Claus Klingshirn

Semiconductor Optics

Second Edition

With 380 Figures and 20 Tables

Springer

Professor Claus Klingshirn
University of Karlsruhe
Institute of Applied Physics
Wolfgang-Gaede-Str. 1
76131 Karlsruhe, Germany

Chapter 27 by R. v. Baltz was taken from Landolt-Börnstein, Group III, Volume 34 / Subvolume C1: "Semiconductor Quantum Structures – Optical Properties" (edited by C. Klingshirn), 2001, Springer-Verlag Heidelberg.

Library of Congress Control Number: 2004107246

ISSN 1439-2674

ISBN 3-540-21328-7 Springer Berlin Heidelberg New York

ISBN 3-540-61687-X 1st Ed. Corr. Printing Springer Berlin Heidelberg New York

Production and typesetting: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig
Cover production: *design & production* GmbH, Heidelberg

Printed on acid-free paper SPIN: 10798841 56/3141/YL 5 4 3 2 1 0

To my parents, my wife and my children

Wahrheit und Klarheit sind komplementär.

E. Mollwo

This aphorism was coined in the nineteen-fifties by E. Mollwo, Professor of Physics at the Institut für Angewandte Physik of the Universität Erlangen during a discussion with W. Heisenberg. The author hopes that, with respect to his book, the deviations from exact scientific truth (Wahrheit) and perfect understandability (Klarheit) are in a reasonable balance.

Just as an illustration of the above statement, the attention of the author has been drawn to the fact, that the same statement has been reported even in German language also from Niels Bohr. See Steven Weinberg, Dreams of a Final Theory, Vintage Books, New York (1994) p. 74.

Preface to the Second Edition

The book on Semiconductor Optics has been favourably received by the students and the scientific community worldwide. After the first edition, which appeared in 1995 several reprints became necessary starting from 1997, one of them for the Chinese market. They contained only rather limited updates of the material and corrections.

In the meantime scientific progress brought a lot of new results, which necessitate a new, seriously revised edition. This progress includes bulk semiconductors, but especially structures of reduced dimensionality. These new trends and results are partly included in existing chapters e.g. for phonons or for time-resolved spectroscopy, partly new chapters have been introduced like the ones on cavity polaritons and photonic structures.

We based the description of the optical properties again on the simple and intuitively clear model of the Lorentz-oscillators and the concept of polaritons as the quanta of light in matter. But since there is presently a trend to describe at least the optical properties of the electronic system of semiconductors by the optical or the semiconductor Bloch equations, a chapter has been added on this topic written by Prof. Dr. R. v. Baltz (Karlsruhe) to familiarize the reader with this concept, too, which needs a bit more quantum mechanics compared the approach used here. The chapter on group theory has been revised by Prof.Dr. K. Hümmer (Karlsruhe/Forchheim)

Karlsruhe,
September 2004

C.F. Klingshirn

Preface to the First Edition

One of the most prominent senses of many animals and, of course, of human beings is sight or vision. As a consequence, all phenomena which are connected with light and color, or with the optical properties of matter, have been focal points of interest throughout the history of mankind. Natural light sources such as the Sun, the Moon and stars, or fire, were worshipped as gods or godesses in many ancient religions. Fire, which gives light and heat, was for many centuries thought to be one of the four elements – together with earth, water, and air. In alchemy, which marks the dawn of our modern science, the Sun and the Moon appeared as symbols of gold and silver, respectively, and many people tried to produce these metals artificially. Some time later, Johann Wolfgang von Goethe (1749–1832) considered his "Farbenlehre" as more important than his poetry. In the last two centuries a considerable fraction of modern science has been devoted to the investigation and understanding of light and the optical properties of matter. Many scientists all over the world have added to our understanding of this topic. As representatives of the many we should like to mention here only a few of them: I. Newton (1643–1727), J.C. Maxwell (1831–1879), M. Planck (1858–1947), A. Einstein (1879–1955), N. Bohr (1885–1962), and W. Heisenberg (1901–1976).

The aim of this book is more modest. It seeks to elucidate one of the numerous aspects in the field of light and the optical properties of matter, namely the interaction of light with semiconductors, i.e., semiconductor optics. The investigation of the properties of semiconductors has, in turn, its own history, which has been summarized recently by H.J. Queisser [85Q1]. In Queisser's book one can find early examples of semiconductor optics, namely the observation of artificially created luminescence by V. Cascariolo in Bologna at the beginning of the 17th century, or by K.F. Braun (1850–1918), inventor of the "Braun'sche Röhre" (Braun's tube) now usually called CRT (cathode ray tube), at the beginning of this century.

Another root of semiconductor optics comes from the investigation of the optical properties of insulators, especially of the color (Farb- or F-) centers in alkali halides. This story has been written down recently by J. Teich-

mann [88T1]. It is inseparably connected with names such as Sir Nelville Mott and A. Smakula, but especially with R.W. Pohl (1884–1976) and his school in Gottingen.

Together with J. Franck (1882–1964) and M. Born (1882–1970) R.W. Pohl was one of the outstanding physicists of the "golden years of physics" at Göttingen before 1933 [77B1,84M1,88H1]. The present author considers himself a scientific grandson of Pohl, with E. Mollwo (1909–1993), F. Stöckmann (1918–1998) and W. Martienssen (*1926) as the intermediate generation, and he owes to them a large part of his scientific education.

Scientific interest in semiconductor optics comprises both fundamental and applied research. It has been an extremely lively, rapidly developing area of research for the last five decades and more, as can be seen from the contributions to the series of International Conferences on the Physics of Semiconductors [50I1] and on Luminescence [81I1] or on Non-linear Optics and Excitation Kinetics [87N1]. It does not need much of a prophetic gift to predict that semiconductor optics will continue to be a major topic of solid state physics far into the next century. Many applications of semiconductor optics are known from everyday life such as light-emitting diodes (LED) in displays, laser diodes in compact-disk (CD) players, laser printers and laser scanners or solar cells.

Karlsruhe, *C.F. Klingshirn*
February 1995

References

[50I1] The Series of Int'l Conferences on the Physics of Semiconductors (ICPS) was started in 1950 in Reading. Proceedings of the more recent ones are
a. 12th ICPS, Stuttgart (1974), ed. by M.H. Pilkuhn (Teubner, Stuttgart 1974)
b. 13th ICPS, Rome (1976), ed. by F.G. Fumi (Tipographia Marves, Rome 1976)
c. 14th ICPS, Edinburgh (1978), ed. by B.L.H. Wilson (The Institute of Physics, Bristol 1979)
d. 15th ICPS, Kyoto (1980), ed. by S. Tanaka, Y. Toyozawa: J. Phys. Soc. Jpn. **49**, Suppl. A (1980)
e. 16th ICPS, Montpellier (1982), ed. by M. Averous: Physica B 117 + 118 (1983)
f. 17th ICPS, San Francisco (1984), ed. by J.M. Chadi, W.A. Harrison (Springer, Berlin, Heidelberg 1984)
g. 18th ICPS, Stockholm (1986), ed. by O. Engstrom (World Scientific, Singapore 1987)
h. 19th ICPS, Warsaw (1988), ed. by W. Zawadzki (The Institute of Physics, Polish Academy of Sciences, 1988)

i. 20th ICPS, Thessaloniki (1990), ed. by E.M. Anastassakis, J.D. Joannopoulos (World Scientific, Singapore 1990)
j. 21st ICPS, Beijing (1992), ed. by Ping Jiang, Hou-Zhi Zheng (World Scientific, Singapore 1993)
k. 22nd ICPS, Vancouver (1994) D.J. Loockwood ed. (World Scientific, Singapore, 1995)
l. 23rd ICPS, Berlin (1996), M. Scheffler and R. Zimmermann (eds.), World Scientific, Singapore (1996)
m. 24th ICPS, Jerusalem (1998), D. Gershoni (ed.), World Scientific, Singapore (1999)
n. 25th ICPS, Osaka (2000), N. Miura and T. Ando (eds.), Springer Proc. In Physics **87**, Springer, Berlin (2001)
o. 26th ICPS, Edinburgh (2002), A.R. Long and J.H. Davies (eds.), Institute of Physics Conf. Series **171**, (2003)
p. 27th ICPS, Flagstaff (2004)

[77B1] A.D. Beyerchen: *Scientists under Hitler,* (Yale Univ. Press, New Haven 1977)
[84M1] E. Mollwo: Physik in unserer Zeit **15**, 110 (1984)
[85Q1] H.-J. Queisser: *Kristallene Krisen* (Piper, München 1985)
[88H1] F. Hund, H. Maier-Leibnitz, E. Mollwo: Eur. J. Phys. **9**, 188 (1988)
[88T1] J. Teichmann: *Zur Geschichte der Festkorperphysik-Farbzentrenforschung bis 1940* (Steiner, Wiesbaden 1988)
[81I1] The proceedings of the Series of Int'l Conferences of Luminescence (ICL) are published in J. Lumin. The more recent ones were
a. ICL, Berlin (1981), ed. by I. Broser, H.-E. Gumlich, R. Broser: J. Lumi. **24/25** (1981)
b. ICL, Madison (1984), ed. by W.M. Yen, J.C. Wright: J. Lumin. **31/32** (1984)
c. ICL, Beijing (1987), ed. by Xu Xurong: J. Lumin. **40/41** (1987)
d. ICL, Lisbon (1990), ed. by S.J. Formosinho, M.D. Sturge: J. Lumin. **48/49** (1990)
e. ICL, Storrs (1993) ed. by D.S. Hamilton, R.S. Meltzer and M.D. Sturge: J. Lumi. **60/61** (1995)
f. ICL, Prague (1996) ed. J. Hala, P. Reinecker, J. Lumin. **72–74** (1997)
g. ICL, Osaka (1999) ed by K. Cho, J. Lumin **87–89** (2000)
h. ICL Budapest (2002), ed by S. Speiser, J. Lumin **102–103** (2003)

[87N1] The Series of International conferences/workshops on Nonlinear Optics and Excitation Kinetics (NOEKS) has been started in the former German Democratic Republic (DDR) and continued successfully after the reunification of Germany. The proceedings have so far been published in
a. NOEKS I Nov. 1987, Bad Stuer phys. stat. sol. (b) **146** and **147** (1988)
b. NOEKS II Dez. 1989, Bad Stuer phys. stat. sol. (b) **159** (1) (1990)
c. NOEKS III Mai 1992, Bad Honnef phys. stat. sol. (b) **173** (1) (1992)
d. NOEKS IV Nov. 1994, Gosen phys. stat. sol. (b) **188** (1) (1995)
e. NOEKS V Sept. 1997, Graal-Müritz phys. stat. sol. (b) **206** (1) (1998)
f. NOEKS VI April 2000, Marburg phys. stat. sol. (b) **221** (1) (2000)
g. NOEKS VII Feb. 2003, Karlsruhe phys. stat. sol. (c) **0** (5) (2003)

Acknowledgements

This book is based on various lectures given by the author at the Universities of Karlsruhe, Frankfurt am Main and Kaiserslautern, at Harvard University and the University of Metz and on contributions given at several of the Summer Schools on Atomic and Molecular Spectroscopy organised by Prof. Dr. B. Di Bartolo in Erice, Sicily.

The sources of the scientific information presented here are partly the references given. Of equal importance, however, is the physics, which I learned from my academic teachers during my studies and PhD work at the University of Erlangen, my post-doc time at the Laboratoire de Spéctroscopie et d'Optique du Corps Soilide in Strasbourg and my Habilitation at the University of Karlsruhe, and later on from fruitful discussions with many collegues and co-workers at the places where I was or still am as Professor (Frankfurt am Main, Kaiserslautern and Karlsruhe) and abroad including guest scientists in my group. Without trying to be complete, I should like to mention my academic teachers Profs. Drs. R. Fleischmann (†), H. Volz (†), E. Mollow (†), R. Helbig and K. Hümmer (Erlangen) and F. Stöckmann (†), W. Ruppel and W. Stößel (Karlsruhe).

From the colleagues I should like to mention with great pleasure fruitful and stimulating discussions e.g. with Profs. Drs. H. Haug, W. Martienssen, E. Mohler and L. Banyai (Frankfurt am Main), J.B. Grun, B. Hönerlage and R. Levy (Strasbourg), B. Stébé (Metz), U. Rössler (Regensburg), E. Göbel, S.W. Koch, S. Schmitt-Rink (†) and P. Thomas (Marburg), J.M. Hvam (Lyngby), D.S. Chemla (Berkeley), K.P. O'Donnel (Glasgow), E. Mazur (Cambridge), I. Bar-Joseph and R. Reisfeld (Israel), I. Broser, R. Zimmermann and F. Henneberger (Berlin), H. Stolz and K. Henneberger (Rostock), A. Reznitsky, A. Klochikhin and S. Permogorov (St. Petersburg), V. Lyssenko (Chernogolovka), O. Gogolin and E. Tsitisishvili (Tiblissi), M. Brodyn and S. Shevel (†) (Kiev), S. Gaponenko and A. Apanasevich (Minsk), H. Kalt, M. Wegener, R. v. Baltz and K. Busch (Karlsruhe), U. Woggon (Dortmund), H. Giessen (Bonn) and last but not least B. Di Bartolo (Boston), also for

running the school in Erice and the special and comfortable atmosphere he creates there for all participants.

My special thanks are due to all my former and present students and co-workers, who produced their Diplom, PhD or Habilitation thesis in my research group and many of the fine results presented in this book and who partly hold in the meantime professorships or equivalent positions of their own (H. Kalt (Karlsruhe), M. Wegener (Karlsruhe), U. Woggon (Dortmund). H. Giessen (Bonn), M. Kuball (Bristol) and W. Langbein (Cardiff)). Beyond that I do not want to give names here, because they are too many and I am afraid to forget somebody.

In this context the financial support for my research is gratefully acknowledged especially from the Deutsche Forschungsgemeinschaft, the Länder Hessen, Rheinland-Pfalz and Baden-Württemberg, the Stiftung Volkswagenwerk, the Bundesministerium für Bildung und Forschung (BMBF) and the European Community.

Especially for this second edition I thank Prof. Dr. Ralph von Baltz (Karlsruhe) for his tremendous help in preparing the chapter on the Bloch equations as well as for his constructive discussions on many other topics of this book. Equally vivid thanks are due to Prof. Dr. Kurt Hümmer (Karlsruhe/Forchheim) for his help and improvement of the chapter on group theory and his lucent comments on crystal optics.

A lot of thanks also to all, who tried to solve "the final problem" in Sect. 27.5 of the first edition. Among these were several of my former and present co-workers. To mention at least a few I should like to name Profs. Drs. Ulrike Woggon, Harald Giessen, Drs. Alexander Jolk, Alexander Dinger and Markus Göppert. Considerable help to solve "the final Problem" came also from colleagues from abroad, who partly used this book for their lectures like Prof. Dr. Carl G. Ribbing (Uppsala), Prof. Dr. B. Stébé (Metz) or Dr. Alexander G. Umnov (at that time at Kawasaki – City) to name just a few.

Last but not least, I should like to thank my secretary Ms Monika Brenkmann for careful and patient typing of corrections and Ms Ursula Bolz for new drawing as well as the Publishing House Springer and there especially Drs. J. Koelsch and Th. Schneider for the excellent cooperation in the production of this new edition.

Karlsruhe, *Claus F. Klingshirn*
September 2004

Contents

1

Introduction

This introductory chapter consists of an outline of the fundamental concepts and ideas on which the text is based, including the rather limited prerequisites so that the reader can follow it and, finally, some hints about its contents.

1.1 Aims and Concepts

The aim of this book is to explain the optical properties of semiconductors, e.g., the spectra of transmission, reflection and luminescence, or of the complex dielectric function in the infrared, visible and near-ultraviolet part of the electromagnetic spectrum. We want to evoke in the reader a clear and intuitive understanding of the physical concepts and foundations of semiconductor optics and of some of their numerous applications. To this end, we try to keep the mathematical apparatus as simple and as limited as possible in order not to conceal the physics behind mathematics. We give ample references for those who want to enter more deeply into the mathematical concepts [62F1, 74B1, 75Z1, 76A1, 81M1, 88Z1, 91D1, 91L1, 93B1, 93H1, 93O1, 95I1, 95M1, 96S1, 97B1, 02S1].

Though many devices are based on the optical properties of semiconductors like photodiodes and solar cells or light emitting and laser diodes, we will not go into the details of such devices except for laser diodes, which are shortly treated in Chap. 22. Information on these topics can be found, e.g., in [65S1, 85P1, 86P1, 92E1, 94C1, 97E1, 97N1, 99B1, 00I1, 02C1, 02S2].

In this spirit, this present textbook is not only suitable for graduate and postgraduate students of physics, but also for students of neighboring disciplines, such as material science and electronics.

The prerequisites for the reader are an introductory or undergraduate course in general physics and some basic knowledge in atomic physics and quantum mechanics. The reader should know, for example, what the Schrödinger equation is, what the words eigen- (or proper-) state and eigen energy mean, and what quantum mechanics predicts about plane waves, the

hydrogen atom or the harmonic oscillator, how to calculate transition probabilities e.g., by Fermi's golden rule. Some basic knowledge of solid state physics will facilitate reading of this book, although the basic concepts will always be outlined here.

At the end of every chapter we give several problems which can be solved with the information given in the text, combined with some basic knowledge of physics, some thinking and some creativity. For fields which are actually rather active we give at the end of the corresponding sections references with only short explanations, which allow the reader to enter such a field of research e.g. in the preparation of a PhD thesis.

1.2 Outline of the Book and a lot of References

In the first part of this book (Chaps. 2–18 we shall present the linear optical properties of semiconductors. We start in Chap. 2 with Maxwell's equations, photons and the density of states and introduce in Chap. 3 the basic concepts of the interaction of light with matter. In Chaps. 4–6 a model system of oscillators is treated with respect to the optical properties which can be expected for such a system. Chapters 7–10 are used to introduce the elementary excitations or quasiparticles in semiconductors, followed by a presentation of the linear optical properties resulting from the interaction of these quasiparticles with light in Chaps. 11–17. Chapter 18 gives a short résumé of the linear optical properties of semiconductors. We include in Chaps. 7–17 modern concepts of semiconductor optics such as the properties of systems of reduced dimensionality, e.g. quantum wells, microcavities, photonic crystals or disordered systems which lead to localization.

At present, more than 600 different semiconductor materials are known. Many of them and their properties are listed in several volumes of Landolt–Börnstein [82L1,01L1]. We shall concentrate here on the most important ones. They are usually tetrahedrally coordinated and comprise, e.g. the group IV elements Si and Ge, the III–V compounds such as GaAs, the IIb–VI semiconductors such as CdS or ZnSe, and the Ib–VII materials such as the Cu halides.

Chapters 19–24 contain the main aspects of the nonlinear optical properties of semiconductors including optical gain and lasing as an example for an application of nonlinear optical properties.

In Chaps. 25–27, which can be considered as a kind of appendix, we shall outline some experimental techniques of semiconductor spectroscopy and some elements of group theory which are relevant for the description of semiconductor optics and a pedestrian approach to semiconductor Bloch equations including some applications of this concept.

In the sections on the linear and on the nonlinear optical properties of semiconductors, the main emphasis is placed on those properties which are

connected with excitations in the electronic system of semiconductors, since these aspects have obtained the widest interest both in fundamental and applied research as can be seen from an inspection of the conference series mentioned in the preface. However, we give also broad information on phonons and other quasiparticles or elementary excitations in semiconductors and on their optical properties.

In most chapters or sections a selection of references will be given for further reading which penetrate deeper into the topic, consider some further aspects, or give a more detailed theoretical description. Since the number of original publications, conference proceedings or summer schools on the topics covered here is "close to infinite", it is definitely only possible to cite a very small fraction of them, the choice of which is partly arbitrary and determined by the author's research interest. Furthermore, we shall not give references at all for things which can be considered to belong to the "general education or culture" in physics but we give references to the sources of original data in the figures. These figures have all been redrawn and generally modified for the didactic purpose of this textbook. We apologize for these deficiencies.

The present book thus complements the textbooks like [75B1, 77L1, 86L1, 89L1, 90K1, 91D1, 03T1] which concentrate more on atomic and molecular spectroscopy and on solid state spectroscopy in general. A rather remarkable series of books on various aspects of optical properties of solids, with some emphasis on insulators, results from the International Schools on "Atomic and Molecular Spectroscopy" held every two years in Erice (Sicily) [81A1]. Other series, which contain a lot of information on solid state optics are listed in [55S1, 62F1, 66S1, 82M1].

A selection of textbooks on general solid state physics are available [73H1, 75Z1, 76A1, 81H1, 81M1, 89K1, 93K1, 95A1, 95C1, 95W1, 00M1]. Semiconductor physics is treated generally, e.g., in [80N1, 91E1, 91S1, 92E2, 93O1, 96Y1, 97S1, 99G1, 00L1, 01H2, 01S1, 02D1], general optics in [01H3, 95L1], optical properties of solids in [59M1, 69O1, 72W1, 75B1, 77L1, 86L1, 89L1, 90K1, 91D1], including some older work.

For semiconductor optics, semiconductor structures of reduced dimensionality or semiconductor growth, including some specialized topics see, e.g., [59M1, 84H1, 86U1, 88Z1, 90G1, 91L1, 92L1, 92S1, 93B1, 93H1, 93O1, 93P1, 93S1, 94C1, 95C1, 96K1, 96O1, 96S2, 97B1, 97W1, 98D1, 98G1, 98J1, 98R1, 98S1, 99B1, 99M1, 00A1, 01C1, 01H1, 02R1, 02S1].

For various aspects of nonlinear optics and spectroscopy see, e.g., [77L1, 84S1, 86L1, 88Z1, 89L1, 93O1, 95M1, 96H1, 96S2, 98M1, 02S1, 04O1].

Recent data collections on bulk semiconductors and on optical properties of quantum structures are compiled in [82L1, 01L1]. The Volumes III 17a, 34C1 and 41A1 also contain condensed treatments of the underlying physics and of experimental techniques.

Complimentary information on the optical properties of metals, which are obviously not a topic of this book, can be found in [72W1, 90K1, 02D1].

1.3 Some Personal Thoughts

At the end of this introduction we want to consider some more general, partly historic or even philosophic aspects in connection with semiconductor optics. We mentioned in the preface of the first edition, that semiconductor optics will be an active and exciting field of research well into the next (now the present) century. We want to dwell on this aspect here a little bit longer. The quantum mechanical understanding of the electronic system of matter started from atoms and developed over molecules to three-dimensional solid, resulting in the beautiful concepts of quasi particles, band structures, etc., which we will outline in a didactic fashion in the following chapters. When these concepts were established, an opposite trend appeared, namely to go backwards from three-dimensional semiconductors (or more generally solids) to structures of reduced dimensionality like quasi two-dimensional quantum wells, quasi one-dimensional quantum wires and finally quasi zero-dimensional quantum dots also known as artificial atoms. We shall also treat these aspects in this book in detail. Presently, a repetition of this development is starting in the sense the that quantum dots are assembled to form one-, two- or three-dimensional arrays and photonic atoms are put together to form photonic crystals in one or two dimensions.

In this sense one has the impression that the field of semiconductor science, including optics, tends towards maturity. It seems difficult to reduce the quasi dimensionality of semiconductor quantum structures below zero, or to do spectroscopy with laser pulses, which are shorter than one or a few cycles of light or with intensities or fluences exceeding those which are sufficient to melt or to evaporate the sample.

On the other hand, at the time of the finishing of this manuscript (end of 2003) there were many open, somewhat controversially discussed and rapidly developing fields of basic and applied semiconductor research, which include, e.g., excitonic Bose–Einstein condensation, photonic crystals (or ∼band gap materials), understanding of the spin properties, THz spectroscopy, organic semiconductors or the development of reliable, long lived semiconductor laser diodes for the whole visible spectrum including the near UV and IR for display purposes, data storage or optical (glass) fiber communication.

Furthermore, one can expect that the spectroscopic techniques developed in semiconductor optics and the theoretical concepts (especially those which are not based on the translational invariance of a crystal) can be used efficently to contribute to the exploration and understanding of materials other than conventional inorganic semiconductors like macromolecules, soft matter, organic semiconductors or (nano-) biophysics, just to mention some of the key words that are presently en vogue.

The author himself has been doing research in semiconductor optics together with his co-workers for more than 30 years.

During this time he has noticed that many topics are in style for some time and then disappear, partly because they are understood, partly because

they are too difficult to be understood or handled and partly simply because something new is being developed.

Strangely enough, the "old" topics tend to reappear after ten or 20 years as something terribly new or modern. To mention only a few recent examples we recall excitonic Bose–Einstein condensation, biexcitons in quantum dots in glass or organic matrices or on the material side GaN or ZnO, which are presently seeing a renaissance. Generally, a new aspect is indeed added, like a reduced dimensionality or a better spatial or temporal resolution. However, often there is a mere reinvention of things that are already known and the new generation of scientists claims some "firsts" because they are not aware of the older work. When this is done not by ignorance, but deliberately, it is especially annoying.

When the author was himself a young PhD student and he or other members of the institute approached their "Doktorvater" Prof. Dr.E. Mollwo with some terribly exciting new results, he often used to state, "Ich wundere mich, aber ich wundere mich nicht sehr" (I am surprised, but I am not very much surprised) and recall some similiar or related phenomenon, which had been investigated some twenty or forty years ago. The author is presently at a stage of age (or possibly wisdom) that he can appreciate this attitude. In this context one could also mention Ben Akiba who cited "Es geschieht nichts Neues unter der Sonne" (There is nothing new under the sun) or more simply phrased, "Alles schon mal da gewesen". This experience should however by no means discourage young (or old) scientists from enthusiastically following their research projects to develop new ideas and concepts and to venture into new fields.

1.4 Problems

1. What are the basic conservation laws in nature?
2. Try to remember some of the basic concepts of quantum mechanics:
   - What is the Hamiltonian in classical and in quantum mechanics?
   - Write down the time-independent and the time-dependent Schrödinger equation for a single particle.
   - What are the eigenenergies and eigenfunctions of a one-dimensional harmonic oscillator and of the hydrogen atom?
   - What can you calculate with time-independent perturbation theory?
   - What does Fermi's golden rule say about transition probabilities?
   - Did you hear terms like density matrix formalism or second quantization? If yes, what do they mean?

References to Chap. 1

[55S1] Solid State Physics, Academic Press, Boston. This series started in 1955 and reached until now approx. 60 Volumes

[59M1] T.S. Moss, Optical Properties of Semiconductors, Butterworth, London (1959)
[62F1] The series Festkörperprobleme/Advances in Solid State Physics, Vol. 1–44 (1962–2005) published by Vieweg (Braunschweig) and recently by Springer, Berlin, contains invited contributions from the Spring Meeting of the German Physical Society
[65S1] E. Spenke, Elektronische Halbleiter, 2nd ed., Springer, Berlin (1965)
[66S1] Semiconductor and Semimetals, Academic Press, Boston. This series started in 1966 and reached until now approx. 75 Volumes
[69O1] Optical Properties of Solids, S. Nudelman and S.S. Mitra eds., NATO ASI series, Plenum Press, New York (1969)
[72W1] F. Wooten, Optical Properties of Solids, Academic Press, New York (1972)
[73H1] H. Haken, Quantenfeldtheorie des Festkörpers, Teubner, Stuttgart (1973)
[74B1] G.L. Bir and G.E. Pikus, Symmetry and Strain Induced Effects in Semiconductors, Wiley, New York (1974)
[75B1] F. Bassani, G.P. Paravicini, Electronic States and Optical Transitions in Solids, Pergamon, Oxford (1975)
[75Z1] J.M. Ziman, Prinzipien der Festkörpertheorie, Harry Deutsch, Zürich (1975) and Principles of the Theory of Solids, 2nd ed. Cambridge University Press, Cambridge (1992)
[76A1] N.W. Ashcroft and N.D. Mermin, Solid State Physics, Holt-Saunders, New York (1976)
[77L1] V.S. Letokhov and V.P. Chebotayev, Nonlinear Laser Spectroscopy, Springer Series in Optical Sciences **4**, Springer, Berlin (1977)
[80N1] S. Nakajima, Y. Toyozawa and R. Abe, The Physics of Elementary Excitations, Springer Series in Solid State Sciences **12**, Springer Berlin (1980)
[81A1] The proceedings of the International School on ,,Atomic and Molecular Spectroscopy" are held every two years in Erice (Sicily), are edited by B. Di Bartolo (Boston) and are actually devoted to optical properties of solids. Some of the recent topics are:
  a. Collective Excitations in Solids (1981), NATO ASI Ser. B **88**, Plenum Press, New York (1983)
  b. Energy Transfer Processes in Condensed Matter (1983), NATO ASI Ser. B **114**, Plenum Press, New York (1984)
  c. Spectroscopy of Solid-State Laser Type Materials (1985), Ettrore Majorana Int. Sci. Ser. **30**, Plenum Press, New York (1987)
  d. Disordered Solids: Structures and Processes (1987), Ettore Majorana Int. Sci. Ser. **46**, Plenum Press, New York (1989)
  e. Advances in Nonradiative Processes in Solids (1989), NATO ASI Ser. B **249**, Plenum Press, New York (1991)
  f. Optical Properties of Excited States in Solids (1991), NATO ASI Ser. B **301**, Plenum Press, New York (1994)
  g. Nonlinear Spectroscopy of Solids: Advances and Application (1993), NATO ASI Ser. B **339**, Plenum Press, New York (1994)
  h. Spectroscopy and Dynamics of Collective Excitations in Solids (1995), NATO ASI Ser. B **356**, Plenum Press, New York (1997)
  i. Ultrafast Dynamics of Quantum Systems (1997), NATO ASI Ser. B **372**, Plenum Press, New York (1998)
  j. Advances in Energy Transfer Processes (1999), World Scientific, Hongkong (2001)

k. Spectroscopy of Systems with Spatially Confined Structures (2001), NATO Science Series II **90**, Kluwer, Dordrecht (2002)
l. Frontiers of Optical Spectroscopy (2003), Kluwer, Dordrecht, in press (2004)
m. New Developments in Optic and Related Fields: Modern Techniques, Materials and Applications (2005) to be published by Kluver

[81H1] K.-H. Hellwege, Einführung in die Festkörperphysik, 2nd ed., Springer, Berlin (1981)

[81M1] O. Madelung, Introduction to Solid State Theory, Springer Series in Solid State Sciences **2**, Springer, Berlin (1981)

[81S1] S.M. Sze, Physics of Semiconductor Devices, John Wiley and Sons, New York (1981) and Semiconductor Devices, 2nd ed. ibid. (2002)

[82L1] Landolt–Börnstein, New Series, Group III, Vol. **17** a to i, **22** a and b, **41** A to D, ed. by O. Madelung and U. Rössler, Springer, Berlin (1982–2001)

[82M1] Some further series contain valuable information on semiconductor optics e.g. Modern Problems in Condensed Matter Sciences; V.M. Agranovich and A.A. Maradudin eds., North Holland, Amsterdam, The series started in 1982. We mention here especially the volumes
a. Vol. 1 Surface Polaritons, V.M. Agranovich and D.L. Mills eds.
b. Vol. 2 Excitons, E.I. Rashba and M.D. Sturge eds.
c. Vol. 6 Electron-Hole Droplets in Semiconductors, C.D. Jeffries and L.V. Keldysh eds.
d. Vol. 9 Surface Excitations, V.M. Agranovich and R. Landon eds.
e. Vol. 10 Electron-Electron Interactions in Disordered Systems, A.L. Efros and M. Pollak eds.
f. Vol. 16 Non equilibrium Phonons in Nonmetallic Crystals, W. Eisenmenger and A.A. Kaplyanski eds.
g. Vol. 22 Spin Waves and Magnetic Excitations, A.S. Borovic-Romanov and S.K. Sinha eds.
h. Vol. 23 Optical Properties of Mixed Crystals, R.J. Elliott and I.P. Ipatova eds.
i. Vol. 24 The Dielectric Function of Condensed Systems, L.V. Keldysh, D.A. Kirzhnitz and A.A. Maradudin eds.

[84H1] H. Haug, S. Schmitt-Rink, Prog. Quantum Electron. **9**, 3 (1984)

[84S1] Y.R. Shen, The Principles of Nonlinear Optics, Wiley, New York (1984)

[85P1] R. Paul, Optoelektronische Halbleiterbauelemente, Teubner, Stuttgart (1985)

[86L1] Laser Spectroscopy of Solids, 2nd ed. Ed. By W.M. Yen, P.M. Seltzer, Topics Appl. Phys. **49**, Springer, Berlin, Heidelberg (1986)

[86P1] R. Paul, Elektronische Halbleiterbauelemente, Teubner, Stuttgart (1986)

[86U1] M. Ueta, H. Kazanki, K. Kobayashi, Y. Toyozawa, E. Hanamura, Excitonic Processes in Solids, Springer Ser. Solid State Sci. Vol. **60**, Springer, Berlin, Heidelberg (1986)

[88Z1] R. Zimmermann, Many-Particle Theory of Highly Excited Semiconductors, Teubner Texte Phys. **18**, Teubner, Leipzig (1988)

[89K1] K. Kopitzky, Einführung in die Festkörperphysik, Teubner, Stuttgart (1989)

[89L1] Laser Spectroscopy of Solids II, ed. By W.M. Yen, Topics Appl. Phys. **65**, Springer, Berlin (1989)

[90G1] Growth and Characterisation of Semiconductors, R.A. Stradling and P.C. Klipstein eds., Adam Hilger, Bristol (1990)

[90K1] H. Kuzmany, Festkörperspektroskopie Springer, Berlin, Heidelberg (1990)
[91D1] W. Demtröder, Laserspektroskopie, 2nd edn. Springer, Berlin, Heidelberg (1991)
[91E1] Electronic Structure and Properties of Semiconductors: R.W. Cahn, P. Haasen, E.J. Kramer (eds.), Materials Science and Technology, Vol. 4, VCH, Weinheim (1991)
[91L1] P.T. Landsberg, Recombination in Semiconductors Cambridge Univ. Press, Cambridge (1991)
[91S1] H. Schaumburg, Halbleiter, B.G. Teubner, Stuttgart (1991)
[92E1] K.J. Ebeling, Integrierte Optoelektronik, 2nd ed., Springer, Berlin (1992)
[92E2] R. Enderlein and A. Schenk, Grundlagen der Halbleiter-Physik, Akademie Verlag, Berlin (1992)
[92L1] Low-Dimensional Electronic Systems, G. Bauer, F. Kuchar and H. Heinrich eds., Springer Series in Solid State Sciences **111**, Springer, Berlin (1992)
[92S1] The Spectroscopy of Semiconductors and Semimetals **36**, D.G. Seiler and Ch.L. Littler eds., Academic Press, Boston (1992)
[93B1] L. Banyai and S.W. Koch, Semiconductor Quantum Dots, World Scientific, Singapore (1993)
[93H1] H. Haug and S.W. Koch: Quantum Theory of the Optical and Electronic Properties of Semiconductors, 2nd ed. World Scientific, Singapore (1993)
[93K1] Ch. Kittel, Festkörperphysik 10th ed., Oldenbourg, Wien (1993)
[93O1] Optics of Semiconductor Nanostructures, F. Henneberger, S. Schmitt-Rink and E.O. Göbel eds., Academic Press, Berlin (1993)
[93P1] N. Peyghambarian, S.W. Koch and A. Mysyrowicz, Introduction to Semiconductor Optics, Prentice Hall, Englewood Cliffs, NJ (1993)
[93S1] J. Singh, Physics of Semiconductors and their Heterostructures, Mc Graw-Hill Inc., New York (1993)
[94C1] W.W. Chow, S.W. Koch and M. Sargent III, Semiconductor Laser Physics, Springer, Berlin (1994)
[95A1] An-Ban Chen and A. Sher, Semiconductor Alloys, Plenum Press, New York (1995)
[95C1] J.R. Christman, Festkörperphysik 2nd ed., Oldenbourg, München (1995)
[95I1] H. Ibach and H. Lüth, Ferstkörperphysik, 4th ed., Springer, Berlin (1995)
[95I2] E.L. Ivchenko and G. Pikus, Superlattices and other Heterostructures, Springer Series in Solid State Sciences **110** (1995)
[95L1] S.G. Lipson, H.S. Lipson and D.S. Tannhauser, Optical Physics 3^{rd} ed., Cambridge University Press, Cambridge (1995) and Optik, Springer, Berlin, Heidelberg, New York (1997)
[95M1] S. Mukamel, Principles of Nonlinear Optical Spectroscopy, Oxford University Press, Oxford (1995)
[95W1] Ch. Weißmantel and C. Hamann, Grundlagen der Festkörperphysik, 4^{th} ed., Johann Ambrosius Barth, Heidelberg (1995)
[96H1] H. Haug and A.-P. Jauho, Quantum Kinetics in Transport and Optics of Semiconductors, Springer Series in Solid State Sciences **123**, Springer, Berlin (1996)
[96K1] H. Kalt, Optical Properties of III–V Semiconductors, Springer Series in Solid State Sciences **120** (1996)
[96O1] Optical Characterization of Epitaxial Semiconductor Layers, G. Bauer and W. Richter eds., Springer, Berlin (1996)

[96S1] A.P. Sutton, Elektronische Struktur in Materialien, VCH, Weinheim (1996)
[96S2] J. Shah, Ultrafast Spectroscopy of Semiconductors and of Semiconductor Nanostructures, Springer Series in Solid State Sciences **115** (1996).
[96Y1] P.Y. Yu and M. Cardona, Fundamentals of Semiconductors, Springer, Berlin (1996)
[97B1] P.K. Basu, Theory of Optical Processes in Semiconductors, Clarendon Press, Oxford (1997)
[97E1] R. Enderlein and N.J.M. Horing, Fundamentals of Semiconductor Physics and Devices, World Scientific, Singapore (1997)
[97N1] S. Nakamura and G. Fasol, The Blue Laser Diode, Springer, Berlin (1997)
[97S1] K. Seeger, Semiconductor Physics, 6th ed., Springer Series on Solid State Sciences **40**, Springer, Berlin (1997)
[97W1] U. Woggon, Optical Properties of Semiconductor Quantum Dots, Springer Tracts in Modern Phyiscs **136**, Springer, Berlin (1997)
[98D1] J.H. Davis, The Physics of Low-Dimensional Semiconductors, Cambridge University Press, Cambridge (1998)
[98E1] S.R. Elliot, The Physics and Chemistry of Solids, Wiley, New York (1998)
[98G1] S.V. Gaponenko, Optical Properties of Semiconductor Nanocrystals, Cambridge University Press, Cambridge (1998)
[98J1] L. Jacah, P. Hawrylak and A. Wójs, Quantum Dots, Springer, Berlin (1998)
[98M1] D.L. Mills, Nonlinear Optics 2nd ed., Springer, Berlin (1998)
[98R1] T. Ruf, Phonon Raman Scattering in Semiconductors Quantum Wells and Superlattices, Springer Tracts in Modern Physics **142**, Springer, Berlin (1998)
[98S1] A. Shik, Quantum Wells, World Scientific, Singapore (1998)
[99B1] K.F. Brennan, The Physics of Semiconductors Applied to Optoelectronic Devices, Cambridge University Press, Cambridge (1999)
[99B2] G. Bastard, Wave Mechanics Applied to Semiconductor Heterostructures, Les éditons de Physique, Les Ulis (1999)
[99G1] H.T. Grahn, Introduction to Semiconductor Physics, World Scientific, Singapore (1999)
[99M1] V.V. Mitin, V.A. Kochelap and M.A. Stroscio, Quantum Heterostructures and Optoelectronics, Cambridge University Press, Cambridge (1999)
[00A1] H. van Amerongen, R. van Grondelle and L. Valkunas, Photosynthetic Excitons, World Scientific, Singapore (2000)
[00I1] Introduction to Nitride Semiconductor Blue Lasers and Light Emitting Diodes, S. Nakamura and S.F. Chichibu eds., Taylor and Francis, London (2000)
[00L1] M.E. Levinshtein and G.S. Simin, Getting to know Semiconductors, World Scientific, Singapore (2000)
[00M1] M.P. Marder, Condensed Matter Physics (corrected printing), Wiley, New York (2000)
[01C1] D.S. Chemla, J. Shah, Nature **411**, 549 (2001)
[01H1] P. Harrison, Quantum Wells, Wires and Dots, John Wiley and Sons, New York (2001)
[01H2] Ch. Hamaguchi, Basic Semiconductor Physics, Springer, Berlin (2001)
[01H3] E. Hecht, Optik, 3rd (ed.), Oldenbourg, München (2001) and Optics, Addison Wesley Longman, (1998)
[01L1] Landolt–Börnstein, New Series, Group III, Vol. **34**C, Parts 1 and 2, ed. by C. Klingshirn, Springer Berlin (2001) and (2004) and Part 3 in preparation

[01S1] U. Schumacher, Halbleiter, 2nd ed., Publicis MCD Corporate Publishing, Erlangen (2001)
[02C1] J.P. Colinge, Physics of Semiconductor Devices, Kluwer, Dordrecht (2002)
[02D1] M. Dressel and G. Grüner, Electrodynamics of Solids, Optical Properties of Electrons in Matter, Cambridge University Press, Cambridge (2002)
[02R1] E. Runge and V. May (eds.), phys. stat. sol. (b) **234**, (1) (2002)
[02S1] W. Schäfer and M. Wegener, Semiconductor Optics and Transport Phenomena, Springer, Berlin (2002)
[02S2] S.M. Sze Semiconductor Device, Physics and Technology, 2^{nd} ed. J. Wiley and Sons, New York (2002)
[03T1] Y. Toyozawa, Optical Processes in Solids, Cambridge, University Press, Cambridge (2003)
[04O1] Optics of Semiconductors and Their Nanostructures, H. Kalt and M. Hetterich (eds.), Springer Series in Solid State Sciences **146**, (2004)

2

Maxwell's Equations, Photons and the Density of States

In this chapter we consider Maxwell's equations and what they reveal about the propagation of light in vacuum and in matter. We introduce the concept of photons and present their density of states. Since the density of states is a rather important property in general and not only for photons, we approach this quantity in a rather general way. We will use the density of states later also for other (quasi-) particles including systems of reduced dimensionality. In addition, we introduce the occupation probability of these states for various groups of particles.

It should be noted, that we shall approach the concept of photons on an elementary level only, in correspondence with the concept of this book. We do not delve into present research topics on photon physics itself like photon-correlation and -statistics, squeezed light, photon anti-bunching, entangled photon states, etc., but give some introductory references for those interested in these fields [89S1, 92M1, 94A1, 01M1, 01T1, 01T2, 02B1, 02D1, 02G1, 02L1, 02Y1]. Einstein, who obtained the Nobel prize for physics in 1921 for the explanation of the photo-electric effect (not for the theory of relativity!), once stated: "Was das Licht sei, das weiß ich nicht" (What the light might be, I do not know). So there still seems to be ample place for research in these fields.

2.1 Maxwell's Equations

Maxwell's equations can be written in different ways. We use here the macroscopic Maxwell's equations in their differential form. Throughout this book the internationally recommended system of units known as SI (systéme international) is used. These equations are given in their general form in (2.1a–f), where bold characters symbolize vectors and normal characters scalar quantities.

$$\nabla \cdot \boldsymbol{D} = \rho\,, \qquad \nabla \cdot \boldsymbol{B} = 0\,, \tag{2.1a,b}$$

$$\nabla \times \boldsymbol{E} = -\dot{\boldsymbol{B}}\,, \qquad \nabla \times \boldsymbol{H} = \boldsymbol{j} + \dot{\boldsymbol{D}}\,, \tag{2.1c,d}$$

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}\,, \qquad \boldsymbol{B} = \mu_0 \boldsymbol{H} + \boldsymbol{M}\,. \tag{2.1e,f}$$

The various symbols have the following meanings and units:

$\boldsymbol{E}$ = electric field strength; $1\,\mathrm{V/m} = 1\,\mathrm{m\,kg\,s^{-3}\,A^{-1}}$
$\boldsymbol{D}$ = electric displacement; $1\,\mathrm{A\,s/m^2} = 1\,\mathrm{C/m^2}$
$\boldsymbol{H}$ = magnetic field strength; $1\,\mathrm{A/m}$
$\boldsymbol{B}$ = magnetic induction or magnetic flux density ; $1\,\mathrm{V\,s/m^2} = 1\,\mathrm{T} = 1\,\mathrm{Wb/m^2}$
ρ = charge density; $1\,\mathrm{A\,s/m^3} = 1\,\mathrm{C/m^3}$
$\boldsymbol{j}$ = electrical current density; $1\,\mathrm{A/m^2}$
$\boldsymbol{P}$ = polarization density of a medium, i.e., electric dipole moment per unit volume; $1\,\mathrm{A\,s/m^2}$
$\boldsymbol{M}$ = magnetization densityof the medium, i.e., magnetic dipole moment per unit volume[1]; $1\,\mathrm{V\,s/m^2}$
$\varepsilon_0 \simeq 8.859 \times 10^{-12}\,\mathrm{A\,s/V\,m}$ is the permittivity of vacuum
$\mu_0 = 4\pi \times 10^{-7}\,\mathrm{V\,s/A\,m}$ is the permeability of vacuum
∇ = Nabla-operator, in Cartesian coordinates $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$
$\dot{} = \partial/\partial t$ i.e., a dot means differentiation with respect to time.

The applications of ∇ to scalar or vector fields are usually denoted by

$$\begin{aligned} \nabla \cdot f(\boldsymbol{r}) &= \operatorname{grad} f, \\ \nabla \cdot \boldsymbol{A}(\boldsymbol{R}) &= \operatorname{div} \boldsymbol{A}, \\ \nabla \times \boldsymbol{A}(\boldsymbol{r}) &= \operatorname{curl} \boldsymbol{A}, \end{aligned}$$

and the Laplace operator Δ is defined as

$$\Delta \equiv \nabla^2.$$

If Δ is applied to a scalar field ρ we obtain

$$\Delta\rho = \frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} + \frac{\partial^2 \rho}{\partial z^2} \tag{2.2}$$

Application to a vector field E results in

$$\Delta \boldsymbol{E} = \begin{pmatrix} \dfrac{\partial^2 E_x}{\partial x^2} + \dfrac{\partial^2 E_x}{\partial y^2} + \dfrac{\partial^2 E_x}{\partial z^2} \\ \dfrac{\partial^2 E_y}{\partial x^2} + \dfrac{\partial^2 E_y}{\partial y^2} + \dfrac{\partial^2 E_y}{\partial z^2} \\ \dfrac{\partial^2 E_z}{\partial x^2} + \dfrac{\partial^2 E_z}{\partial y^2} + \dfrac{\partial^2 E_z}{\partial z^2} \end{pmatrix}. \tag{2.3}$$

[1] Some authors prefer to use $\boldsymbol{M'} = \boldsymbol{M}\mu_0^{-1}$ and thus $\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M'})$. We prefer (2.1e,f) for symmetry arguments.

Further rules for the use of ∇ and of Δ and their representations in other than Cartesian coordinates (polar or cylindrical coordinates) are found in compilations of mathematical formulae [84A1, 91B1, 92S1].

Equations (2.1a,b) show that free electric charges ρ are the sources of the electric displacement and that the magnetic induction is source-free. Equations (2.1c,d) demonstrate how temporally varying magnetic and electric fields generate each other. In addition, the $\boldsymbol{H}$ field can be created by a macroscopic current density $\boldsymbol{j}$. Equations (2.1e,f) are the material equations in their general form. From them we learn that the electric displacement is given by the sum of electric field and polarization, while the magnetic flux density is given by the sum of magnetic field and magnetization. Some authors prefer not to differentiate between $\boldsymbol{H}$ and $\boldsymbol{B}$. This leads to difficulties, as can be easily seen from the fact that $\boldsymbol{B}$ is source-free (2.1b) but $\boldsymbol{H}$ is not, as follows from the inspection of the fields of every simple permanent magnet.

By applying $\boldsymbol{\nabla}\cdot$ to (2.1d) we obtain the continuity equation for the electric charges

$$\operatorname{div} \boldsymbol{j} = -\frac{\partial}{\partial t}\rho\,, \tag{2.4}$$

which corresponds to the conservation law of the electric charge in a closed system.

The integral forms of (2.1) can be obtained from the differential forms by integration and the use of the laws of Gauss or Stokes resulting in

$$\int \rho(\boldsymbol{r})\,\mathrm{d}V = \oint \boldsymbol{D}\cdot \mathrm{d}\boldsymbol{f} \tag{2.5a}$$

$$-\frac{\partial}{\partial t}\int \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{f} = \oint \boldsymbol{E}\cdot \mathrm{d}\boldsymbol{s} \tag{2.5b}$$

where $\mathrm{d}V$, $\mathrm{d}\boldsymbol{f}$ and $\mathrm{d}\boldsymbol{s}$ give infinitesimal elements of volume, surface or area and line, respectively.

In their microscopic form, Maxwells equations contain all charges as sources of the electric field $\boldsymbol{E}_{\mathrm{micro}}$ including all electrons, protons bound in atoms as ρ_{bound} and not only the free space charges ρ. By analogy, not only the microscopic current density $\boldsymbol{j}$ has to be used as a source of $\boldsymbol{H}_{\mathrm{micro}}$ but all spins and $l \neq 0$ orbits of charged particles have to be included as "bound" current density $\boldsymbol{j}_{\mathrm{bound}}$. The transition to macroscopic quantities can then be performed by averaging over small volumes (larger than an atom but smaller than the wavelength of light) and replacing ρ_{bound} by $-\overrightarrow{\nabla}\cdot\boldsymbol{P}$ and $\boldsymbol{j}_{\mathrm{bound}}$ by $\boldsymbol{P} + \operatorname{curl}\boldsymbol{M}/\mu_0$. For more details see [98B1, 98D1] or Chap. 27.

Concerning the units, some theoreticians still prefer the so-called c g s (cm, g, second) system. Though it has only marginal differences in mechanics to the SI system, which is based on the units 1 m, 1 kg, 1 s, 1 A, 1 K, 1 mol and 1 cd, the c g s system produces strange units in electrodynamics like the electrostatic units (esu), which contain square roots of

mass and are therefore unphysical and even ill-defined. For conversion tables see [96L1].

2.2 Electromagnetic Radiation in Vacuum

In vacuum the following conditions are fulfilled

$$\boldsymbol{P} = 0\,; \quad \boldsymbol{M} = 0\,; \quad \boldsymbol{\rho} = 0\,; \quad \boldsymbol{j} = 0. \tag{2.6}$$

With the help of (2.1e,f) this simplifies (2.1c,d) to

$$\nabla \times \boldsymbol{E} = -\mu_0 \dot{\boldsymbol{H}} \quad \text{and} \quad \nabla \times \boldsymbol{H} = \varepsilon_0 \dot{\boldsymbol{E}} \tag{2.7a,b}$$

Applying $\nabla\times$ to (2.7a) and $\partial/\partial t$ to (2.7b) yields

$$\nabla \times (\nabla \times \boldsymbol{E}) = -\mu_0 \nabla \times \dot{\boldsymbol{H}} \quad \text{and} \quad \nabla \times \dot{\boldsymbol{H}} = \varepsilon_0 \ddot{\boldsymbol{E}}\,. \tag{2.8}$$

From (2.8) we find with the help of the properties of the ∇ operator

$$-\mu_0 \varepsilon_0 \ddot{\boldsymbol{E}} = \nabla \times (\nabla \times \boldsymbol{E}) = \nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E}. \tag{2.9}$$

With (2.6), (2.3) and (2.1a) we see that

$$\nabla \boldsymbol{E} = 0 \tag{2.10}$$

and (2.9) reduces to the usual wave equation, written here for the electric field

$$\nabla^2 \boldsymbol{E} - \mu_0 \varepsilon_0 \ddot{\boldsymbol{E}} = 0\,. \tag{2.11}$$

An analogous equation can be obtained for the magnetic field strength. Solutions of this equation are all waves of the form

$$\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}_0 f(\boldsymbol{k}\boldsymbol{r} - \omega t). \tag{2.12}$$

$\boldsymbol{E}_0$ is the amplitude, f is an arbitrary function whose second derivate exists. As can be shown by inserting the ansatz (2.12) into (2.11) the wave vector $\boldsymbol{k}$ and the angular frequency ω obey the relation

$$\frac{\omega}{k} = \left(\frac{1}{\mu_0 \varepsilon_0}\right)^{1/2} = c \quad \text{with} \quad k = |\boldsymbol{k}| = 2\pi/\lambda_\mathrm{v}. \tag{2.13}$$

In the following we use for simplicity only the term "frequency" for $\omega = 2\pi/T$ where T is the temporal period of the oscillation.

In (2.13), c is the vacuum speed of light and λ_v is the wavelength in vacuum. From all possible solutions of the form (2.12) we shall concentrate in the following on the most simple ones, namely on plane harmonic waves, which can be written as

$$\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{k}\boldsymbol{r} - \omega t)]. \tag{2.14}$$

For all waves (not only those in vacuum), the phase and group velocities v_{ph} and v_{g} are given by

$$v_{\mathrm{ph}} = \frac{\omega}{k}; \quad \boldsymbol{v}_{\mathrm{g}} = \frac{\partial \omega}{\partial k} = \mathrm{grad}_{\boldsymbol{k}}\omega, \tag{2.15}$$

where v_{ph} gives the velocity with which a certain phase propagates, (e.g., a maximum of a monochromatic wave) while v_{g} gives the speed of the center of mass of a wave packet with middle frequency ω and covering a small frequency interval $\mathrm{d}\omega$ as shown schematically in Figs. 2.1a,b, respectively. The formulas (2.15) are of general validity. The $\mathrm{grad}_{\boldsymbol{k}}$ on the r.h.s. of (2.15) means a differentiation with respect to $\boldsymbol{k}$; in the sense of $\nabla_k = (\partial/\partial k_x, \partial/\partial k_y, \partial/\partial k_z)$ and has to be used instead of the more simple expression $\partial\omega/\partial k$ in anisotropic media. For the special case of electromagnetic radiation in vacuum we find from (2.13), (2.15)

$$v_{\mathrm{ph}} = v_{\mathrm{g}} = c = (\mu_0 \varepsilon_0)^{-1/2} \tag{2.16}$$

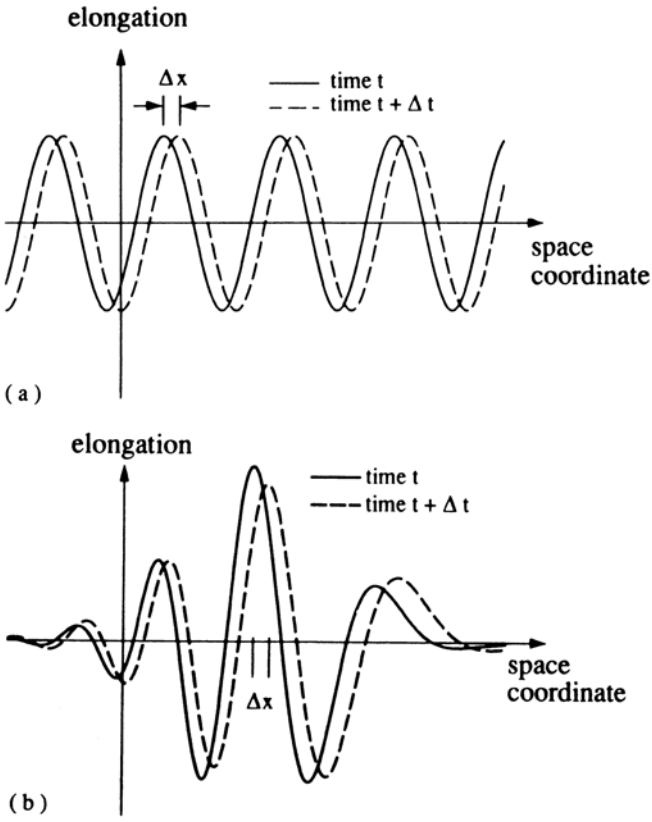


Fig. 2.1. A harmonic wave (**a**) and a wave packet (**b**) shown at two different times t and $t + \Delta t$ to illustrate the concepts of phase and group velocity, respectively

Now we want to see what constraints are imposed by Maxwell's equations on the various quantities such as $\boldsymbol{E}_0$ and $\boldsymbol{k}$. Inserting (2.12) or (2.14) into (2.10) gives

$$\nabla \cdot \boldsymbol{E} = \mathrm{i}\boldsymbol{E}_0 \cdot \boldsymbol{k} \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)] = 0. \tag{2.17}$$

This means that

$$\boldsymbol{E}_0 \perp \boldsymbol{k} \tag{2.18}$$

or, in other words, the electromagnetic wave is transverse in $\boldsymbol{E}$. What can we learn from Maxwell's equations for the other fields? From (2.7) we have for plane waves

$$\boldsymbol{H} = (\omega\mu_0)^{-1}\boldsymbol{k} \times \boldsymbol{E} = \boldsymbol{H}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)] \tag{2.19a}$$

with

$$\boldsymbol{H}_0 = (\omega\mu_0)^{-1}\boldsymbol{k} \times \boldsymbol{E}_0. \tag{2.19b}$$

Furthermore we have with (2.1e,f) and (2.6)

$$\boldsymbol{D} = \boldsymbol{D}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)] = \varepsilon_0 \boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)], \tag{2.19c}$$

$$\boldsymbol{B} = \boldsymbol{B}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)] = \omega^{-1}\boldsymbol{k} \times \boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)]. \tag{2.19d}$$

The electromagnetic wave is, according to (2.19b), also transverse in $\boldsymbol{B}$ and the electric and magnetic fields are perpendicular to each other, that is, we have in general

$$\boldsymbol{D} \perp \boldsymbol{k} \perp \boldsymbol{B} \perp \boldsymbol{D}. \tag{2.19e}$$

In vacuum and isotropic media one has in addition

$$\boldsymbol{E} \parallel \boldsymbol{D} \quad \text{and} \quad \boldsymbol{H} \parallel \boldsymbol{B}. \tag{2.19f}$$

As we shall see later in connection with (2.17) and (2.43),(2.44) one has in matter usually transverse waves, which obey (2.19e) but additionally, longitudinal waves exist under certain conditions.

The momentum density $\boldsymbol{\Pi}$ of the electromagnetic field is given by

$$\boldsymbol{\Pi} = \boldsymbol{D} \times \boldsymbol{B}, \quad \boldsymbol{\Pi} \parallel \boldsymbol{k} \tag{2.20}$$

and the energy flux density by the Poynting vector $\boldsymbol{S}$

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H} \tag{2.21}$$

with $\boldsymbol{S} \parallel \boldsymbol{\Pi}$ in vacuum and isotropic materials.

$\boldsymbol{S}$ is a rapidly oscillating function of space and time. The average value $\langle \boldsymbol{S} \rangle$ is usually called the intensity I or the energy flux density The intensity is proportional to the amplitude squared for all harmonic waves. For the plane monochromatic waves treated here, we obtain

$$\langle \boldsymbol{S} \rangle = \frac{1}{2}|\boldsymbol{E}_0 \times \boldsymbol{H}_0| = \frac{1}{2}\frac{1}{c\mu_0}\boldsymbol{E}_0^2 = \frac{1}{2}\frac{c}{\mu_0}\boldsymbol{B}_0^2 = \frac{1}{2}c\mu_0\boldsymbol{H}_0^2\,. \tag{2.22}$$

Equations (2.20) and (2.21) are also valid in matter.

2.3 Electromagnetic Radiation in Matter; Linear Optics

Now we treat Maxwell's equations in matter. Doing so we have in principle to use the equations in their general from (2.1). However we will still make in several steps some assumptions which are reasonable for semiconductors: we assume first that there are no macroscopic free space charges i.e. $\rho = 0$. Then a treatment of (2.1) in analogy with (2.7)–(2.11) results in

$$\nabla^2 \boldsymbol{E} - \mu_0 \varepsilon_0 \ddot{\boldsymbol{E}} = \mu_0 \ddot{\boldsymbol{P}} + \mu_0 \dot{\boldsymbol{j}} + \nabla \times \dot{\boldsymbol{M}}\,, \tag{2.23}$$

Actually, there also occurs in the derivation of (2.23) a term $\nabla(\nabla \boldsymbol{P})$ or $\nabla(\nabla \boldsymbol{E})$. This term vanishes for transverse waves and is therefore neglected if not mentioned otherwise.

This equation is the inhomogeneous analogue of (2.11) telling us that the sources of an electro-magnetic radiation field can be

- A dipole moment $\boldsymbol{p}$ or a polarization $\boldsymbol{P}$ with a non-vanishing second time derivative
- A temporally varying current density
- The curl of a temporally varying magnetization

Again a similar equation can be obtained for the magnetic field. We continue now with the application of further simplifications and assume that we have a nonmagnetic material, i.e., that the third term on the r.h.s. of (2.23) vanishes. Actually, all matter has some diamagnetism. But this is a rather small effect of the order of 10^{-6} so it can be neglected for our purposes. Paramagnetic and especially ferromagnetic contributions can be significantly larger for low frequencies. However, even these contributions diminish rapidly for higher frequencies. Consequently the assumption of a nonmagnetic material is a good approximation over a wide range of the electromagnetic spectrum starting in the IR even for ferromagnetic materials. Furthermore, the more common semiconductors are not ferro-, ferri- or antiferromagnetic and have only a small concentration of paramagnetic centres which may be seen in electron paramagnetic resonance (EPR), but which have negligible influence on the optical properties. The only exceptions are semiconductors which contain a considerable amount of e.g., Mn or Fe ions as does $Zn_{1-y}Mn_ySe$. We refer the reader to [88D1,91O1,92G1,92Y1,94D1,94G1,96H1,03D1] or to Sect. 16.1 and references therein for this class of materials.

The current term $\boldsymbol{j}$ in (2.1d) deserves some more consideration. The current is driven by the electric field

$$\boldsymbol{j} = \sigma \boldsymbol{E}, \tag{2.24}$$

where σ is the conductivity. For intrinsic or weakly doped semiconductors, the carrier density is small and consequently σ is as well. Then the following inequality holds

$$|\boldsymbol{j}| = |\sigma \boldsymbol{E}| \ll |\dot{\boldsymbol{D}}|. \tag{2.25}$$

In the following we will consider this case and neglect $\boldsymbol{j}$ in (2.1d). For heavily doped semiconductors (2.25) is no longer valid and σ will have some influence on the optical properties at least in the infrared (IR). We come back to this situation in connection with plasmons in Chaps. 10 and 12.

The basic material equation still left in comparison with the vacuum case is now (2.1e) $\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}$.

If we proceed with this equation again in the manner of (2.7)–(2.11) the result is

$$\nabla^2 \boldsymbol{E} - \mu_0 \varepsilon_0 \ddot{\boldsymbol{E}} = \mu_0 \ddot{\boldsymbol{P}} \tag{2.26}$$

Equation (2.26) states the well-known fact that every dipole $\boldsymbol{p}$ and every polarization $\boldsymbol{P}$ with a non vanishing second derivative in time radiates an electromagnetic wave.

As long as we have no detailed knowledge about the relationships between $\boldsymbol{D}$, $\boldsymbol{E}$ and $\boldsymbol{P}$ we cannot go beyond (2.26). Now we make a very important assumption. We assume a linear relationship between $\boldsymbol{P}$ and $\boldsymbol{E}$:

$$\frac{1}{\varepsilon_0} \boldsymbol{P} = \chi \boldsymbol{E} \tag{2.27a}$$

or

$$\boldsymbol{D} = \varepsilon_0 (1 + \chi) \boldsymbol{E} = \varepsilon \varepsilon_0 \boldsymbol{E} \tag{2.27b}$$

with

$$\varepsilon = \chi + 1. \tag{2.27c}$$

This linear relation is the reason why everything that is treated in the following Chaps. 3 to 18 is called linear optics. A linear relation is what one usually assumes between two physical quantities as long as one does not have more precise information. In principle we can also consider (2.27a) as an expansion of $\boldsymbol{P}(\boldsymbol{E})$ in a power series in $\boldsymbol{E}$ which is truncated after the linear term. We come back to this aspect in Chap. 19.[2] The quantities ε and χ are called the dielectric function and the susceptibility, respectively. They can be considered as linear response functions [93S1, 98B1, 98D1].

