization kinetics due to Uhlmann and calculates the cooling rate necessary to bypass the "nose" of the time–temperature-transformation plot thus calculated. (The nose is the minimum time required to initiate homogeneous nucleation, at some temperature well below the thermodynamic freezing temperature). The principal difficulty here is to estimate the viscosity of a supercooled melt, but there are a number of empirical relationships which allow a rough estimate to be made. (There is now available a critical survey of known viscosities of molten metals and alloys, by BATTEZZATI and GREER [1989].) Figure 9 shows some calculated critical quenching rates obtained in this way, and fig. 10 shows how the calculated value of R_c varies with composition across a phase diagram. It will be seen from this second figure that, as a number of people have pointed out and as follows from Davies' theory, the ratio T_g/T_f (where T_g is the glass transition temperature, T_f is the thermodynamic freezing temperature) is a crucial figure of merit in determining GFA, for purely kinetic reasons associated with the need to avoid crystallization. The lower this ratio, the more viscous the melt becomes before it is ever undercooled and the more difficult crystallization becomes, thus enhancing GFA and reducing R_c.

A somewhat related, more thermodynamically biased approach derives from a paper

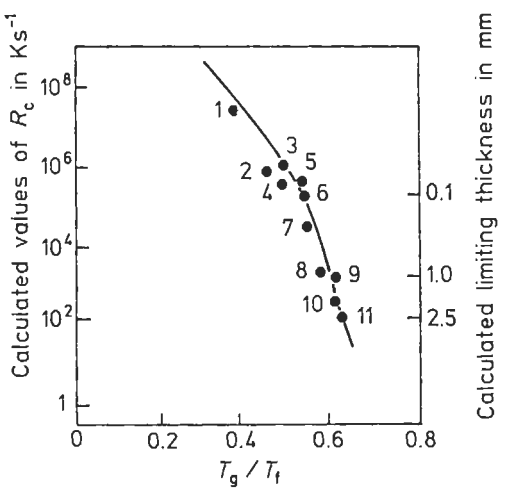


Fig. 9. Calculated critical quenching rates, R_c, for glass formation, and corresponding estimated limiting ribbon thickness for wholly glassy ribbons. Key to alloys: (1) $Fe_{89}B_{11}$; (2) $Au_{78}Ge_{14}Si_5$; (3) $Fe_{83}B_{17}$; (4) $Fe_{41.5}Ni_{41.5}B_{17}$; (5) $Co_{75}Si_{15}B_{10}$; (6) $Fe_{79}Si_{10}B_{11}$; (7) $Fe_{80}P_{13}C_7$; (8) $Pd_{82}Si_{18}$; (9) $Ni_{63}Nb_{37}$; (10) $Pd_{77.5}Cu_6Si_{16.5}$; (11) $Pd_{40}Ni_{40}Pd_{20}$. (After DAVIES [1978]). (An updated version of this plot has just appeared, DAVIES [1995].)

References: p. 1818.

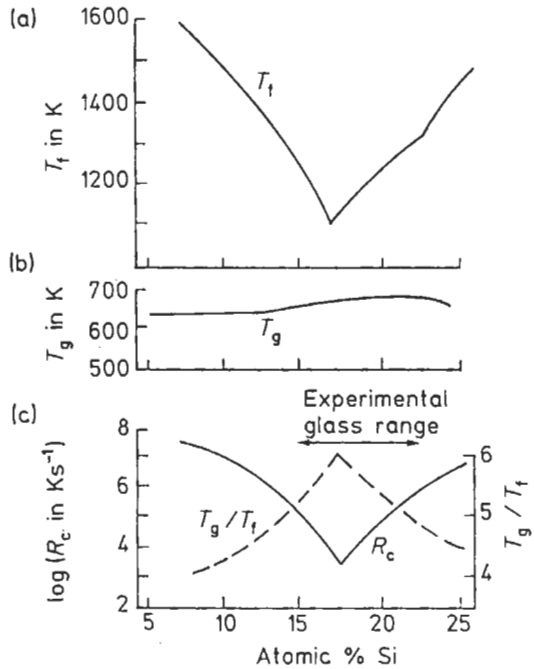


Fig. 10. Equilibrium freezing temperature, T_f, glass transition temperature, T_g, figure of merit, T_g/T_f and calculated critical quenching rate, R_c, for a range of Pd–Si solid solutions.

by J. W. CAHN (1980) on the thermodynamics of metastable equilibria. From this, the role of the relative values of T_g and T_0 becomes clear; here, T_0 is the temperature at which liquid and solid *of the same composition* have the same free energy. A glass forms readily if $T_g > T_0$; what this means is that the melt rigidifies before reaching the temperature at which freezing without composition change (and thus without long-range diffusion), also known as *solute-trapping*, becomes possible. The implications of this are clearly set out in a review by MASSALSKI (1986) and applied to a detailed analysis of GFA across the Cu–Ti system in a paper published about the same time (MASSALSKI and WOYCHIK [1985]). The problem is that, before a glass has actually been made, its T_g is not known. Here, HAFNER's [1983, 1986] ability to calculate glass transition temperatures comes into its own. In fig. 11, for the Ca–Mg system, the calculated T_0 values are shown, and also two estimates for T_g, an upper limit based on the "entropy-crisis" or Kauzmann model, and a lower limit based on a free-volume model. (For a fuller explanation, see HAFNER (1983).) Theoretical composition ranges for glass formation are shown for two cooling rates; here the criterion $T_g > T_0$ is the central consideration; agreement with experiment is quite good. Another detailed thermodynamic analysis of metallic glass formation near eutectic troughs was published by HIGHMORE and GREER [1989].

Another set of theories is based on the postulate that metallic melts are not homo-

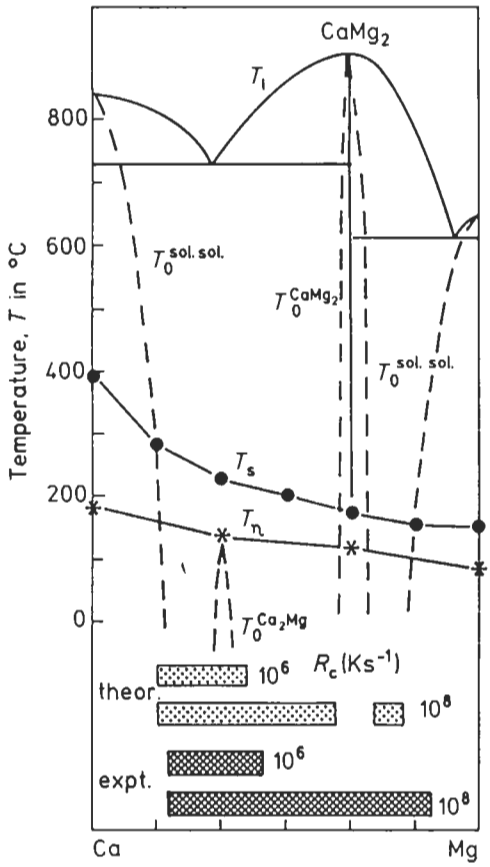


Fig. 11. Calculated phase diagram, T_0 values and upper and lower limits (T_s,T_η) for T_g for the Ca–Mg system, with theoretical and experimental glass-forming ranges. (After HAFNER [1983, 1986]).

geneous in composition but contain compositional clusters. Such clusters are held to aid crystal nucleation, and their absence or weak development to aid glass formation. Contrariwise, short-range order (the converse of clustering), which is believed (though there is a singular absence of diffraction evidence on this point) to increase with falling temperature of a glass, just as it does in a crystalline solid solution, should enhance GFA. (See, for instance, WAGNER's [1986] survey of SRO in metallic glasses; at least, it has been established for Cu–Ti that a glass has higher SRO than the melt from which it is quenched (SAKATA *et al.* [1981]).) Indeed, there is some evidence that some Cu–Ti alloys show clustering in the melt whereas the glass has SRO. This field of research, which has a large literature, has been somewhat beset by controversy: for further details, the reader is referred to RAMACHANDRARAO [1980, 1984].

Finally, we return to a less sophisticated "figure-of-merit" approach which has been

quite successful in rationalizing GFA. DONALD and DAVIES [1978] long ago recognized the awkwardness of theories which related GFA to quantities (such as T_g/T_f) which involve T_g, when normally this is unknown until a glass has been made and examined. They proposed that a good test of GFA is the extent to which the equilibrium freezing temperature of an alloy melt is depressed below the ideal value, which they calculated simply by linear interpolation between the freezing points of the constituent metals. This simple method was then developed by WHANG [1983], who took into account the modification required to allow for possibly extensive solid solution of one metal in the other, in the solid state. This is necessary because solid solubility reduces the slope of the T_0 versus composition curves (like those shown in fig. 11). Whang generated figure-of-merit maps in which one axis gave T_{LR}, defined as $T_{LR} = \Delta T/T_L°$, where ΔT is the difference between the ideal freezing temperature for an alloy ($T_L°$), defined as above, while along the other axis is C_{er}, a simple measure of the amount of the solubility of the minor constituent in the major at the eutectic temperature — a large value of T_{LR} implies a severely depressed liquidus, while a small C_{er} implies a steeply sloping T_0 versus composition curve ... both factors favouring easy glass formation. In fact, the maps so generated show a clear boundary between glass-forming and non-glass-forming alloys — DUBEY and RAMACHANDRARAO [1990] have developed Whang's model to show that most eutectic phase diagrams can be expected to show an asymmetry of GFA, in the sense that glass formation is easier just to one side of the eutectic composition than on the other; the melt-spinning range indicated in fig. 8 shows an example of this, and several of the diagrams in fig. 7 exemplify the same feature.

Whang's theory was then adapted by TENDLER [1986] to show, for a series of Zr-based alloys, that alloys which according to Whang's criterion should be good glass-formers are also those in which there is anomalously *fast diffusion*, in the special sense introduced by WARBURTON and TURNBULL [1975] (see ch. 7, § 4.3.2). Figure 12 shows one of Tendler's figures for a series of Zr–M alloys. All the alloys showing fast diffusion (Zr with Cr, Mn, Fe, Co, Ni, Cu and Be) are also glass-formers. For fast diffusion, the solute atom must be much smaller than the solvent atom (for details see Tendler's paper) and this clearly also favors GFA. In fact, some years ago TURNBULL [1976] had predicted just such an association between GFA and fast diffusion. Care must of course be taken in setting up such diagrams as fig. 12 because, as shown by Bechet and Regazzoni, large radius ratio disparities (leading to strong bonding and to fast diffusion) can modify effective atomic sizes. Fast diffusion is also linked to glass-forming mechanisms through the fact that in the solid-state amorphization reaction (SSAR), one atomic species diffuses very much faster than the other.

This by no means exhaustive overview of the models and theories that have been advanced to make sense of glass-forming systems and ranges might well seem discouraging, because at first sight they are mutually exclusive. In fact, hidden cross-connections undoubtedly exist: the linkage between Whang's thermodynamic approach (related to terminal solid solubilities) and Tendler's association between fast diffusion and GFA clearly comes from a correlation of both solid solubilities and fast diffusion with atomic size ratios. Perhaps in due course even the electronic criteria studied by Nagel and Tauc and by Hafner may prove to be linked with some of the other ideas, e.g., the free volume

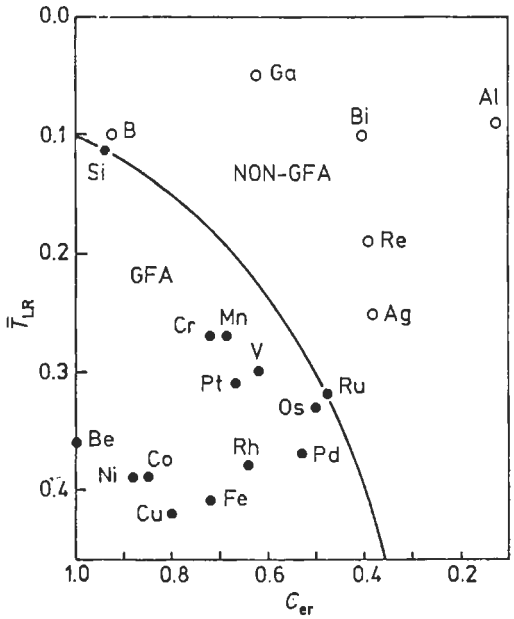


Fig. 12. A number of Zr–M alloys plotted on a Whang graph, separating glass-formers (GFA) from non-glass-formers. (After TENDLER [1986]).

approach due to Ramachandrarao and Yavari. Our own view is that simple geometry — atomic sizes — will prove to be the main criterion that in various subtle ways incorporates the others.

2.3.1. Criteria for amorphization by irradiation, and mechanically aided and induced amorphization

Some alloys can be amorphized either by irradiation with electrons or heavier particles (see § 3.4) or by heavy deformation, if need be followed by annealing (see § 3.8). In § 3.4, below, a brief account is offered of an unresolved controversy regarding the relative roles of two amorphization criteria — the need to destroy the (chemical) long-range order in an irradiated compound to a sufficient degree versus the need to attain a critical concentration of point defects. The criteria for amorphizability in the case of mechanically aided or induced amorphization (see for instance SCHULTZ and HELLSTERN [1987]) are generally assumed to be similar to those applying to irradiation, but no critical comparison of the two families of techniques appears to have been made. There is an important new approach, however, due to BEKE *et al.* [1991]. They estimated the additional elastic (mismatch) energy which is stored in an ordered solid solution when its long-range order is destroyed, whether mechanically or by irradiation; this depends, of course, inter alia on the atomic sizes. This energy is then compared with the ordering energy, and it was concluded that amorphization is possible if (a) the (virtual) critical disordering temperature is above the melting temperature, and (b) the ratio of the

elastic mismatch energy to the ordering energy exceeds a critical amount. A comprehensive test against experimental facts gave widespread though by no means universal agreement.

3. *Practical methods of creating metastable phases and microstructures*

In what follows, various techniques are outlined, both those involving solidification of a melt and those which operate entirely in the solid domain. In exemplifying what can be achieved, emphasis is placed upon amorphization, but most of the methods have of course also been used to make metastable crystalline phases.

3.1. Rapid quenching from the melt

All casting methods involve a chill zone close to the mould surface; the metal in the chill zone cools more rapidly than the interior of the casting, and its grain morphology and composition are accordingly different. Such a zone will normally cool at a rate of up to some hundreds of K/s. The modern techniques of *Rapid Solidification Processing* (RSP), with which this chapter is principally concerned, however, involve cooling rates in the range 10^4–10^7 K/s, and require an approach quite distinct from the traditional casting methods.

Atomization of liquid metal to form small droplets — one of the methods widely used today — goes back to the last century, and as recounted by JONES [1981] and more recently by CAHN [1993] in historical surveys, several investigators in the period 1925–1955 developed variants of chill-casting which gave estimated cooling rates up to 10^5 K/s, while several studied extension of solid solubilities resulting from such cooling rates. However. the remarkable present-day developments in RSP were initiated by Duwez in 1959–60. Duwez, in California, set out to establish whether a *continuous* metastable series of solid solutions could be created in the Cu–Ag system, to bring it into line with the Cu–Au and Au–Ag systems which both show continuous solid solubility, as required according to the familiar Hume–Rothery rules. For this purpose, he argued, the melt would have to be frozen fast enough to inhibit nucleation of two distinct fcc phases. Having unsuccessfully tried a number of other approaches, he designed the *Duwez gun,* a device in which a gaseous shock wave atomizes a drop, ≈ 10 mg, of molten alloy and projects the microdroplets into contact with a copper substrate, or *chill block,* to produce small foils, or *splats.* The technique was by some given the onomatopoeic nickname *splat-quenching*; this term, disapproved of by Duwez, fell out of use together with the original gun, as improved methods were developed.

In 1960, Duwez and his collaborators were successful in making a continuous metastable series of Cu–Ag alloys without any two-phase region, and in the same year also discovered the first *metallic glasses* made by rapid quenching, in the Au–Si and Au–Ge systems (DUWEZ *et al.* [1960]). The essential condition for rapid quenching from the melt, as realized in the Duwez gun, was that small particles of liquid metal should rapidly be flattened into a thin sheet in intimate contact with a good heat sink. More generally — and this was the real importance of the innovation wrought by Duwez —

only by starting from the molten metal was it possible to achieve really high rates of cooling in the *solid.* The importance of RSP arises *both* from the changes brought about in the freezing process *and* from the more rapid cooling of the solid so formed — but it should be noted that not all melt-quenching methods in fact create fast cooling in the solid stage. The inadequacy of melt-spinning in this respect emerges from some striking observations concerning age-hardening kinetics (see § 5.1).

The Duwez gun was used for a decade to investigate the crystallographic and calorimetric characteristics of numerous metastable splat-quenched crystalline phases and of a few glasses, notably $Pd_{80}Si_{20}$ which was for a number of years a standard composition for those interested in the properties of metallic glasses. DUWEZ [1967] has described these pioneering days, and also the background to his original experiments.

The splat-quenching principle was developed to provide faster cooling rates still, in the form of the *rotating-wing quencher,* attempted by various early investigators and designed into an effective form by SOMMER *et al.* [1987]: here a disc fitted with vertical wings rotates very rapidly about a vertical axis, in high vacuum, and a drop of melt is allowed to fall into the rotating wings. The instrument has recently been used to quench out disordered Ti_3Al (BRAUN *et al.* [1994]), a good test of efficacy in fast cooling.

The next stage came when the first methods of making *continuous* rapidly quenched ribbons were invented. All these depend on contact between a thin liquid ribbon and a *moving* chill block. The old method of *single-roller melt-spinning* was re-invented by Pond in 1958 and later published (POND and MADDIN [1969]); *twin-roller melt-spinning* is due to CHEN and MILLER [1970] and was improved by MURTY and ADLER [1982] but is not very efficient and is little used now; *melt-extraction* was developed by MARINGER and MOBLEY [1974]. The fourth major technique, much used in basic research, is the *drop-smasher,* also known as *piston-and-anvil quencher,* in which an alloy drop is levitation-melted inside a conical induction coil, released and quenched between two moving copper surfaces to form a disc (HARBUR *et al.* [1969], BEGHI *et al.* [1969] and CAHN *et al.* [1976]). The principles of these four methods are indicated in fig. 13. Of these methods, melt-spinning and its variant *planar flow-casting* (fig. 14), which allows wide sheets to be made, approaching a meter in width, are in practice the most important; the kinematic and heat-flow aspects of these processes are analysed by SHINGU and ISHIHARA [1993]. Their analysis produces an analytic relationship between ribbon thickness and substrate velocity: ribbon thickness varies as the -0.8 power of substrate velocity, ranging from 70 μm for 15 m/s to 20 μm for 50 m/s. (*See also* ch. 29, § 5.2).

Plasma-spraying is a method which in a sense represents a return to Duwez's microdroplet approach: an electric discharge between two electrodes in argon provides energy to melt alloy globules which are then projected by an argon jet on to a substrate to build up a deposit. The starting material is pre-alloyed powder, 50–100 mm in diameter. The hot gas jet needs to be blown aside to ensure rapid cooling of the impinging droplets. GAGNÉ and ROY [1982] have succeeded in plasma-depositing layers of a Ni–Cr–Si–B glass on to a metallic substrate. A full account of the use of this technique was published by SAFAI and HERMAN [1981], and some of the metastable phases made in this way are discussed by HERMAN and BHAT [1980]. The method can also be used to solidify fine droplets in flight for subsequent compaction, though it has

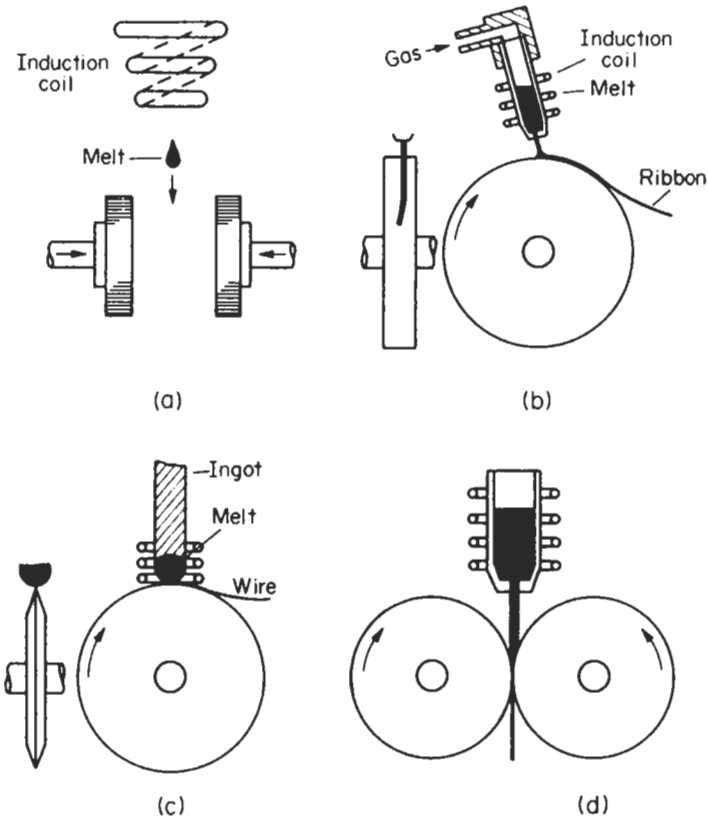


Fig. 13. Principal methods of rapid quenching from the melt: (a) drop-smasher, using levitation-melting by induction; the pistons are pneumatically or electro-magnetically accelerated and come into action when the falling drop breaks a light beam; (b) melt-spinning; (c) pendant-drop melt-extraction (there is a distinct version with the wheel dipping into a melt surface); (d) twin-roller quenching device.

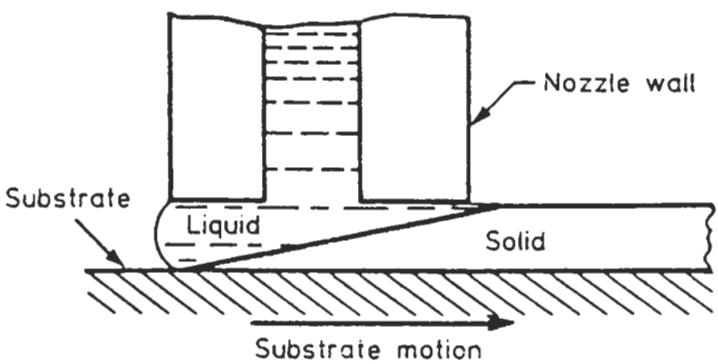


Fig. 14. Planar flow-casting, schematic.

been little used in this mode. A simple flame-heated metal-spraying gun has also been used to deposit glassy layers of $Fe_{40}Ni_{40}P_{14}B_6$ (SHINGU [1979]).

A number of different methods have been used for the production of RSP *wires*. One is *free jet melt-spinning* of a molten alloy, usually a steel, through a nozzle into a gaseous quenching medium (MOTTERN and PRIVOTT [1978]). The problem here is to avoid Rayleigh instabilities which tend to break up the free jet: this is best achieved by alloying the steel with aluminum, quenching into carbon monoxide and relying on the surface oxide layer formed on the freshly formed jet to resist its break-up. Other alloys can be correspondingly treated. An alternative approach (OHNAKA and FUKUSAKO [1978], INOUE *et al.* [1982], OHNAKA [1988]) relies on a variant of melt-spinning: the alloy jet impinges on the inner surface of a rotating vessel containing water which is held in an annular shape by centrifugal force. The method is known as the *in-rotating-water splat-processing* (INROWASP) (fig. 15). This, like the free-jet method, produces wires 100–200 μm in diameter. Rayleigh instabilities are avoided by carefully matching the speed of the liquid metal jet and the rotational speed of the water annulus. There is evidence from comparative relaxation studies (INOUE *et al.* [1983]) that a glass in the form of wire is solidified at about the same rate as the same glass in the form of a melt-spun ribbon several times thinner, but (because of continuing close thermal contact) cools faster *after* solidification (an important aspect of rapid solidification that is too often ignored). In contradiction to this, a very recent study of the microstructure of Ni_3Al converted into wire by the in-rotating-water method (CHIBA and HANADA [1996]) provides some (though not conclusive) microstructural evidence that wires made in this way have solidified more slowly than melt-spun ribbons of equivalent thickness. — Another method of rapid solidification which has recently enjoyed a revival is the *Taylor wire method*: here a thick wire is encapsulated in a glass which is heated and drawn down, together with the molten wire

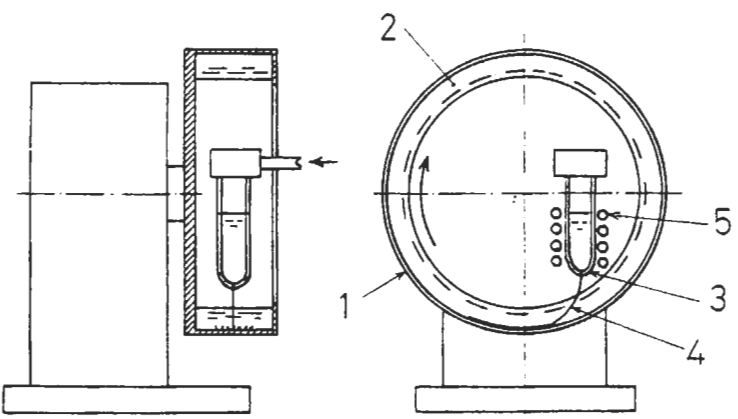


Fig. 15. In-rotating-water splat-processing (INROWASP), schematic. 1, Rotating drum; 2, liquid coolant; 3, nozzle; 4, melt jet; 5, induction coil. (After OHNAKA [1985]).

References: p. 1818.

within; the change in cross-section combined with heat withdrawal by the encapsulating glass provides effective rapid cooling. Apart from the rapid-freezing aspect, this is a good method of making ultrafine wires (see review by DONALD [1987]).

LEBO and GRANT [1974] published a key paper on production of RSP aluminum alloy powder by atomization and also of splats against a cold metal surface, followed by consolidation of the powder at a temperature and for a time designed to prevent deterioration of the desirable properties induced by RSP. Another key paper, describing the compaction and properties of splat-quenched Al—Fe alloys, was by THURSFIELD and STOWELL [1974]. All subsequent research and development of industrial uses of RSP as applied to crystalline alloys has been based on this approach. Atomization methods in use (reviewed by SURYANARAYANA [1991]) include ultrasonic atomization centrifugal atomization in helium and quenching of droplets between rollers to make flakes (SANKARAN and GRANT [1980]).

Consolidation is either by warm extrusion or by hot isostatic pressing. Even cold explosive compaction can be used (MORRIS [1980]). An alternative approach which is increasingly finding favour is to pulverize a brittle crystalline melt-spun ribbon. A small-scale variant, capable of producing extremely fine, submicron powders, is electrostatic extraction of particles from a liquid–alloy tip drawn out into a point by a strong electric field (PEREL *et al.* [1981]). This apparatus has been used for a detailed metallographic study of the microstructures of RSP aluminum alloy globules (LEVI and MEHRABIAN [1982]; the globules were fine enough for TEM examination without further thinning).

Yet another method which is used from time to time is *spark-erosion,* using a standard workshop machine for the purpose. Electrodes and a vibrating charge of blocks, all of the same alloy, are spark-eroded in a dielectric, which can be a hydrocarbon, water or liquid argon. Small drops of the alloy are heated to an extreme degree and turn into a plasma which is then quenched by the surrounding dielectric fluid. The technique is described by BERKOWITZ and WALTER [1987], and CAHN *et al.* [1988] have exemplified the drawbacks arising from carburisation or oxidation of the alloy by the dielectric fluid; this is why liquid argon is to be preferred (BUCKLOW and DRAIN [1964]).

All the foregoing methods have been systematically reviewed and compared in a major overview by SURYANARAYANA [1991]; he also includes recent, advanced atomization methods, some employing lasers, which are now used to make tonnage quantities of atomised alloy powders. The historical background to the early methods of rapid solidification is reviewed by CAHN [1993].

3.1.1. Cooling rates in rapid solidification processing

The rate of cooling of the melt before, and of the solid after solidification can be accessed by direct measurement. Alternatively, the cooling rate at the freezing stage can only be deduced by measurement of a microstructural feature of the resulting crystalline alloy, most commonly the secondary dendrite arm spacing (ch. 8, § 7.1) or, where apposite, eutectic lamellar spacing; but such a microstructural approach is not available when the product is a glass.

Dendrite arm spacings are determined by both the temperature gradient in the solidifying melt, G_L, and the velocity of the melt/solid interface, V. The cooling rate,

$dT/dt, = G_L.V$. The spacing of primary dendrite arms is not suitable for cooling rate measurement because its relationship with G_L and with V follows distinct power laws, whereas for secondary dendrite arms, both variables are related to the arm spacing, l, by an inverse 1/3 power law. This is demonstrated in fig. 16, (JONES [1982]), based on a range of measurements for Al–Cu and Al–Si alloys. For the fastest cooling rates, a microthermocouple projecting into the melt in a piston-and-anvil quenching unit is used (ARMSTRONG and JONES [1979]). The dendrite method is not quite so reliable for nickel- and iron-based alloys; the eutectic-lamella approach, used for melt-spun aluminum alloy ribbons and relying on extrapolation from slower, measured cooling rates, gave an estimate of 3×10^6 K/s in the neighborhood of the freezing temperature (CHATTOPADHYAY and RAMACHANDRARAO [1980]).

The limitations of such indirect methods are: (1) they are not applicable to glass formation; (2) they can give an estimate only of the effective cooling rate (related to the solidification time and specimen thickness) in the temperature range in which solidification takes place, and not of the subsequent cooling rate in the solid state; (3) each family of alloys requires separate calibration; (4) the dendrites may coarsen after solidification and so yield misleading results. For all these reasons, direct measurement is now regarded as preferable, difficult though it is.

The early direct measurements of liquid-quenching rates which have been published have referred to the piston-and-anvil technique. Thus, HARBUR *et al.* [1969] made use of the metals (suitably insulated) of two dissimilar pistons as the legs of an intrinsic thermocouple, whereas DUFLOS and CANTOR [1982, 1987] used a minute intrinsic thermocouple, electrically shorted by the melt itself. Pyrometry of the droplet has also been employed (e.g., KATTAMIS *et al.* [1973]). Melt-spinning poses a more difficult measurement problem altogether: WARRINGTON *et al.* [1982] used calibrated black-and-

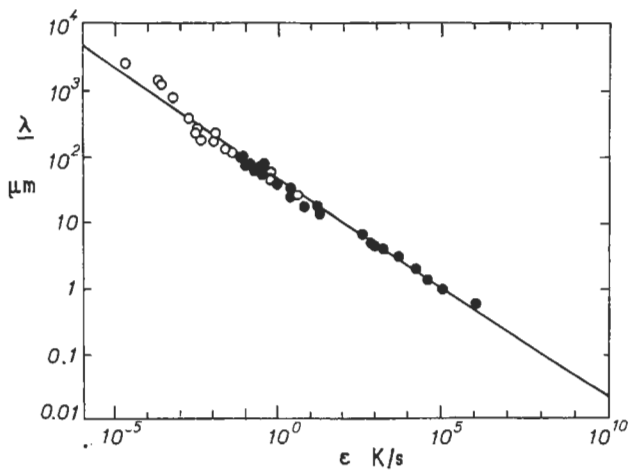


Fig. 16. Dendrite spacing, λ, as a function of cooling rate, ε, near the melting temperature, for Al–4 to 5 wt% Cu (open circles) and Al–7 to 11 wt% Si (solid circles). The plot refers to secondary dendrite arms. The line shown represents $l\varepsilon^{1/3} = \text{const.} = 50$ mm $(K/s)^{1/3}$. (After JONES [1982]).

References: p. 1818.

white photography, while later techniques have focused on the use of calibrated photometry of color photographs of the surface being investigated (GILLEN and CANTOR [1985], BEWLAY and CANTOR [1986]) and the use of infrared imaging devices, which last have proved particularly useful for spray-deposition (GRANT *et al.* [1989], GRANT and CANTOR [1990], BEWLAY and CANTOR [1990]).

Figure 17 shows some results by Duflos and Cantor for a drop-smasher, while fig. 18 shows photometric measurements relating to melt-spinning; the sharp decrease in cooling rate when the ribbon loses contact with the copper wheel should be noted. It is quite clear that the cooling rate varies substantially between the freezing point and ambient temperature. Thus at 1500°C, near the freezing-point of a steel, the cooling rate for piston-quenched steel or iron samples 25–50 μm in half-thickness (the proper measure) is 10^6–10^7 K/s; the higher values apply to the very rapid-acting piston-and-anvil apparatus designed by CAHN *et al.* [1976]. The cooling rate at 1000°C is ≈ 10 times slower than at 1500°C, and at 500°C, 20–30 times slower again. It is evident that the

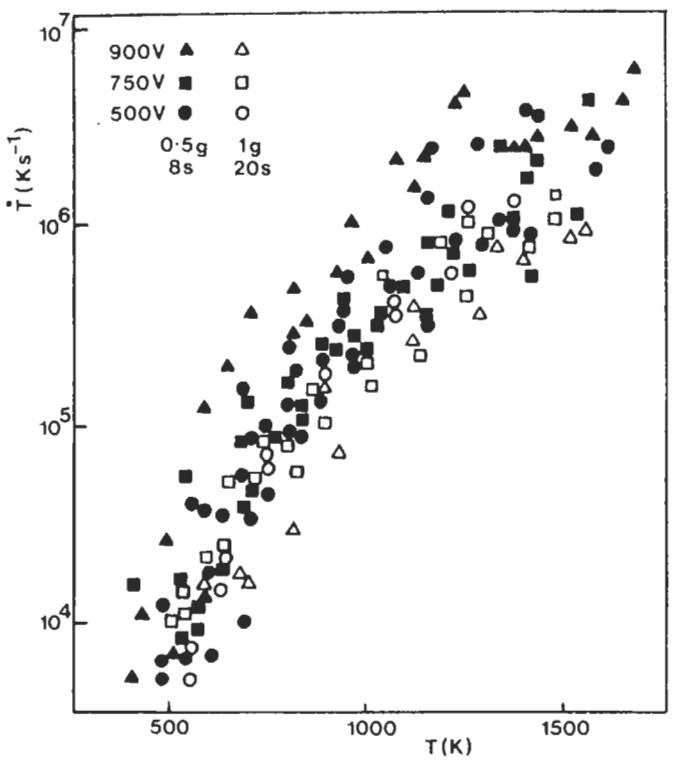


Fig. 17. Data from many cooling curves obtained by embedded microthermocouple and oscilloscope traces during piston-and-anvil drop-smashing of iron, replotted as cooling rate, $\dot{T}$, versus temperature, T. The voltages cited are electromagnetic accelerating voltages; higher voltages imply faster impact. The weights refer to specimen sizes; larger specimens, heated for longer periods, had greater superheat. (After DUFLOS and CANTOR [1982]).

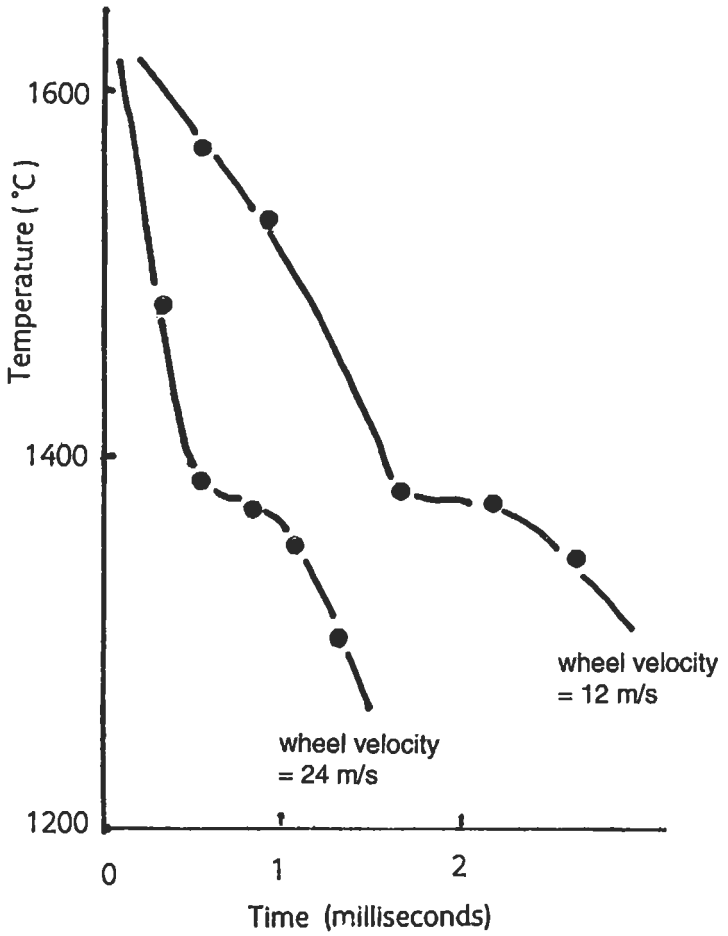


Fig. 18. Temperature measurements during melt-spinning of Ni-5wt% Al, by color-calibrated photometry. (Based on GILLEN and CANTOR [1985]).

custom of characterizing the cooling rate by a single figure is unacceptable, unless the temperature range to which such a figure applies is specified. Single-roller melt-spinning turns out to be slower than piston-and-anvil quenching (HAYZELDEN [1983]): thus, for a 40 mm-thick steel ribbon, the measured cooling rate is 5×10^5 K/s at 1500°C and only 10^4 K/s at 1000°C (because the ribbon has by then left the wheel). It is generally accepted that dual-roller quenching is even slower than single-roller melt-spinning overall, in spite of the presence of double chill block, because the duration of contact between the ribbon and the chill block is so brief. However, the initial cooling rate appears to be very high (DAVIES *et al.* [1978]). In the case of rapid quenching by atomization, cooling rates have been estimated from dendrite spacings: thus, for a gas-atomized superalloy, cooling rates ranging from 10^3–10^6 K/s have been estimated as

particle diameters range from 400 to 10 μm (DUFLOS and STOHR [1982]). The preferred approach at present, however, is by means of heat-transfer calculations, as exemplified for conventional casting in ch. 8., § 2. The crucial parameter here is the *Nusselt* (or *Biot) number, N,* of both specimen and "chill block" (the latter, in the case of a droplet solidifying in flight, is gaseous). $N=hd/k$, where h is the *heat-transfer coefficient, d,* is the effective specimen or chill-block thickness and k is the thermal conductivity. As h increases, cooling changes from Newtonian, via intermediate to "ideal" cooling, with progressively larger temperature gradients in the specimen. h can be estimated from measured cooling rates in the case of piston-and-anvil quenching (CANTOR [1982]), but for powder atomization it is necessary to compute h from first principles for specified boundary conditions (h is very low for pure radiative transfer and several orders of magnitude larger for favorable convective cooling). The theory has become extremely complex: full treatments have been published by LEVI and MEHRABIAN [1982], MEHRABIAN [1982] and CLYNE [1984]; reference should be made to these papers.

The cooling rate varies of course as a function of droplet diameter and depth within the droplet, and is perturbed by the release of latent heat in a way which depends on the undercooling that preceded the start of freezing. A particularly important regime is with heat-transfer rapid enough to permit *hypercooling,* that is, suppression of nucleation till the sample has undercooled so far that the release of the entire latent heat is insufficient to reheat the sample to the equilibrium freezing point: freezing can then go to completion without any further heat extraction. The analysis by CLYNE [1984] concludes that there are two distinct approaches to generating very rapid crystallization from the melt: one can either hypercool, via very small droplets, or else accept a lesser degree of undercooling and arrest the subsequent recalescence by means of highly efficient heat extraction.

In the theory of atomization, it turns out to be more useful to compute solid/liquid interface velocities rather than cooling rates, since the former relate more closely to interface stability. Mehrabian cites values for 50 μm aluminum droplets of 4 m/s at the start of freezing, dropping to 0.04–0.13 m/s in the intermediate stages of freezing.

3.2. Solidification of highly undercooled liquids

When a liquid alloy is cooled, solidification normally starts at a small undercooling below the liquidus temperature, triggered by nucleation on heterogeneities. However, liquid alloys are typically rather resistant to homogeneous nucleation. The most thorough kinetic analysis has been by TURNBULL [1952] of liquid mercury in which homogeneous nucleation appears to occur only at undercoolings as high as 35 to 40% of the absolute melting temperature. Thus there is a possibility of achieving large undercoolings if heterogeneous nucleation can be suppressed or at least inhibited. This is achieved by having a clean melt and ensuring that it does not make contact with a solid container; various methods are illustrated in fig. 19.

In the emulsion method (exploited by TURNBULL [1952] and more recently by PEREPEZKO [1984]) the liquid alloy is dispersed in an inert carrier liquid. When the number of alloy droplets exceeds the number of potent heterogeneous nucleants originally in the alloy, some of the droplets must exhibit substantial undercooling. However, for the

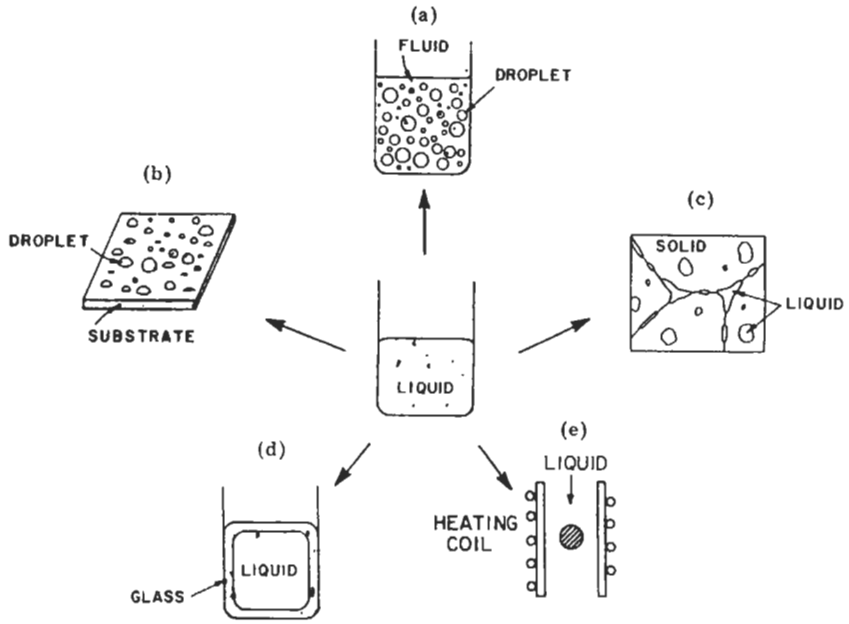


Fig. 19. A schematic illustration of the most common methods for inhibiting heterogeneous crystal nucleation and thereby obtaining highly undercooled liquid alloys: (a) droplet emulsion, (b) substrate, (c) entrained droplet (mushy state quenching), (d) fluxing, and (e) drop-tube or levitation. (After PEREPEZKO [1984]).

emulsion to be stable the droplets must be coated and this coating may itself act as a nucleant. The coating problem is not present in the fluxing technique in which a single alloy drop is surrounded by viscous flux or slag, which, in addition to isolating the drop from contact with solid container, may have a cleaning action (KUI, GREER and TURNBULL [1984]). Truly containerless solidification can be achieved for droplets in free-fall in evacuated drop-tubes (GREER [1994]) and levitated, for example electromagnetically (SCHLEIP *et al.* [1988]). A few examples of maximum undercoolings which have been achieved for various metals are shown in table 2.

When a melt is highly undercooled, solidification when it does occur is rapid and shows all of the microstructural effects outlined in § 1.1. Crystal nucleation can even be avoided completely, yielding a glass (PEREPEZKO and SMITH [1981]). The techniques for obtaining undercooled melts are of interest both for production of metastable states and for fundamental studies. The emulsion and fluxing methods, for example, can readily be used to measure not only undercooling but also the specific heat of the undercooled liquid, an important thermodynamic parameter. Of particular interest is that electromagnetic levitation of a droplet can be used for quantitative studies of solidification velocities at selected undercoolings. Studies of undercooled alloy melts have been comprehensively reviewed by HERLACH *et al.* [1993]. Among the phenomena investigated are grain refinement (GREER [1991]) solute trapping and the nucleation and growth of metastable phases.

These topics are discussed in some detail in several sections of ch. 8, § 4.2.

References: p. 1818.

Table 2
Some examples of maximum undercoolings achieved for pure metals, using various techniques. Much more extensive surveys will be found in KELTON [1991], HERLACH *et al.* [1993] and GREER [1994], from which these data are taken.

Metal	Maximum undercooling ΔT_{max} (K)	Melting point T_m (K)	Normalized undercooling $\Delta T_{max}/T_m$	Experimental method
Ni	480	1728	0.28	electromagnetic levitation
Fe	538	1809	0.30	electromagnetic levitation
Hg	80	234	0.34	emulsion
Ga	174	303	0.57	emulsion
Sn	191	505	0.38	emulsion
Re	975	3453	0.28	drop-tube
Hf	530	2500	0.21	drop-tube

3.3. Deposition by evaporation or sputtering

The rapid-quenching methods described above have been used to make not only crystalline alloys but also metallic glasses. Several glass-making methods which do not depend on melt-quenching will be outlined here. These methods may collectively be termed *pseudo-RSP*.

One group of such methods depends on the deposition of vapor on cold substrates (often held at liquid-nitrogen or even liquid-helium temperature). These methods are in fact of older ancestry than true RSP methods and were used to make various amorphous and supersaturated crystalline materials in the 1930s (Kramer [1934, 1937]) and 1950s (BUCKEL and HILSCH [1956]), generally in the form of thin films. (For more details, see CAHN's 1993 review). Sometimes thermal evaporation is used to produce the vapor (MADER [1965] and SINHA *et al.* [1976]), but cathodic sputtering is more usual; if a powerful high-rate sputtering apparatus is used, glass samples several μm thick can be made, though at very high cost (see review by DAHLGREN [1978]). These techniques are now waning, in spite of the fact that more alloy compositions and wider composition ranges can be made glassy by vapor methods than by liquid-quenching (TURNBULL [1981]). Perhaps a reduction of interest in these methods is a recognition that in metallurgy, as opposed to microcircuitry, thin films have only a modest role.

3.4. Amorphization by irradiation

A further approach is heavy irradiation (to a dose of at least one displacement per atom) of a crystalline alloy (generally an ordered intermetallic compound), using either neutrons or ions (RECHTIN *et al.* [1978], ELLIOTT and KOSS [1981]). As judged by calorimetric tests, such "artificial" glasses are structurally very similar to the same alloys vitrified by RSP. A related "solid" approach to making a glass is by *implantation* into a pure metal of ions of a different element, most usually metalloid ions. Phosphorus has been most effectively used. The bombarding ion both destroys crystalline order and, by

occupying strategic locations in the glass structure, stabilizes it. Self-ion bombardment (e.g., Ni^+ into nickel) never generates a stable glassy structure in a metal (although apparently it can do so in silicon). The technique has been reviewed by BORDERS [1979], and there is later research by RAUSCHENBACH and HOHMUTH [1982], and by RUSSELL [1985]. Very recently, a complete and up-to-date review by HOWE [1995] has been published. Here there is space only for a thumbnail sketch: — A striking group of observations is on *electron-induced* amorphization. MORI and FUJITA [1982] have established the temperature dependence of the critical dose for amorphizing NiTi with 2 MeV electrons, while MORI *et al.* [1983] have observed the surprising phenomenon of initial amorphization preferentially at dislocations, in the same alloy. The theoretical background, including a discussion of critical irradiation doses, is presented by CARTER and GRANT [1982]. It is important to recognize that glassy layers made in this way are less than one micron thick. It is now well established that the intermetallic compound must be *disordered* by the irradiation below a critical level of long-range order; it then becomes unstable towards amorphization (LUZZI *et al.* [1984, 1985, 1986]). MORI *et al.* [1984] nevertheless incline to an alternative (or additional?) criterion, requiring a large critical vacancy concentration. CAHN and JOHNSON [1986] have critically discussed the relative merits of these models. A more recent discussion of these rival models is by MOTTA and LEMAIGNAN [1992]: they assemble evidence that a critical degree of destruction of long-range order is generally a necessary but not a sufficient criterion and that a critical concentration of one type of point defect (either vacancies or interstitials) is also necessary, and they answer earlier criticisms suggesting that the required point defect concentrations are not in practice achievable.

MENDOZA-ZELIS *et al.* [1982] describe the amorphization of nickel films by phosphorus ion implantation. Electrical resistivity indicates that the glass thus prepared is structurally identical to glasses of the same compositions made by evaporation or electrodeposition. (The last two approaches to glass-making, together with the related technique of electroless plating, with a partial but extensive list of compositions prepared in these ways, are outlined by DIETZ [1977].) Ion implantation is also used to amorphize crystalline silicon; some of the consequential research is outlined in the next section. A technique closely related to ion implantation is *ion-beam mixing*, in which two discrete layers at a surface are mixed on an atomic scale by bombardment with "neutral" ions such as A^+ or Xe^+. Thus, Au–Si glasses have been made by ion-mixing of a gold-coated silicon substrate; this same glassy alloy has been made by RSP of the melt, evaporation on to a cold substrate, laser-treatment and ion-mixing. Ion-beam mixing is reviewed by FOLLSTAEDT [1991].

The methods mentioned in the last two paragraphs effectively involve treatments at or close to metallic surfaces, which is the subject of the following section.

3.5. Rapid solidification processing of surfaces

A benefit of the various rapid solidification methods that operate at surfaces is that only a very thin layer of the solid needs to be melted and the underlying solid then can act as a built-in heat sink. The usual instrument employed for this purpose is a laser,

either continuous-acting (CO_2 type), used in scanning mode, or pulsed (ruby type). The most common way of rapid quenching at surfaces involves no local change of composition, but it is possible to coat the surface with a thin layer of another substance and metastably alloy it into the melted surface; in this way, expensive substances need only be used in small quantities.

RSP at surfaces has in practice been used for two very different purposes: (1) There is a large body of research (which can only be briefly exemplified here) on the laser-treatment of semiconductor surfaces, silicon in particular, and (2) laser treatment has been used on a large scale for the hardening of surfaces of steel and some other industrial alloys.

Figure 20 shows the variables involved in the treatment of a surface by a continuous wave laser. Heat-flow calculations, for adequately large heated spots, can be simplified by neglecting sideways heat flux and treating the problem in terms of one-dimensional heat flow: calculations of this kind have been published by MEHRABIAN [1982]. He assumes that the heat which can be usefully absorbed is limited by the onset of vaporization; on this basis, increasing the laser power (which entails faster scanning over the surface) yields a shallower melt, faster refreezing and a steeper thermal gradient. Typical figures, for aluminum, are 650 μm or 6.5 μm melt depth for 5×10^8 or 10^{10} W/m^2 heat flux, respectively. The estimated solid–liquid interface velocities are 0.035 and 3.5 m/s, respectively. KEAR *et al.* [1981] estimated that, again taking vaporization as a limiting factor, quench rates up to 10^6 K/s can be achieved for a melted layer ≈ 10 μm thick. More recently, LIN and SPAEPEN [1984] showed that a pulsed laser can provide cooling rates as high as 10^{14} K/s.

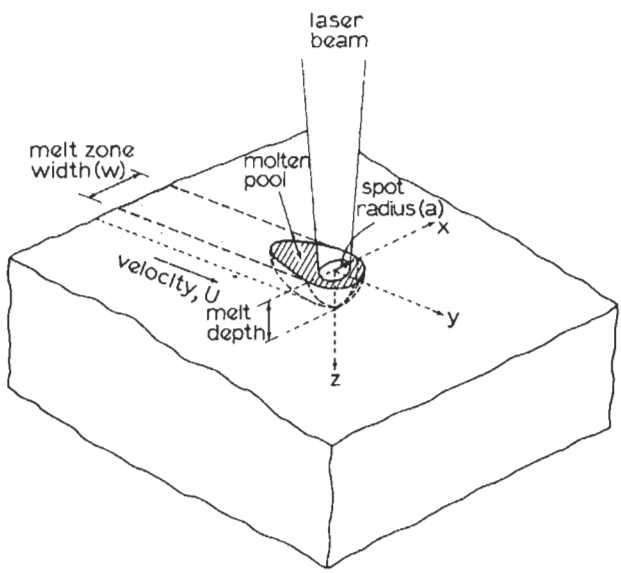


Fig. 20. Schematic illustration of laser beam–substrate geometry during rapid surface melting and refreezing by laser treatment. (After MEHRABIAN [1982]).

The use of Q-switched ruby lasers so designed as to give extremely short pulses in the nanosecond (or even picosecond) range has been highly fruitful in research on silicon. A good example of research on metastable structures in silicon is the body of work on the melting of amorphous silicon. A layer of amorphous silicon can be created at the surface of a silicon crystal slice by ion-implanting either silicon itself or another species such as indium. It was first suggested in 1978, independently by Spaepen and Turnbull and by Bagley and Chen, that such an a-silicon layer might undergo a first-order melting transition at a temperature several hundred degrees lower than the melting temperature of crystalline silicon (although a-silicon cannot have an exact thermodynamically defined melting temperature). The melting of a covalently bonded amorphous solid to give a metallically bonded liquid here is bound to be discontinuous. This was duly confirmed by laser-heating in situ of the a-silicon layer, with a ≈ 3 ns pulse: what is done here is to make transient measurements of properties, in particular, electrical conductance and optical reflectance, which vary with the state of the silicon, with a time resolution of a few nanoseconds. Indeed, the investigators found that the molten layer can then explosively crystallize: the release of latent heat makes the process self-sustaining. Fig. 21, from a review by POATE *et al.* [1987], shows transient conductance and reflectance measurements on such a specimen, and also indicates the evolution with time of the molten and subsequent crystallized layer at the expense of a-silicon. (Because of the large latent heat released, crystallization under these circumstances is "explosive"). The transient measurement approach can even be used to assess whether the crystallizing layer is nucleated at the surface or in the interior of the melt (PEERCY *et al.* [1990]); the conclusion was that nucleation is internal. Figure 22 shows the time-dependent depth of

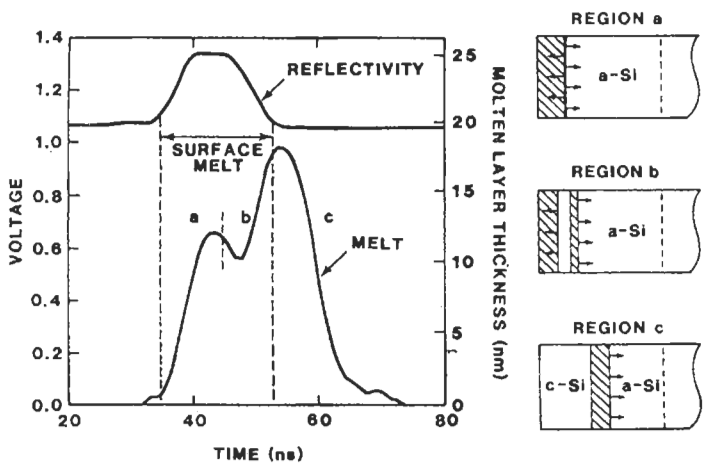


Fig. 21. Transient conductance measurements of molten layer thickness and simultaneous surface reflectance for a 0.20 J/cm^2 pulse incident on a 320 nm thick a–Si film. The schematic diagrams at the right illustrate various stages for the ("explosive") crystallization process. (After POATE *et al.* [1987]).

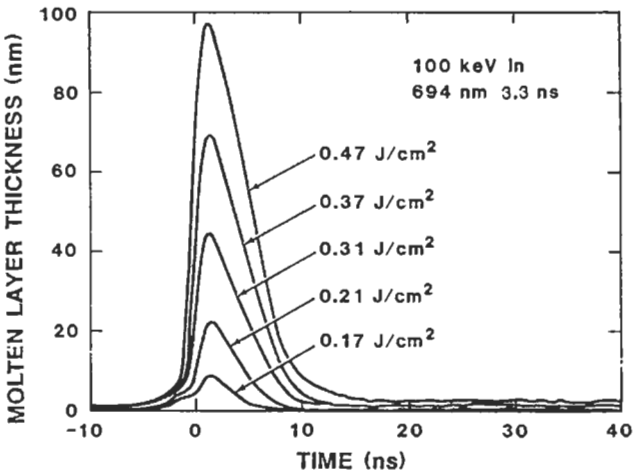


Fig. 22. Molten layer thickness as a function of elapsed time for irradiation of 100 KeV-implanted silicon. (After PEERCY *et al.* [1990]).

the molten layer deduced from reflectance measurements (because reflectance changes as melt thickness changes) for different laser intensities. It can be seen that the dimensions involved are very different from those calculated for typical experiments with metals, as indicated in the preceding paragraph.

At the other extreme, laser treatment of steels, titanium and aluminum alloys has been widely practised, with and without powder injection into the laser-melted surface pool to change local composition. The field has recently been reviewed by MORDIKE [1986, 1991] and by SCHUBERT and BERGMANN [1993]. Thus, aluminum–silicon alloys can be effectively "modified" to refine the scale of the eutectic and enhance mechanical properties thereby, without the use of the usual modifying dopants. Similar effects can be achieved with cast iron. The nature of nucleation of crystalline phases in a surface melt-pool created by laser irradiation can be deduced by examination of the resultant microstructure, and the conclusion was drawn that nucleation at the free surface, at the melt/substrate interface, and homogeneously within the melt, can all be generated by varying the conditions of irradiation. For further details, Mordike's review should be consulted.

A further approach is to make layered alloys and use laser pulses to convert these into a homogeneous metallic glass. Thus, LIN and SPAEPEN [1982] made an iron–boron glass with as little as 5 at% B (much less than is required when less extreme methods are used) by picosecond laser treatment of a finely layered Fe/Fe_3B sandwich.

3.6. Electrochemical, electroless and sonochemical deposition of amorphous phases

It was discovered long ago by Brenner at the NBS that Ni–P alloys containing more than $\approx$ 10 at% P can be made by electrodeposition, to give an ultrahard surface coating.

This industrially useful method has been steadily improved over the years (BRENNER [1963]). The most recent development has been an examination of the influence of pulsed as opposed to steady current (LASHMORE and WEINROTH [1982]), and it appears that Ni–P glass deposits made under these two regimes are structurally distinct (LASHMORE *et al.* [1982]). Little has been published about other compositions that can be made in amorphous form by electrodeposition, but there is evidence that some research has been done under conditions of industrial secrecy: apparently Co–W–B is another alloy family that can be made amorphous in this manner. Also, bright amorphous chromium (presumably stabilised by solutes) can be electrodeposited from chromic acid solution with additives (HOSHINO *et al.* [1986]; TSAI and WU [1990]).

Electroless plating has already been mentioned in passing in § 3.4, in connection with DIETZ's 1977 review. Much research has been done on this technique, which is based on electrochemical principles though no external potential difference is applied. It has very largely been applied to the deposition of Ni–P coatings applied from solutions containing hypophosphite ions, which can be amorphous or nanocrystalline and are very hard. This extensive research, much of it concerned with the control of deposition rates, is reviewed by ROOS *et al.* [1991].

A recently discovered, distinctly exotic way of making an amorphous metal is via *sonochemistry*. This term relates to the technique of injecting energy into a chemically reacting system by means of ultrasound. The key investigators here have been the chemist K. Suslick and his group. Ultrasound is conveyed to a liquid mixture of chemicals in the simple apparatus shown in fig. 23. Suslick has explained how high temperatures and consequential ultrarapid cooling arise in such an apparatus (SUSLICK [1993], [1995]).

Vapor-filled bubbles are formed in the liquid through resonant absorption of sound energy and oscillate in size at, typically, 20 kHz. When a bubble has grown so large that it is no longer resonantly coupled to the ultrasound field, it collapses at extreme speed ("cavitation") under the influence of surface tension and in the process, the vapor is heated adiabatically to a high temperature; immediately after this stage, the surrounding liquid cools the vapor as fast as it was heated. The temperatures reached and the time-scale involved have been measured, among other techniques, by means of *sonoluminescence*, the emission of light flashes during bubble collapse, and temperatures near 5000 K have been estimated (FLINT and SUSLICK [1991]). (The very existence of sonoluminescence has been doubted until, recently, BARBER and PUTTERMAN [1992] measured the time-dependent radii of sonoluminescent bubbles at the same time as the sonoluminescent output. MADDOX [1993] has discussed the implications of these striking observations). The duration of the high temperature is estimated to be of the order of 100 ns and the cooling rate thereafter is around 10^{10} K/s, much faster than any of the other techniques discussed hitherto. "Nearly pure" (> 96% Fe) amorphous iron powder, of 1–2 nm particle size, has been made by the sonochemical decomposition of iron pentacarbonyl, a liquid (SUSLICK *et al.* [1991]). The powder crystallizes at the surprisingly low temperature of ≈ 308°C. This approach does not appear to have been applied yet to nickel carbonyl.

References: p. 1818.

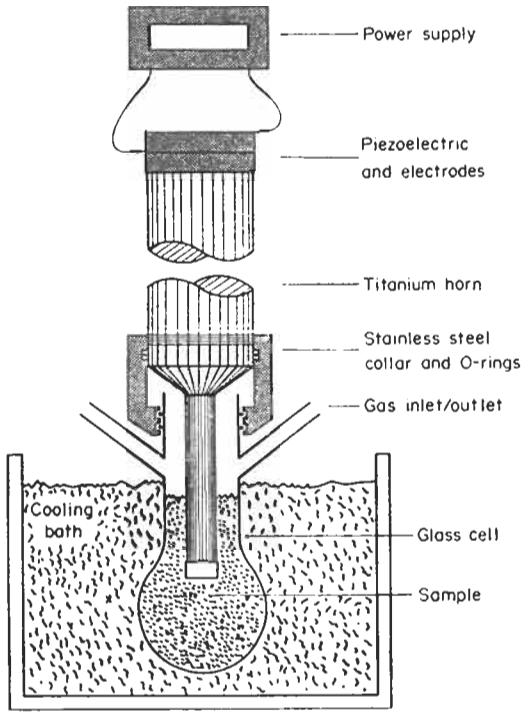


Fig. 23. A typical sonochemical apparatus with directly immersed ultrasonic horn. (After SUSLICK [1993]).

3.7. Solid-state amorphization reactions (SSAR)

Since the amorphous phase in alloys can be a well defined metastable state, it can be produced by reaction, enabled by thermal annealing. The initial "energized" state (see § 1.3) is simply an aggregate of unreacted elements. The first metallurgical demonstration of a solid-state amorphization reaction was by SCHWARZ and JOHNSON [1983], who showed that thin polycrystalline gold and lanthanum films in a multilayered stack would react at 50–100°C to give an amorphous alloy. (This research followed a slightly earlier discovery of a related phenomenon: the compound Zr_3Rh, in a metastable crystalline form, could be amorphized by reacting it with hydrogen to form a metastable amorphous hydride (YEH *et al.* [1983]; this finding has been broadened more recently by AOKI *et al.* [1991], who discovered hydrogen-induced amorphization in numerous binary intermetallic compounds in which one constituent is a hydride former).

A significant number of elemental pairs have since been found to undergo similar reactions similar to that between gold and lanthanum; these include Ni–Zr, Cu–Zr, Fe–Zr, Ni–Ti, Ni–Nb, Cu–Ti among others, as reviewed by JOHNSON [1986] and also in the first conference devoted to this phenomenon (SCHWARZ and JOHNSON [1988]). The amorphization reaction is associated with two characteristics of the elemental pairs, one thermodynamic and one kinetic: a strongly negative heat of mixing in the amorphous state, and

a marked asymmetry in diffusivity, that is, one element diffusing much faster than the other. Since diffusional asymmetry is most likely associated with a size difference, it can be seen that the characteristics needed for the amorphization reaction have much in common with those favoring glass formation generally (§ 2.3). The negative heat of mixing provides the driving force for the reaction, but it is associated with a tendency for the formation of highly cohesive intermetallic compounds. However, at the low annealing temperature of the reaction (≈ 250°C, for example, in the case of Ni–Zr) these compounds have difficulty in nucleating, and then the metastable extrapolated liquidus curves shown in bold in fig. 24 become relevant. The amorphization reaction is then interpretable as eutectic melting occurring below T_g and therefore yielding an amorphous solid rather than a liquid. In the case of Ni–Zr, nickel is a fast interstitial diffuser in crystalline zirconium, whereas zirconium is a slow substitutional diffuser in crystalline nickel. This difference in behavior, in which the Ni can diffuse several orders of magnitude faster than the Zr at the same temperature (GREER *et al.* [1990]), applies also in the amorphous phase. Diffusional asymmetry of this kind appears to apply in the other systems showing the solid-state amorphization reaction, the high mobility of one species enabling the reaction to occur, the low mobility of the other species inhibiting nucleation of the equilibrium intermetallic compounds.

Systematic experiments on Ni–Zr have established that the amorphous phase is nucleated at grain boundaries in the polycrystalline zirconium films (SCHRÖDER *et al.* [1985] and VREDENBURG *et al.* [1986] have confirmed this by a beautiful experiment in which they demonstrated the impossibility of launching an amorphization reaction in a bilayer consisting of polycrystalline nickel deposited on single-crystal zirconium.

It is mostly found that the amorphization reaction proceeds on a planar front at a diffusion-controlled rate which diminishes as the amorphous layer thickens. Beyond a critical thickness of reacted layer, that is, below a critical front velocity, a crystalline

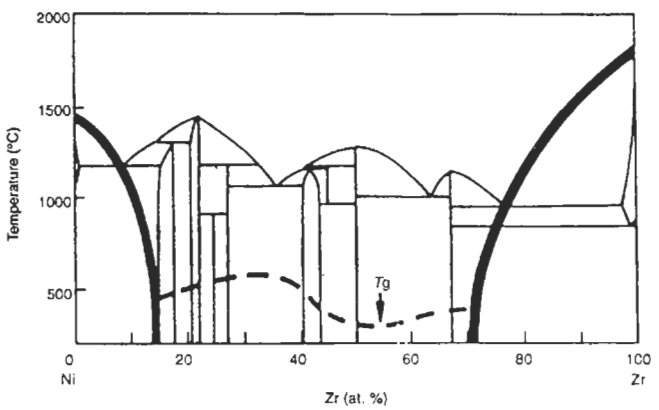


Fig. 24. Ni–Zr equilibrium diagram, a system exhibiting solid-state amorphization. The bold lines show the terminal liquidus curves and their metastable extrapolations. (After HIGHMORE and GREER [1989]).

References: p. 1818.

compound appears. Its appearance has been quantitatively interpreted by HIGHMORE [1990] in terms of transient nucleation theory. The intermetallic nuclei thus formed can grow at the expense of the previously formed amorphous layer (MENG *et al.* [1988]). Thus, if a fully amorphous product is to be obtained by the reaction, initial thicknesses of the elements in a multilayer stack must be small, typically below 100 nm. Multilayers with such thin layers can be made by thin-film deposition, but then the final product is likely restricted to a thin-film form, and this could be obtained more readily by initial co-deposition of the elements. SSAR in multilayers has been illuminatingly examined by JANKOWSKI *et al.* [1995].

The more general issue of the relation between "nucleation of disorder" in SSAR, melting and the destruction of atomic long-range order in a solid solution has been discussed by CAHN and JOHNSON [1986] and brought up to date in a stimulating thermodynamic approach by WOLF *et al.* [1990].

3.8. Amorphization by mechanical processing

When the diffusion-induced amorphization reaction was first discovered, it was soon recognized that deposition by evaporation was not the only way of making the initial multilayers: it was also possible to roll or wire-draw composites of foils or wires of the constituent metals (to reduce the scale of the microstructure and to weld the constituents) and then to anneal the product. In fact, the production of ultrafine filamentary composites by swaging, wire-drawing or rolling of mixtures of two metals, notably for use as superconductors, was pioneered in the 1970s long before the SSAR had been discovered, and reviewed by BEVK [1983]. Soon after the discovery of SSAR, ATZMON *et al.* [1985] rolled stacks of alternating 20–200 mm foils of Ni–Zr, Ni–Er and Cu–Er to very heavy reductions and then annealed them: partial conversion to amorphous phase was observed. SCHULTZ [1987] worked with similar Ni–Zr stacks: his amorphous phase had a composition different from that of the average of the stack, implying that the zirconium had become exhausted before the nickel. It appears that Schultz may have been the first to observe this kind of *mechanically aided amorphization* (SCHULTZ [1989]).

Bordeaux and Yavari conducted an extensive programme of experiments on cold-rolled multilayer stacks of systems such as Ni–Zr, Ni–Nb, Cu–Zr, Cu–Ti and Fe–Ti, in the late 1980s (see review by BORDEAUX and YAVARI [1990a]). X-ray diffraction established that a measure of amorphization took place during the rolling process itself, but annealing was required to complete it. They were able to amorphize Al–Pt stacks completely during heavy rolling (BORDEAUX and YAVARI [1990b]); this kind of process is termed *mechanically induced amorphization*, to distinguish it from mechanically *aided* amorphization They made careful kinetic measurements of amorphization rates, using XRD and DSC, and were able to show (in spite of difficulties occasioned by the fact that the as-rolled layer thicknesses had a broad distribution) that amorphization was much faster than in an interdiffusion experiment at the same temperature, in the absence of cold work. The comparison was also complicated by the fact that deformation causes local heating at the foil interfaces, which may transiently approach +100 K as measured by thin-foil thermocouples (MARTELLI *et al.* [1990]). From their kinetic studies of the rate

of thickness increase of the amorphous layers, Bordeaux and Yavari deduced that the activation energies for nickel diffusion through the layers were much lower than those previously measured for undeformed multilayer stacks. This suggested that "holes" — localised low-density regions in a metallic glass generated by deformation — had accelerated diffusion (see ch. 7, § 9.1.2). It is possible that the extreme defect concentrations generated by very heavy deformation altered the free energy balance between the various phases sufficiently to accelerate the kinetics significantly.

In parallel to these studies, ball mills were introduced as an alternative way of imparting heavy deformation and thus bringing about amorphization without the need for annealing. This was originated by YERMAKOV [1981] in Russia, and soon after taken up by SCHWARZ and KOCH [1983] in America, who prepared Ni–Ti and Ni–Nb amorphous phases by ball-milling. In this process, grinding balls, usually of specially hardened steel, are combined with a small amount of powder, either mixed metallic elements or powder of a prealloyed intermetallic compound, in a container which can be rotated or vibrated or subjected to a combined form of motion. The end-results are found to be sensitive to the particular form of grinding motion employed. Both mixed elements and prealloyed powders often behave in the same way with respect to amorphization. Details of the extensive body of research on this process of mechanically induced amorphization which has been accumulated within a few years can be found in a book chapter by KOCH [1991]. The accumulated evidence suggests that mechanically aided amorphization and mechanically induced amorphization are variants of a single process.

The criteria which govern what systems can be amorphized in this way are discussed at the end of § 2.3.

3.9. Pressure effects

In considering phase transformations, we have mostly been concerned with changes in temperature. Yet pressure is an equally important thermodynamic variable. As we have seen, metastable phases or microstructures can be achieved as a result of rapid temperature changes. Of particular interest, then, is that pressure changes (because they do not involve a conductive process) can, especially for samples of large cross-section, be much more rapid than temperature changes. There are several ways in which pressure and pressure changes can be used to attain large departures from thermodynamic equilibrium.