Both quantities depend on the frequency ω and on the wave-vector $\boldsymbol{k}$, and they both have a real and an imaginary part as shown for ε.

$$\varepsilon = \varepsilon(\omega, \boldsymbol{k}); \quad \chi = \chi(\omega, \boldsymbol{k}) = \varepsilon(\omega, \boldsymbol{k}) - 1, \tag{2.28}$$

$$\varepsilon(\omega, \boldsymbol{k}) = \varepsilon_1(\omega, \boldsymbol{k}) + \mathrm{i}\varepsilon_2(\omega, \boldsymbol{k}). \tag{2.29}$$

The frequency dependence is dominant and will be treated first in Chaps. 3 to 4. We drop the $\boldsymbol{k}$ dependence for the moment but come back to it in

[2] A constant term in this power expansion such as $\boldsymbol{P} = \boldsymbol{P}_0 + \chi E$ would describe a spontaneous polarization of matter which occurs e.g., in pyro- or ferro-electric materials. With arguments similar to the ones given for ferromagnetics we can neglect such phenomena in the discussion of the optical properties of semiconductors.

connection with spatial dispersion in Chap. 5. In Chap. 6 we discuss the properties of ε as a function of frequency and wave vector or as a function of time and space.

The value of $\varepsilon(\omega)$ for $\omega \simeq 0$ is usually called the dielectric constant.

In general ε and χ are tensors. For simplicity we shall consider them to be scalar quantities if not stated otherwise, e.g., in connection with birefringence in Sect. 3.1.7.

Using the linear relations of (2.27) we can transform (2.23) into

$$\nabla^2 \boldsymbol{E} - \mu_0 \varepsilon_0 \varepsilon(\omega) \ddot{\boldsymbol{E}} = 0, \tag{2.30a}$$

where we assumed also that $\varepsilon(\omega)$ is spatially constant on a length scale of the order of the wavelength of light. Deviations of this assumption are treated in Sects. 17.2–4.

If magnetic properties are to be included, a corresponding linear approach would lead to

$$\nabla^2 \boldsymbol{E} - \mu_0 \mu(\omega) \varepsilon_0 \varepsilon(\omega) \ddot{\boldsymbol{E}} = 0, \tag{2.30b}$$

where $\mu(\omega)$ is the magnetic permeability. As outlined above we have in the visible for most semiconductors $\mu(\omega) \simeq 1$.

As for (2.12) the solutions of (2.30) are again all functions of the type

$$\boldsymbol{E} = \boldsymbol{E}_0 f(\boldsymbol{k}\boldsymbol{r} - \omega t), \tag{2.31}$$

or for our present purposes, i.e. again for the case of plane harmonic waves

$$\boldsymbol{E} = \boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{k}\boldsymbol{r} - \omega t)]. \tag{2.32}$$

The relationship between $\boldsymbol{k}$ and ω is however now significantly different from (2.13). It follows again from inserting the ansatz (2.31 or 32) into (2.30) and now reads:

$$\frac{c^2 \boldsymbol{k}^2}{\omega^2} = \varepsilon(\omega). \tag{2.33}$$

This relation appears in Chap. 5 again under the name "polariton equation". It can also be written in other forms:

$$\boldsymbol{k} = \frac{\omega}{c} \varepsilon^{1/2}(\omega) = \frac{2\pi}{\lambda_\mathrm{v}} \varepsilon^{1/2}(\omega) = \boldsymbol{k}_\mathrm{v} \varepsilon^{1/2}(\omega), \tag{2.34}$$

where λ_v and $\boldsymbol{k}_\mathrm{v}$ refer to the vacuum values of the light wave.

For the square root of ε we introduce for simplicity a new quantity $\tilde{n}(\omega)$ which we call the complex index of refraction

$$\tilde{n}(\omega) = n(\omega) + \mathrm{i}\kappa(\omega) = \varepsilon^{1/2}(\omega). \tag{2.35}$$

The equations (2.13) and (2.33–35) can be interpreted in the following way. In vacuum an electromagnetic wave propagates with a wave vector $\boldsymbol{k}_\mathrm{v}$ which is real and given by (2.13). In matter, light propagates with a wave

vector $\boldsymbol{k}$ which can be a complex quantity given by (2.34), or, with the help of (2.35), by

$$k = \frac{\omega}{c}\tilde{n}(\omega) = \frac{\omega}{c}n(\omega) + \mathrm{i}\frac{\omega}{c}\kappa(\omega) = \frac{2\pi}{\lambda_\mathrm{v}}\tilde{n}(\omega) = k_\mathrm{v}\tilde{n}. \tag{2.36}$$

We should notice that k is for complex $\tilde{n}$ not simply $|\boldsymbol{k}|$ since $|\boldsymbol{k}|$ is always a positive, real quantity. Here k means just neglecting the vector character of $\boldsymbol{k}$ but k can still be a real, imaginary or complex quantity according to (2.36). The direction of the real part of $\boldsymbol{k}$, which describes the oscillatory part of the wave, is still parallel to $\boldsymbol{D} \times \boldsymbol{B}$ as in (2.20).

Writing the plane wave explicitly we have:

$$\boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{k}\boldsymbol{r} - \omega t)] = \boldsymbol{E}_0 \exp\left\{\mathrm{i}\left[\frac{\omega}{c}n(\omega)\hat{\boldsymbol{k}}\boldsymbol{r} - \omega t\right]\right\} \exp\left[-\frac{\omega}{c}\kappa(\omega)\hat{\boldsymbol{k}}\boldsymbol{r}\right] , \tag{2.37}$$

where $\hat{\boldsymbol{k}}$ is the unit vector in the direction of $\boldsymbol{k}$, i.e., in the direction of propagation.

Obviously $n(\omega)$ describes the oscillatory spatial propagation of light in matter; it is often called the refractive index in connection with Snells' law of refraction. This means that the wavelength λ in a medium is connected with the wavelength λ_v in vacuum by

$$\lambda = \lambda_\mathrm{v} n^{-1}(\omega). \tag{2.38}$$

In (2.37) $\kappa(\omega)$ describes a damping of the wave in the direction of propagation. This effect is usually called absorption or, more precisely, extinction. We give the precise meaning of these two quantities in Sect. 3.1.5. Here we compare (2.37) with the well-known law of absorption for the light intensity $\boldsymbol{I}$ of a parallel beam propagating in z-direction

$$I(z) = I(z = 0)\,\mathrm{e}^{-\alpha z} \tag{2.39}$$

where $\alpha(\omega)$ is usually called the absorption coefficient, especially in Anglo-Saxon literature. In German literature $\alpha(\omega)$ is also known as "Absorptionskonstante" (absorption constant) and dimensionless quantities proportional to $\kappa(\omega)$ are called "Absorptionskoeffizient" or "Absorptions index" (absorption coefficient or absorption index). So some care has to be taken regarding what is meant by one or the other of the above terms.

Bearing in mind that the intensity is still proportional to the amplitude squared (2.24), a comparison between (2.37) and (2.39) yields

$$\alpha(\omega) = \frac{2\omega}{c}\kappa(\omega) = \frac{4\pi}{\lambda_\mathrm{v}}\kappa(\omega). \tag{2.40}$$

The phase velocity of light in a medium is now given by (2.15)

$$v_\mathrm{ph} = \frac{\omega}{\mathrm{Re}\{k\}} = cn^{-1}(\omega). \tag{2.41}$$

For the group velocity we can get rather complicated dependencies originating from

$$v_{\mathrm{g}} = \frac{\partial \omega}{\partial k}. \tag{2.42}$$

We return to this aspect later.

2.4 Transverse, Longitudinal and Surface Waves

The only solution of (2.9) for light in vacuum is a transverse electromagnetic wave (2.19). This solution exists for light in matter as well. However (2.9) has now with the use of (2.27) the form

$$\nabla \cdot \boldsymbol{D} = \nabla \varepsilon_0 \varepsilon(\omega) \boldsymbol{E} = 0 \tag{2.43}$$

Apart from the above-mentioned transverse solution with $\boldsymbol{E} \perp \boldsymbol{k}$ there is a new solution which does not exist in vacuum ($\varepsilon_{\mathrm{vac}} \equiv 1$), namely

$$\varepsilon(\omega) = 0. \tag{2.44a}$$

This means that we can find longitudinal solutions at the frequencies at which $\varepsilon(\omega)$ vanishes. We call these frequencies correspondingly ω_{L} and note that for

$$\varepsilon(\omega_{\mathrm{L}}) = 0; \quad \boldsymbol{E} \parallel \boldsymbol{k} \text{ is possible} \tag{2.44b}$$

Now let us consider the other fields for this longitudinal wave in matter. From (2.27) we see immediately that we have for the longitudinal modes

$$\boldsymbol{D} = 0 \quad \text{and} \quad \boldsymbol{E} = -\frac{1}{\varepsilon_0} \boldsymbol{P}. \tag{2.45}$$

In matter, the Maxwell's equation $\nabla \times \boldsymbol{E} = -\dot{\boldsymbol{B}}$ is still valid. This leads for plane waves in nonmagnetic material to

$$\boldsymbol{H}_0 = (\omega \mu_0)^{-1} \boldsymbol{k} \times \boldsymbol{E}_0. \tag{2.46}$$

For the longitudinal wave it follows from (2.44) that

$$\boldsymbol{H} = 0 \quad \text{and} \quad \boldsymbol{B} = \mu_0 \boldsymbol{H} = 0 \tag{2.47}$$

The longitudinal waves which we found in matter are not electromagnetic waves but pure polarization waves with $\boldsymbol{E}$ and $\boldsymbol{P}$ opposed to each other with vanishing $\boldsymbol{D}$, $\boldsymbol{B}$ and $\boldsymbol{H}$.

Until now we were considering the properties of light in the bulk of a medium. The boundary of this medium will need some extra consideration e.g., the interface between vacuum (air) and a semiconductor. This interface is crucial for reflection of light and we examine this problem in Sects. 3.1.1–4; 5.4.2 and 5.6. Here we only want to state that the boundary conditions allow a surface mode, that is, a wave which propagates along the interface and has field amplitudes which decay exponentially on both sides. These waves are also known as surface polaritons for reasons discussed in more detail in Sect. 5.6.

2.5 Photons and Some Aspects of Quantum Mechanics and of Dispersion Relations

Maxwell's equations are the basis of the classical theory of light. They describe problems like light propagation and the diffraction at a slit or a grating e.g., in the frame of Huygen's principle or of Fourier optics [93S1].

In the interaction of light with matter, its quantum nature becomes apparent, e.g., in the photoelectric effect which shows that a light field of frequency ω can exchange energy with matter only in quanta $\hbar\omega$. Therefore, the proper description of light is in terms of quantum mechanics or of quantum electrodynamics. However, we shall not go through these theories here in detail nor do we want to address the aspects of quantum statistics of coherent and incoherent light sources, but we present in the following some of their well-known results and refer the reader to the corresponding literature [85G1, 92M1, 94A1, 94B1, 01M1, 01T1, 02B1, 02D1, 02G1, 02L1, 02Y1] for a comprehensive discussion.

The electromagnetic fields can be described by their potentials $\boldsymbol{A}$ and ϕ by

$$\boldsymbol{E} = -\mathrm{grad}\,\phi - \dot{\boldsymbol{A}}; \quad \boldsymbol{B} = \nabla \times \boldsymbol{A}\,. \tag{2.48}$$

where $\boldsymbol{A}$ is the so-called vector potential. Since $\nabla \cdot (\nabla \times \boldsymbol{A}) \equiv 0$ the notation of (2.48) fulfills automatically $\nabla \cdot \boldsymbol{B} = 0$ and reduces the six components of $\boldsymbol{E}$ and $\boldsymbol{B}$ to four.

The vector potential $\boldsymbol{A}$ is not exactly defined by (2.48). A gradient of a scalar field can be added. We can choose the so-called Coulomb gauge

$$\nabla \cdot \boldsymbol{A} = 0\,. \tag{2.49}$$

In this case ϕ is the usual electrostatic potential obeying the Poisson equation:

$$\nabla^2 \phi = -\frac{\rho}{\varepsilon_0 \varepsilon(\omega)}\,. \tag{2.50}$$

In vacuum we still have $\rho = 0$ and we assume the same for the description of the optical properties of matter.

Now we should carry out the procedure of second quantization, for simplicity again for plane waves. A detailed description of how one begins with Maxwell's equations and arrives at photons within the framework of second quantization is beyond the scope of this book see [55S1, 71F1, 73H1, 76H1, 80H1, 85G1, 92M1, 94A1, 94B1]. On the other hand we want to avoid that the creation and annihilation operators appear like a "deus ex machina". Therefore we try at least to outline the procedure.

First we have to write down the classical Hamilton function H which is the total energy of the electromagnetic field using $\boldsymbol{A}$ and ϕ. Then we must find some new, suitable quantities $p_{\boldsymbol{k},s}$ and $q_{\boldsymbol{k},s}$ which are linear in $\boldsymbol{A}$ and which fulfill the canonic equations of motion

$$\frac{\partial H}{\partial q_{\boldsymbol{k},s}} = -\dot{p}_{\boldsymbol{k},s}, \quad \frac{\partial H}{\partial p_{\boldsymbol{k},s}} = -\dot{q}_{\boldsymbol{k},s} \tag{2.51}$$

and are thus canonically conjugate variables. Here $\boldsymbol{k}$ is the wave vector of our plane electromagnetic or $\boldsymbol{A}$-wave and s the two possible transverse polarizations. The Hamilton function reads in these variables:

$$H = \frac{1}{2}\sum_{\boldsymbol{k},s}(p^2_{\boldsymbol{k},s}) + \omega_k^2 q^2_{\boldsymbol{k},s} \tag{2.52}$$

This is the usual form of the harmonic oscillator. The quantization condition

$$p_{\boldsymbol{k},s}q_{\boldsymbol{k},s} - q_{\boldsymbol{k},s}p_{\boldsymbol{k},s} = \frac{\hbar}{\mathrm{i}} \tag{2.53}$$

for all $\boldsymbol{k}$ and $s = 1, 2$ gives then the well-known result for the harmonic oscillator: The electromagnetic radiation field has for every $\boldsymbol{k}$ and polarization s energy steps

$$E_{\boldsymbol{k}} = \left(n_{\boldsymbol{k}} + \frac{1}{2}\right)\hbar\omega_{\boldsymbol{k}} \quad \text{with} \quad n_{\boldsymbol{k}} = 0, 1, 2\ldots \tag{2.54}$$

It can exchange energy with other systems only in units of $\hbar\omega$. These energy units or quanta are called photons. The term $\hbar\omega/2$ in (2.54) is the zero-point energy of every mode of the electromagnetic field.

The so-called particle-wave dualism, that is, the fact that light propagates like a wave showing, e.g., diffraction or interference and interacts with matter via particle-like quanta, can be solved by the simple picture that light is an electromagnetic wave, the amplitude of which can have only discrete values so that the energy in the waves just fulfills (2.54).

From the above introduced, or better, postulated quantities $p_{\boldsymbol{k},s}$ and $q_{\boldsymbol{k},s}$ we can derive by linear combinations operators $a^\dagger_{\boldsymbol{k},s}$ and $a_{\boldsymbol{k},s}$ with the following properties: If $a_{\boldsymbol{k},s}$ acts on a state which contains $n_{\boldsymbol{k},s}$ quanta of momentum $\boldsymbol{k}$ and polarization s it produces a new state with $n_{\boldsymbol{k},s} - 1$ quanta. Correspondingly, $a^\dagger_{\boldsymbol{k},s}$ increases $n_{\boldsymbol{k},s}$ by one. We call therefore $a_{\boldsymbol{k},s}$ and $a^\dagger_{\boldsymbol{k},s}$ annihilation and creation operators, respectively. Since the operators $a_{\boldsymbol{k},s}$ and $a^\dagger_{\boldsymbol{k},s}$ describe bosons (see below), their permutation relation is

$$a_{\boldsymbol{k},s}a^\dagger_{\boldsymbol{k},s} - a^\dagger_{\boldsymbol{k},s}a_{\boldsymbol{k},s} = 1. \tag{2.55a}$$

This holds for equal $\boldsymbol{k}$ and s. The commutator is zero otherwise.

The operator $a^\dagger_{\boldsymbol{k},s}a_{\boldsymbol{k},s}$ acting on a photon state gives the number of photons $n_{\boldsymbol{k},s}$ times the photon state and is therefore called the number operator. Summing over all possible $\boldsymbol{k}$-values and polarizations s gives finally the Hamilton operator

$$H = \sum_{\boldsymbol{k},s}\hbar\omega_{\boldsymbol{k},s}a^\dagger_{\boldsymbol{k},s}a_{\boldsymbol{k},s}. \tag{2.55b}$$

It is clear to the author that the short outline given here is not sufficient to explain the procedure to a reader who is not familiar with it. However,

since the intent is not to write a textbook on quantum electrodynamics, we want to stress here only that the electromagnetic radiation field in vacuum can be brought into a mathematical form analogous to that of the harmonic oscillator, and that quantum mechanics gives for every harmonic oscillator the energetically equidistant terms of (2.54).

The harmonic oscillator is one of the fundamental systems, which has been investigated in physics and is understood in great detail. In theoretical physics a problem can be considered as "solved" if it can be rewritten in the form of the harmonic oscillator. Apart from the electromagnetic radiation field in vacuum, we will come across some other systems which are treated in this way. For those readers who are not familiar with the concept of quantization and who wish to study the procedure in a quiet hour by themselves, we recommend the above given references.

Here are some more results: The two basic polarizations of single quanta of the electromagnetic field, – of the photons – are left and right circular σ^- and σ^+, respectively. A linearly polarized wave can be considered as a coherent superposition of a left and right circularly polarized one with equal frequencies and wave vector $\boldsymbol{k}$. The term coherent means that two light beams have a fixed-phase relation relative to each other. The component of the angular momentum s in the direction of the quantization axis which is parallel to $\boldsymbol{k}$ is for photons thus

$$s_{\parallel} = \pm\hbar. \tag{2.56}$$

This means that photons have integer spin and are bosons. The third possibility $s_{\parallel} = 0$ expected for spin one particles is forbidden, because longitudinal electromagnetic waves do not exist at least in vacuum.

Photons in thermodynamic equilibrium are described by Bose-statistics. The occupation probability f_{BE} of a state with frequency ω is given by

$$f_{\mathrm{BE}} = \left[\exp(\hbar\omega/k_{\mathrm{B}}T) - 1\right]^{-1} , \tag{2.57}$$

where T is the absolute temperature and k_{B} is Boltzmann's constant.

The chemical potential μ which could appear in (2.57) is zero in thermal equilibrium, since the number of photons is not conserved.

An approach to describe non-thermal photon fields e.g. luminescence by a non-vanishing μ and Kirchhoff's law in the sense of a generalized Planck's law is found in [82W1, 92S2, 95D1].

The momentum $\boldsymbol{p}$ of a photon with wave vector $\boldsymbol{k}$ is given, as for all quanta of harmonic waves, by

$$\boldsymbol{p} = \hbar\boldsymbol{k}. \tag{2.58}$$

where $\boldsymbol{k}$ is the real part of the wave vector, which describes as already mentioned the oscillatory, propagating aspect of the plane wave.

To summarize, we can state that photons are bosons with spin $\pm\hbar$, energy $\hbar\omega$ and momentum $\hbar\boldsymbol{k}$ which propagate according to the wave equations.

A very important property of particles in quantum mechanics is their dispersion relation. By this we mean the dependence of energy E or frequency ω

on the wave vector $\boldsymbol{k}$ i.e., the $E(\boldsymbol{k})$ or $\omega(\boldsymbol{k})$ relation. For photons in vacuum we find the classical relation given already in (2.13)

$$E = \hbar\omega = \hbar ck. \tag{2.59}$$

The dispersion relation for photons in vacuum is thus a linear function with slope $\hbar c$ as shown in Fig. 2.2. Correspondingly we find again both for phase and group velocity with (2.15)

$$v_{\mathrm{ph}} = v_{\mathrm{g}} = c\,. \tag{2.60}$$

We conclude this subsection with an explanation of energy units. In the SI system the energy unit is $1\,\mathrm{N\,m} = 1\,\mathrm{kgm^2/s^2}$ with the following identity relations

$$1\,\mathrm{N\,m} = 1\,\mathrm{m\,kg\,s^{-2}} = 1\,\mathrm{V\,A\,s} = 1\,\mathrm{W\,s} = 1\,\mathrm{J}\,. \tag{2.61a}$$

Since the energies of the quanta in optical spectroscopy are much smaller, we frequently use the unit 1 eV. This is the energy that an electron gains if it passes, in vacuum, through a potential difference of one volt, resulting in

$$1\,\mathrm{eV} = 1.60217733 \times 10^{-19}\,\mathrm{J} \approx 1.6 \times 10^{-19}\,\mathrm{J}\,. \tag{2.61b}$$

In spectroscopy another measure of energy is frequently used the wave number. The definition is as follows. One expresses the energy of a (quasi-) particle by the number of wavelengths per cm of a photon with the same energy. So

$$1\,\mathrm{eV} \mathrel{\widehat{=}} 8065.4\,\mathrm{cm^{-1}} \text{ or } 10^4\,\mathrm{cm^{-1}} \mathrel{\widehat{=}} 1.23986\,\mathrm{eV} \tag{2.61c}$$

Another quantity that is sometimes confused with the wave number, which gives the energy and is therefore a scalar quantity, is the wave vector, since it has also the dimension 1/length.

The amount of the (real part of the) wave vector is given by $k = 2\pi/\lambda$, where λ is the wavelength of the corresponding quantum or particle (electron,

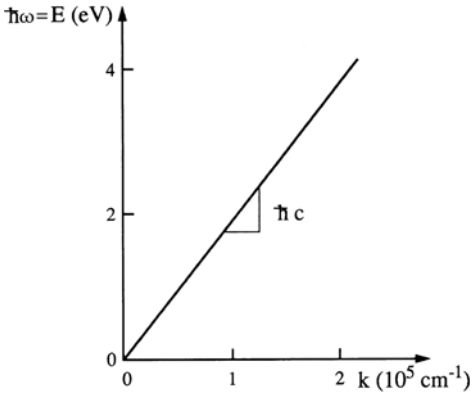


Fig. 2.2. The dispersion relation of photons in vacuum

phonon, photon, etc.). The direction of $\boldsymbol{k}$ is the direction of propagation, i.e., $\boldsymbol{k}$ is normal to the wave-front, in the case of light in vacuum or in matter normal to $\boldsymbol{D} \times \boldsymbol{B}$. The quantity $\boldsymbol{k}$ is very closely related to the (quasi-)momentum of the (quasi-)particle $\boldsymbol{p}$ through

$$\boldsymbol{p} = \hbar \boldsymbol{k} \,. \tag{2.58}$$

For the discussion of the concept of quasi-momentum see, e.g., Sects. 5.2, 3 and 5 or [98B2]. The dispersion relation of (quasi-) particles is thus given by $E(\boldsymbol{k})$. The wave vector of light is in the visible in vacuum, falling in the range of a few times $10^4\,\mathrm{cm}^{-1}$ while the border of the first Brillouin zone (see Sect. 7.2) is of the order of $10^8\,\mathrm{cm}^{-1}$.

It is obvious, that a quantity like a wave-number vector (Wellenzahlvektor) is ill-defined and does not exist!

2.6 Density of States and Occupation Probabilities

A quantity which is crucial in quantum mechanics for the properties of particles is their density of states. It enters, e.g., in Fermi's golden rule which allows one to calculate transition probabilities. We want to discuss this problem in a general way for systems of different dimensionalities $d = 3, 2$ and 1. We shall need these results later on for low-dimensional semiconductor structures. The discussion of the density of states, especially in various dimensions, is not so commonly treated as the harmonic oscillator, and so we shall spend some time on this problem and dwell more on the details. At the end of this section we shall also state the occupation probability in thermodynamic equilibrium for classical particles, for fermions and bosons.

If we consider a particle which is described by a wave function[3] $\phi(\boldsymbol{r})$ then the probability w to find it in a small element of space $\mathrm{d}\tau = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ around $\boldsymbol{r}$ is

$$w(\boldsymbol{r})\tau = \phi^*(\boldsymbol{r})\phi(\boldsymbol{r})\,\mathrm{d}\tau \tag{2.62a}$$

Since the particle has to be somewhere in the system, $w(\boldsymbol{r})$ has to be "normalized", that is,

$$\int_{\text{system}} w(\boldsymbol{r})\,\mathrm{d}\tau = \int_{\text{system}} \phi^*(\boldsymbol{r})\phi(\boldsymbol{r})\,\mathrm{d}\tau = 1. \tag{2.62b}$$

Here, the functions $\phi(\boldsymbol{r})$ are of the form $\exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})$. For normalization purposes a factor has to be added

$$\phi(\boldsymbol{r}) = \Omega^{-1/2} \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r}). \tag{2.63}$$

[3] The letter ϕ has been already used for the electrostatic potential e.g., in (2.50). Since there are more different physical quantities than letters of the alphabet, we sometimes use the same letter for different quantities, but from the context it should be clear what is meant.

The normalization condition (2.62b) results in

$$\Omega^{-1}\int_{\text{system}} \exp(-\mathrm{i}\boldsymbol{k}\boldsymbol{r})\exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})\,\mathrm{d}\tau = \Omega^{-1}\int_{\text{system}} \mathrm{d}\tau = \Omega^{-1}V_{\text{system}} = 1, \tag{2.64}$$

where V_{system} is the volume of our physical system. Consequently Ω is just the volume of the system. To avoid a factor of zero in front of the plane-wave term, one assumes that the system is so big that it contains all physically relevant parts, but that it is not infinite. The simplest choice is a box of length L, or, more precisely speaking, a cube in three dimensions, a square in two, and an interval in one. This procedure is known as "normalization in a box". Consequently we have

$$V_{\text{system}} = L^d \quad \text{with} \quad d = \text{dimensionality of the system} \tag{2.65}$$

and

$$\Omega^{1/2} = L^{d/2} \quad \text{for} \quad d = 3, 2, 1. \tag{2.66}$$

The wave vectors which can exist in such a box are limited by the boundary conditions.

If we assume that we have an infinitely high potential barrier around the box, then the wavefunction must have nodes at the walls (Fig. 2.3a). Consequently the components k_i of $\boldsymbol{k}$ must fulfill

$$k_i = n_i\frac{\pi}{L}; \quad n_i = 1, 2, 3, \ldots; \quad i = 1, \ldots, d, \tag{2.67}$$

where the index i runs over all dimensions.

Such a wave is a standing wave, i.e., a coherent superposition of two waves with $\boldsymbol{k}$ and $-\boldsymbol{k}$ and equal amplitudes. In the following we must consider therefore only positive values of $\boldsymbol{k}$. The various modes are distributed equally spaced over the k_i-axes with a spacing Δk_i,

$$\Delta k_i = \frac{\pi}{L}\,. \tag{2.68}$$

In other words, every state (or mode) needs a volume V_k in $\boldsymbol{k}$-space given by

$$V_k = \left(\frac{\pi}{L}\right)^d. \tag{2.69}$$

Another approach is to impose periodic boundary conditions. Then the plane wave should have equal amplitude and slope on opposite sides of the cube according to Fig. 2.3b. In this case one can fill the infinite space by adding boxes in all dimensions and one finds:

$$k_i = n_i'\frac{2\pi}{L}; \quad n_i' = \pm1, \pm2, \pm3\ldots \tag{2.70}$$

This means for Δk_i

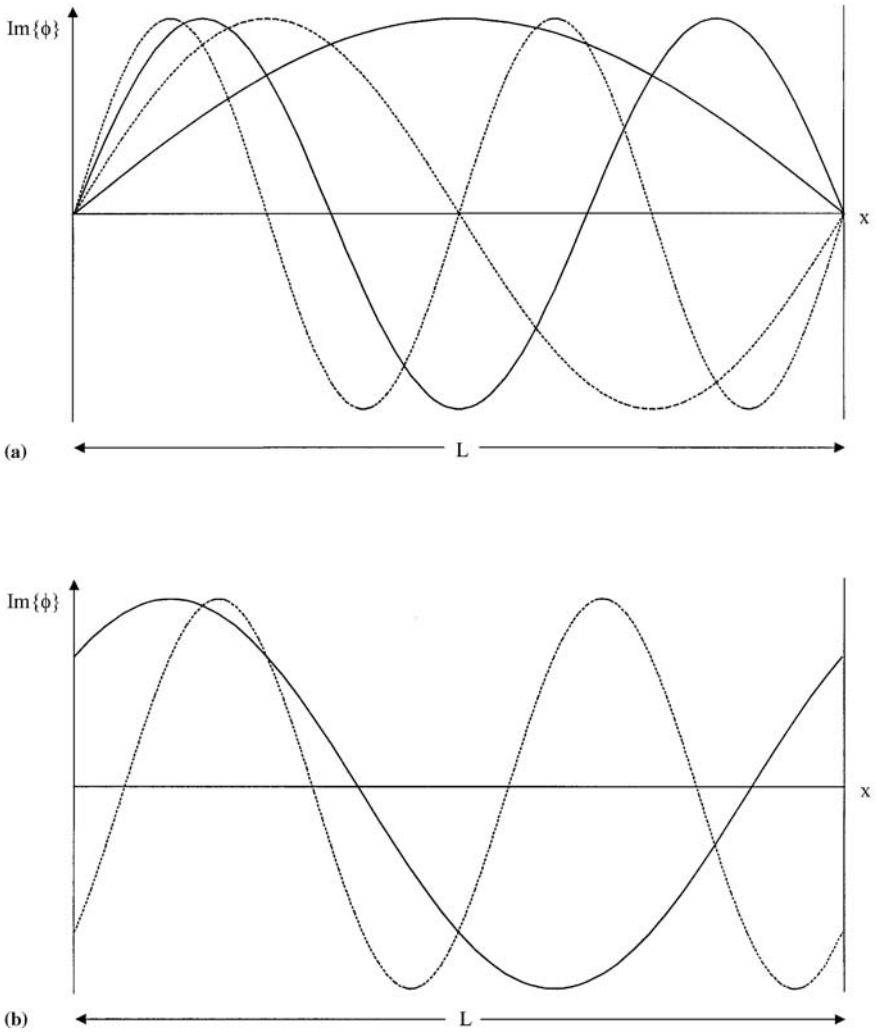


Fig. 2.3. Plane waves which have nodes at the boundaries (**a**) or which obey periodic boundary conditions (**b**)

$$\Delta k_i' = \frac{2\pi}{L}\,. \tag{2.71}$$

In contrast to the case of standing waves, we now have to consider positive and negative values of n_i separately. This procedure results finally in the same density of states.

As a consequence we find that plane waves have in Cartesian coordinates in $\boldsymbol{k}$-space a constant density on all axes.

This result can also be derived qualitatively from the uncertainty principle. If a particle is confined to a length L in direction i its momentum has an uncertainly $\Delta p_i = \Delta\hbar k_i \geq \hbar/L$. Consequently two states, which should be distinguishable or "different", must have k_i values, which differ by roughly $1/L$.

Often one wants to know the number of states in a shell between k and $k + \mathrm{d}k$ independent of the direction of $\boldsymbol{k}$. This question can be answered by introducing polar coordinates in $\boldsymbol{k}$-space. The differential volume $\mathrm{d}V_k$ of a shell of thickness $\mathrm{d}k$ in a d-dimensional $\boldsymbol{k}$-space is given by

$$\begin{aligned} \mathrm{d}V_k &= 2\,\mathrm{d}k \qquad && \text{for } d = 1, \\ \mathrm{d}V_k &= 2\pi k\,\mathrm{d}k \qquad && \text{for } d = 2, \\ \mathrm{d}V_k &= 4\pi k^2\,\mathrm{d}k \qquad && \text{for } d = 3. \end{aligned} \tag{2.72}$$

Depending on the boundary condition we have to take into account only positive (2.67), or positive and negative (2.70), values of $\boldsymbol{k}$ or n_i with corresponding modifications of the prefactors in (2.72).

The number $\hat{D}(\boldsymbol{k})$ of states in $\boldsymbol{k}$-space found between k and $k + \mathrm{d}k$ in polar coordinates is given by dividing $\mathrm{d}V_k$ by the volume for each state and by multiplying by g_s. The quantity g_s considers degeneracies such as the spin degeneracy. For photons we have $g_\mathrm{s} = 2$ according to the σ^+ and σ^- polarizations (see above). The results are

$$\begin{aligned} \hat{D}(k)\,\mathrm{d}k &= g_\mathrm{s}\frac{L}{\pi}\,\mathrm{d}k \qquad && \text{for } d = 1, \\ \hat{D}(k)\,\mathrm{d}k &= g_\mathrm{s}\frac{L^2}{2\pi}k\,\mathrm{d}k \qquad && \text{for } d = 2, \\ \hat{D}(k)\,\mathrm{d}k &= g_\mathrm{s}\frac{L^3}{2\pi^2}k^2\,\mathrm{d}k \qquad && \text{for } d = 3. \end{aligned} \tag{2.73}$$

The derivation of this result is depicted for $d = 2$ in Fig. 2.4. If we neglect constant prefactors and divide by $\mathrm{d}k$ we find

$$\hat{D}(k) \propto g_\mathrm{s} L^d k^{d-1}, \quad d = 1, 2, 3 \ldots . \tag{2.74}$$

If we consider not the number of states in the box of volume L^d but the density of states $D(k)$ per unit of space (e.g., per cm^3 or m^3) the term L^d in (2.74) disappears yielding

$$D(k) \propto g_\mathrm{s} k^{d-1} \tag{2.75}$$

The concept of periodic boundary conditions yields the same result.

This result has to be expected since the density of states per unit volume must be independent of the size of the box which we have in mind provided the box is sufficiently large.

We want to stress here that we assumed only plane waves but did not make any specific assumptions about which type of particles are represented

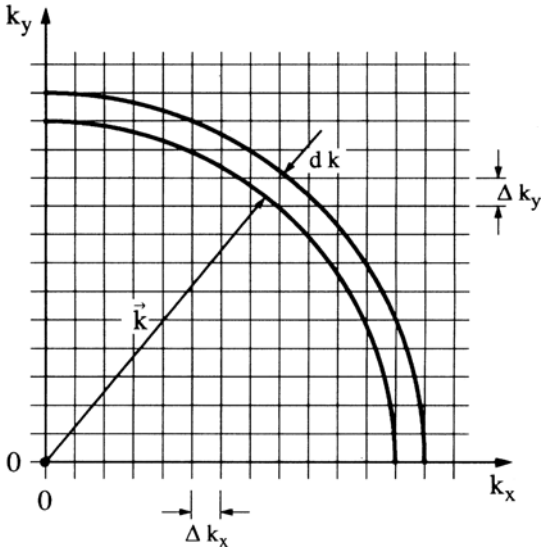


Fig. 2.4. A two-dimensional $\boldsymbol{k}$-space in which the states are equally spaced on the k_x-and k_y-axes to derive (2.73) for $d = 2$

by these plane waves – photons, electrons etc. Therefore this result is valid for all particles described by plane waves.

The next step is now to calculate the density of states on the energy axis, i.e.,

$$D(E)\,\mathrm{d}E. \tag{2.76}$$

This quantity gives the number of states in the energy interval from E to $E + \mathrm{d}E$. To calculate this quantity we need the specific dispersion relation $E(\boldsymbol{k})$ and its inverse $\boldsymbol{k}(E)$ as seen from the identity (2.77):

$$D(E)\,\mathrm{d}E = D[k(E)]\frac{\mathrm{d}k}{\mathrm{d}E}\cdot\,\mathrm{d}E = D[k(E)]\frac{1}{|\mathrm{grad}_k E(\boldsymbol{k})|}\,\mathrm{d}E. \tag{2.77}$$

The term on the right-hand side of (2.77) gives the generalized equation which is also valid for anistropic cases.

In particular for photons in vacuum we have with (2.59)

$$k = \frac{E}{\hbar c} = \frac{\omega}{c};\quad \frac{\mathrm{d}k}{\mathrm{d}E} = \frac{1}{\hbar c} \tag{2.78}$$

Inserting this result in (2.77), for the case $d = 3$ we find

$$D(E)\,\mathrm{d}E = \frac{E^2}{\pi^2(\hbar c)^3}\,\mathrm{d}E$$

or

$$D(\omega)\,\mathrm{d}\omega \propto \omega^2\,\mathrm{d}\omega \tag{2.79a}$$

For massive particles, i.e. particles the dispersion relation of which can be described by an (effective) mass m according to

$$E(\boldsymbol{k}) = \frac{\hbar^2 \boldsymbol{k}^2}{2m} \tag{2.79b}$$

we obtain with (2.77) for a d-dimensional space

$$D(E)\,\mathrm{d}E \propto g_\mathrm{s} E^{\frac{d}{2}-1}\,\mathrm{d}E\,. \tag{2.79c}$$

This formula includes the well-known square root density of states for massive particles in three dimensions.

We repeat again that the density in $\boldsymbol{k}$-space is constant on all Cartesian axes in a d-dimensional space for all particles, which can be described by a plane wave, but the density of states depends on the individual dispersion relation when plotted as a function of energy.

The next quantity, which we need is the occupation probability of the states discussed above. We restrict ourselves in the following to thermodynamic equilibrium. There are three types of statistics which can be considered:

For classical, distinguishable particles, Boltzmann statistics apply:

$$f_\mathrm{B} = \exp[-(E-\mu)/k_\mathrm{B}T]. \tag{2.80a}$$

For bosons, i.e., indistinguishable particles with integer spin, photons being an example, one must use the Bose–Einstein statistics

$$f_\mathrm{BE} = \{\exp[(E-\mu)/k_\mathrm{B}T] - 1\}^{-1}. \tag{2.80b}$$

Fermions, or indistinguishable particles with half-integer spin e.g., electrons obey the Fermi–Dirac statistics f_FD

$$f_\mathrm{FD} = \{\exp[-(E-\mu)/k_\mathrm{B}T] + 1\}^{-1}. \tag{2.80c}$$

The Boltzmann constant is k_B and the chemical potential is μ which gives the average energy necessary to add one more particle to the system. For fermions μ is also known as the Fermi energy E_F. The probability to find a particle in the interval from E to $E + \mathrm{d}E$ is then given by the product of the density of states $D(E)$ and the occupation probability f

$$D(E) f(E, T, \mu)\,\mathrm{d}E \tag{2.81}$$

In Fig. 2.5 we plot f_B, f_BE and f_FD as a function of $(E-\mu)/k_\mathrm{B}T$.

The Boltzmann statistics shows the well-known exponential dependence. The Fermi–Dirac statistics never exceeds one, realizing thus Pauli's exclusion principle. The Bose–Einstein statistics has a singularity for $E = \mu$. This gives rise to Bose–Einstein condensation, or in other words, a macroscopic population of a single state, if μ touches a region with a finite density of states. In this case the species with energies $E = \mu$ and those with $E > \mu$ must be

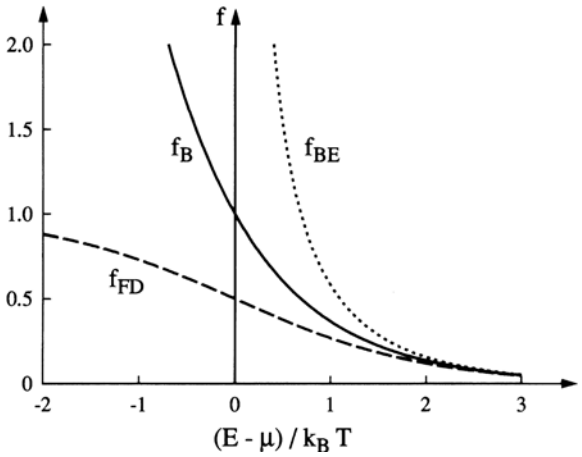


Fig. 2.5. The Boltzmann, Bose–Einstein and Fermi–Dirac distribution functions versus $(E-\mu)/k_{\mathrm{B}}T$

considered separately. Furthermore it is obvious from Fig. 2.5 that f_{BE} and f_{FD} converge to f_{B} for $(E-\mu)/kT > 1$.

In Sects. 7.7, 8.2 and 9 or 20.5 we discuss, when classical Boltzmann statistics is a good approximation and when the use of Fermi–Dirac or Bose–Einstein statistics is obligatory.

The chemical potential μ is zero in thermodynamic equilibrium for quanta whose number is not conserved, for e.g., photons or phonons. We introduce this topic in Chap. 7.

If the number N or density n of particles in a system is known, as is the case for electrons at non-relativistic energies, then μ is well defined by (2.82).

$$\int D(E) f(E, \mu, T)\, \mathrm{d}E = n, \tag{2.82}$$

which says that the density of particles is equal to the integral over the product of the density of states and the probability that a state is occupied.

As an example, we apply now the above statements to photons in a three-dimensional box in thermodynamic equilibrium. With (2.79b) and (2.80b), (2.83) is obtained.

$$N(\omega)\,\mathrm{d}\omega = D(\omega) f_{\mathrm{BE}}(\omega, T)\,\mathrm{d}\omega \propto \omega^2 \left[\exp\left(\frac{\hbar\omega}{k_{\mathrm{B}}T}\right) - 1\right]^{-1} \mathrm{d}\omega\,. \tag{2.83}$$

For the energy content $U(T)$ of the radiation field, the result is

$$U(T) = \hbar \int_0^\infty \omega N(\omega)\,\mathrm{d}\omega \propto \int_0^\infty \omega^3 \left[\exp\left(\frac{\hbar\omega}{k_{\mathrm{B}}T}\right) - 1\right]^{-1} \mathrm{d}\omega\,. \tag{2.84}$$

Equation (2.83) is nothing other than Planck's law of black-body radiation. By substituting x for the variable $\hbar\omega/k_{\mathrm{B}}T$ in (2.84) we immediately find the

Stefan–Boltzmann T^4 law

$$U(T) \propto T^4 \int_0^\infty x^3 (\exp x - 1)^{-1} \, \mathrm{d}x = T^4 A \,, \tag{2.85}$$

where A is a constant.

2.7 Problems

1. The intensity of the sunlight falling on the earth is, for normal incidence and before its passage through the atmosphere, about $1.5\,\mathrm{kW\,m^{-2}}$. Calculate the electric-field strength.
2. Pulsed high power lasers can be easily focussed to a power density I of $10\,\mathrm{GW/cm^2}$. Calculate the E and B fields. Compare them with the electric field in an H atom at a distance of one Bohr radius, and the magnetic field on the surface of the earth, respectively.
3. Calculate the number of photon modes in the visible part of the spectrum ($\approx 400\,\mathrm{nm} \lesssim \lambda_\mathrm{v} \lesssim 800\,\mathrm{nm}$) in a box of $1\,\mathrm{cm^3}$.
4. Calculate the momentum and energy of a photon with $\lambda_\mathrm{v} = 500\,\mathrm{nm}$. At which acceleration voltage has an electron the same momentum?
5. Show qualitatively the $\boldsymbol{B}$, $\boldsymbol{H}$ and $\boldsymbol{M}$ fields of a homogeneously magnetized, brick-shaped piece of iron and for a hollow sphere with inner radius R_0 and outer radius $R_0 + \Delta R$, which is radially magnetized. Use especially for the second case symmetry considerations together with (2.1).
6. Check whether the maximum of $N(\omega)$ in (2.83) shifts in proportion to T (Wien's law), originally formulated as $\lambda_\mathrm{max} \propto T^{-1}$.
7. Compare the contribution of the electric conductivity of a typical semiconductor to that of the polarisation in (2.23) or (2.25). For which frequencies does the second one dominate?
8. Write down the time and space dependence of a spherical wave. Note that the energy flux density varies usually like the amplitude squared. Is it possible to create a spherical vector wave?
9. Inspect (with the help of a textbook or a computer program) the electric field of a static electric dipole and the near and far fields of an oscillating electric dipole. Note that in the near field the electric and magnetic fields are not orthogonal.
10. Show that the definition $\boldsymbol{v}_g = \frac{1}{\hbar} \mathrm{grad_k} E(\boldsymbol{k})$ leads, for massive and massless particles, directly to the relation (quasi-)momentum $\boldsymbol{p} = \hbar \boldsymbol{k}$.
11. How does the density of states as a function of energy vary for a linear dispersion relation (like photons) in 3, 2 and 1 dimensional systems?
12. Consider or find in a textbook the pattern of the collective motion of the H_2O molecules in a surface water wave. Which effects contribute to the restoring force? Are water waves harmonic waves? What happens at a seashore, where the depth of the water decreases gradually? Is there a net transport of matter? Assume that the particles have an electric

charge and move relative to a fixed background of opposite charge. Which charge pattern do you expect close to the surface? Compare with Fig. 4.4b.

References to Chap. 2

We give here and in the following chapters all (co-) authors of a publication to a maximum of three. For papers with four and more (co-)authors we give only the first followed by et al.

[55S1] L.I. Schiff: Quantum Mechanics, 2nd edn. Adv. Solid State Phys. McGraw-Hill, New York (1955)
[71F1] A.L. Fetter, J.D. Walecka: Quantum Theory of Many Particle Systems McGraw Hill, New York (1971)
[73H1] H. Haken: Quantenfeldtheorie des Festkörpers Teubner, Stuttgart (1973)
[76H1] H. Haken: Quantum Field Theory of Solids North Holland, Amsterdam (1976)
[80H1] H. Haken, H.C. Wolf: Atom- und Quantenphysik Springer, Berlin, Heidelberg (1980)
[82W1] P. Würfel: J. Phys. C **15**, 3967 (1982)
[84A1] M. Abramowitz, J.A. Stegun (eds): Pocket Book of Mathematical Functions Deutsch, Thun (1984)
[85G1] W. Greiner: Theoretische Physik, Vols **1**–**10** Deutsch, Frankfurt (1985)
[88D1] Diluted Magnetic Semiconductors, J.K. Furdyna, J. Kossut (eds), Semicond. Semimet. **25** (1988)
[89S1] Squeezed and Nonclassical Light, P. Tombesi ed., NATO ASI Series B **190**, Plenum Press, New York (1989)
[91B1] I.N. Bronstein, K.A. Semendjajew: Taschenbuch der Mathematik, 25th edn. Teubner, Stuttgart (1991)
[91O1] M. von Ortenberg: Festkörperprobleme/Adv. Solid State Phys. **31**, 261 (1991)
[92G1] O. Goede, W. Heimbrodt: Adv. Solid State Phys. **32**, 237 (1992)
[92M1] P. Meystre, M. Sargent III: Elements of Quantum Optics, 2nd ed. Springer, Berlin, Heidelberg (1992)
[92S1] H. Stöcker (ed.): Taschenbuch Mathematischer Formeln und Moderner Verfahren (Deutsch, Frankfurt (1992)
[92S2] K. Schick et al. Appl. Phys. A **54**, 109 (1992)
[92Y1] D.R. Yakovlev: Festkörperprobleme/Adv. Solid State Phys. **32**, 251 (1992)
[93S1] W. Stößel: Fourieroptik, Springer, Berlin (1993)
[94A1] I. Abraham and J.A. Levenson in [81A1]g of Chap. 1 p 251
[94B1] B. Bowlby: in [81A1]g of Chap. 1 p 385
[94D1] T. Dietl: Diluted Magnetic Semiconductors, Handbook of Semiconductors **38**, North Holland, New York (1994)
[94G1] J.A. Gaj et al.: Phys. Rev. B. **50**, 5512 (1994)
[95D1] E. Daub and P. Würfel: Phys. Rev. Lett. **74**, 1020 (1995)
[96H1] R. Hellmann et al.: J. Crystal Growth **159**, 976 (1996)
[96L1] Leitfaden für den Gebrauch des Internationalen Einheitensystems, Physikalisch Technische Bundesanstalt (PTB), Braunschweig (1996) or various textbooks on physics

[98B1] R. v.Baltz in [81A1]i of Chap. 1 p 303
[98B2] R. v.Baltz and C. Klingshirn in [81A1]i of Chap. 1p 381
[98D1] B. Di Bartolo in [81A1]i of Chap. 1, p 1
[98O1] H. Ohno: Science **281**, 951 (1998)
[01M1] G. Messin et al.: Optics Lett. **26**, 1891 (2001)
[01T1] Tao-Hua et al.: Acta Optica Sinica **21**, 1486 (2001)
[01T2] W. Tittel and W. Martienssen: Physikal. Blätter **57**, Heft 7/8, p. 81 (2001)
[02B1] G. Bayer et al.: Physica E **12**, 900 (2002)
[02D1] G.A. Durkin, C. Simon, D. Bouwmeester: Phys. Rev. Lett. **88**, 187902 (2002)
[02G1] O. Gywat, G. Burkard, D. Loss: Phys. Rev. B **65**, 205329 (2002)
[02L1] A.S. Lenihan et al.: Phys. Rev. Lett. **88**, 223601 (2002)
[02Y1] Ye-Liu, Guo-Guangcan: Acta Optica Sinica **22**, 407 (2002)
[03D1] T. Dietl, Europhys. News **34**, 216 (2003)

3

Interaction of Light with Matter

In the next two sections we present some basic interaction processes of light with matter from two different points of view. First we consider matter as a homogeneous medium described by the complex dielectric function $\varepsilon(\omega)$ or by the complex index of refraction $\tilde{n}(\omega)$ (Sect. 3.1). We concentrate especially on the reflection and transmission of light at the plane interface between two media. As an especially simple case we investigate the boundary of matter and vacuum. In the later Sect. 3.2 we will discuss the interaction of the radiation field with individual atoms. In this case quantum mechanics must be used. We will employ what we have explored in Chap. 2.

3.1 Macroscopic Aspects for Solids

3.1.1 Boundary Conditions

Let us start with the macroscopic description of the optical properties of semiconductors. In Fig. 3.1 we show the wave vectors and field amplitudes in the vicinity of the interface between two media for two linear polarizations. In Fig. 3.1a the electric field $\boldsymbol{E}_\mathrm{i}$ of the incident beam is polarized parallel to the plane of incidence, which is defined by the wave vector of the incident light $\boldsymbol{k}_\mathrm{i}$ and the normal to the plane interface $\boldsymbol{e}_\mathrm{n}$. As we will see later from the boundary conditions, the wave vectors and the electric fields of transmitted and reflected beams (indices tr and r, respectively) are in the same plane; the magnetic fields according to (2.19) perpendicular to it. In Fig. 3.1b we have just the opposite situation for $\boldsymbol{E}$ and $\boldsymbol{H}$.

One often assumes for simplicity that the medium I is vacuum (or air), i.e., $\varepsilon_\mathrm{I}(\omega) = \tilde{n}_\mathrm{I}(\omega) \equiv 1$. We do not use this approximation here but we still assume that media I and II are isotropic. This means $\varepsilon_\mathrm{I}(\omega)$ and $\varepsilon_\mathrm{II}(\omega)$ are scalar functions and $\tilde{n}(\omega)$ does not depend on orientation. Phenomena which appear if we drop this assumption are dealt with later. Furthermore, we assume that there is only one reflected and one transmitted beam. This assumption seems

trivial and it is indeed for reflection. In transmission there may be more than one propagating beam, as we shall discuss in Chap. 5. What we first want to know are the dependences of the angles α_r and α_tr on α_i i.e., the laws of reflection and refraction. We then want to know the coefficients of reflection r and transmission t of the interface between media I and II. We can define these coefficients for the field amplitudes E_0.

$$r_\parallel = \frac{E_{0\mathrm{r}}}{E_{0\mathrm{i}}}; \quad t_\parallel = \frac{E_{0\mathrm{tr}}}{E_{0\mathrm{i}}} \tag{3.1a}$$

for the configuration of Fig. 3.1a or

$$r_\perp = \frac{E_{0\mathrm{r}}}{E_{0\mathrm{i}}}; \quad t_\perp = \frac{E_{0\mathrm{tr}}}{E_{0\mathrm{i}}} \tag{3.1b}$$

for the configuration of Fig. 3.1b. However, what is usually measured is the reflectivity R and transmittivity T of an interface for the intensities. We have

$$R_{\perp,\parallel} = \mid r_{\perp,\parallel} \mid^2 \tag{3.2}$$

because incident and reflected beams propagate in the same medium. For the transmission see Fig. 3.7. In order to calculate all the quantities given above we need the corresponding number of equations. They can be deduced from Maxwell's equations as boundary conditions which must be fulfilled at the interface. To do so, we need two general laws of vector analysis which are known as the laws of Gauß and of Stokes, respectively. They read for a given vector field A.

$$\int_\mathrm{volume} \nabla \cdot \boldsymbol{A}\,\mathrm{d}\tau = \oint_\mathrm{surface} \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{f} \tag{3.3a}$$

and

$$\int_\mathrm{surface} (\nabla \times \boldsymbol{A}) \cdot \mathrm{d}\boldsymbol{f} = \oint_\mathrm{line} \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{s} \tag{3.3b}$$

Starting with (2.1)a, with the help of (3.3a), we obtain

$$\int_\mathrm{v} \mathrm{div}\boldsymbol{D}\,\mathrm{d}\tau = \oint \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{f} = \int_\mathrm{v} \rho\,\mathrm{d}\tau. \tag{3.4}$$

We choose the integration volume in the form of a tiny (differentially small) cylinder which contains the interface and has its top and bottom in the media I and II, respectively (Fig. 3.2). Furthermore it is assumed that the ratio of the height to the radius of this cylinder is also infinitesimally small, so that the contribution to the whole integral from the lateral surface of the cylinder is negligible. Then the middle and right-hand-side terms of (3.4) yield

$$(\boldsymbol{D}_\mathrm{I} - \boldsymbol{D}_\mathrm{II}) \cdot \mathrm{d}\boldsymbol{f} = (D_\mathrm{n,I} - D_\mathrm{n,II})\,\mathrm{d}f = \rho_\mathrm{s}\,\mathrm{d}f, \tag{3.5}$$

where the index n means the normal component and ρ_s a surface charge density. The contribution of a volume charge density ρ_v goes to zero with

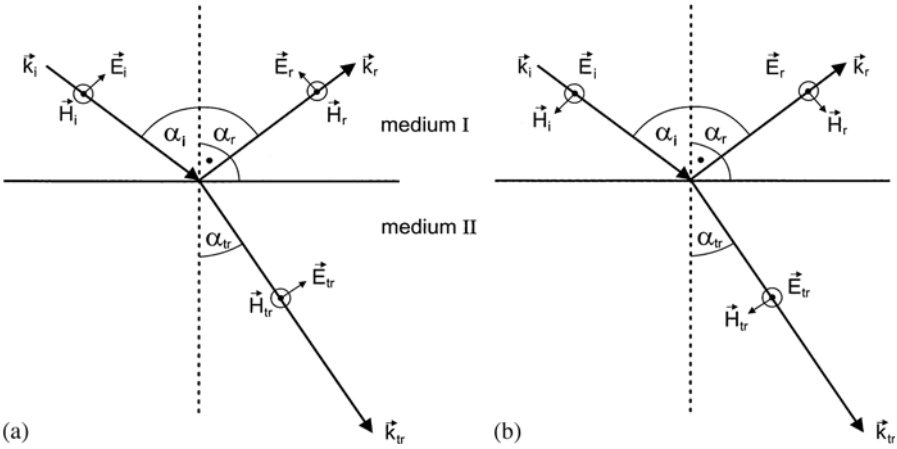


Fig. 3.1. The $\boldsymbol{E}$ and $\boldsymbol{H}$ fields and the wave vectors $\boldsymbol{k}$ for incident, transmitted and reflected beams at an interface between two isotropic media for two different, orthogonal, linear polarizations (**a**) and (**b**) respectively

decreasing height of the cylinder as compared to a possible contibution of a two-dimensional surface charge density ρ_s.

Since we assumed that there are no free charges ρ at all (Sect. 2.3) and consequently no surface charges, the right-hand side of (3.5) vanishes and we find as a boundary condition that the normal component of $\boldsymbol{D}$ is continuous across the interface:

$$D_\mathrm{n}^\mathrm{I} = D_\mathrm{n}^\mathrm{II}. \tag{3.6a}$$

Starting from (2-1)b we find in the same way

$$B_\mathrm{n}^\mathrm{I} = B_\mathrm{n}^\mathrm{II}. \tag{3.6b}$$

Using (2.1c,d) and (3.3b) we get in an analogous way requirements for the tangential components of $\boldsymbol{E}$ and $\boldsymbol{H}$.

$$E_\mathrm{t}^\mathrm{I} = E_\mathrm{t}^\mathrm{II}. \tag{3.7a}$$

$$H_\mathrm{t}^\mathrm{I} = H_\mathrm{t}^\mathrm{II}. \tag{3.7b}$$

Equations (3.6) and (3.7) represent the boundary conditions for electric and magnetic fields. Actually only two of them are independent, the other ones follow directly with the linear approaches (2.27). This is enough to calculate for a given incident beam the properties of the reflected and refracted ones. In order to do so the boundary conditions must be applied to a specific problem. For the configuration of Fig. 3.1 this reads for the incident, reflected and transmitted electric fields

$$\boldsymbol{e}_\mathrm{n} \times \boldsymbol{E}_\mathrm{i} - \boldsymbol{e}_\mathrm{n} \times \boldsymbol{E}_\mathrm{r} = \boldsymbol{e}_\mathrm{n} \times \boldsymbol{E}_\mathrm{tr}, \tag{3.8}$$

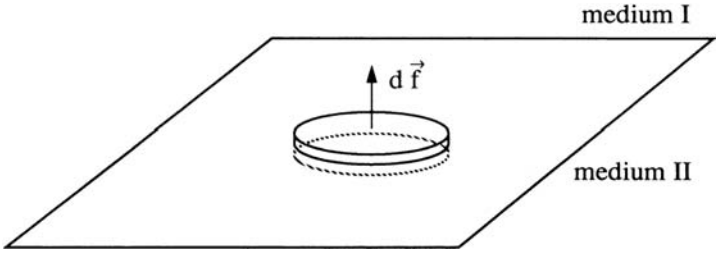


Fig. 3.2. Schematic drawing of the infinitesimally small cylinder used to deduce with (3.4) the boundary condition (3.6)

since the vector products of the various $\boldsymbol{E}_\mathrm{s}$ with the unit vector $\boldsymbol{e}_\mathrm{n}$ normal to the interface generate just the tangential components. Using another one of the set of equations (3.7),(3.8) allows one to calculate the properties of the reflected and transmitted beams. This procedure involves some basically simple but lengthy algebra and does not give further insight into the physics. In accordance with the concept of this book, we consequently skip these calculations which can be found in the literature, (See e.g. [76P1, 77B1, 87H1, 97L1] and references therein) but present the results giving some cross-links to other physical approaches to obtain them.

3.1.2 Laws of Reflection and Refraction

The first, not too surprising, result from the above-mentioned procedure is

$$\omega_\mathrm{i} = \omega_\mathrm{r} = \omega_\mathrm{tr}. \tag{3.9}$$

This means all three beams have the same frequency. This becomes clear from classical physics, as we shall see in Chap. 4 since atoms perform forced oscillations with frequency ω_i under the influence of the incident field and can therefore radiate, according to the linear approach (2.27), only at this frequency. The relation (3.9) is also intelligible from the point of view of quantum mechanics, bearing in mind the law of energy conservation and the fact that a single photon has energy $\hbar\omega$ and can be either reflected or refracted.

The next results are the laws of reflection and Snell's law of refraction. The first one states

$$\alpha_\mathrm{i} = \alpha_\mathrm{r} \tag{3.10a}$$

and;

$$\boldsymbol{k}_\mathrm{i}, \boldsymbol{k}_\mathrm{r} \text{ and } \boldsymbol{e}_\mathrm{n} \text{ are in one plane}, \tag{3.10b}$$

namely in the above-introduced plane of incidence. The second one reads

$$\frac{\sin\,\alpha_\mathrm{i}}{\sin\,\alpha_\mathrm{tr}} = \frac{n_\mathrm{II}}{n_\mathrm{I}};\ \boldsymbol{k}_\mathrm{i}, \boldsymbol{k}_\mathrm{r} \text{ and } \boldsymbol{e}_\mathrm{n} \text{ are in one plane.} \tag{3.10c}$$

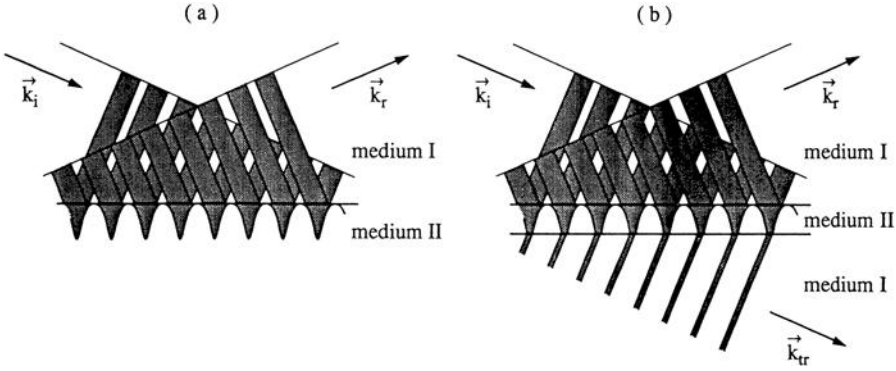


Fig. 3.3. The evanescent wave in the case of total internal reflection (**a**) and the arrangement for the "optical tunneling" effect or attenuated total reflection (**b**)

In Fig. 3.1 we manifest the situation for $n_{\mathrm{I}} < n_{\mathrm{II}}$, i.e., the refraction from an optically thinner into an optically thicker medium. In the opposite case, one reaches a critical angle $\alpha_{\mathrm{i}}^{\mathrm{c}}$ for which $\alpha_{\mathrm{tr}} = 90^{\mathrm{o}}$ given by the condition

$$\alpha_{\mathrm{i}}^{\mathrm{c}} = \arcsin \frac{n_{\mathrm{II}}}{n_{\mathrm{I}}}. \tag{3.11}$$

For $\alpha_{\mathrm{i}} \geq \alpha_{\mathrm{i}}^{\mathrm{c}}$ there is a totally reflected beam but no longer a transmitted one. However, the boundary conditions (3.6) and (3.7) require finite field amplitudes in medium II. Inspection of the boundary conditions shows that a so-called evanescent wave exists in medium II which propagates parallel to the surface. Its field-amplitudes decay exponentially in the direction normal to the interface over a distance of a few wavelengths, as shown schematically in Fig. 3.3a. The reflected wave has under these conditions the same intensity as the incident one. Correspondingly the phenomen is known as total (internal) reflection.

If medium II has only a thickness of the order of a wavelength and is then covered by material I again, then the evanescent wave couples into this medium giving rise to a propagating refracted wave (Fig. 3.3b). Consequently the intensity of the reflected wave decreases. This phenomenon is called attenuated, or frustrated, total reflection (ATR) or the optical tunnel effect in analogy to the quantum-mechanical tunnel effect.

The laws (3.10),(3.11) can be also deduced from the principle of Maupertius or Fermat , which says that for geometrical optics the optical path length, i.e., the product of the geometrical path length and the refractive index n between two points A and B is an extremum, generally a minimum. This is shown schematically for the case of refraction in Fig. 3.4. From all in principle possible ways to travel from A to B,the light propagation is along the one for which (3.12) holds.

$$\delta \int n \mathrm{d}s = \delta(n_{\mathrm{I}}\overline{AC} + n_{\mathrm{II}}\overline{CB}) = 0\,, \tag{3.12}$$

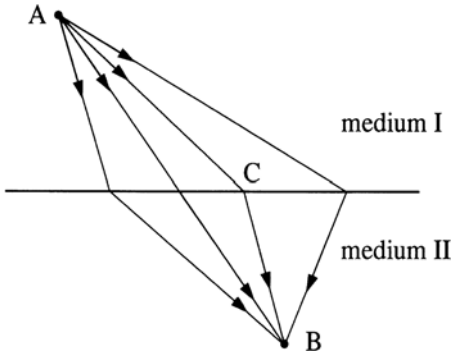


Fig. 3.4. Various possible paths of a light beam travelling from A to B

i.e., the variation δ of the optical path length vanishes. Equation (3.10) can be deduced from (3.12).

3.1.3 Noether's Theorem and Some Aspects of Conservation Laws

A third way of deriving (3.9),(3.10) relies on the law of momentum conservation. Since conservation laws are essential in other fields as well as in physics (See problem 1 of Chap. 1) we shall dwell here on them for some time.

We start with the theorem of E. Noether, which is usually not taught in standard physics courses, though it is of great importance. In simple words it says:

A conservation law follows from every invariance of the Hamilton operator H.

We are not going to prove this statement. Instead we give some well known applications.