When pressure is varied, new phases can appear and pressure is then directly of interest in providing access to a wider range of structures. Some of these structures can, in fact, be obtained without application of pressure. For example, a metastable phase of pure bismuth obtained at high undercooling in a droplet emulsion experiment is a high-pressure phase; thermodynamically it can form because the liquid is cooled below its (metastable) melting point, and it must be kinetically preferred to the stable phase (PEREPEZKO and ANDERSON [1980]). In most cases, however, application of pressure is necessary to cause a transformation (usually in the solid state) to a high-pressure phase. The question is whether this phase will be retained when the pressure is again reduced to ambient. Temperature is commonly used to effect retention under ambient conditions. The general procedure is as follows: Pressure is applied, the temperature is raised to

permit the transformation to the high-pressure phase to take place, the temperature is reduced to lower the atomic mobility, and the pressure is then reduced with the atomic mobility sufficiently low to prevent decomposition of the high-pressure phase. In this way metastable high-pressure phases are obtained. The classic example of such a phase is diamond. Research in this area has concentrated on compounds, such as InSb, GeP, Nb_3Ge, Nb_3Te. Before the advent of oxide-based high critical temperature superconductors, high-pressure phases such as Nb_3Ge were of considerable interest for their superconducting properties. JAYARAMAN [1980] has reviewed metastable phase synthesis by this route. The metastable compounds produced can in some cases decay into amorphous solids (BATTEZZATI [1990]).

A special case of a transformation induced by pressure is melting. This of course occurs only in those systems which show a contraction in volume on melting. Of these, by far the most widely studied is ice. Ice held at 77 K and subjected to a pressure of 10 kbar undergoes amorphization (WHALLEY [1988]). Thermodynamically this result is straightforward: the high pressure takes ice to its melting point despite the low temperature. Although there is some superheating, the ice does melt, albeit to an amorphous phase rather than liquid water, because 77 K is below T_g for water. A similar technique could find application in some semimetal systems showing similar volume changes on melting.

Just the converse effect has been exploited in a wide range of alloys showing the more usual behavior of a volume expansion on melting. For these materials application of pressure *increases* the melting temperature. Thus *rapid pressure application* to the liquid alloy can induce rapid solidification. After the solidification is complete, the temperature is reduced to ambient, not necessarily rapidly. The crystalline products can show all of the features of materials rapidly solidified by temperature reduction, such as refined microstructure, extended solid solubility and metastable crystalline phases. An example of work in this area is that by SEKHAR and RAJASEKHARAN [1986], and the field has been comprehensively reviewed by REDDY and SEKHAR [1989].

The rapid pressure application technique does not lend itself to forming amorphous phases. The application of pressure increases the thermodynamic driving force for crystallization of the liquid without as large a reduction in atomic mobility as would be brought about by the corresponding temperature change. Nevertheless, application of high pressure should decrease atomic mobility and this has been successfully exploited. For example, Cu–Sn alloys held under pressure can be heated by electric resistance, and then quenched by turning off the electrical current. When this is done, Cu–(12–17 at.% Sn) glassy phases can be made, which is not possible under ambient atmospheric pressure (BATTEZZATI [1990]). PONYATOVSKY and BARKALOV [1991] have developed this process, which they call *thermobaric quenching*, and have made glasses in the Cd–Sb, Zn–Sb and Al–Ge systems.

Apart from the pressure effects outlined above, there is a further series of recent observations referring to silicon. Silicon undergoes a transformation to a metastable amorphous structure when subjected to an indentation with a diamond pyramid (e.g., PHARR *et al.* [1992]), a phenomenon linked with the metal/non-metal transition treated by Mott many years ago. The amorphous phase remains metastable after the indenter is

removed. The complicated pedigree of this observation has been explained by CAHN [1992]. Distinct from this stress-induced crystal-to-amorphous transition which is observed at ambient temperature is another, discovered by PIROUZ *et al.* [1990] who found that indentation of silicon in the temperature range 400–650°C leads to a metastable, martensitic transformation from diamond-cubic to a hexagonal crystal structure; germanium behaves similarly.

4. *Metallic glasses: structure and properties*

4.1. Structure

In a single-phase solid alloy the distinctive feature of the glassy state is that it has, in essence, no microstructure. There is no crystal lattice in which to have defects such as grain boundaries or dislocations. But the glass does have a structure in the way its atoms are arranged, and this structure is important in determining its properties. Relaxation (§ 4.2) of the glassy structure on annealing can, for example, lead to increases in viscosity by a factor of up to 10^5. Structure and properties can also show a natural or induced anisotropy. This section reviews the great deal of work which has been done to determine and describe the structures of metallic glasses. Although it is not possible to obtain full information as for a crystal structure, the main features of glass structure have now been established.

The problem of determining the structure of a crystal consists merely of identifying the coordinates of all the atoms in the unit cell; though the task is very difficult for large and complex unit cells, yet in principle (if the "phase problem" can be solved) the structure can be determined precisely. Conversely, for a glass, the structure can described only on a statistical basis: there are no unit cells and the environments of different, chemically identical atoms will necessarily vary. Herein lies the difficulty of determining and describing the structure of any glass (or liquid), and the intellectual challenge no doubt accounts for the attention devoted to the problem. Almost all the analysis of glass structure has been based on *pair distribution functions* (PDFs), because they can be determined from scattering experiments. PDFs give a measure of the probability of finding an atom centre at a distance r from an average central atom; they are often given in a reduced form $G(r)=4\pi r[\rho(r)-\rho_0]$, where $\rho(r)$ is the number of atoms per unit volume at a distance r and ρ_0 is the number of atoms per unit volume in the sample as a whole. The PDF gives statistical information about distances only, and this is insufficient to specify the structure. The angular distribution of interatomic vectors can readily be extracted from structural models, but cannot be determined from conventional scattering experiments.

The determination of glassy structures is based on the comparison of measured PDFs with those calculated from trial structural models. PDFs can be determined with good precision (especially with the advent of synchrotron-based high-intensity sources of X-rays or neutrons (EGAMI [1994])) and can discriminate between models. When an acceptable match is obtained, there is still the problem that the solution is not unique; this is a more severe problem for glasses than for the analogous determination of crystal

structures, because of the angular information missing from PDFs. The discussion which follows is based on PDFs determinable in diffraction experiments, but at the end of the section less conventional methods for structural analysis will be introduced; some of these methods are capable of yielding angular information missing from PDFs.

The PDF of a glass is determined by Fourier inversion of scattering data, for example from normalized X-ray scattering curves of the kind shown in fig. 25. For an alloy glass one set of data is sufficient to determine a *radial distribution function* (RDF) in which the different types of atoms are not distinguished. Alternatively, a *partial pair distribution function* (PPDF) can be determined for each type of atom pair. For a binary alloy there are three types of pairs, and to determine the three PPDFs it is necessary to perform at least three scattering experiments in which the two types of atom have differing relative scattering powers. An example of the type of analysis which can be achieved is given in fig. 26 showing the Ni–Ni, B–B and Ni–B PPDFs in a $Ni_{81}B_{19}$ glass. For a ternary alloy there would be six PPDFs, and an analysis of this, or greater, complexity has not yet been attempted. There are various ways of achieving the differing relative scattering powers for the determination of PPDFs. These are:

(a) Different radiations: X-rays, neutrons and electrons are all possible, but the poor precision available with electrons is a problem.

(b) Anomalous dispersion: For X-rays the effective atomic scattering factors can be altered by using a wavelength close to an absorption edge.

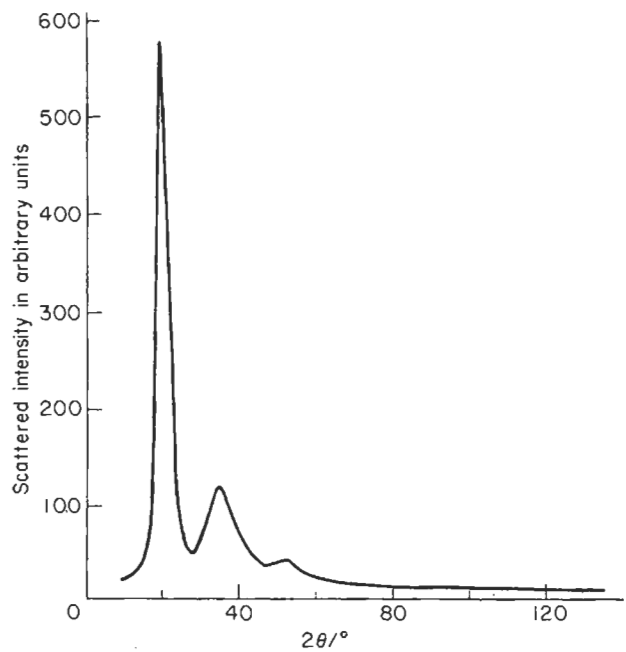


Fig. 25. Scattered x-ray intensity versus scattering angle, θ, for amorphous $Ni_{81}B_{19}$. (After CARGILL, private communication, 1979).

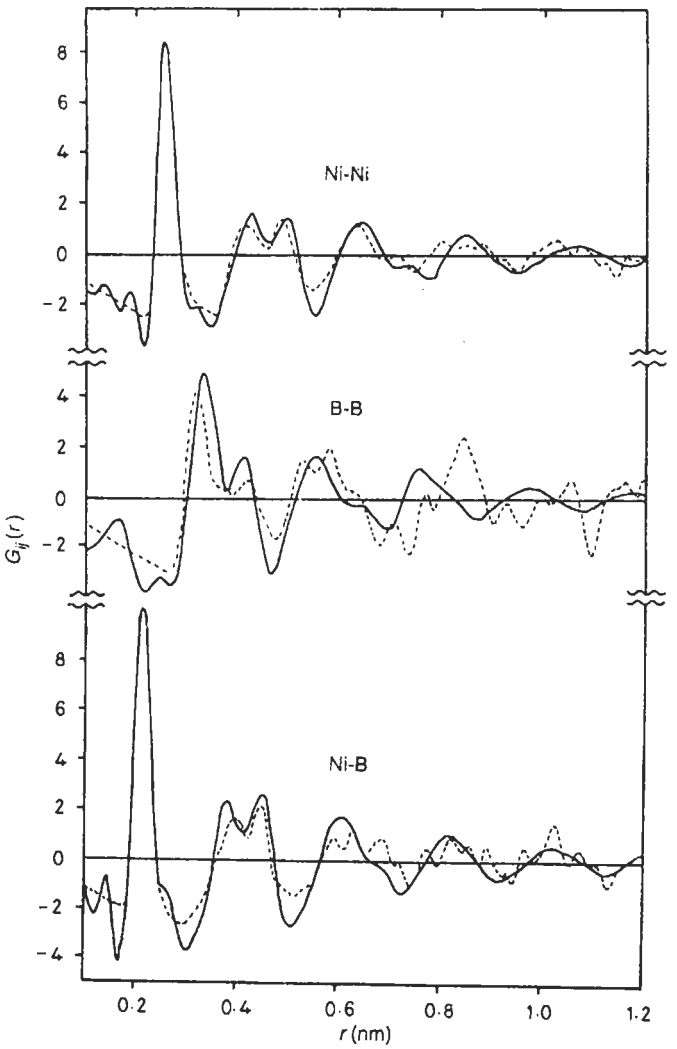


Fig. 26. Partial pair distribution functions for amorphous $Ni_{81}B_{19}$, as measured (full curves) by LAMPARTER *et al.* [1982] and as calculated (broken curves) by DUBOIS *et al.* [1985] from the 'chemical twinning' model.

(c) Polarized neutrons: The scattering from ferromagnetic samples is different for different planes of polarization.

(d) Isotopic substitution: Isotopic substitution leaves the structure of the glass unaffected, but may strongly alter neutron scattering lengths. The PPDFs in fig. 26 were determined in this way. Three $Ni_{81}B_{19}$ samples were prepared, one with natural nickel, one with ^{62}Ni and one with a mixture of ^{60}Ni and ^{62}Ni to give an effective scattering length of zero. In this last case there was no scattering by nickel, and Fourier inversion of the neutron scattering data yielded directly the B–B PPDF.

References: p. 1818.

(e) Isomorphous substitution: It may be possible to substitute one element for another of similar type so that the structure is likely the same. X-ray scattering can then be used with the different atomic scattering factor.

The experimental and computational procedures for determination of RDFs and PPDFs have been described by WAGNER [1983], EGAMI [1981] and SUZUKI [1983]. Assessment of the accuracy of pair distribution functions (obtaining a figure of merit analogous to the *R* (Rietveld) values currently used to assess the quality of a crystal structure determination) has been considered by LIVESEY and GASKELL [1982].

The basis for structure determination is the comparison between measured and calculated PDFs. Yet even without such comparison the measured PDFs reveal basic features of the glassy structure. The first peak in the radial distribution function indicates the nearest neighbor separation and the area under the peak gives the average co-ordination number. The separation is as expected for a condensed phase, and the coordination numbers are typically in the range 11.5 to 14.5, indicating that metallic glasses are densely packed, as would be indicated by an extrapolation of liquid properties. When PPDFs are examined it is found that the distribution of the different elements in a glass can be far from random. This was first established by SADOC and DIXMIER [1976] who found that in a $Co_{81}P_{19}$ glass the relevant PPDF showed that P–P nearest neighbor pairs were completely avoided. This avoidance of metalloid neighbors seems to be a general feature of transition-metal–metalloid alloys (types (i) and (ii) in § 2.2) in their normal composition ranges. When the metalloid content is increased, as for example in $Ni_{64}B_{36}$ studied by COWLAM *et al.* [1984], metalloid contact becomes inevitable, but is still much less prevalent than in a random alloy. Table 3 shows evidence for such chemical ordering in a variety of glasses. The derivation of a suitable order parameter has been considered by SPAEPEN and CARGILL [1985].

For the metal–metalloid glasses showing strong chemical ordering, more complete data from measured PPDFs are given in table 4. The coordination numbers, of metal atoms around a central metalloid, 8 to 9, and around a central metal atom, 10 to 12, are both well defined. Together with the rather small spread in interatomic distances, a limited variability of local atomic environments is suggested. Further progress on elucidating glass structure requires comparison with structural models.

The modelling must use as much incidental information as can be gleaned about the sizes, bonding and coordination of the constituent atoms: observations of density, atomic

Table 3
Observed and hypothetical coordination numbers in some metallic glasses.

Alloy, A–B	A neighbors of B (actual %)	A neighbors of B if there were no ordering (%)
$Fe_{80}B_{20}$	100	80
$Ni_{81}B_{19}$	100	81
$Ni_{64}B_{36}$	88.8	64
$Cu_{57}Zr_{43}$	56.2	57
$Zr_{65}Ni_{35}$	70.1	65
$Zr_{57}Be_{43}$	70	57

Table 4
Atomic coordination in metal-metalloid metallic glasses.

Alloy	Atom pair	$\bar{r}$ (nm)	σ (nm)	Z
$Co_{81}P_{19}$	P–Co	0.232	0.0105	8.9 ± 0.6
	Co–Co	0.254	0.0155	10.0 ± 0.4
$Fe_{80}B_{20}$	B–Fe	0.214	0.0096	8.6
	Fe–Fe	0.257	0.0165	12.4
$Ni_{81}B_{19}$	B–Ni	0.211	0.0142	8.9
	Ni–Ni	0.252	0.0147	10.5
$Fe_{75}P_{25}$	P–Fe	0.238	0.018	8.1
	Fe–Fe	0.261	——	10.7
$Pd_{84}Si_{16}$	Si–Pd	0.240	0.0106	9.0 ± 0.9
	Pd–Pd	0.276	0.0154	11.0 ± 0.7

The mean interatomic spacing, $\bar{r}$, the standard deviation, σ, and the coordination number, from GASKELL [1985].

radii in crystalline forms, interaction potentials (Lennard–Jones, Morse, etc.), interbond angles where appropriate are all taken into account in constructing a model. The actual construction was initially done physically, with balls and springs, or just steel balls, but now is always done by computer simulation. Models for covalently bonded amorphous solids such as silicon involve distinct problems from metallic glass models (only the former incorporate directed atomic bonds) and we discuss only the latter here. There are many difficulties in comparing measured and computed RDFs. The experimental determinations are limited by the range of scattering vector lengths which is accessible, and the modelling is limited by the finite size of the model; the problems and best techniques are comprehensively discussed by FINNEY [1983].

Probably the first structural model suggested for metallic glasses was that based on microcrystallites, i.e., the suggestion that the "glass" is simply a polycrystalline solid in which the grain size is very small. Early work by CARGILL [1970] showed clearly that the microcrystallite model does not provide good quantitative fits to measured RDFs of glasses such as $Ni_{76}P_{24}$. Recently the interest in nanocrystalline materials has led to a re-examination of this question. There are examples of apparently glassy alloys for which the structure can be successfully modelled assuming nm-sized grains of an icosahedral phase. A key observation, however, is that such materials do not transform in the way expected for true glasses; on annealing they transform to a polycrystalline aggregate of larger grains by a process of continuous grain growth rather than by nucleation and growth (CHEN and SPAEPEN [1991]). For true metallic glasses, which form the great majority of cases, structural modelling has mostly been based on either the *dense random packing* (DRP) model of BERNAL [1964], or the *local-coordination model* of GASKELL [1979].

Bernal's empirical model originally referred to the DRP of hard spheres all of the same diameter and was based on physical measurements with steel balls kneaded in a rubber bag. The packing is statistically reproducible, and the structure can be considered to be made up of only five simple polyhedra (canonical holes), shown in fig. 27, the

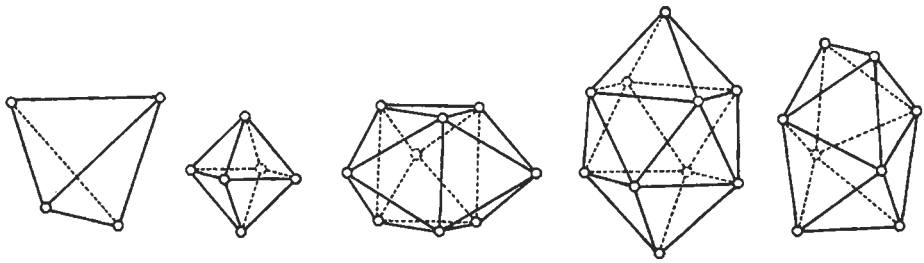


Fig. 27. Polyhedra formed by dense random packing of hard spheres, according to Bernal.

smaller holes (tetrahedra and octahedra) being in the great majority. The topology of DRP is related to that of the continuous random network first proposed for tetrahedrally bonded oxide glasses by ZACHARIASEN [1932] (CONNELL [1975], CHAUDHARI *et al.* [1976]). The average coordination number in the DRP of hard spheres is well defined and is 12.3 (SCOTT and MADER [1964]). Models based on the DRP of hard spheres give fair agreement between computed and measured RDFs. Two deficiencies in the model can be identified: the spheres in the modelling are infinitely hard, and only one type of atom is included. With realistic interatomic potentials, sequential computer relaxation of *soft* spheres can be used to increase the packing density of the model in better agreement with measured densities. In this way the fit between computed and measured RDFs becomes rather good in some cases, notably for vapor-deposited nominally pure metals (VON HEIMENDAHL [1975]). The extension of this type of work to two-component glasses was undertaken by BOUDREAUX and GREGOR [1977]. For glasses of the metal-metalloid type they constructed a DRP with a mixture of metal and metalloid atoms and imposed a rule that there should be no nearest-neighbor metalloid pairs. While a model of this type gives a much better fit to experimental RDFs, there are still significant problems, particularly in matching PPDFs.

The modelling of metal-metalloid glasses is now based on the local coordination model of GASKELL [1979]. The coordination of the metalloid typically by nine nearest-neighbor metal atoms in glasses of this type (table 3) is the same as in many crystalline metal-metalloid compounds. In these compounds the basic coordination polyhedron is a trigonal prism (giving six nearest neighbors, with three more slightly further away). The prisms are linked in one of two ways (fig. 28). GASKELL [1979] proposes that the glass structure is a non-periodic array of such prisms. Computer models assembled in this way and then relaxed under appropriate atomic potentials give rather good fits to measured distribution functions. The second nearest neighbor metal–metalloid distance in measured PPDFs can correspond distinctly to one or the other of the distances expected from the types of linkage in fig. 28. Thus there is evidence for medium-range order in the glasses in which the packing of the prismatic units resembles that in particular crystal structures. More evidence for this is provided by the structure in the B–B PPDF in fig. 26 out to quite large distances; this structure cannot be reproduced by a model based on randomly packed trigonal prisms. The prisms themselves appear to arise not from any directed bonding, but from a compromise between the tendency to the closest possible packing and the need for

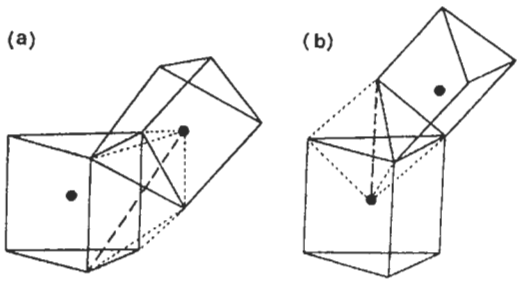


Fig. 28. The packing of trigonal prisms as found in (a) Fe_3C and (b) Fe_3P. Each prism has a metalloid atom at its center. The long-dashed line indicates the second-nearest metalloid–metal distance in each case. (After GASKELL [1983]).

space to insert metalloid atoms between the preponderant metal atoms (GASKELL [1985]).

The local coordination model can be adapted for cases other than metal-metalloid glasses around the 20 at% metalloid composition. MEHRA *et al.* [1983] have applied it to $(Ru_{84}Zr_{16})_{1-x}B_x$ glasses with $0.40 < x < 0.53$ — i.e., an unusually high metalloid content. They show that no DRP model can give even moderate agreement with the scattering data from such glasses, while a form of irregular prismatic packing based on the Ru_2B crystal structure gives good, though not perfect agreement of RDFs.

The local coordination model has been developed into the *chemical twinning model* by DUBOIS *et al.* [1985a], who point out that trigonal prismatic coordination is common not only in metal-metalloid compounds but also in many metal–metal compounds in which the minor component has a radius 65 to 85% that of the major component. The model then can be applied to a wide range of glass types. In crystalline compounds trigonal prisms can be considered to result from chemical twinning; for example the cementite (Fe_3C) structure is obtained in this way from the hcp lattice. The smaller atoms of the minor component are on the twinning planes, and the frequency of these planes is adjusted to match the composition. In modelling glassy structures, regular twinning is permitted only over a short correlation length of the order of 1 to 2 nm. The type of twinning is chosen to match the counterpart crystalline compound. As shown by the dotted curves in fig. 26, the fits to PPDFS obtained in this way can be quite good. The chemical twinning model has the distinction of being perhaps the only structural model for metallic glasses which has led to the discovery of new glass-forming compositions. According to the model, glass-forming ability should be greatest when there are competing types of twinning in the alloy. This idea has been used in developing glass-forming compositions in aluminum-based alloys (DUBOIS *et al.* [1985b]).

As exemplified in fig. 7, most glass-forming alloy systems are able to form glasses over a range of compositions. In many such systems, physical properties vary substantially with composition. Structural models of the kind discussed above give some reason to expect that there could be clear discontinuities in some properties and fig. 29 gives some measured examples.

It has been emphasized that pair distribution functions are limited in the information

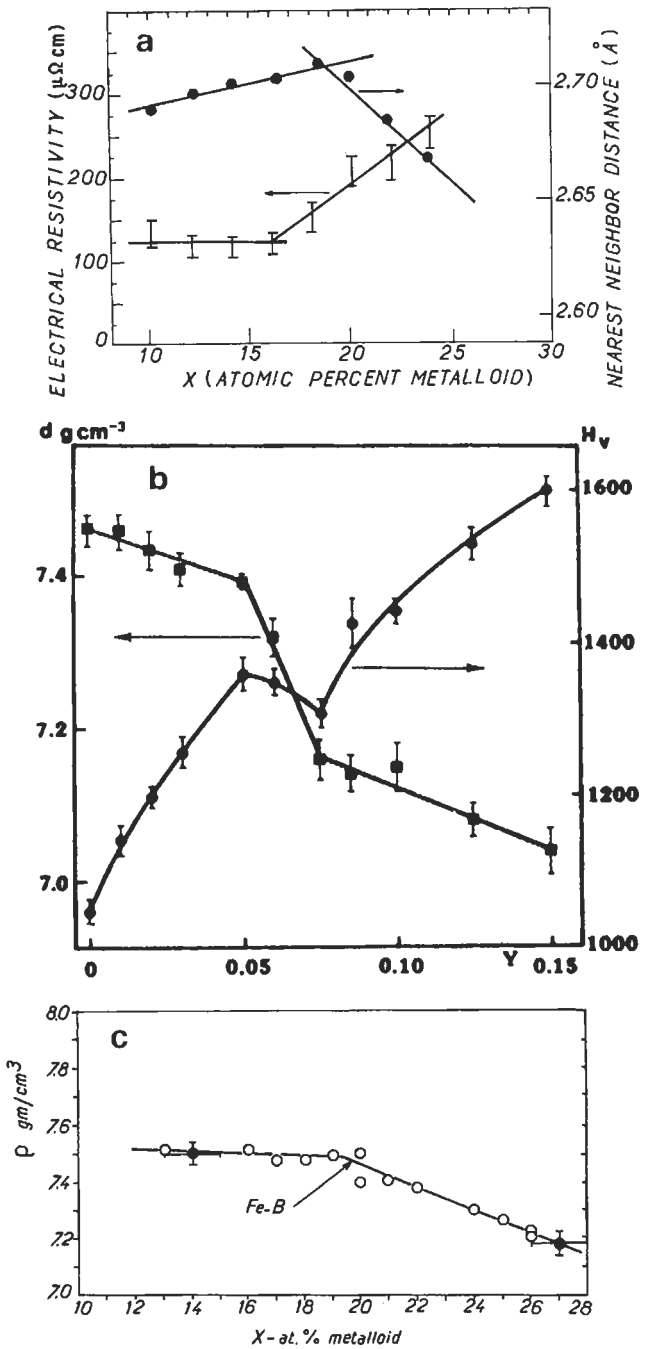


Fig. 29. Instances of discontinuities in the variation of properties with composition in metallic glasses: (a) electrical resistivity and nearest-neighbor distance in $(Mo_{0.6}Ru_{0.4})_{1-x}B_x$ glasses (after JOHNSON [1982]); (b), density, d, and Vickers microhardness, H_v, in $Fe_{0.85-y}Si_yB_{0.15}$ glasses (after AL BIJAT *et al.* [1982]; (c) densities of Fe–B glasses (after RAY *et al.* [1977]).

which they provide. For this reason there has been much interest in more local probes of atomic environments. One such technique is EXAFS in which the fine structure of an X-ray absorption edge is analysed to obtain information on the local environment of absorbing atoms. Unfortunately a structural model is still required and the details of the model are important (so that the influence of atom positions and atom identities is hard to separate). Nonetheless, useful results have been obtained; generally they confirm the conclusion from PPDF studies that the nearest neighbor environment in a glass is well defined and similar to that in the corresponding crystal (WONG [1981], GURMAN [1985]). We turn now to techniques which can provide some information on the angular distribution of nearest neighbors. Nuclear magnetic resonance (NMR) and Mössbauer spectroscopy, through the interaction of the nuclear quadrupole moment of the probe atom with the local field gradient, can provide information on the symmetry of the local environment. The extraction of further information is unfortunately very model-dependent. NMR experiments in general show local symmetry similar to that found in the corresponding crystalline phases. In Ni–B glasses, for example, PANNISOD *et al.* [1983] have shown that the variation of the environments around ^{11}B nuclei is modelled well by a combination of the environments in crystalline Ni_3B and Ni_4B_3. Each of these phases has trigonal prismatic coordination of the boron. Specifically the environment of the boron in the glasses is not similar to that in crystalline Ni_2B (which does not have trigonal prisms). Mössbauer spectroscopy is a less discriminating test of structural models, but again there is evidence for crystal-like structural units (WHITTLE *et al.* [1985]) and for local chemical order (SCHOLTE *et al.* [1985]).

Field-ion microscopy (FIM) permits the imaging of individual atoms on the surface of a needle tip, and in combination with an atom probe the identity of the atoms can be established. By field evaporation, atomic layers can be stripped from the tip, and in this way a catalogue of atomic dimensions in three dimensions can be built up. Even though this has not been fully achieved for a metallic glass, there have been significant results. For example, PILLER and HAASEN [1982] have detected that a $Fe_{40}Ni_{40}B_{20}$ glass is inhomogeneous on a scale of 2 to 4 nm. One phase is boron-poor, the other contains ~ 25% B, and Piller and Haasen interpret the structure as arising from a competition between DRP and defined local coordination. DE HOSSON [1983] has determined planar RDFs from FIM images of the palladium atoms in $Pd_{77}Cu_5Si_{18}$glass, and has shown that they are consistent with RDFs obtained from scattering experiments. Of greater importance is that he has determined the angular distribution of nearest-neighbor interatomic vectors (unobtainable from scattering experiments). This distribution has peaks at 60°, 90°, 120° and 150°. The strong peak at 60° is evidence for tetrahedral packing, but the other peaks suggest that other polyhedra may be present.

High-resolution transmission electron microscopy (HREM) is not a useful tool for studying the structure of amorphous materials at the nearest-neighbor level; the resolution is barely sufficient and there is a severe problem with superposition even in thin foils. On the other hand, HREM can be a useful probe of medium-range structure. GASKELL and SMITH [1980] have shown that in a Pd–Si glass there are domains about 2 nm across in which atomic positions appear to be correlated. This is consistent with the correlations observed in PDFs and may lend support to the chemical twinning model discussed above.

References: p. 1818.

4.2. Relaxation

When a metallic glass is annealed in the vicinity of its glass transition temperature, many physical and mechanical properties change, in pursuit of *metastable equilibrium*.. (The equilibrium must be metastable because the equilibrium state in this temperature range is crystalline). All such changes entail atomic *diffusion,* which has been extensively studied in metallic glasses. The phenomenology and mechanism of diffusion in metallic glasses is fully treated in ch. 7, § 9, and will not be repeated here. However, that treatment omits reference to one striking characteristic of diffusion in metallic glasses, i.e., the diminution in diffusivity when a metallic glass is annealed: in effect, the process of diffusion tends to choke off that process itself. This phenomenon is analogous to the gradual drop in diffusivity when a crystalline metal has been quenched from a high temperature so that excess vacancies are retained, and these vacancies gradually anneal out (§ 5.1, below).

Figure 30 shows an instance of an anneal-induced fall in diffusivity, in this instance, of boron in a nickel–niobium glass (KIJEK *et al.* [1986]). Very precise measurements of x-ray absorption (CHASON and MIZOGUCHI [1987]) have shown (fig. 31) that the fall in diffusivity in such a glass is correlated with an increase in density. i.e., with a reduction in free volume, which is the "carrier" of diffusion. The drop in diffusivity saturates after

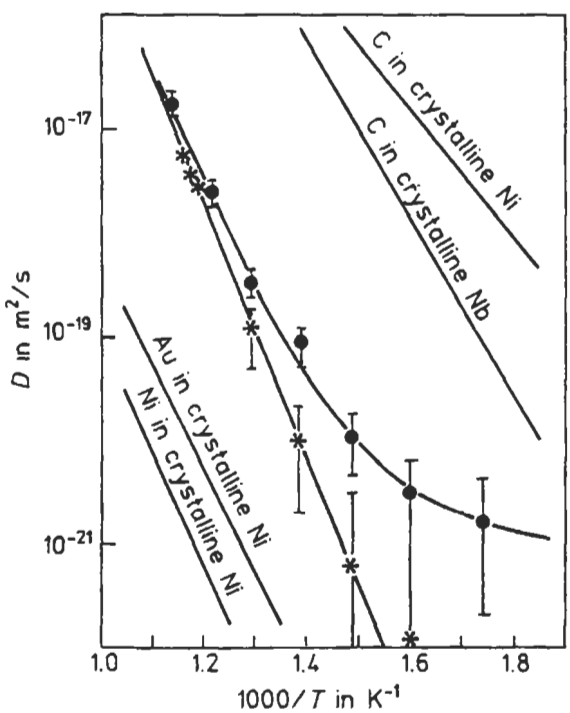


Fig. 30. Diffusivity of ^{11}B in $Ni_{60}Nb_{40}$ glass. O unrelaxed; * relaxed 420 s at 878 K. (After KIJEK *et al.* [1986]).

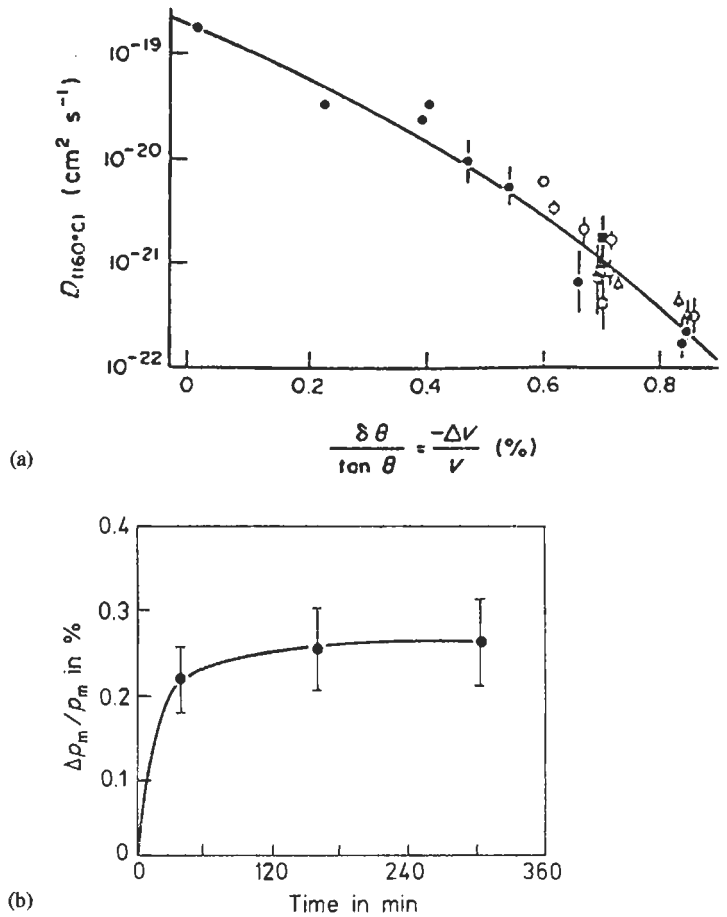


Fig. 31. Diffusion in a metallic glass as affected by a relaxation anneal. (a) Interdiffusivity, $\tilde{D}$, in a Fe–Ti amorphous multilayer at 160°C, plotted against change in atomic volume, $\Delta V/V$. The solid line is a fit to the free-volume theory of diffusion. (After CHASON and MIZOGUCHI [1987]. (b) Relative change in the density of a $Pd_{82}Si_{18}$ glass as a function of annealing time at 260°C, determined from changes in x-ray absorption. (After CHASON *et al.* [1985]).

a time if the annealing temperature is not too low: this indicates that the free volume tends towards a metastable equilibrium level characteristic of the annealing temperature. This aspect of diffusion is further discussed by CAHN [1990].

The change in properties that accompanies an anneal of a metallic glass in the neighborhood of T_g is generically called *relaxation*. This is by no means specific to metallic glasses: for instance, the refractive index of an oxide glass, of the kind used for lenses, changes steadily as it is annealed close to its glass transition.

Properties which have been pressed into service to examine the consequences of structural relaxation include, in particular, density or length changes, elastic modulus,

Table 5
Structural relaxation in amorphous alloys: properties affected and types of relaxation.