If H is invariant against infinitesimal translations in time $\mathrm{d}t$, i.e., if H does not depend explicitly on time, then the total energy E of the system described by H is conserved

$$H(t) = H(t + \mathrm{d}t) \longrightarrow E = \text{const.} \tag{3.13a}$$

If H is invariant against an infinitesimal translation along an axis x, then the x-component of the total momentum $\boldsymbol{p}$ is conserved

$$H(x) = H(x + \mathrm{d}x) \longrightarrow p_x = \text{const.} \tag{3.13b}$$

If H is invariant against an infinitesimal rotation $\mathrm{d}\phi$, e.g., around an axis z, that is, $\boldsymbol{d\phi} = (0, 0, \mathrm{d}\phi)$ then the z-component of the angular moment $\boldsymbol{L}$ is conserved

$$H(\boldsymbol{r}) = H(\boldsymbol{r} - \boldsymbol{r} \times \boldsymbol{d\phi}_z) \longrightarrow L_z = \text{const.} \tag{3.13c}$$

The axis along which (3.13c) is valid is called the quantization axis.

For the problem of light reflection and refraction at an interface, (3.13a) is still valid, resulting in energy conservation which we used, e.g., in connection with (3.9). Concerning (3.13b) the problem is only invariant against infinitesimal translation parallel to the interface. Correspondingly only the momentum parallel to this interface is conserved. We learned with (2.58) that the momentum of photons (and of all other free particles) is $\hbar\boldsymbol{k}$. Consequently at the interface the conservation laws must be fufilled.

$$k_{\mathrm{i}\parallel} = k_{\mathrm{r}\parallel}\,, \tag{3.14a}$$

$$k_{\mathrm{i}\parallel} = k_{\mathrm{tr}\parallel}\,. \tag{3.14b}$$

Since incident and reflected beams propagate in the same medium, the lengths of the wave vectors are equal, too.

$$|\boldsymbol{k}_{\mathrm{i}}| = |\boldsymbol{k}_{\mathrm{r}}|\ , \tag{3.15}$$

The only solution for (3.14a) and (3.15) is then obviously the law of reflection (3.10). For the relation of $\boldsymbol{k}_{\mathrm{i}}$ and $\boldsymbol{k}_{\mathrm{t}}$ we find accordingly in addition to (3.15)

$$k_j = k_{\mathrm{vac}} n_j = \frac{\omega}{c} n_j\,;\ \ j = \mathrm{I},\,\mathrm{II}. \tag{3.16}$$

The simultaneous solution of (3.15) and of (3.16) gives just (3.10c).

When we describe damping by a complex wave vector (2.36),(2.37) then the above conservation laws applied to the real, i.e., oscillatory parts of $\boldsymbol{k}$. For clarity, the situation is depicted (again) in Fig. 3.5.

The conservation law (3.13c) still holds for a quantization axis perpendicular to the interface.

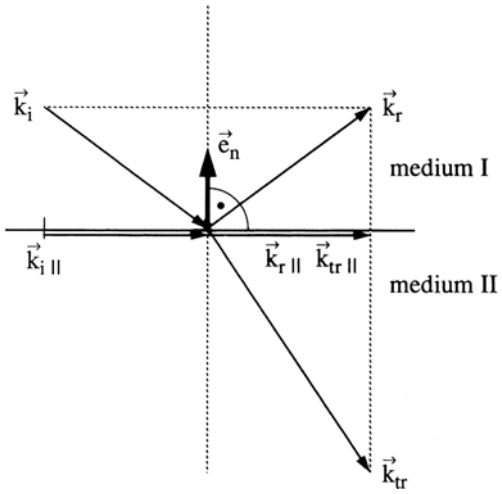


Fig. 3.5. The momenta $\hbar\boldsymbol{k}$ of incident, reflected and refracted beams at an interface

3.1.4 Reflection and Transmission at an Interface and Fresnel's Formulae

Continuing to exploit the boundary conditions (3.6),(3.7) we give now the results for the transmission t and the reflection r of an interface for the simplifying assumption that both media are transparent. This means that the imaginary part κ of the complex index of refraction $\tilde{n}$ is negligible, i.e., much smaller than the real part n. This is the case of weak absorption.

$$|\kappa| \ll |n| \backsimeq 1 \text{ weak absorption}, |\kappa| \gg |n| \backsimeq 1 \text{ strong absorption.} \tag{3.17}$$

The resulting equations are known as the Fresnel formulae for the regime of weak absorption. They read, according to [76P1,77B1,87H1,97L1] (as usual our treatment is for non-magnetic materials)

$$r_\perp = \frac{n_\mathrm{I}\cos\alpha_\mathrm{i} - n_\mathrm{II}\cos\alpha_\mathrm{tr}}{n_\mathrm{I}\cos\alpha_\mathrm{i} + n_\mathrm{II}\cos\alpha_\mathrm{tr}} = -\frac{\sin(\alpha_\mathrm{i} - \alpha_\mathrm{tr})}{\sin(\alpha_\mathrm{i} + \alpha_\mathrm{tr})}, \tag{3.18a}$$

$$r_\parallel = \frac{-n_\mathrm{II}\cos\alpha_\mathrm{i} + n_\mathrm{I}\cos\alpha_\mathrm{tr}}{n_\mathrm{I}\cos\alpha_\mathrm{tr} + n_\mathrm{II}\cos\alpha_\mathrm{i}} = -\frac{\tan(\alpha_\mathrm{i} - \alpha_\mathrm{tr})}{\tan(\alpha_\mathrm{i} + \alpha_\mathrm{tr})}, \tag{3.18b}$$

$$t_\perp = \frac{2n_\mathrm{I}\cos\alpha_\mathrm{i}}{n_\mathrm{I}\cos\alpha_\mathrm{i} + n_\mathrm{II}\cos\alpha_\mathrm{tr}} = \frac{2\sin\alpha_\mathrm{tr}\cos\alpha_\mathrm{i}}{\sin(\alpha_\mathrm{i} + \alpha_\mathrm{tr})} \tag{3.18c}$$

$$t_\parallel = \frac{2n_\mathrm{I}\cos\alpha_\mathrm{i}}{n_\mathrm{I}\cos\alpha_\mathrm{tr} + n_\mathrm{II}\cos\alpha_\mathrm{i}} = \frac{2\sin\alpha_\mathrm{tr}\cos\alpha_\mathrm{i}}{\sin(\alpha_\mathrm{i} + \alpha_\mathrm{tr})\cos(\alpha_\mathrm{i} - \alpha_\mathrm{tr})}. \tag{3.18d}$$

The relation between α_i and α_tr according to (3.10c) is used to progress from one set of formula to the other. The signs in (3.18) depend on the way in which we defined the field amplitudes in Fig. 3.1. However it is obvious from (3.18a,b), that there is for normal incidence ($\alpha_\mathrm{i} = 0°$) a phase shift of 0 or of π between the incident and reflected field for reflection at the optically thinner or thicker medium, respectively, i.e.,

$$r_{\parallel,\perp}(\alpha_\mathrm{i} = 0) = \frac{n_\mathrm{I} - n_\mathrm{II}}{n_\mathrm{I} + n_\mathrm{II}};\ t_{\parallel,\perp}(\alpha_\mathrm{i} = 0) = \frac{2n_\mathrm{I}}{n_\mathrm{I} + n_\mathrm{II}}\,. \tag{3.18e}$$

These signs are, however, not a serious problem since it is usually not possible to measure directly the field amplitudes in the optical regime but only the intensities $I = \langle S \rangle$.[See (2.21),(2.22)]

We display in Fig. 3.6 formulae (3.18) graphically and the phase shift between the various reflected components, assuming that the incident ones are "in phase". Furthermore we show the results for strong absorption not covered by (3.18).

The experimentally accessible quantities are R and T which can be calculated from (2.22),(3.2).

We discuss first the reflectivity R. For the orientation $R_\perp$, R increases monotonically with α_i. The limiting values for $\alpha_\mathrm{i} = 0°$ and $\alpha_\mathrm{i} = 90°$ are given in (3.19).

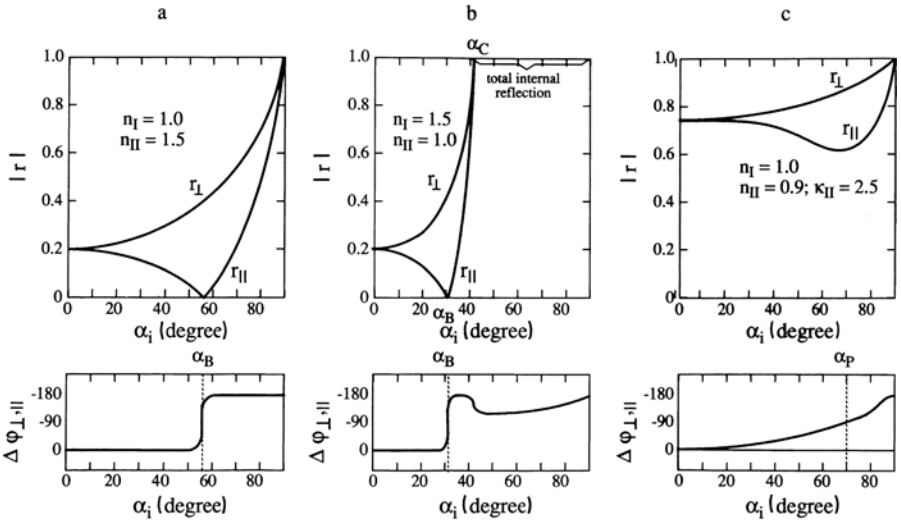


Fig. 3.6. The reflection coefficient r for the (electric) field amplitudes according to (3.18) and the phase difference between the reflected amplitudes for the orientations $r_\perp$ and $r_\parallel$, and reflection at an optically thicker (**a**) and thinner (**b**) medium and for reflection at a strongly absorbing medium (**c**). According to [76P1]

For weak absorption ($n \gg \kappa$)

$$R(\alpha_\mathrm{i} = 0) = \left(\frac{n_\mathrm{II} - n_\mathrm{I}}{n_\mathrm{II} + n_\mathrm{I}}\right)^2 \tag{3.19a}$$

For $n_\mathrm{I} = 1$ the answer, again for weak absorption, is

$$R(\alpha_\mathrm{i} = 0) = \left(\frac{n_\mathrm{II} - 1}{n_\mathrm{II} + 1}\right)^2; \tag{3.19b}$$

and for strong absorption and $n_\mathrm{I} = 1$

$$R(\alpha_\mathrm{i} = 0) = \frac{(n_\mathrm{II} - 1)^2 + \kappa_\mathrm{II}^2}{(n_\mathrm{II} + 1)^2 + \kappa_\mathrm{II}^2}, \tag{3.19c}$$

while for grazing incidence

$$R(\alpha_\mathrm{i} = 90^\mathrm{o}) = 1 \tag{3.19d}$$

in all cases.

$R_\perp$ increases monotonically with increasing α_i in all cases. In contrast $R_\parallel$ goes through a minimum at a certain angle α_B with

$$R_\parallel(\alpha = \alpha_\mathrm{B}) = 0 \tag{3.20}$$

for the case of weak absorption. The angle α_B is known as Brewster's angle or the polarization angle. For $\alpha = \alpha_B$ only the component polarized perpendicularly to the plane of incidence is reflected. So this angle can be used to polarize light, if unpolarized light is directed to the interface. Note that the transmitted beam is not strictly polarized, but has only some preference for the orientation parallel to the plane of incidence.

The condition which comes from (3.18b) for $r_\parallel = 0$ is

$$n_{II}\cos\alpha_i = n_I\cos\alpha_{tr} \text{ or } \tan(\alpha_i + \alpha_{tr}) = \infty. \tag{3.21}$$

Equation (3.21) has, apart from the trivial solution (no interface $\rightarrow n_I = n_{II}$; $\alpha_i = \alpha_{tr}$), the solution

$$\alpha_i + \alpha_{tr} = 90^o\ , \tag{3.22}$$

i.e., the reflected and refracted beams propagate perpendicularly to each other. This fact can be easily understood. As we shall see in Chap. 4 the reflected beam is radiated from the forced oscillations of the atoms close to the surface which forms the optically thicker medium (assuming for the moment that medium I is vacuum). Since dipoles do not radiate in the direction of their axis, and since the polarization in the medium is perpendicular to $\boldsymbol{k}_{tr}$ (transverse wave) we find directly (3.22).

For the case $n_{II} < n_I$ we also find the critical angle α_c for total internal reflection in Fig. 3.6b which we mentioned already earlier.

If we send light on the interface polarized differently than $E_\parallel$ or $E_\perp$ to the plane of incidence, we can decompose it always into two components with the above orientations, we calculate their reflected or transmitted amplitudes with (3.18) or Fig. 3.6 and superpose them again, taking into account the relative phase shifts given in Fig. 3.6. In the general case of a phase shift different from 0^o or 180^o the reflected light will be elliptically polarized for a linearly polarized incident beam. However, in experimental investigations of the optical properties of semiconductors, one tries to avoid this additional complication, usually by choosing the simplest geometries.

For strong absorption, R does not reach zero for any polarization (Fig. 3.6c), and starts already for $\alpha_i = 0$ rather close to one. This leads to a statement which may seem contradictory in itself at first glance: strongly absorbing materials absorb only a small fraction of the incident light. The solution is clear, since the bigger fraction is reflected. The smaller fraction which actually enters the medium is absorbed however over a short distance. In a weakly absorbing medium, the major portion of the light is transmitted through the surface and may be completely absorbed if the medium is thick enough. Indian ink is in the sense of (3.17) a weakly absorbing medium, metals are strongly absorbing over wide spectral ranges and have R close to unity.

With increasing α_i, $R_\perp$ increases monotonically while $R_\parallel$ goes through a shallow minimum as seen in Fig. 3.6c. The principle angle of incidence α_p is defined by some authors as the α_i for which the slopes of the curves $R_\parallel(\alpha_i)$ and $R_\perp(\alpha_i)$ are equal. Other authors prefer to use the minimum of $R_\perp$ as

definition of α_p. The difference is marginal. The phase shift between the two components is just $\frac{\pi}{2}$ for α_p.

A consequence of the smooth variation of the phase shift with α_i is the fact that linearly polarized light impinging on a metallic mirror is usually elliptically polarized after reflection except for the simple orientations $E_\parallel$, and $E_\perp$. This fact should be remembered when building an optical setup in the lab.

To conclude this subsection, we shall shortly consider transmission through a single interface. For a lossless interface, energy conservation requires for the incident, reflected and transmitted powers P of the light

$$P_\mathrm{i} = P_\mathrm{r} + P_\mathrm{tr}, \tag{3.23a}$$

where the power is defined as energy per unit of time.

Despite (3.23a), $\mid r \mid^2$ and $\mid t \mid^2$ do not add up to unity, since these quantities give information about the reflected and transmitted light intensities. This quantity gives, as stated already earlier, the energy flux density, i.e., the energy per unit of time and of area. Since the cross-sections of the incident and reflected beams are equal, but different from the transmitted one for $\alpha_\mathrm{i} \neq 0$ as shown in Fig. 3.7, a corresponding correction factor has to be added to $T = |t|^2$ to fulfill (3.23a):

$$T_{\parallel,\perp} = \frac{I_\mathrm{tr}}{I_\mathrm{i}} \frac{\cos\alpha_\mathrm{tr}}{\cos\alpha_\mathrm{i}} = \mid t_{\parallel,\perp} \mid^2 \frac{\cos\alpha_\mathrm{tr}}{\cos\alpha_\mathrm{i}} \ . \tag{3.23b}$$

For more details about transmission and reflection at a plane interface see [76P1, 77B1, 87H1, 97L1] or the references given therein.

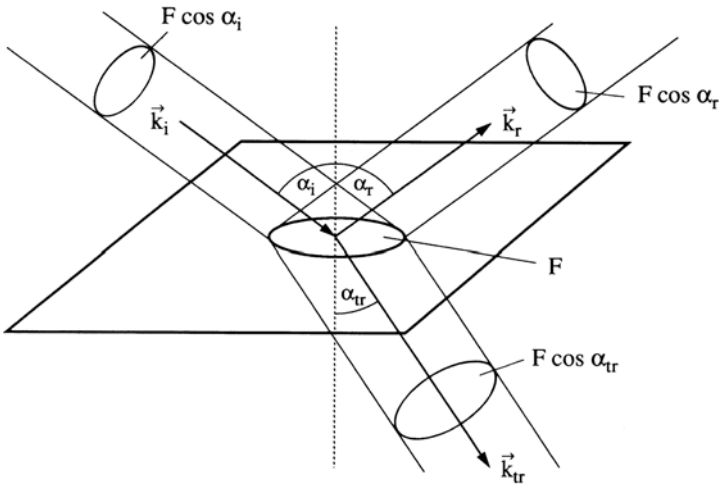


Fig. 3.7. The relation of the cross sections of incident, transmitted and reflected beams at a plane interface between two media

3.1.5 Extinction and Absorption of Light

Until now we have considered mainly what happens in the immediate surroundings of a plane interface between two media. Now we shall spend a few minutes on the propagation of a beam within a medium, continuing the discussion in Sect. 2.3 in connection with (2.37).

If light propagates in a medium other than vacuum, its intensity or field amplitude decreases in most cases with increasing distance, even if we (un-physically) assume a strictly parallel beam and neglect all diffraction losses connected with its finite diameter. In reality, both assumptions above can be fulfilled only to a good approximation but never in a strict, mathematical sense. The decay is usually exponential with increasing distance d [See (2.37)–(2.40)].

$$I = I_0 \mathrm{e}^{-\alpha(\omega)d} \ . \tag{3.24}$$

If energy is pumped into a suitable material and in a suitable way $\alpha(\omega)$ may even become negative for a certain range of frequencies and consequently light is amplified. We call these materials active, laser or inverted materials in contrast to passive materials with $\alpha \geq 0$.

The attenuation of light according to (3.24) is called "extinction". It comprises two groups of phenomena:

$$\alpha_{\text{extinction}}(\omega) = \alpha_{\text{absorption}}(\omega) + \alpha_{\text{scattering}}(\omega) \tag{3.25}$$

Extinction is the more comprehensive term. It enters in the damping γ in Chap. 4 or in the phase relaxation time T_2 discussed in more detail in Chap. 23 via $\gamma = 2\hbar {T_2}^{-1}$ and contains two contributions. Absorption is the transformation of the energy of the light field into other forms of energy like heat, chemical energy or electromagnetic radiation which is not coherent and generally also frequency shifted with respect to the incident beam. This latter phenomenon is usually called (photo-) luminescence. The other contribution to extinction is attenuation by (coherent) scattering of light. The unshifted component is called Rayleigh Scattering and requires some disorder in the medium. We come back to this phenomenon in Chap. 23. The frequency shifted (coherent) parts are known as Raman- or Brillouin Scattering. See Chap. 11 for these effects or [75C1]. If the scattering particles do not show absorption in the visible and have typical sizes large compared to the wavelength of the light λ, the material usually looks white. Examples are the powder of ZnO (just to start with a semiconductor), ground sugar and salt, clouds, snow, the foam of beer, milky quartz or the bark of a birch tree. The reason is that the scattering of light at the interfaces by reflection and refraction is roughly wavelength independent and thus the same for all colors.

If the particles are small compared to λ, the scattering at these particles becomes wavelength dependent. Often one finds an ω^4 law

$$\frac{I_{\text{scatter}}}{I_{\text{incident}}} \propto \omega^4 \tag{3.26}$$

This relation follows from a combination of (2.26) and (2.22). It explains that the sunlight propagating through clear atmosphere preferentially looses the high frequencies, i.e., short wavelength or blue parts of its spectrum by scattering from N_2, O_2 and other molecules of the atmosphere. Consequently the sun itself looks yellow to red depending on the thickness of air through which the sunlight has to travel and the sky appears blue from the scattered light. The above mentioned disorder necessary for this type of Rayleigh scattering comes from density fluctuations of the particles constituting the air within the coherence volume of (sun-) light. For more details of light scattering, including particles with sizes comparable to the wavelength (the so-called Mie-scattering) see e.g. [77B1, 87H1, 97L1].

If the scattering or absorbing particles are diluted and do not interact with each other, one finds proportionality between their concentration n_p and α

$$\alpha(\omega) = n_\mathrm{p}\alpha_\mathrm{s}(\omega), \tag{3.27}$$

where α_s is the specific extinction constant. Equation (3.27) is also known as Beer's law. Sometimes not the proportionality between α and n_p is denoted as Beer's or Lambert–Beer's law, but the exponential dependence of the light intensity on the distance (3.24).

Though there is evidently a rather clear definition of the terms "extinction" and "absorption", one uses often in "every day" language in the lab and also in many books including this one the word absorption instead of extinction, sometimes for convenience, and sometimes because it is not always clear which group of phenomena is responsible for the attenuation of a light beam along its path through matter.

3.1.6 Transmission Through a Slab of Matter and Fabry Perot Modes

We discuss now in connection with Figs. 3.8 and 3.9 the transmission and reflection of a plane-parallel slab of matter of geometrical thickness d with ideal, lossless surfaces. The surrounding material **I** is air or vacuum ($n_\mathrm{I} = 1$, $\kappa_\mathrm{I} = 0$). The total transmission $\hat{T}$ or reflection $\hat{R}$ does not only depend on material **II** and on the angle α_i but also on the properties of the incident light field, e.g., on its polarization and on its coherence length l_c, i.e., the distance over which there is a fixed phase-relation. We can discuss here only some limiting cases. See also e.g. [76P1, 77B1, 85C1, 87H1, 97L1] or [82L1] of Chap. 1, especially Vol. 17a p 11.

For strong absorption in the sense discussed in connection with (3.18) it is easily possible to detect the beam reflected from the front surface. The transmitted beam is strongly attenuated for $d \gg \lambda \simeq 0.5\,\mu$m in the visible part of the spectrum, i.e., for an optical density $\alpha(\omega)d \gg 1$. Consequently $\hat{R}$ reduces to R given for normal incidence by (3.19c). T is difficult to measure since sufficiently thin samples are often not easily available and

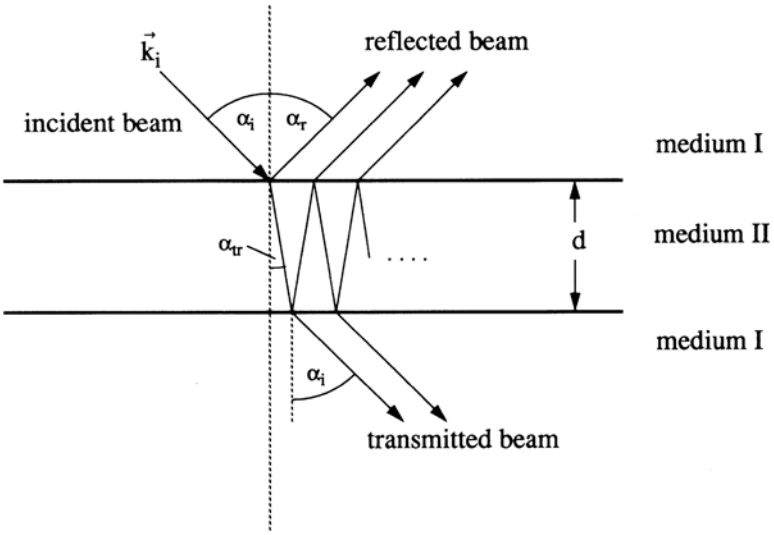


Fig. 3.8. Multiple reflection of an incident light beam in a plane-parallel slab of matter

sometimes have optical properties different from the bulk material. Similar statements for $\hat{R}$ are true for $\alpha(\omega)\mathrm{d} \gg 1$ even if $\alpha(\omega)$ is small. This can be expressed by

$$\hat{R}(\omega) = R(\omega) \text{ for } \alpha(\omega)d \gg 1 \ . \tag{3.28}$$

The most convenient regime in which to measure $\alpha(\omega)$ is $1 \leqslant \alpha(\omega)d \leqslant 5$. The reflection has to be taken into account only once at the front and rear surfaces since multiply reflected beams are very weak due to the absorption. We find with Fig. 3.8

$$\begin{aligned}\hat{T}(\omega) = &\left[1 - R^{\mathrm{I}\to\mathrm{II}}(\omega, \alpha_\mathrm{i})\right] \exp\left[-\alpha(\omega)d\cos^{-1}\alpha_\mathrm{tr}\right] \\ &\times \left[1 - R^{\mathrm{II}\to\mathrm{I}}(\omega, \alpha_\mathrm{tr})\right] \ .\end{aligned} \tag{3.29a}$$

This simplifies for normal incidence to (3.29b) bearing in mind that $R^{\mathrm{I}\to\mathrm{II}} = R^{\mathrm{II}\to\mathrm{I}} = R(\omega)$ for weakly absorbing material (See (3.19a)).

$$\hat{T}(\omega) \simeq \left[1 - R(\omega)\right]^2 \mathrm{e}^{-\alpha(\omega)d}. \tag{3.29b}$$

For the conditions of (3.28) we can write

$$\hat{R}(\omega) \simeq R(\omega) + \left[1 - R(\omega)^2\right] R(\omega)\mathrm{e}^{-2\alpha(\omega)d} \simeq R(\omega) \ . \tag{3.29c}$$

For materials with an optical density $\alpha(\omega)d \leq 1$ things become more complicated again, because we must consider multiple reflection. The behavior depends strongly on the relation of the optical pathlength and the coherence

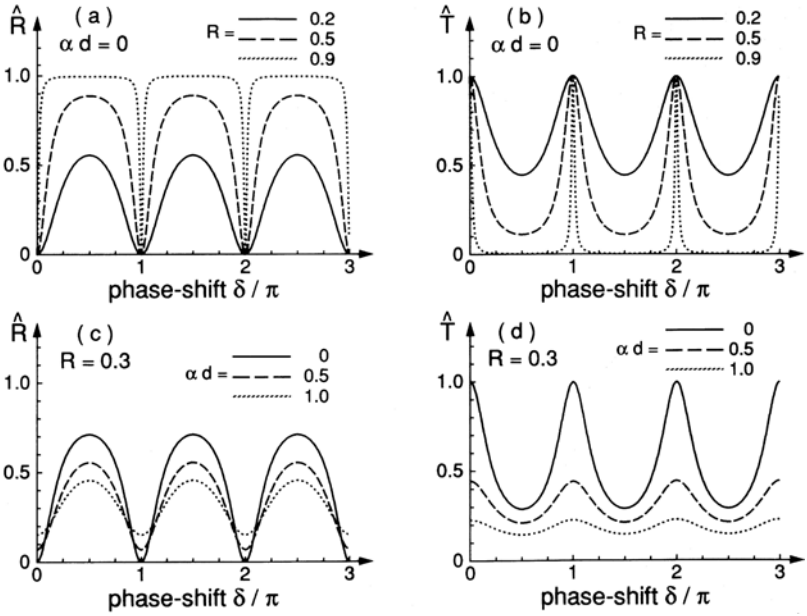


Fig. 3.9. Transmission $\hat{T}$ (**b**, **d**) through and reflection $\hat{R}$ (**a**, **c**) from a Fabry–Perot resonator as a function of the phase shift δ for one half round trip for various values of the reflectivity R of a single surface or the optical density αd of the medium in the resonator

length l_c of the light beam. For $dl_c^{-1} \gg 1$ we have to add intensities resulting in (in this case, for normal incidence)

$$\hat{T}(\omega) \simeq \frac{[1-R(\omega)]^2\,\mathrm{e}^{-\alpha(\omega)d}}{1-R^2(\omega)\,\mathrm{e}^{-2\alpha(\omega)d}} \simeq [1-R(\omega)]^2\,\mathrm{e}^{-\alpha(\omega)d} \text{ for } \alpha d \lesssim 1. \qquad (3.29\mathrm{d})$$

For long coherence lengths, the field amplitudes interfere with approriate phases and the two plane-parallel interfaces form a Fabry–Perot resonator. There are two limiting cases: in one case all partial waves reflected at the two surfaces interfere constructively in the resonator. This condition is fulfilled if an integer number m of half waves fits in the resonator, expressed mathematically as

$$\lambda_m = \frac{d2n(\omega)}{m} \text{ or } \omega_m = \frac{m\pi c}{dn(\omega_m)} \text{ or } k_m = m\frac{\pi}{d} \text{ with } m = 1, 2, 3 \ldots . \qquad (3.30)$$

In this case we have a large field amplitude in the resonator which may surpass even the amplitude of the incident beam, a total transmission $\hat{T}$ close to unity and correspondingly a weak total reflection $\hat{R}$ (even if R is close to unity!). In the opposite case of mainly destructive interference of the partial waves in the resonator, we find just the opposite situation,

that is, $\hat{R} \leq 1$; $\hat{T} \ll 1$. Such a device is called etalon or Fabry–Perot resonator (or FP resonator). Sometimes the two names Fabry and Perot are interchanged.

The general formula for the FP resonator reads approximately [85C1]

$$\hat{T} = \frac{A}{1 + F \sin^2 \delta} \; , \quad \hat{R} = \frac{B + F \sin^2 \delta}{1 + F \sin^2 \delta} \tag{3.31a}$$

with

$$\begin{aligned} F &= \frac{4R_\alpha}{(1 - R_\alpha)^2} \; , \\ A &= \frac{\mathrm{e}^{-\alpha d}(1 - R_\mathrm{F})(1 - R_\mathrm{B})}{(1 - R_\alpha)^2} \; , \\ B &= \frac{R_\mathrm{F}(1 - R_\alpha / R_\mathrm{F})}{(1 - R_\alpha)^2} \; , \end{aligned} \tag{3.31b}$$

R_F: reflectivity of front surface of the Fabry–Perot resonator;
R_B: reflectivity of back surface of the Fabry–Perot resonator;
$R_\alpha = (R_\mathrm{F} R_\mathrm{B})^{1/2} \mathrm{e}^{-\alpha d}$; $\delta = n(\omega) k^\mathrm{vac} d = n(\omega) \omega d / c$.

One often has $R_\mathrm{F} = R_\mathrm{B} = R$. For this condition and $\alpha d = 0$ one has the simplification A = 1, B = 1.

For a lossless medium $\alpha = 0$, and normal incidence these expressions reduce to 1

$$\hat{T} = \frac{1}{1 + F' \sin^2 \delta} \; , \tag{3.32a}$$

$$\hat{R} = \frac{F' \sin^2 \delta}{1 + F' \sin^2 \delta} = 1 - \hat{T} \; . \tag{3.32b}$$

with the phase shift δ given in (3.31) and the finesse F' given by

$$F' = \frac{4R}{(1 - R)^2} \; . \tag{3.32c}$$

In Fig. 3.9 we show $\hat{T}$ and $\hat{R}$ as a function of δ for various values of R and αd. For vanishing damping $\hat{T}(\omega)$ reaches unity and $\hat{R}(\omega)$ zero for the conditions of (3.32). Increasing F' makes the FP resonances narrower. $\hat{T}$ and $\hat{R}$ always add up to unity. For finite damping or α this is no longer the case. $\hat{T}$ remains below unity and the height of the resonance decreases with decreasing F', i.e., increasing α for constant values of R. Sometimes lossless, high finesse Fabry–Perot etalons are used as high resolution monochromators. In this case not only the width of the resonance is of importance, but also the spectral distance between the resonances, the so-called free spectral range given by $\delta = \pi$. A more detailed treatment of the Fabry–Perot and of related problems like dielectric single and multilayer coatings is beyond the scope of this book and [76P1, 77B1, 85C1, 87H1, 89T1, 97L1] are suggested for further reading. We shall come back to Fabry–Perot resonator including oblique incidence in Chap. 17.

3.1.7 Birefringence and Dichroism

Until now we have assumed, for the sake of simplicity, that our sample is isotropic, i.e., that $\varepsilon(\omega)$ is a scalar function and that means with (2.20),(2.21) and (2.27) that $\boldsymbol{D}$ and $\boldsymbol{E}$ as well as $\boldsymbol{S}$ and $\mathit{\Pi}$ are parallel. Later on we shall in general use this assumption again. Here we want to have a short look at what happens if we have an anisotropic material. Indeed, many crystals are anisotropic, including the hexagonal wurtzite structure of several semiconductors. Even cubic crystals can show a weak anisotropy for a finite wave vector $\boldsymbol{k} \neq 0$ since cubic symmetry is lower in symmetry than spherical symmetry, see also [96T1] and Chap. 26. This latter aspect will not be considered for the moment. In the mechanical model which we shall treat in Chap. 4, we can already understand such anisotropies if we assume that some oscillators can be excited (i.e., elongated) only in one direction, e.g., in the x-direction but not in the others. Such an oscillator would react only on the component of an incident electric field polarized $\boldsymbol{E} \parallel \hat{\boldsymbol{x}}$. In the microscopic model the same approach means that the oscillator strength f introduced also in Chap. 4 depends on the direction of polarization and is, e.g., finite for light polarized parallel to a crystallographic axis and zero perpendicular to it. Indeed it is already sufficient that the oscillator strength is different for different orientations of the polarization with respect to the crystallographic axis in order to obtain birefringence.

To describe such situations it is necessary to remember that the dielectric function $\varepsilon(\omega)$ is generally a tensor. It describes the connection between the two vectors $\boldsymbol{D}$ and $\boldsymbol{E}$.

$$D_i = \varepsilon_0 \sum_j \varepsilon_{ij} E_j \;\; ; \;\; i,j = x,y,z \tag{3.33}$$

In principle the ε-tensor has nine components. It can be shown ([77B1, 83K1, 87H1, 97L1] and [74B1] of Chap. 1) that the conservation law for the electromagnetic field energy requires $\varepsilon_{ij} = \varepsilon_{ji}$ so we are left with a maximum of six independent components. Furthermore, it can be shown that every symmetric 3×3 tensor can be brought into diagonal form by a suitable rotation of the cartesian coordinate system. If we choose this coordinate system, the $\varepsilon_{ij} = 0$ for $i \neq j$ and we are left with the three elements on the main diagonal ε_{xx}, ε_{yy}, ε_{zz}, which are all different in the general case, i.e., for so-called biaxial crystals (see below).

Usually one tries to align the cartesian coodinates for $\varepsilon(\omega)$ in a simple way with respect to the crystallographic axes. In uniaxial systems one identifies the z-axis with the crystallographic $\boldsymbol{c}$-axis and the x-y plane with the (usually almost isotropic) plane perpendicular to $\boldsymbol{c}$.

If transformed on these main axes, the tensor $\varepsilon(\omega)$ has therefore in the main diagonal two equal elements

$$\varepsilon_{xx}(\omega) = \varepsilon_{yy}(\omega) \neq \varepsilon_{zz}(\omega) \tag{3.34a}$$

and zeros otherwise.

For biaxial systems of even lower symmetry, such simple connections are often no longer possible and one finds

$$\varepsilon_{xx}(\omega) \neq \varepsilon_{yy}(\omega) \neq \varepsilon_{zz}(\omega) \neq \varepsilon_{xx}(\omega) \ . \tag{3.34b}$$

For cubic crystals one finds for $\boldsymbol{k} = 0$: $\varepsilon_{xx}(\omega) = \varepsilon_{yy}(\omega) = \varepsilon_{zz}(\omega)$. This situation allows us to treat $\varepsilon(\omega)$ as a scalar quantity. For $\boldsymbol{k} \neq 0$ birefringence and dichroism (though weak) may appear for certain orientations as already mentioned [96T1].

Since all of the important semiconductors crystallize either in cubic systems (diamond structure with point-group O_h, zincblende structure T_d or hexagonal ones (wurtzite structure $C_{6\mathrm{v}}$)) we will not go below uniaxial symmetry and refer the reader for these problems to books on crystal optics [76P1,77B1,84A1,97L1] or on crystallography [83K1]. The meaning of "point groups" will be explained Chap. 26.

The consequences of the tensor character of $\varepsilon(\omega)$ are birefringence and dichroism. We briefly outline both effects below. Dichroism means literally that a crystal has two different colors depending on the direction of observation. In a more general sense one describes with the word dichroism every dependence of the absorption spectra on the direction of polarization. In Fig. 3.10 we show schematically transmission spectra for a dichroitic, uniaxial material of a certain thickness. The sample is transparent for both polarizations below $\hbar\omega_1$. The resonance at $\hbar\omega_1$ is assumed to couple more strongly to the light field (i.e., to have larger "oscillator strength") for the orientation $\boldsymbol{E} \perp c$ than for $\boldsymbol{E} \parallel \boldsymbol{c}$. Above $\hbar\omega_2$ light is absorbed almost completely for both orientations. The dichroitic region obviously lies between $\hbar\omega_1$ and $\hbar\omega_2$.

In some materials this region covers a wide spectral range, in some cases the whole visible part of the spectrum. In such a case this material can be used as a polarizer. As an example polaroid films contain long organic molecules which are oriented parallel to each other by the stretching of the film during the manufacturing process. These molecules absorb radiation polarized parallel to the chain, and transmit for the perpendicular orientation over most of the visible spectrum. Another material which is known for its dichroism are some colored varieties of tourmaline. In many semiconductors the dichroitic spectral range is rather narrow and amounts often only to a few tens of meV. These materials are, of course, of no use as commercial polarizers, but the investigation of the dichroism gives very important information on the symmetries and selection rules of the resonances. We will see some examples of this effect in Chap. 13.

If we assume that the eigenfrequencies and/or oscillator strengths of some resonance(s) depend on polarization, then we know immediately from the Kramers-Kronig relations, presented in Chap. 6, that the real part of the refractive index $n(\omega)$ depends also on the orientation of $\boldsymbol{E}$ relative to $\boldsymbol{c}$, i.e., the material is birefringent. We can even state that every dichroitic material must show birefringence and that birefringent material must have some spectral range in which dichroism occurs. For uniaxial materials (e.g., crystals

with uniaxial symmetry C_{6v} or as a prototype calcite with point group D_{3d}) an incident light beam can always be decomposed into two components of the electric field polarized parallel and perpendicular to the main section. The main section is the plane defined by the crystallographic axis and the incident wave vector. The beam polarized perpendicular to the main section is called the ordinary (o) beam. Its refraction is described by Snells' law and the refractive index $\tilde{n}(\omega)$ is independent of orientation. This fact can be understood since the ordinary beam is always polarized perpendicular to the $\boldsymbol{c}$-axis and we assume that uniaxial materials are isotropic in the plane $\perp \boldsymbol{c}$. This is strictly correct only for vanishing wave vectors (and corresponds just to the situation for the dipole approximation in Sect. 3.2.2 and to a very good approximation for small but finite $\boldsymbol{k}$ values. The so-called extraordinary beam (eo), the polarization of which falls in the main section, has components $\boldsymbol{E} \parallel \boldsymbol{c}$ and $\boldsymbol{E} \perp \boldsymbol{c}$ the weights of which depend on the angle $\gamma = \angle(\boldsymbol{k}, \boldsymbol{c})$. It is not surprising that the refractive index experienced by the extraordinary beam depends an γ, since the relative coupling to the oscillators active for the orientations $\parallel \boldsymbol{c}$ or $\perp \boldsymbol{c}$ changes with γ. For a general direction of incidence an unpolarized (or elliptically polarized) beam will be decomposed into two beams polarized perpendicular to each other – the ordinary and the extraordinary ones - which will be separated in space, as shown schematically in Fig. 3.11. This is the concept, which allows us to use birefringent materials as polarizers. We note here already that the wave vector $\boldsymbol{k}$ and the Poynting vector $\boldsymbol{S}$ of the eo beam are not necessarily parallel to each other. The reason will be given in connection with Fig. 3.13. In the case of Fig. 3.11 the refractive index of the

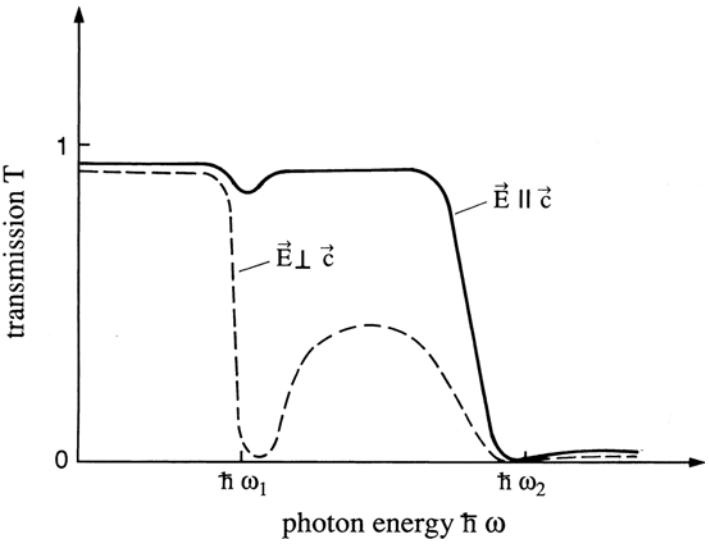


Fig. 3.10. Schematic drawing of the transmission spectra of a dichroitic material for two different polarizations of light with respect to the crystallographic axis

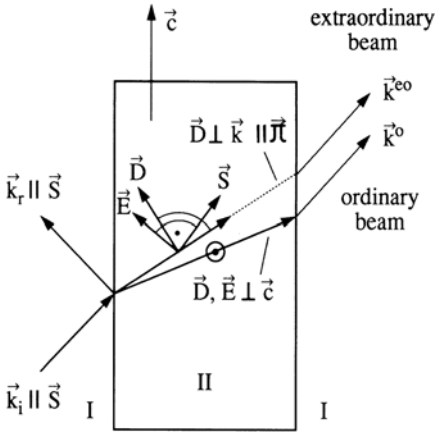


Fig. 3.11. Birefringence for an unpolarized beam falling on a birefringent material at a given angle with the crystallographic axis $\boldsymbol{c}$ parallel to the interface

ordinary beam is greater than that of the extraordinary one. This situation is called negative-uniaxial birefringence, the opposite situation corrrespondingly positive-uniaxial birefringence, since Δn is defined as

$$\Delta n := n_{\mathrm{eo}} - n_{\mathrm{o}} \tag{3.34c}$$

where the indices o and eo stand for ordinary and extraordinary.

There are two limiting orientations which result in rather clear and simple situations. Therefore these orientations are usually investigated in semiconductor optics. One situation is $\boldsymbol{k} \parallel \boldsymbol{c}$. In this case the $\boldsymbol{E}$ field can be only perpendicular to c, this means one observes the ordinary beam only, independent of the polarization of the incident beam. The other clear orientation is $\boldsymbol{k} \perp \boldsymbol{c}$. In this case one can choose by a polarizer the orientation $\boldsymbol{E} \perp \boldsymbol{c}$ for the ordinary beam or $\boldsymbol{E} \parallel \boldsymbol{c}$ for the extraordinary beam. In the latter situation, the $\boldsymbol{E}$ field acts only on oscillators which can be elongated parallel to $\boldsymbol{c}$.

Oblique incidence on a surface cut parallel or perpendicular to c or normal incidence on a surface cut under an arbitrary angle (Fig. 3.12) with respect to $\boldsymbol{c}$ are much more complicated to evaluate concerning the spectra of reflection or transmission. The worst situation is, of course, oblique incidence on a plane at an arbitrary angle to the $\boldsymbol{c}$-axis. Scientists working on semiconductor optics usually try to avoid these situations, scientists devoting their work to crystal optics find it challenging and even prefer biaxial systems to others.

A situation which allows us to discuss various aspects of birefringence is perpendicualr incidence on a plane at an oblique angle with respect to $\boldsymbol{c}$. We shall dwell a few minutes on this topic.

The experimental result is shown in Fig. 3.12. The incident beam is split into two when entering the birefringent material. The ordinary beam continues to propagate normal to the surface as expected from Snells' law (3.10b)

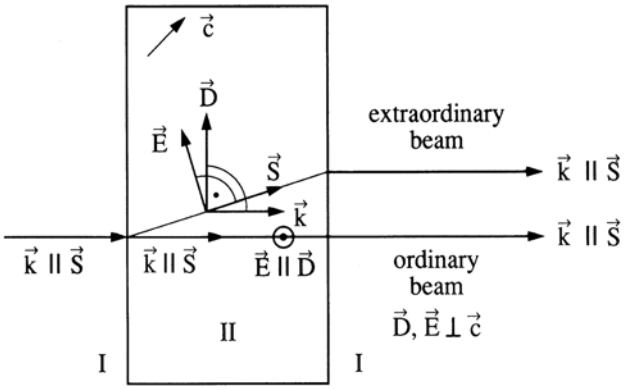


Fig. 3.12. Same as Fig. 3.11 but for normal incidence on a surface cut under an arbitrary angle with respect to $\boldsymbol{c}$

or from the conservation of momentum parallel to the surface (3.14). The extraordinary one seems to violate these two rules. Since a violation of the law of momentum conservation would be very serious, not only for physics, we have to look more closely at this problem.

There are two rather simple ways to present birefringence. One, which we shall outline towards the end of this subsection is in terms of the indicatrix. In the other method, one uses in polar coordinates a plot which gives the phase velocity $v_{\mathrm{ph}} = cn^{-1}(\omega)$. This is basically the inverse of the real part of $\tilde{n}(\omega)$ as a function of the direction of propagation. In a uniaxial system, the figures produced when we include all directions are a sphere for the ordinary beam and a figure with rotational symmetry with respect to $\boldsymbol{c}$ for the extraordinary one. They touch for propagation of light parallel to the crystallographic axis as shown in Fig. 3.13 since there are only o-waves for this orientation, as mentioned above. Now recall Huygens' principle for the propagation of light which says that every point illuminated by an incident primary wave front becomes the source of a secondary wavelet such that the primary wavefront at a later time results from the superposition of the amplitudes of all wavelets. In addition, we must discard the back-travelling waves in the way shown by Fresnel and Kirchhoff which would appear otherwise from the above given principle. With this amendment we can construct wavefronts when we identify the shape of the wavelets with Fig. 3.13. For more details see e.g. [83K1,96T1, 02T1] and references given therein.

Shown in Fig. 3.12 are the $\boldsymbol{S}, \boldsymbol{\Pi}, \boldsymbol{k}, \boldsymbol{E}$ and $\boldsymbol{D}$ vectors. The vectors $\boldsymbol{k}$ and $\boldsymbol{\Pi}$ are always parallel to each other. The magnetic vectors $\boldsymbol{B}$ and $\boldsymbol{H}$ are parallel to each other and normal to the electric vectors and are, in this context, of no further interest since we are dealing with nonmagnetic material.

In Fig. 3.14a the situation is shown for the o-beam. The wavelets are spheres, the resulting wave front is parallel to the vacuum-medium interface. The vectors of energy-flux density $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$, of momentum density

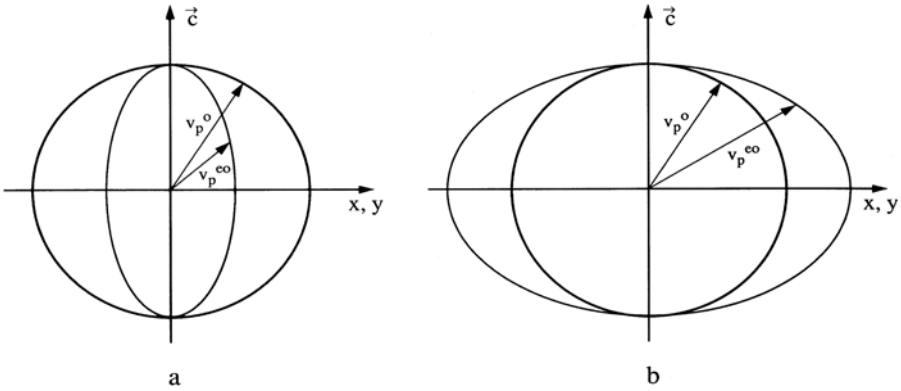


Fig. 3.13. The phase velocity of light in an optically uniaxial material shown for one frequency ω in polar coordinates for positive (**a**) and negative (**b**) birefringence for the ordinary (o) and extraordinary (eo) beams

$\boldsymbol{\Pi} = \boldsymbol{D} \times \boldsymbol{B}$ and the wavevector $\boldsymbol{k}$ ($\hat{=}$ momentum $\hbar\boldsymbol{k}$) are parallel. (Diffraction effects caused by the finite beam diameter are neglected here, though they are obviously also described by Huygens' principle.) The situation for the eo-beam is presented in Fig. 3.14b. The wave front produced by the superposition (or interference) of the wavelets, and constructed as the tangent to the wavelets, is still parallel to the interface. This wave front describes the $\boldsymbol{D}$ field because we know from Maxwell's equations and the boundary conditions of Sect. 3.1.1 that the boundary condition for $\boldsymbol{D}$ is that the normal component D_n is continuous over the interface. In our case, $D_\mathrm{n} = 0$ on both sides. As a consequence, the classical momentum density Π and the momentum $\hbar\boldsymbol{k}$ of the light quanta are still perpendicular to the interface, as required by the conservation of the momentum component parallel to the interface, which is here obviously zero.

On the other hand, we can see that the whole wave front is shifting sideways with continuing propagation into the medium. This shift is described by the Poynting vector $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$. The direction of this vector is just given by the origin of the wavelet and the point where the tangent touches it. The $\boldsymbol{E}$ field is necessarily perpendicular to $\boldsymbol{S}$. As required by the boundary condition for $\boldsymbol{E}$ as deduced from Maxwell's equations, the tangential component $\boldsymbol{E}_\mathrm{t}$, must be the same on both sides of the interface (including incident, refracted and reflected beams). The normal component of $\boldsymbol{E}$ can change, and that is what happens in the orientation of Fig. 3.14b. To summarize, we observe that there are no violations of any conservation laws. The tangential components of the momentum are conserved at the interface and for this quantity Snells' law is still valid. However, the direction of energy propagation given by $\boldsymbol{S}$ changes, but there are no conservation laws for this direction. The law of energy conservation itself has, of course, to be fulfilled, this means in this case

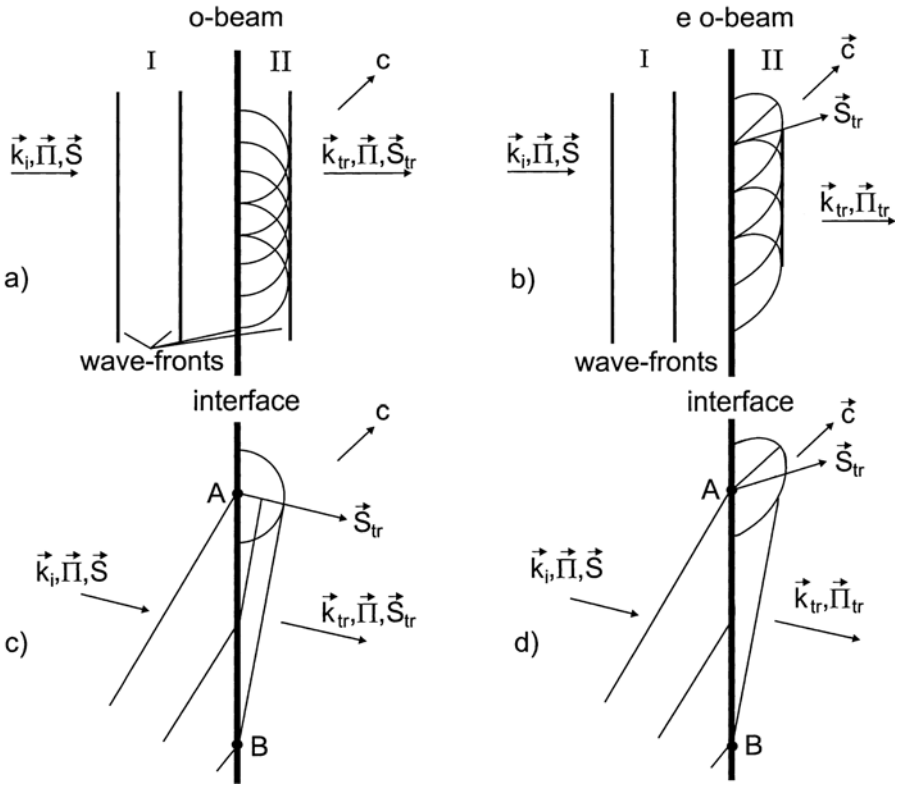


Fig. 3.14. The construction of the wave fronts of the ordinary (**a**, **c**) and the extraordinary (**b**, **d**) beams for various geometries using Fig. 3.13

that the total amount of energy falling per unit time on the interface equals the sum of transmitted and reflected energies.

Figures 3.14c and d finally shows schematically the general situation i.e., oblique incidence on a surface cut at an arbitrary angle with respect to $\boldsymbol{c}$. The construction simply uses the size of the wavelets around point A when the incident wave front just reaches the interface at point B. We see that the refractive index for $\boldsymbol{k}$ is different for both polarisations and that the directions of $\boldsymbol{k}_{\rm tr}$ and $\boldsymbol{S}_{\rm tr}$ differ for the eo beam.

The basic idea of the representation of birefringence by the indicatrix is explained in connection with Fig. 3.15. We plot from the origin lines in all directions with a length equal to the refractive index $n(\omega)$ of a wave with $\boldsymbol{D}$ polarized in this direction. In doing so, we get a sphere for an isotropic material (Fig. 3.15a) and ellipsoids with rotational symmetry for uniaxial materials. In Figs. 3.15b and c we show the situation for positively and negatively birefrigent uniaxial crystals, respectively. The rotation axis coincides in this case with the crystallographic $\boldsymbol{c}$-axis.

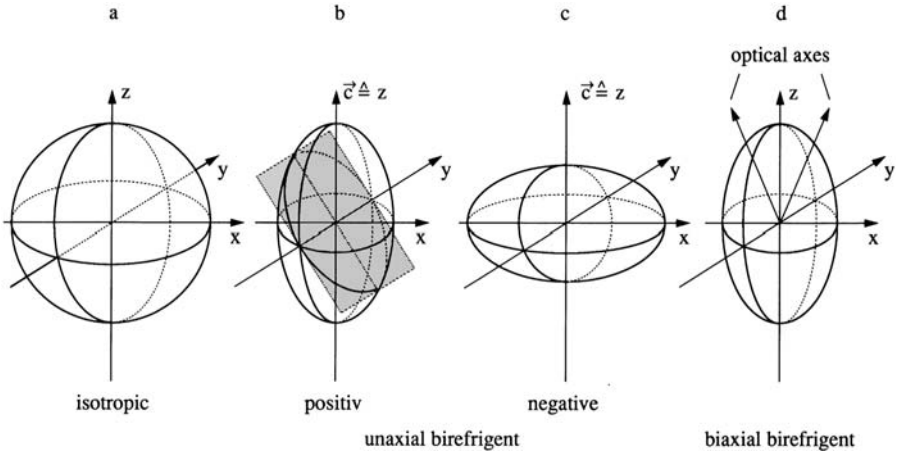


Fig. 3.15. The indicatrix for isotropic (**a**) uniaxial (**b**, **c**) and biaxial materials (**d**) according to [83K1, 87H1]

For crystals of even lower symmetry, all three axes of the indicatrix have different lenghts (Fig. 3.15d) in agreement with the situation for $\varepsilon(\omega)$ in (3.34b).

The use of the indicatrix is now the following: For a given direction of propagation $\boldsymbol{k}$ one cuts the indicatrix with a plane normal to $\boldsymbol{k}$ which contains the origin. In general, this cross section is an ellipse except for isotropic materials. The distance from the origin to this ellipse in a certain direction gives the refractive index of a beam with $\boldsymbol{D}$ polarized in this direction and propagating in the direction of $\boldsymbol{k}$. The optical axes are for anisotropic materials (Figs. 3.15b–d) now defined as those directions of propagation for which the cross section of the indicatrix is a circle. For the uniaxial materials in Figs. 3.15b and c there is obviously only one optical axis, which coincides with the $\boldsymbol{c}$-axis as stated already above. For systems of lower symmetry, where all three main axes of the indicatrix (and all three $\varepsilon_{ii}(\omega)$, $i = x, y, z$ of the dielectric tensor) are different, one finds two optical axes, which generally do not coincide with any of the cartesian coordinates (Fig. 3.15d) and the direction of which can even vary with ω.

It should be mentioned that birefringence can be created in non-birefringent materials, e.g., by the application of electric- or strain-fields. These phenomena are known as Pockels and Kerr effect and as mechanical or stress birefringence or "photoelasticity", respectively. The Pockels effect occurs in crystals without centers of inversion and the effects vary proportially with $\boldsymbol{E}$. The Kerr effect occurs in isotropic solids like glass, in fluids like nitrobenzol and varies therefore proportionally with $\boldsymbol{E}^2$. While these effects are often deliberately investigated, especially the last ones can produce unwanted, spurious effects if a sample is not mounted strain-free, e.g., in a cryostat.

All the phenomena of birefringence which we discussed here for a uniaxial material are as mentioned at the beginning of this subsection, simply due

to the fact that the dielectric function $\varepsilon(\omega)$ is a tensor for crystallographic systems of symmetry lower than cubic and that $\boldsymbol{D}$ and $\boldsymbol{E}$ need therefore no longer be parallel to each other. More information on the topic of this subsection is found in [75C1, 76P1, 77B1, 83K1, 84A1, 87H1, 96T1, 97L1, 02T1] and references therein.

3.1.8 Optical Activity

The last aspect we want to mention in this first subchapter on macroscopic optical properties is optical activity. In this case a linearly polarized light beam remains linearly polarized when propagating through an optically active medium but the direction of polarization rotates slowly by an amount δ_{rot} that varies with increasing distance l, which is the distance the light travels through the medium, i.e.

$$\delta_{\mathrm{rot}} = al \tag{3.35}$$

where a depends on the material and on ω.

While a birefringent material splits an incident beam in two, orthogonally polarized ones, which propagate with different phase velocities through the sample and thus accumulate a phase shift with increasing distance (see problem 8 in Sect. 3.3). An optically active material splits a linearly polarized incident beam into a left (σ^-) and right (σ^+) circularly polarized one with equal amplitude. Again both components travel with different phase velocities. A superposition after a certain distance results again in linearly polarized light as long as there is no difference in the absorption for σ^+ and σ^- polarized light, but with a tilted direction of polarization.

This phenomenon already occurs in amorphous solids including liquids and gases if they contain molecules with a well defined chirality or handiness like dextrose or laevulose which have both the chemical composition $C_6H_{12}O_6$. In crystalline solids, including semiconductors, optical activity occurs if either the molecules in the unit cell have a chirality or if atoms or isotropic molecules are arranged in the unit cell in such a way that a screw-axis appears. Examples are quartz or wurtzite-type crystals parallel to their c-axes, which are both essentially built from tetrahedrons, but are arranged in a way that a shift and a rotation along the axis transforms the crystal into itself.

The optical activity induced in a material by a magnetic field in the direction of light propagation is known as the Faraday effect. In magnetic dichroism one investigates the difference in the absorption spectra for σ^+ and σ^- polarized light induced by the B-field.

3.2 Microscopic Aspects

In contrast to the preceding section, we present now the basic interaction processes between light and matter from a microscopic point of view. We use

here the perturbative or weak coupling approach for the interaction between light and matter, which is in most cases sufficient for dilute systems such as gases. For solids the strong coupling approach is often necessary, which leads to the concept of polaritons and which is introduced in Chap. 5. We describe first in words the basic interaction mechanisms between light and matter, namely absorption, spontaneous and stimulated emission (Sect. 3.2.1), then we proceed to the treatment of linear optical properties in the framework of perturbation theory (Sect. 3.2.2). Since these topics are also treated in many textbooks (e.g. References [81M1,90K1,93H1,96Y1] of Chap. 1 or [55S1,71F1, 73H1, 76H1, 92M1] of Chap. 2) it is not necessary to go into too much detail here.

3.2.1 Absorption, Stimulated and Spontaneous Emission, Virtual Excitation

For simplicity we assume that we have a certain number of two-level "atoms" as shown in Fig. 3.16. Every atom has one electron which can be either in the ground or in the excited state. Later we will extend the model from a two-level system to bands in semiconductors, but the basic interaction processes remain the same.

In Fig. 3.16a an incident photon hits an atom in its ground state. With a certain probability the photon is annihilated and the electron gains enough energy to reach the excited state. For reasons of energy conservation, the photon has to fulfill the condition

$$\hbar\omega = E_{\mathrm{ex}} - E_{\mathrm{g}} \ , \tag{3.36}$$

where $E_{\mathrm{ex}} - E_{\mathrm{g}}$ is the energy difference between the ground and excited states. We call this process absorption in agreement with the definition in Sect. 3.1.5 if the energy of the photon is soon converted into other forms of energy, that is, if the electron undergoes some scattering processes, which destroy its coherence or more precisely the coherence of the electric polarization connected with this transition to the incident light field. See Chap. 23. The electron eventually returns to its ground state and looses its energy e.g., as phonons i.e., as heat or as a photon which is not coherent with the incident one.

If an incident photon hits an atom with its electron in the excited state it can induce with a certain probability a transition of the electron from the excited to the ground state. In this process a second photon is created which is identical in momentum, energy, polarization and phase to the incident one. This process is called induced or stimulated emission. This process can be used to amplify a photon field. It is therefore the basic mechanism for all lasers (Light Amplification by Stimulated Emission of Radiation). Absorption and stimulated emission are closely related events (Fig. 3.16b).

An electron in the excited state can also with a certain probability reenter the ground state by itself, either by emitting a photon (Fig. 3.16c) or by

loosing the transition energy through phonons or collisions. In the present context, the first mechanism is of interest. It is called spontaneous emission or spontaneous radiative recombination, while the second possibility is known as non-radiative recombination. Spontaneous emission can also be understood in another way. In Sect. 2.5 we saw in connection with (2.54) that photons are similar to harmonic oscillators and have consequently a zero-point energy. This zero-point energy exists for all photon modes. It cannot be absorbed because a harmonic oscillator does not have any states below the zero-point energy, but it can induce a transition in the way as discussed in connection with Fig. 3.16b. So we can consider spontaneous emission as a process induced by the zero-point vibrations of the electro-magnetic field, which are also called vacuum fluctuations (Fig. 3.16c).

The last process presented here is virtual excitation. Understanding this phenomenon often causes some problems for students. So we develop this topic slowly and try to explain it from various points of view in context with Fig. 3.16d. Virtual excitation means the creation of a state with the same wave function as the excited state, but with an energy which is different from the eigenenergy of this excited state. This process becomes possible through the uncertainty principle of quantum mechanics which can be written in space and momentum coordinates.

$$\Delta x_i \Delta p_i \geq h \text{ for } i = 1, 2, 3 \,. \tag{3.37a}$$

A similar relation exists for energy and time

$$\Delta E \Delta t \approx \hbar \,. \tag{3.37b}$$

We need here (3.37b). It says that it is possible to violate energy conservation by an amount ΔE up until a maximum time Δt which fulfills the above condition. Or, in other words, if we want to define the energy with a certain precision ΔE, the state has to exist at least for a time Δt. In principle (3.37b) is valid also in simple classical wave theory (e.g., acoustics) and is very

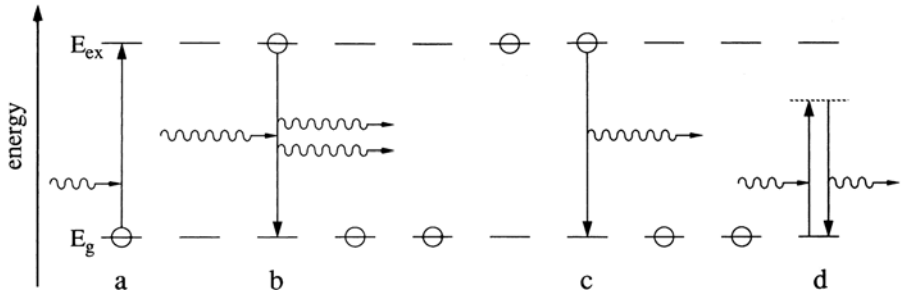


Fig. 3.16. Schematic representations of various interaction processes of light with matter namely absorption (**a**), stimulated emission (**b**), spontaneous emission (**c**) and virtual excitation (**d**)

well known from Fourier transformations. A harmonic oscillation with central frequency ω which lasts only for a time Δt, has a spectral width $\Delta\omega$ given by

$$\Delta\omega\Delta t \geq 1 \; . \tag{3.37c}$$

The connection between (3.37b) to (3.37c) comes simply from the relation

$$E = \hbar\omega \; . \tag{3.37d}$$

It should be mentioned that a similar set of arguments holds for (3.37a), too. If we send now a photon with energy $\hbar\omega'$ to the atom, we can excite the electron for a maximum time Δt given by (3.37b) or for our specific case by

$$\Delta t \approx \hbar[\,|\,(E_{\mathrm{ex}} - E_{\mathrm{g}}) - \hbar\omega'\,|\,]^{-1} \; . \tag{3.38}$$

At the latest after the time Δt has elapsed, the excited state must collapse. The simplest way to collapse is to emit a photon identical to the one which caused the virtual excitation. This "new" photon has however a certain phase delay with respect to the incident photon, because the energy was stored for a maximum time Δt in the atom. As a consequence, an electromagnetic wave propagates with a lower phase-velocity through an ensemble of atoms than through vacuum. The same effect is described phenomenologically by the refractive index $n(\omega)$. (See (2.15) and (2.41).) So we get a first hint of how $n(\omega)$ can be understood and calculated in quantum mechanics. Obviously Δt increases if we approach the resonance condition in (3.38) and consequently wave propagation through a material will deviate more strongly from that through vacuum. This is indeed the case, as we shall see in Sect. 4.3.