Property affected by relaxation	Type of relaxation[a]					Selected references
	(i)	(ii)	(iii)	(iv)	(v)	
Density	I		Y			GERLING *et al.* [1988], TOULOUI *et al.* [1985]
Length changes	D		Y			HUIZER and VAN DER BEUKEL [1987]
Thermal expansion coefficient	D					KOMATSU *et al.* [1983]
Enthalpy	D		Y		Y	WOLDT [1988], INOUE *et al.* [1983]
Young's modulus	I		Y	Y	Y	KOEBRUGGE and VAN DER BEUKEL [1988], CHEN *et al.* [1983]
Hardness/yield stress	I				Y	BRESSON *et al.* [1982], DENG and ARGON [1986]
Viscosity	I	Y		Y	Y	VOLKERT and SPAEPEN [1989]
Relaxation time for stress relief	I	Y		Y	Y	WILLIAMS and EGAMI [1976]
Diffusivity	D				Y	AKHTAR and MISRA [1986]
Curie temperature	I, D	Y	Y	Y	Y	GREER and LEAKE [1979], EGAMI [1978a]
Electrical resistivity	I, D		Y			KELTON and SPAEPEN [1984], BALANZAT *et al.* [1985]
Temperature coefficient of resistivity	D, I				Y	KELTON and SPAEPEN [1984], BALANZAT *et al.* [1985]
Internal friction	D					POSGAY *et al.* [1985]
Strain at fracture	D		Y		Y	INOUE *et al.* [1983], MULDER *et al.* [1983]
Saturation magnetization	D				Y	NISHI *et al.* [1986]
Magnetic after-effect	D		Y	Y	Y	GUO *et al.* [1986]
Induced magnetic anisotropy			Y	Y		GUO *et al.* [1986]
Structure factor (X-ray, neutron)	Y				Y	CACIUFFO *et al.* [1989]
Extended X-ray absorption fine structure	Y					BAXTER [1986]
Positron lifetime	D					DE VRIES *et al.* [1988]
Superconducting transition temperature	D		Y			JOHNSON [1981]
Mössbauer spectra	Y	Y			Y	STRÖM-OLSEN *et al.* [1988]
Corrosion resistance	I					MASUMOTO *et al.* [1986]
Density of states at Fermi level	I, D				Y	ZOUGMORÉ *et al.* [1988]
Debye temperature	I				Y	ZOUGMORÉ *et al.* [1988]
Magnetostriction coefficient	I				Y	CUNAT *et al.* [1988], TARNÓCZI *et al.* [1988]

[a] The types of relaxation are: (i) monotonic relaxation; (ii) reversible relaxation near the glass transition temperature T_g; (iii) reversible relaxation well below T_g; (iv) memory effects; and (v) effects of production conditions. I and D indicate an increase or decrease, respectively, of the proper value in monotonic relaxation; Y indicates occurrence of that type of relaxation.

viscosity, diffusivity (as we have seen), Curie temperature (for ferromagnetic glasses), resistivity, and a number of others as shown in table 5.

One can make a distinction between *monotonic* relaxation and wholly or partially *reversible* relaxation. Fig. 32 shows some examples of observed relaxation kinetics. In this figure, (a), showing change of density deduced from a reduction in length at 623 K, is a good instance of monotonic relaxation: if the annealing temperature is raised above 623 K, the length does not again increase. The elastic (Young's) modulus, however,

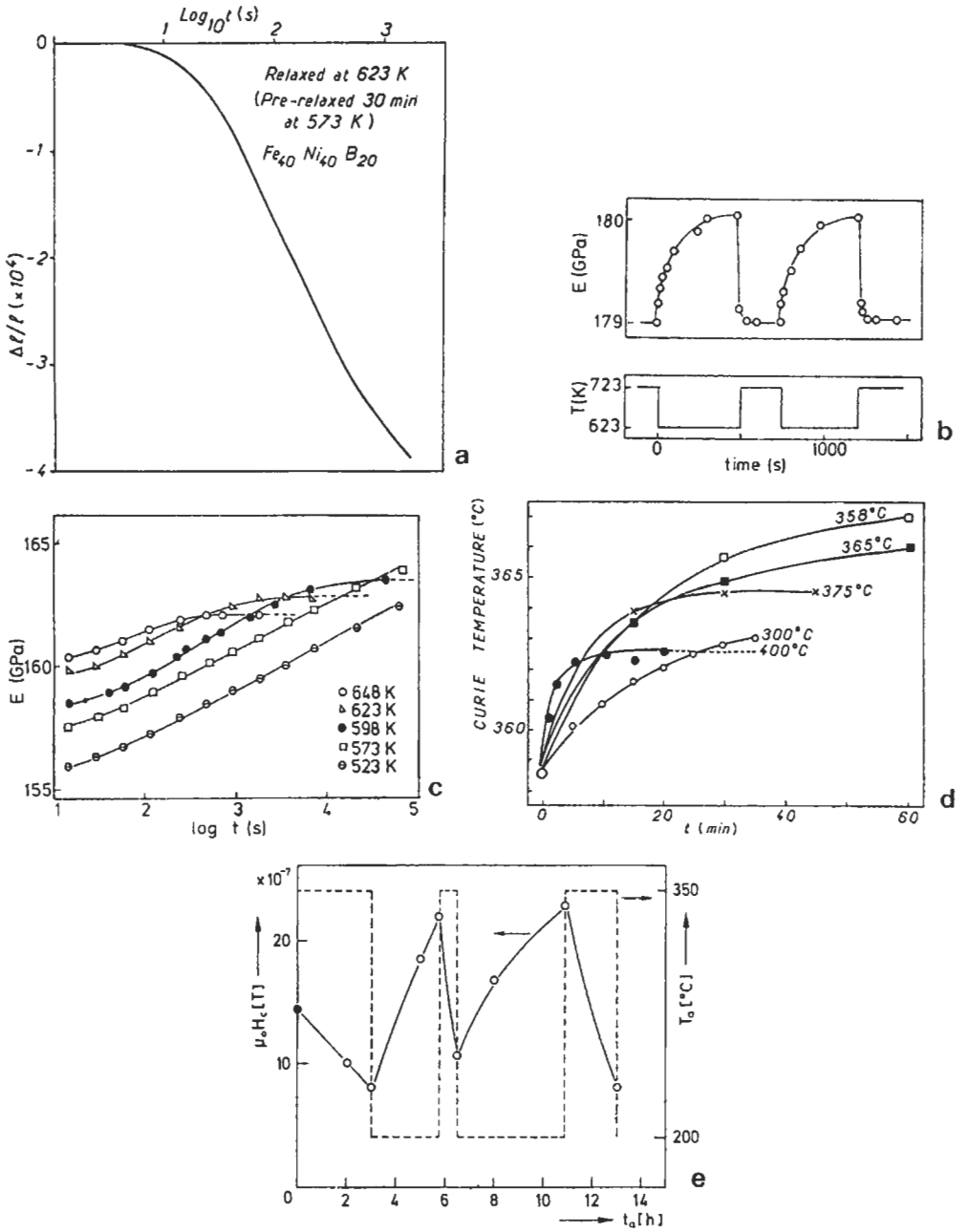


Fig. 32. Manifestations of relaxation of metallic glasses: (a) Irreversible length change of a $Fe_{40}Ni_{40}B_{20}$ glass; (b) short-term reversible changes in Young's modulus of $Co_{58}Fe_5Ni_{10}B_{16}Si_{11}$ glass cycled between 623 and 723 K; (c) irreversible changes in Young's modulus in the same glass at various temperatures, starting from the as-quenched state; (d) changes in the Curie temperature of $Fe_{80}B_{20}$ glass as a function of holding time at various temperatures above the Curie temperature ... substantially reversible; (e) reversible changes of coercive field of the above-mentioned Co-rich glass, a composition of low magnetostriction, during cyclic annealing.

References: p. 1818.

shows both monotonic behavior (graph (c)) and reversible behavior once the monotonic process has gone to completion (graph (b)). The ferromagnetic Curie temperature shows this same double behavior, though in fig. 32 there is only space to show the monotonic form. These two main types of relaxation are further subdivided in the table, according in particular to the temperature range in question: the meaning of the columns is indicated at the bottom of the table. It is striking that (column (iii)) even though the kinetics of structural rearrangement well below T_g are so slow that monotonic relaxation cannot go to completion (i.e., to metastable equilibrium), yet reversible cycles of the type seen in fig. 32b are possible in this low temperature range. EGAMI [1978] first observed this with respect to the Curie temperature; he found cycles just like those in fig. 32b, cycling between 250° and 300°C. Egami proposed that there are two distinct types of atomic process during structural relaxation: *topological short-range ordering* (TSRO) involves a rearrangement of atomic sites and a reduction of free volume (increase of density); it is this process that is associated with normal relaxation near T_g. The other process, *compositional (or chemical) short-range ordering* (CSRO), is a statistical rearrangement of chemically distinct species on fixed sites, and can occur at lower temperatures. This classification has been widely accepted as providing the best interpretation of the many observations of relaxation in metallic glasses. EGAMI [1986] published a critical examination of his hypothesis some years after he first put it forward. — It would provide useful confirmation of Egami's central hypothesis, that reversible property changes are associated with changes in CSRO, to compare the magnitude of such changes in those glasses which are known to have substantial amounts of CSRO in equilibrium, like Cu–Ti for instance, and others, such as Fe–Ti, which have been found to have little or no equilibrium CSRO. No such comparisons have yet been performed. However, there is mounting evidence, when the magnitude of reversible property relaxation is studied as a function of composition for a series of glasses in a ternary system, that the magnitude is greatest for compositions which permit the maximum exchange of species, and this supports the identification of reversible relaxation with changes in CSRO (GREER [1993b], STRÖM-OLSEN *et al.* [1988]).

Direct structural studies by XRD have provided only limited evidence. Changes in TSRO are so small that any changes in x-ray scattering measurements, like those shown in fig. 25, are so slight that they often vanish in the experimental scatter. However, EGAMI [1978b] did succeed in demonstrating significant changes in the x-ray diffraction pattern of a metallic glass, and showed that the densification on annealing involves a change in coordination number and not a reduction in the mean nearest-neighbor distance. The same features have been found for relaxation of amorphous silicon, where changes in the x-ray pattern are more substantial (ROORDA [1990]). — Reversible CSRO changes have not been detected by XRD, but rather by Mössbauer measurements, which are highly sensitive to nearest-neighbor geometry ((STRÖM–OLSEN *et al.* [1988]).

X-ray diffraction has, however, been successfully exploited in relation to another form of annealing — *stress-induced directional ordering*. (HILZINGER [1982]). This is closely related to field-induced directional ordering, in which a magnetic field is applied while the metallic (ferromagnetic) glass is heated: the glass so treated possesses an induced uniaxial magnetic anisotropy, and it has been well established since the 1950s

(GRAHAM [1959]), for crystalline solid solutions which behave similarly, that this effect is caused by directional order, that is, a very slightly different proportion of chemically unlike nearest neighbor pairs measured parallel and perpendicular to the originally applied field. In fact, the rate at which the anisotropy is established provides a rough measure of diffusion rates in glasses (LUBORSKY [1980]). It appears that the stress-induced effect is more pronounced: SUZUKI *et al.* [1987], in a striking experimental tour de force, were able to show directly, by energy-dispersive x-ray diffraction from a stress-annealed $Fe_{40}Ni_{40}Mo_3Si_{12}B_5$ glass, done successively with two orthogonal diffraction vectors, that directional short-range order (called "bond-orientational anisotropy" by them) was induced by the stress-annealing and decayed gradually on subsequent stress-free annealing. These investigators presumed that what they had observed was really a change in TSRO, with changes in the absolute numbers of bonds in the two directions; however, a later study from the same laboratory (TOMIDA and EGAMI [1991]) recognized that chemical as well as topological anisotropy must be involved in the phenomenon. These exact studies show how difficult it is to secure definitive information about TSRO and CSRO.

There is no space here to discuss in any detail the effects of irradiation on TSRO and CSRO: The effects of neutron irradiation on density and associated changes in ductility (see first entry in table 5) are linked with changes in TSRO, while the studies of Hillairet and his group (e.g., HILLAIRET *et al.* [1985]), using electrical resistivity as a marker, have generated good though indirect evidence of substantial changes in CSRO within atomic displacement cascades in glasses such as $Ti_{50}Be_{40}Zr_{10}$.

The kinetics of relaxation became a disputed problem when the *crossover effect* was discovered, in respect of the Curie temperature and elastic modulus (GREER and SPAEPEN [1981]; SCOTT and KURŠUMOVIĆ [1982], respectively). Another name for what was discovered is *memory effect.* What this involves can conveniently be illustrated by an older study of changes in refractive index of an oxide glass (MACEDO and NAPOLITANO [1967]). Figure 33 and its caption explain how such an experiment is done. If there were only a single relaxation process, the plot would simply show a horizontal line when the

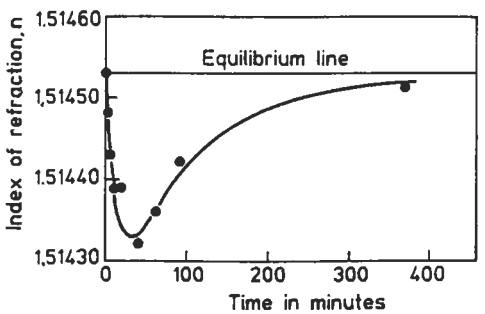


Fig. 33. Refractive index *vs* time in a "crossover" experiment on a borosilicate glass. The glass was held at temperature T_1 until the index reached a value corresponding to equilibrium at a second, higher, temperature T_2. The temperature was then increased abruptly to T_2: instead of remaining constant, the index followed the curve shown, returning eventually to the T_2 equilibrium value. (After MACEDO and NAPOLITANO [1967]).

temperature was raised from T_1 to T_2. Experiments like this show that there is not a one-to-one relationship between a property and the internal state of the glass, in the sense that the same value of the property can be associated with distinct structural states generated by different heat-treatment programs. There must be at least two distinct processes, but in fact with respect to metallic glasses, detailed analysis has shown (GIBBS *et al.* [1983], LEAKE *et al.* [1988]) that an entire spectrum of processes with distinct activation energies must be involved. This kind of experiment, again, shows the near-impossibility of associating different aspects of relaxation with precisely defined structural changes. Nevertheless, a number of investigators have attempted to interpret relaxation kinetics in terms of the reduction of free volume, and this kind of theory (e.g., VAN DEN BEUKEL and SIETSMA [1990]) is linked with theories of the glass transition, which are also concerned with ideas of critical free volume.

Some remarks about how structural relaxation influences the creep behavior of metallic glasses will be found in § 4.4, below.

4.3. Crystallization

As emphasized in § 2.1, a metallic glass can be considered to be an undercooled liquid, and as such it is susceptible to crystallization. While there may be alloy compositions for which a metallic glass is more stable than any single crystalline phase, the glass always has a higher free energy than a combination of crystalline phases. Crystallization of metallic glasses has been widely studied, and a number of reviews are available (e.g., KÖSTER and HEROLD [1981], KÖSTER [1984], KÖSTER and SCHÜNEMANN [1993]). There are several reasons for interest. It is important to characterize the stability of metallic glasses in order that desirable properties may be preserved under service conditions. Crystallization of a glass (also known as *devitrification*) is very much slower than conventional solidification of a liquid, and it can readily be interrupted by lowering the temperature. This facilitates fundamental studies of the processes of nucleation and growth, in particular of nucleation kinetics difficult to study quantitatively in the liquid state. A glass is formed only if crystal nucleation and growth are substantially avoided during liquid quenching; thus crystallization studies are of relevance for understanding glass formation. Finally, it is increasingly recognised, and is now a matter of commercial practice, that useful partially or fully crystalline microstructures can be obtained by crystallizing metallic glasses; this processing route is analogous to that used for glass-ceramics.

We compare first the basic kinetics of crystallization from the glassy state and solidification from the liquid state. Figure 34 is a schematic illustration, for a glass-forming alloy, of the variation of the homogeneous crystal nucleation frequency and the crystal growth rate over the temperature range from the liquidus down to below the glass-transition temperature T_g. As the undercooling is increased, the thermodynamic driving force for transformation to the crystalline state increases and the atomic mobility decreases. The forms of the temperature dependences of homogeneous nucleation and growth are readily derived from standard expressions; both show maxima at intermediate temperatures because of the lack of driving force at high temperature and the lack of

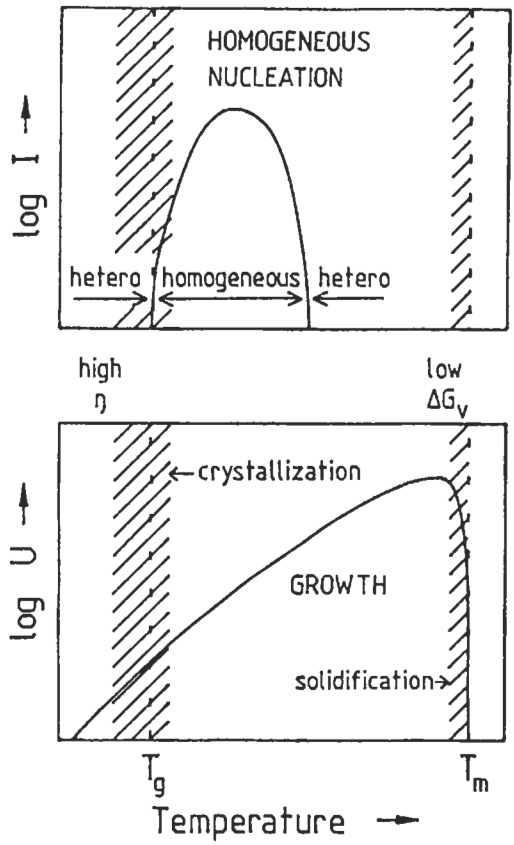


Fig. 34. Schematic representation of the homogeneous nucleation frequency and the growth rate of crystals in an undercooled melt or glass. The melting temperature, T_m, and the glass transition temperature, T_g, are shown, together with (shaded) the regions for conventional solidification and for crystallization of amorphous alloys, The temperature ranges in which homogeneous or heterogeneous nucleation dominate are marked, indicating that either can dominate for crystallization of amorphous alloys.

mobility at low temperature. The maximum rate occurs at a lower temperature for nucleation than for growth, because of the need for the driving force to overcome the solid/liquid interfacial energy. As seen in fig. 34, conventional solidification occurs at rather low undercooling with moderate growth rates on heterogeneous nuclei. Crystallization of a glass occurs near T_g at very low growth rates on nuclei which could be heterogeneous (predominant at lower temperature) or homogeneous. In conventional solidification the rate of latent heat production is matched by the rate of external heat extraction, but there can be significant temperature gradients. As liquids are undercooled further, nucleation may occur at temperatures for which the atomic mobility is still large; crystal growth is then rapid and the release of latent heat is more rapid than the heat extraction, leading not only to temperature gradients but to a rapid overall temperature

rise (*recalescence*). For crystallization of the glass, however, the crystal growth is so slow that the transformation can be isothermal, with negligible temperature gradients. The solidification processes considered in chapter 8 mostly occur close to the liquidus, at temperatures above the range in which recalescence is significant. The crystallization processes of concern in this chapter mostly occur near the glass transition temperature at temperatures below the recalescence range. While solidification and crystallization can be viewed within one framework as indicated in fig. 34, significant differences can be expected. Taking dendritic growth as an example, the rate-limiting processes in solidification are heat and solute redistribution at the growing tip, whereas in crystallization the rate is limited by the mobility of the crystal/glass interface itself and by solute redistribution.

Nevertheless, the common aspects of the processes in the two regimes have for example enabled heterogeneous nucleation in a metallic glass to be used to study particle inoculation (grain refinement) processes in conventional solidification processing. In such a study, SCHUMACHER and GREER [1994] have embedded inoculant TiB_2 particles in an amorphous aluminum alloy. Crystallization of the alloy to α-Al occurs on the particle surfaces, and novel microscopical studies of the nucleation mechanism are possible because the crystallization proceeds so much more slowly than conventional solidification.

Figure 35 shows a typical differential scanning calorimeter (DSC) trace obtained on heating a metallic glass. In the DSC the power necessary to heat the sample through any

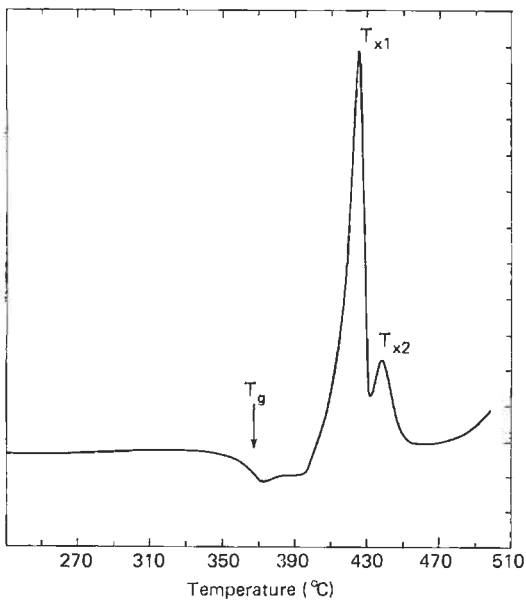


Fig. 35. Differential scanning calorimeter record during the heating, at 20K/min, of $Pd_{77.5}Cu_6Si_{16.5}$ glass. 6 mg sample. The ordinate represents power released (exotherm, positive displacements) or absorbed (endotherms, negative displacements). T_g is clearly differentiated from the first crystallization peak, T_{x1}.

small temperature interval is automatically compared with the power necessary to heat a reference sample (undergoing no transformations) through the same temperature interval. As plotted in fig. 35, an upward deflection corresponds to heat release from the sample. In this example of a comparatively easy glass-former, the glass transition is evident and shows the increase in specific heat when the glass changes to liquid. Before the glass transition is reached, temperature cycling experiments can reveal irreversible heat releases corresponding to structural relaxation in the glassy state (§ 4.2). The glass transition is followed by crystallization. As shown in this case, there are often separate peaks in the DSC trace revealing distinct stages in the crystallization of the sample, the integrated area under each peak giving the associated enthalpy change. The different stages may correspond to successive crystallization of different parts of the sample (perhaps arising from composition changes induced by the crystallization processes themselves), or to transformations from metastable crystalline states towards the final equilibrium state. As the thermodynamic driving force for crystallization of the glass is very high and the atomic mobility is low, it is likely that there will be metastable phases (or combinations of phases, for example a metastable eutectic) before equilibrium is reached.

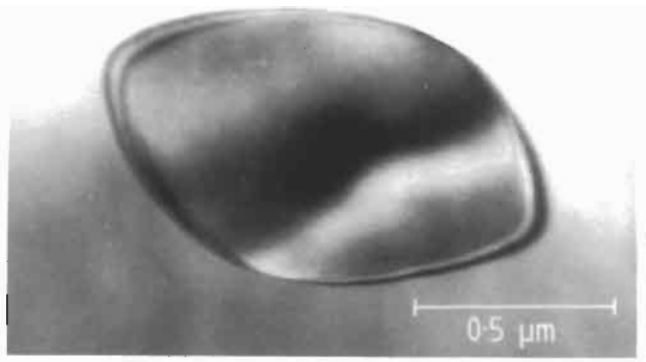
In thermodynamic terms, three types of initial crystallization can be distinguished. In *polymorphic* crystallization the glass transforms to a single crystalline phase of the same composition. The crystalline phase is not normally a solid solution of the base element because glass formation would be difficult if such a crystallization were possible and because any crystallization which did occur would likely involve solute partitioning. The phase formed in polymorphic crystallization is normally an intermetallic compound. In *primary* crystallization a single crystalline phase is formed which has a composition different from that of the glass. Solute partitioning is involved and the crystal growth is often dendritic. Complete transformation to the initial phase may not be possible and the remaining glassy matrix of changed composition may undergo a distinct crystallization. The third type of crystallization is *eutectic*, in which two crystalline phases grow cooperatively. There is no overall composition difference between the glass and a eutectic colony. Within the colony straightforward lamellar structures are only rarely observed, and the microstructure is on a fine scale which does not vary strongly with crystallization conditions. Microstructures arising from the three types of crystallization are shown in fig. 36. Ternary eutectic crystallization is also known, in which three phases form cooperatively. The growth rate of the eutectic colonies is particularly slow in such cases, which are usually associated with very easy glass formation (GARCIA-ESCORIAL and GREER [1987]). Another form of decomposition of a glassy phases is separation into two glassy phases (TANNER and RAY [1980]). These phases can then crystallize separately. The driving forces for the various forms of decomposition are illustrated schematically in fig. 37.

Experimental studies of crystallization can be based on isothermal anneals or on continuous heating of the kind used for fig. 35. In either case, DSC (which in effect measures enthalpy changes in the sample) or measurement of other property changes (notably electrical resistivity) can be used to derive the volume fraction x transformed as a function of time t. Isothermal x vs t measurements are often analysed according to

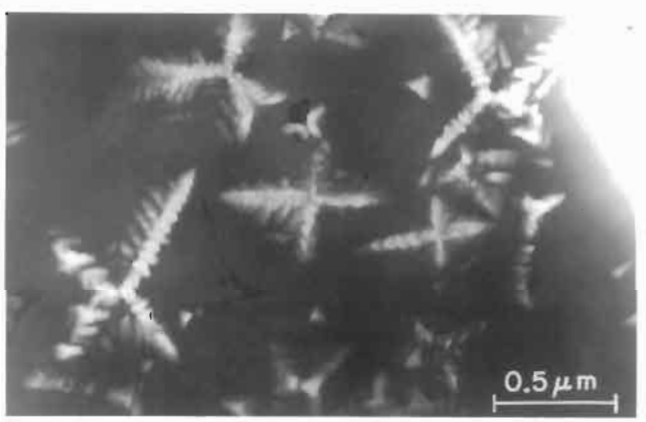
References: p. 1818.

Johnson–Mehl–Avrami kinetics, $x(t) = 1 - \exp[-(Kt)^n]$, where K is a temperature-dependent rate constant and n is an exponent which can relate to the modes of nucleation and growth (see, e.g., ch. 28, § 3.2). It has been frequently pointed out, however, that such analyses are subject to errors of interpretation and must be supplemented by microscopical studies. In particular, if the transformation in a sample differs from place to place, quite erroneous conclusions may be reached, as has been elucidated by HOLZER and KELTON [1991]. X-ray diffraction and related techniques are useful for identifying the crystalline products, but microstructural techniques are essential for understanding transformation kinetics. The scale of crystalline volumes developing on annealing a metallic glass can vary enormously, from a few nanometers up to as much as a millimeter in a very good glass-forming alloy. Depending on the scale, transmission electron microscopy (TEM), scanning electron microscopy of polished and etched surfaces or optical metallography may be appropriate. Recently, intense radiation sources, such as X-rays from a synchrotron, have permitted structural studies with very fine time resolution. Time-resolved studies during crystallization can reveal transient phases and other important aspects of the process (KÖSTER *et al.* [1991]).

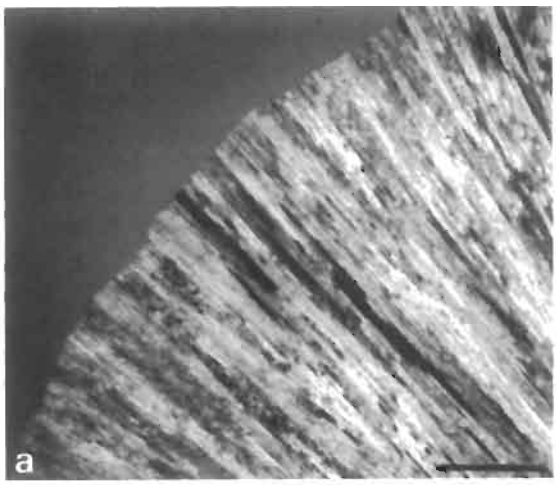
An example of a kinetic analysis is given in fig. 38 from the work of GREER [1982] on $Fe_{80}B_{20}$ glass which crystallizes eutectically to a mixture of α-Fe and a metastable boride,



(a)



(b)



(c)

Fig. 36. Transmission electron micrographs showing: (a) Polymorphic crystallization: $Fe_{76}B_{24}$ glass annealed at 400°C for 10 min (reproduced with the kind permission of U. Köster). (b) Primary crystallization: $Fe_{81.5}B_{14.5}Si_4$ glass annealed at 445°C for 20 min, showing α-iron dendrites. (c) Eutectic colony in $Fe_{40}Ni_{40}B_{20}$ glass, annealed 1 h at 390°C, showing γ-(Fe, Ni) + (Fe, Ni)$_3$B. Marker = 10μm.

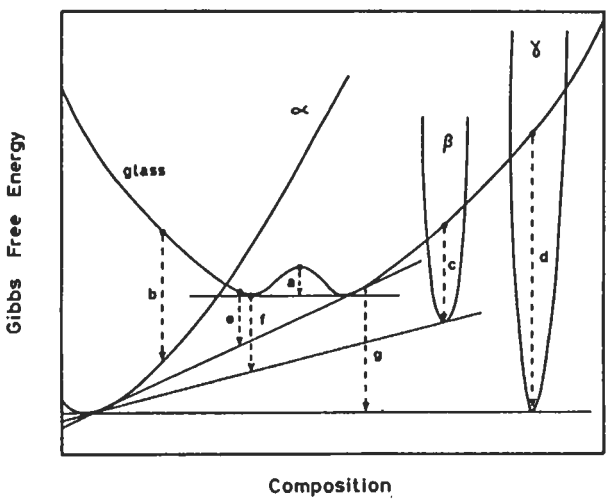


Fig. 37. Schematic free-energy curves for a binary alloy system, showing how the glass or amorphous phase can lower its free energy by undergoing: (a) separation into two amorphous phases; (b) polymorphic crystallization into the supersaturated solid solution α; (c) polymorphic crystallization to the metastable compound β; (d) polymorphic crystallization to the equilibrium phases γ; (e) primary crystallization of α; (f) eutectic crystallization to α and metastable β; and (g) equilibrium eutectic crystallization to α and γ.

References: p. 1818.

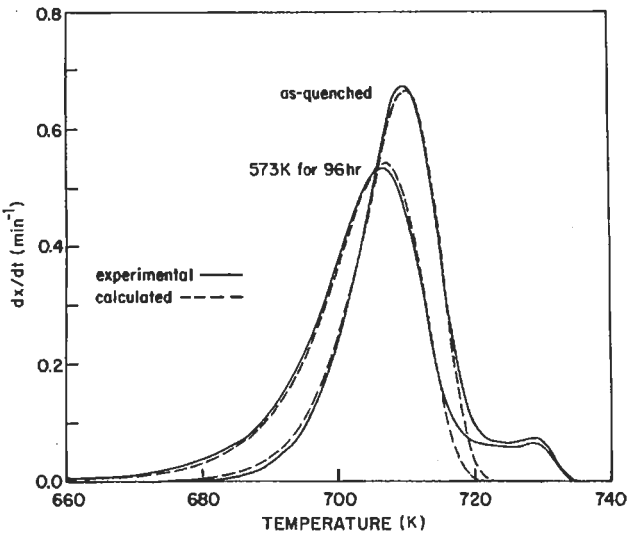


Fig. 38. Crystallization kinetics, dx/dt, deduced from continuous DSC heating data, for $Fe_{80}B_{20}$ glass 32 mm thick (solid curves), compared with fitted curves derived from a model. The lower peak corresponds to a preannealed sample. (After GREER [1982]).

tetragonal Fe_3B. The traces in fig. 38 are derived from DSC data and show the rate of change of transformed fraction during continuous heating. The data (main peaks) can be well matched by a numerical model which assumes eutectic growth from a fixed number of quenched-in nuclei. The population of these nuclei can be determined from the numerical fitting and was found to be of the order of $10^{18}/m^3$, in agreement with direct TEM observations. It was suggested that these nuclei most likely arose by homogeneous nucleation in the melt during quenching. The higher-temperature small peaks in fig. 38 arise from a thin layer on one side of the melt-spun ribbons which contains no quenched-in nuclei.

Nucleation kinetics in metallic glasses can be determined by direct counting of the numbers of crystals, by inference from overall transformation kinetics, or by measuring the crystal size distribution in a partially transformed sample. Examples of the latter are shown in fig. 39. The types of nucleation behavior which have been observed are illustrated schematically in fig. 40. Nucleation can occur during an anneal and can be homogeneous or heterogeneous, showing site saturation (see ch. 15, § 2.1). In either case transient effects may be evident. Alternatively, in some cases annealing merely leads to crystal growth on pre-existing nuclei. Homogeneous nucleation is perhaps most clearly shown in $Fe_{40}Ni_{40}P_{14}B_6$ glass, in which nucleation kinetics of type (b) (fig. 40) are found. Even for this well studied glass, however, a fully quantitative fitting of the kinetics to nucleation theory has remained elusive. The steady-state homogeneous nucleation frequency I^S (m^{-3} s^{-1}) should be of the form (KELTON [1991]):

$$I^S = \frac{A}{\eta} \exp\left[-\frac{16\pi}{3k_B} \frac{\sigma^3}{T\Delta G_V^{\,2}} \right]$$

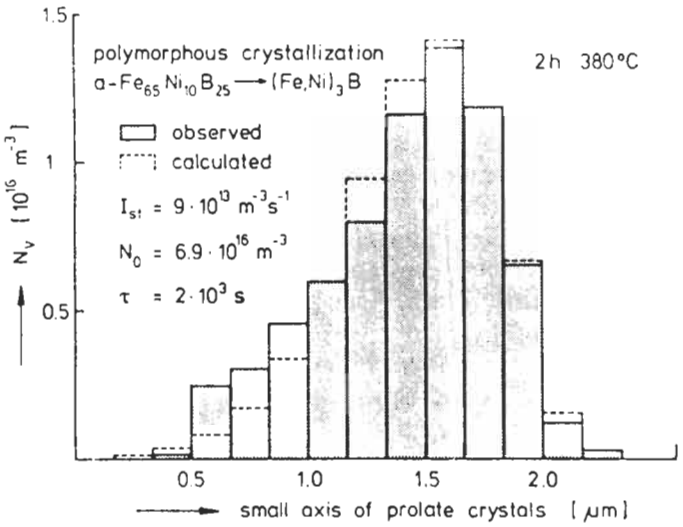


Fig. 39. Typical histogram of crystal diameter distribution in a partially crystallized metallic glass: polymorphous crystallization in $Fe_{65}Ni_{10}B_{25}$. (After KÖSTER and SCHÜNEMANN [1993]).

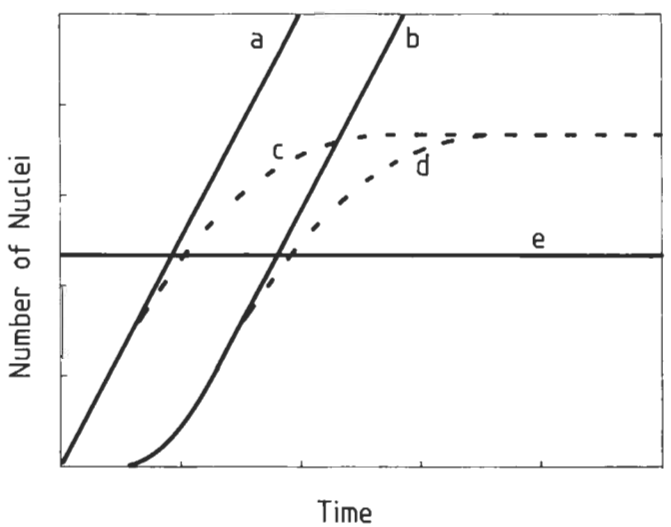


Fig. 40. Schematic variation in the number of nuclei with time in an isothermal anneal for the following nucleation types: (a) steady-state homogeneous; (b) transient homogeneous; (c) steady-state heterogeneous; (d) transient heterogeneous; (e) quenched-in active nuclei. (After GREER [1988]).