If the virtually excited state emits a photon $\hbar\omega'$ in a direction different from the incident one, we have a scattering process as discussed in connection with (3.26). If $\hbar\omega'$ approaches the resonance energy, this scattering process is also known as resonance fluorescence.

In connection with this scattering, we may ask how can light propagate at all in a clear or transparent medium? The answer is that we have in dense media many scattering atoms or centers in the coherence volume of light, independent of whether it is light from an incandescent lamp, a laser or another source. As a consequence, every scattered wave finds another one which has a phase difference, resulting in destructive interference. The only way that all scattered waves interfere constructively is just the usual propagating wave. The explanation for blue sky, which we gave in Sect. 3.1.5 fulfills, apart from the condition that the diameter of the scattering centers is small compared to the wavelength of the scattered radiation, another condition, namely that there are only a few scattering centers i.e., gas molecules, density fluctuations or other inhomogenities per coherence volume of sunlight, so that the mutually destructive interference of the scattered waves is not complete. For a detailed discussion of spatial and temporal coherence see e.g. [93S1] of Chap. 2.

If the virtually excited state disappears under simultaneous emission of a photon and the creation or annihilation of a phonon (i.e., a quantum of the lattice vibrations, as in Sect. 11.1), energy conservation for the emitted photon $\hbar\omega_{\mathrm{R}}$ implies

$$\hbar\omega_{\mathrm{R}} = \hbar\omega' \pm \hbar\Omega_{\mathrm{phonon}} \ . \tag{3.39}$$

This phenomenon is called Raman scattering for optical phonons, and Brillouin scattering for acoustic phonons. The "−" sign gives the Stokes and the "+" sign the anti-Stokes emission. Similar processes are also possible with more than one phonon or with excitations in the electronic or spin system of the semiconductor.

From the few phenomena outlined briefly above which involve virtual excitation, it is obvious that this mechanism is of some importance for the optical properties of matter. Therefore we want to examine it from another point of view and outline the well-known classical analog of the virtually excited states. In addition this analogy gives some justification for the calculation of the dielectric function used in Chap. 4.

Virtual excitation in quantum mechanics corresponds to a driven or forced oscillation in classical mechanics. If we have an oscillator of eigen frequency ω_0 (corresponding to the energetic differences $E_{\mathrm{ex}} - E_{\mathrm{g}}$ in quantum mechanics) and if we excite it with an external frequence ω, it will oscillate with frequency ω after a short damped transient feature of oscillations with ω_0. The amplitude of these steady oscillations increases with decreasing detuning $|\omega - \omega_0|$ depending on the properties of the oscillator, e.g., its damping. This increase of the amplitude of the classical oscillator corresponds to the increase of Δt in (3.39) in the picture of virtual excitation, and it is qualitatively understandable that we will get the strongest deviations of $\varepsilon(\omega)$ or $\tilde{n}(\omega)$ from the vacuum value $\varepsilon = \tilde{n} = 1$ in the vicinity of the resonance ω_0. We elaborate this concept in detail in Chap. 4.

However, before doing so, we shall demonstrate how the various transitions shown in Fig. 3.16 and some others can be treated quantitatively in quantum mechanics by perturbation theory.

3.2.2 Perturbative Treatment of the Linear Interaction of Light with Matter

In this section we present first the Hamiltonian of the system elaborating the perturbation terms. Then we outline shortly how a perturbation causes transitions between various eigenstates. Finally we join these two things together ending with an understanding of the theoretical description of absorption and stimulated or spontaneous transitions.

The Hamiltonian of the total system consisting of the electron states in Fig. 3.16 of the two level atoms (or of the bands of the semiconductor), the radiation field and the interaction of these two systems can be written as

$$H = H_{\mathrm{el}} + H_{\mathrm{rad}} + H_{\mathrm{interac}} \ . \tag{3.40}$$

In the picture of second quantization outlined in Sect. 2.5 H_{interac} contains terms which describe, e.g., the annihilation of a photon and of an electron in the ground state and the creation of an electron in the excited state for the process shown in Fig. 3.16a, weighted with a factor which contains the transition matrix element.

The exact solution of the total Hamiltonian leads to the polariton concept (Chap. 5) in the field of linear optics and describes among other things the changes of the electronic states introduced by the presence of the radiation in nonlinear optics for which we will see some examples in Sects. 20.1 to 20.4.

For our present purposes we follow an approach which is widely used and which treats the radiation field as a small perturbation, that is, we assume that the eigenstates φ and the eigenenergies E_{n} of H_{el}, do not change much in the presence of the electromagnetic field, and that the eigenstates of H_{rad} are the photons described already in Sect. 2.5. The approximation which we use now is known as the semiclassical treatment of radiation. It consists of replacing the canonical conjugate momentum $\boldsymbol{p}$ in the Hamilton function by

$$\boldsymbol{p} \rightarrow \boldsymbol{p} - e\boldsymbol{A} \ , \tag{3.41}$$

where $\boldsymbol{A}$ is the vector potential (2.48).

If we replace $\boldsymbol{p}$ by its operator

$$\boldsymbol{p} = \frac{\hbar}{\mathrm{i}} \nabla \tag{3.42}$$

the single particle Hamiltonian reads

$$H = \frac{1}{2m} \left(\frac{\hbar}{\mathrm{i}} \nabla - e\boldsymbol{A} \right)^2 + V(\boldsymbol{r}) \tag{3.43}$$

including any electrostatic potential into $V(\boldsymbol{r})$. Making use of the Coulomb gauge (2.49) we can evaluate (3.43) to obtain

$$H = -\frac{\hbar^2}{2m} \nabla^2 + V(\boldsymbol{r}) - \frac{e}{m} \boldsymbol{A} \frac{\hbar}{\mathrm{i}} \nabla + \frac{e^2}{2m} \boldsymbol{A}^2 \ , \tag{3.44}$$

$$H = H_{\text{el}} - \frac{e}{m} \boldsymbol{A} \frac{\hbar}{\mathrm{i}} \nabla + \frac{e^2}{2m} \boldsymbol{A}^2 \ , \tag{3.45}$$

$$H = H_{\text{el}} + H^{(1)} + H^{(2)} \ . \tag{3.46}$$

In (3.46) there are two perturbation terms $H^{(1)}$ and $H^{(2)}$. If we assume that A and thus the light intensity are small, and in the regime of linear optics they are small by definition, then $H^{(1)}$ is a perturbation term of first order and $H^{(2)}$ is small of second order. Consequently $H^{(1)}$ must be used in first-order perturbation theory. In the second-order approximation we have to use $H^{(1)}$ in second-order perturbation theory and $H^{(2)}$ in

first-order perturbation theory, etc. We shall come back to this latter aspect in Chap. 19.

In order to arrive at Fermi's golden rule for the transition rate w_{ji} from an initial state i (e.g., the ground state g in Fig. 3.16) to another state j (e.g., the excited state ex in Fig. 3.16) one uses the time-dependent Schrödinger equation

$$H\psi = \mathrm{i}\hbar\frac{\partial\psi}{\partial t} \tag{3.47a}$$

with the stationary solutions when $H = H_0$

$$\psi_\mathrm{n}(\boldsymbol{r},t) = \varphi_\mathrm{n}(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(E_\mathrm{n}/\hbar)t} \ . \tag{3.47b}$$

For the solution in the presence of a perturbation $H^{(1)}$ we make the ansatz

$$\psi(\boldsymbol{r},t) = \sum_\mathrm{n} a_\mathrm{n}(t)\varphi_\mathrm{n}(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(E_\mathrm{n}/\hbar)t} \ . \tag{3.48}$$

We assume that the perturbation is switched on at $t = 0$. Before the system is in state i, i.e.,

$$\left.\begin{aligned} a_\mathrm{i}(t) &= 1 \\ a_{\mathrm{n}\neq\mathrm{i}}(t) &= 0 \end{aligned}\right\} \quad \text{for } t \leq 0 \ . \tag{3.49}$$

For $t > 0$ the $a_{\mathrm{n}\neq\mathrm{i}}(t)$ start to grow and under these conditions the transition rate w_{ij} of Fermi's golden rule becomes

$$w_{ij} = \frac{2\pi}{\hbar}\left|H_{ij}^{(1)}\right|^2 D(E) \ , \tag{3.50a}$$

where $D(E)$ is the density of the final states modified by momentum conservation if applicable. $H_{ij}^{(1)}$ is the transition matrix element given by

$$H_{ij}^{(1)} = \int \psi_j^*(\boldsymbol{r})H^{(1)}\psi_i(\boldsymbol{r})\,d\tau =: < \psi_j \mid H^{(1)} \mid \psi_i > \tag{3.50b}$$

For a non-degenerate two level system $D(E)$ is simply one per atom. The square of the transition matrix element $|H_{ij}|^2$ is known as the transition probability. Later on we will assume, for simplicity of writing, that some constant factors as the term $2\pi/\hbar$ are incorporated in this $|H_{ij}|^2$. Transition probabilities are given apart from some coefficients by the square of the respective transition matrix elements of (3.50) in the case of first order perturbation and by the terms (3.51) for second order.

The transition rate w_{ij} is proportional to the transition probability multiplied by the square of the amplitude of the perturbation $H^{(1)}$, i.e., here by $|\boldsymbol{A}_0|^2 \sim I$ where I is the light intensity, i.e. as already mentioned the energy flux per unit of area and time.

If the first-order perturbation term (3.50b) vanishes, then according to what we stated above the second-order contribution reads

$$w_{ij}^{(2)} = \frac{2\pi}{\hbar} \left| \sum_{k \neq i,j} \frac{H_{jk}^{(1)} H_{ki}^{(1)}}{E_i - E_k} + H_{ij}^{(2)} \right|^2 D(E) \ . \tag{3.51}$$

We restrict ourselves for the moment to the first order according to (3.50) and discuss the term $H_{ij}^{(1)}$ in some more detail for the perturbations of (3.45). With the vector potential $\boldsymbol{A}$

$$\boldsymbol{A} = \boldsymbol{A}_0 e^{i(\boldsymbol{k}\boldsymbol{r} - \omega t)} \tag{3.52a}$$

we find, e.g., for the absorption process (3.50)

$$\begin{aligned} w_{\mathrm{g}\to\mathrm{ex}} =& \frac{2\pi}{\hbar} \left| \frac{-e\hbar}{im} A_0 \cdot \int \varphi_{\mathrm{ex}}^*(\boldsymbol{r})\, \mathrm{e}^{\mathrm{i}(E_{\mathrm{ex}}/\hbar)t} \boldsymbol{e}_{\mathrm{A}}\, \mathrm{e}^{\mathrm{i}(\boldsymbol{k}\boldsymbol{r} - \omega t)} \nabla \varphi_{\mathrm{g}}(\boldsymbol{r})\, \mathrm{e}^{-\mathrm{i}(E_{\mathrm{g}}/\hbar)t} \mathrm{d}\tau \right|^2 D(E_{\mathrm{g}} + \hbar\omega) \\ & \propto |A_0 < \varphi_{\mathrm{ex}}|\, H^{(1)}\, |\varphi_{\mathrm{g}} >|^2\, D(E_{\mathrm{g}} + \hbar\omega) =: \\ & A_0^2 \mid H_{\mathrm{eg}}^{(1)} \mid^2 D(E_{\mathrm{g}} + \hbar\omega)\,, \end{aligned} \tag{3.52b}$$

where e_{A} is the unit vector in the direction of $\boldsymbol{A}$. We also give in (3.52) a generally used abreviation for the integral. A significant transition rate occurs only if the time dependent exponential functions vanish, or, mathematically

$$E_{\mathrm{ex}} - E_{\mathrm{g}} - \hbar\omega = 0\,. \tag{3.53a}$$

This is again the law of energy conservation. If the φ_i have plane-wave character and are described by a wave vector $\boldsymbol{k}$, a similar argument results in $\boldsymbol{k}$ conservation.

$$\hbar\boldsymbol{k}_{\mathrm{ex}} - \hbar\boldsymbol{k}_{\mathrm{g}} - \hbar\boldsymbol{k} = 0 \tag{3.53b}$$

This is not the case for the two-level atoms discussed here but is true for most of the eigenstates of crytalline semiconductors.

We see that the transition rate is proportional to A_0^2, and thus to the light intensity $I = \langle S \rangle$ or the density of photons $N_{\mathrm{ph}}(\omega)$ in a certain mode:

$$w_{ij} \propto A_0^2 \left| H_{\mathrm{eg}}^{(1)} \right|^2 \propto I \left| H_{\mathrm{eg}}^{(1)} \right|^2 \propto N_{\mathrm{ph}} \left| H_{\mathrm{eg}}^{(1)} \right|^2 \ . \tag{3.54}$$

By partial integration using the fact that the eigenfunctions form an othonormal set, or that the $H^{(1)}$ is Hermitian adjoint, or by the argument of the microscopic reversibility of a transitions from state $i \to j$ and from $j \to i$ induced by some perturbation $H^{(1)}$, we find that

$$\left| \int \varphi_j^* H^{(1)} \varphi_{\mathrm{i}}\, \mathrm{d}\tau \right|^2 = \left| \int \varphi_{\mathrm{i}}^* H^{(1)} \varphi_j\, \mathrm{d}\tau \right|^2 \ . \tag{3.55}$$

We see that the probabilities for induced emission and absorption are the same and that the rates differ only by factors containing the number of atoms in

the upper and lower states. This fact is the basis for the relation between the Einstein coefficients. See Problem 10 in Sect. 3.3. Spontaneous emission has to be treated in the sense mentioned above as emission stimulated by the zero field. We return to this aspect in a moment. First the interaction operator $H^{(1)}$ should be simplified to reach the so-called dipole approximation.

We note that the radius of an atom ($r \simeq 0.1$ nm) and the distance between neighboring atoms in a solid ($a \simeq 0.3$ nm) are small compared to the wavelength in the visible ($\lambda \simeq 500$ nm). Therefore there is practically no phase shift of the electromagnetic radiation over one atom or between one atom and its neighbors. Thus we can expand the term $\mathrm{e}^{\mathrm{i}\boldsymbol{kr}}$ in (3.52) in a power series and stop after the constant term

$$\mathrm{e}^{\mathrm{i}\boldsymbol{kr}} = 1 + \frac{\mathrm{i}\boldsymbol{kr}}{1!} + \frac{\mathrm{i}\boldsymbol{kr}^2}{2!} + \ldots \simeq 1 \ . \tag{3.56}$$

This is the first step towards the dipole approximation. It means, that the momentum of the photon $\hbar\boldsymbol{k}$ in (3.53b) is negligible.

The matrix element $H_{ij}^{(1)}$ still contains the momentum operator $\boldsymbol{p} = \frac{\hbar}{\mathrm{i}}\boldsymbol{\nabla}$

$$H_{ij}^{(1)} \sim \langle \varphi_\mathrm{i} \, |\boldsymbol{p}| \, \varphi_\mathrm{i} \rangle =: \langle \boldsymbol{p}_{i,j} \rangle \ .$$

With the semiclassical relation

$$\frac{\hbar}{\mathrm{i}} \nabla = p = m\dot{\boldsymbol{r}} \tag{3.57}$$

and some arguments of plausibility (See e.g. [55S1] of Chap. 2), we find that

$$\int \varphi_j^* \frac{\hbar}{\mathrm{i}} \nabla \varphi_\mathrm{i} \, \mathrm{d}\tau = m \frac{\mathrm{i}}{\hbar} \left(E_i - E_j\right) \int \varphi_j^* \boldsymbol{r} \varphi_\mathrm{i} \, \mathrm{d}\tau = m\omega \int \varphi_j^* \boldsymbol{r} \varphi_i \, \mathrm{d}\tau \, . \tag{3.58}$$

For a detailed derivation of this relation actually some knowledge of the wave functions is required. (For details see e.g. [55S1, 85G1, 92M1] of Chap. 2).

We note that in this so-called dipole approximation (3.56) to (3.59) the transition rate is given by

$$w_{ij} \sim I\omega^2 \left| \mathrm{e_A} < e\boldsymbol{r}_{ij} > \right|^2 D(E) = Iw^2 \left| H_{ij}^D \right|^2 D(E) \, . \tag{3.59}$$

This result can also be obtained in a more intuitive way if we remember that the energy of a dipole $e\boldsymbol{r}$ in an electric field $\boldsymbol{E} = \dot{\boldsymbol{A}}$ is given by

$$e\boldsymbol{r} \cdot \boldsymbol{E} = H^{(1)} \tag{3.60}$$

Using this approach in combination with (3.50) yields directly (3.59).

From now on we will call $\left|H_{ij}^D\right|^2$ the dipole-transition probability and the operator $e\boldsymbol{r}$ the dipole operator H^D.

This result can be obtained still in another way using for $\boldsymbol{A}$ a suitable gauge different from the Coulomb gauge ($\nabla \cdot \boldsymbol{A} = 0$) (see (2.49)) via $\boldsymbol{A'} = \boldsymbol{A} + \nabla\chi$

and $\Phi' = \Phi - \dot{\chi}$ where χ is a scalar field with existing second derivative. It can be shown that $\boldsymbol{A}'$ results in identical electric and magnetic fields as $\boldsymbol{A}$. The choice $\chi(\boldsymbol{r}, t) = i\omega^{-1}\boldsymbol{r}\boldsymbol{E} = i\omega^{-1}\boldsymbol{r} \cdot \boldsymbol{E}_0 \exp\left[i(\boldsymbol{k}r - \omega t)\right]$ results for $H^{(1)}$ directly in (3.60).

Transitions using higher terms in (3.44), (3.51), (3.56) correspond to quadrupole, octupole and higher order transitions.

To conclude this section, we calculate the net rate of the transitions shown in Fig. 3.16a–c.

We assume that we have a density of photons N_{ph} which populate only one mode in the sense used for the calculation of the density of states in Sect. 2.6 that is, all photons have the same wave vector $\boldsymbol{k}$, polarization $\boldsymbol{e}_{\mathrm{A}}$ and energy $\hbar\omega$.

Furthermore $\hbar\omega$ fulfills the energy conservation law according to (3.54). The density of identical two-level atoms is N_{A} where a fraction α_{g} is in the ground state and correspondingly $(1 - \alpha_{\mathrm{g}})$ are in the excited state. The net rate of the change of N_{ph} with time is then given using (3.54,59)

$$\frac{\partial N_{\mathrm{ph}}}{\partial t} = -N_{\mathrm{A}}\alpha_{\mathrm{g}}N_{\mathrm{ph}}\left|H^{D}_{\mathrm{g}\to\mathrm{e}}\right|^2 + N_{\mathrm{A}}(1-\alpha_{\mathrm{g}})(1+N_{\mathrm{ph}})\left|H^{D}_{\mathrm{g}\to\mathrm{e}}\right|^2 . \tag{3.61}$$

The first term on the r.h.s. describes the absorption of photons, the second one the spontaneous and stimulated emission in the factor $(1 + N_{\mathrm{ph}})$.

From (3.55) we see that

$$\left|H^{D}_{\mathrm{g}\to\mathrm{e}}\right|^2 = \left|H^{D}_{\mathrm{e}\to\mathrm{g}}\right|^2 = \left|H^{D}\right|^2 \tag{3.62}$$

and hence

$$\frac{1}{\left|H^{D}\right|^2}\frac{\partial N_{\mathrm{ph}}}{\partial t} = N_{\mathrm{ph}} \cdot N_{\mathrm{A}}(1-2\alpha_{\mathrm{g}}) + N_{\mathrm{A}}(1-\alpha_{\mathrm{g}}) \ . \tag{3.63}$$

The first term on the right-hand side depends linearly on N_{ph} and describes the net rate of absorption and stimulated emission. The second terms gives the spontaneous emission since it is independent of N_{ph}.

There is net absorption for $\alpha_{\mathrm{g}} > 1/2$ (absorption coefficient $\alpha(\omega) > 0$ or $\partial N_{\mathrm{ph}}/\partial t < 0$) and amplification or optical gain for $\alpha_{\mathrm{g}} < 1/2$, i.e., for gain, more than half of the atoms have to be in the upper state. This situation cannot be reached in thermal equilibrium, but only under the influence of a suitable source of pump power. Usually one or more additional energy levels are required (three- and four-level lasers). This fact can be easily elucidated with the following argument. We start with a situation where all atoms are in the ground state i.e., $\alpha_{\mathrm{g}}(t = 0) = 1$. If we send for $t > 0$ a photon field with frequency ω fulfilling (3.52a) into the system we initially have absorption since $\alpha(\omega)$ is given by

$$-\alpha(\omega) \sim N_{\mathrm{A}}(1-2\alpha_{\mathrm{g}}) \ . \tag{3.64}$$

With increasing time and pump power $\alpha(\omega)$ decreases because α_{g} decreases. For the situation $\alpha_{\mathrm{g}} = 1/2$ the absorption vanishes and the material becomes

transparent. This means no more pumping is possible under (quasi-) stationary excitation to reach $\alpha_g > 1/2$. In Chaps. 21 to 23 we will reexamine the above considerations using the proper statistics introduced in Sect. 2.6. In Chaps. 23 and 27 we shall see further that without dephasing processes or for sufficiently intense pulses shorter than the dephasing time it is possible to drive the system periodically from the ground state to the completely inverted state ($\alpha_g = 0$) and back again. This process is known as Rabi flopping.

3.3 Problems

1. Consider the interface between vacuum (or air) and glass ($n = 1.45$) at a wavelength λ_{vac} of $0.5\,\mu$m.
   Calculate for an angle of incidence $\alpha_i = 45^o$ the incident, reflected, and transmitted wave vectors, and the transmitted and reflected intensities for both polarizations.
   Calculate Brewster's angle for the transition air→glass and glass→air and the angle for the onset of total internal reflection.
2. Find a piece of polarizing material (polaroid) and observe the light reflected from a nicely polished floor or scattered from the blue sky using different orientations of the light propagation and of the polarization. Do not look into the sun! Try to explain your findings.
3. Play with a piece of clear calcite and the polaroid.
4. From Figs. 3.11 and 3.12 one can understand that a birefringent crystal can be used as a polarizer if the lateral diameter of the beams is smaller than their lateral displacement after the passage. Usually one uses slightly more complex arrangements known as Glan-Thomson or Taylor polarizers. Make yourself familiar with the way of operation of these optical components. What can happen if you use them under oblique incidence or with a di- or convergent light beam with large angle of aperture?
5. Derive the laws of reflection and refraction from the principle of Maupertins (see Fig. 3.4) and from momentum conservation (3.16). Does this law also hold if you kick a soccer ball against a wall?
6. Consider Fig. 3.9c. Explain in words why only in this diagram the three curves intersect, i.e., why zero absorption gives both the highest maxima and the lowest minima in reflectance.
7. Verify the energy conservation law (3.23a) with the help of (3.23b) and (3.18).
8. What is a quarter ($\lambda/4$) or half ($\lambda/2$) wave plate? How thick is it? What is the state of polarisation of the transmitted light beam, when the incident beam is linearly polarized? Does it depend on the orientation of the incident linear polarization with respect to the crystallographic axis. Make a simple sketch. (Generally one uses a uniaxial material and normal incidence for theses devices.) What is a low order $\lambda/4$ plate? Why are achromatic $\lambda/4$ plates rather expensive?

9. Make a simple sketch to make yourself familiar with optical activity. Can you imagine, that the absorption can be different for σ^+ and σ^- polarized light?
10. Find in a textbook the definition and meaning of Einstein's coefficients.
11. Inform yourself on time independent perturbation theory. The perturbed wave function contains in first order virtually excited states and the perturbed energy in second order. Inspect time dependent perturbation theory for virtual excitations. Derive Fermi's golden rule.
12. Verify some of the prominent features in Fig. 3.6 with (3.18), e.g., normal and grazing incidence, α_{B} or α_{tot}.

References to Chap. 3

[75C1] R. Claus, L. Merten, J. Brandmüller: Light Scattering by Phonon Polaritons, Springer Tracts, Mod. Phys. Vol. **75**, Springer, Berlin, Heidelberg (1975)
[76P1] R.W. Pohl: Optik und Atomphysik, 13th edn., Springer, Berlin, Heidelberg (1976)
[77B1] M. Born, E. Wolf: Principles of Optics, Pergamon, Oxford (1977)
[83K1] W. Kleber: Einführung in die Kristallographie, 16th edn., VEB Verlag Technik, Berlin (1983) or more recent editions
[84A1] V.M. Agranovich, V.L. Ginzburg: Crystal Optics with Spatial Dispersion and Excitons, Springer Ser. Solid-State Sci. **42**, 2nd ed., Springer, Berlin, Heidelberg (1984)
[85C1] D. Craig: et al. IEEE J. QE-**21**, 1363 (1985)
[87H1] E. Hecht: Optics, 2nd edn., Addison-Wesley, Reading (1987)
[89T1] A. Thelen: Design of Optical Interference Coatings, McGraw Hill, New York (1989)
[96T1] E. Tsitsishvili, Applied Physics A **62**, 255 (1996) and Solid State Commun. **100**, 541 (1996)
[97L1] S.G. Lipson, H.S. Lipson, D.S. Tannhauser: Optik, Springer, Berlin (1997)
[02T1] E. Tisitishvili in Ref [81A1]k of Chap. 1, p 357 (2002)

4

Ensemble of Uncoupled Oscillators

The optical properties of matter are determined by the coupling of various types of oscillators in matter to the electromagnetic radiation field. In other words, an incident electromagnetic field will cause these oscillators to perform driven or forced oscillations. The amplitude of these driven oscillations depends on the angular frequency ω of the incident field, on the eigenfrequency ω_0 of the oscillators, on the coupling strength f between electromagnetic field and oscillator, and on its damping γ. In semiconductors these oscillators or resonances include optical phonons, excitons including their ionisation continuum and higher band-to-band transitions or plasmons. They will be explained in some detail in Chaps. 7–10. We can anticipate that many basic features of the optical properties of these resonances are similar. Therefore it is reasonable to discuss first, in a general way, the optical properties of an ensemble of model oscillators. By using the results of Chaps. 4–6 we shall obtain in Chaps. 11–15 a quite simple and straightforward access to the optical properties of semiconductors.

It turns out that a treatment of the optical properties of an ensemble of model oscillators in terms of classical mechanics and electrodynamics yields results which are, in many respects, very close to reality see Chap. 27. This is especially true for the spectra of the complex dielectric function or refractive index, or of the spectra of reflection and transmission. All four are closely connected (Chap. 6). We shall therefore follow this classical approach for some while, and explain at the appropriate places what modifications appear if quantum mechanics is applied.

These model oscillators are known as Lorentz oscillators. A treatment of these Lorentz oscillators, or if a finite electrical conductivity $\sigma(\omega)$ is included [see, e.g., (2.24)] of the Drude–Lorentz model, is found in many textbooks on optics and solid state physics like those mentioned in Chaps. 1 and 2. See also [63H1, 72W1].

As mentioned in the preface to this second edition, there is presently a trend to describe the optical properties, especially those of the electronic system of semiconductors, by the optical or the semiconductor Bloch equa-

tions. In this approach, there is a great risk that the physics is concealed behind a rather complex mathematical formalism or it disappears partly in numerical solutions. We start therefore with the simple and intuitively clear concept of the Lorentz oscillators, but present the other concept in Chap. 27 so that the reader can familiarize himself with it.

We will now consider the optical properties of an ensemble of oscillators. We begin with the simplest case of uncoupled oscillators and refine the concept in various steps up to Chap. 8.

4.1 Equations of Motion and the Dielectric Function

We assume that we have an ensemble of identical uncoupled harmonic oscillators. For simplicity we choose a periodic one-dimensional array in the direction of light propagation with a lattice-constant a as shown in Fig. 4.1a. These harmonic oscillators all have the same eigenfrequency ω_0'. If we neglect damping for the moment, then ω_0' can be expressed in a mechanical model by the mass m and the force constant β of the springs as

$$\omega_0'^2 = \beta m^{-1}. \tag{4.1}$$

If we elongate the oscillators in phase (Fig. 4.1b) the whole ensemble oscillates with ω_0' and the same will be true if we excite neighboring oscillators in antiphase (Fig. 4.1c). The first case corresponds to

$$\lambda = \infty \quad \text{or} \quad k = 0, \tag{4.2a}$$

and the second one to

$$\lambda_{\min} = 2a \quad \text{or} \quad k_{\max} = \frac{\pi}{a}. \tag{4.2b}$$

The point $k = 0$ is usually called the Γ-point in $\boldsymbol{k}$-space and the condition (4.2b) gives the boundary of the first Brillouin zone for a simple linear chain or cubic lattice. The values given by (4.2b) are the shortest physically meaningful wavelength and the largest $\boldsymbol{k}$-vector in our system. The eigenfrequency will also be ω_0' for all λ or $\boldsymbol{k}$-values in between and so we get the horizontal dispersion relation for our system in Fig. 4.2a, i.e., the width of the band of eigenfrequencies is zero (so-called flat band). A wave packet created in our system by elongating only one oscillator (Fig. 4.1d) will not propagate, since there is no coupling to the neighboring oscillators. In agreement we find from (2.15) that the group velocity $\partial\omega/\partial k$ is zero.

What we have seen here is an example of the general rule that in an ensemble of identical oscillators [or of atoms or of other (quasi-) particles] a finite coupling between neighbors results in a finite bandwidth and a non-vanishing group velocity while vanishing coupling results in vanishing bandwidth and group velocity

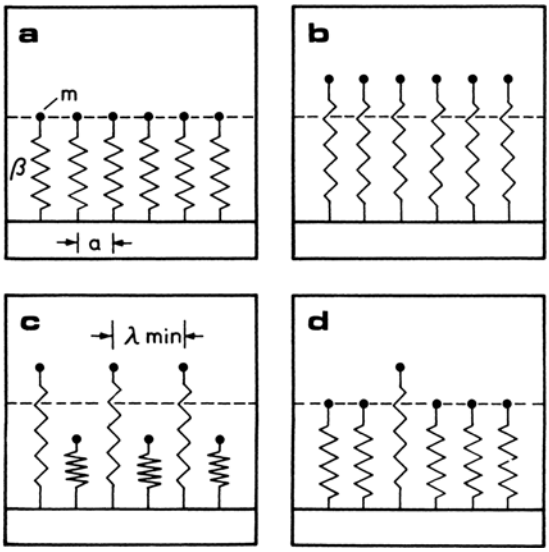


Fig. 4.1. A part of a periodic array of uncoupled oscillators in their equilibrium position (**a**); elongated with a wavelength $\lambda \Rightarrow \infty$ (**b**); with the shortest physically meaningful wavelength $\lambda_{\min}$ (**c**), and a non-propagating wave packet (**d**). After [63H1]

$$\begin{aligned} &\text{zero coupling} \longleftrightarrow \text{zero bandwidth} \longleftrightarrow v_{\mathrm{g}} = 0, \quad \text{Fig.4.2a} \\ &\text{finite coupling} \longleftrightarrow \text{finite bandwidth} \longleftrightarrow v_{\mathrm{g}} \neq 0. \quad \text{Fig.4.2b} \end{aligned} \tag{4.3}$$

We will discuss the implications of relaxing the assumption "uncoupled" later on, proceeding to the more realistic assumption of coupled oscillators in Sect. 5.4.

In the next step we couple the independent oscillators to the electric field of the electromagnetic radiation given by

$$\boldsymbol{E} = (E_0, 0, 0) \exp\left[\mathrm{i}(k_z z - \omega t)\right]. \tag{4.4a}$$

This means that the light wave propagates in the z-direction and is polarized in the x-direction, parallel to the elongation of the oscillators. By considering the oscillator at $z = 0$ or making use of the dipole approximation of (3.56−60) (i.e., $a \ll \lambda$) we can drop the spatial variation in (4.4a) to obtain

$$\boldsymbol{E} = (E_0, 0, 0)\, \mathrm{e}^{-\mathrm{i}\omega t}. \tag{4.4b}$$

For the coupling, we assume that the mass m of every oscillator carries a charge e. For neutrality reasons, we need then a charge $-e$ fixed at the equilibrium position of every mass. $\boldsymbol{E}$ will then exert a force on the oscillator, and the elongation x is connected with an electric dipole moment via

$$p_x = ex. \tag{4.5}$$

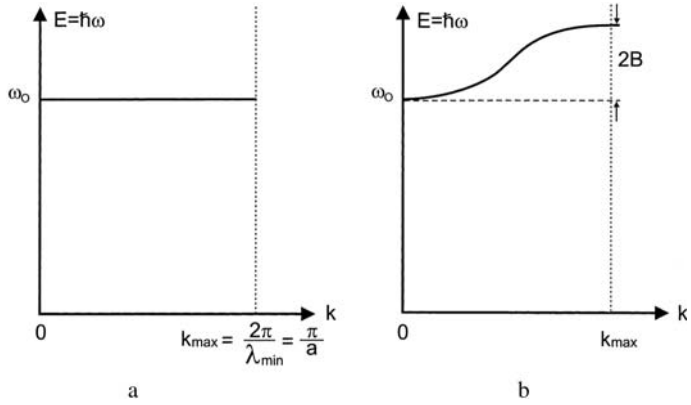


Fig. 4.2. The dispersion relation of an ensemble of uncoupled (**a**) and of coupled (**b**) oscillators

This or similar models and/or the resulting equations for $\varepsilon(\omega)$, go by various names such as Lorentz or Selmaier oscillator, Helmholtz–Ketteler or Kramers–Heisenberg formula.

The equation of motion is given by

$$m\ddot{x} + \gamma m\dot{x} + \beta x = eE_0\,\mathrm{e}^{-\mathrm{i}\omega t}, \tag{4.6}$$

where we have included a phenomenological damping constant γ. For a more elaborate theory e.g. in the frame of the fluctuation-dissipation theorem or for the aspect of Markov-damping used here compared to non-Markovian damping see e.g. [67K1, 85C1, 01W1, 02S1] or to some extend Sect. 23.2 and the references given there. Equation (4.6) is a linear, inhomogeneous differential equation of second order. The general solution is a sum of the general solution of the corresponding homogeneous equation and of a special solution of the inhomogeneous one. Correspondingly we use the following ansatz:

$$x(t) = x_0 \exp\left[-\mathrm{i}\left(\omega_0'^2 - \gamma^2/4\right)^{1/2} t\right] \exp\left(-t\gamma/2\right) + x_\mathrm{p}\,\mathrm{e}^{-\mathrm{i}\omega t}\,. \tag{4.7}$$

The first term on the right-hand side is the solution of the homogeneous equation and describes a transient feature. For

$$\omega_0'^2 - \gamma^2/4 > 0\,, \tag{4.8}$$

one finds a damped oscillation with a damping-dependent eigenfrequency $\left(\omega_0'^2 - \gamma^2/4\right)^{1/2}$. The inequality (4.8) defines the regime of weak damping. For stronger damping one gets essentially an exponentially decaying term.

This transient feature disappears in any case after $t > \gamma^{-1}$. It is thus of importance for (ultra-) fast, time-resolved spectroscopy treated in Chap. 23. For the stationary, linear optics regime, in which we are presently interested,

we may safely omit this term. What is then left is a forced oscillation with amplitude x_p. Inserting the ansatz (4.7) into (4.6) we find

$$x_\mathrm{p} = \frac{eE_0}{m}\left(\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma\right)^{-1}\,, \tag{4.9}$$

i.e., the usual resonance term.

This oscillation is connected with a dipole moment at every oscillator of

$$p_x = ex_\mathrm{p} \tag{4.10}$$

and a polarizability $\hat{a}(\omega)$. The hat on $\hat{a}$ and the prime on ω_0' indicate that we shall introduce two corrections here and in Sect. 4.2

$$\hat{a}(\omega) = \frac{ex_\mathrm{p}}{E_0} = \frac{e^2}{m}\left(\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma\right)^{-1}\,. \tag{4.11}$$

If we use a three-dimensional array of oscillators with density N, the result is the same and we shall get a preliminary polarization density $\boldsymbol{P}$

$$\boldsymbol{P} = N\hat{a}\boldsymbol{E}_0 = \frac{Ne^2}{m}\left(\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma\right)^{-1}\boldsymbol{E}_0\,. \tag{4.12}$$

This means that we describe the ensemble of oscillators as an effective medium, an approach which is well justified for $a \ll \lambda$. See (4.4b) or Ref. [97W1] of Chap. 1 and [97S1] of Chap. 5.

From (4.12) and (2.27) we get the following expressions for the dielectric displacement $\boldsymbol{D}$, the dielectric function $\varepsilon(\omega)$, and the susceptibility $\chi(\omega)$.

$$\boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P} = \varepsilon_0\left[1 + \frac{Ne^2}{\varepsilon_0 m}\left(\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma\right)^{-1}\right]\boldsymbol{E} \tag{4.13}$$

and

$$\varepsilon(\omega) = 1 + \frac{Ne^2}{\varepsilon_0 m}\left(\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma\right)^{-1} = \chi(\omega) + 1\,. \tag{4.14}$$

where N is the number of oscillators per unit volume. Now we want to address two corrections to the above set of equations in order to obtain the final expression for $\varepsilon(\omega)$.

4.2 Corrections Due to Quantum Mechanics and Local Fields

The term Ne^2m^{-1}, e.g., in (4.12) gives the coupling strength of the electromagnetic field to the oscillators in our mechanical model.

In quantum mechanics, this coupling is given by the transition matrix element squared. For dipole-allowed transitions this reads, as mentioned already in Sect. 3.2.2, Eqs. (3.58)–(3.60):

$$\left|H_{ij}^D\right|^2 = \left|\langle j\left|H^D\right|i\rangle\right|^2\,. \tag{4.15}$$

where i and j stand for initial and final state and H^D for the dipole operator $e\boldsymbol{r}$ or ex. For dipole-forbidden transitions, magnetic dipole or electric quadrupole matrix elements can become relevant. They are usually orders of magnitude smaller. There are different conventions for introducing the transition matrix element into the dielectric function:

Some authors define a dimensionless quantity $\hat{f} = \frac{2m\omega_0'}{\hbar e}\left|H_{ij}^D\right|^2$, call it the oscillator strength, and use it to multiply the term $Ne^2m^{-1}\varepsilon_0^{-1}$. Others call the whole numerator oscillator strength f; i.e.,

$$f = \frac{2N\omega_0'}{\varepsilon_0\hbar}\left|H_{ij}^D\right|^2 \tag{4.16}$$

We follow the second way for simplicity of notation, but stress that there is no physical difference between the methods. We obtain

$$\varepsilon(\omega) = 1 + \frac{f}{\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma}\,. \tag{4.17}$$

For the next correction to (4.17) we have to consider what is the electric field $\boldsymbol{E}$ [e.g., in (4.6)] acting on the oscillators. For dilute systems, $\boldsymbol{E}$ is just the external incident field and we can use (4.17) as it is. For dense systems, i.e., systems with a high density N of oscillators, the local field E^{loc} acting on the oscillators consists of two parts, namely the external field and the field created by all the other dipoles. Taking into account this effect leads for cubic or amorphous materials to the so-called Clausius–Mosotti or Lorenz–Lorentz equation, which relates the polarizability $\hat{a}$ to $\varepsilon(\omega)$ through

$$\frac{\varepsilon(\omega)-1}{\varepsilon(\omega)+2} = \frac{N\hat{a}(\omega)}{3\varepsilon_0} = \frac{\frac{1}{3}N\frac{e^2}{m\varepsilon_0}}{\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma}\,. \tag{4.18a}$$

See also [02B1]. Obviously (4.18a) recovers the form (4.17) for dilute systems (small N) for which $N\hat{a}$ is small and $\varepsilon(\omega)$ deviates only a little from unity. If this approximation is not valid, $\varepsilon(\omega)$ can be rewritten for small damping in the form (4.17) but with a shifted eigenfrequency

$$\omega_0^2 = \omega_0'^2 - \frac{Ne^2}{3m\varepsilon_0} = \omega_0'^2 - f/3\,. \tag{4.18b}$$

resulting in

$$\varepsilon(\omega) = 1 + \frac{f}{\omega_0^2 - \omega^2 - \mathrm{i}\omega\gamma}\,. \tag{4.19}$$

This formula now incorporates local field effects and the quantum mechanical transition probabilities. The new eigenfrequency ω_0 is the only physically relevant and experimentally accessible one. Therefore only this quantity appears in experiments. A similar procedure is also valid for crystal symmetries other than the cubic one, but in these cases the tensor character of $\varepsilon(\omega)$ comes into play (see Sect. 3.1.7), resulting in parameters f, ω_0 and γ depending on the orientation of the electric field relative to the crystal axes.

4.3 Spectra of the Dielectric Function and of the Complex Index of Refraction

As we shall see later, a semiconductor contains not only one type of oscillators and one resonance frequency ω_0, but many of them – like phonons, excitons etc. In linear optics, i.e. in linear response theory, we can just to sum over all resonances leading to

$$\varepsilon(\omega) = 1 + \sum_j \frac{f_j}{\omega_{0j}^2 - \omega^2 - \mathrm{i}\omega\gamma_j} \,. \tag{4.20a}$$

This is essentially the so-called Helmholtz–Ketteler formula or Kramers–Heisenberg dielectric function.

If the eigenfrequencies ω_0 form a continuous band, like in the ionisation continuum of the excitons or in an inhomogeneously broadened system where every oscillator has a slightly different frequency ω_0, then the Σ in (4.20a) is better replaced by an integral over these contributions

$$\int \frac{f(\omega_0)}{\omega_0 - \omega^2 - \mathrm{i}\omega\gamma(\omega_0)} \,\mathrm{d}\omega_0 \tag{4.20b}$$

where $f(\omega_0)$ is a oscillator strength per frequency interval.

We keep this in mind, but for simplicity of writing use the form (4.20a) in the following text.

We now want to discuss the contribution of an isolated resonance at $\omega_{0j'}$ in (4.20a) which allows some simplification. For closely spaced resonances the whole formula (4.20a) has to be used. We discuss this situation seperately in Sect. 4.5. For the present purposes a single resonance well seperated from all other resonances, we note that the contribution of a single resonance at $\omega_{0j'}$, is constant for $\omega \ll \omega_{0j'}$ with a contribution to $\varepsilon(\omega)$

$$\frac{f_{j'}}{\omega_{0j'}^2} = \text{const} \tag{4.21}$$

and tends to zero for $\omega \gg \omega_{0j'}$.

This means that in the vicinity of a resonance $\omega_{0j'}$ we can neglect the contributions from all lower resonances $\omega_{0j} \ll \omega_{0j'}$ and the constant contributions of all higher resonances $\omega_{0j} \gg \omega_{0j'}$ can be summarized in a so-called background dielectric constant ε_b. Obviously ε_b is unity for the highest resonance in the system, i.e. in the sum of (4.20a). For our purposes this highest resonance would be in the X-ray regime connected e.g. with the K absorption edge. So we finally get the simplified expression in the spectral surroundings of $\omega_{0j'}$

$$\varepsilon(\omega) = \varepsilon_\mathrm{b} + \frac{f_{j'}}{\omega_{0j'}^2 - \omega^2 - \mathrm{i}\omega\gamma_{j'}} \tag{4.22a}$$

In the following we drop the index j'. Equation (4.22a) can be separated into real and imaginary parts

$$\varepsilon(\omega) = \varepsilon_\mathrm{b}\left(1 + \frac{f(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2} + \mathrm{i}\frac{\omega\gamma f}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2}\right)$$
$$= \varepsilon_1(\omega) + \mathrm{i}\varepsilon_2(\omega)\,. \tag{4.22b}$$

For vanishing damping $\gamma \Rightarrow 0$ and in the vicinity of ω_0 (4.22a) can be simplified by noting $(\omega_0^2 - \omega^2) = (\omega_0 + \omega)(\omega_0 - \omega) \approx 2\omega_0(\omega_0 - \omega)$ to

$$\varepsilon(\omega) = \varepsilon_\mathrm{b} + \frac{f}{2\omega_0}\frac{1}{\omega_0 - \omega} \tag{4.22c}$$

In Fig. 4.3 we show the real and imaginary parts of $\varepsilon(\omega)$. For negligible damping $\gamma \to 0$ we find a pole in $\mathrm{Re}\{\varepsilon(\omega)\}$, and $\mathrm{Im}\{\varepsilon(\omega)\}$ converges to a δ function at ω_0. Finite damping results in a broadening of $\mathrm{Im}\{\varepsilon(\omega)\}$ to the Lorentzian lineshape of (4.22b) and a smooth connection of the two branches of $\mathrm{Re}\{\varepsilon(\omega)\}$.

We concentrate now on the case of small damping. One of the two frequencies of special interest for $\varepsilon(\omega)$ is the eigenfrequency ω_0, which is connected with the singularity. The other one corresponds to the point at which $\mathrm{Re}\{\varepsilon(\omega)\}$ crosses zero. Going back to (2.17) and (2.43) we find

$$\nabla \cdot \boldsymbol{D} = \varepsilon_0 \nabla \cdot \varepsilon(\omega) \cdot \boldsymbol{E} = 0\,. \tag{4.23}$$

As already discussed, this equation is usually used to argue that electromagnetic waves are transverse, since $\nabla \cdot \boldsymbol{E} = 0$ is zero for this case. The other solution $\varepsilon(\omega = \omega_\mathrm{L}) = 0$ gives the frequency of a longitudinal mode. This mode is a pure polarization mode with $\varepsilon_0 \boldsymbol{E} = -\boldsymbol{P}$, i.e., antiparallel polarization and electric field (see Sect. 2.4).

For $\gamma \Rightarrow 0$ we find

$$\varepsilon(\omega = \omega_\mathrm{T} \equiv \omega_0) = \infty,$$
$$\varepsilon(\omega = \omega_\mathrm{L}) = 0. \tag{4.24}$$

The relation between these two frequencies is given by

$$\omega_\mathrm{L}^2 - \omega_\mathrm{T}^2 = f/\varepsilon_\mathrm{b} \sim \left|H_{ij}^D\right|^2, \tag{4.25}$$

and we can make the following statements. A finite oscillator strength f is necessarily connected with a finite longitudinal-transverse splitting energy Δ_LT and vice versa.

$$\left|H_{ij}^D\right|^2 \neq 0 \Leftrightarrow f \neq 0 \Leftrightarrow \Delta_\mathrm{LT} = \hbar(\omega_\mathrm{L} - \omega_\mathrm{T}) \neq 0\,. \tag{4.26}$$

The physical reason for this energy splitting is that the longitudinal polarisation wave produces, in contrast to the transverse wave a longitudinal

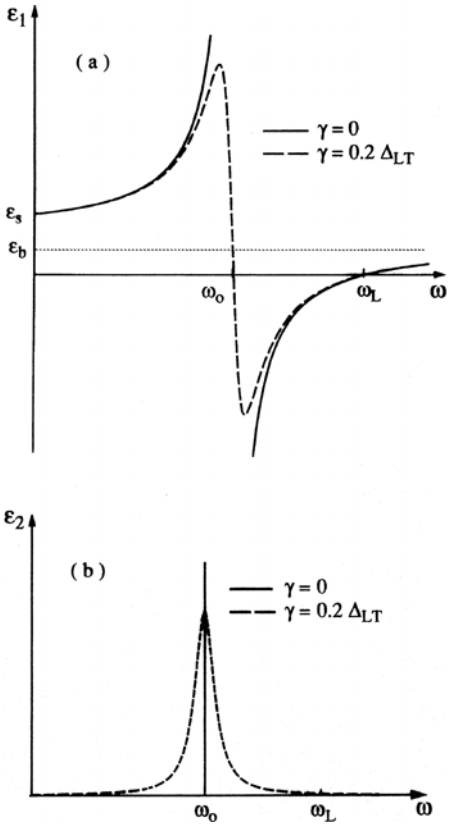


Fig. 4.3. The real (**a**) and imaginary (**b**) parts of the dielectric function for zero and finite damping

electric field (2.45), which acts as an additional restoring force and increases the longitudinal eigenfrequency ω_{L} above the transverse eigenfrequency. This is illustrated in Fig. 4.4a for a slab of matter with finite thickness. In the case of Fig. 4.4b no additional restoring force results from the surface polarization charges in contrast to Fig. 4.4a. The polarisation charges in Fig. 4.4b gives rather rise to a surface mode, to which we come back to in Sect. 5.6. The spreading of the field lines of this surface mode on both sides of the interface explain qualitatively the fact that the frequency of the surface mode is situated between ω_0 and ω_{L}. See Sects. 2.4 and 5.6.

If we call the roughly constant value of $\varepsilon(\omega)$ well below ω_0 the static dielectric constant ε_{s} and the value above ε_{b}, as already mentioned, we find

$$\varepsilon_{\mathrm{s}} = \varepsilon_{\mathrm{b}} + f/\omega_0 \tag{4.27}$$

and the Lyddane-Sachs-Teller relation

$$\frac{\varepsilon_{\mathrm{s}}}{\varepsilon_{\mathrm{b}}} = \frac{\omega_{\mathrm{L}}^2}{\omega_0^2} > 1 \quad \text{for} \quad f > 0. \tag{4.28}$$

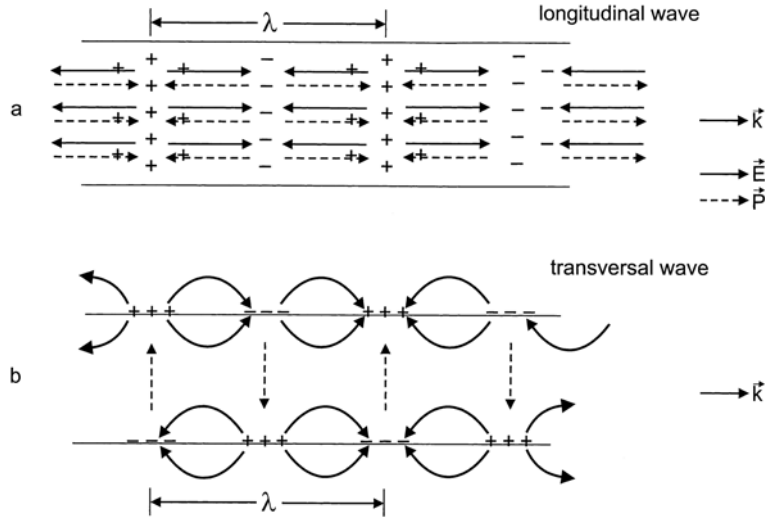


Fig. 4.4. The polarisation charges and the resulting electric and polarisation fields (*solid* and *dashed arrows*, respectively) in a slab of matter for a longitudinal (**a**) and a transversal wave (**b**)

Furthermore we can make the slightly tricky statement that the background dielectric constant ε_b of a resonance ω_{0j} is simultaneously the static dielectric constant for the next higher resonance ω_{0j+1}.

Finally we can define the term "small damping" already used several times above by

$$\gamma < \hbar^{-1}\Delta_\mathrm{LT} = \omega_\mathrm{L} - \omega_\mathrm{T} \tag{4.29}$$

and frequencies ω far above or below ω_0 means

$$\hbar\,|\omega - \omega_0| \gg \Delta_\mathrm{LT}. \tag{4.30}$$

Now we concentrate on the complex index of refraction $\tilde{n}(\omega) = n(\omega) + \mathrm{i}\kappa(\omega)$, which is connected to the dielectric function via (2.35)

$$\tilde{n}(\omega) = \varepsilon^{1/2}(\omega) \tag{4.31}$$

and consequently

$$\begin{aligned}\varepsilon_1(\omega) &= n^2(\omega) - \kappa^2(\omega)\\ \varepsilon_2(\omega) &= 2n(\omega)\kappa(\omega)\end{aligned}$$

or

$$\begin{aligned}n(\omega) &= \left(\frac{1}{2}\left\{\varepsilon_1(\omega) + \left[\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega)\right]^{1/2}\right\}\right)^{1/2}\\ \kappa(\omega) &= \left(\frac{1}{2}\left\{-\varepsilon_1(\omega) + \left[\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega)\right]^{1/2}\right\}\right)^{1/2}.\end{aligned} \tag{4.32}$$

In Fig. 4.5 we show the real and imaginary parts of $\tilde{n}$ for vanishing and for finite γ.

When approaching the resonsance from the low-frequency side, one sees that both $\mathrm{Re}\{\tilde{n}\} = n$ and thus also the real part of $k = \frac{\omega}{c} n(\omega)$ increase drastically. Between ω_0 and ω_L we find

$$\omega_0 \leq \omega \leq \omega_\mathrm{L} \begin{cases} n = 0 & \text{for} \quad \gamma = 0 \\ n \ll 1 & \text{for} \quad \gamma \neq 0 \\ \kappa > n \end{cases} \tag{4.33}$$

The imaginary part of $\tilde{n}(\omega)$; i.e. $\kappa(\omega)$ starts with a singularity at $\omega = \omega_0$ for $\gamma = 0$ and then drops to small values for ω approaching ω_L.

This means that for $\gamma = 0$ there is no propagating, i.e., no spatially oscillating or wave-like solution in the medium for the spectral region addressed in (4.33). Instead we have only a spatially exponentially decaying amplitude similar to the type known for total reflection in the medium with the lower index of refraction. For finite γ we get a small real part of $\tilde{n}$ which means that some light can penetrate into the medium, but this light is damped over

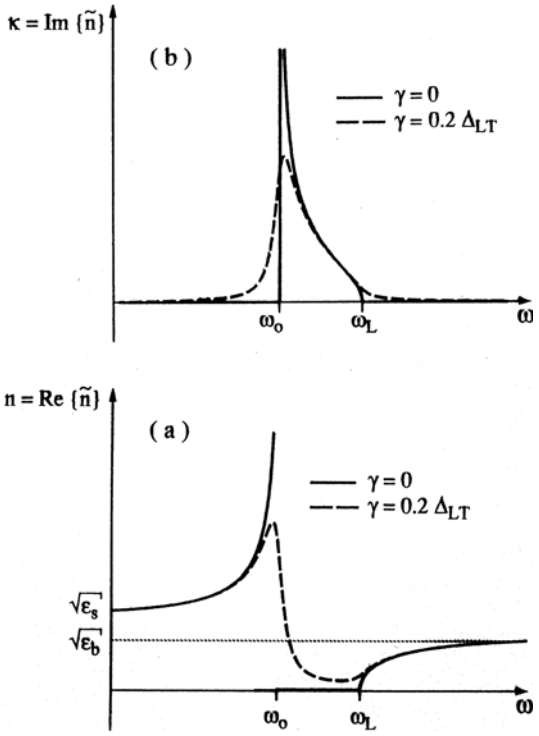


Fig. 4.5. Real (**a**) and imaginary (**b**) parts of the complex index of refraction for zero and finite damping

a distance shorter than the wavelength in the medium since we have $\kappa > n$ i.e., "strong absorption" according to (3.17).

If all oscillators have the same eigenfrequency and the broadening in $\varepsilon_2(\omega)$ results only from a finite damping γ, the system is said to be homogenously broadened and $\varepsilon_2(\omega)$ shows the usual Lorentzian shape. If the system contains many oscillators with slightly different frequencies ω_0, the system is said to be inhomogeneously broadened. The shape of $\varepsilon_1(\omega)$ and $\varepsilon_2(\omega)$ is then determined by the distribution function of eigenfrequencies and/or of oscillator strengths in (4.20b). The width of $\varepsilon_2(\omega)$ is then a convolution of a Lorentzian $L(\omega)$ for every single oscillator and of the distribution. In the case of a Gaussian $G(\omega)$ this convolution leads to a Voigt lineshape in the sense of (4.34)

$$V(\omega) = \int_{-\infty}^{+\infty} L(\omega - \omega')G(\omega')\,\mathrm{d}\omega' = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} \frac{\Gamma}{\sigma} \frac{\exp\left\{-\frac{(\omega'-\omega_0)^2}{2\sigma^2}\right\}}{(\omega-\omega')^2 + \left(\frac{\Gamma}{2}\right)^2}\,\mathrm{d}\omega' \tag{4.34}$$

where σ and Γ give the inhomogeneous and homogenous broadening, respectively.

It should also be noted that even for homogeneously broadened resonances $\kappa(\omega)$ and $\alpha(\omega)$ do not exhibit intrinsically a Lorentzian lineshape in contrast to $\varepsilon_2(\omega)$. See, e.g., Fig. 4.4b or (4.32). However for small oscillator strength and/or concentration of the oscillators and finite damping, $\alpha(\omega)$ can recover a Lorentz-like lineshape. See, e.g., [76P1] of Chap. 3.

4.4 The Spectra of Reflection and Transmission

With a knowledge of $\tilde{n}(\omega)$ we can now discuss the spectra of reflection and transmission. We start with the reflectivity $R(\omega)$ of a single interface between vacuum (or air) and the medium. We discuss only the situation of normal incidence [see Sect. 3.1.4 and (3.19)].

$$R(\omega) = \frac{I_\mathrm{r}}{I_\mathrm{i}} = \frac{[n(\omega)-1]^2 + \kappa^2(\omega)}{[n(\omega)+1]^2 + \kappa^2(\omega)} \tag{4.35a}$$

Oblique incidence is described by the Fresnel formula (3.18).

$R(\omega)$ is plotted in Fig. 4.6. First we discuss the situation for $\gamma = 0$. When approaching the resonance from the low frequency side, R starts with an almost constant value determined from (3.19b) and (4.27), (4.28):

$$R = \frac{\left(\varepsilon_\mathrm{s}^{1/2} - 1\right)^2}{\left(\varepsilon_\mathrm{s}^{1/2} + 1\right)^2} \quad \text{for} \quad \omega_0 - \omega \gg \frac{1}{\hbar}\Delta_\mathrm{LT}\,. \tag{4.35b}$$

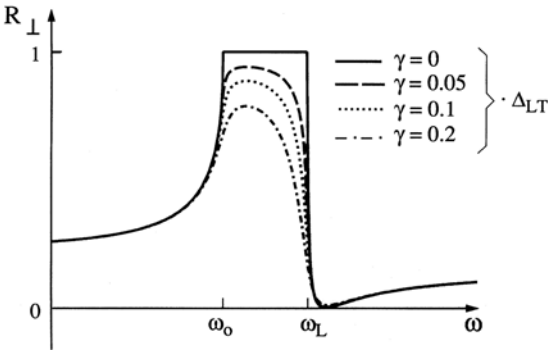


Fig. 4.6. The reflection spectrum of a single resonance with zero and finite damping for normal incidence

Above ω_{L}, the reflectivity converges towards a lower constant value

$$R = \frac{\left(\varepsilon_{\mathrm{b}}^{1/2} - 1\right)^2}{\left(\varepsilon_{\mathrm{b}}^{1/2} + 1\right)^2} \quad \text{for} \quad \omega - \omega_0 \gg \hbar^{-1}\Delta_{\mathrm{LT}} \; . \tag{4.35c}$$

For vanishing damping, the reflectivity R increases just below ω_0 and reaches the value $R = 1$ for $\omega = \omega_0$. Between ω_0 and ω_{L} we have $R = 1$. For $\omega > \omega_{\mathrm{L}}$ R drops rapidly and reaches zero at the frequency where $n(\omega) = 1$. Then R increases again towards a constant value given by (4.35c). We see now that all light is reflected in the region where we have no propagating mode in the medium in agreement with the discussion of (4.33). Such a band is also called stop-band because the light is "stopped" and sent back.

An alternative approach to understanding the stop-band is via total internal reflection. See (3.18) or Fig. 3.6b. For vacuum we have $n \equiv 1$. Between ω_0 and ω_{L} the value of n is below unity. Consequently, matter is optically "thinner" than a vacuum. In such a situation, the total "internal" reflection occurs for angles of incidence exceeding α_c.

In general, α_c is defined such that $\sin\alpha_c$ cannot be larger than unity resulting in (3.10c) in the general case of

$$\sin\alpha_c = \frac{n_{\mathrm{II}}}{n_{\mathrm{I}}} \quad \text{for} \quad n_{\mathrm{II}} < n_{\mathrm{I}} \tag{4.36}$$

where $n_{\mathrm{I}} \equiv 1$. For $n_{\mathrm{II}} \to 0$ we find $\alpha_c \to 0$ resulting in total internal reflection already for normal incidence.

For finite γ the reflection spectrum is smoothed out due to the finite values of $n(\omega)$ for $\omega_0 \le \omega \le \omega_{\mathrm{L}}$ and of $\kappa(\omega)$ for $\omega \le \omega_0$ and for $\omega_{\mathrm{L}} \le \omega$. It should be noted, that the reflection minimum occurs also for weak damping only slightly above ω_{L}, but the reflection maximum is no longer related to ω_0. We shall see in Sect. 5.4 that this is even true for $\gamma = 0$ in the case of resonances with spatial dispersion.

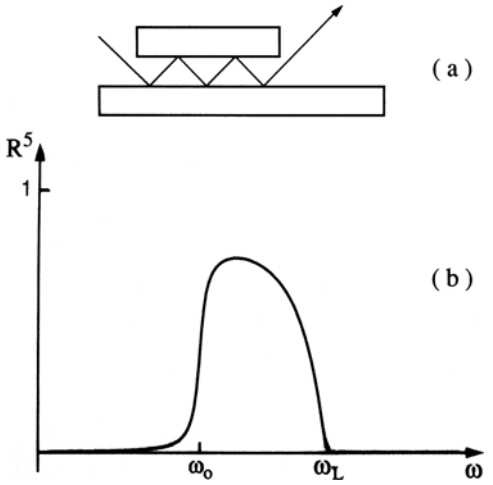


Fig. 4.7. An arrangement for multiple reflection from a medium (**a**) and the resulting "Reststrahlbande" (**b**)

The spectral region between ω_0 and ω_L is often called "Reststrahlbande"[1] for the following reason. If we send a light beam with a broad, essentially flat, spectral distribution of the intensity around ω_0 onto a sample and allow it to be reflected several times, e.g., in the configuration given in Fig. 4.7, then significant intensity will remain as a "rest" only in the region $\omega_0 \leq \omega \leq \omega_\mathrm{L}$. It should be noted that we do not assume any coherent superposition of the beams, i.e., the Reststrahlbande comes only from $\varepsilon(\omega)$ and has nothing to do with Fabry–Perot resonators (Sect. 3.1.6) or a Lummer–Gehrke plate. See [76P1] of Chap. 3 for this device.

Up to now we have discussed the reflectivity of a single interface. In experiments it is usually easier to handle semiconductor plates with two surfaces and a geometrical thickness d rather than samples filling a semi-infinite half-space. To handle this problem, we use the results of Sect. 3.1.6.

In the vicinity of the resonance the reflection spectrum will remain the same as in Fig. 4.6 since multiple reflection is suppressed due to the large value of κ and thus of the absorption coefficient (2.40)

$$\alpha(\omega) = 2\,\frac{\omega}{c}\,\kappa(\omega). \tag{4.37}$$

Away from the resonance we have to consider for the total transmission $\hat{T}$ contributions due to multiple reflections at the two surfaces. There are two cases, as already discussed in Sect. 3.1.6; if d is longer than the coherence length of the light source, or if the two surfaces of the slab are not exactly plane-parallel (e.g., due to steps on the surfaces or due to a small angle between them) we have to add intensities. This results in a total transmission $\hat{T}$ and reflection $\hat{R}$ of the slab given by (3.29).

[1] This term has been introduced in 1987 by H. Rubens and is used also in the english literature as a synonym for stop-band.

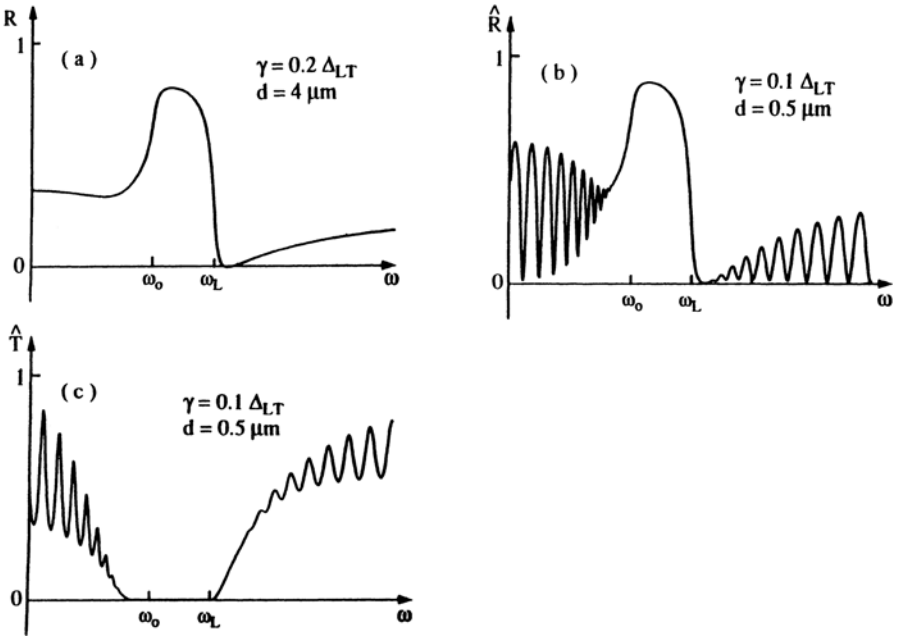


Fig. 4.8. The reflectivity $\hat{R}$ of a thin slab of matter in the vicinity of a resonance for the cases where intensities (**a**) or field amplitudes (**b**) have to be added and the transmission (**b**) for the case where amplitudes add

In Fig. 4.8a we show $\hat{R}$ for this situation. We start with a reflectivity given by the combined contributions of front and rear side. Approaching the resonance, $\hat{R}$ drops first slightly because the absorption reduces the influence of the rear side. The reflection around the resonance remains the same as above in Fig. 4.6. If the sample becomes transparent above the resonance, we again find a contribution from the rear side to the reflectivity.

In Figs. 4.8b and c we give an example for the $\hat{T}$ and $\hat{R}$ spectra for the second case when the sample forms a FP resonator using the set of equations (3.31),(3.32). We see that the finesse F is reduced close to the resonance due to increasing absorption. The spectral spacing between the resonances decreases as the resonance is approached from lower frequencies, due to the steep increase of $n(\omega)$ or of $\boldsymbol{k}$. Above resonance, the FP mode structure reappeares.

Every directional dependence of the quantities ω_0, f and γ which describe a resonance will lead to a directional dependence of $n(\omega)$ and $\kappa(\omega)$, i.e., to birefringence and dichroism, discussed in Sect. 3.1.7. In Chap. 5 we shall see that a dependence of the dispersion $\omega_0(\boldsymbol{k})$ can lead to a weak birefringence, and indeed this effect is observed. For a cubic crystal with point group T_d (see Chap. 26) one finds that light propagating with $\boldsymbol{k}$ parallel to the [110] direction may display slightly different optical properties when the electric field is polarized in the directions [110] and [001].

We will not consider the spectra of luminescence in this section since they require more detailed knowledge of the excitations in the semiconductor and especially of the radiative and nonradiative decay channels. Examples will be given in Chaps. 14 and 15.

4.5 Interaction of Close Lying Resonances

Until now we have considered a single resonance, well separated from others. Now we discuss what happens for two close-lying resonances. Such a situation occurs, for instance, for the A and B Γ_5 excitons in ZnO. See [81K1, 93F1, 02B1] and references therein. To this end we consider two resonances with equal oscillator strength f and slightly different eigenfrequencies $\omega_{\mathrm{A}} < \omega_{\mathrm{B}}$. For simplicity we neglect damping and spatial dispersion, which we introduce only in Sect. 5.4, but which would be adequate for exciton resonances.

The equations displayed graphically in Fig. 4.9 are

$$\varepsilon_{\mathrm{A}}(\hbar\omega) = \varepsilon_{\mathrm{b}} + \frac{f}{\hbar^2(\omega_{\mathrm{A}}^2 - \omega^2)}\,, \tag{4.37a}$$

$$\varepsilon_{\mathrm{B}}(\hbar\omega) = \varepsilon_{\mathrm{b}} + \frac{f}{\hbar^2(\omega_{\mathrm{B}}^2 - \omega^2)}\,, \text{ and} \tag{4.38a}$$

$$\varepsilon_{tot}(\hbar\omega) = \varepsilon_{\mathrm{b}} + \frac{f}{\hbar^2(\omega_{\mathrm{A}}^2 - \omega^2)} + \frac{f}{\hbar^2(\omega_{\mathrm{B}}^2 - \omega^2)} \tag{4.38b}$$

$$= \varepsilon_{\mathrm{A}}(\hbar\omega) + \varepsilon_{\mathrm{B}}(\hbar\omega) - \varepsilon_{\mathrm{b}}\,. \tag{4.38c}$$

We see in Fig. 4.9a and b what we expect already from Fig. 4.3a. Especially we see that the longitudinal transverse splitting Δ_{LT} is about 5 meV, the same for both resonances, because we used the same oscillator strength f and background dielectric constant ε_{b} for both resonances.

The situation becomes significantly different for (4.38c) shown in Fig. 4.9c. Though the oscillator strengths are still the same for both resonances, the value of Δ_{LT} is now smaller for resonance A and larger for resonance B, as indicated by the arrows while the singularities ω_{A} and ω_{B} did not shift. Δ_{LT} decreased to about 2 meV for the resonance A and increased to about 8 meV for resonance B. This is a first example of level repulsion, which we will encounter more frequently in the following chapters of this book. The "moral of the story" is, that one cannot deduce from Δ_{LT} or the spectral width of the stop-band, in a straight-forward way, the oscillator strengths of two close-lying resonances, but has to go through a careful analysis. As an outlook to Sect. 13.1.3 we state here that the AΓ_5 and BΓ_5 exciton resonances of CdS are sufficiently separated spectrally, i.e., several times Δ_{LT} to treat them to a good approximation as independent resonances, while those of ZnO are not.

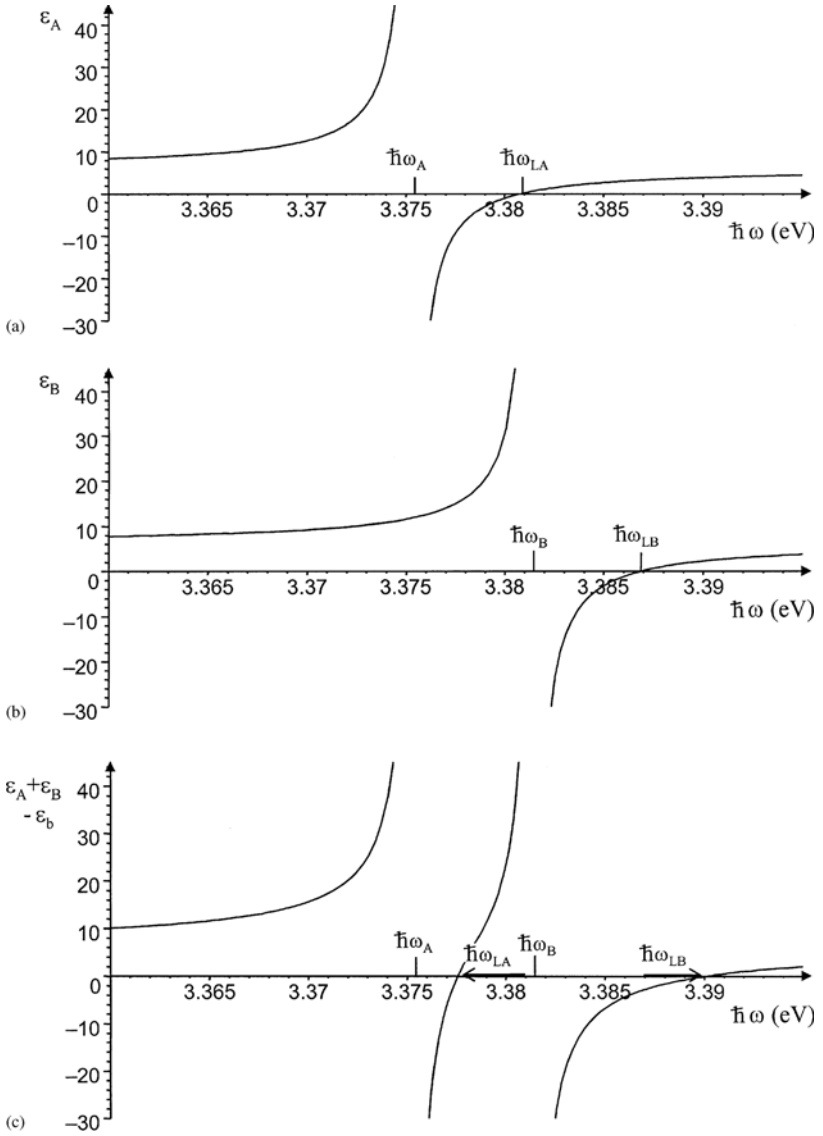


Fig. 4.9. The real part of $\varepsilon(\hbar\omega)$ for two resonances separately (**a**, **b**) and for two close lying resonances (**c**). Numbers on the x-axis refer roughly to ZnO

4.6 Problems

1. Study for the case of weak damping, some reflection spectra in the infrared (optical phonons) of at least partly ionic bound semiconductors, and compare with the data for ω_0 and ω_L given there.

2. Calculate the spectra of reflection for a single surface for weak and for strong damping with otherwise constant parameters. Compare the shift of the reflection maxima and minima with respect to the transverse and longitudinal eigenenergies ω_0 and ω_L, respectively. Which quantity can be deduced with reasonable accuracy from a first inspection of the reflection spectra?
3. Make a qualitative sketch of the electric fields for normal incidence of light on a medium with higher or lower index of refraction and weak or vanishing absorption, for strong absorption and for frequency $\omega_0 < \omega < \omega_L$ and vanishing damping.

References to Chap. 4

[63H1] J.J. Hopfield, D.G. Thomas: Phys. Rev. **132**, 563 (1963).
[67K1] R. Kubo: Dynamical Processes in Solid State Optics, Syokabo, Tokyo (1967) and further textbooks by this author.
[72W1] F. Wooten, Optical Properties of Solids, Academic Press, New York (1972)
[81K1] R. Kuhnert, R. Helbig, K.Hümmer: phys. stat. sol. (b) **107**, 83 (1981).
[85C1] H.B. Callen: Thermodynamics and an Introduction to Thermostatistics, 2nd ed., Wiley, New York (1985).
[93F1] M. Fiebig, D. Fröhlich, Ch. Pahlke-Lerch: phys. stat. sol. (b) **177**, 187 (1993).
[01G1] B. Gil: Phys. Rev. B **64**, 201310 (R) (2001).
[01W1] M. Wegener, in [81A1]j, p. 215 (2001).
[02B1] Broadband Dielectric Spectroscopy, F. Kremer and A. Schönhals (eds.), (Springer, Berlin 2002).
[02S1] W. Schäfer and M. Wegener, Semiconductor Optics and Transport Phenomena, Springer, Berlin (2002)

5

The Concept of Polaritons

Here we want to discuss in more detail what is actually propagating when "light" travels through matter. In vacuum the situation was quite clear on our present level of understanding in Sect 2.2. Light in vacuum is a transverse electromagnetic wave, the quanta of which are known as photons.

There are two levels at which one can describe the interaction of light with matter. One is the so-called perturbative treatment or weak coupling case. In this approach, which we used in Sect. 3.2, the electromagnetic field and the excitations of the matter are treated as independent quantities. As shown in Fig. 3.16a a photon is absorbed and the matter goes from the ground state to the excited state, and that's it. This approach is sufficient for many purposes but, if we look closer, we see that this is not the whole story. The optically excited state of the matter is necessarily connected with some polarization $\boldsymbol{P}$. Otherwise the transition would be optically forbidden, i.e., it would not couple to the electromagnetic field e.g. via the dipole-operator. On the other hand, we know that every oscillating polarization emits an electromagnetic wave, which may act back onto the incident electromagnetic field. This interplay will lead us in the following to the strong coupling limit between light and matter and to the concept of polaritons, see [58H1, 63H1, 75C1, 84A1, 85H1, 98K1] and references therein. In later chapters we shall see that many of the experimentally observed phenomena can be described quantitatively only in the strong coupling limit. In this chapter we introduce the concept of polaritons including the aspects of spatial dispersion and of surface polaritons. The concept of polaritons was originally introduced for crystalline solids [63H1], but it can be easily generalized to non-crystalline materials including gases. We want to stress here the general concept and give examples from different fields, which are otherwise not the topic of this book. A large number of examples from the field of semiconductor optics can be found in later chapters, e.g., for phonon-, plasmon or exciton polaritons including systems of reduced dimensionality, cavity (i.e. Fabry–Perot) polaritons or photonic crystals (see Chaps. 11–13, 15, 17).

5.1 Polaritons as New Quasiparticles

Due to the relation (2.28) repeated here:

$$\boldsymbol{P} = \varepsilon_0 \left[\varepsilon(\omega) - 1\right] \boldsymbol{E}, \tag{5.1}$$

the electric field in matter is always accompanied by a polarization wave. This statement is true as long as $\varepsilon(\omega)$ or $\tilde{n}(\omega)$ deviate from one, and holds in the whole spectral range from $\omega = 0$ well beyond the highest eigenfrequencies of the system. These highest eigenfrequencies are actually situated in the X-ray region, for example, the K or L absorption edges. Above this region $\varepsilon(\omega)$ and $\tilde{n}(\omega)$ approach unity from below according to (4.19). In other words, we can state that, for the whole spectral range discussed in this book, i.e., for the IR, VIS and UV, light travelling in a solid is always a mixture or superposition of an electromagnetic wave and a "mechanical" polarization wave and this is not only true for semiconductors, but also for other dense media like metals, insulators and liquids and close to resonance even for gases. This mixed state of electromagnetic and polarization waves is quantized (see below) in the sense that it can exchange energy with other systems, including the photon–field in vacuum only by integer multiples of $\hbar\omega$.

The name of these energy quanta is polariton. This name is composed of polarization and photon, thus directly describing what it is, especially since the ending -on is generally used for (quasi-) particles (for an explanation of "quasi-" see Sect. 7.6) like photon, electron, gluon, proton, meson, plasmon, phonon and many others, but with a few exceptions (neutrino, quark). As we shall later see, these polarization waves include motions of different ions in a semiconductor relative to each other, two particle, i.e., electron–hole pair excitations, and collective motions of the electron cloud with respect to the nuclei and the inner filled electron shells. Later (Chaps. 7–10), we shall discuss these excitations in more detail and see that they can also be quantized to form quasiparticles with energy $\hbar\omega$ and momentum $\hbar\boldsymbol{k}$ in a similar way as for photons. The names of these quanta, or quasiparticles, are phonons, excitons and plasmons, respectively. We can therefore repeat the above statement about what is propagating as "light" in matter, namely that it is a mixture of photons and other quasiparticles that describe the quanta of the polarization field, with the new aspect, that the mixed state is quantized, not its constituents.

In the mechanical model used in Chap. 4 we would describe this phenomenon as follows: an incident electromagnetic wave excites the oscillators. This oscillation is connected with a polarization which itself radiates again an electromagnetic wave; this in turn excites the oscillators, etc.

In the following we outline shortly the concept of polaritons in the picture of second quantization for a crystal in which $\boldsymbol{k}$ is a good quantum number.

Anyone who is not familiar with the concept of second quantization can simply skip the following treatment. There will be no problems in understanding the rest of the book. It must just be remembered that light propagating

in matter is a mixture of an electromagnetic and a polarization wave. This mixed wave can be quantized and the energy quanta are known as polaritons. They are the quasiparticles of "light" in matter.

The Hamiltonian, which, in the picture of second quantization, describes the interacting system of the photons and other quasiparticles (or excitations of matter) reads according to e.g. [93H1, 93P1] of Chap. 1.

$$H = \sum_k \hbar\omega_k a_k^+ a_k + \sum_{k'} E(\boldsymbol{k}') B_{k'}^+ B_{k'} + \mathrm{i}\hbar \sum_k g_k \left(B_k^+ a_k + \text{h.c.}\right). \tag{5.2}$$

The first two terms on the right-hand side contain the number operators $a_k^+ a_k$ and $B_{k'}^+ B_{k'}$of the photons and of the other quasiparticles (see also (2.55b)) which are in most cases also bosons, and represent the Hamiltonian of the non-interacting systems. The third term describes the interaction, e.g., the annihilation of a photon a_k and creation of another quasiparticle B_k^+ (under momentum conservation) or vice versa. The prefactors g_k simply contain the transition matrix elements H_{ij} discussed in Sect. 3.2.2. If we use these terms as a perturbation, we arrive back at the weak-coupling limit.

The crucial point now is that the whole Hamiltonian (5.2) can be diagonalized by a proper choice of linear combinations p_k of creation and annihilation operators of photons and of the quasiparticles representing the matter. In the following, we merely outline briefly this Bogoliubov-transformation-like procedure. For details the reader is referred again to [93H1, 93P1] of Chap. 1.

The above procedure brings the Hamiltonian of (5.2) into the following form:

$$H = \sum_k E_k p_k^+ p_k \tag{5.3}$$

with suitable coefficients u_k and v_k:

$$p_k = u_k B_k + v_k a_k. \tag{5.4}$$

The p_k and p_k^+ are the annihilation and creation operators for the quanta of the mixed state of photon and polarization wave with wave vector $\boldsymbol{k}$, which are consequently called polaritons and $E(\boldsymbol{k})$ is their energy and dispersion relation.

5.2 Dispersion Relation of Polaritons

All wave-like excitations can be described by two quantities, namely their (angular) frequency ω which is connected with the quantum energy simply by $E = \hbar\omega$ and their wave vector $\boldsymbol{k}$ which gives the (quasi-) momentum $\hbar\boldsymbol{k}$. The relation which connects ω and $\boldsymbol{k}$ is usually called dispersion relation $E(\boldsymbol{k})$ or $\omega(\boldsymbol{k})$

The dispersion relations $E(\boldsymbol{k})$ which we have encountered until now were very simple. For photons in vacuum it was a straight line with slope $\hbar c$

(Fig. 2.2) and for the polarization wave of an ensemble of uncoupled oscillators a horizontal line (Fig. 4.2a). The relation between ω and $\boldsymbol{k}$ for polaritons, i.e., for the light quanta in matter, can be derived from classical physics and agrees with the results of the quantum-mechanical treatment outlined above.

We remember that the wave vector in matter $\boldsymbol{k}$ is connected with the wave vector in vacuum $\boldsymbol{k}_\mathrm{v}$ by the complex refractive index $\tilde{n}(\omega)$ (2.36). To get rid of the vector character, we consider the squares

$$\boldsymbol{k}^2 = k^2 = \tilde{n}^2(\omega) k_\mathrm{v}^2 \tag{5.5}$$

Now we also recall (2.35), (2.13) saying that $\tilde{n}^2(\omega) = \varepsilon(\omega)$ and $\boldsymbol{k}_\mathrm{v}^2 = (2\pi/\lambda_\mathrm{v})^2 = (\omega/c)^2$ and obtain again (2.33)

$$\frac{c^2 k^2}{\omega^2} = \varepsilon(\omega). \tag{5.6}$$

This is the so-called polariton equation. On the other hand, we know $\varepsilon(\omega)$ which is given in the vicinity of a single resonance by (4.20), (4.22). Putting (5.6) and (4.22) together we find

$$\frac{c^2 k^2}{\omega^2} = \varepsilon_\mathrm{b} + \frac{f}{\omega_0^2 - \omega^2 + \mathrm{i}\omega\gamma}. \tag{5.7}$$

This is an implicit representation of $\omega(\boldsymbol{k})$ for the polaritons. For the simplest case, namely vanishing damping γ and no dependence of ω_0 or f on k, it is quite easy to calculate $\boldsymbol{k}(\omega)$ and $\omega(\boldsymbol{k})$. We do not give the formulas here because they bring no further physical insight, but in Fig. 5.1 we give the dispersion relation for the case just mentioned including only one

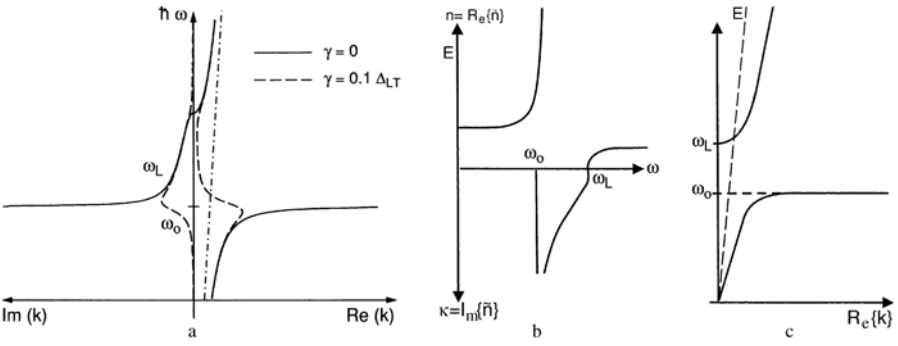


Fig. 5.1. The polariton dispersion in the vicinity of a single resonance for vanishing damping (*solid lines*) and finite damping (*dashed lines*) for $\varepsilon_\mathrm{b} = 1$. The *dashed-dotted line* gives the disperson of photons in vacuum (**a**); real and imaginary parts of $\tilde{n}(\omega)$ for vanishing damping (**b**) and the creation of the polariton dispersion (*solid lines*) from those of excitons and photons (*dashed lines*) and the non-crossing rule; ($\varepsilon_\mathrm{b} > 1$) (**c**)

resonance. For several resonances the r.h.s. of (4.22) has to be replaced by (4.20). The dispersion relation starts for $\omega = 0$ and $\boldsymbol{k} = 0$ as a straight line. This part is called the lower polariton branch (LPB). Since the dispersion of photons in vacuum is also a straight line, but with a slope $\hbar c$ instead of $\hbar c \varepsilon_{\mathrm{s}}^{-1/2}$, the dispersion relation is said to be "photon-like" as long as it is a straight line.

It bends over when we approach the resonance frequency ω_0, and in this region the polariton dispersion is called phonon-like or exciton-like, depending on whether the resonance corresponds to a phonon or an exciton, respectively. Between the transverse and longitudinal eigenfrequencies there is no propagating mode for the present approximation of uncoupled oscillators and for vanishing damping, i.e., we have again the stop-band or Reststrahlbande discussed in Sect. 4.4. There may be a longitudinal branch at $\hbar\omega_{\mathrm{L}}$, which does usually not couple to the electromagnetic field. At $\hbar\omega_{\mathrm{L}}$ also the upper polariton branch (UPB) begins. This bends upwards again displaying a photon-like behaviour, but now with a slope $\hbar c \varepsilon_{\mathrm{b}}^{-1/2}$ compared to $\hbar c \varepsilon_{\mathrm{s}}^{-1/2}$ for the LPB. Between ω_{T} and ω_{L}, $\boldsymbol{k}$ is purely imaginary since $\tilde{n}(\omega)$ is imaginary in this range; the consequences for the optical properties were discussed already in Sect. 4.4 above.

Actually, the dispersion relation shown in Fig. 5.1a is not so surprising as it looks at first glance. If we take the spectral dependences of $n(\omega)$ and of $\kappa(\omega)$, i.e., of the real and imaginary parts of $\tilde{n}(\omega)$, from Fig. 4.4, turn the ω-axis from the x-direction into the y-direction (see Fig. 5.1b) and multiply $n(\omega)$ and $\kappa(\omega)$ by ωc^{-1}, i.e., essentially by a straight line through the origin, according to

$$\mathrm{Re}\{k\} = n(\omega)\omega c^{-1}; \quad \mathrm{Im}\{k\} = \kappa(\omega)\omega c^{-1}, \tag{5.8}$$

we obtain Fig. 5.1a.

The dispersion relation of the polariton can also be deduced from the quantum-mechanical "non-crossing rule". This non-crossing rule says roughly the following: There are two energy levels E_1 and E_2, which depend on some parameter p. This parameter can be the wave vector, a particle density, a constant electric or magnetic field, a strain field, etc. We assume that these two levels cross as a function of the parameter as sketched in Fig. 5.2 by the dashed lines. If there is any coupling between these two levels, then the cross-over point disappears, and the two levels repel each other in the way shown by the solid lines in Fig. 5.2. The splitting at the former cross-over point is just proportional to the coupling strength between the two levels.

Actually this behaviour is not new or characteristic for quantum mechanics. It is basically only the splitting of the eigenfrequencies of two coupled identical pendula (or more generally of two harmonic oscillators) caused by a weak coupling found in every undergraduate textbook on classic mechanics including the beating if both eigenfrequencies are excited. In [98J1] it has been shown in detail that the classical and the quantum-

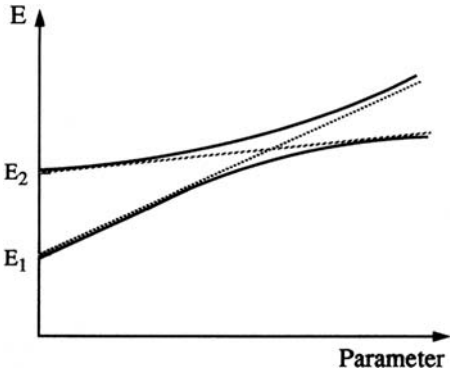


Fig. 5.2. A sketch to illustrate the quantum-mechanical non-crossing rule for two non-interacting (*dashed line*) and interacting (*solid line*) energy levels

mechanical splitting of coupled oscillators are indeed governed by the some set of equations.

We now apply the general knowledge obtained above and with Fig. 5.2 to the polariton problem. See Fig. 5.1c. The "parameter" in Fig. 5.2 is now obviously the $\boldsymbol{k}$ vector. We have the horizontal dispersion of the ensemble of uncoupled oscillators and the steep straight line for the photons. If the ensemble of oscillators does not couple to the electromagnetic photon field, the dashed lines already describe the system. If there is, however, a finite coupling (e.g., a non-vanishing dipole or quadrupole coupling matrix element) the crossing is avoided and the solid lines describe the dispersion relation of the coupled system. Note that (apart from a possible longitudinal branch, not shown in Fig. 5.1c for clarity) the solid lines in Fig. 5.1c are the only states of this system. There is no state at $\hbar\omega_0$ for $\boldsymbol{k} = 0$, as we shall see later in various realistic examples.

If we relax the assumption of vanishing damping, the polariton dispersion is changed as shown in Fig. 5.1a by the dashed line. Now the propagating modes below ω_{T} and above ω_{L} experience damping too. Between ω_{T} and ω_{L} there appears a region with negative group velocity $\mathrm{d}\omega/\mathrm{d}k < 0$. At first glance, this situation seems unphysical in the sense that a light pulse moves out of a sample ($v_{\mathrm{g}} < 0$) when we send it onto the sample. A first way out of this dilemma is given by the strong absorption which is necessarily connected with the region of negative v_{g}. It allows one to interpret the negative v_{g} in this spectral region as a wave that is damped out faster than it can propagate.

Another interpretation can be found later in Sect. 5.4, where we take into account spatial dispersion. For this more general case, the problem of negative v_{g} essentially disappears and we obtain in addition an understanding, why e.g. the TO-phonon branch can be observed by inelastic neutron scattering over the whole Brillouin zone, though the dashed line in Fig. 5.1a turns around for rather small k-values already.

For a discussion of the fact, that both the phase- and group velocities can exceed for some frequencies the vacuum speed of light see [94R1]. It should be noted, however, that the signal and energy transport velocities never exceed c.

We have stated several times that at $\hbar\omega_{\mathrm{L}}$, i.e., at the frequency where $\varepsilon(\omega = \omega_{\mathrm{L}}) = 0$, a longitudinal branch may exist. We want to give now some insight into when this will be the case with the help of Fig. 5.3. In Fig. 5.3a we assume that the ensemble of oscillators can be excited with the same eigenfrequency in all directions of space. This situation frequently occurs to a good approximation in cubic materials. If a light beam is shone on the sample, e.g., in the direction k_x, then the oscillators oriented along y and z couple to the light field and form a twofold degenerate, transverse upper and lower polariton branch, while the oscillators oriented in the x direction are parallel to k_x and give rise to a singly degenerate longitudinal branch.

If we have a uniaxial material, we frequently find oscillators, which can either be excited with the same eigenfrequency in the xy plane normal to the crystallographic $c \equiv z$ axis or with another eigenfrequency parallel to z only. In the first of these two cases shown in Fig. 5.2b, we obtain only a twofold degenerate upper and lower polariton branch, but no longitudinal branch for $\boldsymbol{k} \| z$ since both oscillators in the xy plane are then transversal. For $\boldsymbol{k}$ in xy plane, e.g., $\boldsymbol{k} \| x$ there is a nondegenerate transverse polariton branch for $\boldsymbol{E} \| y$ and a longitudinal branch from the x-oriented oscillators. For $\boldsymbol{k} \| x$ and $\boldsymbol{E} \| z$ the oscillators do not couple and we obtain intersecting dispersion curves.

Finnally, if we consider an oscillator that can be elongated only in one direction (e.g., z) we obtain only one nondegenerate transverse lower and upper polariton branch for $\boldsymbol{k} \perp z$ and $\boldsymbol{E} \| z$, an intersecting longitudinal branch for $\boldsymbol{k} \| z$ and no interaction at all for $\boldsymbol{k} \perp z$, $\boldsymbol{E} \perp z$.

If a light beam propagates with angle α different from 0° or 90° relative to the c-axis in a uniaxial material, we obtain the ordinary beam for $\boldsymbol{E}$ perpendicular to the plane defined by $\boldsymbol{k}$ and z, since $\boldsymbol{E}$ is then always in the xy plane, i.e., normal to the main section and couples to the oscillators. For $\boldsymbol{E} \perp \boldsymbol{k}$ but in the plane defined by $\boldsymbol{k}$ and z we obtain the extraordinary beam with an orientation dependent refractive index. (See Sect. 3.1.7).

In the language of polaritons this is called a mixed-mode polariton for the following reason. The "quantization axis" for the light field is the crystallographic $\mathbf{c}$ or z axis. For a light beam propagating under an angle α to the $\boldsymbol{c}$ axis is different from 0° and 90° the electric field can be decomposed into two components $\boldsymbol{E} \| \boldsymbol{c}$ and $\boldsymbol{E} \perp \mathbf{c}$, the relative strength of which depends on α. For the situation of Fig. 5.3b only the component $\boldsymbol{E} \perp \boldsymbol{c}$ couples to the oscillators and in Fig. 5.3c only $\boldsymbol{E} \| \boldsymbol{c}$. In Fig. 5.3d we show the dispersion of these mixed mode polaritons for an ensemble of oscillators, which can be excited only in the $\boldsymbol{c}$ direction, and light polarized linearly in the $\boldsymbol{k}$, z plane. The oscillator strength and thus the $L-T$ splitting depend on the ratio $\boldsymbol{E} \| \boldsymbol{c} / \boldsymbol{E} \perp \mathbf{c}$. It is maximal for $\beta = \angle(\boldsymbol{E}, z) = 0°$ corresponding to the situation on the l.h.s. of Fig. 5.3c and is zero for $\angle(\boldsymbol{E}, z) = 90°$ corresponding to the middle one. In between the oscillator strength and Δ_{LT} vary according to

$$\Delta_{\mathrm{LT}}(\beta) = \Delta_{\mathrm{LT}}(\beta = 0) \cos^2 \beta \,. \tag{5.9}$$

It should be noted there that we always used the angles α and β of $\boldsymbol{k}$ and $\boldsymbol{E}$ relative to $\boldsymbol{c}$ in the medium, i.e., the Snellius law with a β-dependent refractive index has to be used in addition when considering the external angles.

For biaxial materials the situation becomes even more complex, but as stated in Sect. 3.1.7, this leads to crystal optics, which is beyond the scope of this book.

Similarly to the dielectric function in Fig. 4.9 the polariton dispersion is modified in the case of two close-lying resonances, i.e., the transition matrix element cannot be deduced directly from the width of the stopband.

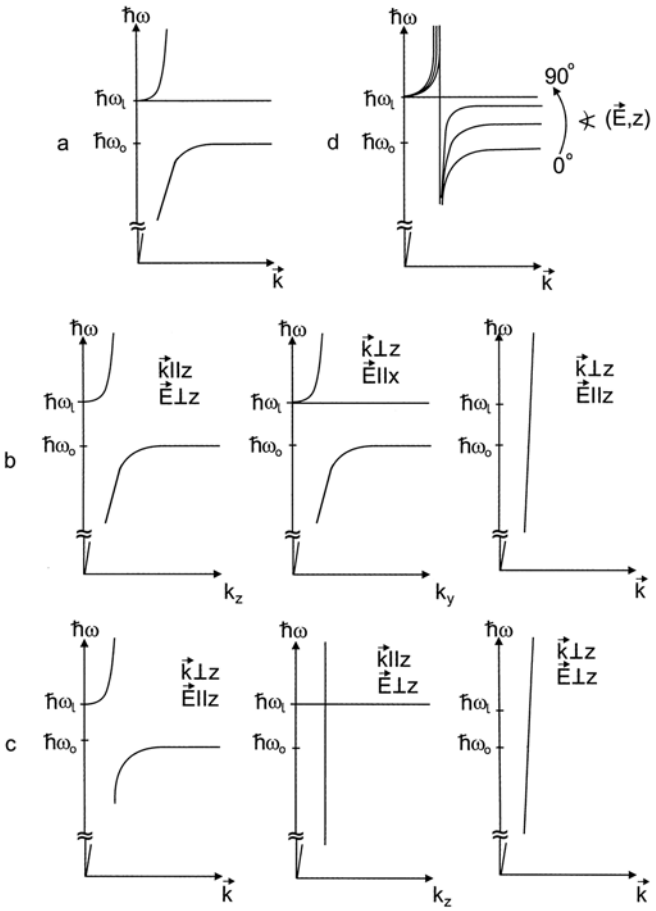


Fig. 5.3. Sketches of the polariton dispersion and of possible longitudinal branches for various orientations and an ensemble of oscillators that can be excited with the same eigenfrequency and oscillator strength in all three directions of space (**a**) only in the xy plane (**b**) and only along $z = c$ (**c**). The angle-dependent mixed-mode polariton dispersion for the situation as in (**c**) is shown in (**d**)

5.3 Polaritons in Solids, Liquids and Gases and from the IR to the X-ray Region

We shall see in later chapters many examples for polaritons in semiconductors in the infrared (IR) part of the spectrum (phonon-, magnon- or plasmon-polaritons) and in the visible (VIS) and near ultraviolet (exciton-, cavity- and plasmon-polaritons). Here we want to stress that the polariton concept is by no means an academic problem or restricted to semiconductors. It works well for the description of the optical properties of insulators or metals, but also for disordered systems like amorphous solids, liquids and gases, and it is also valid for any resonance coupling to the electromagnetic field, including intranuclear transitions in the X or γ-ray region.

5.3.1 Common Optical Properties of Polaritons

To demonstrate that a spectrum of the refractive index as shown, e.g., in Fig. 5.4 is not a characterisation of solids we present an 100-year-old example for Na vapor in Fig. 5.4. See [02W1, 03L1, 09M1] or [76P1] from Chap. 3 and references therein. If a continuous spectrum (a so-called white light spectrum) is transmitted through a cell filled with Na vapor of homogenous density (Fig. 5.4a,b) one finds the transmitted light absorption bands. One of the most prominent bands corresponds to the well-known yellow emission (and absorption) line with is actually a double line at 589.35930 nm and 588.99631 nm corresponding to the 2.1 eV range. The absorption process is due to the transition of the outer electron from $^2S_{1/2}$ to $^2P_{3/2}$ and $^2P_{1/2}$, respectively. See [91D1] of Chap. 1. The emission is just the reverse process. If the white light beam is sent through a prism of Na vapor, the light is defected around this resonance as shown in Fig. 5.4c,d. The behavior is just the one we expect for the refractive index and the polariton dispersion in the vicinity of a resonance. When we approach the resonance from lower photon energies, n and $\boldsymbol{k}$ increase rapidly. At the resonance absorption occurs and above the refractive index $\sqrt{\varepsilon_b}$ approaches from below, explaining the deflection of light in opposite directions for $\omega < \omega_0$ and $\omega > \omega_L$, respectively. In contrast to solids, gases and vapors are very dilute. Consequently, the oscillator strength is small and n deviates considerably from unity only in the very vicinity of the resonance. Similar experiments have been performed with many other vapors of, e.g., Sr, Ca, Ba [02W1, 03L1, 09M1]. The results allowed to develop and to verify the concept of the Lorentz oscillators. At that time neither the concept of photons nor of polaritons was known. The quantization of the electromagnetic field in photons was first introduced by Planck in 1905 and allowed him to explain black body radiation [see (2.85)]. Later on it was beautifully confirmed in the explanation of the (external) photoelectric effect by Einstein, which had been earlier observed by Milikan and Hertz.

From our present point of view, it is clear that the light field in Na vapor is also quantized in energy packets of $\hbar\omega$. The fact that the dispersion relation

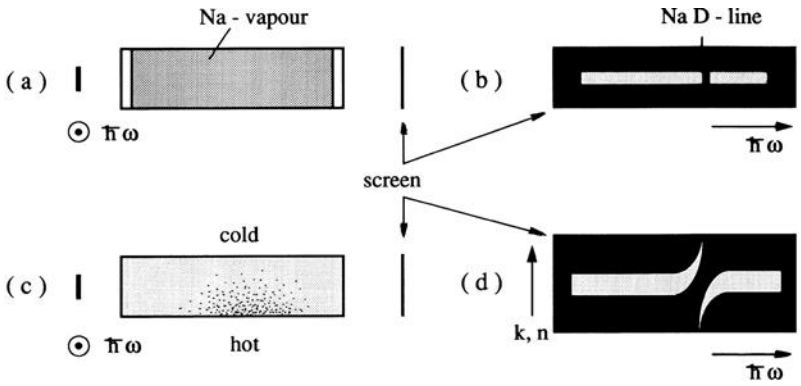


Fig. 5.4. Absorption and dispersion of Na vapour in the vicinity of the yellow sodium line. Schematic arrangement for the absorption (**a**) in a homogeneous vapour and the resulting absorption spectrum (**b**). Deflection in a "prism" of Na vapor (**c**) and the dispersion relation (**d**). The energy axis is normal to the paper in (**a**) and (**c**); (**b**) and (**d**) are photographic positives ([00K1] and [76P1] of Chapt. 3 and references therein)

$\hbar\omega(\boldsymbol{k})$ and the refractive index deviate in the vicinity of the resonance clearly from $\hbar\omega = \hbar ck$ and $n = 1$, respectively, is proof that these quanta consist not only of electromagnetic radiation, but that they are necessarily accompanied by a polarisation wave. Therefore it is straightforward to generalize the concept of polaritons beyond ordered or crystalline systems. One point which has to be considered, however, is the wave vector. We come back to this aspect in Sect. 5.3.2.

To conclude the discussion of Fig. 5.4 it should be noted that the vapor prisms were generally not obtained by keeping the Na vapor in a prism shaped cell, but that a concentration gradient has been created laterally over the light beam by a transverse temperature gradient and results in a density gradient which in turn gives a gradient in oscillator strength and finally a variation of the optical path length across the beam in a similar way as a usual (glass) prism does.

Now we present beautiful experiments that further prove the validity of the polariton concept in seemingly completely different systems.

The basic idea is illustrated in Fig. 5.5a where we show the polariton dispersion in the vicinity of a resonance. We now send a short light pulse on the sample. The duration of the pulse τ_P is assumed to be so short that its spectral width $\Delta\omega$ covers the vicinity of the resonance via $\tau_P \cdot \Delta\omega \approx 1$ and $\Delta\omega > \Delta_{\mathrm{LT}}$. We can now discuss the propagation of this pulse through the medium in different ways. We know, that a pulse propagates with its group velocity $v_g = \partial\omega/\partial k$ [see (2.15)]. Evidently we find on the lower and the upper polariton branches pairs of states with equal v_{g}, i.e., equal slope as shown for one case in Fig. 5.5a. Since these states are both excited coherently by the incident pulse and since they have slightly different frequencies, they accumulate a relative

phase shift, which increases linearly with time and traveling distance and changes the superposition periodically from constructive to destructive and back again. This is nothing but the well know beating phenomenon. Depending on the sample thickness and v_g, some of the polariton pairs will arrive at the end of the sample with constructive interference, and adjacent pairs will arrive in a destructive superposition, the next with a constructive one, etc. The consequence is that there will be no simple pulse at the end of the sample but one will find an envelope of the amplitude and intensity of the transmitted beam oscillating in time. The period of this oscillation increases with time, since the polariton pairs with large v_g, i.e., a steep slope arrives first. They have a bigger frequency difference and consequently a faster beating period. The pairs with smaller v_g also have a smaller beat period and arrive later. This is exactly the behavior observed for an incident ns pulse centered around the yellow resonance in Na vapor (Fig. 5.5b and [84R1]) and propagating through a cell with Na vapor with a length of 40 mm. The temporal decay of the transmitted signal is governed by the damping of the resonance and by the energy density per frequency interval of the incident pulse. The experiment in Fig. 5.5b thus confirms again the validity of the polariton concept in this system.

An alternative way to describe the experiment is to Fourier transform the incident pulse from the time to the frequency domain (see e.g. [99B1]). Then one allows each frequency component to propagate through the sample according to its complex index of refraction $\tilde{n}(\omega) = n(\omega) + i\kappa(\omega)$ resulting for each frequency ω after the sample in a well-defined phase and amplitude. All components are then coherently superimposed and this sum is back transformed to the time domain, resulting in a very good description of the experimental finding. Actually the theoretical curve shown in Fig. 5.5b has been calculated along these lines.

With Fig. 5.5c,d we demonstrate that identical results have been obtained with completely different systems. Figure 5.5c shows these propagation quantum beats for the well known exciton resonance in the yellow spectral range of Cu_2O. (More precisely, in the 1s ortho exciton of the yellow series, which is only quadrupole allowed. For details see Sect. 13.2.) The spectral region is approximately the same as for Na vapor, but Cu_2O is a crystalline solid more precisely a semiconductor.

The third example in Fig. 5.5d from [99B1] is, by contrast, in a completely different spectral rage. The resonance is the 14.4 keV Mösbauer transition of ^{57}Fe. The resonance is not even in the electronic system of the solid, but in the nucleus and the light pulse has been produced in a synchrotron. Nevertheless the phenomenon is identical, revealing that the polariton concept is also applicable in this spectral range of X or γ-rays.

A further common feature of the polariton dispersion is the dramatic decrease of the slope of the dispersion curve just below ω_0. Consequently, the group velocity should also obtain in this spectral region, values considerably below the vacuum speed of light $c \approx 3 \times 10^8$ m/s. Indeed this phenomenon

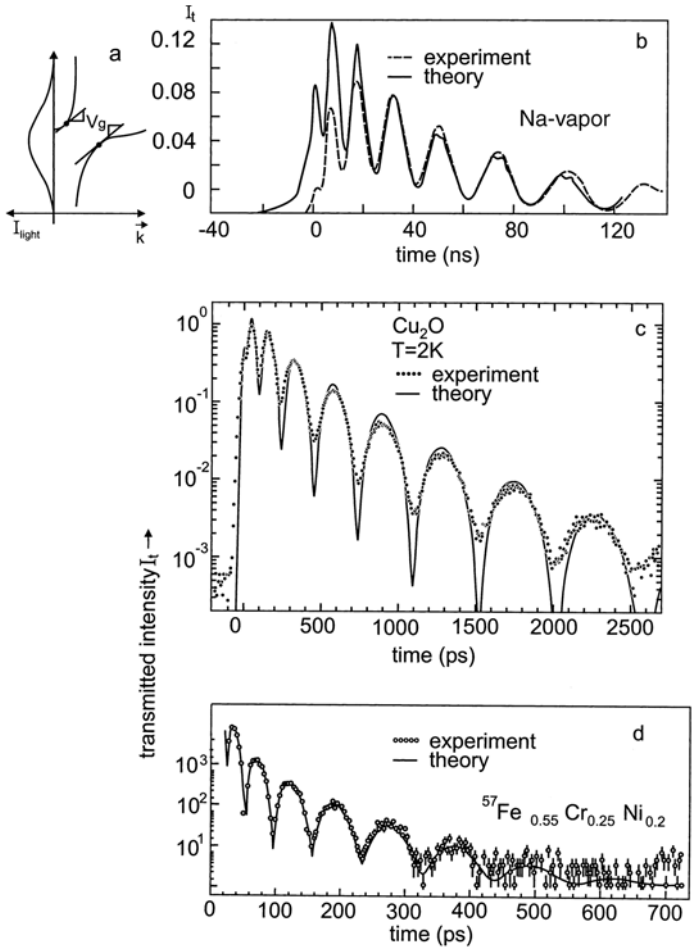


Fig. 5.5. The schematic explanation for the occurrence of propagation quantum beats (**a**) and their observation in Na vapor (**b**) in the ortho exciton resonance of the yellow series in Cu_2O (**c**) and in an ensemble of ^{57}Fe nuclei (**d**) ([84R1,91F1,99B1])

has been observed, e.g., for the exciton polariton in CuCl, GaAs or InSe with values of v_g down to 10^4 m/s. See Sect. 13.1 or [79M1, 79U1, 81I1, 97N1] and more recently also in Cu_2O [04F1]. In the vicinity of the Mösbauer transition of ^{57}Fe nuclei at 14.4 keV, even values down to the 10 m/s range have been observed [99B1]. Comparably low values and even a complete stop of light for a short time have been found in Na vapor [99H1,99K1,01K1]. In the latter case an additional trick has been used, namely excitation induced transparency in a small spectral interval of the upper of the two yellow Na lines, which results via Kramers–Kronig relations in a very steep structure in the group-velocity spectrum.

5.3.2 How the *k*-vector Develops

In nature, there are several conservation laws. See also Sect. 3.1.3. The most basic ones state that in a closed system (c.s.), i.e., a system that has no interaction with the surroundings whatsoever, the sum of all energies (including the rest-mass energies m_0c^2), of all momenta $\boldsymbol{p}$, of all angular momenta $\boldsymbol{L}$ (including spins) and of all electric charges ρ are constant. The sum of all entropies S_i is either constant, if there are no or only reversible processes, or it increases in the case of an irreversible processes. These conservation laws are presented in (5.10).

$$\sum_{\text{c.s.}} E_i = const \tag{5.10a}$$

$$\sum_{\text{c.s.}} \boldsymbol{P}_i = \boldsymbol{const} \tag{5.10b}$$

$$\sum_{\text{c.s.}} \boldsymbol{L}_i = \boldsymbol{const} \tag{5.10c}$$

$$\sum_{\text{c.s.}} \rho_i = const \tag{5.10d}$$

$$\sum_{\text{c.s.}} S_i \geq const \tag{5.10e}$$

These conservation laws are facts of experience that cannot be proven from first principles, but can be traced back to symmetries as outlined already in Sect. 3.1.4.

The only closed system in a strict sense is the whole universe. But much smaller parts on a laboratory scale are frequently a very good approximation of closed systems as we use them, e.g., in the lectures or practice to demonstrate elastic and inelastic scattering processes.

We have already used a subsystem in the above sense, without mentioning it, in derivation of the law of reflection and refraction in Sect. 3.1.4.

As a preparation for what follows, we show in Fig. 5.6 the reflection or diffraction of a single photon at a plane interface between the vacuum and matter repeating in Fig. 3.5.

In Sect. 3.1.4 we argued that energy conservation requires $\omega_i = \omega_r = \omega_{tr}$ and that there is only translational invariance along the interface plane, consequently only the parallel component of $\hbar\boldsymbol{k}$ is conserved. This argument is correct for the considered subsystem. If we consider the whole closed system, the total momentum has to be conserved. This means that momentum is transferred to the matter, namely $2\hbar\boldsymbol{k}_{i\,normal}$ in the case of reflection and $\hbar\,(\boldsymbol{k}_{i\,normal} - \boldsymbol{k}_{tr,\,normal})$ for refraction. The energy transfer to the matter connected with this momentum transfer is completely negligible, however, due to the mass of the matter M, which is many orders of magnitude larger than the "mass" of the photon $\hbar\omega c^{-2}$.

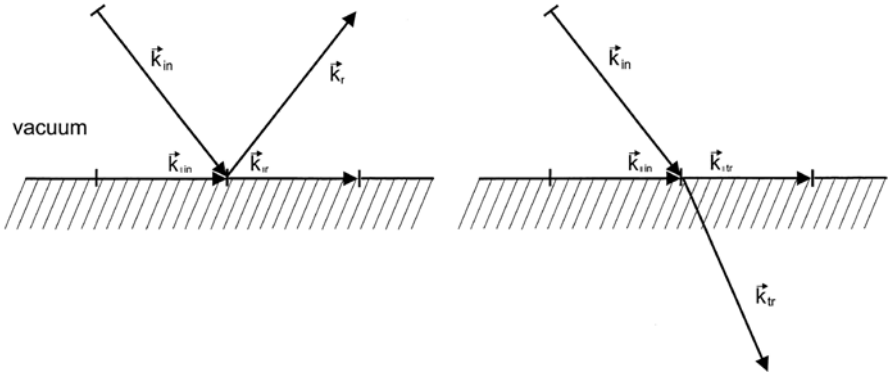


Fig. 5.6. Reflection and refraction of a single photon at an interface between vacuum and matter

So it is safe to say for the subsystem that the matter can "take up" momentum without energy transfer or to restrict ourselves to the subsystem "interface" with the restricted translational symmetry. For a more detailed discussion of this aspect see [98B1] and the references given therein.

Now we want to use a similar set of arguments to show how $\boldsymbol{k}$ conservation develops in optics or the law that a light beam propagates along a straight line in a homogenous medium.

We start with the situation of light scattering at a single atom in Fig. 5.7.

An incident plane electromagnetic wave with $\boldsymbol{k}_{in}$ hits an atom. In classical physics this atom performs a forced oscillation with an amplitude that depends on the detuning between ω_i and the resonance frequency ω_0 of the atom and on damping. This oscillation is connected with an oscillating dipole, radiating a dipole wave as shown schematically in Fig. 5.7. Evidently, the emission goes into all directions with the angular dependent amplitude of an oscillating

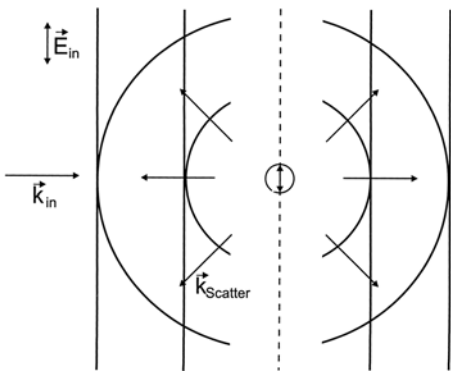


Fig. 5.7. Schematic drawing of the wave fronts in a light scattering event at a single atom or oscillator ([98J1])

.

dipole. This means that there is nothing like a straight propagation of the light beam in the interaction with a single atom.

The same result is obtained in the quantum mechanical description. A photon comes in with $\hbar\boldsymbol{k}_i, \hbar\omega_i$ and virtually excites the atom. After a time Δt, limited by $\hbar(\omega_i - \omega_0)\Delta t \leq \hbar$, where $\hbar\omega_0$ is now a transition energy in the atom, the atom again radiates a photon $\hbar\boldsymbol{k}_s, \hbar\omega_s$ in an arbitrary direction. If we repeat this experiment many times, the integrated angular intensity distribution will be the one of a radiating dipole. The process is known as light scattering or if ω_i approaches ω_0 as resonance fluorescence.

Momentum conservation can now be discussed in the two ways outlined above. We can either state that space with a single atom, e.g., at the origin of our coordinate system, no longer has translational invariance. Consequently, there is no longer momentum conservation and an incident photon $\hbar\boldsymbol{k}_i$ can be scattered in any direction $\hbar\boldsymbol{k}_s$.

The law of global momentum conservation can be satisfied by stating that the atom takes over some momentum $\boldsymbol{p}_{atom}$

$$\boldsymbol{p}_{atom} = \hbar(\boldsymbol{k}_i - \boldsymbol{k}_s)\,. \tag{5.11a}$$

The energy transfer connected with this momentum transfer

$$E_i = \boldsymbol{p}_{atom}^2/2M \tag{5.11b}$$

is in the neV regime for the visible and therefore for most practical purposes completely negligible, except for laser cooling of atoms. So we are again safe with both statements that either the atom takes up momentum but no energy or that the translational invariance is broken by the presence of the atom and momentum conservation is violated. This situation may change if we go to the X or γ-ray regime, where $\boldsymbol{k}$ increases. In the case of extremely narrow resonances, the frequency shift in the scattering process can be measurable. See problem 5.7.4.

Now we discuss the transition to a dense medium.

When we have a few of such atoms in the coherence volume of the incident light, like in a dilute gas, the scattered waves interfere constructively or destructively as a function of their direction dependent relative phases. So we obtain (momentarily) a rather complex scattering pattern, similar to a Speckle pattern, but is averaged out over time if the scattering centers move statistically as the atoms in a dilute gas do. The coherence volume of light is the volume over which the light wave has a well defined phase relation. The coherence volume depends on the light source and is generally smaller for incandescent light as compared to laser light. For details see, e.g., references [93S1] of Chap. 2 or [98B2] of Chap. 1.

The situation becomes simpler again if we have many oscillators or atoms in the coherence volume of light. It can be shown, that all scattered waves interfere essentially destructively except for the forward scattered waves. This

means that the straight propagation of a light beam, which implies $\boldsymbol{k}$ conservation, is trivial in vacuum and develops in matter again with increasing density of oscillators N.

A lower limit for the coherence volume of light V_{coh} is given by the wavelength λ

$$V_{coh} \geq \lambda^3 \,. \tag{5.12a}$$

The condition to have many oscillators in the coherence volume of light is fulfilled in any case if we have

$$N \cdot \lambda^3 \gg 1 \,. \tag{5.12b}$$

This is equivalent with the statement that the average distance d between the atoms or oscillators is small compared to the wavelength

$$d \gg \lambda \,. \tag{5.12c}$$

Consequently the light field cannot "resolve" the individual oscillators and a description in the sense of an effective medium is adequate [97S1] or reference [97W1] of Chap. 1. Equation (5.12) can be also used as a definition of a "dense" medium. See also the derivation of the dipole approximation in Sect. 3.2.2.

To summarize we can state that in a dense medium, $\boldsymbol{k}$ conservation is recovered, including its application in the laws of reflection and refraction at the boundary of a dense medium, and the electromagnetic wave propagating in this medium is accompanied by a polarization wave, bringing us back again to the polariton concept.

If the medium is completely ordered, i.e., a defect-free crystal, then there is only the straight propagating beam, since in such a situation we recover translational invariance modulo integer multiples of the reciprocal lattice vectors, (see Sect. 7.2).

If there are any inhomogeneities like defects in a crystal or inhomogeneous strain or fluctuations of N in a gas, there is always a small amount of light that is scattered without frequency shift, the so-called Rayleigh scattering. We come back to this topic in Sect. 23.2.

While $\boldsymbol{k}$ conservation (modulo reciprocal lattice vector $\boldsymbol{G}$ see Sect. 7.2) is a strict law in perfect crystals the situation is different in disordered systems like gases, liquids, amorphous solids or alloys as outlined below.

For waves with a wavelength λ that fulfills (5.12), holds the law of $\boldsymbol{k}$ conservation and we make use of it whenever we look through a glass window, onto a clear lake or through atmosphere. If we go to shorter and shorter wavelengths, however, the conditions (5.12) are less and less fulfilled and $\boldsymbol{k}$ conversation is more and more relaxed. This transition can be nicely followed by comparing, e.g., crystalline quartz (c-SiO_2) on one side and amorphous fused silica (a-SiO_2) or glass on the other.

In the visible, (5.12) is fulfilled in both cases, light propagates along a straight line, $\boldsymbol{k}$ conservation holds and we have well-defined laws of reflection and refraction at the surface.

If we go to the X-ray regime we have

$$\lambda \leq a\,. \tag{5.13}$$

For the c-SiO_2 $\boldsymbol{k}$ conservation modulo reciprocal, lattice vectors is still valid for elastic scattering, resulting in the Ewald's-construction and well-defined diffraction spots fulfilling

$$\boldsymbol{k}_i - \boldsymbol{k}_s = \boldsymbol{G}; \quad \hbar\omega_i = \hbar\omega_s \Rightarrow |\boldsymbol{k}_i| = |\boldsymbol{k}_s|\,. \tag{5.14}$$

In contrast in this regime for a-SiO_2 we obtain only blurred ring structures, indicating a substantial relaxation of $\boldsymbol{k}$ conversation at these short wavelengths.

So we can state that $\boldsymbol{k}$ conservation is valid in disordered systems if (5.12) is fulfilled, but is increasingly relaxed if we approach the condition (5.13).

5.4 Coupled Oscillators and Polaritons with Spatial Dispersion

For our model system we have assumed until now zero coupling between neighboring oscillators as shown in Fig. 4.1. The consequence of this assumption was a dispersion relation which is simply a horizontal line (Fig. 4.2a). We now relax this assumption by using a concept introduced in [58H1,62P1,63H1,74A1] and elaborated in detail e.g. in [75C1,75L1,78B1,78H1,78S1,79B1,79S1,81B1, 81L1, 81S1, 82O1, 82R1, 82S1, 83M1, 84A1, 84H1, 84S1, 85H1, 89R1, 91R1] and consider a more realistic coupling between neighboring oscillators realized, e.g., by weak springs, as shown schematically in Fig. 5.8. The most important consequence of this coupling is that the eigenfrequency is now a function of $\boldsymbol{k}$ as shown in Fig. 4.2b. For very long wavelength, i.e., $\boldsymbol{k} \to 0$, neighboring oscillators are still in phase and the coupling springs are not elongated. Therefore they still oscillate with the same frequency as the uncoupled ones. For decreasing wavelength, the coupling springs are elongated and increase the "effective" spring constant. As a consequence, ω_0 increases with increasing $\boldsymbol{k}$. The resulting band width 2B indicated in Fig. 4.2b is directly proportional to the coupling strength between neighboring oscillators. A wave packet created by elongating one or a few oscillators, as in Fig. 4.1d, will now propagate with a finite group velocity $v_\mathrm{g} = \mathrm{d}\omega/\mathrm{d}k$ and will show some dispersion. This means that the width of the spatial envelope of the wave packet will increase with time as indicated in Fig. 2.1b. This phenomenon can be easily observed by throwing a stone onto the still surface of a lake. The expanding ring-like wave structure shows a drastic increase of the width of its envelope function during propagation. The detailed shape of $\omega(\boldsymbol{k})$ depends on the physical nature of the oscillators and the coupling mechanism.

For our model system, ω_0 increases with $\boldsymbol{k}$ and we shall see that this is also true, e.g., for excitons. For optical phonons one has usually a decrease of

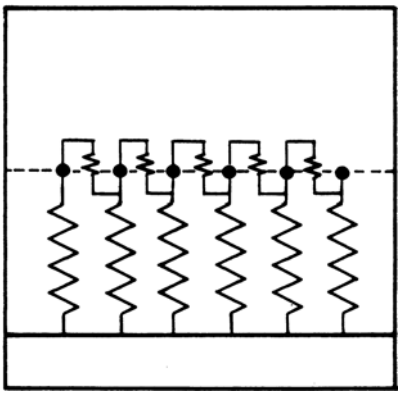


Fig. 5.8. An ensemble of coupled oscillators. Compare with Fig. 4.1. After [63H1]

the eigenfrequency with increasing $\boldsymbol{k}$. These points will be presented in more detail in Chaps. 7, 9 to 13. For the following discussion of our model system we shall use the dispersion relation of Fig. 4.2b, but the conclusions will also be qualitatively valid for other dependences of $\omega_0(\boldsymbol{k})$.

The fact that the eigenfrequency ω_0 of some excitation of a solid depends on $\boldsymbol{k}$ is often called "spatial dispersion", for reasons given later. However, the $\omega(\boldsymbol{k})$ dependence of the photons themselves, or of the polaritons, is not usually classed as spatial dispersion.

Since we have used the word dispersion now in various connections, we shall summarize the meanings here:

The term "dispersion relation" or simply "dispersion" means the relation $E(\boldsymbol{k})$ or $\omega(\boldsymbol{k})$ for all wave-like excitations independent of the functional dependence. It can be simply a horizontal line, a linear or parabolic relation, or something more complicated. Every excitation which has a wave-like character has a dispersion relation.

The term "spatial dispersion" means that the eigenfrequency of one of the elementary excitations in a solid depends on $\boldsymbol{k}$ and is not just a horizontal line (Fig. 4.2a,b).

In technical optics, dispersion often refers more specifically to the dependence of the refractive index n on the wavelength. Materials such as the glass used to make lenses, generally have a decreasing refractive index with increasing wavelength i.e., $\mathrm{d}n/\mathrm{d}\lambda < 0$. This behavior is called "normal dispersion" and is the usual behavior in the transparent spectral region, while an anormal dispersion, i.e., $\mathrm{d}n/\mathrm{d}\lambda > 0$ is limited to strongly absorbing regions (Fig. 4.4).

Finally, the word dispersion is also used for the fact that the envelope of a wavepacket, e.g., of a short light pulse in matter, becomes spatially broader with time.

The last two meanings of the word dispersion are actually consequences or special examples of the general definition of the term given first. If not stated otherwise, we use the term dispersion (-relation) in this book to mean the $E(\boldsymbol{k})$ or $\omega(\boldsymbol{k})$ relation.

5.4.1 Dielectric Function and the Polariton States with Spatial Dispersion

The dielectric function given in (4.22) for the simple case of an isolated resonance has to be modified if we want to take spatial dispersion effects into account. The eigenfrequency ω_0 has to be replaced by $\omega_0(\boldsymbol{k})$ and the oscillator strength f and the damping γ may also depend on $\boldsymbol{k}$ resulting in

$$\varepsilon(\omega, \boldsymbol{k}) = \varepsilon_\mathrm{b} = \frac{f(\boldsymbol{k})}{\omega_0^2(\boldsymbol{k}) - \omega^2 - \mathrm{i}\omega\gamma(\boldsymbol{k})} \tag{5.15}$$

The most significant change is the fact that ε is now a function of two independent variables, ω and $\boldsymbol{k}$.

Along with the transverse eigenfrequency $\omega_0(\boldsymbol{k})$ the longitudinal eigenfrequency ω_L defined as $\varepsilon(\omega_\mathrm{L}) = 0$ also becomes $\boldsymbol{k}$-dependent, i.e., $\omega_\mathrm{L} = \omega_\mathrm{L}(\boldsymbol{k})$. The same is true for the longitudinal-transverse splitting Δ_LT which is connected with $f(\boldsymbol{k})$. In principle all of the above quantities have to be given as a function of $\boldsymbol{k}$.

What one usually does is to give – if possible – an analytic expression for $\omega_0(\boldsymbol{k})$ and to still consider f and γ as $\boldsymbol{k}$-independent. Though there is clear experimental evidence that f and γ depend on $\boldsymbol{k}$(see e.g. [82O1, 83M1, 84S1, 85K1, 89R1, 91R1] and references therein), these dependences are usually less critical for the correct description of the optical properties of semiconductors than the $\boldsymbol{k}$-dependence of ω_0 and will be neglected in the following. For $\omega_0(\boldsymbol{k})$ we again use for simplicity a parabolic relation, which monitors nicely the onset of the dispersion relation in Fig. 4.2b but also that of excitons (see Chap. 9). But we stress once more that the consequences are qualitatively similar for other relations. We have

$$\varepsilon(\omega, \boldsymbol{k}) = \varepsilon_\mathrm{b} + \frac{f(\boldsymbol{k})}{\omega_0^2 + 2\omega_0 A\boldsymbol{k}^2 - \omega^2 - \mathrm{i}\omega\gamma}\,. \tag{5.16a}$$

with

$$\omega_0^2(\boldsymbol{k}) = \left(\omega_0 + A\boldsymbol{k}^2\right)^2 = \omega_0^2 + 2\omega_0 A\boldsymbol{k}^2 + A^2\boldsymbol{k}^4 \approx \omega_0^2 + 2\omega_0 A\boldsymbol{k}^2$$
$$\text{for} \quad |\boldsymbol{k}| \ll \pi/a\,. \tag{5.16b}$$

The approximation used in (5.16b) is usually valid for massive and for effective mass particles (see Sect. 8.5) like excitons.