References: p. 1818.

where ΔG_V is the free energy driving force (J/m^3) for crystallization, σ is the crystal/glass (or liquid) interfacial energy (J/m^2), η is the glass or liquid viscosity (Pa s), k_B is Boltzmann's constant, and A is a prefactor. Of the three main materials parameters in this expression, η can be measured, ΔG_V can be estimated by various methods, but σ is not known. As reviewed by KELTON [1991], all reasonable assumptions about the form of the temperature dependence of η, ΔG_V, and σ, when substituted in the above equation, fail to fit the full set of measured steady-state nucleation frequencies, shown in fig. 41. If σ is independent of temperature, then a plot of $I^S\eta$ *vs* $1/T\Delta G_{V2}$ should be straight, but as can be seen in the figure this is so only over the central portion of the data. The value of A derived from fitting the central portion is reasonable for homogeneous nucleation. It is possible to fit the entire data set by assuming a temperature-dependent σ, but then the fitted value of A is much reduced and characteristic of heterogeneous nucleation (GRÁNÁSY *et al.* [1994]). Thus, even for this well studied glass the steady-state nucleation kinetics cannot be interpreted with certainty. While the crystallization of metallic glasses offers great potential for a quantitative test of nucleation theory in metallic systems, the potential remains largely unrealised, particularly because of problems with determining the materials parameters. Some aspects are already clear, however. The non-linear variation of free energy of crystallization with undercooling (arising from the excess specific heat of the liquid, as discussed in § 2.1) must be taken into account to explain quite basic experimental findings, for example that the activation energy for nucleation is typically greater than that for crystal growth (GREER [1994]).

In addition to $Fe_{40}Ni_{40}P_{14}B_6$ there are other metallic glasses for which the nucleation kinetics suggest homogeneous nucleation. For some glasses (for example the aluminum–

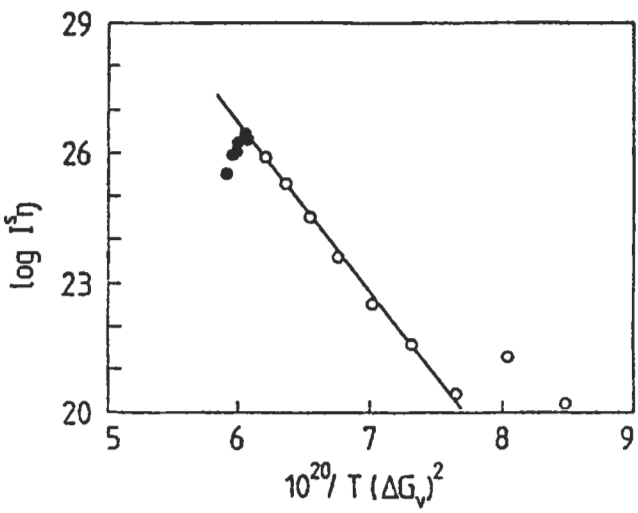


Fig. 41. Crystal nucleation in amorphous $Fe_{40}Ni_{40}P_{14}B_6$; the graph points are from two experimental studies. The plot of $\log I^S\eta$, *vs* $1/T\Delta G_V^2$ (where I^S is the homogeneous nucleation frequency and η is the viscosity) should be linear for a constant (i.e., temperature-independent) solid/liquid interfacial energy.

rare-earth–transition metal glasses discussed later in this section) the population density of nuclei, up to $10^{25}/m^3$, is so high that homogeneous nucleation seems the only plausible hypothesis. In yet other cases (for example $Fe_{80}B_{20}$ as shown in fig. 38) crystallization is by growth on a large population of quenched-in nuclei. The large population of such nuclei and its strong dependence on quench rate suggest transient homogeneous nucleation during the rapid quenching of the alloy. Analysis of the transient effects suggests that in some cases, inhibition of nucleation is essential if glass formation is to be possible (KELTON and GREER [1986]). This is illustrated for the case of $Au_{81}Si_{19}$ in fig. 42. The dashed line in the lower figure indicates the low level of volume fraction transformed (10^{-6}) which is usually taken as the condition for successful formation of a (metallic) glass; it can be seen that this level is achieved with transient nucleation for a quench rate of 10^5 K/s, whereas for steady state nucleation the unrealistic rate of 10^8 K/s would be needed. Thus, crystallization studies and their analysis reveal that for the formation of some metallic glasses from the liquid-high quench rates are absolutely necessary.

In marked contrast, there are other metallic glasses for which crystallization is clearly dominated by heterogeneous nucleation, the homogeneous nucleation frequency being very small. Examples are the good glass-formers $Pd_{40}Ni_{40}P_{20}$ (DREHMAN and GREER

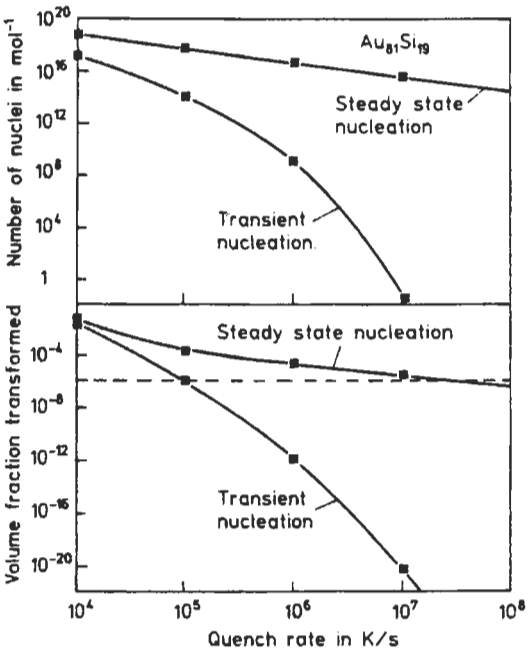


Fig. 42. (a) The number of nuclei and (b) the transformed crystal fraction calculated for quenching molten Au81Si19 at various rates, assuming either steady-state or transient homogeneous nucleation. (After KELTON and GREER [1986]).

[1984]) and $Pd_{77.5}Cu_6Si_{16.5}$ (KIMINAMI and SAHM [1986]. That the predominant nucleation is heterogeneous can be exploited, as processing under very clean conditions can permit bulk glass formation at very low cooling rates (as low as ~ 1 K/s for $Pd_{40}Ni_{40}P_{20}$ (DREHMAN and GREER [1984]).

Surfaces are often preferred sites for crystallization of a metallic glass. The main reason is modification of surface composition, for example by oxidation. There are clear possibilities to manipulate surface composition to enhance or impede nucleation, or to select crystallizing phases different from those which would be formed in the bulk. Comprehensive work in this area has been by KÖSTER [1984]. Surface crystallization can be used to improve bulk properties of a metallic glass (particularly magnetic properties) by inducing a compressive in-plane stress in a thin ribbon (OK and MORRISH [1981]). WEI and CANTOR [1989] made a detailed study of surface crystallization of $Fe_{40}Ni_{40}B_{20}$ and found that it is enhanced if the alloy is first phase-separated by relaxation and then abraded at the surface, or else the surface is enriched in nickel by electroplating and the glass is then annealed. Sometimes, surface oxidation can have the opposite effect of inhibiting crystallization at the surface: thus, with $Pd_{40}Ni_{40}P_{20}$, a thin NiO layer protects the glass from local loss of phosphorus; such loss is here the prime cause of preferential surface nucleation. Removal of the original surface can drastically modify crystallization behavior: thus, $Ni_{66}B_{34}$ glass (which normally crystallizes from the surface with a strong accompanying [100] fiber texture), loses the quenched-in nuclei responsible for this if the surface is etched off (KÖSTER and SCHÜNEMANN [1993]); their chapter can be consulted for more detailed information concerning crystallization at surfaces.

Finally in this section we consider crystallization of metallic glasses as a route to desirable microstructures. This concept is well known for glass-ceramics. For metallic glasses it has been comparatively neglected, but is receiving increasing attention. It may seem strange to adopt an extreme processing route (such as melt-spinning) to avoid crystallization and form a glass, only then to crystallize it. The important point is that the crystalline product made from the glass can have a very fine and uniform microstructure, fine because the high effective undercooling favors nucleation, and uniform because of the uniformity of the glass itself and the lack of macrosegregation during crystallization. Devitrification offers a particularly attractive way of making nanocrystalline material (with grain size less than 100 nm) cheaply, controllably, and in comparative bulk without the need for consolidation.

The first serious work on the exploitation of devitrification was by RAY [1981] on alloys of Fe, Ni, Al, Cr, Mo, Co and W in multiple combinations, with 5 to 12 at.% of B or other metalloids as aids to glass formation. Similar alloy developments have been made by DAS *et al.* [1985] and VINEBERG *et al.* [1985]. Ni–Mo–B and Ni–Al–Ti–X–B and later other Ni–Mo–B alloys with added Cr were made by melt-quenching, comminution and consolidation by extrusion or HIPping. During the processing, ordered phases including Ni_4Mo, Ni_3Mo, Ni_2Mo and $Ni_3(Al,Ti)$ are precipitated from the crystallized matrix, together with stable boride precipitates. This family of alloys is now manufactured commercially under the trade name Devitrium® (VINEBERG *et al.* [1985]), and the best examples have impressive high-temperature properties, exceeding high-grade tool steels. Work in this area is continuing; for example, ARNBERG *et al.* [1991] have

developed a range of tool steels by crystallizing Fe–Cr–Mo–C–B or Fe–Cr–Mo–C–V glasses.

The most important industrial application of metallic glass crystallization has been in the production of the modern family of permanent magnets based on the high-coercivity phase $Nd_2Fe_{14}B$. These materials can be made either by sintering an alloy powder, by melt-spinning the alloy to form a glass and then devitrifying, or by direct quenching to a nanocrystalline structure. The melt-quenching method was introduced by CROAT *et al.* [1984], and has been fully commercialized. The process has been reviewed by BUSCHOW [1986]. Figure 43 provides an overview of the production possibilities by direct quenching or by heating a glass. The hard magnetic phase is desired, the alternative soft magnetic phase must be completely avoided, and a fine grain size is desirable for high coercivity. In the materials of most interest, grain sizes range from 14 nm to 50 nm.

Interest in nanocrystalline magnetic materials now extends to the opposite extreme: very soft magnets. Glassy alloys based on Fe–B–Si are well known as soft magnetic materials. If the composition is modified by addition of copper (apparently to enhance crystal nucleation) and niobium (to impede crystal growth), crystallization of such alloys leads to an isotropic microstructure with a very fine grain size of 10 to 20 nm (YOSHIZAWA *et al.* [1988]). At this scale the local magnetocrystalline anisotropies are averaged out, and soft magnetic properties result. The materials, now marketed under the trade name Finemet®, are of interest in comparison to glassy soft magnetic alloys because they can have very low magnetostriction; the magnetostrictions of α-Fe crystallites and of the residual amorphous phase in the microstructure are of opposite sign and substantially cancel (HERZER [1991]).

Nanocrystalline materials made by devitrification of metallic glasses may also be of interest for their mechanical properties. Aluminum-based alloys with rare earth and late transition metal additions up to a total of 10 to 15 at. % can either by direct quenching

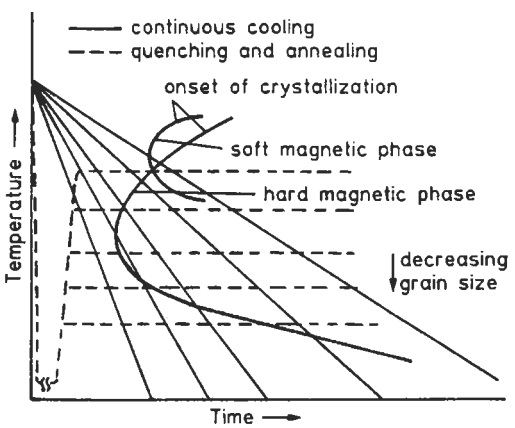


Fig. 43. Schematic time–temperature-transformation diagram to rationalize optimum conditions for producing Nd–Fe–B permanent magnets. (After WARLIMONT [1985]).

or by devitrification yield microstructures — nanometer scale — with exceptional strength, discussed further in § 4.4.1. Consolidation of glassy or partially devitrified material may lead to a fully crystalline product, but one in which a fine-scale microstructure with desirable properties is retained. A particular property of interest is superplasticity. An alloy of the aluminum – rare earth – transition metal type has shown superplastic elongations of more than 500% at strain rates as high as 1/s (Higashi *et al.* [1992]). Even after consolidation and deformation, the mechanical properties (for example, a tensile strength ~ 900 MPa) are superior to those of conventional high-strength aluminum alloys.

4.4. Plastic deformation and fracture

The principal plastic characteristics of metallic glasses are, first, their great strength (the flow stress is typically as much as 1/50 of the shear modulus); second, their fracture toughness, which in the normal state is very much greater than that of comparably strong fibrous materials such as graphite; and third, the mechanisms of flow and fracture which are entirely distinct from those found in crystalline materials.

Table 6 compares the flow stresses of a number of ductile metallic glasses, both on an absolute basis and normalized for density, with those of some other ultra-high-strength wires or fibres. Metallic glasses can be seen to be among the strongest practical materials known (the adjective is to distinguish them from single-crystal whiskers).

Metallic glasses deform plastically by two entirely different regimes. Under tension, at high stresses near their fracture limit, they deform extremely inhomogeneously: shear

Table 6
Tensile properties of various continuous filaments and ribbons.

Material	Yield stress σ_y/GPa	Relative density d	Young's modulus		
			$\sigma_y d^{-1}$/GPa	E/GPa	$E d^{-1}$/GPa
S-glass (SiO_2–Al_2O_3–MgO)	5.0	2.5	2.0	85	34
C fibre (high-yield PAN type)[a]	3.2	2.6	1.7	490	153
B filament (on W core)[a]	2.5–4.5	2.6	1.0–1.7	380	146
SiC microcrystalline filament[a]	3.5	2.6	1.4	200	77
Kevlar fibre (organic polymer)[b]	2.8	1.5	1.9	135	90
High-C steel wire	4.2	7.9	0.55	210	27
$Fe_{80}B_{20}$ met. glass[b]	3.6	7.4	0.5	170	23
$Ti_{50}Be_{40}Zr_{10}$ met. glass	2.3	4.1	0.55	105	26
$Ti_{60}Be_{35}Si_5$ met. glass[b]	2.5	3.9	0.65	110	28
$Cu_{50}Zr_{50}$ met. glass[b]	1.8	7.3	0.25	85	12

[a] Indicates heat-resistant materials.
[b] Indicates materials with some ductility.

bands form on planes close to the orientation of maximum resolved shear stress. These are very thin (typically 10–20 nm thick according to DONOVAN and STOBBS [1981] and high-speed cinematography and acoustic emission indicate that the bands form extremely rapidly. NEUHAUSER [1978] found a formation time for individual bands in $Pd_{80}Si_{20}$ < 2 ms.) At low stresses, there are no shear bands and deformation is by homogeneous creep. A schematic deformation map separating the two deformation regimes is shown in fig. 44. The ductility in inhomogeneous deformation is usually only about 1% in tension, whereas many glasses can be rolled to reductions exceeding 50% if sufficiently hard and stiff miniature rolls are used. Creep rates in most creep experiments have been so low that total creep strains in reasonable times are less than 1%, but it is possible at temperatures close to T_g to hot-press a metallic glass foil such as $Fe_{40}Ni_{40}P_{14}B_6$ to strains

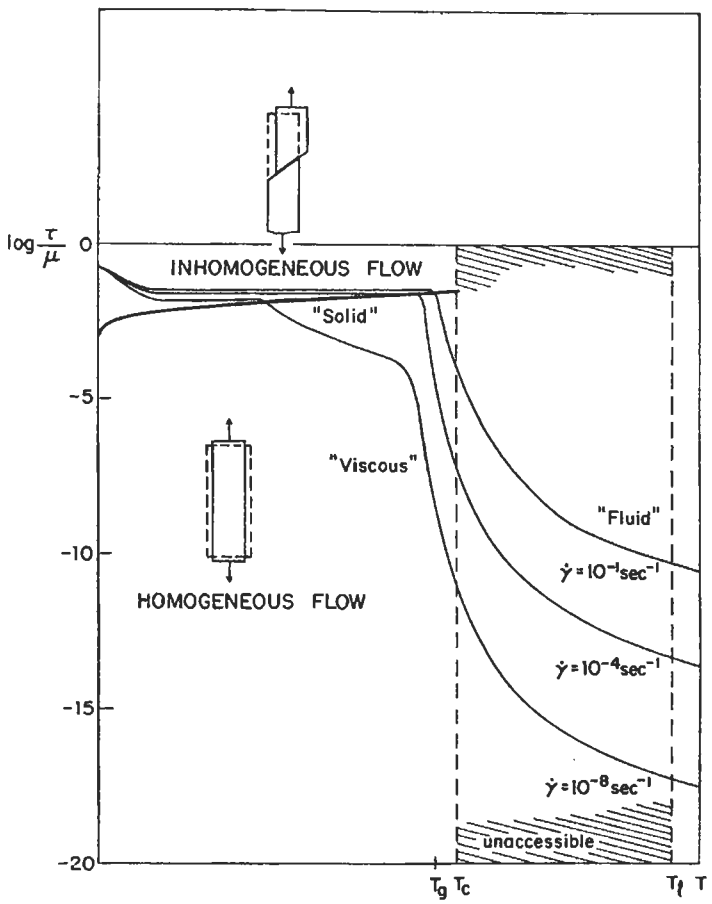


Fig. 44. Deformation map for an idealized metallic glass. T_g is the glass transition, T_c the crystallization temperature, T_l the liquidus temperature, τ is the shear stress, μ the shear modulus, $\dot{\gamma}$ is shear rate. (After SPAEPEN [1977]).

References: p. 1818.

of the order of unity (PATTERSON *et al.* [1978]). This approach has been used in Japan (MASUMOTO [1994]) to shape minute but complex objects such as a micro-gearwheel from a metallic glass.

The mode of ductile fracture of a metallic glass is unique. Fracture following inhomogeneous flow always takes place by separation along a previously active shear band — i.e., it is strictly a shear failure — and the fracture surface shows a "river pattern" of fine ridges which can be exactly mimicked by simultaneously shearing and pulling apart a pair of glass slides coated with vaseline. Accordingly, it has been proposed that the material in a shear band which is about to fail behaves in a liquid-like manner. Shear along one of the shear bands clearly generates additional free volume and this lowers the viscosity locally and can eventually lead to separation along the shear band. The creation of free volume can be deduced from the measured decrease ($\approx 0.15\%$) in density after extensive inhomogeneous shear (CAHN *et al.* [1984]). The complex fracture mechanics of metallic glasses are reviewed by KIMURA and MASUMOTO [1983]. There is no record of the appearance of a fracture in creep; apparently glasses subject to *slow* homogeneous flow undergo perfect viscous flow with no structural deterioration that could lead to fracture.

When a metallic glass is stressed at a high temperature, as remarked, it undergoes homogenous flow by a purely viscous process. As creep proceeds, free volume is progressively annealed out and since the viscosity is directly linked to free volume, the creep rate under constant stress progressively diminishes. The relation between viscosity, η, stress, σ, and strain rate, $\dot{\varepsilon}$, is: $\eta = (1/3)\,\sigma/\dot{\varepsilon}$. Instead of a constant-stress creep test, one can also measure stress relaxation at constant strain or perform a bend-stress relaxation experiment (a ribbon is elastically bent and held at a fixed radius while it is annealed, and then released to see how far it springs back). A brilliant series of experiments to study these processes, mostly with $Pd_{82}Si_{18}$ glass, was performed by Spaepen and Taub and reviewed by them (SPAEPEN and TAUB [1983]).

According to the Stokes–Einstein relation, $\eta = kT/Db$, where D is a self-diffusivity and b is a characteristic distance of the same order as the interatomic distance (NACHTRIEB [1976]). The rise in viscosity on relaxation due to reduced free volume should thus be associated with a fall in diffusivity. However, there have been increasing doubts as to the applicability of the Stokes–Einstein equation to metallic glasses, and a number of experiments claimed to show that the equation breaks down have recently been critically discussed by VAN DEN BEUKEL and SIETSMA [1994]. Van den Beukel has recently proposed that, for metallic glasses, ηD^2 is constant and has advanced reasons for preferring this relationship (DUINE *et al.* [1992, 1993]).

The free volume model of flow and viscosity, originated by Cohen and Turnbull, can be outlined as follows: Each atom is confined in a cage defined by its nearest neighbors; that cage can vary in volume, and the free volume is that part of a cage in which the atom can move without an energy change. Normally the atom can only oscillate, but if there is a density fluctuation there is a computable probability (dependent of course on temperature) that the atom will jump to a neighboring cage if the fluctuation generates an instantaneous local free volume exceeding a critical amount. Such jumps define both viscosity and creep rate under stress. Relaxation diminishes the average free volume and

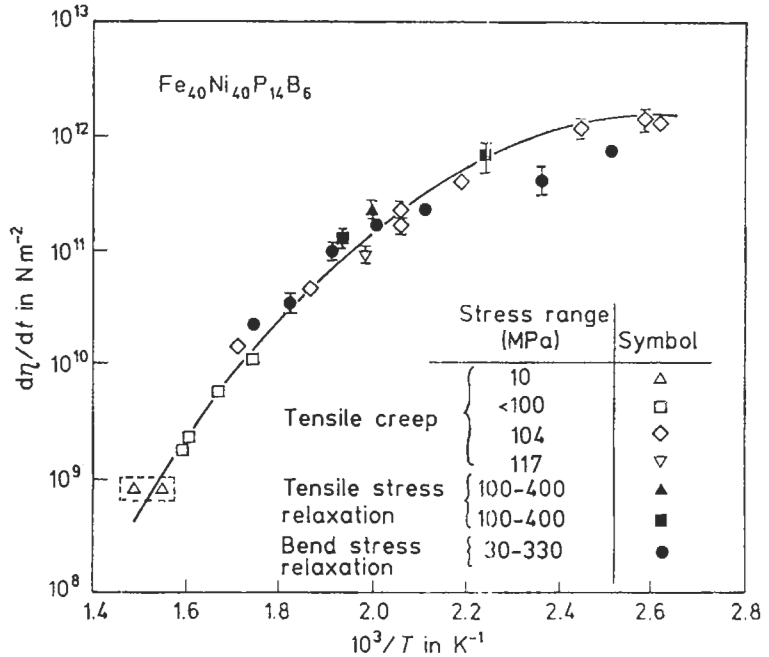


Fig. 45. Rate of change of viscosity of $Fe_{40}Ni_{40}P_{14}B_6$ glass with time as a function of temperature, as measured by three different experimental techniques. (After SPAEPEN and TAUB [1983]).

thus increases the viscosity, linearly with time as it turns out. Figure 45 shows how the (linear) rate of increase of viscosity varies with temperature for a metal-metalloid glass.

Attempts to interpret such relaxation-induced changes quantitatively are still beset by some difficulties, notably the need to assume that the concentration of flow defects (unlike free volume collectively) to some extent changes reversibly with temperature. A "flow defect" here is a particular configuration of free volume of a kind which favors atomic jumps, as sketched above. (The notion that the reversible element in creep behavior is linked, as suggested for other aspects of relaxation, with changes in CSRO, does not seem to have been seriously analysed.) For details of the theoretical considerations involved, the reader is referred to Spaepen and Taub's review, and to other detailed reviews by CHAUDHARI *et al.* [1983] and SPAEPEN *et al.* [1986]. There is a surprising measure of reversibility in the viscosity (a value characteristic of a given temperature T_1 can be attained directly by annealing a quenched sample at that temperature or by annealing at T_1 a sample previously equilibrated at a lower temperature (VOLKERT and SPAEPEN [1989])). It is undoubtedly difficult to define in an acceptable operational manner what is meant by a defect in an amorphous structure. Very recently, SPAEPEN [1994] has proposed regularities governing the evolution of a population of defects (density fluctuations) in a metallic glass of different sizes (singlets ... i.e., effectively a single missing atom ... doublets, triplets and quadruplets), with a range of shifts from one to another among these categories. The idea is that singlets govern

diffusion, doublets govern plastic flow, triplets can annihilate partially and quadruplets are large enough to collapse and disappear entirely. This model is consistent with the apparent inapplicability of the Stokes–Einstein relationship and the fact that different regularities describe diffusion and flow.

The figures cited in table 6 indicate that there is scope for exploiting the very high strength of metallic glass ribbons or wires in composites (especially in view of their usually excellent corrosion resistance), but surprisingly not much has been done to pursue this. HAGIWARA *et al.* [1985] shown that iron-based INROWASPed glass *wires,* strengthened by additions of niobium, tantalum, chromium or molybdenum, combine high yield strengths with particularly good fatigue strengths. At about the same time, OHNAKA [1985], the main originator of the INROWASP process, suggested the use of wires made by this process as motor tire reinforcement, and subsequently an American company showed that Fe–Si–B, with an appropriate coating, fulfil the requirements of tire reinforcement, which are: a high tensile strength, good adhesion to rubber, and excellent resistance to fatigue and corrosion (OGINO [1986]). A recent study (VAIDYA and SUBRAMANIAN [1990]) has shown the potential of metallic glass ribbons made by melt-spinning to reinforce glass-ceramics (i.e., devitrified oxide glasses). The melt-spun ribbons are assembled with the oxide glass precursor in powdered form; the assembly is sintered at 400–450°C, to full density, and sintering is followed by devitrification of the powder without devitrification of the metallic glass. Large increases of strength and toughness were achieved with as little as 1 vol% of reinforcing ribbons.

Apart from the good strength of uncrystallized (undevitrified) metallic glasses, recent researches on aluminum-base glasses have revealed the remarkable properties that can be attained by *partial* devitrification. This topic is covered in the next section.

4.4.1. Partially devitrified (nanostructured) metallic glasses

As emphasized in the previous section, metallic glasses can have very high yield strengths, approaching the ideal strength (that is, the strength shown by a dislocation-free material). In steels, because of the microstructural changes associated with the ferrite-austenite transition, conventional materials can also reach a significant fraction of the ideal strength. However, for light alloys based on Al or Mg such a transition does not exist, and the hardening based on precipitates or dispersoids tends not to be so effective. For light alloys, therefore, the very high specific strengths (that is, strength/density) of amorphous alloys are of particular interest. As outlined in § 2.2, there are amorphous alloys based on aluminum with rare earth and transition metal additions and these alloys are ductile. Their flow strengths can be as high as 1250 MPa (BLANK–BEWERSDORFF [1991]), about twice as high as the best precipitation-hardened conventional alloy, and immensely harder than the commercial purity element.

Partial or complete devitrification usually embrittles a metallic glass. In some cases, however, the ductility can even be improved (HILLENBRAND *et al.* [1982]). In the case of the Al-based glasses, partial devitrification leads to increased tensile strength — as much as a 50% increase on the corresponding amorphous alloy — while ductility is retained. The partially devitrified materials, introduced in § 4.3, have a fine dispersion of 3- to 10-nm Al crystallites in an amorphous matrix. They are obtained by quenching at below

the critical rate for full glass formation or by annealing an initially fully glassy material. The devitrification can also be induced by mechanical working. These materials, developed in Japan (KIM *et al.* [1990]) and the United States (CHEN *et al.* [1991]) show optimum tensile strength for a crystalline volume fraction of ~ 25%, and the values are quite exceptional — as high as 1560 MPa for $Al_{88}Ni_9Ce_2Fe_1$ (INOUE *et al.* [1992]). The strengths of the devitrified materials are compared with other aluminum alloys, including amorphous alloys, in fig. 46. This shows that as the microstructural scale of aluminum alloys is decreased the strength increases. The reasons why a nanometer-scale dispersion of crystallites should give such a pronounced hardening effect are not yet clear. However, the aluminum crystals appear to be too small to contain any dislocations.

4.4.2. Thermal embrittlement

Many metallic glasses are capable of a moderate degree of deformation at room temperature: the process is heterogeneous and involves the formation of shear bands

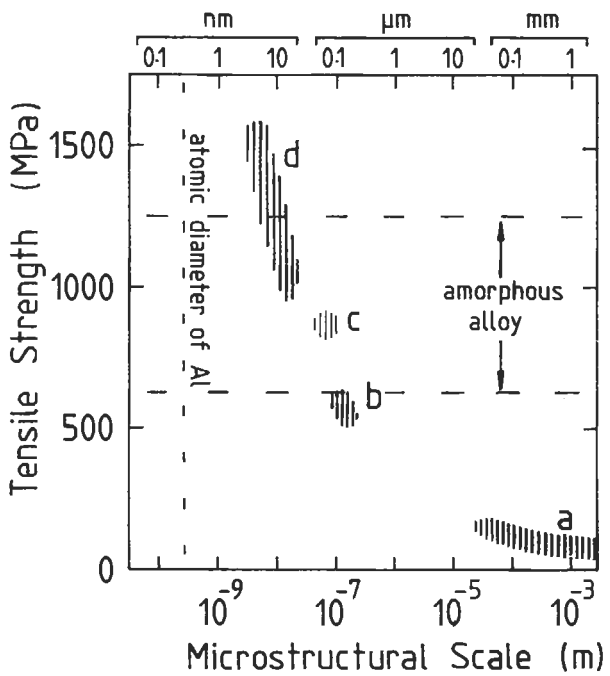


Fig. 46. The tensile strengths of aluminum alloys compared as a function of microstructural scale: (a) commercial purity aluminum, (b) the strongest conventional precipitation-hardened aluminium alloys, (c) amorphous, or partly amorphous, aluminum-rare earth-transition metal alloys, consolidated and crystallized (HIGASHI *et al.* [1992], OHTERA *et al.* [1992]), and (d) amorphous aluminum-rare earth-transition metal alloys partially devitrified (either during quenching or by annealing) to give nanometer-scale crystallites (CHEN *et al.* [1991], INOUE *et al.* [1992], KIM *et al.* [1991]). The diameter of an aluminum atom, and the wide range of measured tensile strengths for *fully amorphous* Al–TM–Ln alloys are shown for comparison (TM = transition metal, Ln = lanthanide).

References: p. 1818.

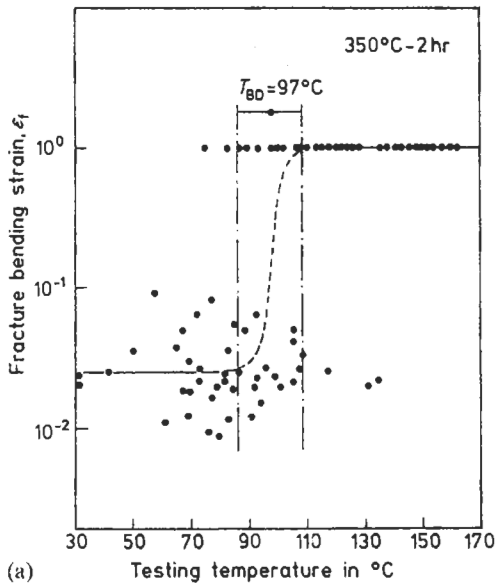
approximately along a plane of resolved shear stress. A ductile metal ribbon can generally be bent sharply back on itself without fracturing. Some glasses, mostly those based on iron (one might call these "glassy steels") show (1) a ductile–brittle transition temperature, and (2) the phenomenon of *thermal embrittlement*, which is marked by an increase in the ductile–brittle transition temperature on annealing well below the crystallization temperature. Figure 47 exemplifies these phenomena. The mechanism of this embrittlement has given rise to much investigation and argument.

SPAEPEN [1977] first put forward the idea that shear band formation (morphologically similar to slip bands in crystalline metals) is to be attributed to a multiplication of free volume at an incipient band, which reduces the flow stress locally so that the shear band develops preferentially. As mentioned above, CAHN *et al.* [1984] showed that the overall density reduction in a Pd–Cu–Si glass due to a 40% rolling reduction was 0.14%, which implied a much larger density reduction within the narrow shear bands themselves. This notion was further analyzed by ARGON *et al.* [1985], and they concluded, from measured overall density changes in heavily deformed glasses (rolled or wire-drawn) and the observed shear-band geometry, that within a shear band itself the local dilatation due to free volume can be as large as 50%. CAHN *et al.* [1984] further showed that the excess free volume in the shear bands was dissipated during a relaxation anneal, while KRISHNANAND and CAHN (1975) had earlier shown that after such a relaxation anneal and renewed plastic deformation, the new population of shear bands was not in the same sites as the old, i.e., the old shear bands had been locally "healed" by the relaxation anneal.

It can thus be taken as well established that free volume plays a central role in both homogeneous (creep) and heterogeneous plastic flow in metallic glasses.

The theories to explain thermal embrittlement fall into two categories: (1) the hypothesis that annealing removes excess free volume and (2) the hypothesis that the homogeneous glass structure separates into two distinct amorphous phases. The basis of the first hypothesis is clear from the foregoing: further, WU and SPAEPEN [1986] have assembled detailed evidence on thermal embrittlement consistent with this model. The second hypothesis leans on observations such as that due to WALTER *et al.* [1976], who by Auger electron spectroscopy established an enrichment of phosphorus at fracture surfaces of a $Fe_{40}Ni_{40}P_{14}B_6$ glass, and PILLER and HAASEN [1982] who used field ion microscopy to demonstrate the formation, during relaxation annealing of $Fe_{40}Ni_{40}B_{20}$, of minute zones enriched in boron to 25 at.%. However, tests by X-ray small-angle scattering of some other glasses have failed to demonstrate phase separation of this kind, whereas other experiments have given contrary results. The literature is quite extensive.

A possible marriage between the two types of hypothesis has been proposed in an interesting paper by YAVARI [1986]. By analysing the volume per metal atom in Fe–B glasses of different compositions, he shows that a local enrichment of B will need to "suck" free volume from the adjacent matrix, reducing the free volume concentration there. He suggests, therefore, that the phase separation hypothesis resolves itself, indirectly, into the free volume hypothesis. According to Yavari, the fact that some additives (e.g., Ce) delay embrittlement to higher temperatures, whereas others such as Sb promote it, can also be interpreted on the basis of his model. — Another observation



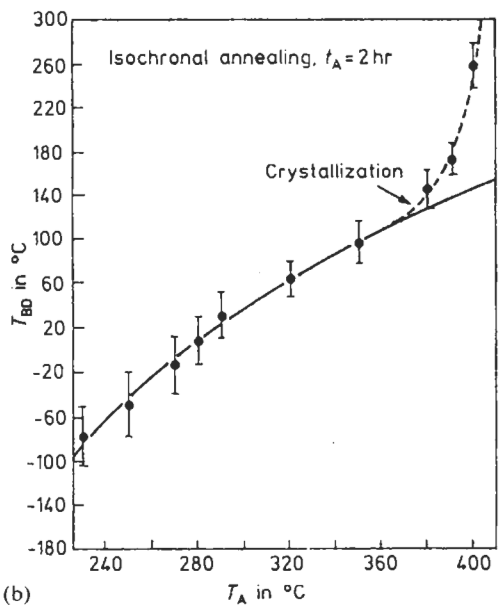


Fig. 47. Thermal embrittlement: (a) Strain at fracture in a bending test as a function of temperature for a $Fe_{79.3}Be_{16.4}Si_{4.0}C_{0.3}$ glass, annealed for 2 h at 350°C after melt-spinning. T_{BD} is the ductile–brittle transition temperature. (b) Change of T_{BD} with isochronal anneals at progressively increasing temperatures. (After SPAEPEN *et al.* [1986]).

References: p. 1818.

(YAMASAKI *et al.* [1985]) to the effect that Fe-based glasses whose compositions fall close to an equilibrium eutectic composition embrittle much more sluggishly, if at all, is attributed to a reduced tendency to phase separation for such compositions.

The clearest demonstration to date of the central role of free volume in determining the incidence of thermal embrittlement comes from an important series of publications by GERLING *et al.* [1985, 1988, 1989, 1990] relating to $Fe_{40}Ni_{40}B_{20}$ glass. Briefly, they have established in circumstantial detail how this glass is embrittled by the loss of free volume on annealing: for instance, the embrittlement behavior varies with ribbon thickness, because the quenched-in free volume varies also. They have further shown that a glass which has been thermally embrittled can be reductilized by neutron irradiation, which creates fresh free volume. The restored ductility can in turn be removed by a second relaxation anneal.

4.5. Other properties of metallic glasses

Apart from their mechanical properties, some metallic glasses are beginning to be exploited for their magnetic, chemical and diffusion properties, and as precursors for crystallization (devitrification). Ferromagnetic (mostly iron-based) glasses, as used nowadays in transformers, are treated in chapter 29 and will not be discussed here. Precursors for devitrification are treated in §§ 4.3 and 4.4.1, and this includes a brief account of permanent magnets made by this route. There is space here only for a summary account of the other properties.

Many metallic glasses have exceptionally good corrosion resistance, even in quite strong acids, and this is clearly due to their structural homogeneity, which obviates the creation of the electrolytic microcells that enhance corrosion of crystalline metals. For some uses, metallic glasses can be used "as is": thus, Fe–Cr–P–C ribbons were used for the active element in electromagnetic filters to remove rust from water (KAWASHIMA *et al.* [1985]); since the field is strongest at the edges of ribbons, narrow ribbons were used. Their high corrosion resistance ensured a long life for the filter elements.

For other purposes, it is necessary to put a glassy coating on a crystalline metallic substrate; this can be done by either laser-glazing or sputter-deposition. DIEGLE and HASHIMOTO [1988] and HASHIMOTO [1993] have shown that deposits made by sputtering can provide quite remarkable corrosion protection: thus, a–Cu–Ta (anywhere from 20 to 80 at% Ta) passivates spontaneously even in 12N HCl, and behaves much better than pure crystalline tantalum, itself notably resistant to acids.

HASHIMOTO [1993] has also surveyed the use of metallic glass coatings in electrocatalytically active electrodes, for instance, for use in chlorine manufacture by electrolysis. However, for lack of space the reader must be referred to Hashimoto's excellent review. There has also been extensive research, published in the chemical literature, on the use of metallic glasses (in the devitrified condition) as catalysts for gas-phase reactions such as the hydrogenation of carbon monoxide: here again, the reader is referred to an excellent review by MOLNAR *et al.* [1989], as well as to Diegle and Hashimoto's conference proceedings.

Finally, metallic glasses have found an exotic use as diffusion barriers in microelec-

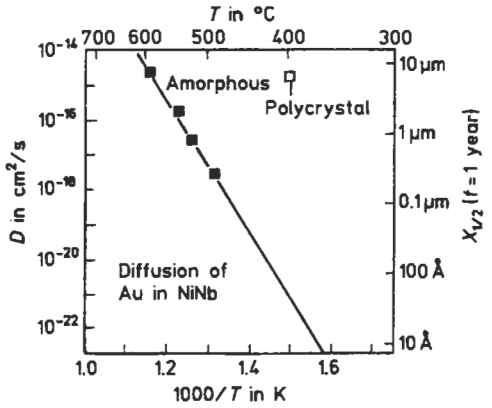


Fig. 48. Diffusivity of gold in amorphous and in polycrystalline $Ni_{55}Nb_{45}$. The righthand ordinate shows the diffusion distance in a period of one year. (Aftyer DOYLE *et al.* [1982]).

tronic circuits, especially for those to be used at temperatures of several hundreds of degrees Celsius, for which they are superior to most crystalline alloys (CAHN [1986], DE REUS [1995]). This works because in the temperature range of interest, atomic transport in polycrystalline metals is predominantly by grain-boundary diffusion, which is absent in glasses. Figure 48 shows this: the diffusion of gold in a–$Ni_{55}Nb_{45}$ is seen to be seven orders of magnitude slower than in the same alloy after it has been devitrified (which however requires higher temperatures to do). Ta–Ir glass, which crystallizes at over 900°C, has been effectively used to make diffusion barriers between silicon substrates and Y–Ba–Cu–O ceramic superconductors layers (DE REUS *et al.* [1988]).

5. *Rapid-solidification-processed (RSP) and other metastable crystalline alloys*

5.1. Age-hardening of light alloys

Before we treat the properties of RSP alloys, it is appropriate here to remind the reader that metastable crystalline states are by no means restricted to alloys which have undergone extreme cooling rates and often indeed, solid-state quenching is all that is required. The two such states that have been known and studied for longest are martensite in steels and the age-hardened state in light alloys, Al–Cu in particular: both these are products of solid-state quenching. Both were cited as archetypes in the first paragraph of this chapter. Martensite is discussed in chapters 16 and 17; here we devote some paragraphs to the key phenomenon of age-hardening, since the structural changes (as distinct from the hardening mechanism) are not treated elsewhere in *Physical Metallurgy*.

Age-hardening was discovered accidentally in 1906 by a German engineer, Wilm,

who examined an aluminum alloy containing 3.5 wt% copper, 0.5 wt% magnesium and 0.5 wt% manganese, and his findings were eventually published in 1911. His alloy, quenched from 500°C at which temperature it had been homogenised, hardened substantially over a period of days at room temperature. This process came to be called *age-hardening*. The original paper is not cited here, though it contains much of interest; it is too difficult to locate. Wilm's paper (in translation) and extracts from many other of the early key papers have been reproduced in a most useful book about age-hardening by MARTIN [1968], which also includes an extended commentary about the meaning of the early discoveries. Wilm did not understand why his alloy hardened and this was left to three American metallurgists, Merica, Waltenburg and Scott, who in 1919 established the shape of the solvus curve for copper in aluminum, establishing that the solubility increased sharply with temperature, and they were the first to suggest that the precipitation of the equilibrium phase, $CuAl_2$, was the basis of the hardening, although it could not be seen in the microstructure. (In fact, the age-hardening of Wilm's alloy, which he called "duralumin", was later found to be due not only to the copper but also to the precipitation of a form of the phase Mg_2Si, the silicon being an unintended impurity in the alloy).

Understanding did not develop further until in 1938 two independent notes appeared side by side in the journal *Nature,* by GUINIER [1938] and PRESTON [1938], soon followed by more substantial presentations by these same authors. They showed by Laue or oscillating-crystal x-ray diffraction photographs that age-hardening of Al–Cu alloys was associated with a *pre-precipitation* process: thin "zones" parallel to {100} became enriched in copper, without change as yet in crystal structure, and the lattice spacing was concurrently locally reduced because copper atoms are smaller than aluminum atoms. They also indicated that these zones (which later came to be called Guinier–Preston, or G–P, zones) developed into better defined intermediate crystal structures at a later stage of ageing. These later, distinct, phases in Al–Cu came to be called θ'', θ' and θ (all tetragonal), and all coherent with the matrix on {100}. θ is in fact the equilibrium $CuAl_2$ phase; θ' and θ are formed only when ageing is done above room temperature. The association of the successive stages of hardening with the development in the pre-precipitation structures was soon demonstrated, first by Gayler in 1940. Research on both the structural and the hardening stages proceeded intensively in the two following decades: an important synthesis was by SILCOCK *et al.* [1953–54]. Soon after this, the age-hardening stages began to be examined by transmission electron micrography, and such matters as forms of heterogeneous nucleation of zones and transition phases were examined. A definitive review of what had been discovered about these stages came a few years later (KELLY and NICHOLSON [1963]). This review covered not only Al–Cu alloys but also others, such as Al–Ag in which the initial age-hardened zones are spherical and Al–Mg–Si alloys in which they are rod-shaped. After that, the fast pace of research in this field slackened, but x-ray studies, especially by small-angle scattering (see ch. 12, § 5.2) have continued. An up-to-date outline of age-hardening in light alloys (aluminum, magnesium and titanium base) has been published by POLMEAR [1995].

Figure 49 shows the metastable solubility curves for the various intermediate phases as deduced from the researches in the 1950s and 1960s. GP zones and θ'' nucleate

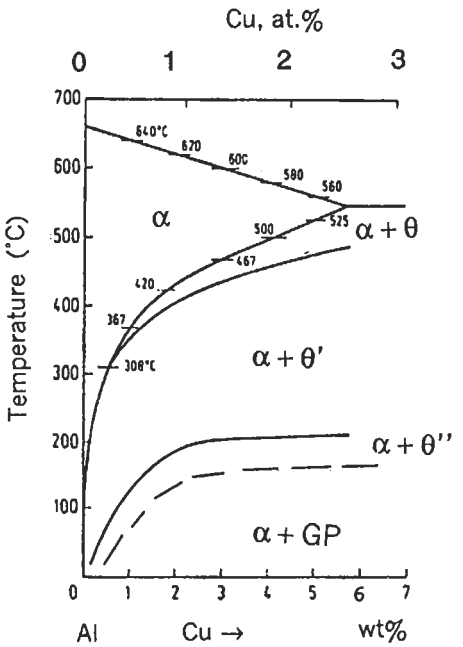


Fig. 49. Solubilities of copper in aluminum in the presence of various, mostly metastable phases as a function of temperature.

homogeneously, θ' forms preferentially at dislocations and the equilibrium phase, θ, at grain boundaries. The succession of stages follows Ostwald's (and Turnbull's) "rule of stages": the zones or particles of each stage undergo *reversion* (i.e., re-solution) before the next phase is formed; the initial presumption that each stage formed by atomic rearrangement of the earlier form proved to be mistaken. Reversion was studied in detail by experiments in which an age-hardening treatment at one temperature was interrupted by a shift to a higher or lower temperature (e.g., KÖSTER and KAM [1938]); the sudden shifts of hardness, later supplemented by structural studies, showed that structural reversion took place.

The kinetics of aging were studied not only by x-ray diffraction and hardness measurements but also by resistivity measurements, conducted by Turnbull and his associates in the late 1950s on Al–Cu alloys. The resistivity is closely linked to the process of formation of GP zones. These systematic studies were of great importance because they proved unambiguously that the rate of zone formation (or clustering, as they called it) at room temperature was orders of magnitude faster than would have been predicted by extrapolating measured diffusivities of copper in aluminum (DESORBO *et al.* [1958]), and the same thing was found for the age-hardening kinetics of Al–Ag (TURNBULL and TREAFTIS [1957]) and later of Al–Zn. The activation energy of the process clearly is that for vacancy migration, showing that there was a metastable excess of quenched-in vacancies which accounted for the rapid aging. The striking effect of these

References: p. 1818.

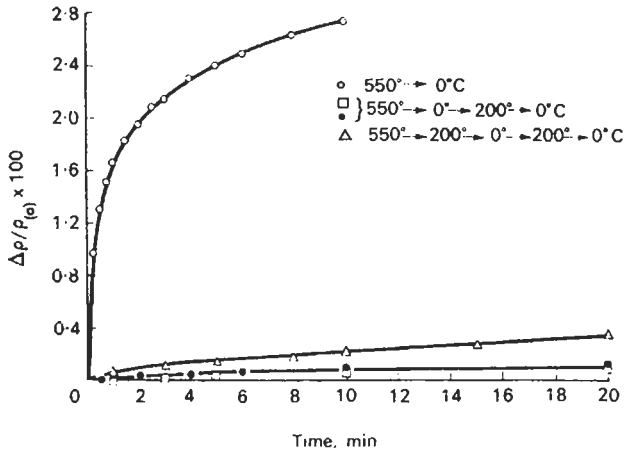


Fig. 50. Fractional change in resistivity *vs* aging time at 0°C for an Al-2at% Cu alloy after the indicated quenching sequences. (Holding times were 1 h at 0°C, 2 min at 200°C). (After TURNBULL *et al.* [1960]).

excess vacancies emerges from fig. 50, taken from a later study by TURNBULL *et al.* [1960]: here it is seen that a reversion treatment at 200°C, interspersed with aging at 0°C, greatly slows down aging kinetics because such reversion quickly gets rid of a large fraction of the excess vacancies (the authors believed that this happens because the reversion treatment creates dislocation loops which act as efficient vacancy sinks). Another feature of these studies was a comparison of aging after quenching into different fluids: the faster the quench, the more rapid the ultimate aging kinetics. This series of studies constitutes one of the great classics of physical metallurgy. (See also ch. 7, § 8.1).

Turnbull's experiments stimulated, after a long interval, some very interesting investigations of concentrated Al–Mg alloys, by VAN ROOYEN *et al.* [1988]. Using scanning differential calorimetry, these authors compared, using both melt-quenched and solid-quenched specimens, the kinetics of both the formation of GP zones at room temperature, and their subsequent redissolution during heating. (The melt-quenching took place by melt-spinning, producing 30 μm thick ribbons, the solid quenching was carried out by reusing these same ribbons). It was shown that in the solid-quenched specimens, both GP zone formation and GP zone redissolution were much faster than in the melt-quenched specimens. This was attributed to the greater efficacy of the solid-state quench in "freezing" in excess vacancies. That, in turn, is no doubt due to the greater effectiveness of a quench in water in maintaining a fast cooling rate down to quite modest temperatures, whereas melt-spinning leads to very fast solidification but slow cooling in contact with air after solidification is complete. This reinforces the message, conveyed in § 3.1, that a quenching method must take into account the measured cooling rate *in the temperature range of interest,* and is in harmony with the observation cited there, that in-rotating-water wire quenching is more effective in maintaining fast cooling in the solid state than is melt-spinning.

In recent years, research has been focused on the Al–Zn alloys, mostly by means of the small-angle scattering of x-rays or neutrons (see also ch. 12, § 5.2). These alloys

decompose by a spinodal mechanism (see ch. 15, § 3.1) below a well defined spinodal temperature, though the similarity of aging kinetics above and below this temperature (HENNION *et al.* [1982]; GUYOT and SIMON [1982]) indicate that the spinodal singularity does not play any substantial part in the decomposition kinetics. Many other relevant references, as well as a detailed explanation of spinodal decomposition as observed in Al–Zn alloys, can be found in a book chapter by BINDER [1991].

5.2. General characteristics of rapid-solidification-processed (RSP) crystalline alloys

The principal changes which can be brought about in crystalline alloys consequent on RSP have already been outlined in § 1.1. They include:

(1) Extension of solid-state solubility;
(2) refinement of grain size, with possible modification of grain shapes and textures;
(3) reduction or elimination of microsegregation;
(4) formation of metastable phases;
(5) achievement of high point-defect concentrations.

All of these have been surveyed in a general review by HONEYCOMBE [1978], and various aspects, grain refinement and metastable phase formation in particular, are treated elsewhere in this chapter.

We have already seen, in § 2.1, how thermodynamic and kinetic considerations, the variation of T_0 with composition in particular, affect solute trapping — that is, metastable extension of solid solubility. Aluminum alloys have been far more studied than any others, especially those with transition-metal solutes, and fig. 51, taken from MIDSON and JONES [1982], summarizes available results. It is plain that the solubility enhancement factor is in general much larger for solutes with a small equilibrium solubility. (This is reminiscent of the inverse correlation between intensity of segregation at grain boundaries and equilibrium solubility, see ch. 13, § 4.2.) Midson and Jones analyze these results in terms of the morphological criterion and Jones has also sought to interpret the lattice parameters of some greatly supersaturated solid solutions in terms of rival physical models (JONES [1983]). An early but substantial collection of experimental data on extended solubility in many different solvents was assembled by ANANTHARAMAN and SURYANARAYANA [1971].

WOOD and HONEYCOMBE [1981] review the available information on extension of solubility in iron-base alloys. Their figures together with some additional ones assembled by CANTOR [1982] are collected in table 7. In addition, C solubility in martensites is enhanced up to 2.0 wt% and in austenites up to 3.5 wt% before cementite is detected; but results are very variable and dependent on the measurement technique used. Boron, which is almost insoluble in equilibrium in pure α-iron (≈ 50 ppm) was forcibly dissolved by RSP up to 1 wt %; as with aluminum alloys, it appears that the least soluble elements have the highest enhancement factor.

Extended metastable solid solubility of course leads to greater volume fractions of pre-precipitation structures (GP zones) and consequently enhanced hardening. Much of the work on this has been purely empirical and there is rather a dearth of basic studies of pre-precipitation and precipitation mechanisms in RSP alloys. The rejection of solute

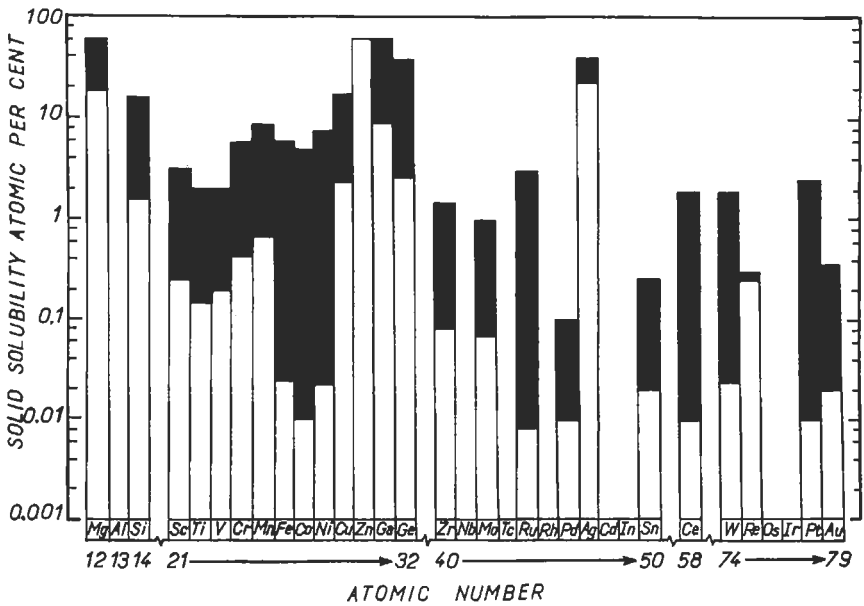


Fig. 51. Solubility extensions of various solutes in aluminum. The black bars indicate the increase beyond equilibrium room-temperature solubilities achieved by rapid solidification processing. (After MIDSON and JONES [1982]).

from the matrix and the identity and density of zones and phases need to be studied and compared with those for solid-quenched alloys of the same type. The rejection of solute from supersaturated RSP Al-12 at% Cu alloy has been studied by KRISHNANAND and CAHN [1976] and the decomposition products of RSP Al–17 at% Cu by SCOTT and LEAKE [1975]. The behavior of supersaturated RSP Al–8 wt % Fe alloys was examined by JACOBS *et al* [1974], Al–Fe–Cr alloys by YEARIM and SHECHTMAN [1982], while a range of supersaturated Al–Zr alloys were examined in detail by SAHIN and JONES [1978]

Table 7
Extensions of solubility for substitutional solutes in rapidly solidified binary iron-base alloys.

Solute	Max. equil. solubility (at %)	Extended solubility (at %)
Cu	7.2	15.0
Ga	47.0	50.0
Ge	20.0	25.0
Mo	26.0	40.6
Rh	50.0	100.0
Sn	9.8	20.0
Ti	9.8	16.0
W	13.0	20.8, 18.0

and HORI *et al.* [1982].

It is important to emphasize that to obtain the full benefits of improved mechanical properties from RSP of alloys, it is necessary to develop special solute-rich compositions and not to restrict oneself to standard compositions used for normal casting.

One of the most common consequences of RSP is refinement of microstructure, including such features as dendrite arm spacings, eutectic lamellar spacings, particle dispersions and grain size. When there is solute partitioning, rapid solidification leads to shorter diffusion distances and thereby to a refined structure. Even without partitioning, however, for example in pure metals, there is refinement of grain size. Sizes in the range 1 to 5 μm are quite common, and in splat-quenched foils can be as small as 0.01 μm. The refinement of grain size by RSP may have a number of distinct origins (quite apart from refinement brought about by the addition of inoculant particles). The possible origins have been reviewed by GREER [1991]. As shown in fig. 34, solidification at high undercooling should favor homogeneous nucleation and yield finer grains. BOSWELL and CHADWICK [1977] attempted thus to rationalize micrograin sizes in aluminum. In general, however, recalescence precludes grain refinement occurring in this way, except in glass-forming alloys near T_g. After rapid quenching, for example by melt-spinning, the grains are not necessarily equiaxed. Melt-spun ribbons have a have a fine equiaxed chill zone at the surface next to the quenching wheel, but this is followed by columnar growth, with column width increasing through the thickness of the ribbon because of competitive growth of differently oriented grains. The initial grain size at the wheel-side of the ribbon is governed by the heterogeneous nucleation frequency and the crystal growth rate. A simple analysis based on the temperature dependence of these quantities suggests that the grain size should be inversely proportional to cooling rate, in good agreement with experiment (GREER [1991]). Rapid solidification also occurs in highly undercooled bulk melts, and it is known that grain-refined structures are obtained for undercoolings beyond a critical value (as seen, for example, in the work of KATTAMIS and FLEMINGS [1970]) (see also ch. 8, § 11.2). There has been much investigation of this phenomenon, and it appears that it is not due to copious nucleation in the melt, but rather to a break-up of the initial dendrites (Schwarz *et al.* [1994]).

Consequences of a fine grain size include substantial "Petch-strengthening" at ambient temperature (DUFLOS and CANTOR [1978]) as well as an enhanced liability to superplasticity in RSP alloys (e.g., CHEESE and CANTOR [1980]), in the sense that superplasticity persists to higher strain rates. Quite generally, fine grains enhance hot-formability and thus RSP alloys are easier to shape as well as often having better mechanical properties after forming and heat-treatment, than slowly solidified analogues.

Fine grains also have consequences for martensitic transformations in steels and for creep in superalloys; these are discussed below.

Few studies have been reported of preferred orientations in as-cast structures produced by RSP. BLAKE *et al.* [1982] found remarkably perfect alignment of (0001) planes of Zn and Cd parallel to the surface of melt-spun ribbons; several other metals formed weaker textures of various types. The texture of high-purity (99.999%) zinc was found to deteriorate with time, and this was tentatively attributed to selective grain growth at ambient temperature. KAN *et al.* [1982] have melt-spun ribbons of highly

References: p. 1818.

alloyed Fe–Si (6.5 wt% Si, unlike the usual 3 wt%) and examined the preferred orientation. Grains with (100) [0*kl*] orientations grew preferentially to give a sharp planar texture, leading to much reduced alternating current losses. The texture-sharpening is no doubt an instance of surface-controlled grain growth (ch. 28, § 4.4.2).

The general technological implications of the combination of extended solubility, microcrystallinity and reduced segregation have been lucidly reviewed by GRANT [1978]. KÖSTER and SCHÜNEMANN [1993] provide a broad-based overview of phase transformations in rapidly solidified alloys.

5.3. RSP light alloys

A particularly important practical objective is the long-standing requirement of substantially increasing the elastic stiffness of aluminum alloys for airframe applications. Elastic moduli are not structure-sensitive and so this objective is best approached by introducing large amounts of stiff intermetallic compounds; initially, those based on manganese were preferred. When this was first attempted (DUDZINSKI *et al.* [1947–48]), the segregation associated with the slow freezing used led to coarse, uneven particle distributions and the alloys were hopelessly brittle. Once RSP methods had become established, numerous investigators recognized that this approach was the answer to the original difficulties. It should be noted here that while there is evidence that melt-quenching traps fewer excess vacancies than does solid-state quenching (§ 5.1), yet melt-quenching is more effective in permitting large supersaturations of solute; this is helped by fast cooling through the freezing temperature.

One company began using a form of planar flow-casting to produce alloy ribbons which were then chopped into small fragments and compacted by hot extrusion or by one of several rapidly acting high-temperature processes, such as impact forging or explosive consolidation, which restrict or eliminate structural coarsening during compaction (ADAM and LEWIS [1985]). These authors, for instance, give elastic properties of hot-extruded Al–(8–12)Fe–2V alloys: the Young's modulus for pure aluminum is 70 GPa, and the alloying enhanced this value by up to 37% without substantial embrittlement. The enhancement is due to a combination of enhanced solubility and fine dispersion of intermetallic compounds. Since the 1985 account, development of this kind of alloy has continued and further improvements of properties has been achieved (DAS and FROES [1993]). Thus, silicon has been added to make Al–Fe–V–Si alloys, in which the chief dispersoid is $Al_{13}(Fe, V)_3Si$. Figure 52 shows the high-temperature strength of different members of this family, with different contents of this dispersoid. The top plot refers to an alloy with a stiffness almost 40% up from that of pure aluminum, though its fracture toughness is only moderate; the bottom curve refers to a somewhat weaker alloy which is highly formable; it has been drawn into wire.

Enhanced stiffness is a special requirement of modern airframe aluminum alloys, but so is strength, notably at elevated temperatures, because of aerodynamic heating in flight). Figure 53, from Adam and Lewis's 1985 review, shows the success which had already been achieved 10 years ago in attaining this objective. All these advanced alloys contain iron: ever since Thursfield and Stowell's 1974 study, this has been the primary

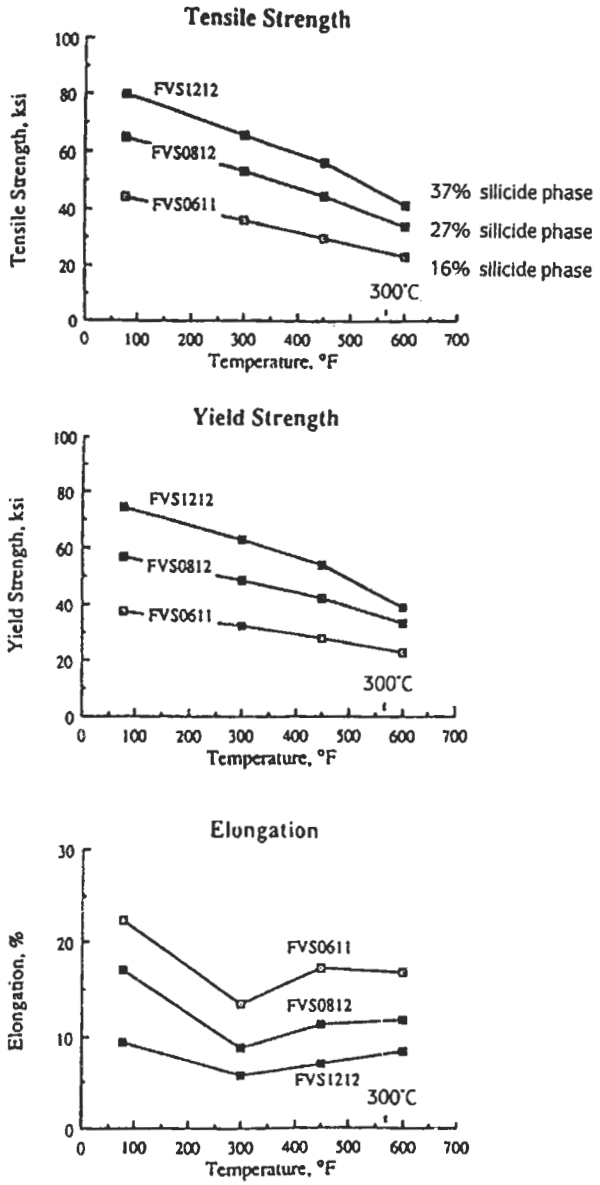


Fig. 52. Tensile properties of RSP and consolidated Al–Fe–V–Si alloys as a function of temperature. (After DAS and FROES [1993]).

solute of choice (iron has a small equilibrium solubility and a large enhancement of metastable solubility through RSP). In the figure, 70775–T6 and 2219–T851 are non-RSP high-grade alloys, and the "AFWAL program goal" was a military objective of that time. It can be seen that, even then, the new alloys came close to matching the state-of-the-art

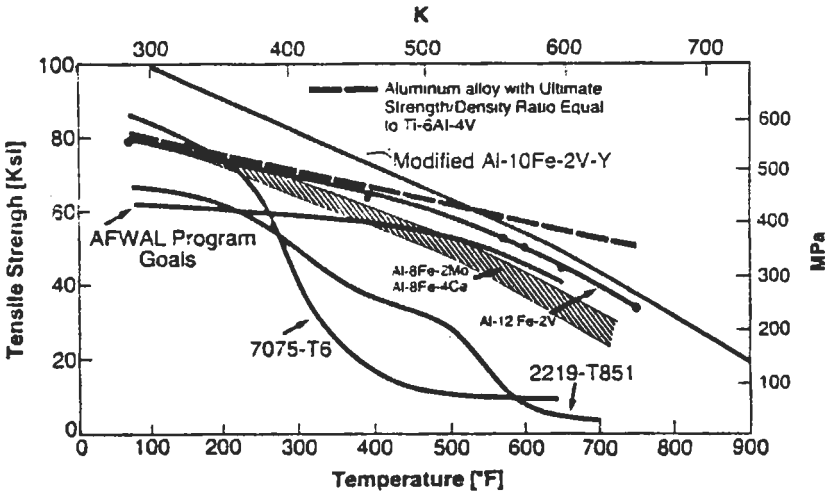


Fig. 53. Elevated-temperature strengths of advanced aluminum alloys. (After ADAM and LEWIS [1985]).

aeronautic titanium alloy, Ti–6Al–4V. A notable feature of such alloys, arising from their microstructural homogeneity, is their good corrosion resistance.

For the range of alloys based on Al–Li, which has received intense attention in the past decade because of the prospect of much enhanced specific stiffness (lithium reduces density and enhances stiffness), RSP has been extensively explored since the early work of SANKARAN and GRANT [1980], and at present there is still much dispute as to the relative benefits of standard ingot processing and RSP. It has been shown that RSP allows larger lithium contents than does ingot metallurgy (5 wt% as opposed to 2.5) without causing embrittling coarse dispersoids (MESCHTER *et al.* [1988]). Also, RSP alloys show less anisotropy than do those made by ingot metallurgy. However, the economics of the process may rule it out. A recent overview of Al–Li-based alloys (GRIMES [1990]) asserts that early worries about the adequacy of the ingot route and claims that powder metallurgy was essential, had largely vanished by 1990.

Das and Froes' overview of 1993 also covers the striking progress that has been made with RSP magnesium-base alloys: these alloys are always hovering on the verge of large-scale application and then do not make progress, largely because of corrosion problems. RSP alloys have improved corrosion properties as wĕll as improved strength, and their prospects may now improve. The overview also covers the considerable work done with RSP titanium-base alloys. The review must be consulted for further details.

5.4. Steels

The quite extensive early work on microstructure of RSP steels has been fully reviewed by WOOD and HONEYCOMBE [1981]. Extension of solid solubilities in iron was outlined in § 5.2. The findings concerning the precipitation morphology of various

carbides, both in simplified "model" steels and in commercial compositions, do not lend themselves to summary and the review should be consulted.

The principal driving force for research in this field has been the desire to improve the characteristics of tool ("high speed") steels. Much of the research has centred on simplified model alloys such as Fe–W–C. The most striking finding (see for instance SARE and HONEYCOMBE [1978] and RAYMENT and CANTOR [1981]) is the fact that for the fastest quenches and for solute-rich steels, δ-iron is retained; it is only at intermediate quench rates that austenite is retained in such steels, and at conventional cooling rates, martensite is the major phase. The kinetics of hardening during heat-treatment are not much affected though the pattern of carbide precipitation is refined and, in particular, carbide segregation, a serious problem, is minimised (KASAK and DULIS [1978]). The peak hardness is however raised for optimum treatments (RAYMENT and CANTOR [1978]). Practical developments based on RSP of tool steels are surveyed by JONES [1982].

In scientific terms, one of the most striking sets of findings concerns the modification of the martensitic transformation by RSP. INOKUTI and CANTOR [1976] found a variation of microhardness values for a series of RSP Fe–Ni alloys, as shown in fig. 54a; values for the solid-quenched alloys are shown for comparison. This was traced to a very pronounced change in martensite start (M_s) temperatures on cooling (fig. 54b) This curve, and also the $\alpha \rightarrow \gamma$ start temperatures on reheating, were obtained calorimetrically: the martensitic alloy was heated slowly in a DSC until it had all transformed back to γ and then slowly cooled again; M_s ($\gamma \rightarrow \alpha$) was marked by a pronounced exothermic peak, $\alpha \rightarrow \gamma$ by an endotherm (INOKUTI and CANTOR [1982]). The striking feature was that the depressed M_s caused by the original rapid quench from the melt does *not* disappear on repeated cycling between α and γ; only a prolonged γ anneal restores the alloy to normal behavior (labelled "solid quench" in fig. 54b). It is therefore not the rapid cooling per se which reduces the Ms, so it can only be the austenite structure resulting from that quench. In fact, various earlier studies not involving RSP have shown a relationship between austenite grain size, d, and M_s (e.g., for Fe–32% Ni, $\Delta M_s = 40$–100 K for d reduced from 200 to 10 μm). In the RSP alloys, d is only 4–5 μm and this leads to a drop in M_s of up to 200K for Fe–17% Ni. Thermal cycling, via the reverse shear transformation, restores the fine austenite grains, and only a long γ anneal gives the γ grains a chance to grow substantially. Such austenite grain growth seems often to be particularly sluggish in RSP steels (SUGA *et al.* [1980]); a micrographic study suggested that stable, extremely fine sulphide dispersions at austenite grain boundaries in RSP steels are responsible for such retardation of grain growth. (KELLY and VANDER SANDE [1980]). A slight reduction in the cooling rate (substituting melt-spinning for piston-and-anvil quenching) generated massive transformation instead of martensite (HAYZELDEN and CANTOR [1984–85]). A study of Fe–Ni–C alloys, with lower martensite start temperatures (HAYZELDEN and CANTOR [1982], using a scanning calorimetry approach, similarly showed M_s depressed by a very fine grain size generated by RSP.

Figure 29a shows that pure iron can also form martensite. The structure, formation conditions and crystallography of this martensite were studied by DUFLOS and CANTOR [1978]; hardness goes up with martensite/ferrite ratio, and for 100% martensite, the microhardness of pure iron is 600–700! A Hall-Petch plot (ch. 17, § 5.1) for pure iron is

References: p. 1818.

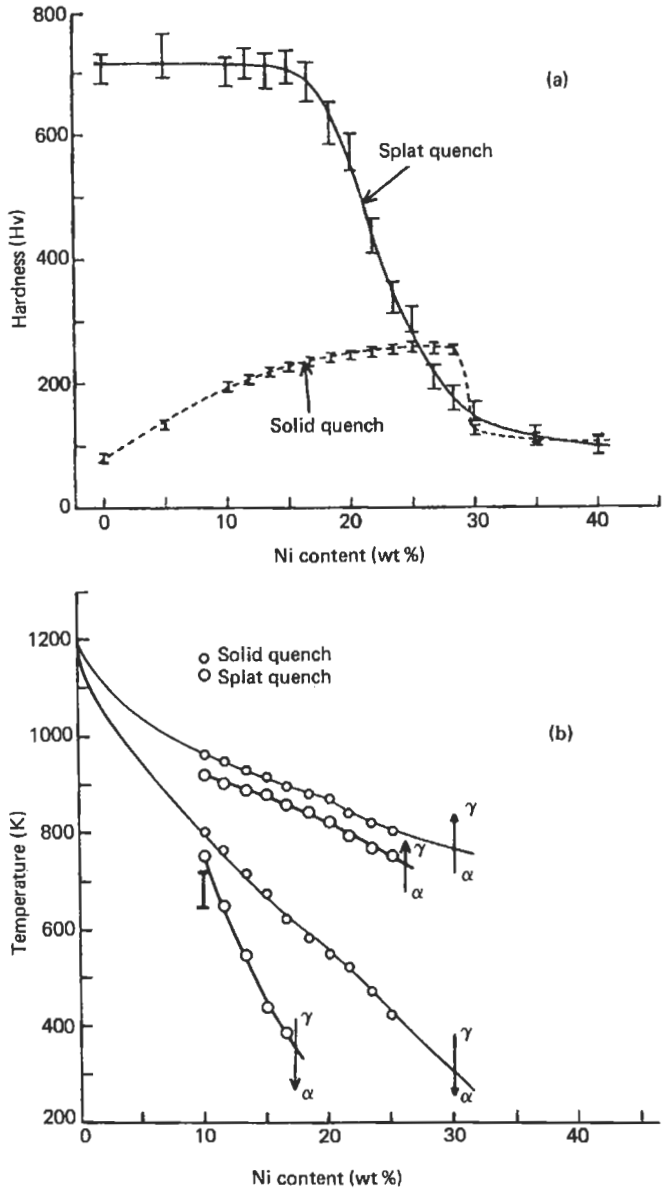


Fig. 54. (a) Microhardness values for liquid-quenched and solid-quenched Fe–Ni alloys. (b) $\gamma \rightarrow \alpha$ (M_s) and $\alpha \rightarrow \gamma$ transformation temperatures for a range of Fe–Ni alloys. (Both after INOKUTI and CANTOR [1982]).