To determine the dispersion relation of the polariton we must again combine the polariton equation with the dielectric function resulting again in an implicit relation for $\omega_0(\boldsymbol{k})$:

$$\frac{c^2\boldsymbol{k}^2}{\omega^2} = \varepsilon_\mathrm{b} + \frac{f}{\omega_0^2 + 2\omega_0 A\boldsymbol{k}^2 - \omega^2 - \mathrm{i}\omega\gamma}\,. \tag{5.16c}$$

We consider first the case of vanishing damping in Fig. 5.9a and start with the real part of $\boldsymbol{k}$. The transverse lower polariton branch starts (LPB)

photon-like and then bends over to asymptotically approach the parabolic dispersion relation of the resonance. If there is a constant Δ_{LT}, i.e., $f \neq f(\boldsymbol{k})$, then the longitudinal branch shown by the dashed line in Fig. 5.9a starts at ω_{L} and is then essentially parallel to the exciton-like part of the LPB. At ω_{L} the transverse upper polariton branch (UPB) also begins, going over into a photon-like (i.e., linear) dispersion relation with slightly steeper slope, as shown already in Fig. 5.1a. The imaginary part of $\boldsymbol{k}$ starts at ω_{L} and bends downwards to reach asymptotically a curve which is produced by reflecting the transverse eigenfrequency $\omega_0(\boldsymbol{k})$ through the point $(\omega_0, \boldsymbol{k} = 0)$. This means that the UPB has a purely imaginary continuation below ω_{L}.

If we include a small but finite damping, we end up with the situation shown in Fig. 5.9b. Compare to Fig. 5.1a without spatial dispersion. It can be seen that the LPB and the UPB now extend over the whole energy range. They have both a real and an imaginary part at all frequencies. The imaginary part of the LPB is very small below ω_0, peaks between ω_0 and ω_{L}, and

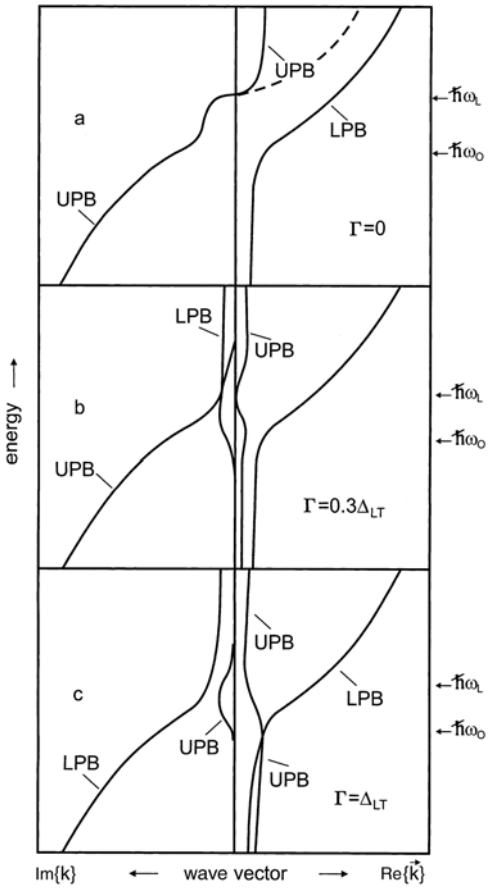


Fig. 5.9. The real and imaginary parts of a polariton dispersion relation in the vicinity of a resonance with spatial dispersion for vanishing (**a**) small $\hbar\Gamma < \Delta_{\mathrm{LT}}$) (**b**) and strong $\hbar\Gamma \approx \Delta_{\mathrm{LT}}$) damping (**c**). After [83M1]

decreases gradually for higher energies. The UPB has a very small imaginary part above ω_L which increases continuously below, but there is also a small real part of the UPB below ω_L. The problem of negative group velocity in the spectral region of the resonance, mentioned already in Sect. 5.2, has become less severe and appears only in the form of a small hump in the dispersion relation for finite damping. For the case of strong damping shown in Fig. 5.9c the role of the LPB is gradually taken over by the UPB below ω_0 [83M1] and one approaches a situation similar to the case without spatial dispersion (compare with Fig. 5.1a). In other words this means that one can describe the optical properties of a strongly damped resonance by neglecting the influence of spatial dispersion. However, it should be stressed, that this effect is still present in principle, but the strong damping is what dominates the optical properties.

Even with strong damping there remains at least one transversal branch that extends over the whole Brillouin zone in contrast to Fig. 5.1a for $\gamma \neq 0$ and possibly also a longitudinal one.

These branches are the ones observed, e.g., in inelastic neutron scattering when determining the dispersion relation of TO- and LO-phonons in the whole Brillouin zone.

5.4.2 Reflection and Transmission and Additional Boundary Conditions

From the polariton dispersion with spatial dispersion shown in Fig. 5.9 we can easily recover the real and imaginary parts of $\tilde{n}$ by just reversing the procedure given in (5.8). We will not spend time on this procedure (See e.g. the formulae given in [75L1, 78B1, 78H1, 78S1, 79B1, 81L1, 82R1, 82S1] but discuss directly the optical spectra, especially the consequences of spatial dispersion on the reflection as compared to the situation of Fig. 4.5 where spatial dispersion was still neglected. The two most important points are, first, that there is no more strict stop-band, i.e., there is at least one propagating mode for every frequency ω (with and without damping) and, second, that in some spectral regions there is more than one propagating mode. This second point is especially obvious for $\omega > \omega_L$ in Fig. 5.9a.

We now want to discuss the consequences of these two new phenomena arising from spatial dispersion on the optical spectra, especially on the reflection spectrum.

The fact that we have at least one propagating mode – generally with real and imaginary part – for all frequencies even between ω_T and ω_L, means that the reflectivity for (normal) incidence no longer reaches unity, even for the case of vanishing damping, as sketched in Fig. 5.10.

The fact that we have more than one mode (propagating or evanescent, i.e., $\mathrm{Im}(\boldsymbol{k}) \ll \mathrm{Re}(\boldsymbol{k})$ or $\mathrm{Im}(\boldsymbol{k}) \gg \mathrm{Re}(\boldsymbol{k})$ respectively, in the solid for one frequency means that the two independent boundary conditions deduced from Maxwell's equations in connection with (3.6), (3.7) and Figs. 3.1, 3.2 are no

longer sufficient. For a given incident wave we could deduce the amplitudes of one reflected and one transmitted wave. If there are two or even more states in the medium at the same frequency coupling to the incident field, we need one or more additional boundary conditions (abc). To make the situation clear, we show in Fig. 5.11 the wave vectors for such a case for normal and oblique incidence. The incident and reflected beams obey the usual law of reflection, their components parallel to the surface are equal. The same is true for the transmitted beams, but the total length of the wave vectors on the LPB and the UPB are different, in agreement with Fig. 5.9. Obviously the two beams are travelling in different spatial directions and this is a reason why the $\boldsymbol{k}$-dependence of ω_0 is called "spatial dispersion". Though Fig. 5.11 has some similarity with the picture for birefringence (See Sect. 3.1.7) we point out that the reasons are quite different. In Fig. 3.12 the o and eo beams have orthogonal polarization and the phenomenon needs uniaxial or lower symmetry. Spatial dispersion occurs independent of the symmetry and even for cubic symmetry and the two beams are polarized in the same direction. Spatial dispersion also occurs for crystals of lower symmetry which then may show birefringence in addition. In this case the dispersion curves of Fig. 5.9 have to be drawn twice with different parameters for the o and the eo beams.

As already mentioned, the abc cannot be deduced from Maxwell's equations. Their capacity is exhausted with one reflected and one transmitted beam. Since the complex index of refraction around the resonance is rather different for the LPB and the UPB, which therefore contribute differently to the reflection spectrum according to (3.20), the abc should contain information about the "branching ratio", i.e., which fractions of the incident beam couple in the medium to the LPB and to the UPB as a function of frequency.

The abc are somewhat arbitrary (we shall explain later on why) and are based mainly on arguments of physical plausibility. On the vacuum side of the

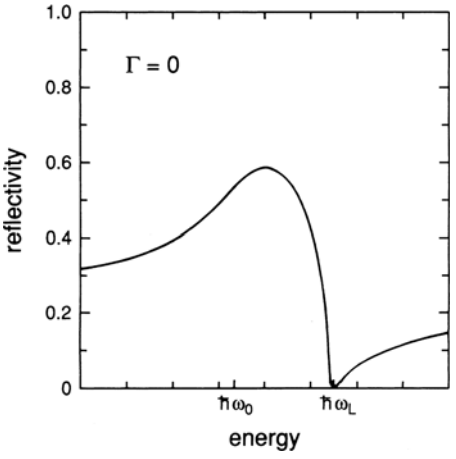


Fig. 5.10. Sketch of the reflection spectrum in the vicinity of a resonance with spatial dispersion and without damping. Compare with Fig. 4.5. After [82R1]

interface the polarization is trivially zero. To avoid an unphysical discontinuity in the polarization, one possible abc is that the polarization of the medium must be zero at the interface

$$\boldsymbol{P}(z=0)=0\,. \tag{5.17a}$$

Another argument says that the polarization should vary smoothly across the interface, implying that the derivative with respect to the normal direction has to vanish, resulting in

$$\left.\frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}z}\right|_{z=0}=0\,. \tag{5.17b}$$

Some authors favor a linear combination of the two conditions:

$$\boldsymbol{P}|_{z=0}+\beta\left.\frac{\partial\boldsymbol{P}}{\partial z}\right|_{z=0}=0 \quad \text{with} \quad -1\le\beta\le+1\,. \tag{5.17c}$$

The reflection spectrum shown in Fig. 5.10 is actually calculated for an exciton resonance using the material parameters of CdS and the abc (5.10), the so-called Pekar–Hopfield abc. See also Chap. 13 and [62P1, 74A1, 75L1, 78B1, 78H1, 78S1, 79B1, 79S1, 81B1, 81L1, 81S1, 82O1, 82R1, 82S1, 83M1, 84H1, 84S1, 85H1, 89R1, 91R1] for further details of the abc.

It turns out that experimentally observed spectra, e.g., of exciton resonances, can be fitted with all the above-mentioned abc, but with slightly different values for the other parameters, such as f and γ, that describe the resonance.

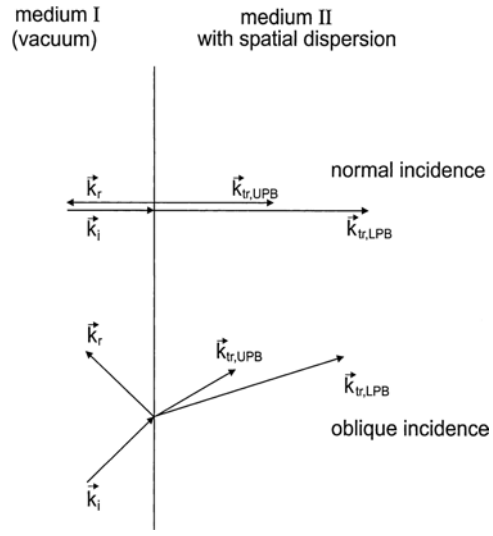


Fig. 5.11. The wave vectors of a resonance with spatial dispersion for $\omega > \omega_{\mathrm{L}}$. Note that there is more than one propagating wave in the medium

As a rule of thumb, one can state for all abc and weak damping that the light propagating in matter at frequencies ω sufficiently below the transverse eigenfrequency and above the longitudinal one travels almost completely on the LPB and on the UPB, respectively. "Sufficiently" means in this context

$$\hbar\,|\omega - \omega_{\mathrm{T,L}}| \gtrsim 10\Delta_{\mathrm{LT}}\ . \tag{5.18}$$

The crucial spectral range where spatial dispersion and the problem of abc are of importance is thus the resonance and its vicinity.

Now let us have a look at the transmission, including spatial dispersion. Since both the LPB and the UPB have substantial imaginary parts in the vicinity of the resonance, we again expect a dip in the transmission in the region around the resonance. At considerably lower frequencies, the light couples almost completely to the LPB – which has an almost negligible imaginary part in this region – so that the sample is transparent. The same is true significantly above ω_{L} for the UPB, at least if there are no other resonances. The shape of the effective absorption coefficient $\alpha(\omega)$ or of $\kappa(\omega)$ can look somewhat different from Fig. 4.4 possibly developing a spike at the longitudinal eigenfrequency, as indicated in Fig. 5.12. In addition, one can see from Fig. 5.9 that the imaginary parts of $\boldsymbol{k}$ and thus κ or α are different for the two polariton branches. The amplitudes and light intensities of each polariton branch therefore decay exponentially with thickness, but the sum of both, which is the only experimentally accessible quantity, may show a non-exponential decay with sample thickness. Furthermore the damping may in some cases be higher close to the surface compared to the bulk of the sample due to imperfections introduced into the lattice close to the surface.

Strong damping, i.e., $\hbar\gamma \geq \Delta_{\mathrm{LT}}$, reduces the importance of the influence of spatial dispersion on the optical spectra as already mentioned above. The resonance in the reflection spectrum is then already so strongly washed out by damping that the details, whose description relies on spatial dispersion, are no longer observable.

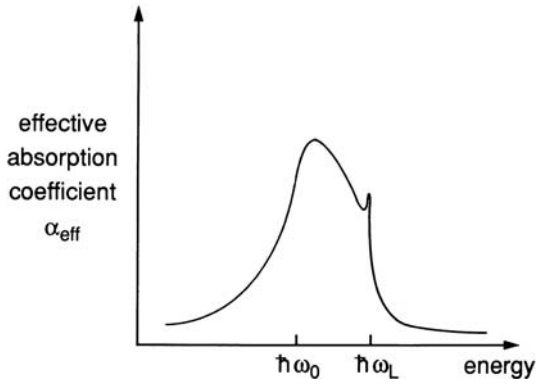


Fig. 5.12. A possible spectrum of the effective absorption coefficient in the vicinity of a resonance with spatial dispersion ([82R1])

It is physically not completely satisfying that the choice of the abc is to some extent arbitrary (5.17a–c). The parameters of the resonance deduced with the use of different abc are (slightly) different. On the other hand, the value of quantities such as f cannot depend on the abc chosen by the physicist running the fitting program. The answer is that the problem of the abc is an artificial one. The dielectric function $\varepsilon(\omega, \boldsymbol{k})$ describes the optical properties in the bulk of the sample. In our derivation of $\varepsilon(\omega, \boldsymbol{k})$ in Chap. 4 no surface was included. Later on, however, we use this dielectric function to describe the optical properties of the interface between two semi-infinite half-spaces, one usually being vacuum, the other the semiconductor under consideration. The price we have to pay for this "improper" use of the dielectric function is the problem of the abc. If we were to calculate the optical properties of a half-space from the outset, the problem of the abc could be avoided. Indeed, some calculations have used this idea [79S1, 81B1, 81S1, 82S1] or [98H1, 99H2]. However, this procedure has its own problems. For example, the band-structure and the exciton states have to be calculated for the half-space, which means that we can in principle no longer use Bloch's theorem for the direction normal to the interface. This causes complications which can be overcome only with difficulty and by using various approximations and simplifications. Therefore most authors prefer to use the bulk dielectric function and some of the abc to evaluate the optical spectra. More information about the problem of abc and the rather lengthy formulas for calculating the spectra of reflection can be found in the references given already above.

More recently a very interesting solution to the problem of abc's was put forward by [98H1], which triggered some discussion [99H2], but allows one to avoid the abc-problem.

5.5 Real and Imaginary Parts of Wave Vector and Frequency

Until now, we have assumed that we can describe the light wave or polariton wave propagating in matter by a real frequency and a wave vector which has a real and an imaginary part according to the complex index of refraction. In principle, however, one could take $\boldsymbol{k}$ to be real and introduce a complex frequency $\tilde{\omega}$ by including the damping term $\mathrm{i}\omega\gamma$ as in (4.14), (4.22):

$$\tilde{\omega} = \omega - \mathrm{i}\gamma \tag{5.19}$$

At first glance, there is no reason to prefer one approach over the other. Actually both approaches are possible and it is the experiments performed that decide which model is the more appropriate one.

If we shine a monochromatic wave with well-defined frequency ω (e.g., a spectrally narrow laser beam) on the sample, or if we select such a frequency with a monochromator, then we have the situation of (4.14), (4.32)

i.e., a forced oscillation, and here we have to use a purely real ω, but a complex $\boldsymbol{k}$ to describe the decrease in amplitude of the polariton wave as it travels through the crystal.

If, on the other hand, we could by some means create at a certain time (e.g., $t = 0$), a polariton wave in the sample with constant amplitude everywhere and let it evolve for $t > 0$, then the amplitude would decay with time, but would remain the same everywhere in space. This is just the situation described by a real $\boldsymbol{k}$ only and a complex $\tilde{\omega}$. The reader might think that a wave with constant amplitude from $-\infty$ to $+\infty$ in space is highly unphysical, but in fact this approximation is as valid as that of a strictly monochromatic wave, which necessarily endures from $t = -\infty$ to $t = +\infty$. Everything with a finite temporal duration has a finite spectral halfwidth.

These are just consequences of the fact that the time t and frequency ω domains and the space $\boldsymbol{r}$ and wave vector $\boldsymbol{k}$ domains are connected with each other by one and three-dimensional Fourier transforms, respectively.

Multiplication of ω or $\boldsymbol{k}$ with $\hbar$ then immediately gives the "uncertainty relations" for energy and time or momentum and space (3.37).

Since the experimental situation discussed first is much more frequently used than the second one, we will restrict ourselves for the rest of this book to the situation of complex $\boldsymbol{k}$ and real ω. Bearing in mind, however, that there is a third uncertainty relation for truth and understandability of a text as stated at the beginning of the book, we should inform the reader that, in principle, one needs to use both a complex $\boldsymbol{k}$ and a complex ω since all excitations usually have a finite lifetime or phase relaxation time. Chapter 23 gives further details on this topic.

5.6 Surface Polaritons

For almost every wave-like excitation in the bulk of a solid or even liquid sample, there exists a surface or interface mode including Rayleigh-waves in earthquakes or the usual waves on the surface of water. There are, e.g., surface acoustic phonons, surface plasmons, etc. For recent reviews see, e.g., [82M1]a of Chap. 1 or [98B2] and the references therein.

Here we want to say a few more words about the surface polaritons already mentioned briefly in Sect. 2.4.

Surface polaritons are also quanta or quasiparticles of the mixed state of an electromagnetic and a polarization wave. They are distinguished from bulk modes by the fact that they can only propagate along the interface between two different media. The amplitudes decay exponentially with distance from the interface on both sides, as shown schematically in Fig. 5.13a, i.e., surface polaritons are evanescent waves on both sides of the interface, in contrast to the one-sided evanescent wave in the case of total internal reflection of Fig. 3.3a. See also Fig 4.4. For every volume polariton there exists a surface polariton.

We want now to discuss the conditions for which surface polaritons can exist. For simplicity we restrict ourselves to the case of vanishing damping and no spatial dispersion: $\gamma = 0$ and $A = 0$ in (5.16). We assume that the interface is formed by an essentially non-dispersive medium I described by a constant real index of refraction $n_{\mathrm{I}}^2 = \varepsilon_{\mathrm{I}}$ on one side (for vacuum $n_{\mathrm{I}} = 1$) and the medium under consideration with $\varepsilon(\omega) = \tilde{n}^2(\omega)$ on the other (medium II). If the surface polariton cannot propagate into medium I or II, there must be some physical reasons preventing its decay by radiating into the halfspaces I and II.

As a first condition we may state that in medium II there are no propagating waves between the transverse and longitudinal eigenfrequencies ω_{T} and ω_{L}, as discussed for example in connection with Figs. 4.4, 4.5 and 5.1a. The propagation into medium I can be excluded if the wave vector $\boldsymbol{k}_{\mathrm{s}}$ of the surface polariton, which is directed parallel to the interface, is larger than $\boldsymbol{k}_{\mathrm{I}}$ of any wave propagating in medium I. Under such a condition the conservation law for $\boldsymbol{k}$ parallel to the interface results with Fig. 5.13b in

$$\boldsymbol{k}_{\mathrm{s}}^2 + \boldsymbol{k}_{\perp}^2 = \boldsymbol{k}_{\mathrm{I}}^2 \quad \text{and} \quad \boldsymbol{k}_{\mathrm{s}}^2 > \boldsymbol{k}_{\mathrm{I}}^2 \Rightarrow \boldsymbol{k}_{\perp}^2 < 0. \tag{5.20}$$

The r.h.s of (5.20) simply says that $\boldsymbol{k}_{\perp}$ is purely imaginary and this is what we need for an evanescent wave normal to the interface.

We can summarize these arguments by saying that surface polaritons can be expected in the spectral region given by

$$\omega_0 \equiv \omega_{\mathrm{T}} \leq \omega_{\mathrm{s}} \leq \omega_{\mathrm{L}} \quad \text{in medium II and} \tag{5.21a}$$

$$\boldsymbol{k}_{\mathrm{s}} \geq n_{\mathrm{I}} \omega c^{-1}, \quad \text{in medium I} \tag{5.21b}$$

or

$$\mathrm{Re}\{\varepsilon_{\mathrm{II}}(\omega)\} < 0 \quad \text{and} \quad |\mathrm{Re}\{\varepsilon_{\mathrm{II}}(\omega)\}| > \varepsilon_{\mathrm{I}}. \tag{5.21c}$$

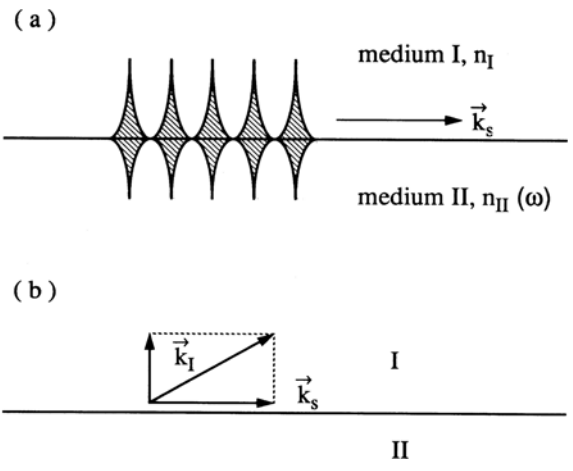


Fig. 5.13. Schematic drawing of the decay of the amplitudes of a surface polariton propagating along an interface (**a**) and a sketch to illustrate the derivation of (5.20) (**b**)

The dispersion relation of surface polaritons $\omega_s(\boldsymbol{k}_s)$ can be deduced for the present assumptions from the boundary conditions given by Maxwell's equations. We will not go through the procedure here, but merely give the result and refer the reader for its derivation to the literature e.g. [82M1]a,d of Chap. 1 or [74O1, 81L1] and references therein.

$$\boldsymbol{k}_s = \left(\frac{\varepsilon_I \cdot \varepsilon_{II}(\omega)}{\varepsilon_I + \varepsilon_{II}(\omega)} \right)^{1/2} \frac{\omega}{c}. \tag{5.22}$$

We should note that $\varepsilon_{II}(\omega)$ is negative in the region of (5.21a). In order to get a real value of $\boldsymbol{k}_s$ we evidently have, in addition to (5.21a,b) to fulfill also (5.21c).

The polarization of the surface polaritons is as follows. If the interface is the xy-plane and the surface wave propagates in the x-direction, i.e., $\boldsymbol{k}_s \| x$ then the electric field is in the xz-plane and the magnetic induction is along the y-axis.

For large $\boldsymbol{k}_s$ one finds that the surface polariton occurs at a frequency ω_s where

$$\mathrm{Re}\left\{\varepsilon_{II}\left(\omega_s\right)\right\} = -1 \quad \text{for} \quad k_s \gg \frac{\omega}{c} \quad \text{and} \quad \varepsilon_I = 1\,. \tag{5.23}$$

In the case of a simple plasmon we have

$$\varepsilon(\omega) = \varepsilon_b + \frac{\omega_{PL}^2}{-\omega^2 - i\omega\gamma} \tag{5.24}$$

as will be shown later in Chap. 10. For simplifying conditions, $\gamma \Rightarrow 0$ and $\varepsilon_b = 1$ (the latter being realistic only for some simple metals), we find from (5.24) the frequently given relation [80R1, 88R1]

$$\omega_s = \omega_{PL}/\sqrt{2} \quad \text{for large} \quad \mathrm{k_s}\,. \tag{5.25}$$

In Fig. 5.14 we show as a summary of this chapter the dispersion of polaritons and of surface polariton for a resonance without (a), and with (b), spatial dispersion and vanishing damping. In the latter case the calculated dispersion relation $\omega_s(\boldsymbol{k}_s)$ is slightly influenced by the abc used. The condition (5.23) is reached for large $\boldsymbol{k}_s$ generally slightly below the longitudinal branch as seen, e.g., in Fig. 4.3. The dispersion relation for large $\boldsymbol{k}_s$ is parallel to the longitudinal branch without reaching it. For more details see [74O1, 81L1].

To conclude this section we should briefly stress one point. Since surface polaritons cannot propagate into medium I (generally vacuum) they cannot be created by shining light of an appropriate frequency on the sample. The same is true for the other side. As a consequence, it is not possible to excite surface polaritons directly. A frequently used method involves attenuated total reflection. This technique will be outlined briefly in Sects. 11.1.5, 12.2, 13.1.5 and 25.1.

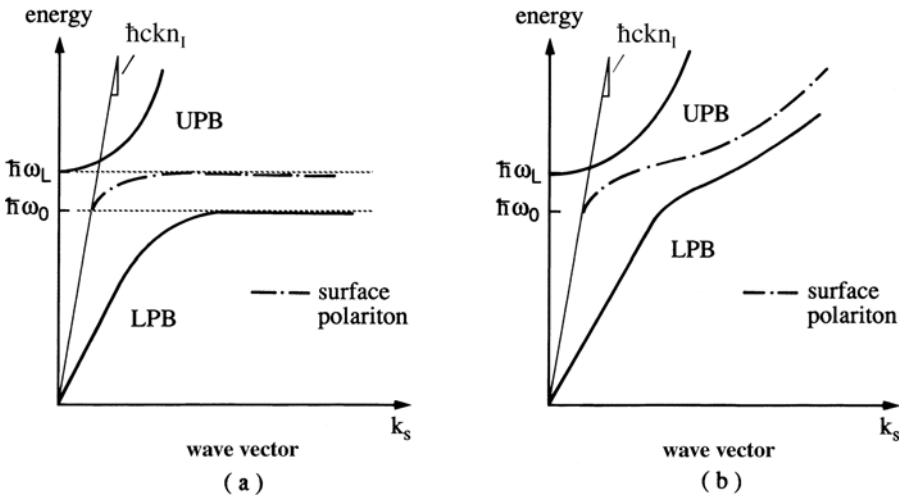


Fig. 5.14. The dispersion of the surface polariton for a resonance without (**a**), and with (**b**), spatial dispersion. The damping is assumed to be negligible in both cases.

5.7 Problems

1. What dispersion relations would you expect for the polariton resulting from oscillators with the dispersion relations of Figs. 4.2a and b? Do not forget that a finite coupling between photons and the oscillator necessarily implies a finite Δ_{LT}.
   Check if you were right when you come to the chapters on phonon and on exciton polaritons.
2. Calculate the frequency shift a photon experiences when it is scattered off an atom in a backward direction. Compare with the homogenous linewidth of luminescence lines in semiconductor optics, which hardly fall below 0.1 meV.
3. Inform yourself about the possibility of cooling atoms by absorption and emission of photons.
4. What is the Mösbauer effect? How does it work?
5. In Na vapor it is possible to slow light down to an almost complete stop. Inform yourself with the help of some literature. Effects apart from the extremely flat dispersion relation for large k-vectors contribute to this phenomenon.
6. Calculate the dispersion relation $\omega(\boldsymbol{k})$ from (5.7) for vanishing damping. What changes if $\omega_0 = \omega_0(\boldsymbol{k})$ and/or if a small but finite damping γ are introduced?
7. Write down the equations of motion of two coupled harmonic oscillators and try to solve them. Find or imagine examples in classical physics and in quantum mechanical systems.

8. Which dependence of $\Delta_{\mathrm{LT}}(\boldsymbol{k})$ do you expect for a transition that is dipole forbidden but allowed in quadrupole approximation.
9. Can you give an intuitive explanation as to why a photon with spin $\pm\hbar$ can excite a quadrupole transition, e.g., from an atomic s state to a d (or s) state? Hint: Place the atom in the origin of the coordinate system and vary the impact parameter of the photon.
10. Sketch the dispersion of a polariton resonance with spatial dispersion and an oscillator strength which increases with $\boldsymbol{k}$. (Assume for simplicity zero damping). Does Δ_{LT} then also depend on $\boldsymbol{k}$?
11. Sketch the dispersion of the polariton for two close lying resonances A and B, with and without spatial dispersion for an order of the energies at $\boldsymbol{k}=0$ $\hbar\omega_0^A < \hbar\omega_{\mathrm{L}}^A < \hbar\omega_0^B < \hbar\omega_{\mathrm{L}}^B$. Is it possible for a single orientation of the polarization to have the sequence $\hbar\omega_0^A < \hbar\omega_0^B < \hbar\omega_{\mathrm{L}}^A < \hbar\omega_{\mathrm{L}}^B$?
12. Find out something about the method of attenuated total reflection (ATR) from Chap. 26. Compare this method with the "optical tunnel effect" of Fig. 3.3b.
13. Apart from the use of ATR methods, it is possible to excite surface polaritons optically if a periodic structure, i.e., a grating, is formed at the interfaces. What is the principle behind this? Compare this with the statements about momentum conservation in Sects. 3.1.3 and 9.2.

References to Chap. 5

[02W1] J. Wood: Phil. Mag. **3**, 128 (1902) and Phys. Zeitschrift **3**, 230 (1902)
[03L1] O. Lummer, E. Pringsheim: Phys. Zeitschrift **4**, 430 (1903)
[09M1] Müller-Pouillets: Lehrbuch der Physik und Meteorologie 2. Band, 3. Buch p 227 Vieweg, Braunschweig (1909) and references given therein
[58H1] J.J. Hopfield: Phys. Rev. **112**, 1555 (1958)
[62P1] S.J. Pekar: Sov. Phys. Sol. State **4**, 953 (1962)
[63H1] J.J. Hopfield, D.G. Thomas: Phys. Rev. **132**, 563 (1963)
[74A1] G.S. Agarwal: Phys. Rev. B **10**, 1447 (1974)
[74O1] A. Otto: Festkörperprobleme/Adv. Solid State Phys. **14**, 1 (1974)
[75C1] R. Claus, L. Merten, J. Brandmüller: Light Scattering by Phonon Polaritons, Springer Tracts Mod. Phys. **75**, Springer, Berlin, Heidelberg (1975)
[75L1] J. Lagois, K. Hümmer: phys. stat. sol. (b) **72**, 393 (1975)
[78B1] I. Broser et al.: phys. stat. sol. (b) **90**, 77 (1978)
[78H1] K. Hümmer, P. Gebhardt: phys. stat. sol (b) **85**, 271 (1978)
[78S1] W. Stössel, H.J. Wagner: phys. stat. sol. (b) **89**, 403 (1978)
[79B1] I. Broser, M. Rosenzweig: phys. stat. sol. (b) **95**, 141 (1979)
[79M1] Y. Masumoto et al.: J. Phys. Soc. Japan **47**, 1844 (1979)
[79S1] A. Stahl, Ch. Uhilein: Festkörperprobleme/Adv. Solid State Phys. **XIX**, 159 (1979)
[79U1] R.G. Ulbrich, G.W. Fehrenbach: Phys. Rev. Lett. **43**, 963 (1979)
[80B1] I. Broser, M. Rosenzweig: Phys. Rev. B **22**, 2000 (1980)
[80R1] H. Raether: Excitation of Plasmons and Interband Transitions by Electrons, Springer Tracts in Mod. Phys. **88** (1980)

[81B1] I. Balslev: Phys. Rev. B **23**, 3977 (1981)
[81I1] T. Itho et al.: Solid State Commun. **37**, 925 (1981)
[81L1] J. Lagois: Phys. Rev. B **23**, 5511 (1981)
[81L2] J. Lagois: Oberflächenpolaritonen, Habilitation Thesis, Erlangen (1981) and references therein
[81S1] A. Stahl: phys. stat. sol.(b) **106**, 575 (1981)
[82O1] Y. Onodera: J. Phys. Soc. Jpn. **51**, 2194 (1982)
[82R1] M. Rosenzweig: Excitonische Polaritonen – optische Eigenschaften räumlich dispersiver Medien, Dissertation, Berlin (1982)
[82S1] A. Stahl, I. Balslev: phys. stat. sol. (b) **111**, 531 (1982) and ibid. **113**, 583 (1982)
[82S2] T. Skettrup: phys. stat. sol. (b) **109**, 663 (1982)
[83M1] M. Matsushito, J. Wicksted, H.Z. Cummins: Phys. Rev. B **29**, 3362 (1983)
[83S1] Y. Segawa et al.: J. Phys. Soc. Japan **52**, 3664 (1983)
[84A1] V.M. Agranovich, V.L. Ginzburg: Crystal Optics with Spatial Dispersion and Excitons, Springer Ser. Solid-State Sci. **42**, 2nd edn., Springer, Berlin Heidelberg (1984)
[84H1] P. Halevi, R. Fuchs: J. Phys. C. **17**, 3869 and 3889 (1984)
[84R1] J.E. Rothenberg, D. Grischkovsky, A.C. Balant: Phys. Rev. Lett. **53**, 552 (1984)
[84S1] T. Shigenari, X.Z. Lu, H.Z. Cummins: Phys. Rev. B **30**, 1962 (1984)
[85H1] B. Hönerlage et al.: Phys. Rep. **124**, 161 (1985)
[88R1] H. Raether: Surface Plasmons on Smooth and Rough Surfaces and on Gratings, Springer Tracts in Modern Physics **111**, Springer Berlin (1988)
[89R1] V. Ya Reznichenko, M.I. Strashnikova, V.V. Cherny: phys. stat. sol. (b) **152**, 675 (1989)
[91F1] D. Fröhlich et al.: Phys. Rev. Lett **67**, 2343 (1991)
[91R1] V. Ya Reznichenko, M.I. Strashnikova, V.V. Cherny: phys. stat. sol.(b) **167**, 311 (1991)
[94R1] G. Raithel: Phys. Bl. **50**, 1149 (1994) and references given therein
[96H1] T. Ha et al.: Phys. Rev. Lett. **77**, 3979 (1996)
[97N1] S. Nüsse et al.: Phys. Rev. B **55**, 4620 (1997)
[97S1] A.J. Sievers: in [81A1]h of Chap. 1, p 227
[98B1] R. v. Baltz, C. Klingshirn in [81A1]i of Chap. 1, p 381
[98B2] G. Benedek in [81A1]i of Chap. 1, p 295
[98H1] T. Ha et al.: Phys. Rev. Lett. **80**, 2093 (1998)
[98H2] K. Henneberger: Phys. Rev. Lett. **80**, 2889 (1998)
[98J1] A. Jolk, C. Klingshirn, R. v. Baltz in [81A1]i of Chap. 1 p 397
[98K1] C. Klingshirn, in [81A1]i of Chap. 1, p 143
[99B1] U. van Bürck: Hyperfine Interaction **123/124**, 483 (1999) and in Nuclear Resonant Scattering of Synchrotron Radiation, E. Gerdau and H. de Waavel, eds, Baltzer, Oxford (1999)
[99H1] L.V. Hau et al.: Nature **397**, 594 (1999)
[99H2] K. Henneberger: Phys. Rev. Lett. **83**, 1265 (1999)
[99K1] M.M. Kash et al.: Phys. Rev. Lett. **82**, 5229 (1999)
[99N1] D.F. Nelson, B. Chen: Phys. Rev. Lett. **83**, 1263 (1999)
[99Z1] R. Zeyher: Phys. Rev. Lett. **83**, 1264 (1999)
[00K1] C. Klingshirn: Physik in unserer Zeit **31**, Issue 4, p 144 (2000)
[01K1] O. Kocharovskaya, Y. Rostovtsev, M.O. Scully: Phys. Rev. Lett. **86**, 628 (2001)
[04F1] D. Fröhlich et al., Solid State Communications, to be published (2004)

6

Kramers–Kronig Relations

In this chapter we want to investigate some general relations between the real and imaginary parts of $\tilde{n}$ or ε. For more details see e.g. [72A1, 78H1, 95P1] or [72W1, 82L1, 90K1, 96Y1] of Chap. 1 and further references given therein.

6.1 General Concepts

We stated in Sect. 2.3 that the susceptibility χ and the dielectric function $\varepsilon = \chi + 1$ are response functions of the medium which describe the response (in this case the polarization) to a stimulus (in this case the incident electric field) for the special case of an incident monochromatic wave with frequency ω. We now leave ω- and $\boldsymbol{k}$-space for a moment and go to t and $\boldsymbol{r}$ space, i.e., to time and real space.

The most general expression for a linear response function is

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(\boldsymbol{r},t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \chi\left(\boldsymbol{r},\boldsymbol{r}',t,t'\right)\boldsymbol{E}\left(\boldsymbol{r}',t'\right)\,\mathrm{d}t'\,\mathrm{d}\boldsymbol{r}'. \tag{6.1}$$

This means that the polarization $\boldsymbol{P}$ at point $\boldsymbol{r}$ and time t depends on the electric field at all other places and at all times.

Similar arguments can be given for the current density $\boldsymbol{j}\left(\boldsymbol{r},t\right)$ replacing the susceptibility by the electric conductivity $\sigma\left(\boldsymbol{r},\boldsymbol{r}',t,t'\right)$. We are now going to simplify (6.1) in various steps.

First we assume that the sample is homogeneous in time, i.e., its properties do not depend on t explicitly. Then χ depends only on the time difference $t - t'$. Since our medium consists of atoms, it is not homogenous in space, but if we assume that all wavelengths present in $\boldsymbol{E}\left(\boldsymbol{r}',t'\right)$ are much longer than the lattice constant, then an analogous approach holds for $\boldsymbol{r} - \boldsymbol{r}'$. This is essentially the same assumption as made in the dipole approximation (3.2).

Equation (6.1) then transforms into

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(\boldsymbol{r},t) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} \chi\left(\boldsymbol{r}-\boldsymbol{r}',t-t'\right)\boldsymbol{E}\left(\boldsymbol{r}',t'\right)\mathrm{d}t'\,\mathrm{d}\boldsymbol{r}'. \tag{6.2}$$

The response function χ is said to be non-local, i.e., the polarization at $\boldsymbol{r}$ also depends on the electric field at other places $\boldsymbol{r}'$. In other words, a polarisation created at one place $\boldsymbol{r}'$ by the electric field at this place contributes at a later time to the polarisation at another place $\boldsymbol{r}'$. The finite propagation time is known as retardation.

This is just the phenomenon of spatial dispersion, as becomes clear by inspecting the models of Fig. 4.1 or of Sect. 5.2.

The integral (6.1),(6.2) is a convolution both in space and in time. It simplifies to a product under Fourier transformation (see [98B1,98D1] of Chap. 2) that is one dimensional in time and three dimensional in space. Executing this Fourier transform we obtain

$$\frac{1}{\varepsilon_0}\boldsymbol{P}\left(\boldsymbol{k},\omega\right) = \chi\left(\boldsymbol{k},\omega\right)\boldsymbol{E}\left(\boldsymbol{k},\omega\right)\,. \tag{6.3}$$

We now simplify the expression (6.3) in various steps.

First we neglect spatial dispersion, i.e., we go from the situation in Fig. 5.8 back to Fig. 4.1. In this case the response function is local, i.e., $\boldsymbol{P}(\boldsymbol{r})$ depends only on $\boldsymbol{E}(\boldsymbol{r})$, reducing the dependence on $\boldsymbol{r}-\boldsymbol{r}'$ in (6.2) to a δ function $\delta\left(\boldsymbol{r}-\boldsymbol{r}'\right)$ and simplifying

$$\chi\left(\boldsymbol{k},\omega\right) \Rightarrow \chi\left(\omega\right) \tag{6.4}$$

or (6.5)

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(t) = \int\limits_{-\infty}^{+\infty} \chi(t-t')\boldsymbol{E}(t')\,\mathrm{d}t'\,. \tag{6.5}$$

Now we use a very important physical argument, namely causality. This argument is, in this context, valid both in classical physics and in quantum mechanics and means that the response $\boldsymbol{P}$ cannot come before the stimulus and thus

$$\chi(t-t') \equiv 0 \quad \text{for} \quad t' > t \tag{6.6}$$

or

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(t) = \int\limits_{-\infty}^{t} \chi(t-t')\boldsymbol{E}(t')\,\mathrm{d}t'\,. \tag{6.7}$$

We now execute the above mentioned Fourier transform with respect to time, resulting in

$$\boldsymbol{P}(\omega) = \int\limits_{-\infty}^{+\infty} \boldsymbol{P}(t)\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t\,, \tag{6.8a}$$

$$\boldsymbol{E}(\omega) = \int_{-\infty}^{+\infty} \boldsymbol{E}(t)\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t\ , \tag{6.8b}$$

$$\chi(\omega) = \int_{-\infty}^{+\infty} \chi(t-t')\,\mathrm{e}^{\mathrm{i}\omega(t-t')}\,\mathrm{d}t\ . \tag{6.8c}$$

Here we see that a complex $\chi(\omega)$ results from a real $\chi(t-t')$. Inserting (6.7) into (6.8a) results in

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(\omega) = \int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\omega t}\left[\int_{-\infty}^{t} \chi(t-t')\boldsymbol{E}(t')\,\mathrm{d}t'\right]\,\mathrm{d}t\ . \tag{6.9}$$

Introducing

$$1 = \mathrm{e}^{-\mathrm{i}\omega t'}\,\mathrm{e}^{\mathrm{i}\omega t'} \tag{6.10}$$

in the inner integral and rearranging the terms gives

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(\omega) = \int \boldsymbol{E}(t')\,\mathrm{e}^{\mathrm{i}\omega t'}\left[\int \chi(t-t')\,\mathrm{e}^{\mathrm{i}\omega(t-t')}\,\mathrm{d}t\right]\,\mathrm{d}t' = \chi(\omega)\boldsymbol{E}(\omega)\ . \tag{6.11}$$

This is identical to (2.27),(2.28).

With the knowledge of (6.11) we can apply Cauchy's theorem, which connects the real and imaginary parts of the Fourier transforms of analytic functions. This theorem leads us to

$$\varepsilon_1(\omega) - 1 = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\varepsilon_2(\omega')}{\omega' - \omega}\,\mathrm{d}\omega'$$

and

$$\varepsilon_2(\omega) = -\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\varepsilon_1(\omega') - 1}{\omega' - \omega}\,\mathrm{d}\omega', \tag{6.12}$$

where P in front of the integral means the principal value. Equation (6.12) can be rewritten as

$$\varepsilon_1(\omega) - 1 = \mathrm{Re}\{\chi(\omega)\} = \frac{2}{\pi} P \int_{0}^{\infty} \frac{\omega'\varepsilon_2(\omega)}{\omega'^2 - \omega^2}\,\mathrm{d}\omega'$$

and

$$\varepsilon_2(\omega) = \mathrm{Im}\{\chi(\omega)\} = -\frac{2\omega}{\pi} P \int_{0}^{\infty} \frac{\varepsilon_1(\omega') - 1}{\omega'^2 - \omega^2}\,\mathrm{d}\omega'. \tag{6.13}$$

since $\varepsilon(\omega) = \varepsilon^*(-\omega)$.

Similar relations hold for the phase $\phi(\omega)$ and amplitude $\rho(\omega)$ of the reflectivity $r(\omega)$ given for normal incidence by

$$r_\perp(\omega) = \frac{n(\omega) - 1 + \mathrm{i}\kappa(\omega)}{n(\omega) + 1 + \mathrm{i}\kappa(\omega)} = \rho(\omega)\,\mathrm{e}^{\mathrm{i}\phi(\omega)}. \tag{6.14}$$

Compare (6.14) with the expression for $R_\perp$ in (3.19).

$$\ln\{\rho(\omega)\} = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\phi(\omega')}{\omega' - \omega}\,\mathrm{d}\omega',$$

$$\phi(\omega) = -\frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\ln \rho(\omega')}{\omega' - \omega}\,\mathrm{d}\omega' = \frac{-2\omega}{\pi} P \int_0^{\infty} \frac{\ln \rho(\omega')}{\omega'^2 - \omega^2}\,\mathrm{d}\omega'. \tag{6.15}$$

Relations between $n(\omega)$ and $\kappa(\omega)$, i.e., between the real and imaginary parts of $\tilde{n}(\omega)$ can be deduced from the relations

$$\begin{aligned} &\mathrm{Im}\{\chi(\omega)\} = \varepsilon_2(\omega) = 2n(\omega)\kappa(\omega) \\ &\mathrm{Re}\{\chi(\omega)\} = \varepsilon_1(\omega) - 1 = n^2(\omega) - \kappa^2(\omega) - 1\,. \end{aligned} \tag{6.16}$$

For systems, where $\chi(\omega)$ is small compared to unity, the following approximate relations hold

$$\begin{aligned} n(\omega) - 1 &= \frac{2}{\pi} P \int_0^{\infty} \frac{\omega' \kappa(\omega')}{\omega'^2 - \omega^2}\,\mathrm{d}\omega' \\ \kappa(\omega) &= -\frac{2\omega}{\pi} P \int_0^{\infty} \frac{n_r(\omega')}{\omega'^2 - \omega^2}\,\mathrm{d}\omega'\,. \end{aligned} \tag{6.17}$$

Furthermore, the following relations hold [95P1].

$$\begin{aligned} n(\omega) &= \frac{1 - \rho^2(\omega)}{1 + \rho^2(\omega) - 2\rho(\omega)\cos\phi(\omega)} \\ \kappa(\omega) &= \frac{2 - \rho(\omega)\sin\phi(\omega)}{1 + \rho^2(\omega) - 2\rho(\omega)\cos\phi(\omega)} \quad \text{and} \\ f &= \frac{2}{\pi} \int_0^{\infty} \omega \varepsilon_2(\omega)\,\mathrm{d}\omega\,. \end{aligned} \tag{6.18}$$

The transverse and longitudinal eigenfrequencies can be deduced from the maxima of $\varepsilon_2(\omega)$ and of the so-called loss function $\mathrm{Im}\left\{-\frac{1}{\varepsilon(\omega)}\right\}$, respectively. See also [80R1, 88R1] of Chap. 5.

The relations (6.12),(6.13),(6.15),(6.17) are known as Kramers–Kronig relations. They are of very general nature and rely only on causality and locality

of the response. The two most important consequences are first, that if $\varepsilon_1(\omega)$ or $n(\omega)$ deviate in some frequency range from 1, then there must necessarily be absorption structures somewhere, i.e. spectral regions must exist with $\varepsilon_2(\omega) \neq 0$ or $\kappa(\omega) \neq 0$ and vice versa and, second, that if either the real or imaginary part of $\varepsilon_(\omega)$, $\tilde{n}(\omega)$ or $r(\omega)$ is known over the whole spectral range, then the other part can be calculated. Due to the denominator it is in practice sufficient to know the real or imaginary part over only a finite but not too small region around ω to be able to calculate the other part.

If spatial dispersion i.e. a non local response are included, the Kramer–Kronig relations become more complicated. This topic is beyond the scope of this book and for further details we refer the reader to [78H1] or [98B1, 98D1] of Chap. 2 and references therein.

We now leave the subject of ensembles of oscillators and proceed to consider the elementary excitations characteristic of semiconductors. They will later replace the model oscillators considered so far.

6.2 Problem

If you are interested in the analysis of complex functions, derive the Kramers-Kronig relations from the properties of analytic complex functions $f(z)$ with $z \in \mathbb{C}$ and give the restrictions imposed on $f(z)$.

References to Chap. 6

[72A1] abc Physik (Brockhaus, Leipzig 1972)
[78H1] K. Hümmer: Exczitonische Polaritonen in einachsigen Kristallen, Habilitation Thesis, Erlagen (1978)
[95P1] C. Porter, D.B. Tanner: Lab Programs, Dept. of Physics, University of Florida (1995)

7

Crystals, Lattices, Lattice Vibrations and Phonons

In this chapter we start to discuss topics that are specific to crystalline solids and, starting with Chap. 8, to semiconductors. We shall inspect the elementary excitations and quasi particles in semiconductors in Chaps. 7 to 10. These will be needed to describe and understand the linear optical properties in Chaps. 11 to 18. More details about these elementary excitations are found in textbooks on solid state physics; see for examples the References given in Chap. 1 like [81A1]a and h or [75Z1, 81M1, 89K1, 93K1, 95C1, 95I1, 97S1] and many others.

7.1 Adiabatic Approximation

If we want to describe a semiconductor, all we have to do in principle, is to solve the Schrödinger equation for the problem. It depends on the coordinates of the ion cores, consisting of the nucleus and the tightly bound electrons in the inner shells and the outer or valence electrons with coordinates $\boldsymbol{R}_j$ and $\boldsymbol{r}_i$, and masses M_j and m_0, respectively. The Hamiltonian reads:

$$H = -\frac{\hbar^2}{2}\sum_{j=1}^{M}\frac{1}{M_j}\Delta_{\boldsymbol{R}_j} - \frac{\hbar^2}{2m_0}\sum_{i=1}^{N}\Delta_{\boldsymbol{r}_i} + \frac{1}{4\pi\varepsilon_0} \times \left(\sum_{j>j'}\frac{e^2 Z_j Z_{j'}}{|\boldsymbol{R}_j - \boldsymbol{R}_{j'}|} + \sum_{i>i'}\frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_{i'}|} + \sum_{i,j}\frac{e^2 Z_j}{|\boldsymbol{R}_j - \boldsymbol{r}_i|}\right) . \tag{7.1}$$

Z_j is the effective charge of the ion core j and the indices j and i run over all M ion cores and N electrons, respectively.

We want to stress here that out of the four fundamental interactions so far known, namely strong, electromagnetic, weak, and gravitational interaction, only the electromagnetic one is of importance for all chemical prop-

erties including binding and thus also for the typical properties of semiconductors such as the transport and the optical properties discussed in this book. Within the electromagnetic interaction we restrict ourselves here to the electric ones (including exchange interaction) since electric interactions are usually much stronger than magnetic ones, basically since electric interactions begin with monopole-monopole (i.e., Coulomb) interaction, whereas magnetic interactions start only with dipole-dipole interactions, due to the absence of magnetic monopoles [see (2.1a)]. Magnetic interactions, however, do have a certain subtle importance e.g. in (dilluted) magnetic semiconductors.

The wavefunction solving (7.1) depends on all coordinates $\boldsymbol{R}_j$ and $\boldsymbol{r}_i$ including spins.

$$H\phi(\boldsymbol{r}_i, \boldsymbol{R}_j) = E\phi(\boldsymbol{r}_i, \boldsymbol{R}_j) \tag{7.2}$$

Since the indices j and i running from one to M and N, respectively, both count of the order of 10^{23} particles per cm^3 of semiconductor, it is obvious that there is at present no realistic chance of solving (7.1),(7.2) though a proper solution would, in principle, contain all information about a given semiconductor. If we do not want to get stuck at this point we must use some approximations to simplify (7.1). The most important one is the so-called adiabatic or Born–Oppenheimer approximation. It starts from the fact that the mass of an ion core is three to five orders of magnitude heavier than a free electron, i.e.,

$$M_j \simeq 1836 \cdot A_j m_0 \tag{7.3}$$

where A_j is the mass number of ion j. Since the electric forces that bind the outer electrons to the atom, and which can be described by a force constant β, are comparable to the ones which bind neighboring atoms or ions, we can easily see, even from classical arguments, that the highest resonance frequencies Ω with which ions can oscillate are much lower than the corresponding values ω for electrons

$$\Omega \simeq \left(\beta M_j^{-1}\right)^{1/2} \ll \omega = \left(\beta m_0^{-1}\right)^{1/2} \ . \tag{7.4}$$

Consequently, the electrons can practically instantaneously follow the motion of the ion cores, but not vice versa. This is the essence of the adiabatic approximation. On this basis we can separate $\phi(\boldsymbol{r}_i, \boldsymbol{R}_j)$ into a product of a wavefunction which depends only on the $\boldsymbol{R}_j$ and describes the motion of the ion cores, and another one which gives the wavefunction of the electron system depending on the momentary values of the $\boldsymbol{R}_j$. In a next step we will further assume that all ions are fixed at their equilibrium positions $\boldsymbol{R}_{j0}$ resulting finally in

$$\phi(\boldsymbol{r}_i, \boldsymbol{R}_j) = \phi(\boldsymbol{r}_i)\phi(\boldsymbol{R}_{j0}) \ , \tag{7.5}$$

and treating both the interaction between electrons and the deviation of the ions from their equilibrium positions in perturbation theory.

Before we start to inspect both factors of (7.5), we shall briefly outline how we describe a periodic lattice.

7.2 Lattices and Crystal Structures in Real and Reciprocal Space

In most cases we shall consider crystalline semiconductors. Disordered or amorphous systems will be mentioned explicitly. Crystalline solids have a periodic spatial arrangement of atoms, i.e., they show long-range order. We can define in such a case three non-coplanar elementary translation vectors $\boldsymbol{a}_i (i = 1, 2, 3)$ with the property that if we start at a special atom, e.g., a Ga atom in a GaAs crystal, we reach an identical atom if we move by a vector $\boldsymbol{R}$ given by

$$\boldsymbol{R} = n_i \boldsymbol{a}_1 + n_2 \boldsymbol{a}_2 + n_3 \boldsymbol{a}_3 \tag{7.6}$$

with $n_i = 0, \pm 1, \pm 2, \ldots$

The vector $\boldsymbol{R}$ is called a translation vector of the lattice. If we shift the lattice by $\boldsymbol{R}$ it comes to a position which is identical to the starting one.

The vectors $\boldsymbol{a}$, define a parallelepiped which is called the unit cell (see also Figs. 7.1,2). The whole volume of a crystal is completely filled with identical unit cells. The unit cell and the vectors $\boldsymbol{a}_i$ are called primitive if the unit cell has the minimum possible volume. This definition is not unique as we explain for a two-dimensional cubic lattice in Fig. 7.1, where we show a non-primitive unit cell and two primitive ones. By convention, a special primitive unit cell is agreed upon. In our case the one defined by $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$.

The vectors $\boldsymbol{R}$ evidently form for an infinite crystal an Abelian group which is called the translational group (Chap. 26). The positions of the atoms in the unit cell are given by the so-called basis. In Fig. 7.1 the basis consists of two atoms, one atom A at $(0, 0)$ and one atom B at $(1/2\boldsymbol{a}_1, 1/2\boldsymbol{a}_2)$. The translation vectors $\boldsymbol{a}_i$ and the basis is all that we need to describe a crystal structure.

The translation vectors $\boldsymbol{a}_i$ define an abstract, translation invariant lattice, the basis gives the information where the atoms are really located in the primitive unit cell. The lattice and the basis define together the crystal structure.

Apart from the translational group there is another type of symmetry operation which transforms the lattice into itself, but for which at least one point is kept fixed. These symmetry operations also form a group which is called the point group. The elements of this group are for example reflections at mirror planes, rotations around axes with two-, three-, four- or six-fold symmetry or the inversion through the origin.

Furthermore there may be screw axes or glide planes, which combine either a rotation axis or a mirror plane with a translation by a rational fraction of the $\boldsymbol{a}_i$. The abstract translation lattices can be grouped into 14 Bravais lattices namely one triclinic, two monoclinic, four orthorhombic, two tetragonal, three cubic, one trigonal (rhombohedral) and one hexagonal lattice.

If we include the positions of the atoms and the translational invariance we find from all possible combinations of the symmetry operations which

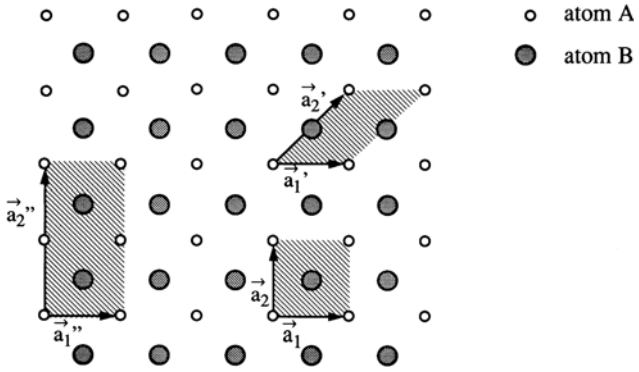


Fig. 7.1. Two primitive (*r.h.s*) and a non-primitive (*l.h.s.*) unit cells in a two-dimensional cubic lattice with a basis consisting of two different atoms per primitive unit cell

transform an infinite crystal into itself a total of 230 so-called space groups, out of which 73 can be written as a product of the translation group and the point group. For details see Chap. 26 and the references therein. The most important point groups for semiconductors are O_h (realized e.g. in the crystal structures of diamond, Si, Ge, Cu_2O or NaCl), T_d (realized e.g. in the zincblende type crystal structure of ZnS, ZnSe, GaAs, InP, CuCl or AgBr) and C_{6v} (realized e.g. in the wurtzite type crystal structure of ZnS, ZnO, CdS or GaN). We give in Fig. 7.2 the crystal structures of diamond, zinblende and wurtzite. The diamond crystal structure consist of C atoms occupying the lattice points of two face-centered cubic lattices shifted by 1/4 of the space diagonal of the cubic unit cell. For zincblende one has the same principle, however one of the two sublattices is occupied by atoms A, the other by B. The wurtzite crystal structure is hexagonal with a polar crystallographic **c**-xis. In all three cases, one atom is surrounded tetrahedrally by its four nearest neighbours. The difference between zincblende and wurtzit structures is in the positions of the next nearest neighbours only. Therefore several of the above mentioned compound semiconductors can crystallize in both structures like ZnS (which is notorious for these two polytypes), CdS or GaN. It is recommended that the reader visualizes these differences using some crystal models.

The chemical binding of the semiconductors is covalent for the elements (C, Si, Ge) with sp^3 hybridization, and acquires an increasing and finally dominant ionic admixture when going to the III–V, II^b–VI and I^b–VII compounds.

Now we want to introduce the so-called reciprocal lattice. It is defined by its elementary translation vectors $\boldsymbol{b}_i$ in the same way as the lattice in real space. The $\boldsymbol{b}_i$ are given by:

$$\boldsymbol{b}_1 = \frac{2\pi}{V_{\mathrm{uc}}} \boldsymbol{a}_2 \times \boldsymbol{a}_3 \tag{7.7}$$

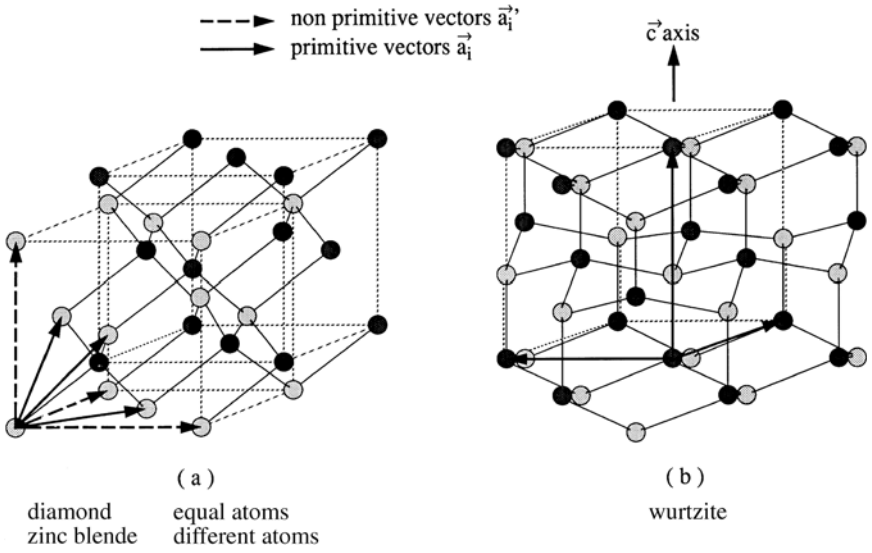


Fig. 7.2. The unit cells for the diamond and zinc-blende-type crystal structures (**a**) and of the wurtzite crystal structures (**b**) (see [82L1] of Chap. 1)

and cyclic permutations of the indices V_{uc} is the volume of the unit cell given by

$$V_{\mathrm{uc}} = \boldsymbol{a}_1(\boldsymbol{a}_2 \times \boldsymbol{a}_3). \tag{7.8}$$

A general translation vector of the reciprocal lattice is usually called $\boldsymbol{G}$

$$\boldsymbol{G} = l_1\boldsymbol{b}_1 + l_2\boldsymbol{b}_2 + l_3\boldsymbol{b}_3 \quad l_i = 0, \pm 1, \pm 2, \ldots \quad i = 1, 2, 3\,. \tag{7.9}$$

Without trying to be complete, we give some properties of the reciprocal lattice and its connections with the real one.

Every periodic function in real space which is sufficiently smooth and has a periodicity given by $f(\boldsymbol{r}+\boldsymbol{R}) = f(\boldsymbol{r})$ and $\boldsymbol{R}$ defined by (7.6) can be written as a Fourier series summing over all vectors of the reciprocal lattice

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G}} f_{\boldsymbol{G}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{G}\boldsymbol{r}} \tag{7.10}$$

with

$$f_{\boldsymbol{G}} = \boldsymbol{V}_{\mathrm{uc}}^{-1} \int_{\mathrm{uc}} f(\boldsymbol{r})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\boldsymbol{r}}\, \mathrm{d}\tau\,.$$

The scalar product of $\boldsymbol{R}$ and $\boldsymbol{G}$ always fulfills

$$\boldsymbol{R} \cdot \boldsymbol{G} = 2\pi m; \quad m = 0, \pm 1, \pm 2, \ldots \tag{7.11}$$

As a consequence, we can choose to describe effects occurring in periodic lattices in real space or in reciprocal space. The latter is the appropriate space

for wave vectors $\boldsymbol{k}$ or (quasi-)momenta $\hbar\boldsymbol{k}$. The "translation" from one space into the other is given by the three-dimensional Fourier series of (7.10).

In a crystal lattice we no longer have invariance with respect to infinitesimal translations in space (Sect. 3.1.3) but only invariance with respect to translations by integer multiples of $\boldsymbol{a}_i$. The conservation law for the momentum $\hbar\boldsymbol{k}$ which follows from an invariance with respect to infinitesimal translations according to Noether's theorem (3.14b) is modified for a periodic lattice so that $\hbar\boldsymbol{k}$ is conserved only to within integer mupltiples of the $\boldsymbol{b}_i$, i.e., we can add to a given $\boldsymbol{k}$-vector a vector from the reciprocal lattice $\boldsymbol{G}$:

$$\boldsymbol{k} \leftrightarrows \boldsymbol{k} + \boldsymbol{G} \; . \tag{7.12}$$

This is a very important statement which forms, together with energy conservation, the basis, for example, of Ewald's construction for the diffraction of X-rays or neutrons from a periodic lattice.

From (7.12) it is clear that we do not have to consider the whole $\boldsymbol{k}$-space, but can restrict ourselves to a "unit-cell" defined by the vectors $\boldsymbol{b}_i$. Every $\boldsymbol{k}$-vector that is outside the unit cell can be shifted inside the unit cell by adding an appropriate $\boldsymbol{G}$. Usually one does not work in reciprocal space with unit cells defined as in Fig. 7.1 or 7.2, but uses another construction explained in Fig. 7.3 for the two-dimensional case. One constructs the planes perpendicularly bisecting the lines connecting one point of the reciprocal lattice, which is chosen as the origin, with all others. The figure thereby enclosed around the origin is called the first Brillouin zone; the equivalent pieces which are next together form the second Brillouin zone, and so on. All Brillouin zones have equal area or volume in two or three dimensions, respectively. All higher Brillouin zones can be shifted into the first one by adding appropriate $\boldsymbol{G}$ vectors. The Brillouin zones also form a type of elementary cells, but constructed

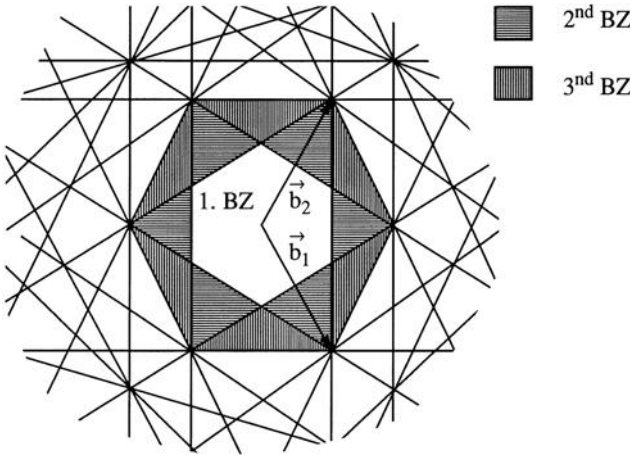


Fig. 7.3. The first Brillouin zones of a two-dimensional, hexagonal lattice

according to Fig. 7.3 and not according to Fig. 7.1. The first cell constructed in real space according to Fig. 7.3 is known as a Wigner–Seitz cell. The names of points and lines of high symmetry in the first Brillouin zone are indicated in Fig. 7.4.