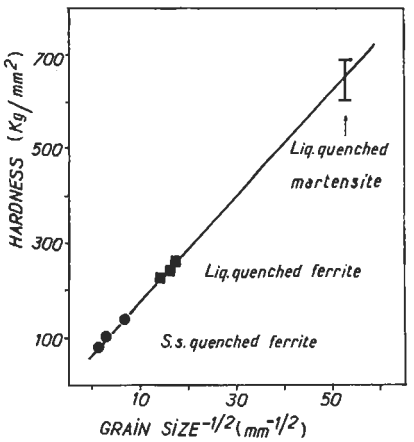


Fig. 55. Microhardness *vs* (grain size)$^{-1/2}$ for pure iron. (After DUFLOS and CANTOR [1978]).

shown in fig. 55, taking the thickness of a martensite lath as "grain size". It can be seen that the remarkably high hardness of martensitic iron can be largely, or even wholly, attributed to "grain-size strengthening". Recently, the limit of applicability of the Hall–Petch straight-line plot for very small grain sizes, because of insufficient space for dislocation pileups, has been discussed by Smith *et al.* [1995]; clearly, a grain size of about 0.35 μm is still within that limit.

A study by MAWELLA *et al.* [1982] of a melt-spun ultra-strong steel indicated that the considerable strength increase generated by melt-spinning was primarily due to refinement of the martensite structure. It has also been established that high-carbon cast iron (2.4–3 wt% C) following RSP by centrifugal atomization consists largely of retained austenite. Subsequent annealing produces a microcrystalline (α + carbide) structure. Even at 650°C, this microcrystalline structure is so stable that the material behaves superplastically (EISELSTEIN *et al.* [1983]). The conversion of brittle cast iron into a superplastic material is a particularly striking indicator of the changes in metallurgical characteristics obtainable by RSP.

5.5. Superalloys

Apart from tool steels and aluminum alloys, the other principal category of alloys to which RSP has been extensively applied is that of the superalloys, both those used for turbine discs and those destined for blades. The two Reston Conferences, in 1977 and 1979, are the principal source of information up to now on these studies; some other publications are by TRACEY and CUTLER [1981] and by SMUGERESKY [1982]. JONES [1982] also provides a lucid summary.

Most investigators up to now have used atomized powders, especially those made by centrifugal atomization in helium. However, DUFLOS and STOHR [1982] have demon-

strated that melt-spinning can provide higher quench rates and thus comminuted melt-spun ribbon provides an alternative material for subsequent hot consolidation, most commonly done by hot isostatic pressing. The benefits of RSP include the possibility of forcing more solute into the γ and γ' phases, producing alloys fine-grained and homogeneous enough to permit easy hot-working, even superplastic forming in extreme cases, and uniformity of structure throughout a large piece. The microstructural implications of a change from gas atomization to melt-spinning of superalloys have been examined in some detail by LIEBERMANN *et al.* [1983]. SHOHOJI *et al.* [1982] have also examined the structure of melt-spun superalloys and estimated cooling rates.

INOUE *et al.* [1983b] melt-quenched a number of superalloys consisting of γ' only. They found that these RSP alloys, instead of being almost perfectly ordered with large antiphase domains, were imperfectly ordered with APDs well below 0.1 mm in diameter, and were free of grain-boundary segregation. These structural modifications lead, in several alloys, to greatly enhanced ductility and some increase in strength. However, this much-cited research has recently been called into question, because the authors used melt-spun ribbons for microstructural examination and INROWASP-wires for mechanical testing, presuming that both forms had similar microstructures. CHIBA and HANADA [1996] have now demonstrated that the wires had quite a different microstructure from the ribbons, apparently because of slower cooling through the freezing-temperature.

There is, however, a major problem. The fine grain sizes produced by RSP, while they facilitate forming, leave the alloy very liable to creep at high temperatures but low stresses. This can be overcome by a process of *directional recrystallization* of the RSP alloy after it has been consolidated and formed: the process consists of a slow passage through a steep thermal gradient, to form large elongated grains rather like those generated by directional solidification. However, as SMUGERESKY [1982] indicates, the prolonged period spent at high temperatures during consolidation, forming and directional recrystallization may damage the desirable phase structure generated by RSP, and this problem is attracting further attention. The technique is currently applied primarily to oxide-dispersion-strengthened superalloys like MA 6000 (which contains 1.1 wt% Y_2O_3). Such alloys cannot be turned into coarse or single crystals by directional solidification, and thus directional recrystallization is the solution of choice (ROBERTS [1993]). The mechanism of this process has been examined in detail by BALOCH and BHADESHIA [1990], and computer simulation has recently cast further light on it (GODFREY and MARTIN [1995]).

References

ADAM, C. M. and R. E. LEWIS, 1985, in: Rapidly Solidified Crystalline Alloys, eds. S. K. Das, B. H. Kear and C. M. Adam (The Metallurgical Society, Warrendale, PA), p. 157.

AGYEMAN, K., R. MULLER and C. C. TSUEI, 1979, Phys. Rev. **B19**, 193.

AKHTAR, D. and R. D. K. MISRA, 1986, Scripta Metall. **20**, 627.

AL BIJAT, S., R. IRALDI, J. M. DUBOIS, G. LE CAËR and C. TÈTE, 1982, in: Proc. 4th Int. Conf. on Rapidly Quenched Metals. Eds. T. Masumoto and K. Suzuki (Japan Institute of Metals, Sendai) p. 375.

ANANTHARAMAN, T. R. and C. SURYANARAYANA, 1971, J. Mater. Sci. **6**, 1111.

ANGELL, C. A., 1995, Science **267**, 1924.

AOKI K., X. G. LI and T. MASUMOTO, 1991, Mater. Sci. Eng. **A133**, 565. (7th Conf. RQM).
ARMSTRONG, G. R. and H. JONES, 1979, in: Solidification and Casting of Metals (The Metals Society, London; now Institute of Materials), p. 456.
ATZMON, M., K. M. UNRUH and W. L. JOHNSON, 1985, J. Appl. Phys. **58**, 3865.
BALANZAT, E., J. T. STANLEY, C. MAIRY and J. HILLAIRET, 1985, Acta Metall. **33**, 785.
BALOCH, M. M. and H. K. D. H. BHADESHIA, 1990, Mat. Sci. Techn. **6**, 1236.
BARBER, B. P. and S. J. PUTTERMAN, 1992, Phys. Rev. Lett. **69**, 3839.
BATTEZZATI, L., 1990, Phil. Mag. **B61**, 511.
BATTEZZATI, L. and A. L. GREER, 1987, Int. J. Rapid Solidification **3**, 23.
BATTEZZATI, L. and A. L. GREER, 1989, Acta Metall. **37**, 1791.
BAXTER, D. V., 1986, J. Non-Cryst. Solids **79**, 41.
BECHET, D. and G. REGAZZONI, 1991, Mater. Sci. Eng. **A134**, 1120.
BEGHI, G., R. MATERA and G. PIATTI, 1969, J. Nucl. Mater. **31**, 259.
BEKE, D. L., P. I. LOEFF and H. BAKKER, 1991, Acta Metall. et Mater. **39**, 125, 1267.
BENJAMIN, J. S., 1970, Metall. Trans. **1**, 2943.
BERKOWITZ, A. E. and J. L. WALTER, 1987, J. Mater. Res. **2**, 275.
BERNAL, J. D., 1964, Proc. R. Soc. (Lond.) **A284**, 299.
BEVK, J., 1983, Annu. Rev. Mater. Sci. **13**, 319.
BEWLAY, B. P. and B. CANTOR, 1986, Int. J. Rapid Solidification **2**, 107.
BEWLAY, B. P. and B. CANTOR, 1990, Metall. Trans. **21B**, 899.
BINDER, K., 1991, in: Phase Transformations in Materials, ed. P. Haasen, Vol. 5 of Materials Science and Technology, eds. R. W. Cahn, P. Haasen and E. J. Kramer (VCH, Weinheim) p. 405.
BLAKE, N. W., F. A. HAMES and R. W. SMITH, 1982, in: Rapidly Solidified Amorphous and Crystalline Alloys, Proc. Mater. Research Soc. Annual Meeting, Boston, 1981, eds. B. H. Kear, B. C. Giessen and M. Cohen (North-Holland, Amsterdam) p. 363.
BLANK-BEWERSDORFF, M., 1991, J. Mater. Sci. Lett. **10**, 1225.
BORDEAUX, F. and A. R. YAVARI, 1990a, J. Phys. France **51**, C-249.
BORDEAUX, F. and A. R. YAVARI, 1990b, J. Appl. Phys. **67**, 2385.
BORDERS, J. A., 1979, Annu. Rev. Mater. Sci. **9**, 313.
BORMANN, R., 1994, Mater. Sci. Eng. **A179/A180**, 31. (Proc. 8th Int. Conf. on Rapidly Quenched and Metastable Materials, Sendai, 1993).
BORMANN, R. and K. ZÖLTZER, 1992, Phys. Stat. Sol. (a) **131**, 691.
BORMANN, R., F. GÄRTNER and K. ZÖLTZER, 1988, J. Less-Comm. Metals **145**, 283.
BOSWELL, P. G. and G. A. CHADWICK, 1977, Scripta Metall. **11**, 459.
BOUDREAUX, D. S., and J. M. GREGOR, 1977, J. Appl. Phys. **48**, 152 and 5057.
BRAUN, J., M. ELLNER and B. PREDEL, 1994, Z. Metallkde. **85**, 855.
BRENNER, A.. 1963, Electrodeposition of Alloys: Principles and Practice, 2nd edition (Academic Press, New York).
BRESSON, L., J. P. CHEVALIER and M. FAYARD, 1982, Scripta Metall. **16**, 499.
BUCKEL, W. and R. HILSCH, 1956, Z, Phys. **146**, 27.
BUCKLOW, I. A. and L. E. DRAIN, 1964, J. Sci. Instr. **41**, 614.
BUSCHOW, K. H. J., 1979, J. Less-Common Metals **66**, 89.
BUSCHOW, K. H. J., 1982, Solid State Commun. **43**, 171.
BUSCHOW, K. H. J., 1986, Mater. Sci. Reports **1**, 1.
CACIUFFO, R., M. STEFANON, W. S. HOWELLS, A. K. SOPER, P. ALLIA, F. VINAI, S. MELONE and F. RUSTICHELLI, 1989, Physica **B156 & 157**, 220.
CAHN, R. W., 1979, in: Alliages et Matériaux Amorphes. 21ème Colloque de Métallurgie, Saclay, 1978 (Centre d'Etudes Nucléaires de Sacley. Gif-sur-Yvette) p. 137.
CAHN, R. W., 1986, J. Vac. Sci. Technol. **A4**, 3071.
CAHN, R. W., 1990, in: Basic Features of the Glassy State, eds. J. Colmenero and A. Alegria (World Scientific, Singapore), p. 282.
CAHN, R. W., 1992, Nature **357**, 645.
CAHN, R. W., 1993, in: Rapidly Solidified Alloys, ed. H. H. Liebermann (Dekker, New York), p. 1.

CAHN, R. W. and W. L. JOHNSON, 1986, J. Mater. Res. **1**, 724.
CAHN, R. W., J. L. WALTER and D. W. MARSH, 1988, Mater. Sci. Eng. **98**, 33.
CAHN, R. W., K. D. KRISHNANAND, M. LARIDJANI, M. GREENHOLZ and R. HILL, 1976, Mater. Sci. Eng. **23**, 83.
CAHN, R. W., N. A. PRATTEN, M. G. SCOTT, H. R. SINNING and L. LEONARDSSON, 1984, in: Mater. Res. Soc. Symp. Proc. **28**, 241.
CALKA, A., M. MADHAVA, D. E. POLK, B. C. GIESSEN, H. MATYJA and J. B. VANDER SANDE, 1977, Scripta Metall. **11**, 65.
CANTOR, B., 1982, Rapidly Solidified Amorphous and Crystalline Alloys, Proc. Mater. Research Soc. Annual Meeting, Boston, 1981, eds. B. H. Kear, B. C. Giessen and M. Cohen (North-Holland, Amsterdam), p. 317.
CARGILL, G. S., 1970, J. Appl. Phys. **41**, 12.
CARTER, G. and W. A.. GRANT, 1982, Nucl. Instr. and Meth. **199**, 17.
CHASON, E. H., A. L. GREER, K. F. KELTON, P. S. PERSHAN, L. SORENSEN, F. SPAEPEN and A. H. WEISS, 1985, Phys. Rev. **B32**, 3399.
CHASON, E. H. and T. MIZOGUCHI, 1987, Mat. Res. Soc. Symp. Proc. **80**, 61.
CHATTOPADHYAY, K. and P. RAMACHANDRARAO, 1980, J. Mater. Sci. **15**, 685.
CHAUDHARI, P., J. F. GRACZYK, O. HENDERSON AND P. STEINHARDT, 1976, Phil. Mag. **31**, 7272.
CHAUDHARI, P., F. SPAEPEN and P. J. STEINHARDT, 1983, in: Glassy Metals II, eds. H. Beck and H.-J. Güntherodt (Springer, Berlin) p. 127.
CHEESE, R. and B. CANTOR, 1980, Mater. Sci. Eng. **45**, 83.
CHEN, H. S. and C. E. MILLER, 1970, Rev. Sci. Instr. **41**, 1237.
CHEN, H. S., J. T. KRAUSE, A. INOUE And T. MASUMOTO, 1983, Scripta Metall. **17**, 1413.
CHEN, H., Y. HE, G. J. SHIFLET and S. J. POON, 1991, Scripta Metall. Mater. **25**, 1421.
CHEN, L. C. and F. SPAEPEN, 1991, J. Appl. Phys. **69**, 679.
CHIBA, A. and S. HANADA, 1996, Intermetallics **4** (in press).
CLEMENS, B. M., W. L. JOHNSON and R. B. SCHWARZ, 1984, J. Non-Cryst. Solids **61&62**, 817.
CLYNE, T. W., 1984, Metall. Trans. **15B**, 369.
CONNELL, G. A. N., 1975, Solid State Commun. **16**, 109.
COWLAM, N., W. GUOAN, P. P. GARDNER and H. A. DAVIES, 1984, J. Non-Cryst. Solids **61-62**, 337.
CROAT, J. J., R. W. HERBST, R. W. LEE and F. E. PINKERTON, 1984, J. Appl. Phys. **55**, 2078.
CUNAT, Ch., H. R. HILZINGER and G. HERZER, 1988, Mater. Sci. Eng. **97**, 497.
DAHLGREN, S. D., 1978, in: Proc. 3rd Int. Conf. on Rapidly Quenched Metals, Brighton, ed. B. Cantor (Metals Society, London), vol 2, p. 36.
DAS, S. K. and F. H. FROES, 1993, in: Rapidly Solidified Alloys, ed. H. H. Liebermann (Marcel Dekker, New York), p. 339.
DAS, S. K., K. OKAZAKI and C. M. ADAM, 1985, in: High Temperature Alloys – Theory and Design, ed. J. O. Stiegler (The Metallurgical Society, Warrendale, PA) p. 451.
DAVIES, H. A., 1976, Phys. and Chem. of Glasses **17**, 159.
DAVIES, H. A., 1978, in: Proc. 3rd Int. Conf. on Rapidly Quenched Metals, Brighton, 1978, ed. B. Cantor (Institute of Metals, London), vol. 1, p. 1.
DAVIES, H. A., 1995, in: Nanostructured and Non-Crystalline Materials, eds. M. Vázquez and A. Hernando (World Scientific, Singapore), p. 3.
DAVIES, H. A., B. G. LEWIS and I. W. DONALD, 1978, in: Rapid Solidification Processing, Principles and Technologies. Proc. Conf. on Rapid Solidification Processin, Reston, 1977, eds. R. Mehrabian *et al.* (Claitor, Baton Rouge, LA), p. 78.
DE HOSSON, J.Th. M., 1983, in: Amorphous Materials: Modeling of Structure and Properties, ed. V. Vitek, (Met. Soc. AIME, Warrendale PA) p. 147.
DENG, D. and A . S. ARGON, 1986, Acta Metall. **34**, 2011.
DE REUS, R., 1995, in: Intermetallic Compounds, eds. J. H. Westbrook and R. L. Fleischer (Wiley, Chichester), Vol. 2, p. 603.
DE REUS, R., F. W. SARIS and T. S. BALLER, 1988, J. Less-Common Metals **145**, 621.
DESORBO, W., H. N. TREAFTIS and D. TURNBULL, 1958, Acta Metall. **6**, 401.
DE VRIES, J., G. W. KOEBRUGGE and A. VAN DEN BEUKEL, 1988, Scripta Metall. **22**, 637.

DIEGLE, R.B. and K. HASHIMOTO (eds.), 1988, Proc. Symposium on Corrosion, Electrochemistry and Catalysis of Metallic Glasses (The Electrochemical Society, Pennington).
DIETZ, G., 1977, J. Magn. Magn. Mater. **6**, 47.
DONALD, I.W., 1987, J. Mater. Sci. **22**, 2661.
DONALD, I.W. and A.H. DAVIES, 1978, J. Non-Cryst. Solids **30**, 77.
DONALD, I.W. and A.H. DAVIES, 1980, J. Mater. Sci. **15**, 2754.
DONG, X.Z., W. FERNENGEL and H. KRONMÜLLER, 1982, Appl. Phys. **A28**, 103.
DONOVAN, P.E. and W.M. STOBBS, 1981, Acta Metall. **29**, 1419.
DÖRNER, W., H. MEHRER, P. J. POKELA, E. KOWALA and M.-A. NICOLET, 1991, Mater. Sci. Eng. **B10**, 165.
DOYLE, B.L., P.S. PEERCY, J.D. WILEY, J.H. PEREPEZKO and J.E. NORDMAN, 1982, J. Appl. Phys. **5**, 6168.
DREHMAN, A.J., and A.L. GREER, 1984, Acta Metall. **32**, 323.
DUBEY, K.S. and P. RAMACHANDRARAO, 1984, Int. J. Rapid Solidif. **1**, 1.
DUBOIS, J.M., 1992, in: Trends in Non-Crystalline Solids, eds. A. Conde, C.F. Conde and M. Millán (World Science, Singapore) p. 343.
DUBOIS, J.M., P.H. GASKELL and G. LE CAËR, 1985a, in: Rapidly Quenched Metals (V), eds. S. Steeb and H. Warlimont (North-Holland, Amsterdam) p. 567.
DUBOIS, J.M., G. LE CAËR and K. DEHGHAN, 1985b, in: Rapidly Quenched Metals (V), eds. S. Steeb and H. Warlimont (North-Holland, Amsterdam) p. 197.
DUDZINSKI, N., J.R. MURRAY, B.W. MOTT and B. CHALMERS, 1947-48, J. Inst. Metals **74**, 291.
DUFLOS, B. and B. CANTOR, 1978, in: Proc. 3rd Int. Conf. on Rapidly Quenched Metals, Brighton, 1978, ed. B. Cantor (Institute of Metals, London), vol. 1, p. 110.
DUFLOS, B. and B. CANTOR, 1982, Acta Metall. **30**, 323.
DUFLOS, B. and B. CANTOR, 1987, J. Mater. Sci. **22**, 119.
DUFLOS, B. and J.F. STOHR, 1982, J. Mater. Sci. **17**, 3641.
DUINE, P.A., J. SIETSMA and A. VAN DEN BEUKEL, 1992, Acta Metall. Mater. **40**, 743.
DUINE, P.A., J. SIETSMA and A. VAN DEN BEUKEL, 1993, Phys. Rev. B **48**, 6957.
DUWEZ, P., 1967, Trans. Am. Soc. Metals **60**, 607.
DUWEZ, P., R.H. WILLENS and W. KLEMENT, 1960, Nature (London) **187**, 869; J. Appl. Phys. **31**, 1136 and 1500.
EGAMI, T., 1978, Mater. Res. Bull. **13**, 557.
EGAMI, T., 1978b, J. Mater. Sci. **13**, 2587.
EGAMI, T., 1981, in: Glassy Metals I: Ionic Structure, Electronic RTRANSPORT and Crystallization (Springer, Berlin) p. 25.
EGAMI, T., 1986, in: Amorphous Metals and Semiconductors, eds. P. Haasen and R.I. Jaffee (Pergamon, Oxford) p. 222.
EGAMI, T, 1994, Mater. Sci. Eng. **A179/A180**, 17.
ELLIOTT, R.O. and D.A. KOSS, 1981, J. Nucl. Mater. **97**, 339.
EISELSTEIN, C.E., O.A. RUANO and O.D. SHERBY, 1983, J. Mater. Sci. **18**, 483.
EVANS, P.V., A. GARCIA-ESCORIAL, P.E. DONOVAN AND A.L. GREER, 1987, Mater. Res. Soc. **57**, 239.
FECHT, H. and W.L. JOHNSON, 1988, Nature **334**, 50.
FECHT, H., P.J. DESRÉ and W.L. JOHNSON, 1989, Phil. Mag. B **59**, 577.
FINNEY, J.L., 1983, in: Amorphous Metallic Alloys, ed. F.E. Luborsky (Butterworth, London) p. 42.
FISCHER, M., D.E. POLK and B.C. GIESSEN, 1978, in: Proc. 1st Int. Conf. on Rapid Solidification Processing (eds. R. Mehrabian, B.H. Kear and M. Cohen (Claitor, Baton Rouge), p. 140.
FLINT, E.B. and K.S. SUSLICK, 1991, Science **253**, 1397.
FOLLSTAEDT, D.M., 1991, in: Processing of Metals and Alloys, ed. R.W. Cahn, Vol. 15 of Materials Science and Technology, eds. R.W. Cahn, P. Haasen and E.J. Kramer (VCH, Weinheim) p. 247.
GAGNÉ, H. and C. ROY, 1982, High Temp. Technology, **1**, 93.
GARCIA-ESCORIAL, A. and A.L. GREER, 1987, J. Mater. Sci. **22**, 4388.
GARRONE, E. and L. BATTEZZATI, 1985, Phil. Mag. **B52**, 1033.
GASKELL, P.H., 1979, J. Non-Cryst. Solids **32**, 207.
GASKELL, P.H., 1983, in: Amorphous Materials: Modeling of Structure and Properties, ed. V. Vitek (The Metallurgical Society, Warrendale, PA) p. 47.

GASKELL, P. H., 1985, in: Glass, Current Issues, eds. A. F. Wright and J. Dupuy (Martinus Nijhoff, Dordrecht) p. 54.
GASKELL, P. H. and D. J. SMITH, 1980, J. Microscopy **119**, 63.
GERLING, R., F. P. SCHIMANSKY and R. WAGNER, 1985, in: Rapidly Quenched Metals (V), eds. S. Steeb andH. Warlimont (North-Holland, Amsterdam) p. 1377; 1988, Acta Metall. **36**, 575; 1989, Acta Metall. **36**, 575; 1990, Int. J. Rapid Solidification **5**, 137.
GIBBS, M. R. J., J. E. EVETTS and J. A. LEAKE, 1983, J. Mater. Sci. **18**, 278.
GIESSEN, B. C., M. MADHAVA, D. E. POLK and J. B. VANDER SANDE, 1976, Mater. Sci. Eng. **23**, 145.
GIESSEN, B. C., 1981, Proc. 4th Int. Conf. on Rapidly Quenched Metals. Eds. T. Masumoto and K. Suzuki (Japan Institute of Metals, Sendai) p. 213.
GILLEN, A. G. and B. CANTOR, 1985, Acta Metall. **33**, 1813.
GODFREY, A. W., and J. W. MARTIN, 1995, Phil. Mag. **A72**, 737.
GRAHAM, Jr., C. D., 1959, in: Magnetic Properties of Metals and Alloys. Seminar volume. (Amer. Soc. Metals, Cleveland) p. 288.
GRÁNÁSY, L., I. EGRY, L. RATKE and D. M. HERLACH, 1994, Scripta Metall. Mater. **31**, 601.
GRANT, N. J., 1978, in: Rapid Solidification Processing, Principles and Technologies. Proc. Conf. on Rapid Solidification Processing, Reston, eds. R. Mehrabian *et al.* (Claitor, Baton Rouge, LA), p. 230.
GRANT, P. S. and B. CANTOR, 1990, Powder Metallurgy **33**, 144.
GRANT, P. S., W. T. KIM, B. P. BEWLAY and B. CANTOR, 1989, Scripta Metall. **23**, 1651.
GREER, A. L., 1982, Acta Metall. **30**, 171.
GREER, A. L., 1988, J. Less-Common Metals **140**, 327.
GREER, A. L., 1991, Mater. Sci. Eng. **A133**, 16.
GREER, A. L., 1993a, Nature **366**, 303.
GREER, A. L., 1993b, in: Rapidly Solidified Alloys, ed. H. H. Liebermann (Dekker, New York), p. 269.
GREER, A. L., 1994, Mater. Sci. Eng. **A178**, 113.
GREER, A. L. and J. A. LEAKE, 1979, J. Non-Cryst. Solids **33**, 291.
GREER, A. L. and F. SPAEPEN, 1981, in: Structure and Mobility in Molecular and Atomic Glasses, Annals N. Y. Acad. Sci. **371**, 218.
GREER, A. L., K. DYRBYE, L.-U. AAEN–ANDERSEN, R. E. SOMEKH, J. BØTTIGER and J. JANTING, 1990, Mat. Res. Soc. Symp. Proc. **371**, 3.
GREGAN, G. P. J., Y. D. DONG and M. G. SCOTT, 1981, in: Proc. Int. Conf. on Metallic Glasses, Vol. 2 (Kultura, Budapest, Hungary) p. 197.
GRIMES, R., 1990, in: Second Supplementary Volume to the Encyclopedia of Materials Science and Engineering, ed. R. W. Cahn (Pergamon, Oxford), p. 667.
GUINIER, A., 1938, Nature **142**, 569.
GUO, H. Q., H. KRONMÜLLER, N. MOSER and A. HOFMANN, 1986, Scripta Metall. **20**, 185.
GURMAN, S. J., 1985, in: Proc. 5th Int. Conf. on Rapidly Quenched Metals, eds. S. Steeb and H. Warlimont (North-Holland, Amsterdam) p. 427.
GUYOT, P. and J. P. SIMON, 1982, in: Dynamics of Ordering Processes in Condensed Matter, eds. S. Komura and H. Furukawa (Plenum, New York), p. 211.
HAFNER, J., 1981, in: Glassy Metals I, ed. H.-J. Güntherodt. (Springer, Berlin) p. 93.
HAFNER, J., 1983, Phys. Rev. B **28**, 1734.
HAFNER, J., 1986, in: Amorphous Metals and Semiconductors. Eds. P. Haasen and R. I. Jaffee (Pergamon, Oxford) p. 151.
HAGIWARA, M., A. INOUE and T. MASUMOTO, 1985, in: Rapidly Quenched Metals (V), eds. S. Steeb andH. Warlimont (North-Holland, Amsterdam) p. 1779.
HARBUR, D. R., J. W. ANDERSON and W. J. MARAMAN, 1969, Trans. Met. Soc. AIME **245**, 1055.
HASEGAWA, S. and R. RAY, 1978, J. Appl. Phys. **49**, 4174.
HASEGAWA, D. R. and L. E. TANNER, 1977, J. Appl. Phys. **49**, 3211.
HASHIMOTO, K., 1993, in: Rapidly Solidified Alloys, ed. H. H. Liebermann (Dekker, New York), p. 591.
HAYZELDEN, C., 1983, D. Phil. Thesis, Sussex University.

HAYZELDEN, C., and B. CANTOR, 1982, in: Proc. Int. Conf. on Solid→Solid Phase Transformations, Pittsburgh, 1981, eds. H.I. Aaronson, D.E. Laughlin, R.F. Sekerka and C.M. Wayman (The Metallurgical Society, Warrendale, PA), p. 1397.

HAYZELDEN, C. and B. CANTOR, 1984-85, Int. J. of Rapid Solidification **1**, 237.

HENNION, M., D. RONZAUD and P. GUYOT, 1982, Acta Metall. **30**, 599.

HERLACH. D.M., R.F. COCHRANE, I. EGRY, H.J. FECHT and A.L. GREER, 1993, Int. Mater. Rev. **38**, 273.

HERMAN, H. and H. BHAT, 1980, in: Synthesis and Properties of Metastable Phases, eds. E.S. Machlin and T. J. Rowland (The Metallurgical Society, Warrendale, PA) p. 115.

HERZER, G., 1991, Mater. Sci. Eng. **A133**, 1.

HIGASHI, K., T. MUKAI, S. TANIMURA, A. INOUE, T. MASUMOTO, K. KITA, K. OHTERA and J. NAGAHORA, 1992, Scripta Metall. Mater. **26**, 191.

HIGHMORE, R.J., 1990, J. de Physique France (Colloques) **51**, C-37.

HIGHMORE, R.J. and A.L. GREER, 1989, Nature **339**, 363.

HILLAIRET, J., C. MAIRY, E. BALANZAT, J. HAMANT and M. MASSON [1983], Radiation Eff. **70**, 131.

HILLENBRAND, H.-G., E. HORNBOGEN and U. KÖSTER, 1982, in: Proc. 4th Int. Conf. on Rapidly Quenched Metals, eds. T. Masumoto and K. Suzuki (Japan Institute of Metals, Sendai) p. 1369.

HILZINGER, H.R., 1982, in: Proc. 4th Int. Conf. on Rapidly Quenched Metals, eds. T. Masumoto and K. Suzuki (Japan Inst. of Metals, Sendai), vol. 2, p. 791.

HOLZER, J.C. and K.F. KELTON, 1991, Acta Metall. Mater. **39**, 1833.

HONEYCOMBE, R.W.K., 1978, in: Proc. 3rd Int. Conf. on Rapidly Quenched Metals, Brighton, 1978, ed. B. Cantor (Institute of Metals, London), vol. 1, p. 73.

HORI, S., S. SAJI and A. TAKEHARA, 1982, in: Proc. 4th Int. Conf. on Rapidly Quenched Metals, Sendai 1981, eds. T. Masumoto and K. Suzuki (Japan Inst. MetAls, Sendai), p. 1545.

HORVÁTH, J., J. OTT, K. PFAHLER and W. UHLFERT, 1988, Mater. Sci. Eng. **97**, 409.

HOSHINO, S., H.A. LAITINEN and G.B. HOFLUND, 1986, J. Electrochem. Soc. **133**, 681.

HOWE, L.M., 1995, in: Intermetallic Compounds: Principles and Practice, eds. J.H. Westbrook and R.L. Fleischer (Wiley, Chichester), Vol. 1, p. 791.

HUIZER, E. and A. VAN DEN BEUKEL, 1987, Acta Metall. **35**, 2843.

INOKUTI, Y. and B. CANTOR, 1976, Scripta Metall. **10**, 655.

INOKUTI, Y. and B. CANTOR, 1982, Acta Metall. **30**, 343.

INOUE, A., 1995, in: Proc. Int. Symp. on Metastable, Mechanically Alloyed and Nanocrystalline Materials (1994), Materials Science Forum, **179/180**, 691.

INOUE, A., and T. MASUMOTO, 1980, Sci. Rep. Res. Inst. Tohoku Univ., **28A**, 165.

INOUE, A., M. HAGIWARA and T. MASUMOTO, 1982, J. Mater. Sci. **17**, 580.

INOUE, A., M. HAGIWARA, T. MASUMOTO and H.S. CHEN, 1983, Scripta Metall. **17**, 1205.

INOUE, A., H. TOMIOKA and T. MASUMOTO, 1983b, Metall. Trans. **14A**, 1367.

INOUE, A., K. OHTERA, A.P. TSAI and T. MASUMOTO, 1988a, Japan J. Appl. Phys. **27**, L280.

INOUE, A., K. OHTERA and T. MASUMOTO, 1988b, Japan J. Appl. Phys. **27**, L736.

INOUE, A., T. ZHANG and T. MASUMOTO, 1989, Mater. Trans. Japan Inst. Met. **30**, 965.

INOUE, A., N. MATSUMOTO and T. MASUMOTO, 1990, Mater. Trans. Japan Inst. Met. 31, 493.

INOUE, A., Y. HORIO, Y.H. KIM and T. MASUMOTO, 1992, Mater. Trans. JIM **33**, 669.

INOUE, A., T. NAKAMURA, N. NISHIYAMA, T. SUGITA and T. MASUMOTO, 1993a, Key. Eng. Mater. **81-83**, 147.

JACOBS, M.H., A.G. DOGGETT and M.J. STOWELL, 1974, J. Mater. Sci. **9**, 1631.

JANKOWSKI, A.F., P. SANDOVAL and J.P. HAYES, 1995, Nonstruct. Mat. **5**, 497.

JAYARAMAN, A., 1980, in: Synthesis and properties of metastable phases, eds. E.S. Machlin and T.J. Rowland (Met. Soc. AIME, Warrendale, PA) p. 75.

JOHNSON, W.L., 1981, in: Glassy Metals I, eds. H.-J. GÜNTHERODT and H. BECK (Springer, Berlin) p. 191.

JOHNSON, W.L., 1982, in: Metastable Materials Formation by Ion Implantation, Proc. Mater. Res. Soc. Annual Meeting, Boston, 1981, eds. S.T. Picraux and W.J. Choyke (North-Holland, Amsterdam) p. 183.

JOHNSON, W.L., 1986, Progr. Mater. Sci. **30**, 81.

JOHNSON, W.L., S.J. POON and P. DUWEZ, 1975, Phys. Rev. B **11**, 150.

JONES, H., 1981, in: Ultrarapid Quenching of Liquid Alloys, Treatise on Materials Science and Technology, vol. 20, ed. H. Herman, (Academic, New York), p. 1.