For a simple cubic lattice with

$$\boldsymbol{a}_1 = (a, 0, 0), \quad \boldsymbol{a}_2 = (0, a, 0), \quad \boldsymbol{a}_3 = (0, 0, a)\ , \tag{7.13}$$

the $\boldsymbol{b}_i$ are also orthogonal with

$$\boldsymbol{b}_1 = \left[\frac{2\pi}{a}, 0, 0\right],\ \boldsymbol{b}_2 = \left[0, \frac{2\pi}{a}, 0\right],\ \boldsymbol{b}_3 = \left[0, 0, \frac{2\pi}{a}\right]\ , \tag{7.14}$$

and the first Brillouin zone is a cube which extends in all three directions from

$$-\frac{\pi}{a} \leq k_i \leq +\frac{\pi}{a}, \quad i = x, y, z\ . \tag{7.15}$$

In Fig. 7.4 we give the first Brillouin zones for this simple cubic lattice and for the point groups T_d, 0_h and $C_{6\mathrm{v}}$ using the primitive unit cell including the notation for some special points and directions. The center of the first Brillouin zone $\boldsymbol{k} = (0, 0, 0)$ is always called the Γ-point, other points of high symmetry are labelled by capital Latin letters and directions of high symmetry by capital Greek letters. As an example: in T_d symmetry, when leaving the Γ-point in the Σ-direction one arrives in at the point K at the border of the first Brillouin zone.

The quantity $\hbar\boldsymbol{k}$ of an excitation in a periodic lattice is usually called quasi-momentum if one wants to stress the difference compared to the momentum $\boldsymbol{p} = \hbar\boldsymbol{k}$ of a free particle in vacuum, e.g., a photon or an electron, where, in contrast to (7.12) no reciprocal lattice vector may be added. Actually it is possible to make a transition from one case to the other: if the lattice

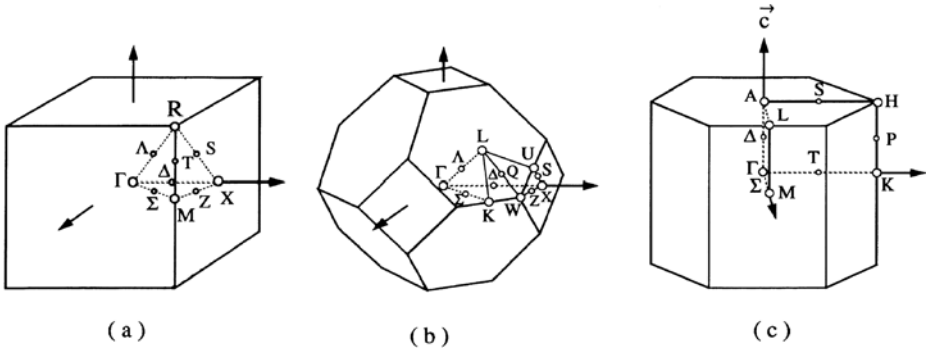


Fig. 7.4. The first Brillouin zones for a simple cubic lattice (**a**), of the diamond and zinc-blende-type structures (point groups 0_h, and T_d, respectively) (**b**) and of the hexagonal wurtzite-type structure ($C_{6\mathrm{v}}$) (**c**). The names of points and of directions of high symmetry are indicated. According e.g. to Reference [82L1] of Chap. 1

constant a goes to zero, the system regains translational invariance with respect to infinitessimally small shifts in real space. On the other hand, the $\boldsymbol{b}_i$ go to infinity in this limit (7.7) and the first Brillouin zone fills the whole $\boldsymbol{k}$-space, so that reciprocal lattice vectors become physically meaningless. For a more detailed discussion of the term "quasi"-momentum see e.g. Sect. 7.6 or [98B1] of Chap. 2.

7.3 Vibrations of a String

In Sects. 7.3 to 7.6 we treat the lattice vibrations and the resulting quanta, the phonons, in the way adopted in many textbooks, i.e., we start with a homogeneous string, proceed to monatomic and diatomic chains and finally arrive at the three-dimensional solid.

Let us first consider a quasi one-dimensional string, as shown schematically in Fig. 7.5. Two types of waves can propagate along it, transverse and longitudinal ones. The direction of the elongation is perpendicular to the direction of propagation, i.e., in the x–y plane, or parallel to it, i.e., in the z-direction, respectively. We start with the latter case. The mass density of the string is ρ its cross-section A, and the elongation of an infinitesimally small piece $\mathrm{d}z$ of the string at z from its equilibrium position is $u(z)$. Then, Newton's equation of motion reads

$$\mathrm{d}m\frac{\partial^2 u}{\partial t^2} = \rho A \cdot \mathrm{d}z \cdot \frac{\partial^2 u}{\partial t^2} = F \,. \tag{7.16}$$

The force F is connected to the elasticity modulus E via

$$F = A \cdot E\frac{\partial^2 u}{\partial z^2} \,. \tag{7.17}$$

The appearance of the second derivative in (7.17) is for some students surprising, bearing in mind Hooke's law. However, we must consider that the stress σ is indeed given by

$$\sigma(z) = E\frac{\partial u}{\partial z} \,. \tag{7.18}$$

If the stress is the same on both sides of the infinitesimal element of length $\mathrm{d}z$, the resulting forces at z and $z + \mathrm{d}z$ compensate each other to zero. The restoring force F is therefore given by $\mathrm{d}\sigma/\mathrm{d}z$ leading to (7.17).

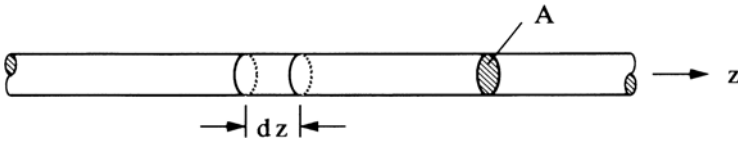


Fig. 7.5. A piece of a string as a model to explain the derivation of (7.19)

Putting (7.16) and (7.17) together leads to the standard harmonic wave equation

$$\rho \frac{\partial^2 u}{\partial t^2} = E \frac{\partial^2 u}{\partial z^2} \,. \tag{7.19}$$

With the ansatz

$$u = u_0 \exp\left[\mathrm{i}(kz - \omega t)\right] \tag{7.20}$$

for a plane wave we find the dispersion relation for longitudinal waves

$$\omega_{\mathrm{L}} = (E/\rho)^{1/2} k \,. \tag{7.21}$$

This is a linear relation as shown in Fig. 7.6. Consequently phase and group velocity are constant and equal, namely, with (2.13):

$$v_{\mathrm{ph}}^{\mathrm{L}} = v_{\mathrm{g}}^{\mathrm{L}} = (E/\rho)^{1/2} \,. \tag{7.22}$$

For the two degenerate, transverse waves we find in a similar way

$$\omega_{\mathrm{T}} = (G/\rho)^{1/2} k \tag{7.23}$$

or

$$v_{\mathrm{ph}}^{\mathrm{T}} = v_{\mathrm{g}}^{\mathrm{T}} = (G/\rho)^{1/2} \,. \tag{7.24}$$

where G is the shear or torsion modulus.

Since it is known from the theory of elasticity that

$$G \leq E \,, \tag{7.25}$$

we find

$$v_{\mathrm{ph}}^{\mathrm{T}} \leq v_{\mathrm{ph}}^{\mathrm{L}} \,, \tag{7.26}$$

a result comparable to (4.28).

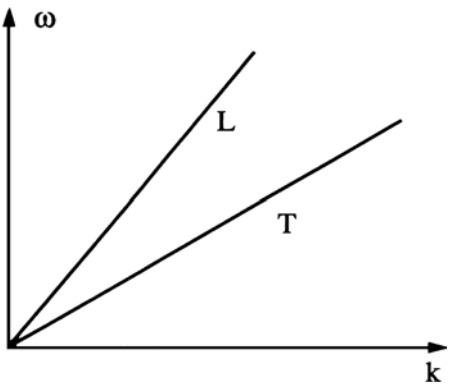


Fig. 7.6. The dispersion relation of waves on a homogeneous string

7.4 Linear Chains

We now should consider the regime of validity of the above calculation. We assumed a homogeneous string, neglecting the fact that a solid is made up from atoms. Therefore the above approximation can only be valid for wavelengths much longer than the lattice constant or for wave vectors close to the center of the first Brillouin zone, i.e.,

$$\lambda \gg a \quad \text{or} \quad k \ll \frac{\pi}{a} \,. \tag{7.27}$$

For shorter wavelengths we have to consider the atomic structure of solids. The interaction potential between neighboring atoms as a function of the lattice constant a looks approximately like Fig. 7.7 for all types of binding, e.g., covalent, ionic or metallic. For sufficiently large lattice constants there is no interaction between the atoms, i.e., $V = 0$; then comes an attractive regime (without which there would be no solids); and this is finally followed by a steep repulsive increase due to Pauli's exclusion principle when the filled inner shells of neighboring atoms start to overlap. Different analytic approximations to $V(a)$ are known for example the Born-Mayer or Lennard-Jones potentials. These details have at present no relevance for us. We note that a crystal left to itself will come to a state close to the energetic minimum, i.e., to the equilibrium position a_0. In the vicinity of a_0, $V(a)$ can be approximated by a parabola, that is by a harmonic potential, shown by the dashed line. It is at least qualitatively clear that this harmonic approximation is valid only very close to a_0. For larger deviations from a_0 significant anharmonicities (i.e., deviations from the harmonic potential) have to be expected. The anharmonicities are characteristic for lattice vibrations and manifest themselves, among other things, in the thermal expansion of solids and in phonon–phonon interaction.

For the moment, however, the harmonic approximation is good enough and we consider a linear model solid in which every atom with mass M is connected to its neighbors by a "spring" with a force constant D, representing the harmonic potential, leading to the linear-chain model of Fig. 7.8a in which

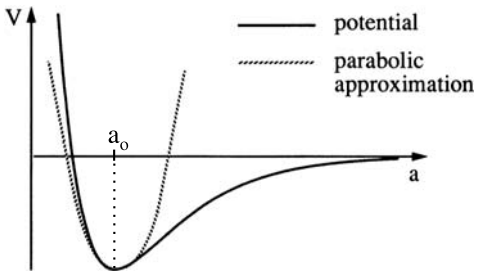


Fig. 7.7. Sketch of the potential V between neighboring atoms in a solid as a function of the lattice constant a

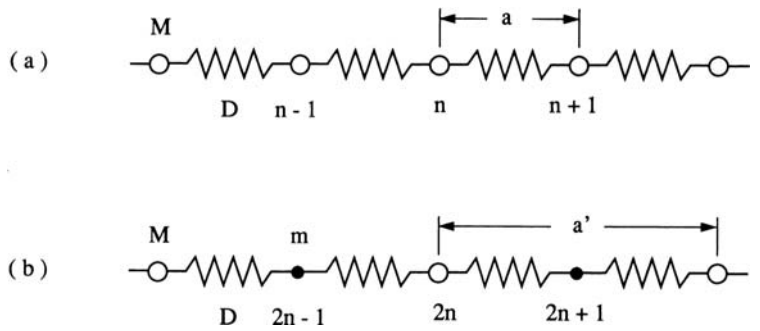


Fig. 7.8. The linear chain model for the cases of one atom per unit cell (**a**) and two atoms per unit cell (**b**)

we indicate also the lattice constant a. Evidently we have a basis consisting of one atom per primitive unit cell. At this point it is important to stress the difference between the models of Figs. 4.1 and 5.8 on one side and of Fig. 7.8 on the other side. In the first case we had independent oscillators, and a weak coupling between them was introduced only to simulate spatial dispersion. Here, in Sect. 7.4, the coupling springs from the only forces acting on the atoms. We introduce now the displacement of atom u_n from the equilibrium position and obtain the equation of motion, again for the longitudinal mode

$$M\frac{\partial^2 u_n}{\partial t^2} = D\left[(u_{n+1} - u_n) - (u_n - u_{n-1})\right] = D(u_{n+1} - 2u_n + u_{n-1})\ . \quad (7.28)$$

Instead of the second differential quotient in the homogeneous approximation we are now left with a second order difference equation.

As a solution of (7.28) we again try a plane wave which reads, in this discrete case,

$$u_n = u_{n,0}\exp\left[\mathrm{i}(kna - \omega t)\right] \quad (7.29a)$$

$$u_{n\pm1} = u_{n\pm1,0}\exp\left\{\mathrm{i}\left[k(n \pm 1)a - \omega t\right]\right\}\ . \quad (7.29b)$$

For a plane wave we conclude that the amplitudes of the various atoms are equal, i.e.,

$$u_{n,0} = u_{n+1,0} = u_{n-1,0}\ . \quad (7.29c)$$

Inserting (7.29a,b,c) into (7.28) gives

$$-M\omega^2 = D\left(\mathrm{e}^{-\mathrm{i}ka} - 2 + \mathrm{e}^{\mathrm{i}ka}\right) = -2D\left(1 - \cos ka\right) \quad (7.30)$$

or

$$\omega = \left(\frac{4D}{M}\right)^{1/2}\left|\sin\frac{ka}{2}\right|\ . \quad (7.31)$$

The dispersion relation according to (7.31) is shown in Fig. 7.9 together with the phase and group velocities in the first Brillouin zone.

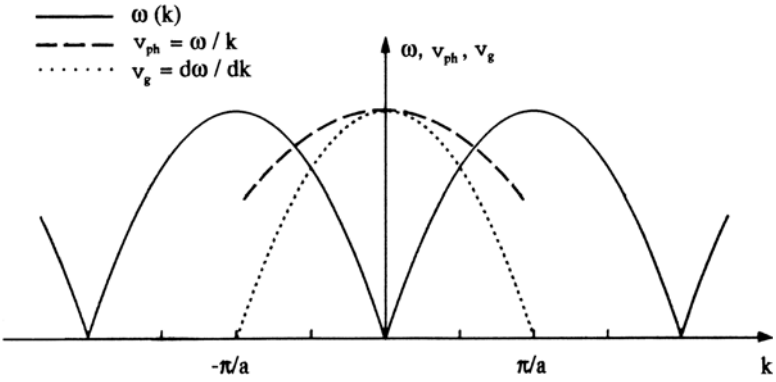


Fig. 7.9. The dispersion relation and the phase and group velocities of the longitudinal vibrations of the monatomic chain of Fig. 7.8a

As can be expected from the discussion in Sect. 7.2, the dispersion relation outside the first Brillouin zone repeats just what is inside, or, in other words, the branches inside and outside can be shifted into each other by adding or subtracting reciprocal lattice vectors $l \cdot 2\pi/a$, where l is a positive or negative integer.

The fact that there is nothing new outside the first Brillouin zone can be easily elucidated for the case of lattice vibrations in connection with Fig. 7.10.

If adjacent atoms are displaced in antiphase, we end up with the shortest physically meaningful wavelength (solid line) $\lambda_{\min}$ by

$$\lambda_{\min} = 2a \rightarrow k_{\max} = \frac{2\pi}{\lambda_{\min}} = \frac{\pi}{a} \,. \tag{7.32}$$

Of course we can define a shorter wavelength as indicated by the dashed line resulting here in $k = 3\pi/a$. But this definition is physically meaningless since we have no atoms at the positions between $z = na$ and $z = (n+1)a$. On the other hand, $k = 3\pi/a$ corresponds to π/a by just adding $G = -2\pi/a$.

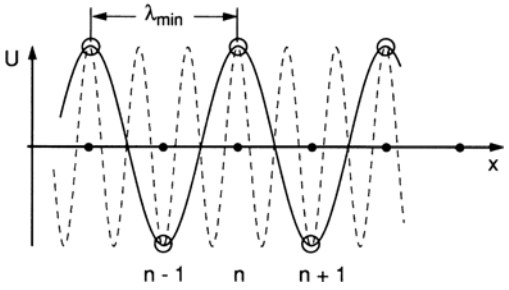


Fig. 7.10. A vibration of the linear chain of Fig. 7.8a with a wave vector inside and outside the first Brillouin zone

We now inspect v_{ph} and v_{s} in the first Brillouin zone. Except at $k = 0$ the two quantities are no longer equal and change with k. Consequently a wave packet will become broader during propagation, i.e., it will show "dispersion" in the sense discussed in Chap. 5. The fact that the dispersion relation is horizontal at the border of the first Brillouin zone in the direction normal to this border is generally the case.

Since most semiconductors have more than one atom per primitive unit cell, we address this situation with Fig. 7.8b where we evidently have a basis consisting of two atoms with masses M and m. The lattice constant is now a'.

Using the nomenclature of Fig. 7.8b we obtain the following equations of motion in analogy to (7.28):

$$M\frac{\partial^2 u_{2n}}{\partial t^2} = D\left(u_{2n+1} - 2u_{2n} + u_{2n-1}\right) , \tag{7.33a}$$

$$m\frac{\partial^2 u_{2n+1}}{\partial t^2} = D\left(u_{2n+2} - 2u_{2n+1} + u_{2n}\right) . \tag{7.33b}$$

Using again the ansatz

$$u_{2n} = u_{2n,0}\exp\left[\mathrm{i}(2nak - \omega t)\right] , \tag{7.34a}$$

$$u_{2n+1} = u_{2n+1,0}\exp\left\{\mathrm{i}\left[(2n+1)ak - \omega t\right]\right\} , \tag{7.34b}$$

and noting again that the amplitudes of equal masses are equal

$$u_{2n,0} = u_{2n+2,0} = A_M , \tag{7.35a}$$

$$u_{2n+1,0} = u_{2n-1,0} = A_m , \tag{7.35b}$$

we get the following set of linear equations

$$\left.\begin{aligned}(2D - \omega^2 M)A_M - 2D\cos(ka)A_m = 0 ,\\ -2D\cos(ka)A_M + (2D - \omega^2 m)A_m = 0 .\end{aligned}\right\} \tag{7.36}$$

These have a non-trivial solution (i.e., one other than $A_M = A_m = 0$) only if the determinant of the coefficients vanishes. The dispersion relation resulting from the corresponding secular equation reads:

$$\omega^2 = D\left(\frac{1}{m} + \frac{1}{M}\right) \pm D\left[\left(\frac{1}{m} + \frac{1}{M}\right)^2 - \frac{4}{Mm}\sin^2\frac{ka'}{2}\right]^{1/2} \tag{7.37}$$

The dispersion relation has now two branches, as shown in Fig. 7.11, where we give also the values at some special points. The lower branch is usually called the acoustic branch since sound waves propagate according to its modes. The upper branch is called the optical one, for reasons given below.

We can enter the solution (7.37) into (7.36) and calculate the ratio A_M/A_m. The procedure is straightforward but lengthy and so we present

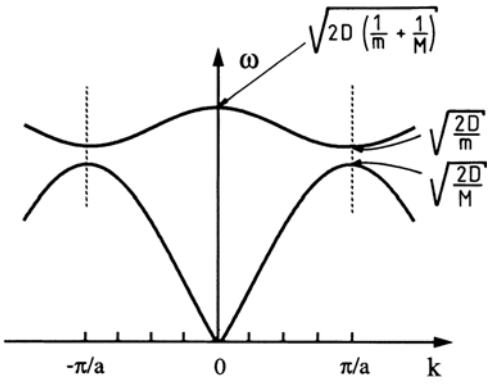


Fig. 7.11. The dispersion relation for the diatomic chain

the result only graphically in Fig. 7.12 in agreement with our statement in Sect. 1.1, and discuss it in connection with Fig. 7.13. For the acoustic branch, the two different atoms are displaced in the same direction, Fig. 7.13a. For very long wavelengths (i.e., $k \simeq 0$) the amplitudes are equal. Actually the case $k = 0$ corresponds to a simple displacement of the whole crystal. For increasing k the amplitude of the heavy mass M gets larger than that of m for $M > m$, and, at the boundary of the first Brillouin zone, only the heavy masses oscillate (Fig. 7.13b), resulting in an eigenfrequency $(2D/M)^{1/2}$ as indicated in Fig. 7.11.

For the optical branch, the two atoms are displaced in anti-phase (Fig. 7.13c). At the boundary of the first Brillouin zone only the light masses oscillate (Fig. 7.13d).

If the two different atoms carry an electric charge, i.e., if the binding is at least partly ionic, then the oscillation according to Fig. 7.13c is connected with an oscillating electric dipole. This allows it to couple to the electromagnetic light field at least for the transverse eigenmodes and this is why these oscillations are called "optical" modes. We come back to this aspect in Chap. 11.

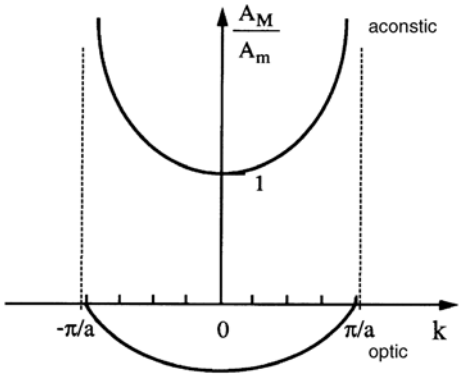


Fig. 7.12. The ratio of the amplitudes $A_{\mathrm{M}}/A_{\mathrm{m}}$ of the two different masses for the two dispersion branches of Fig. 7.11 as a function of k

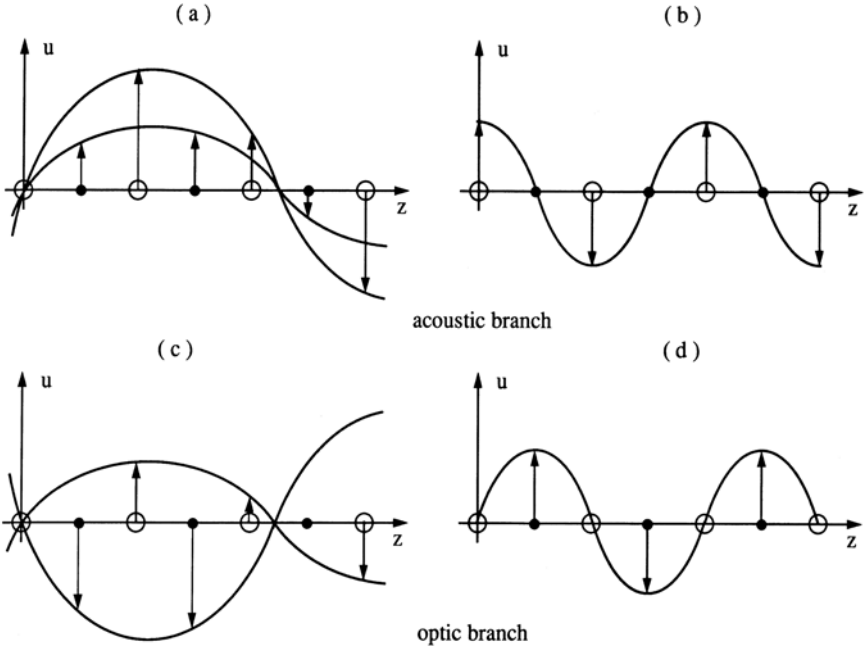


Fig. 7.13. Sketch of the displacements of the atoms on the acoustic (**a**, **b**) and the optical branches for two different wave vectors (**c**, **d**)

Here we have discussed in detail only the longitudinal modes, but it is obvious that both for the monatomic and the diatomic chain, for every wave vector and every branch, two (degenerate) transverse oscillations also exist.

The dispersion relation of Fig. 7.11 can also be deduced in another way starting from the one of Fig. 7.9, i.e., from the monatomic chain which we repeat in Fig. 7.14a. Now we imagine that we paint the atoms of the monatomic chain in two different alternating colors, but without changing their physical properties. As a consequence we have increased the length of the primitive unit cell by a factor of two and the new lattice constant a' is given by

$$a' = 2a \tag{7.38}$$

and this in turn reduces the length of the first Brillouin zone by one-half as shown in Fig. 7.14a,b. Consequently, we can shift the outer parts of the dispersion relation by vectors of the new reciprocal lattice into the first Brillouin zone. This situation is shown in Fig. 7.14b. Since the atoms are still identical, the two branches cross at the border of the first Brillouin zone. If we now also introduce differences in the physical properties of the atoms, for example giving them different masses, then we end up with the situation of Fig. 7.14c which is identical to Fig. 7.11. We shall use this set of arguments again in connection with superlattices later on in Sect. 8.10. First we want to extend this discussion of the lattice vibrations of three-dimensional systems.

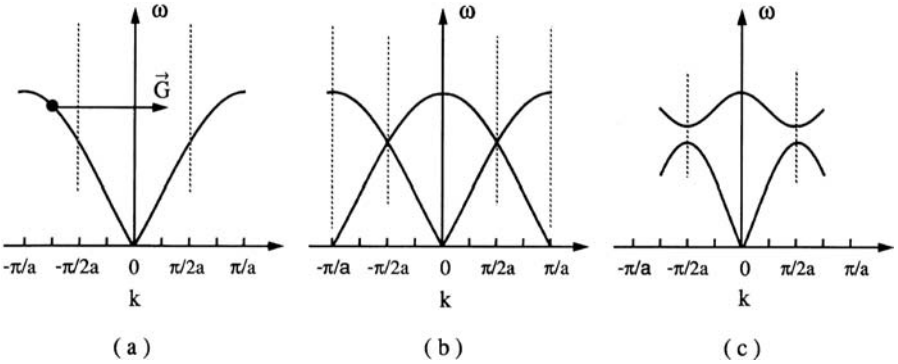


Fig. 7.14. The transition from the monatomic (**a**) to the diatomic (**c**) chain. (See text for explanation)

7.5 Three-Dimensional Crystals

If we consider a three-dimensional crystal, not too much changes with respect to the chain model of Sect. 7.4, at least for the dispersion relations. The atoms are connected with nearest (and possibly next-nearest) neighbors in the three-dimensional crystal stucture. The set of equations analogous to (7.28) or (7.33) will become correspondingly more complex, but the result will be qualitatively the same. There are still the acoustic branches and, in addition, optical ones if we have more than one atom in the primitive unit cell. There are always three acoustic branches, namely one longitudinal and two transverse ones for every $\boldsymbol{k}$-vector and $3s-3$ optic ones:

$$\left.\begin{array}{ll} \text{number of acoustic branches:} & 3 \\ \text{number of optical branches:} & 3s-3 \end{array}\right\} \tag{7.39}$$

where s is the number of atoms per primitive unit cell.

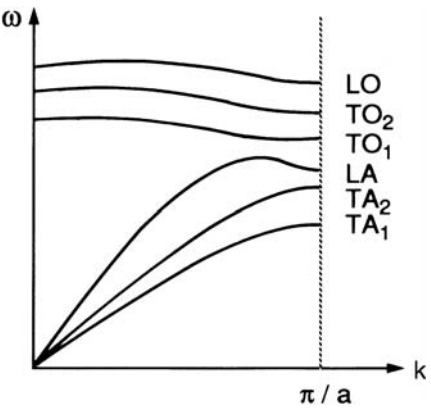


Fig. 7.15. Sketch of the dispersion relation of lattice vibrations for a three-dimensional anisotropic crystal with partly ionic binding and two atoms per primitive unit cell

If the crystal is anisotropic the dispersion relations will be different for different directions in the Brillouin zone and the degeneracy between the two transverse modes for every $\boldsymbol{k}$ may be lifted. We show such a situation schematically in Fig. 7.15.

7.6 Quantization of Lattice Vibrations: Phonons and the Concept of Quasiparticles

If we look again at (7.28) or (7.33) for the $u_n(t)$ we see that these equations are rather similar to the equation of motion for an harmonic oscillator

$$m\frac{\partial^2 v_n}{\partial t^2} = -Dv_n \; . \tag{7.40}$$

The only difference are the terms with indices $n \pm 1$, i.e., the off-diagonal terms in the language of matrix representation. It is now possible to find appropriate linear combinations of the u_n, so-called normal coordinates v_n

$$v_n = \sum_{n'} a_{n'} u_{n'} \tag{7.41}$$

such that the v_n obey equations like (7.40) or (2.52) or, in other words, to diagonalize the problem. In doing so one usually closes the linear chain (or the three-dimensional equivalent) to a huge ring in order to close the set of differential equations.

On the other hand a quantum mechanical treatment of the harmonic oscillator leads to discrete energy levels given by

$$E_n = \hbar\omega_n \left(n + \frac{1}{2}\right), \quad n = 0, 1, 2, 3 \ldots . \tag{7.42}$$

It is now important to note that the dispersion relation shown for example in Fig. 7.15 is not influenced by a linear transformation of the coordinates.

As a consequence we may say that the lattice vibrations (more precisely the vibrations of the crystal structure) consist of quanta according to (7.42) for each wave vector $\boldsymbol{k}$ and branch i. The total energy of the lattice vibrations can be written as

$$E = \sum_{\boldsymbol{k},i} \hbar\omega_{\boldsymbol{k},i} \left(n_{\boldsymbol{k},i} + \frac{1}{2}\right) \tag{7.43}$$

in analogy to (2.55b) or (5.3), where the index i labels the various branches.

In the framework of second quantization it is possible to define creation and annihilation operators $b^+_{\boldsymbol{k},i}$ and $b_{\boldsymbol{k},i}$ respectively, and the Hamiltonian can

then again be written in terms of the number operator, in a similar way to that used already for photons

$$H = \sum_{\boldsymbol{k},i} \hbar\omega_{\boldsymbol{k},i} \left(b^{+}_{\boldsymbol{k},i} b^{+}_{\boldsymbol{k},i} + \frac{1}{2} \right) . \tag{7.44}$$

Obviously there is a close analogy between the quanta or particles of the electromagnetic radiation in vacuum, the photons, and the quanta of the lattice vibrations. The quanta of the lattice vibrations are called phonons and are considered as quasi-particles. The attribute "quasi" has two reasons. In contrast to "real" particles like photons, electrons or protons, phonons can exist only in matter and not in vacuum. They are characterized by their energy and momentum $\hbar\boldsymbol{k}$ but, in contrast to vacuum, $\hbar\boldsymbol{k}$ is a "quasi-momentum" which is defined only modulo integer multiplies of the elementary translation vectors $\boldsymbol{b}_i$ of the reciprocal lattice for a crystalline solid (Sect. 7.2). The concept of quasi-particles, which as been introduced by Landau about 50 years ago, is a very important one for the understanding of solids. In the next chapters we shall become familiar with several other quasi-particles in solids. The phonons are just the first example.

The existence and the dispersion relation of phonons can be investigated by inelastic neutron scattering. An incident neutron from a mono-energetic beam with

$$E_{\mathrm{i}} = \frac{\hbar^2 \boldsymbol{k}_{\mathrm{i}}^2}{2m_{\mathrm{n}}} \tag{7.45}$$

is scattered under creation or annihilation of a phonon, resulting in a neutron in the final state with E_{f}, $\boldsymbol{k}_{\mathrm{f}}$, which are given via the conservation laws of energy and quasi-momentum

$$E_{\mathrm{f}} = \frac{\hbar^2 \boldsymbol{k}_{\mathrm{f}}^2}{2m_{\mathrm{n}}} = E_{\mathrm{i}} \pm \hbar\omega_{\mathrm{Phonon}}$$

and

$$\boldsymbol{k}_{\mathrm{f}} = \boldsymbol{k}_{\mathrm{i}} \pm \boldsymbol{k}_{\mathrm{Phonon}} + \boldsymbol{G} \tag{7.46}$$

By measuring the properties of the incident and scattered neutrons it is possible to prove the existence of phonons and to determine their dispersion relation. Note that the scattering probability for the neutrons may depend on $\boldsymbol{G}$.

Figure 7.16 shows the dispersion relation of the phonons in two different semiconductors, Si and CdS, for various directions in $\boldsymbol{k}$-space (see Fig. 7.4). All well-known semiconductors, including the elemental ones like Si and Ge, have more than one atom per unit cell and therefore support both acoustic and optical branches. Si has only covalent binding in contrast to CdS which has a mixed ionic-covalent binding. Therefore the Si atoms do not carry an electric charge and, as a result, even the optical phonons do not couple directly to the radiation field, resulting at $\boldsymbol{k} = 0$ in an oscillator strength $f = 0$ and consequently in $\Delta_{\mathrm{LT}} = 0$ as a result of (4.26).

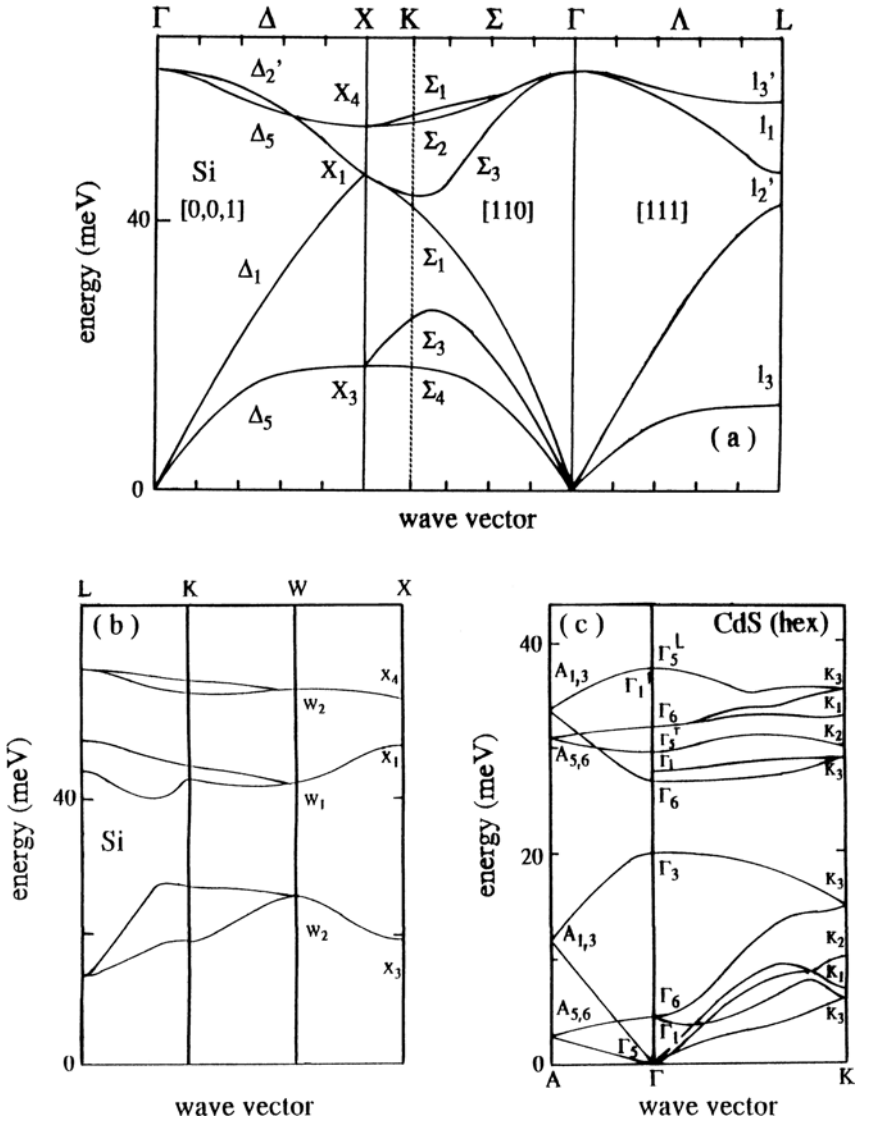


Fig. 7.16. Dispersion relation of phonons in Si and CdS for various directions in $\boldsymbol{k}$-space (**a**) and (**c**) show essentially directions in $\boldsymbol{k}$ space, which originate, from the Γ point and (**b**) on the surface of the first Brillouin zone After [82L1] of Chap. 1

In connection with Fig. 7.7 we have introduced a harmonic interaction potential between atoms. Actually there are strong anharmonicities, as already mentioned, which are, among others, due to the fact that the electron distribution changes almost instantaneously with the changing positions $\boldsymbol{R}_j$ of the atoms, c.f. the adiabatic approximation of Sect. 7.1. This fact results in

a variation of the "spring-constant" D with the lattice constant a, i.e., in an anharmonicity.

The resulting anharmonicity manifests itself, for example, in scattering processes between phonons. We show some of them schematically in Fig. 7.17. However, one should bear in mind that all combination possibilities are exhausted only when the three-dimensional $\boldsymbol{k}$-space is considered.

In (a) a TO phonon decays under energy and momentum conservation into two acoustic phonons in (b) two transverse acoustic phonons combine to form a longitudinal acoustic phonon. Energy and momentum conservation read e.g. for (a) and (b):

$$\text{(a)}\ \hbar\omega_\mathrm{i} = \hbar\omega_\mathrm{f1} + \hbar\omega_\mathrm{f2}, \quad \boldsymbol{k}_\mathrm{i} = \boldsymbol{k}_\mathrm{f1} + \boldsymbol{k}_\mathrm{f2} \tag{7.47a}$$

$$\text{(b)}\ \hbar\omega_\mathrm{i1} + \hbar\omega_\mathrm{i2} = \hbar\omega_\mathrm{f}, \quad \boldsymbol{k}_\mathrm{i1} = \boldsymbol{k}_\mathrm{i2} + \boldsymbol{k}_\mathrm{f}\ . \tag{7.47b}$$

These are so-called n or normal processes. In the decay process (c) we end up in the second Brillouin zone with one phonon and fold the phonon back with a vector of the reciprocal lattice $\boldsymbol{G}$.

$$\hbar\omega_\mathrm{i} = \hbar\omega_\mathrm{f1} + \hbar\omega_\mathrm{f2}; \qquad \boldsymbol{k}_\mathrm{i} = \boldsymbol{k}_\mathrm{f1} + \boldsymbol{k}_\mathrm{f2} + \boldsymbol{G}\ . \tag{7.47c}$$

The situation of (c) is known as a u or Umklapp process.

7.7 The Density of States and Phonon Statistics

Now we want to calculate the density of states of phonons as a first application of what we have learned in Sect. 2.6 for a three-dimensional crystal.

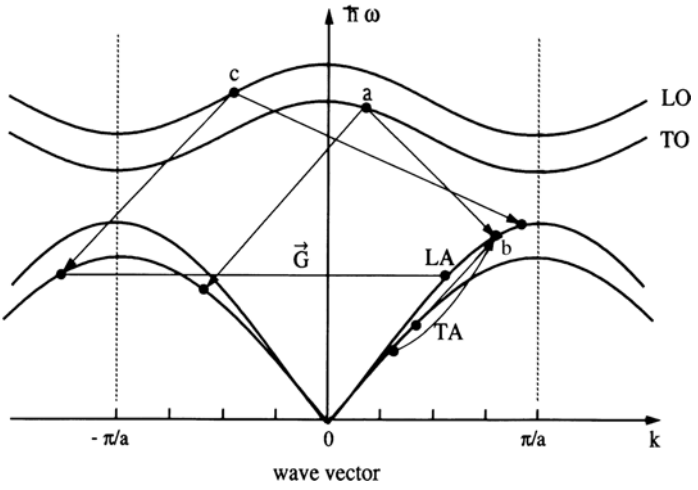


Fig. 7.17. Sketch of three possible decay or fusion processes of phonons

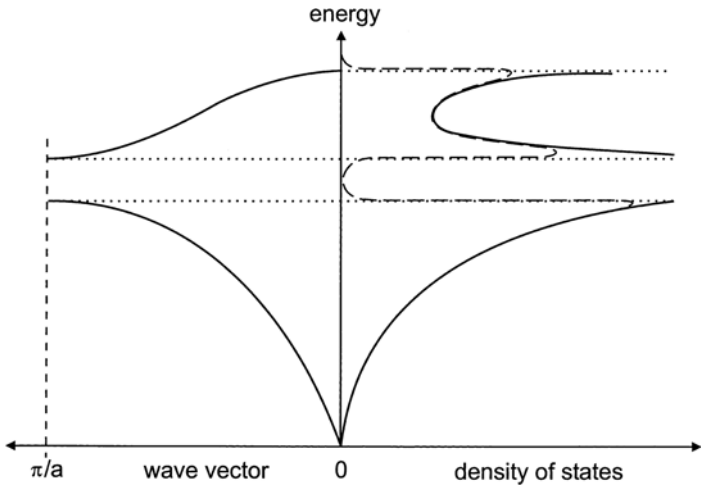


Fig. 7.18. Schematic drawing of the density of states of a three-dimensional, isotropic semiconductor with two atoms per unit cell and degenerate optical and acoustic branches

We assume for simplicity that we have two atoms per unit cell and that the resulting three acoustic and optical branches are degenerate i.e. transverse and longitudinal branches coincide. Furthermore we assume an isotropic semiconductor. This results in Fig. 7.18 where we plot on one side of the x-axis $\boldsymbol{k}$ and on the other the density of states $D(\omega)$. As stated in Sect. 2.6 we need the dispersion relation to calculate $D(\omega)$. In the linear part of the acoustic branch we have

$$\omega = v_{\mathrm{s}} k \ , \tag{7.48}$$

where v_{s} is the constant velocity of sound. This linear relation results immediately in

$$D(\omega) = \mathrm{const} \cdot \omega^2 \tag{7.49}$$

i.e., a parabolic density of states similar to that for photons in vacuum. The difference is only in the proportionality factor of (7.48),(7.49).

When deviations from the linear dispersion relation start, the calculations get more lengthy and we will not go into details. The denominator on the right of (2.77) tells us, however, that the density of states has in principle a singularity and in practice a steep maximum when we have a horizontal slope of the $\omega(\boldsymbol{k})$ relation, as indicated in Fig. 7.18.

Note that the situation is in reality more complex than in in Fig. 7.18 due to the facts, that the dispersion relation may depend in the direction of $\boldsymbol{k}$, the transverse and longitudinal branches are not degenerate, and that the first Brillouin zone is not a sphere but a polyhedron, with the consequence, that the range of $\boldsymbol{k}$ vectors depends on the direction in the Brillouin zone. For realistic densities of states see e.g. [82L1] of Chap. 1.

If we want to know the number of phonons we must integrate over the density of states weighted with the probability that the state is populated. The commutation relations for phonons obey

$$\left[b_{\boldsymbol{k},i}, b^{+}_{\boldsymbol{k}',i'}\right]^{-} = \partial_{\boldsymbol{k},\boldsymbol{k}'}\partial_{i,i'} \tag{7.50}$$

where i is the label of the branch. This means that phonons are bosons. So we have in thermal equilibrium to use Bose–Einstein statistics (2.80b)

$$f_{\mathrm{BE}}(\hbar\omega, T) = \left\{\exp\left[(\hbar\omega - \mu)/k_{\mathrm{B}}T\right] - 1\right\}^{-1} \tag{7.51}$$

Since the number of phonons is not conserved, due e.g., to phonon–phonon interaction (see above), the chemical potential μ is zero in thermal equilibrium (Sect. 2.6). For the total number of phonons, $N(T)$, we obtain

$$N(T) = \int_0^\infty D(\hbar\omega)\left[\exp(\hbar\omega/k_{\mathrm{B}}T) - 1\right]^{-1}\,\mathrm{d}\omega \tag{7.52a}$$

and for the energy of the phonon system

$$U(T) = \int_0^\infty \hbar\omega D(\hbar\omega)\left[\exp(\hbar\omega/k_{\mathrm{B}}T) - 1\right]^{-1}\,\mathrm{d}\omega\,. \tag{7.52b}$$

Starting from (7.52b) it is easy to calculate the specific heat of the phonon system

$$c_p \simeq c_v = \frac{\partial U}{\partial T} \tag{7.53}$$

if appropriate approximations are made for $D(\omega)$. Einstein assumed that all phonons have the same frequency ω_{E} i.e., he approximated $D(\hbar\omega)$ by a δ-function

$$D(\hbar\omega) = \delta(\omega - \omega_{\mathrm{E}})3Ns\,, \tag{7.54a}$$

and Debye continued the linear part of the dispersion relation up to a frequency which is also chosen to accommodate all $3Ns$ degrees of freedom of the atoms, where N is the number of unit cells and s the number of atoms per unit cell

$$3Ns = \int_0^{\omega_{\mathrm{D}}} D(\hbar\omega)\,\mathrm{d}\omega \quad \text{with} \quad D(\omega) \propto \omega^2\,. \tag{7.54b}$$

For high temperatures (i.e., $k_{\mathrm{B}}T > \hbar\omega_{\mathrm{D}}$ or $\hbar\omega_{\mathrm{E}}$) both approximations give the classical limit, namely the law of Dulong and Petit

$$c_v = 3Nsk_{\mathrm{B}} \tag{7.55}$$

and a continuous approach to zero for $T \to 0$. In the case of Debye's approximation one finds the well-known T^3 law:

$$c_v \sim (T/\theta)^3 \quad \text{with} \quad k_\mathrm{B}\theta = \hbar\omega_\mathrm{D} \quad \text{and} \quad T < \theta \,. \tag{7.56}$$

Since the discussion of specific heat c_v is outside the scope of this book, we do not take it any further, having introduced it simply to illustrate the applicability of the concept of the density of states, and mention only that the approach in (7.52b),(7.54b) is, apart from some constants and $\omega_\mathrm{D} \to \infty$ identical to the one used to describe Planck's law of blackbody radiation as seen by comparing with (2.84).

Finally, we want to point out the following. Since phonons are bosons, processes which involve phonons can be stimulated in the same way as was discussed in connection with photons in Sect. 3.2.1, i.e., we get in transitions involving the emission of a phonon in a certain mode, apart from other terms, a factor like

$$W_{\mathrm{i}\to\mathrm{f}} \sim (N^k_\mathrm{Ph} + 1) \tag{7.57}$$

To get a feeling we assume a lattice temperture of 77 K (i.e., liquid N_2) and an acoustic phonon with an energy around 0.2 meV (such phonons will be used in Brillouin scattering in Sect. 13.3) and find for these conditions with (7.51)

$$N^k_\mathrm{Ph} \simeq 30 \gg 1 \,. \tag{7.58}$$

This means that the occupation number is much larger than one and processes which involve the emission of phonons with energies smaller than the thermal energy are stimulated by the phonons. Depending on the process under consideration, it is however necessary to consider the reverse process, too, which depends on N^k_Ph.

7.8 Phonons in Alloys

An alloy is a random binary, ternary or higher mixture of atoms. In contrast to amorphous materials, the atoms in an alloy are still sitting on well defined, periodically arranged lattice sites. Two very widely investigated semiconductor alloys are $Al_{1-y}Ga_yAs$ or $CdS_{1-x}Se_x$. The first one cristallizes in the cubic zinc-blende structure the second usually in the hexagonal wurtzite type. However the two different kations Al and Ga or the two anions S and Se are randomly distributed over the respective sublattice with an average concentration y or x. See for the concept of the virtual crystal approach Sect. 8.15.

What are the consequences of alloying for the phonons.

For long-wavelenth acoustic phonons the situation is generally rather simple. The dispersion relation starts linear as in Figs. 7.9, 7.11 or 7.14 with a slope or velocity of sound which is a weighted average of the two parent components. Long wavelength means large compared to the interatomic distance. When approaching the boarder of the first Brillouin zone, things become

less clear. Since only one sublattice is elongated (see Fig. 7.13b) the phonon frequency may be broadened depending on which type of atom oscillates due to their different masses and "spring constants" to their neighbours. In this sense even a chemically pure crystal, containing however different isotopes of the same element, has to be considered as an alloy. For details of this topic see e.g. [94C1, 94R1].

For the optical phonons, two different situations may arise. If the bands of e.g. the LO-phonons overlap energetically in the two parent materials as is the case e.g. in $Zn_{1-y}Cd_yS$ one finds the so called amalgamation type behaviour, i.e. the LO eigenfrequency shifts continuously with the composition y from the one of AB to the one of AC in an alloy $AB_{1-x}C_x$ as shown in Fig. 7.19a.

If the bands do not overlap energetically an AB oscillation mode does not find a resonant partner of AC modes for any $\boldsymbol{k}$ vector and vice versa and the alloys exhibit generally the so-called persistent mode behaviour. This means, that the eigenfrequencies of the pure AB and AC modes vary only marginally under alloying as a function of x, but their relative weights or oscillator strengths change as shown schematically in Fig. 7.19b.

Examples of the optical properties will be given in Chap. 11.

7.9 Defects and Localized Phonon Modes

In our discussion of phonons (and of our model substance in Chaps. 4 and 5), we assumed until now a perfect arrangement of atoms. In fact, every real semiconductor contains a lot of crystal or lattice defects. There are point defects, one-dimensional defects like dislocations, two-dimensional

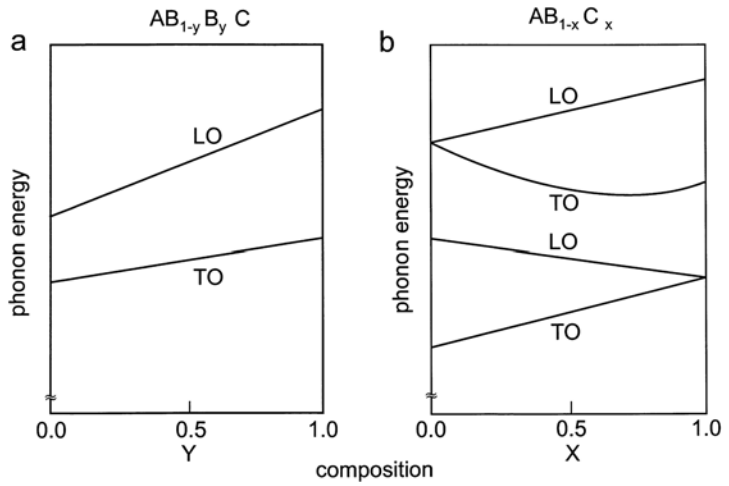


Fig. 7.19. Composition dependence of the LO phonon energies in alloys for the cases of amalgamation type (**a**) and persistent mode type (**b**) behaviour of the phonons

defects like stacking faults and (small angle) grain boundaries, or three-dimensional ones like precipitates or voids. In the present context the most important are the point defects. For an introduction to this topic see e.g. textbooks on solid state physics like [75Z1, 81H1, 81M1, 93K1, 95C1, 95I1, 95W1] of Chap. 1. Point defects include vacancies, interstitital and substitutional atoms. We give some examples for a compound semiconductor AB in Fig. 7.20. We come back to this topic when we discuss electron states in semiconductors. What we need here is the fact that a point defect may have a different mass m' and/or chemical binding, i.e., spring constant D' as compared to the atom which would be at this place in a perfect lattice.

A consequence of such a point defect is that a localized phonon mode may appear. This is a mode which cannot propagate through the sample with a plane-wave factor as in ((7.29a,b) or (7.34)). Instead the amplitude has a maximum at the place of the defect and decays exponentially with increasing distance from it. Obvisouly such a mode is localized at or in the vicinity of the point defect.

If the eigenfrequency ω_{loc} of such a localized phonon mode falls into the bands of the intrinsic acoustic or optic modes and couples to them, it will not produce a big effect. Once such a localized mode is excited, it decays rapidly into bulk modes. The situation is different if ω_{loc} falls either in a spectral region where the pure material has no eigenfrequences at all, or couples only weakly to the bulk modes. Then the localized mode can produce, for example, an additional absorption band or Raman satellite. We come to this point later. The situation

$$\omega_{\mathrm{loc}} > \omega_{\mathrm{bulk}}^{\mathrm{LO}} \tag{7.59}$$

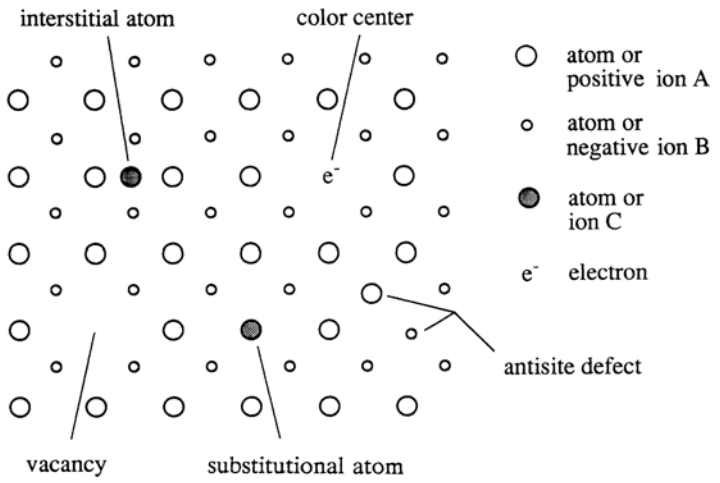


Fig. 7.20. The lattice of a semiconductor AB containing various types of point defects

can be realized for example by incorporating a substitutional atom, which has approximately the same "spring constant" D but a much lighter mass m' than the atom which it replaces, according to (7.37) or Fig. 7.11.

Since the translation invariance of the lattice is destroyed at the point defect, the eigenfrequency ω_{loc} cannot be connected with a well-defined wave vector. There is, however, a possibility to incorporate a localized mode in a dispersion relation based on the following consideration. A localized mode can be constructed in a Fourier-transform-like method by a superposition of bulk modes with appropriate coefficients:

$$u_{\mathrm{loc}}(\boldsymbol{r}) = \sum_{\boldsymbol{k},i} a_{\boldsymbol{k},i} u_{\boldsymbol{k},i}(\boldsymbol{r}) \;, \tag{7.60}$$

where the index i runs over the various branches. Modes which are localized to one unit cell will need contributions from the whole Brillouin zone, while those which are more extended in real space involve contributions from smaller wave vectors only.

We can now indicate a localized mode in the dispersion relation by a horizontal line covering the region of $\boldsymbol{k}$ values that make substantial contributions to the expansion of (7.62). We show such a situation in Fig. 7.21 where the thickness of the horizontal line is related to the amount of the coefficients $|a_{\boldsymbol{k}}|$ in (7.62).

It should be noted that every defect is a scattering center for phonons and contributes, together with the anharmonicity mentioned above, to the finite phase relaxation time of phonons.

To conclude this discussion we would like to stress that even different isotopes act as scattering centers due to their different masses. If, for example, the intrinsic ballistic propagation of phonons is to be investigated, it is

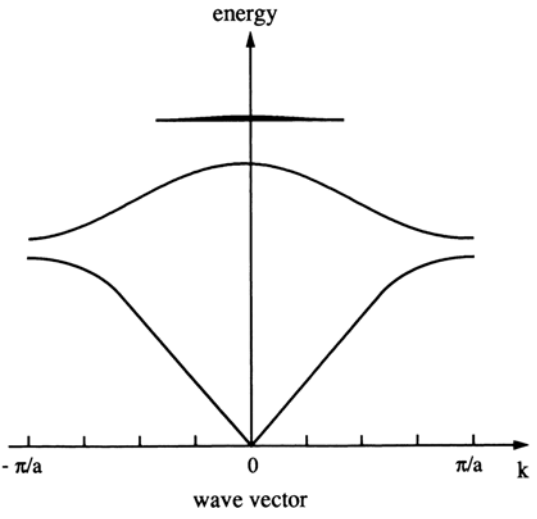


Fig. 7.21. A localized phonon mode represented schematically in a dispersion diagram

desirable to use materials which consist of one isotope only. For more information of this topic see Ref. [81A1]g of Chap. 1 or [94C1] and references given therein.

7.10 Phonons in Superlattices and in other Structures of Reduced Dimensionality

Until now we were mainly considering homogeneous, three-dimensional semiconductors. For the phonons, however, we started with a one-dimensional chain, but we stated that the dimensionality does not have significant influence on the dispersion relation but it does on the density of states, according to what we learned in Sect. 2.6. For the regime of a linear dispersion relation, i.e., for acoustic phonons with not too large wave vectors see Figs. 7.15 – 7.18 we get with (2.77)

$$D(\omega) \sim \omega^{d-1} \tag{7.61}$$

where d is the dimensionality of the system.

Now we want to address in the context of phonons for the first time a rather modern topic in semiconductor physics, namely superlattices.

A superlattice is a man-made periodic structure which consists of thin alternating layers of two different materials, as shown in Fig. 7.22. The different layers are only a few lattice constants thick and can be prepared by various techniques like molecular-beam epitaxy (MBE), metal-organic chemical vapor-phase deposition (MOCVD), hot-wall epitaxy (HWE) or atomic layer epitaxy (ALE). A description of these methods is beyond the scope of this book, but we give some references for the interested reader, e.g., [80M1, 89H1, 89R1]. See also Sects. 8.11 – 8.13. Especially well-suited materials for growing superlattices are the III–V compounds GaAs, AlAs

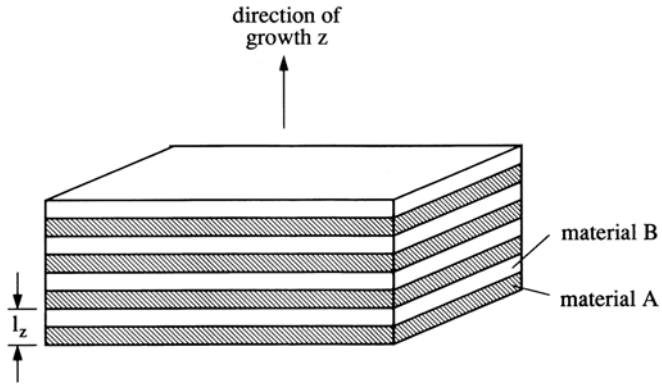


Fig. 7.22. A superlattice consisting of alternating thin layers of two different materials

and their alloys Al_{1-y}, Ga_yAs, since the lattice constant is almost independent of y. Other systems under investigation involve the two elemental semiconductors Si and Ge or ZnSe and $ZnS_{1-x}Se_x$. Due to their different lattice constants, the fabrication of superlattices is in this case more difficult and the layers are strained (strained layer superlattice). For other structures of reduced dimensionality like quantum wells, ∼wires or ∼dots see again Sects. 8.11 – 8.13.

The formation of a superlattice results in strong modifications of the dispersion relation for phonons in the growth direction, i.e. normal to the planes, while the in-plane dispersion is generally at least not quantitatively modified. We concentrate therefore on this direction, which is frequently identified with the z-direction.

As can be seen from Fig. 7.22 we created a new artificial periodicity l_z which is an integer multiple of the (strained) lattice constants of materials A and B

$$l_z = n_1 a_A + n_2 a_B \tag{7.62}$$

This is why such structures are called (artificial) superlattices. Now we use a similar argument as in connection with Fig. 7.14.

The Brillouin zone in the z-direction extends no longer from $-\pi/a \leq k_z \leq \pi/a$ but only from

$$-\frac{\pi}{l_z} \leq k_z \leq \frac{\pi}{l_z} \tag{7.63a}$$

All parts of the dispersion curve outside this interval can be shifted into the first mini-Brillouin zone by vectors of the new reciprocal lattice

$$G_z = l_3 \frac{2\pi}{l_z}; \quad l_3 = 0, \pm 1, \pm 2, \ldots \tag{7.63b}$$

For a single acoustic phonon branch this procedure results in a dispersion relation like in Fig. 7.23, where we show the first few of these so-called backfolded acoustic phonon branches.

The dispersion relation which enters is an average of the dispersion relations of the two materials, weighted by their relative thicknesses (the so-called Rytov-model) [89J1, 96G1, 98G1, 01D1].

Usually one can observe only the few lowest, backfolded acoustic phonon branches. The higher ones are broadened and washed out e.g. due to thickness-fluctuations of the layers.

For the optical phonon branches two different things can happen. If the branches in the two different materials overlap energetically, one obtains in a similar way as for the acoustic phonons mentioned above backfolded optic phonons.

If the do not overlap energetically, TO or LO optical phonon oscillation in one material do not find a resonant partner in the other (and vice versa) similarly to the case of alloys. Consequently the phonons become localized or confined to the respective layer. The elongation-pattern of the atoms has nodes

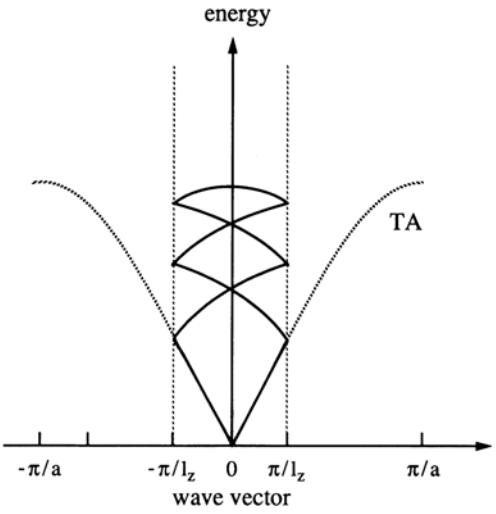


Fig. 7.23. The influence of superlattice formation on the acoustic phonon branch

at the interface or extends only marginally with an exponentially decaying amplitude into the other material as shown schematically in Fig. 7.24a. With minor corrections it can therefore be stated that an integer number of half-waves has to fit into the layer of one material resulting in possible values for k_z in one layer of thickness $n_i a_i$

$$k_z^{(m)} = m\frac{\pi}{n_i a_i}; m = 1, 2, 3, \ldots; \quad i = A, B \tag{7.64}$$

These discrete k_z values are mapped in Fig. 7.24b on the dispersion relation of an optical phonon branch of e.g. material B. The observation of the higher orders of such confined optical phonons allows to investigate the dispersion relation $\hbar\omega\,(k_z)$ over large parts of the first bulk Brillouin zone [89J1, 94C1, 94R1, 98G1, 01D1] and references therein.

If the superlattice is made of a ternary system, e.g. layers AB and AC, generally no new valences and oscillation modes will appear at the interface, which do not already exist in one or the other material. The situation is different for quaternary systems of alternating layers AB and DC. In this case new modes may appear at the interface (here AC and BD oscillations) which exist in neither of the parent compounds. A recently studied system of this type are e.g. CdS/ZnSe superlattices [01D1]: Depending on the chemical termination of the layers, there may be CdSe and ZnS modes. If these modes are situated energetically in a region, where no eigenmodes (confined or backfolded) of AB or DC (here of CdS or ZnSe) occur, these interface modes can exist only at the interface with amplitudes decaying to both sides as shown schematically in Fig. 7.24a. A similar system is BeTe/ZnSe. Interface phonons may also occur at the boundary between a polar material (e.g. GaAs) and a non-polar one (e.g. Si). Details of this situation are discussed e.g. in [96Y1] of Chap. 1.

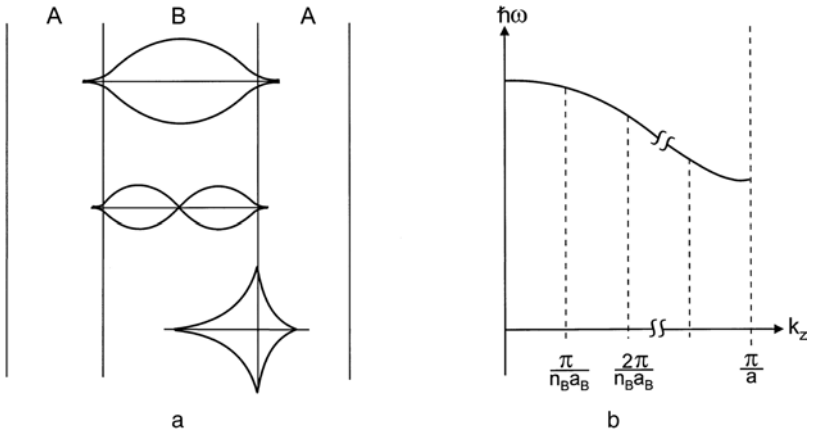


Fig. 7.24. Schematic representation of the envelope of an optical phonon mode confined to material B in a superlattice and of an interface phonon (**a**) and the discrete values of k_z on an optical phonon branch of material B. (**b**)

As a result, we may summarize that in superlattices there may be as new phenomena backfolded or confined phonon modes as well as interface modes.

To conclude this subsection we mention that in the extreme confinement case of phonons in all three directions of space, which may occur in so called quantum dots (also known as quantum boxes, nano-crystallites or nano-islands, depending partly on the way of creation (see Sect. 8.13), the dispersion relation of the phonons breaks up into a series of discrete modes, which is governed by the boundary conditions (e.g. free or clamped). See e.g. [92N1, 93T1, 95O1, 96W1, 98A1] and references therein.