JONES, H., 1982, Rapid Solidification of Metals and Alloys (Institution of Metallurgists, London; now Institute of Materials).
JONES, H., 1983, Scripta Metall. **17**, 97.
KAN, T., Y. ITO and H. SHIMANAKA, 1982, J. Magn. Magn. Mater. **26**, 127.
KASAK, A. and E. J. DULIS, 1978, Powder Metallurgy **21**, 114.
KATTAMIS, T. Z. and M. C. FLEMINGS, 1970, Metall. Trans. **1**, 1449.
KATTAMIS, T. Z., W. E. BROWER and R. MEHRABIAN, 1973, J. Cryst. Growth **19**, 229.
KAWASHIMA, A., K. ASAMI, T. SATO and K. HASHIMOTO, 1985, in: Rapidly Quenched Metals (V), eds. S. Steeb and H. Warlimont (North-Holland, Amsterdam) p. 1671.
KEAR, B. H., J. N. MAYER, J. M. POATE and P. R. STRUTT, 1981, in: Metallurgical Treatises, eds. J. K. Tien and J. F. Elliott (TMS-AIME, Warrendale, PA), p. 321.
KELLY, A. and R. B. NICHOLSON, 1963, Progr. Mater. Sci. **10**, 149.
KELLY, T. F. and J. B. VANDER SANDE, 1980, Rapid Solidification Processing: Principles and Technologies II. Proc. (2nd) Conf. on Rapid Solidification Processing, Reston, 1979, eds. R. Mehrabian *et al.* (Claitor, Baton Rouge, LA, p. 100.
KELTON, K. F., 1991, Solid State Phys. **45**, 75.
KELTON, K. F. and F. SPAEPEN, 1984, Phys. Rev. **B30**, 5516.
KELTON, K. F. and A. L. GREER, 1986, J. Non-Cryst. Solids **79**, 295.
KIJEK, M. M., D. W. PALMER and B. CANTOR, 1986, Acta Metall. **34**, 1455.
KIM, S. G., INOUE, A. and MASUMOTO, T., 1990, Mater. Trans. Japan Inst. Met. **31**, 929.
KIM, Y. H., A. INOUE and T. MASUMOTO, 1990, Mater. Trans. Japan Inst. Met. **31**, 747.
KIM, Y. H., A. INOUE and T. MASUMOTO, 1991, Mater. Trans. Japan Inst. Met. **32**, 334.
KIMINAMI, C. S. and P. R. SAHM, 1986, Acta Metall. **34**, 2129.
KIMURA, H. and T. MASUMOTO, 1983, in: Amorphous Metallic Alloys, ed. F. E. Luborsky (Butterworth, London) p. 187.
KISHORE, K. and H. K. SHOBHA, 1994, J. Chem. Phys. **101**, 7037.
KOCH, C. C., 1991, in: Processing of Metals and Alloys, ed. R. W. Cahn, Vol. 15 of Materials Science and Technology, eds. R. W. Cahn, P. Haasen and E. J. Kramer (VCH, Weinheim) p. 193.
KOEBRUGGE, G. W. and A. VAN DEN BEUKEL, 1988, Scripta Metall. **22**, 589.
KOMATSU, T., K. MATUSITA and R. YOKOTA, 1983, J. Non-Cryst. Sol. **72,** 279.
KÖSTER, U., 1984, in: Amorphous Metals and Non-Equilibrium Processing, ed. M. von Allmen (Editions de Physique, Paris) p. 175; Z. Metallkde. **75**, 691.
KÖSTER, W. and K. KAM, 1938, Z. Metallkde. **30**, 320.
KÖSTER, U. and U. HEROLD, 1981, in: Glassy Metals I, eds. H.-J. Güntherodt and H. Beck (Springer, Berlin) p. 225.
KÖSTER, U. and SCHÜNEMANN, 1993, in: Rapidly Solidified Alloys, ed. H. H. Liebermann (Dekker, New York), p. 303.
KÖSTER, U., U. SCHÜNEMANN, M. BLANK-BEWERSDORFF, S. BRAUER, M. SUTTON and G. B. STEPHENSON, 1991, Mater. Sci. Eng. **A133**, 611.
KRAMER, J., 1934, Ann. Phys. **19**, 37.
KRAMER, J., 1937, Z. Phys. **106**, 675.
KRISHNANAND, K. D. and R. W. CAHN, 1975, Scripta Metall. **9**, 1259.
KRISHNANAND and CAHN , 1976, in: Proc. 2nd Int. Conf. on Rapidly Quenched Metals, Cambridge, MA, 1975, eds. N. J. Grant and VB. C. Giessen, part 1 (MIT Press, Cambridge, MA) p. 67.
KUI, H.-W. and D. TURNBULL, 1985, Appl. Phys. Lett. **47**, 796.
KUI, H.-W., A. L. GREER and D. TURNBULL, 1984, Appl. Phys. Lett. **45**, 615.
KURŠUMOVIČ, A., M. G. SCOTT, E. GIRT and R. W. CAHN, 1980, Scripta Metall. **14** 1303.
LAMPARTER, P., W. SPERL, S. STEEB and J. BLÉTRY, 1982, Z. Naturforsch. **37a**, 1223.
LASHMORE, D. S. and J. F. WEINROTH. 1982, Plating and Surface Finishing (August) p. 72.
LASHMORE, D. S., L. H. BENNETT, H. E. SCHONE, P. GUSTAFSON and R. E. WATSON, 1982, Phys. Rev. Lett. **48**, 1760.
LEAKE, J. A., E. WOLDT and J. E. EVETTS, 1988, Mater. Sci. Eng. **97**, 469.
LEBO, M. and N. J. GRANT, 1974, Metall. Trans. **5**, 1547.

LEVI, C. G. and R. MEHRABIAN, 1982, Metall. Trans. **13A**, 13.
LEWIS, B. G. and H. A. DAVIES, 1976, Mater. Sci. Eng. **23**, 179.
LI, Y., H. JONES and H. A. DAVIES, 1992, Scripta Metal. et Mater. **26** ,1371.
LIEBERMANN, H. H., R. E. MAXWELL, R. W. SMASHEY and J. L. WALTER, 1983, Metall. Trans. **14A**, 1817.
LIN, C.-J. and F. SPAEPEN, 1982, Appl. Phys. Lett. **41**, 721.
LIN, C.-J. and F. SPAEPEN, 1984, Mat. Res. Soc. Symp. Proc. **28**, 75.
LIVESEY, A. K. and P. H. GASKELL, 1982, in: Proc. 4th Int. Conf. on Rapidly Quenched Metals, Sendai, 1981, eds. T. Masumoto and K. Suzuki (Japan Inst. of Metals, Sendai) p. 335.
LUBORSKY, F. E., 1977, in: Amorphous Metallic Alloys, ed. F. E. Luborsky (Butterworth, London) p. 360.
LUBORSKY, F. E., 1980, in: Ferromagnetic Materials, ed. E. P. Wohlfahrt (Butterworth, London), vol. 1, p. 451.
LUZZI, D. E., H. MORI, H. FUJITA and M. MESHII, 1984, Scripta Metall. **18**, 957.
LUZZI, D. E., H. MORI, H. FUJITA and M. MESHII, 1985, Scripta Metall. **19**, 897.
LUZZI, D. E., H. MORI, H. FUJITA and M. MESHII, 1986, Acta Metall. **34**, 629.
MACEDO, P. B. and A. NAPOLITANO, 1967, J. Res. Nat. Bur. Standards **71A**, 231.
MADDOX, J., 1993, Nature **361**, 397.
MADER, S., 1965, J. Vac. Sci. Technol. **2**, 35.
MADER, S., A. S. NOWICK and H. WIDMER, 1967, Acta Metall. **15**, 203, 215.
MAITREPIERRE, P. L.,1970, J. Appl. Phys. **41**, 498.
MAKINO, A., INOUE, A. and T. MASUMOTO, 1990, Mater. Trans. Japan Inst. Met. **31**, 884.
MARINGER, R. E. and C. E. MOBLEY, 1974, J. Vac. Sci. Technol. **11**, 1067.
MARTELLI, S., G. MAZZONE, A. MONTONE and M. VITTORI ANTISARI, 1990, J. Phys. France, **51**, 241.
MARTIN, J. W., 1968, Precipitation Hardening (Pergamon, Oxford).
MARTIN, J. W., B. CANTOR and R. D. DOHERTY, 1996, chapter 3 in: Stability of Microstructure in Metallic Systems (2nd edition) (Cambridge University Press).
MASSALSKI, T. B., 1986, in: Amorphous Metals and Semiconductors. Eds. P. Haasen and R. I. Jaffee (Pergamon, Oxford) p. 178.
MASUMOTO, T., 1994, Mater. Sci. Eng. **A179/A180**, 8. (Proceedings of 8th Int. Conf. on Rapidly Quenched and Metastable Materials, Sendai, 1993).
MASUMOTO, Y., A. INOUE, A. KAWASIMA, K. HASIMOTO, A. P. TSAI and T. MASUMOTO, 1986, J. Non-Cryst. Solids **86**, 121.
MAWELLA, K. J. A., R. W. K. HONEYCOMBE and P. R. HOWELL, 1982, J. Mater. Sci. **17**, 2850.
MEHRABIAN, R., 1982, Int. Metall. Rev. **27**, 185.
MENDOZA-ZELIS, L., L. THOMÉ, L. BROSSARD, J. CHAUMONT, Q. KROLAS and H. BERNAS, 1982, Phys. Rev. **B26**, 1306.
MENG, W. J., C. W. NIEH, E. MA, B. FULTZ and W. L. JOHNSON, 1988, Mater. Sci. Eng. **97**, 87.
MESCHTER, P. J., R. J. LEDERICK and J. E. O'NEAL 1988, in: Proc. Int. Conf. on PM Aerospace Materials, A Metal Powder Report Conf., Luzern, 1987 (MPIF, Amer. Powder Met. Inst., Princeton, NJ) p. 27. 1.
MIDSON, S. P. and H. JONES, 1982, in: Proc. 4th Int. Conf. on Rapidly Quenched Metals, Sendai, 1981, eds. T. Masumoto and K. Suzuki (Japan Inst. of Metals, Sendai) p. 1539.
MOLNÁR, A., C. V. SMITH, and M. BARTOK, 1989, Advances in Catalysis **36**, 329.
MORDIKE, B. L., 1986, ed., Laser Treatment of Materials (DGM Informationsgesellschaft Verlag, Oberursel, Germany).
MORDIKE, B. L., 1991, in: Processing of Metals and Alloys, ed. R. W. Cahn, Vol. 15 of Materials Science and Technology, eds. R. W. Cahn, P. Haasen and E. J. Kramer (VCH, Weinheim) p. 111.
MORI, H. and H. FUJITA, 1983, Jap. J. Appl. Physics **21**, L484.
MORI, H., FUJITA. H. and M. FUJITA, 1983, Jap. J. Appl. Phys. **22**, L94.
MORI, H., H. FUJITA, M. TENDO and M. FUJITA, 1984, Scripta Metall. **18**, 783.
MORRIS, D. G., 1980, in: Rapid Solidification Processing: Principles and Technologies II. Proc. (2nd) Conf. on Rapid Solidification Processing, Reston, 1979, eds. R. Mehrabian *et al.* (Claitor's, Baton Rouge, LA, p. 372.
MOTORIN, V. I., 1983, phys. stat. sol. (a) **80**, 447.
MOTTA, A. T. and C. LEMAIGNAN, 1992, in: Ordering and Disordering in Alloys, ed. A. R. Yavari (Elsevier Applied Science, London) p. 255.

MOTTERN, J. W. and W. J. PRIVOTT, eds., Spinning Wire from Molten Metal (Amer. Inst. Chem. Engrs., New York), A. I.Ch. E. Symp. Series **74**, no. 180.
MULDER, A. L., S. VAN DER ZWAAG and A. VAN DEN BEUKEL, 1983, Scripta Metall. **17**, 1399.
MURTY, Y. V. and R. P. I. ADLER, 1982, J. Mater. Sci. **17**, 1945.
NACHTRIEB, N. H., 1976, Ber. Bunsengesellschaft Phys. Chem. **80**, 678.
NAGEL, S. R. and J. TAUC, 1977, Solid State Comm. **22**, 129.
NEUHÄUSER, H, 1978, in: Proc. 3rd Int. Conf. on Rapidly Quenched Metals, Brighton, 1978, ed. B. Cantor (Institute of Metals [now Materials], London), vol. 2, p. 431.
NISHI, Y., T. KAI, M. TACHI, T. ISHIDAIRA and E. YAJIMA, 1986, Scripta Metall. **20**, 1099.
OGINO, T., 1986, French patent 2 569 143.
OHNAKA, I., 1985, Int. J. Rapid Solidification **1**, 219.
OHNAKA, I., 1988, in: Supplementary Volume 1 to the Encyclopedia of Materials Science and Engineering, ed. R. W. Cahn (Pergamon, Oxford), p. 584.
OHNAKA, I. and T. FUKUSAKO, 1978, J. Japan Inst. Metals **42**, 415.
OHTERA, K., A. INOUE, T. TERABAYASHI, H. NAGAHAMA and T. MASUMOTO, 1992, Mater. Trans. Japn Insty. Met. § 33, 775.
OK, H. N. and A. H. MORRISH, 1981, Phys. Rev. **B23**, 1835.
OSTWALD, W., 1893, Lehrbuch der allgemeinen Chemie, **2** (part 1), p. 517.
PANNISOD, P., I. BAKONYI and R. HASEGAWA, 1983, Phys. Rev. B **28**, 2374.
PATTERSON, J., A. L. GREER, J. A. LEAKE and D. R. W. JONES, 1978, in: Proc. 3rd Int. Conf. on Rapidly Quenched Metals, Brighton, 1978, ed. B. Cantor (Institute of Metals [now Materials], London), vol. 2, p. 293.
PEERCY, P. S., J. Y. TSAO and M. O. THOMPSON, 1990, J. Mater. Res. **5**, 1463.
PEKER, A. and W. L. JOHNSON, 1993, Appl. Phys. Lett. **63**, 2342.
PEREL, J., J. F. MAHONEY, B. E. KALENSHER and P. DUWEZ, (1981), in: Advances in Metal Processing, eds. J. J. Burke, R. Mehrabian and V. Weiss (Plenum, New York), p. 79.
PEREPEZKO, J. H., 1984, Mater. Sci. Eng. **65**, 125.
PEREPEZKO, J. H. and I. E. ANDERSON, in: Synthesis and Properties of Metastable Phases, eds. E. S. Machlin and T. J. Rowland (Met. Soc. AIME, Warrendale, PA) p. 31.
PEREPEZKO, J. H. and J. S. SMITH 1981, J. Non-Cryst. Solids **44**, 65.
PHARR, G. M., W. C. OLIVER, R. F. COOK, P. D. KIRCHNER, M. C. KROLL. T. R. DINGER and D. R. CLARKE, 1992, J. Mater. Res. **7**, 961.
PILLER, J., and P. HAASEN, 1982, Acta Metall. **30**, 1.
PIROUZ, P., R. CHAIM, U. DAHMEN and K. H. WESTMACOTT, 1990, Acta Metall. et Mater. **38**, 313, 323, 329.
POATE, J. M., P. S. PEERCY and M. O. THOMPSON, 1987, Mat. Res. Soc. Symp. Proc. **57**, 465.
POLK, D. E. and H. S. CHEN, 1974, J. Non-Cryst. Solids, **15**, 165.
POLK, D. E. and B. C. GIESSEN, 1978, in: Metallic Glasses. Eds. J. J. Gilman and H. J. Leamy (Amer. Soc. Metals, Metals Park, Ohio) p. 1.
POLMEAR, I. J., 1995, Light Alloys, 3rd edition (Edward Arnold, London)
POND, R. and R. MADDIN, 1969, Mater. Sci. Eng. **23**, 87.
PONYATOVSKY, E. G. and O. I. BARKALOV, 1991, Mater. Sci. Eng. **A133**, 726.
POSGAY, G., F. J. KEDRES, B. ALBERT, S. KISS and I. Z. HARANGUZO, 1985, in: Proc. 5th Int. Conf. on Rapidly Quenched Metals, eds. S. Steeb and H. Warlimont, (North-Holland, Amsterdam)) p. 723.
PRESTON, G. D., 1938, Nature **142**, 570.
RAMACHANDRARAO, 1980, Z. Metallkde. **71**, 172.
RAMACHANDRARAO, P., 1984, in: Metallic Glasses – Production, Properties and Applications. Ed. T. R. Anantharaman (Trans Tech, Aedermannsdorf, Switzerland) p. 31.
RAUSCHENBACH, B. and K. HOHMUTH, 1982, phys. stat. sol. (a) **72**, 667.
RAY, R., 1981, J. Mater. Sci. **16**, 1924, 2927.
RAY, R., B. C. GIESSEN and N. J. GRANT, 1968, Scripta Metall. **2**, 359.
RAY, R. and E. MUSSO, 1976, US Patent 3 981 722.
RAY, R., R. HASEGAWA, C. P. CHOU and L. A. DAVIS, 1977, Scripta Metall. **11**, 973.
RAYMENT, J. J., and B. CANTOR 1978, Metal Sci. **12**, 156.

RAYMENT, J.J., and B. CANTOR, 1981, Metall. Trans. **12A**, 1557.
RECHTIN, M.D., J.B. Vander Sande and P.M. BALDO, 1978, Scripta Metall. **12**, 639.
REDDY, G.S. and J.A. SEKHAR, 1989, in: Principles of Solidification and Materials Processing, eds. R. Trivedi, J.A. Sekhar and J. Mazumdar, (Oxford & IBH, New Delhi) Vol. 1, p. 83.
ROBERTS, W.T., 1993, in: Third Supplementary Volume to the Encyclopedia of Materials Science and Engineering, ed. R.W. Cahn (Pergamon, Oxford), p. 1616.
ROORDA, S., 1990, Doctoral thesis, University of Utrecht, Netherlands.
ROOS, J.R., J.-P. CELIS and M. de BONTE, 1991, in: Processing of Metals and Alloys, ed. R.W. Cahn, Vol. 15 of Materials Science and Technology, eds. R.W. Cahn, P. Haasen and E.J. Kramer (VCH, Weinheim), p. 481.
RUHL, R.C., B.C. GIESSEN, M. COHEN and N.J. GRANT, 1967, Acta Metall. **15**, 1693.
RUSSELL, K.C., 1985, in: Proc. Int. Seminar on Solute-Defect Interaction, Kingston, Ontario, August 1985, eds. G. Saimote, G. Kidson and G. Purdy (Pergamon, Oxford) p. 317.
SADOC, J.F. and J. DIXMIER, 1976, Mater. Sci. Eng. **23**, 187.
SAFAI, S. and H. HERMAN, 1981, in: Ultrarapid Quenching of Liquid Alloys, Treatise on Materials Science and Technology, vol. 20, ed. H. Herman, (Academic, New York), p. 183.
SAHIN, E. and H. JONES, 1978, in: Proc. 3rd Int. Conf. on Rapidly Quenched Metals, Brighton, 1978, ed. B. Cantor (Institute of Metals, London), vol. 1, p. 138.
SAKATA, M., N. COWLAM and H.A. DAVIES, 1981, J. Phys. F **11**, L157.
SANKARAN, K.K. and N.J. GRANT, 1980, Mater. Sci. Eng. **44**, 213.
SARE, I.R. and R.W.K. HONEYCOMBE, 1978, J. Mater. Sci. **13**, 1991.
SAUNDERS, N. and A.P. MIODOWNIK, 1986, J. Mater. Res. **1**, 38.
SCHLEIP, E., R. WILLNECKER, D.M. HERLACH and G.P. GÖRLER, 1988, Mater. Sci. Eng. **98**, 39.
SCHOLTE, P.M.L.O., M. TEGGE, F. VAN DER WOUDE, K.H.J. BUSCHOW AND I. VINCZE, 1985, in: Proc. 5th Int. Conf. on Rapidly Quenched Metals, eds. S. Steeb and H. Warlimont (North-Holland, Amsterdam) p. 541.
SCHRÖDER, H., K. SAMWER and U. KÖSTER, 1985, Phys. Rev. Lett. **54**, 19.
SCHUBERT, E. and H.W. BERGMANN, 1993, in: Rapidly Solidified Alloys, ed. H.H. Liebermann (Marcel Dekker, New York), p. 195.
SCHULTZ, L., 1987, Mat. Res. Soc. Symp. Proc. **80**, 97.
SCHULTZ, L., 1989, in: New Materials by Mechanical Alloying Techniques, eds. E. Arzt and L. Schultz, (DGM Informationsgesellschaft, Oberursel, Germany) p. 53.
SCHULTZ, L. and E. HELLSTERN, 1987, Mat. Res. Soc. Symp. Proc. **80**, 3.
SCHUMACHER, P. and A.L. GREER, 1994, Mater. Sci. Eng. **A181/A182**, 1335.
SCHWARZ, M., A. KARMA, K. ECKLER and D.M. HERLACH, 1994, Phys. Rev. Lett. **73**, 1380 (correction, ibid. 2940).
SCHWARZ, R.B. and W.L. JOHNSON, 1983, Phys. Rev. Lett. **51**, 415.
SCHWARZ, R.B. and C.C. KOCH, 1983, Appl. Phys. Lett. **49**, 146.
SCHWARZ, R.B. and W.L. JOHNSON (eds.), 1988, Proc. Conf. on Solid-State Amorphizing Reactions, J. Less-Common Metals **140**.
SCOTT, G.D. and D.L. MADER, 1964, Nature **201**, 382.
SCOTT, M.G. and J.A. LEAKE, 1975, Acta Metall. **23**, 503.
SCOTT, M.G. and A. KURŠUMOVIĆ, 1982, Acta Metall. **30**, 853.
SEKHAR, J.A. and T. RAJASEKHARAN, 1986, Nature **320**, 153.
SHINGU, P.H. and K.N. ISHIHARA, 1993, in: Rapidly Solidified Alloys, ed. H.H. Liebermann (Marcel Dekker, New York), p. 103.
SHOHOJI, N., H.A. DAVIES, H. JONES and D.H. WARRINGTON, 1982, in: Proc. 4th Int. Conf. on Rapidly Quenched Metals, eds. T. Masumoto and K. Suzuki (Japan Inst. of Metals, Sendai) p. 1529.
SILCOCK, J.M., T.J. HEAL and H.K. HARDY, 1953-54, J. Inst. Metals **82**, 239.
SIMPSON, A.W. and P.A. HODKINSON, 1972, Nature **237**, 320.
SINHA, A.K., B.C. GIESSEN and D.E. POLK, 1976, in: Treatise on Solid-State Chemistry, ed. N.B. Hannay (Plenum, New York) vol. 3, p. 1.
SINNING, H.-R., L. LEONARDSSON and R.W. CAHN, 1985, Int. J. of Rapid Solidification **1**, 175.

SMITH, C. H., 1993, in: Rapidly Solidified Alloys, ed. H. H. Liebermann (Marcel Dekker, New York), p. 617.
SMITH, T. R., R. W. ARMSTRONG, P. M. HAZZLEDINE, R. A. MASUMURA and C. S. PANDE, 1995, in: Mat. Res. Soc. Syp. Proc. **362**, 31.
SMUGERESKY, J. E., 1982, Metall. Trans. **13A**, 1535.
SOMMER, F., G. DUDDEK and B. PREDEL, 1978, Z. Metallkde. **69**, 587.
SOMMER, F., T. LANG and B. PREDEL, 1987, Z. Metallkde. **78**, 596.
SPAEPEN, F., 1977, Acta Metall. **25**, 407.
SPAEPEN, F., 1994, Mater. Sci. Eng. **A179/A180**, 81.
SPAEPEN, F. and G. S. CARGILL, 1985, in: Proc. 5th Int. Conf. on Rapidly Quenched Metals, eds. S. Steeb and H. Warlimont (North-Holland, Amsterdam) p. 581.
SPAEPEN, F. and A. I. TAUB, 1983, in: Amorphous Metallic Alloys, ed. F. E. Luborsky (Butterworth, London) p. 231.
SPAEPEN, F., S. S. TSAO and T. W. WU, 1986, in: Amorphous Metals and Semiconductors, eds. P. Haasen and R. I. Jaffee (Pergamon, Oxford) p. 365.
SRIVASTAVA, P. K., B. C. GIESSEN and N. J. GRANT, 1972, Metall. Trans. **3**, 977.
STIFFLER, S. R., P. V. EVANS and A. L. GREER, 1992, Acta Metall. 40, 1617.
STRÖM-OLSEN, J. O., R. BRÜNING, Z. ALTOUNIAN and D. H. RYAN, 1988, J. Less-Common Metals **145**, 327.
SUGA, M., J. L. GOSS, G. R. OLSEN and J. B. VAN DER SANDE, 1980, in: Rapid Solidification Processing: Principles and Technologies II. Proc. (2nd) Conf. on Rapid Solidification Processing, Reston, 1979, eds. R. Mehrabian *et al.* (Claitor, Baton Rouge, LA, p. 364.
SURYANARAYANA, C., 1980, Rapidly Quenched Metals – A Bibliography (Plenum, New York).
SURYANARAYANA, C., 1991, in: Processing of Metals and Alloys, ed. R. W. Cahn, Vol. 15 of Materials Science and Technology, eds. R. W. Cahn, P. Haasen and E. J. Kramer (VCH, Weinheim), p. 57.
SUSLICK, K. S., 1993, in: Supplementary Volume 3 of the Encyclopedia of Materials Science and Engineering, ed. R. W. Cahn (Pergamon, Oxford), p. 2094.
SUSLICK, K. S., 1995, Mater. Research Bull. **20** (April), 29.
SUSLICK, K. S., S. B. CHOE, A. A. CICHOWLAS and M. W. GRINSTAFF, 1991, Nature **353**, 414.
SUZUKI, K., 1983, in: Amorphous Metallic Alloys, ed. F. E. Luborsky (Butterworths, London), p. 74.
SUZUKI, Y., J. HAIMOVICH and T. EGAMI, 1987, Phys. Rev. **B35**, 2162.
TAKAYAMA, S., 1976, J. Mater. Sci. **11**, 164.
TANNER, L. and R. RAY, 1980, Scripta Metall. **28**, 633.
TARNÓCZI, T., A. LOVAS and C. KOPASZ, 1988, Mater. Sci. Eng. **97**, 509.
TENDLER, R. H. de, 1986, J. Mater. Sci. **21**, 64.
THURSFIELD, G. and M. J. STOWELL, 1974, J. Mater. Sci. **9**, 1644.
TOLOUI, B., A. KURŠUMOVIĆ and R. W. CAHN, 1985, Scripta Metall. **19**, 947.
TOMIDA, T. and T. EGAMI, 1991, Mater. Sci. Eng. **A133**, 931.
TRACEY, V. A. and C. P. CUTLER, 1981, Powder Metallurgy **24**, 37.
TSAI, R.-Y. and S.-T. WU, 1990, J. Electrochem. Soc. **137**, 2803.
TURNBULL, D., 1952, J. Chem. Phys. **20**, 411.
TURNBULL, D., 1976, J. de Physique **37**, C4-1.
TURNBULL, D., 1981, Metall. Trans. **12A**, 695.
TURNBULL, D. and H. N. TREAFTIS, 1957, Acta Metall. Mater. **5**, 534.
TURNBULL, D., H. S. ROSENBAUM and H. N. TREAFTIS, 1960, Acta Metall. **8**, 277.
VAN DEN BEUKEL, A. and J. SIETSMA, 1990, Acta Metall. et Mater. **38**, 383.
VAN DEN BEUKEL, A. and J. SIETSMA, 1994, Mater. Sci. Eng. **A179/A180**, 86. (Proc. 8th Int. Conf. on Rapidly Quenched and Metastable Materials, Sendai, 1993).
VAN ROOYEN, M., J. A. SINTE MAARTENSDIJK, and E. J. MITTEMEIJER, 1988, Metall. Trans. **19A**, 2433.
VINEBERG, E. J., E. K. OHRINER, E. P. WHELAN and G. E. STAPLEFORD, 1985, in: Rapidly Solidified Crystalline Alloys, eds. S. K. Das, B. H. Kear and C. M. Adam (The Metallurgical Society, Warrendale, PA) p. 306.
VOLKERT, C. A. and F. SPAEPEN, 1989, Acta Metall. **37**, 1355.
VON HEIMENDAHL, L., 1975, J. Phys. F **5**, L141.
VREDENBURG, A. M., J. F. M. WESTENDORP, F. W. SARIS, N. M. van der PERS and Th. D. KEIJSER, 1986, J. Mater. Res. **1**, 775.

WAGNER, C. N. J., 1983, in: Amorphous Metallic Alloys, ed. F. E. Luborsky (Butterworth, London) p. 58.
WAGNER, C. N. J., 1986, in: Amorphous Metals and Semiconductors. Eds. P. Haasen and R. I. Jaffee (Pergamon, Oxford) p. 54.
WALTER, J. L., F. BACON and F. E. LUBORSKY, 1976, Mater. Sci. Eng. **24**, 239.
WARBURTON, W. K. and D. TURNBULL, 1975, in : Diffusion in Solids: Recent Developments, eds. A. S. Nowick and J. J. Burton (Academic Press, New York) p. 171.
WARLIMONT, H., 1985, in: Rapidly Quenched Metals (Proc. 5th Conf.), S. Steeb, H. Warlimont, eds. (North-Holland, Amsterdam), p. 1599.
WEI, Gao and B. CANTOR, 1989, Acta Metall. **37**, 3409.
WHANG, S. H., 1983, Mater. Sci. Eng. **57**, 87.
WHALLEY, E., 1988, J. Less-Common Metals, **140**, 361.
WHITTLE, G. L., P. WELLS, S. J. CAMPBELL, A. CALKA AND A. M. STEWART, 1985, in: Proc. 5th Int. Conf. on Rapidly Quenched Metals, eds. S. Steeb and H. Warlimont (North-Holland, Amsterdam) p. 545.
WILLIAMS, R. S. and T. EGAMI, 1976, IEEE Trans. Mag. **MAG12**, 927.
WOLDT, E., 1988, J. Mater. Sci. **23**, 4383.
WOLF, D., P. R. OKAMOTO, S. YIP, J. F. LUTSKO and M. KLUGE, 1990, J. Mater. Res. **5**, 286.
WONG, J., 1981, in: Glassy Metals I, eds. H. J. Güntherodt and H. Beck, (Springer, Berlin).
WOOD, J. V. and R. W. K. HONEYCOMBE, 1981, in: Ultrarapid Quenching of Liquid Alloys. Treatise on Materials Science and Technology, vol. 20, ed. H. Herman (Academic, New York) p. 117.
WU, T. W. and F. SPAEPEN, 1986, in: Mechanical Behavior of Rapidly Solidified Material, eds. S. M. L. Sastry and B. A. Macdonald (TMS, Warrendale, PA) p. 293.
YAMASAKI, T, M. TAKAHASHI and Y. OGINO, 1985, in: Rapidly Quenched Metals (V), eds. S. Steeb and H. Warlimont (North-Holland, Amsterdam) p. 1381.
YAVARI, A. R., 1986, J. Mater. Res. **1**, 746.
YAVARI, A. R., P. HICTER and P. DESRÉ, 1983, J. Chim. Physique **79**, 572.
YEARIM, R. and D. SHECHTMAN, 1982, Metall. Trans. **13A**, 1891.
YEH, X. L., K. SAMWER and W. L. JOHNSON, 1983, Appl. Phys. Lett. **42**, 242.
YERMAKOV, A. Ye., V. A. BARINOV and Ye. Ye. YURCHIKOV, 1981, Phys. Met. Metall. (translation of Fiz. Met. Met.) **54**, 935.
YOSHIZAWA, Y., S. OGUMA and K. YAMAUCHI, **1988**, J. Appl. Phys. **64**, 6044.
ZACHARIASEN, W. H., 1932, J. Am. Ceram. Soc. **54**, 3841.
ZOUGMORÉ, F., J. C. LASJAUNIAS and O. BÉTHOUX, 1988, J. Less-Common Metals **145**, 367.

Further reading

Numerous international conferences have been devoted to the field of rapidly solidified alloys since 1970, recently at 3-yearly intervals; initially the conference title referred to "rapidly quenched metals", latterly to "rapidly quenched and metastable materials". The proceedings of most of these are cited in the Bibliography, in relation to specific papers, as are other important proceedings such as those of the Reston Conferences on Rapid Solidification Processing and other 'one-off' meetings. Most of these will not be listed separately here, but it should be recognized that these proceedings are prime sources in this broad field.

A few selected overviews are listed below. Others are recommended in the body of the chapter.

T. R. Anantharaman and C. Suryanarayana, Review: A decade of quenching from the melt, J. Mater. Sci. **6** (1971) 1111.

R. W. Cahn, Metallic Glasses, Contemp. Phys. 21 (1980) 43.

F. E. Luborsky, in: Ferromagnetic Materials, vol. 1, ed. P. E. Wohlfahrt (North-Holland, Amsterdam. 1980), ch. 6, p. 451. (Detailed quantitative data on magnetic properties, densities, etc., of many metallic glasses are included).

H. Herman, ed., Ultrarapid Quenching of Liquid Alloys, Treatise on Materials Science and Technology, vol. 20 (Academic Press, New York, 1981). (The chapter by Li, on mechanical properties, tabulates T_g, T_x, densities, elastic moduli, etc., of many metallic glasses).

H. Jones, Rapid Solidification of Metals and Alloys (Monograph No. 8, The Institution of Metallurgists, London; now the Institute of Materials; 1982).
S. R. Nagel, Metallic Glasses, Adv. Chem. Phys. 51 (1982) 227.
R. Mehrabian, Rapid Solidification, Int. Metals Review **27** (1982) 185.
F. E. Luborsky, ed., Amorphous Metallic Alloys (Butterworth, London, 1983). The most complete survey of metallic glasses.
F. Spaepen and D. Turnbull, Metallic Glasses, Ann. Rev. Phys. Chem. **35** (1984) 241.
S. K. Das, B. H. Kear and C. M. Adam (eds.), Rapidly Solidified Crystalline Alloys (The Metallurgical Society, Warrendale, PA, 1985).
S. M. L. Sastry and B. A. MacDonald (eds.), Mechanical Behavior of Rapidly Solidified Materials (The Metallurgical Society, Warrendale, PA, 1985).
T. R. Anantharaman and C. Suryanarayana, Rapidly Solidified Metals (Trans Tech. Publications, Switzerland, 1987).
R. W. Cahn, Metallic Glasses, in: Glasses and Amorphous Materials, ed. J. Zarzycki, Vol. 9 of Materials Science & Technology (VCH, Weinheim, Germany, 1991), p. 493.
C. Suryanarayana, Rapid Solidification, in: Processing of Metals and Alloys, ed. R. W. Cahn (Vol. 15 of Materials Science and Technology, ed. R. W. Cahn, P. Haasen and E. J. Kramer (VCH, Weinheim, Germany) p. 57.
H. H. Liebermann (ed.), Rapidly Solidified Alloys (Wiley, New York, 1993). (Including, in particular, chapters by R. W. Cahn on history, A. L. Greer on relaxation and diffusion, T. Egami on metallic glass structure, and U. Köster and U. Schünemann on phase transformations in RSP alloys.)
A. L. Greer, Amorphous phases, in: Intermetallic Compounds, ed. J. H. Westbrook and R. L. Fleischer (Wiley, New York, 1995), p. 731. Also: W. L. Johnson, Metastable Phases, ibid., p. 687.
A. L. Greer, Metallic Glasses, Science **267,** (1995) 1947.
A. L. Greer, Crystallization of Amorphous Alloys, Metall. Trans. **26A** (1995) (in press).
H. Bakker, G. F. Zhou and H. Yang, Mechanically Driven Disorder and Phase Transformations in Alloys, Progr. Mat. Sci. **39** (1995) 159.
J. W. Martin, B. Cantor and R. D. Doherty, Stability of Microstructure in Metallic Systems (2nd edition) (Cambridge University Press, 1996).

Classified bibliographies of papers concerned with rapid solidification (covering both crystalline and amorphous materials):

H. Jones and C. Suryanarayana, Annotated Bibliography 1958–72. J. Mater. Sci. 8 (1973) 705.
C. Suryanarayana, Rapidly Quenched Metals: A Bibliography, 1973–79 (Plenum, New York, 1981).
C. Suryanarayana, Bibliography on Mechanical Alloying and Milling (Cambridge International Science Publishing, 1995).

– SUBJECT INDEX –

1st Volume: Pages 1–942
2nd Volume: Pages 943–1830
3rd Volume: Pages 1831–2740

"page number ff" signifies that treatment of the topic continues on the following page(s).

The late Prof. Peter Haasen

Colour picture on front cover:

Simulation of an alloy dendrite growing into a supercooled liquid using the phase-field method. The colours show variation of composition (atomic fraction Cu) in the liquid and solid for parameters approximating a Ni-Cu alloy with 0.41 atomic fraction Cu. See ch. 8, par. 7.5 (Courtesy of William J. Boettinger and James A. Warren).

ISBN 0 444 89875 1