7.11 Problems

1. Inspect or build some lattice and crystal models to become familiar with the topics presented in Sect. 7.2.
2. Show that the reciprocal lattice of a face-centered cubic (fcc) lattice is a body centered cubic (bcc) lattice and vice versa.
3. Calculate and draw the unit cell and the Wigner-Seitz cell in real space and the first three Brillouin zones in reciprocal space for a simple cubic and a hexagonal two-dimensional lattice.
4. Study with the data found in [82L1] of Chap. 1 the transverse and longitudinal eigen-frequencies of optical phonons of a series of semiconductors like ZnO, ZnS, ZnSe, and ZnTe, or of a similar series with the same anion and different cations. What do you conclude?
5. Compare the longitudinal-transverse splitting at $k = 0$ for Ge and the corresponding III–V, II–VI and I–VII compounds on the same line of

the periodic table of the elements. Can you qualitatively explain the findings?

6. Inspect a model of a cubic crystal (e.g. zinc-blende). Find the non primitive cubic unit cell and the primitive one. Explain qualitatively that such a crystal should be optically isotropic for light propagating ($\boldsymbol{k} \neq 0$!) e.g. in the directions (100) or (111) but not in (110).

References to Chap. 7

[80M1] Molecular Beam Epitaxy, B.R. Pamplin (ed.), Pergamon, Oxford (1980)

[89H1] M.A. Hermann, H. Sitter, Molecular Beam Epitaxy, Springer, Berlin, Heidelberg (1989)

[89J1] B. Jusserand and M. Cardona, in Topics in Applied Physics **66**, Springer, Berlin (1989)

[89R1] M. Razeghi: The MOCVD Challenge, Hilger, Bristol (1989) and references therein

[92N1] S. Nomara, T. Kobayashi, Solid State Commun. **82**, 335 (1992)

[93T1] A. Tomaka, S. Onari and T. Arai, Phys. Rev. B **47**, 1237 (1993)

[94C1] M. Cardona, Festkörperprobleme/Advances in Solid State Physics **34**, 35 (1994)

[94R1] T. Ruf et al., Festkörperprobleme/Advances in Solid State Physics **34**, 237 (1994)

[95O1] S. Okamoto, Y. Masumot, J. Luminesc. **64**, 253 (1995)

[96G1] J. Geurts, Progr. Cryst. Growth Charact. **32**, 185 (1996)

[96W1] U. Woggon et al., Phys. Rev. B **54**, 1506 (1996)

[98A1] A.P. Alivisatos et al., J. Chem. Phys. **89**, 4001 (1998)

[98G1] M. Göppert et al., Phys. Rev. B **57**, 13068 (1998)

[01D1] A. Dinger et al., Phys. Rev. B **64**, 245310 (2001)

8

Electrons in a Periodic Crystal

Let us now return for a moment to the Hamiltonian (7.1) which describes a crystal. By introducing the phonons we have treated the motion of the atoms. We now assume that the atoms are fixed at their equilibrium positions, i.e., that no phonons are excited. Then the sum over the potentials of all atoms forms a periodic potential for the electrons; but we are still left with a Hamiltonian for about 10^{23} interacting outer electrons per cubic centimeter, which should be properly treated in a many particle formalism. Unfortunately it is extremely difficult to handle this approach. Instead one generally uses a so-called one-electron approximation. The idea is the following: One assumes that the periodically arranged atoms and all interaction potentials between electrons together form a periodic potential $V(\boldsymbol{r})$ with

$$V(\boldsymbol{r}+\boldsymbol{R}) = V(\boldsymbol{r})\,. \tag{8.1}$$

where $\boldsymbol{R}$ is a translation vector of the lattice (see Sect. 7.2).

This is a valid approach especially for semiconductors or simple metals. In semiconductors there is only a low density (generally between 10^9 and $10^{19}\,\mathrm{cm}^{-3}$) of free electrons (or holes) that can react easily on their mutual interaction. In simple metals the density of free carriers is higher, but the density of those reacting on their mutual interaction is limited to the fraction situated in an interval of width $\Delta E \approx 4k_\mathrm{B}T \;<<< E_\mathrm{F}$ energetically around the Fermi energy. It must be noted that there are other systems in which the mutual interaction between carriers plays a crucial role like in the Hubbard model, in heavy Fermion systems or in connection with the Kondo effect. These topics are, however, beyond the scope of this book but we provide some references for the interested reader [93F1, 93H1].

We return now to the approximation (8.1). One calculates the eigenstates for one electron of the corresponding Schrödinger equation and populates these states with electrons according to Fermi–Dirac statistics until all electrons have been accommodated. The potential $V(\boldsymbol{r})$ should ideally be calculated in a self-consistent way e.g. by Hartree or Hartree–Fock approaches. But simpler

and more feasible approaches are often used such as a periodic arrangement of screened Coulomb potentials or a "muffin-tin" potential.

It is beyond the scope of this book to present details of band structure calculations. There are excellent reviews and (text)books on this topic with various degrees of sophistication, some of which were already cited in Chap. 1 like [75B1, 75Z1, 81M1, 81S1, 91E1, 93K1, 95C1, 95I1, 96S1, 96Y1, 97S1, 99B1, 01H2,01S1] in addition to [70M1,88C1,88R1,91E1] of this chapter, from which much of the following material has been taken. On the other hand, the topic is crucial for the understanding of the optical properties of the electronic subsystem of semiconductors. Therefore we first summarize in the following the main results and explain the existence of metals, semiconductors and insulators for those who do not want to go deeper into details, then we give an overview on the various approaches to the bandstructure calculations starting, e.g., from free electrons or from atomic orbitals and including the basic concept of $\boldsymbol{k} \cdot \boldsymbol{p}$ theory. The next sections are devoted to consequences of the band structure of semiconductors like the concepts of crystal electrons and holes as new quasiparticles, the effective mass approach, realistic semiconductor band structures and others, before we concentrate on structures of reduced dimensionality.

8.1 Bloch's Theorem

There are essentially two approaches to the problem of band structure calculations. In one case one starts with free electrons, which have the simple parabolic dispersion relation in the non-relativistic regime

$$E(\boldsymbol{k}) = \frac{\boldsymbol{p}^2}{2m_0} = \frac{\hbar^2 \boldsymbol{k}^2}{2m_0} \tag{8.2}$$

shown in Fig. 8.1a by the dashed line. One introduces a weak periodic (see (8.1)) potential as a perturbation. These techniques include the "empty lattice", the nearly-free-electron (NFE), augmented plane wave (APW) and orthogonalized plane wave (OPW) approaches. In the last two cases, terms are added to make sure that the plane waves are orthogonal to the deeper atomic levels. These terms are treated as potentials and are known as pseudo potentials. These pseudo potentials may be energy dependent.

The introduction of such a weak periodic potential does not alter the dispersion relation very much, but results in the formation of energy gaps at the boarders of the Brillouin zones in which no stationary electron states exist (see Fig. 8.1a). We will explain this point in more detail in Sect. 8.7.

The other group of methods starts with the atomic orbitals of the atoms forming the solid or more specifically the semiconductor. They involve the summation over one or more atomic orbitals placed at every atom site in the crystal and treat the interaction between the orbitals at neighbouring sites, i.e., the wave function overlap as perturbation. These techniques include, e.g.,

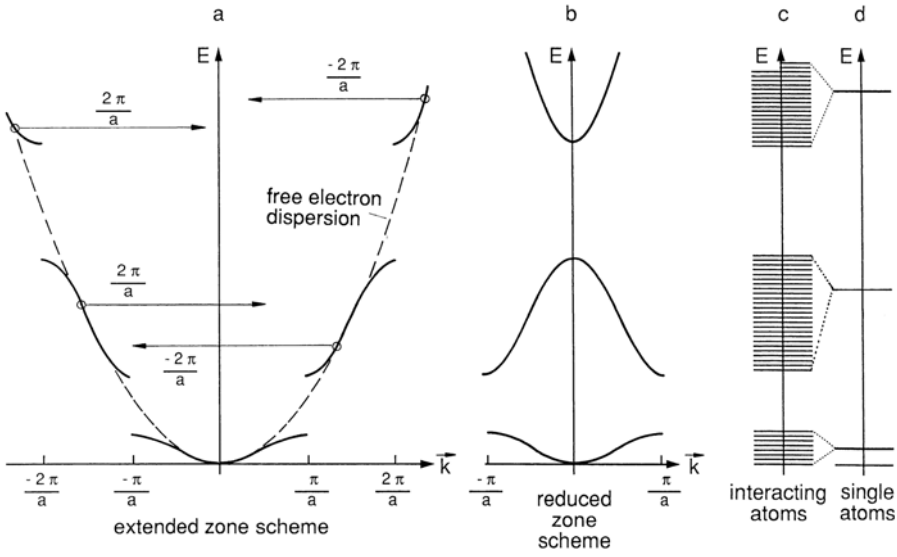


Fig. 8.1. The reduced-zone scheme (**b**) of electronic energy bands in a crystalline solid reached either starting from (nearly) free electrons (**a**) or from atomic orbitals (**c**, **d**)

the tight-binding approximation, the linear combination of atomic orbitals (LCAO) or the extended Hückel method (see Fig. 8.1c,d).

This interaction between atoms leads to a splitting of the atomic level in as many states as there are atoms (Fig. 8.1c,d) similarly as in classical mechanics where coupling between two identical oscillators leads to two close-lying eigenfrequencies, three oscillators to three eigenfrequencies and 10^{23} oscillators to 10^{23} eigenfrequencies. The level spacing is so small that for all practical purposes the result is a continuous energy band. For some more details also see Sect. 8.7.

The first group of band structure calculation methods is usually more suitable for conduction band states of semiconductors the second one for valence band states. For the meaning of these two terms in the context of semiconductors see Sect. 8.2.

Furthermore, there is an increasing trend to start less with intuitive physical considerations but to trust more the increasing power of computers to handle many particle problems.

Independent from the individual approach, one finds in all cases that electron states in a periodic potential are energetically arranged in energy bands of a certain width, which may be separated by gaps in which no stationary eigenstates exist (Fig. 8.1). The electron eigenstates of a periodic potential are so-called Bloch waves $\phi_{\boldsymbol{k},i}(\boldsymbol{r})$,

$$\phi_{\boldsymbol{k},i}(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{r}} u_{\boldsymbol{k},i}(\boldsymbol{r})\,, \tag{8.3a}$$

where

$$u_{\boldsymbol{k},i}(\boldsymbol{r}) = u_{\boldsymbol{k},i}(\boldsymbol{r} + \boldsymbol{R})\,. \tag{8.3b}$$

The $\phi_{\boldsymbol{k},i}$ are evidently a product of a plane wave and a lattice periodic term $u_{\boldsymbol{k},i}$ (8.3b), where $\boldsymbol{k}$ is the wave vector and i the index of the band, as shown in Fig. 8.2. The eigenenergies in the bands depend both on $\boldsymbol{k}$ and i and are periodic in $\boldsymbol{k}$-space, i.e.,

$$E_i(\boldsymbol{k}) = E_i(\boldsymbol{k} + \boldsymbol{G}) \tag{8.4}$$

In a similar way one finds that

$$\phi_{\boldsymbol{k},i}(\boldsymbol{r}) = \phi_{\boldsymbol{k}+\boldsymbol{G},i}(\boldsymbol{r})\,. \tag{8.5}$$

These are the fundamental points of Bloch's theorem. Since the $u_{\boldsymbol{k},i}(\boldsymbol{r})$ remember the wave function of the parent atoms and the exponential term the plane wave character, the Bloch functions (8.3) incorporate the two limiting approaches (e.g., NFE and LCAO) mentioned above.

The Bloch theorem is also known as the Ewald–Bloch theorem since Ewald has found almost identical rules for the propagation of X-rays in crystals (see textbooks on solid state physics).

The statement of (8.4) once more allows the dispersion relation $E(\boldsymbol{k})$ to be reduced to the first Brillouin zone (Fig. 8.1a,b) in a similar way as for phonons in Sect. 7.4. Actually there are various possibilities to present the electronic band structure, which are shown in Fig. 8.3. See also Fig. 8.5.

Figure 8.3a shows the extended zone scheme in which various branches of the dispersion relation $E(\boldsymbol{k})$ are situated in various Brillouin zones. Equation (8.4) allows to shift the outer branches into the first Brillouin zone with suitable vectors of the reciprocal lattice $\boldsymbol{G}$ (here $\boldsymbol{G} = \pm 2\pi/a$) resulting in the reduced zone scheme of Figs. 8.1b or 8.3b. Alternatively we may repeat all branches periodically over the various Brillouin zones as shown in Fig. 8.3c.

Usually one uses the reduced scheme of Fig. 8.3b. The extrema of the various bands tend to occur either at the centre of the first Brillouin zone,

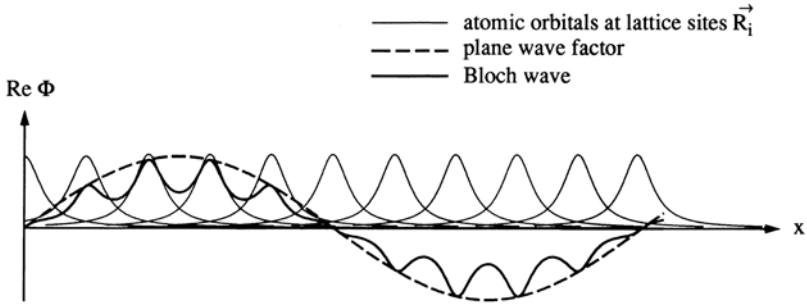


Fig. 8.2. Schematic drawing of the real (or imaginary) part of a Bloch wave in one dimension

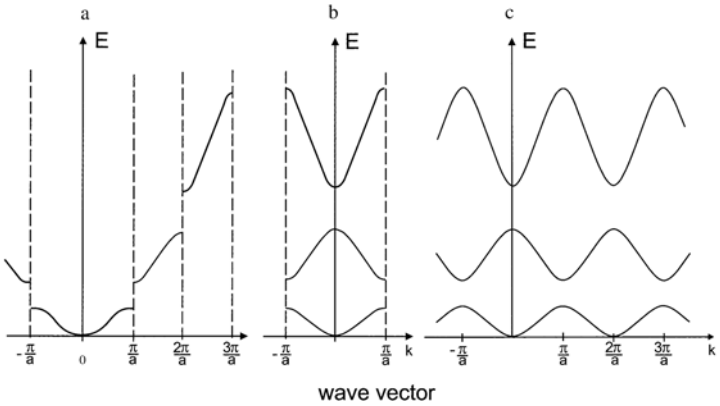


Fig. 8.3. The extended (**a**) the reduced (**b**) and the periodic band structure (**c**)

i.e., at $\boldsymbol{k} = 0$ (the so-called Γ-point) or at the boarders of the Brillouin zone as shown in Fig. 8.3b, but as we shall see later, extrema may also occur somewhere else on the interval $0 \leq |k| < \pi/a$. The bands tend to be to a good approximation parabolic in the vicinity of the extrema but the positive or negative curvature can be different at every extremum and may even depend on the direction of $\boldsymbol{k}$ in systems of higher dimensionality than one including three. When approaching the boarder of the Brillouin zone, the slope of $E(\boldsymbol{k})$ normal to this boarder is generally zero.

With increasing energy the width of the allowed bands tends to increase, while the width of the forbidden gaps decreases. It is even possible that various bands overlap energetically.

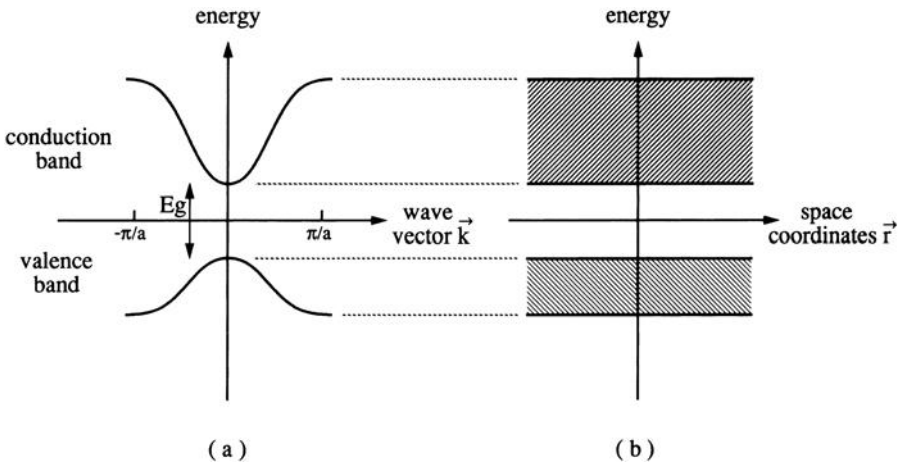


Fig. 8.4. The representation of bands in $\boldsymbol{k}$-space (**a**) and in real space (**b**)

We will give some arguments for this behavior later in this chapter.

To conclude this section we want to mention that the bands are usually represented in one of two different ways. The first is the dispersion relation (Figs. 8.1b, 8.3, 8.4a). In the other case one plots the width of the bands and of the gaps between them as a function of the space coordinate $\boldsymbol{r}$ (Fig. 8.4b). The latter is especially useful to demonstrate spatial inhomogeneities or localized states. We shall meet some examples of both cases later.

8.2 Metals, Semiconductors, Insulators

Having obtained a first insight into the electronic bandstructure in the above section, we now want to make the second step and put all electrons into the bands using, of course, Fermi–Dirac statistics. We further assume zero temperature.

Some of the energetically lower-lying bands will be completely filled. We call bands which are completely filled at $T = 0\,\mathrm{K}$ "valence bands", while all partly filled or empty bands are "conduction bands".

The important region for our purposes is that around the highest valence and the lowest conduction bands. This region determines not only the optical properties around the fundamental absorption edge, but also the magnetic properties and the electronic contributions to the conductivities of electricity and of heat.

If the filling procedure of the bands ends in such a way that, at $T = 0\,\mathrm{K}$ there are one or more partly-filled conduction bands (Fig. 8.5a,b), we have a metal. This situation arises for example if the atomic orbital which forms the band is itself only partly occupied (e.g., the outer s-level of the alkali metals Li, Na ...) (Fig. 8.5a) or if a completely filled orbital forms a band which overlaps with a band stemming from an empty atomic orbital, as is the case in the rare earth metals (Ca, Mg ...) in Fig. 8.5b. If, on the other hand, the filling procedure gives one or more completely-filled valence bands which are separated by a gap E_g from completely empty conduction bands, we have a semiconductor for

$$0 < E_\mathrm{g} \leq 4\,\mathrm{eV} \tag{8.6a}$$

and an insulator for

$$E_\mathrm{g} \geq 4\,\mathrm{eV}\,. \tag{8.6b}$$

The "boarder line" of 4 eV is set by convention and is not sharp.

Diamond has, e.g., a gap of 5.5 eV, but is still considered to be a semiconductor, especially because it can be n or p doped.

Materials with

$$E_\mathrm{g} = 0\,, \tag{8.7a}$$

i.e., materials in which the lowest conduction and uppermost valence band "touch" each other but do not overlap are called semimetals.

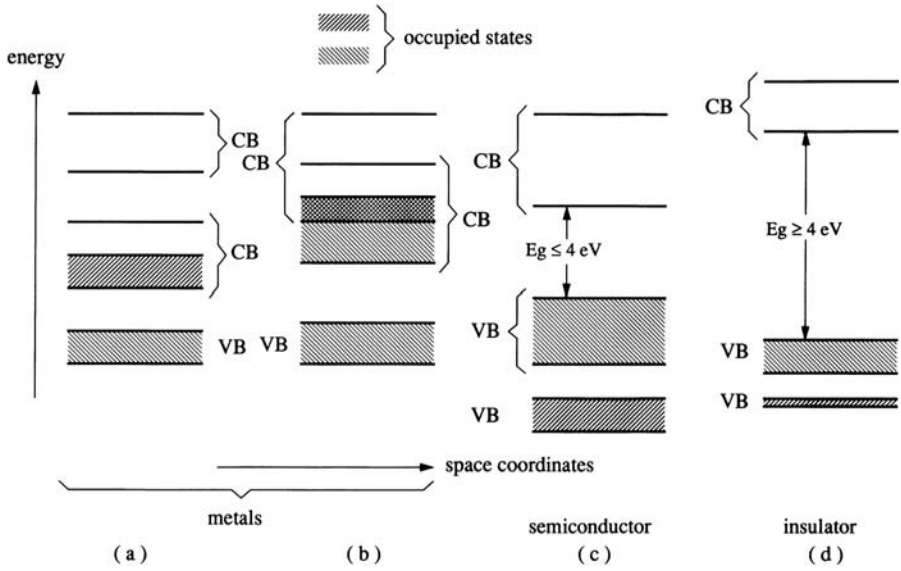


Fig. 8.5. Occupation of the bands for a metal (**a**,**b**) a semiconductor (**c**) and an insulator (**d**)

Semiconductors with

$$0 < E_g \leq 0.5\,\text{eV} \tag{8.7b}$$

are called narrow gap semiconductors.

In the range

$$0.5\,\text{eV} < E_g \leq 2\,\text{eV}$$

we find the usual semiconductors, from which many are of wide technical importance like Ge, Si or GaAs. The regime

$$2\,\text{eV} < E_g \leq 4\,\text{eV} \tag{8.7c}$$

characterizes the wide gap semiconductors, which lead continuously to the insulators.

To conclude this section we want to introduce the terms "direct" and "indirect" gap semiconductors. A semiconductor is said to have a direct gap, or to be a direct semiconductor, if the global maximum of the highest valence band and the global minimum of the lowest conduction band occur at the same point of the first Brillouin zone in the reduced zone scheme as in Fig. 8.4a. This point is frequently but not necessarily the Γ-point (i.e., $\boldsymbol{k} = 0$). There are also direct gap semiconductors where the gap occurs at a specific point on the boarder of the Brillouin zone. Semiconductors, in which the two above-mentioned extrema occur at different $\boldsymbol{k}$ values in the Brillouin zone have an indirect gap or are called indirect semiconductors. The reason for this terminology comes from optical transitions between the band extrema.

A photon with an energy equivalent to the typical width of a semiconductor band gap has (almost) vanishing momentum on the scale of the first Brillouin zone. Optical transitions are in this sense vertical in a $E(\boldsymbol{k})$ relation. If the band extrema occur at the same point in $\boldsymbol{k}$-space, a transition between the extrema is directly possible by absorption (or emission) of a photon. In the other case this transition is forbidden by the $\boldsymbol{k}$-conservation law and only indirect transitions are possible, which also involve the absorption or emission of a phonon for momentum conservation.

We shall see examples for both cases in Sect. 8.8.

8.3 An Overview of Semiconducting Materials

We now understand that semiconductors are solids that have at $T = 0\,\mathrm{K}$ a series of completely filled valence bands. The uppermost band is separated by an energy gap E_g from the empty conduction bands.

Since a crystal can be considered a huge molecule, we should also give the terms used in chemistry. The uppermost valence band is known in this scientific discipline as highest occupied molecular orbital (HOMO). The lowest conduction band is called consequently the lowest unoccupied molecular orbital (LUMO). If not stated explicitly otherwise, when we use valence and conduction band we always mean the highest and lowest ones, respectively.

Presently more than 600 element and compound semiconductors are known. They are listed with their propeties in the volumes by Landolt–Börnstein ([82L1] of Chap. 1).

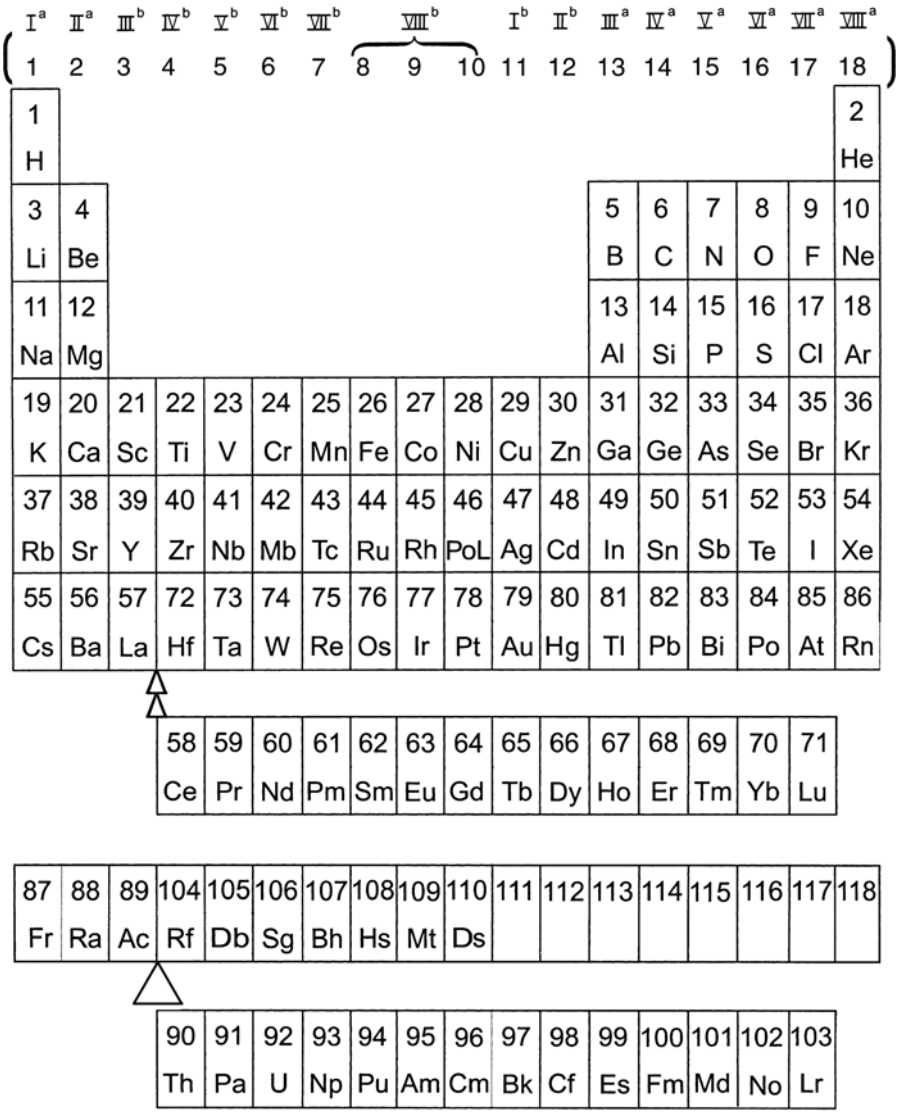
With the help of the periodic Table 8.1 we try to localize the most important semiconductors. In the first two lines we give the current internationally recommended numbering system for the columns used by chemists, which runs from 1 to 18, and the old one, which we and most other semiconductor physicists use. It runs from I o VIII and has subclassifications such as $\mathrm{II^A}$ and $\mathrm{II^B}$.

The technically most important semiconductor is Si. It is found in column $\mathrm{IV^A}$. Conduction and valence bands are formed from the antibonding and bonding sp^3 hybrid orbitals. The binding is completely covalent and E_g is around 1.1 eV at room temperature. See, e.g., Table 8.2 where we give some data, which we partly explain later. Silicon crystallizes in the diamond structure with point group 0_h. The modification diamond of carbon is a semiconductor as well as Ge.

Tin is usually a metal, but there exists a modification, which is a semimetal namely grey tin, which also crystallizes in the diamond structure. Lead, finally, is a metal.

We can already see a general trend, namely that the width of the band gap decreases if one goes down the columns. SiC also belongs to the group IV semiconductors. SiC is notorious for forming many different polytypes. The

Table 8.1. Periodic table of the elements. The names of elements 104 to 110 might change and others may be discovered in the future.



indirect band gaps of all of them are around 3 eV, i.e., they lie between the gaps of diamond and Si.

Carbon has two further modifications namely graphite and fullerene (C_{60}). Graphite is a semimetal with rather strong covalent sp^2 hybrid binding within the hexagonal layers and weak van der Waals binding between the layers. C_{60} is a semiconductor with $E_g \approx 2.2$ eV.

In the diamond structure every atom is tetrahedrally surrounded by its nearest neighbors. See Fig. 7.2. We can now replace, e.g., the Ge atoms on one sublattice by Ga, which has one outer electron less, and the others by As, which has one electron more than Ge, leaving the total number of electrons per unit unchanged, but adding a little bit of ionic binding to the still dominant covalent one. This procedure leads to the so-called zincblende type crystal structure with point group T_d and to the group of III–V semiconductors. To this group of semiconductors belong the compounds of B, Al, Ga or In with N, P, As or Sb. We see again the general trend of the band gap with BN being an insulator, AlN, GaN being wide gap semiconductors while InSb is a narrow gap semiconductor.

The group III nitrides crystallize preferentially in the hexagonal wurtzite type structure (point group C_{6v}). In this case every atom of one type is still tetrahedrally surrounded by the others, but the arrangement of the next-nearest neighbours is such, that a hexagonal structure evolves (see Fig. 7.2).

If we repeat this step, which leads from the group IV semiconductors to the III–V compounds, once more or even twice more, we come to the II–VI (more precisely $\mathrm{II^B}$–$\mathrm{VI^A}$) and I–VII (more precisely $\mathrm{I^B}$–$\mathrm{VII^A}$) semiconductors, with increasing and finally dominating ionic binding, but still generally tetrahedral coordination. The $\mathrm{II^A}$–$\mathrm{VI^A}$ and $\mathrm{I^A}$–$\mathrm{VII^A}$ components are usually insulators.

The II–VI semiconductors comprise the compounds of Zn, Cd or Hg with O, S, Se or Te. The gap generally decreases when going down the columns ZnS, ZnO and CdS are wide gap semiconductors, while the mercury compounds are usually semimetals. The II–VI semiconductors usually crystallize in zincblende or wurtzite type structures with a few exceptions. Both structures are partly possible with only minor energetic differences, as is the case for ZnS.

Some compounds have other structures like CdO (rocksalt structure), HgO (rhombohedral). HgS is a semimetal in the zincblende structure but a semiconductor with a gap around 2.2 eV in the trigonal modification (red cinnabar).

The main I–VII compounds are listed in Table 8.1. Not much is known about the fluorides or the Au^+ halides concerning their properties as semiconductors.

Until now the list included only elements or binary compounds. In a similar way as above one can come to ternary semiconductors like $CuGaSe_2$ or even to quaternary ones like Ag_2CdGeS_4.

Furthermore, some of the elements and many of the binary compounds partially form alloys even without a miscibility gap like $Si_{1-x}Ge_x$, $Ga_{1-y}Al_yAs$, $CdS_{1-x}Se_x$, $ZnSe_{1-x}Te_x$ or $Cd_{1-y}Hg_yTe$. In an alloy one still has, in principle, nice crystals with a periodic lattice structure, but the lattice sites of one sublattice are randomly occupied by the two different atoms ($Si_{1-x}Ge_x$), anions (e.g., $CdS_{1-x}Se_x$) or cations (e.g., $Ga_{1-y}Al_yAs$). However, on a microscopic scale, the concentration fluctuations of the composition x introduce some disorder.

Some alloys tend to form ordered structures for compositions close to 0.5 like $Ga_{0.5}In_{0.5}P$ adopting the so-called CuPt structure [95C1].

Table 8.2. Values of the band gap energy of some group IV, III–V, II–VI and I–VII semiconductors (SC) at low temperature ($T \lesssim 10\,\mathrm{K}$), Sy: Symmetry (point group). According to [82L1] of Chap. 1

Group IV				Group III–V				Group II–VI				Group I–VII			
SC	Sy	E_g (eV)	dir/indir	SC	Sy	E_g (eV)	dir/indir	SC	Sy	E_g (eV)	dir/indir	SC	Sy	E_g (eV)	dir/indir
C	O_h	5.48	i	AlN	C_{6v}	6.28	d	ZnO	C_{6v}	3.437	d	CuCl	T_d	3.395	d
Si	O_h	1.17	i	AlP	T_d	2.53	i	ZnS	C_{6v}	3.91	d	CuBr	T_d	3.077	d
Ge	O_h	0.744	i	AlAs	T_d	2.228	i	ZnS	T_d	3.78	d	CuI	T_d	3.115	d
grey Sn	O_h	0	semimetal	AlSb	T_d	1.696	i	ZnSe	T_d	2.82	d	AgCl	T_d	3.249	i
				GaN	C_{6v}	3.503	d	ZnTe	T_d	2.391	d	AgBr	T_d	2.684	i
				GaP	T_d	2.350	i	CdO	O_h	0.8	i	AgI	C_{6v}	3.024	d
				GaAs	T_d	1.518	d	CdS	C_{6v}	2.583	d				
				GaSb	T_d	0.812	d	CdSe	C_{6v}	1.841	d				
				InN	C_{6v}	2.2	d	CdTe	T_d	1.60	d				
				InP	T_d	1.424	d	HgS	T_d	0	d				
				InAs	T_d	0.418	d	HgSe	T_d	0	d				
				InSb	T_d	0.237	d	HgTe	T_d	0	d				

Alloying is also possible on both sublattices like in $Ga_{1-y}In_yN_xAs_{1-x}$.

Most of the examples in the rest of this book will be taken from these more common semiconductors, but there are many more, some of which we mention below.

The IV–VI compounds (also known as lead salts) include the compounds of Pb or Sn with S, Se and Te. They serve partly as IR laser diodes.

There are further elemental semiconductors like S, Se, Te some modifications of P or I (As and Sb are considered semimetals). There exist various oxides as semiconductors apart from the group II^B oxides like GeO_2, SnO_2 (SiO_2 = quartz is an insulator); Cu_2O; TiO_2 in its various modifications (anatase, rutil, brookit) or the highly poisonous Tl–halides. To conclude this section we mention organic semiconductors like crystals of anthracene ($C_{14}H_{10}$), pentacene ($C_{22}H_{14}$), dibenzothiophene $C_{12}H_8S$ and hexathiophene. Organic semiconductors do not fall within the focus of this book, but we will occasionally give examples of their optical properties.

There is a general trend that the band gap of semiconductors decreases with increasing temperature. The decrease $E_g(T = 0) - E_g(T) = \Delta E_g(T)$ tends to vary quadratically with temperature at lower temperatures ($T \leq 100\,\mathrm{K}$) and linearly above. This behavior is often described by the Varshni formula [67V1]

$$\Delta E_g\,(T) = \frac{\alpha T^2}{\beta + \gamma T}\,. \tag{8.8}$$

More complex formula were recently suggested, e.g., in [94A1, 02G1, 03G1] and references therein.

Some semiconductors like CuCl or some lead salts also show an increase in E_g with increasing temperature, for some others like CuBr $E_g(T)$ goes through a maximum with increasing temperature. For data see, e.g., in [82L1] of Chap. 1 and references therein.

8.4 Electrons and Holes in Crystals as New Quasiparticles

As we shall see later, the optical properties of the electronic system of semiconductors are largely determined by transitions of electrons between the upper valence bands and the lower conduction bands.

The bandstructure as presented until now, i.e., in connection with Figs. 8.1 – 8.5, describes the so-called $N \pm 1$ particle problem in the following sense: if we consider a semiconductor with a completely filled valence band containing N electrons per cm^3

$$N \simeq 10^{22} - 10^{23}\,\mathrm{cm}^{-3} \tag{8.9}$$

and a completely empty conduction band and add one more electron, we find that this electron can be placed into exactly the conduction band states. If

Table 8.3. Properties of a hole in the valence band compared to the properties of the electron that has been removed from the valence band to create the hole

Property	Hole	Removed electron
electric charge q	q_h	$= -q_{re}$, $q_{re} \approx -1.6 \times 10^{-19}$ As
wave vector	$\boldsymbol{k}_\mathrm{h}$	$= -\boldsymbol{k}_{re}$
spin	σ_h	$= -\sigma_{re}$
eff. mass	$m_\mathrm{h} > 0$	$= -m_{re}$, $m_{re} < 0$

we remove one of the N electrons and ask from which state it came, we find the valence band states.

An obvious step now is to consider the one or few electrons in an otherwise empty conduction band (CB). For an almost filled valence band, however, it is easier to consider the few empty states and their properties instead of the many occupied ones. This idea leads to the concept of "defect-electrons" or "holes". The properties of the hole are connected in the following way (Table 8.3) with the properties of the electron that has been removed from the valence band (VB).

From Table 8.3 we see that the hole has a positive charge and that its wave vector and spin are opposite to those of the electron removed from the valence band. The two latter statements are easy to understand. A semiconductor with a completely filled valence band has total momentum and spin equal to zero. If we take one particle out, the remainder acquires for the above quantities values exactly opposite to those of the removed particle. For clarity Figs. 8.6a,b show the bandstructure containing one electron in the conduction band and one hole in the valence band, respectively. The states are equidistant in $\boldsymbol{k}$ (see Sect. 2.6) but we should note that there are usually 10^{22}–10^{23} states in each band per cm^3 and not only the few shown in Fig. 8.6.

The electrons and holes in a semiconductor crystal are quasi-particles. They can exist only in the crystal and not in vacuum, in contrast to normal electrons and positrons with which they have a lot in common, except the magnitude of the energy gap which is $\simeq 1$ MeV for normal electrons and positrons i.e., twice the rest mass of 511 keV $= m_0 c^2$. The dispersion relations of electrons and holes are different from those of free electrons and positrons which for the non-relativistic case are given by

$$E_\mathrm{e,p} = \pm \left(m_0 c^2 + \frac{\hbar^2 k^2}{2m_0} \right) , \tag{8.10}$$

where m_0 is the free electron mass.

The quantity $\hbar \boldsymbol{k}_\mathrm{e,h}$ of crystal electron and hole is a quasi-momentum, since it is conserved only modulo reciprocal lattice vectors – see (8.4),(8.5) – and since the Bloch waves of (8.2),(8.3) are not proper eigenstates of the momentum operator $\frac{\hbar}{\mathrm{i}} \nabla$. For more details of the concept of quasi-momentum see [98B1] of Chap. 5.

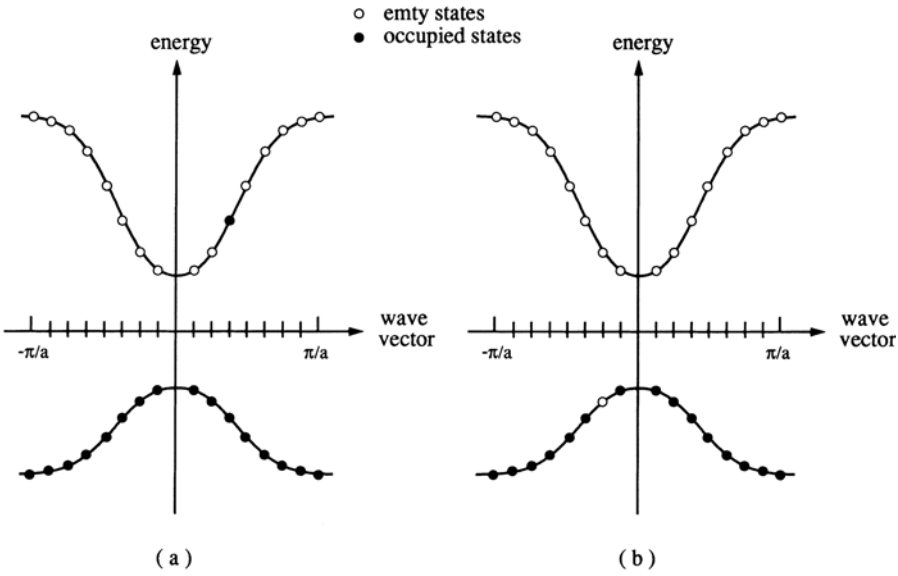


Fig. 8.6. One electron in the conduction band (**a**) and one hole in the valence band (**b**). *Full circles:* occupied states; *open circles:* empty states

One should note that the energy of a hole increases if it is brought deeper into the valence band.

8.5 The Effective-Mass Concept

If we want to describe the motion of an electron or hole in a semiconductor under the influence of an external field (e.g., an electric or magnetic field) it is intuitively clear that we ought to consider a wave packet rather than the infinitely extended Bloch waves. To describe such a wave packet we superpose Bloch waves of a certain range of $\boldsymbol{k}$-vectors around a $\boldsymbol{k}_0$ as described schematically in (8.11)

$$\phi_{\boldsymbol{k}_0}(\boldsymbol{r}) = \sum_{\boldsymbol{k}} a_{\boldsymbol{k}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} u_{\boldsymbol{k}}(\boldsymbol{r})\,. \tag{8.11}$$

These types of wave packet are known as Wannier-functions. In order to keep the $\boldsymbol{k}$-vector reasonably well defined, we localize the wavefunction only to a volume larger than a unit cell. Due to the uncertainty relation we would need wavefunctions from the whole Brillouin zone if we wanted to localize an electron to within one unit cell.

An external force $\boldsymbol{F}$ changes the energy of the wave packet according to

$$\mathrm{d}E(\boldsymbol{k}) = \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{s} = \boldsymbol{F}\cdot\boldsymbol{v}_{\mathrm{g}}\,\mathrm{d}t \tag{8.12}$$

with the group velocity according to (2.13)

$$\boldsymbol{v}_{\mathrm{g}} = \frac{1}{\hbar}\mathrm{grad}_{\boldsymbol{k}} E(\boldsymbol{k}) \tag{8.13a}$$

or simplified

$$\boldsymbol{v}_{\mathrm{g}} = \frac{1}{\hbar}\frac{\mathrm{d}E}{\mathrm{d}k} \tag{8.13b}$$

Since, on the other hand, we have

$$\mathrm{d}E(\boldsymbol{k}) = \frac{\mathrm{d}E(\boldsymbol{k})}{\mathrm{d}k} \cdot \mathrm{d}\boldsymbol{k} = \hbar \boldsymbol{v}_{\mathrm{g}}\, \mathrm{d}\boldsymbol{k}\,, \tag{8.13c}$$

we find from combining (8.12) and (8.13c)

$$\hbar\frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}t} = \boldsymbol{F} = \dot{\boldsymbol{p}} \tag{8.13d}$$

This expression corresponds to Newton's law of motion, but now for the quasi-momenta $\hbar\boldsymbol{k}$ of the crystal electrons or holes.

From (8.13b), (8.13d) we get for the acceleration $\boldsymbol{a}$ of the wave packet

$$\boldsymbol{a} = \frac{\mathrm{d}\boldsymbol{v}_{\mathrm{g}}}{\mathrm{d}t} = \frac{1}{\hbar}\frac{\partial^2 E}{\partial k \partial t} = \frac{1}{\hbar}\frac{\partial^2 E}{\partial k^2}\frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}t} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k^2}\boldsymbol{F}\,. \tag{8.13e}$$

Comparing with the trivial form of (8.13d)

$$\boldsymbol{a} = \frac{1}{m}\boldsymbol{F} \tag{8.13f}$$

we find that the crystal electron and hole move under the influence of external fields through the crystal like a particle, however, with an effective mass given by

$$\frac{1}{m_{\mathrm{eff}}} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k^2} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k_i \partial k_j}\,; \quad i,j = x,y,z\,. \tag{8.14}$$

For free electrons and positrons (8.14) leads to $m_{\mathrm{eff}} = m_0$.

The right-hand side of (8.14) snows that the effective mass is actually a tensor and can depend on the direction in which the electron or hole moves.

The important point now is that the bands of semiconductors tend to be parabolic in the vicinity of the band extrema, as shown schematically in Fig. 8.7 or in the real band structures discussed below with Figs. 8.9–8.12. These extrema are most important for the optical and transport properties. The effective masses are constant in these regions. This leads to the so-called effective-mass approximation. Electrons and holes in a semiconductor are simply treated as free particles, but with an effective mass given by (8.14).

We should mention that the mass of an electron is positive if the curvature of the band is positive. Due to the change of the sign of the properties of holes compared to those of the missing electron in the valence band (Table 8.3) the

mass of the hole is positive at the maximum where the curvature is actually negative.

The explanation behind the effective masses of electrons and holes is basically the following. The forces which act on an electron are the ones from the other ions and electrons in the crystal and the externally applied ones. For simplicity, we condensed the first of these forces to yield the periodic potential of (8.1) using the one-electron approximation. Then we chose to consider only the external forces acting on the electrons and the price we have to pay is the fact that the electrons and holes react with an effective mass. This mass is fortunately, in some regions at least, constant, allowing the effective-mass approximation mentioned above, but it can change as a function of $\boldsymbol{k}$ and become negative or infinite, as can be seen in Fig. 8.7.

It is important to note that an increasing curvature of a band is necessarily connected with an increasing width of the band. Therefore we find the qualitative relation:

$$\begin{aligned} &\text{electron easy to move and accelerate} \longleftrightarrow \text{low effective mass} \\ &\longleftrightarrow \text{large curvature of the band} \longleftrightarrow \text{large band width} \\ &\longleftrightarrow \text{strong coupling between adjacent atoms and vice versa}. \end{aligned} \tag{8.15}$$

Thus we have here another example of the more general discussion given in the introduction to Sect. 5.4.

Furthermore it can be stated as a trend that a large band width leaves less space on the energy scale for gaps. Therefore we can also state as a rule of thumb, that narrow gap semiconductors tend to have small effective masses and wide gap semiconductors tend to have carriers with larger effective masses.

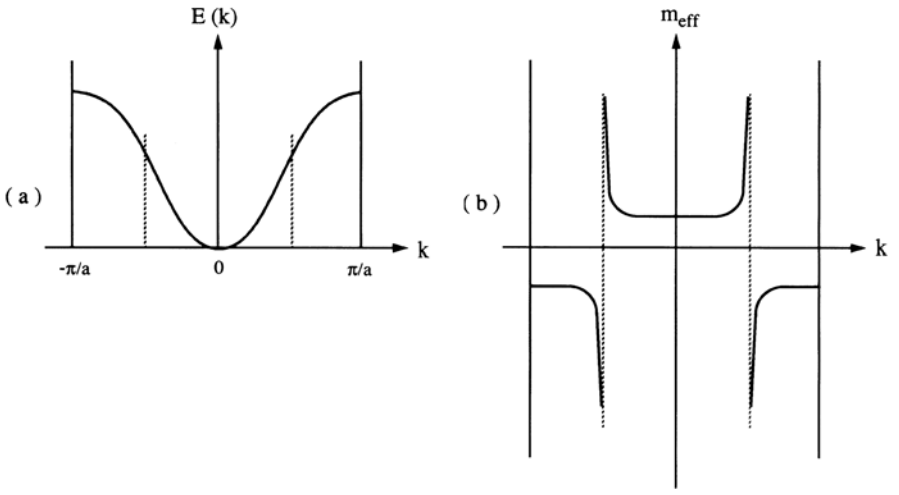


Fig. 8.7. Schematic dispersion (**a**) of a conduction band and the resulting effective mass (**b**). After [95I1] of Chap. 1

In the rest of this book we shall use the effective-mass approximation if not stated otherwise, including the polaron correction discussed in the next section.

We will in future omit the prefix "crystal" when talking about electrons and holes in semiconductors and we indicate the effective masses by indices e and h:

$$\begin{aligned} m_{\mathrm{e}} &= \text{effective mass of electron}\,, \\ m_{\mathrm{h}} &= \text{effective mass of hole}\,, \end{aligned} \tag{8.16}$$

sometimes with an additional index to distinguish different bands, but we always remember, that these charge carriers are quasi-particles, characterized by their band index, energy and (quasi-!) momentum.

It should be mentioned that the Fermi energy is often situated in metals with partly-filled bands in regions where the band structure $E(\boldsymbol{k})$ is no longer parabolic. On the other hand, the electrons close to the Fermi energy are in metals the relevant ones for many physical properties like the electrical conductivity, the specific heat or the paramagnetic behavior of the electron gas. Therefore the use of the effective mass concept as given with (8.14) is less frequently used in metal physics and replaced by another concept that is based on the fact that the first derivative $\mathrm{grad}_{\boldsymbol{k}} E(\boldsymbol{k})$ vanishes per definition close to the extrema of the conduction and valence bands of semiconductors but usually not around E_{F} in a metal.

One can then define a "momentum effective mass" m_i in contrast to the "force effective mass" m_{eff} in (8.13) via

$$\hbar \boldsymbol{k} = \boldsymbol{p} = m \boldsymbol{v} \tag{8.17}$$

resulting in $\boldsymbol{v}_{\mathrm{g}} = \frac{1}{\hbar} \mathrm{grad}_{\boldsymbol{k}} E(\boldsymbol{k})$ in the vicinity of $E_{\mathrm{F}}(\boldsymbol{k}_{\mathrm{F}})$ in

$$\frac{1}{m_i} = \frac{1}{\hbar k_{\mathrm{F}}} \mathrm{grad}_{\boldsymbol{k}} E\left(\boldsymbol{k} = \boldsymbol{k}_{\mathrm{F}}\right) . \tag{8.18}$$

8.6 The Polaron Concept and Other Electron–Phonon Interaction Processes

Before proceeding to some basic concepts of bandstructure calculations and to the band structures of real semiconductors, we shall discuss here various aspects of electron–phonon interaction (see also [69R1,76C1,77P1,79T1,80M1, 80P1]).

If we introduce an electron or a hole into a semiconductor which has at least partially ionic binding, the additional charge carrier will polarize the lattice provided we relax the assumption that the ions are "fixed" at their lattice sites. An electron will attract atoms with a positive charge and repel those with a negative one (Fig. 8.8). For holes, the situation is just the opposite.

We can describe this lattice distortion as a superposition of preferentially longitudinal optical phonons, i.e., a free carrier is accompanied by a "phonon cloud". The entity of charge carrier plus phonon cloud is called a polaron. In semiconductors, the radius of the phonon cloud is larger than the lattice constant, resulting in a so-called "large polaron" in contrast to the "small polaron" which occurs in ionic insulators and may lead to self-localization within a unit cell. The effective mass $m_{\mathrm{e,h}}$ of a polaron is greater than that of an electron in a rigid lattice $m_{\mathrm{e,h}}^{\mathrm{r.l.}}$ See e.g. [69R1, 70M1, 77P1, 79T1, 80M1, 80P1, 88C1] or [75Z1, 76A1, 81M1, 82L1, 99B1] of Chap. 1.

$$m_{\mathrm{e,h}} \simeq m_{\mathrm{e,h}}^{\mathrm{r.l.}} \left(1 + \frac{\alpha_{\mathrm{e,h}}}{6}\right) , \tag{8.19a}$$

where the index r.l. stands for "rigid lattice" and α is a dimensionless quantity which describes the Fröhlich coupling of carriers to the LO phonons. One finds

$$\alpha_{\mathrm{e,h}} = \frac{\mathrm{e}^2}{8\pi\varepsilon_0\hbar\omega_{\mathrm{LO}}} \left(\frac{2m_{\mathrm{e,h}}^{\mathrm{r.l.}}\omega_{\mathrm{LO}}}{\hbar}\right)^{1/2} \left(\frac{1}{\varepsilon_{\mathrm{b}}} - \frac{1}{\varepsilon_{\mathrm{s}}}\right) . \tag{8.19b}$$

where ε_{s} and ε_{b} are the low-and high-frequency values of the dielectric function below and above the optical phonon resonances, respectively; see also Chaps. 4 and 12.

Obviously semiconductors without an ionic binding contribution have $\varepsilon_{\mathrm{b}} = \varepsilon_{\mathrm{s}}$ [see (4.26), (4.28)] and thus $\alpha = 0$. For all normal semiconductors one finds

$$\alpha \leq 1 . \tag{8.19c}$$

Additionally the lattice relaxation leads to a decrease of the width of the gap by an amount ΔE_{g} with contributions for electrons and holes

$$\Delta E_{\mathrm{g}}^{\mathrm{e,h}} = \alpha_{\mathrm{e,h}}\hbar\omega_{\mathrm{LO}} . \tag{8.19d}$$

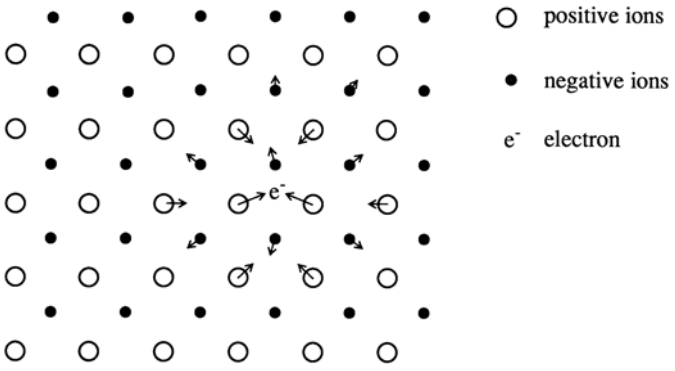


Fig. 8.8. The lattice distortion around a carrier in a (partly) ionic semiconductor illustrating the polaron concept

The radius of the phonon cloud in the polaron r_p is finally given by

$$r_\mathrm{p}^{\mathrm{e,h}} = \left(\hbar/2m_{\mathrm{e,h}}^{\mathrm{r.l.}}\omega_\mathrm{LO}\right)^{1/2} . \tag{8.19e}$$

Basically it is very difficult to "fix the atoms at their lattice sites". Therefore all common experimental techniques to determine the effective masses of electrons and holes, e.g., by cyclotron resonance, or the value of the gap, e.g., by optical spectroscopy, will give polaron values. We therefore continue to use the effective mass approximation and simply bear in mind that all values given for $m_\mathrm{e,h}$ or E_g are actually polaron values.

Apart from the renormalization of mass and energy there are other consequences of carrier–phonon interaction, the most important being scattering phenomena. An electron or hole can be scattered inelastically, e.g., by emitting or absorbing a phonon under energy and momentum conservation.

$$E_\mathrm{e}^\mathrm{i} = E_\mathrm{e}^\mathrm{f} \pm \hbar\omega_\mathrm{Phonon} , \tag{8.20a}$$

$$\boldsymbol{k}_\mathrm{e}^\mathrm{i} = \boldsymbol{k}_\mathrm{e}^\mathrm{f} \pm \boldsymbol{k}_\mathrm{Phonon}(+\boldsymbol{G}) . \tag{8.20b}$$

The interaction Hamiltonian which describes such processes can have various origins as shown e.g. in [73H1, 82L1, 96Y1] of Chap. 1.

- Optical phonons are often accompanied by an electric field (Sect. 7.4, 7.5 and 11.1). The interaction of carriers with the electric field of preferentially longitudinal optical phonons is known as the Fröhlich interaction.
- Since the width of the gap depends on the lattice constant and on the arrangement of the atoms in the basis, a change of these quantities will influence the bandstructure. On the other hand, a phonon can be considered as a periodic deformation of the arrangement of atoms, and the carriers "feel" the resulting modulation of the bands. The resulting interaction between carriers and phonons is called deformation-potential scattering. The deformation potential scattering occurs for both acoustic and optical phonons. The deformation potential Ξ connects usually the relative displacement $\Delta a/a$ of an atom with a = lattice constant with the energetic shift of a band ΔE resulting schematically in

$$\Delta E = \Xi\frac{\Delta a}{a} \tag{8.21}$$

  Typical values of deformation potentials are in the range of 5 to 10 eV. Note that (8.21) is valid only for $\Delta a/a <<< 1$.
- Finally, it is known that many non-centrosymmetric crystals show the piezo-electric effect, i.e., the appearance of an electric field as a consequence of strain, i.e., of lattice distortion. Again we can consider an acoustic (or optical) phonon as a periodic modulation of the lattice parameters, which produces, via the piezo-electric effect, a varying electric field which interacts with the electrons and holes. This effect is the so-called piezo (acoustic) coupling.

More details about the polaron concept and on carrier-phonon coupling can be found in the references given above.

In the so-called semimagnetic semiconductors, usually II–VI compounds in which the cations are partly replaced by Mn or Fe ions, one finds so-called magnetic polarons, i.e., a spin alignment of the paramagnetic ions in the vicinity of a carrier. For details see e.g. references given in Chap. 16

8.7 Some Basic Approaches to Band Structure Calculations

In this section we will take a closer look at band structure calculations. Since this topic is already treated in many (text)books and review articles, it does not make too much sense to simply write it down here again. Instead we want to present the basic concepts and ideas but we are not aiming towards a stringent mathematical treatment of the topic. For this aspect we refer the reader rather to the literature like the textbooks on solid state and semiconductor physics of Chap. 1 and references given therein.

We start with the concept of nearly free electrons, proceed to the approach starting from atomic orbitals and finish this section with the so-called $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory, which is very useful to calculate band structures in the vicinity of the extrema.

As already stated, the dispersion relation of a free electron in vacuum is a parabola in the nonrelativistic regime as shown in Fig. 8.9a or 8.1a, i.e.,

$$E(\boldsymbol{k}) = \frac{\hbar^2 \boldsymbol{k}^2}{2m_0} \tag{8.22}$$

We start, for simplicity, with a one-dimensional system and represent the relation (8.22) for this case in Fig. 8.9a.

Next we introduce a very weak periodic potential as shown schematically in Fig. 8.10 with a periodicity a. Weak means, in this context, that the dispersion relation of Fig. 8.9a is only marginally influenced by this potential. Nevertheless, we have introduced a periodicity by this potential in real space, which evokes the (Ewald–)Bloch theorem. As a consequence, the dispersion relation can be repeated in $\boldsymbol{k}$-space by shifting it by integer multiples of $2\pi/a$ resulting in the extended zone scheme of Fig. 8.9b (dashed lines). Compare with Fig. 8.1. In the limit that the amplitude of the periodic potential goes to zero, this is all. If the potential is small but finite, a further modification occurs. Small gaps open at the borders of the Brillouin zones as shown by the solid lines. There are various possibilities to explain the appearance of these gaps and the fact that two different energy eigenvalues exist for the same $\boldsymbol{k}$-vector at the Brillouin zone boundary.

In a first argumentation this splitting can be considered as an example of the non-crossing rule already discussed in Sect. 5.1. The coupling between

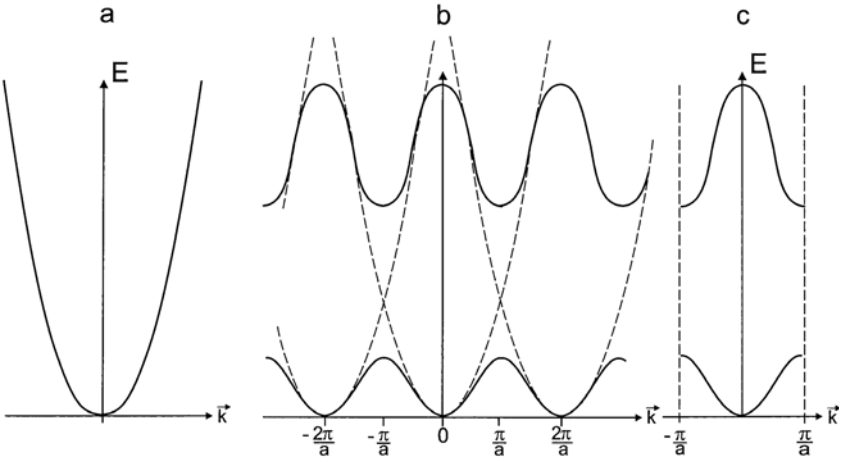


Fig. 8.9. The (one-dimensional) dispersion relation of a free electron in vacuum (**a**) in a weak periodic potential (**b**) and the resulting reduced zone scheme (**c**)

the degenerate states at $k = n\pi/a$ is mediated by the periodic potential. Indeed it can be shown that the width of the band gap is determined by the Fourier component of the periodic potential (see, e.g., (7.10)) with the reciprocal lattice vector $\boldsymbol{G}$, which couples the bands that intersect for vanishing potential.

Another set of arguments works as follows. Assume a plane electron wave is impinging on the periodic potential. Then the potential will produce scattered waves. For an arbitrary $\boldsymbol{k}$-vector, i.e., wavelength and (in dimensions higher than one) direction of incidence, the scattered waves will have various phase shifts relative to each other and will essentially cancel. However, there are special wave vectors, namely just the ones at the borders of the Brillouin zones (in our present one-dimension model $k = \pm n\pi/a$) where all scattered waves interfere constructively to form a backscattered wave with large amplitude, i.e., an incident wave with, e.g., $k_i = \pi/a$ creates a backscattered wave with $k_\mathrm{b} = -\pi/a$ (and $k_i - k_\mathrm{b} = 2\pi/a = G$). Since the backscattered wave with k_b scatters again off the periodic potential contributing to the wave with k_i we end up under stationary conditions with two counterpropagating waves of equal amplitudes and $k_i = -k_\mathrm{b}$. As is well known, the superposition of two waves with $\boldsymbol{k}$ and $-\boldsymbol{k}$ and equal amplitudes produces a standing wave. This fact has in our context two consequences:

- A wave packet built up from a standing wave and its surroundings in $\boldsymbol{k}$-space has zero group velocity. Consequently the slope of the dispersion relation must be zero, normal to the border of the Brillouin zone.
- The standing wave can be the sine- or the cosine-type with the same wavelength and $\boldsymbol{k}$-vector as shown in Fig. 8.10, where the probabilities are plotted. These two waves have their maxima at the positions of the maxima

or of the minima of the periodic potential, respectively. Consequently they see different average potential energies but have equal kinetic energy. This fact results in the energy splitting, i.e., in the appearance of the gaps at the borders of the Brillouin zones.

We proceed now, still in the context of nearly free electrons, to two- and three-dimensional potentials.

In Fig. 8.11a we show the parabolic dispersion relation of (8.22) over a two dimensional $\boldsymbol{k}$-space assuming a hexagonal, two-dimensional lattice but vanishing potential. The differently hatched areas in the k_x, k_y plain correspond to the different Brillouin zones. Figure 8.11a thus corresponds to the extended zone scheme. Figure 8.11b shows the resulting reduced zone scheme, which is obtained by shifting the parts outside the first Brillouin zone by suitable vectors $\boldsymbol{G}$ of the reciprocal lattice into the first zone. The various energy surfaces or bands should touch, e.g., at the border of the first Brillouin zone for vanishing potentials like in Fig. 8.9b for the one-dimensional case, but have been displaced in Fig. 8.11b and c for clarity along the energy axis to allow an "insight" into the first Brillouin zone.

In Fig. 8.11c a weak potential has been switched on. As we expect from our knowledge from the one dimensional case the overall appearance of the $E(\boldsymbol{k})$ relation does not change very much but kinks are smoothed out, energy gaps open, e.g., at the boarder or the center of the first Brillouin zone and the slope of the dispersion relation vanishes at the zone boundary in the direction normal to it. Figure 8.11d finally gives the dispersion relation along various directions in the two-dimensional $\boldsymbol{k}$-space for vanishing and for weak potentials.

In Fig. 8.12a and b we finally go through the same procedure as in Fig. 8.11d, but now for a real three-dimensional crystal structure, namely that for diamond (point group O_h).

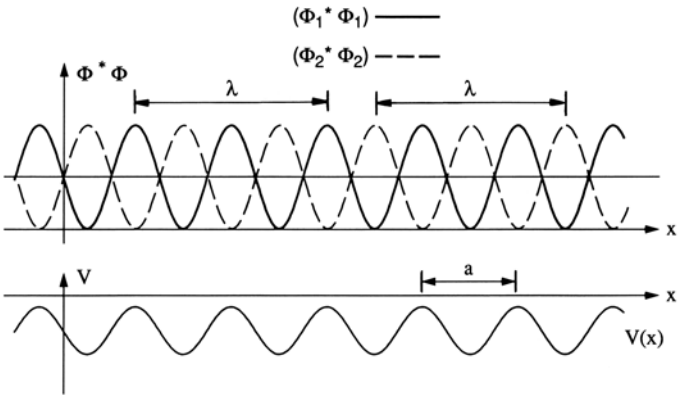


Fig. 8.10. Schematic drawing of two Bloch waves with equal $\boldsymbol{k}$-vector but different potential energies at the boundary of the first Brillouin zone over a periodic potential $V(x)$

The case of the empty lattice in Fig. 8.12a already shows a rather complex band structure and we understand that the complexity of a three-dimensional band structure comes to a large extent from the back-folding of the parabolic dispersion into the first Brillouin–zone.

Actually there is a possibility to test various band structure calculations, the so-called empty lattice datetest. One runs any type of band structure calculation but with vanishing potential. Since the solution is known exactly, e.g., for the diamond structure in Fig. 8.12a the band structure calculation must give this result. If it fails, something might be wrong, at least for the calculation of the bands arising from the outer, weakly bound atomic orbitals.

In Fig. 8.12b a realistic band structure calculation is shown for Si which crystallizes in the diamond structure, as already mentioned in Sect. 8.3, but still neglecting spin. The dispersion relation is given for the direction from the Γ-point towards the X and L points as in Fig. 7.4b, but also from Γ along the Σ direction to the U, K points and from there along the surface of the first Brillouin zone to the X point. The indirect band-gap in Si occurs between the

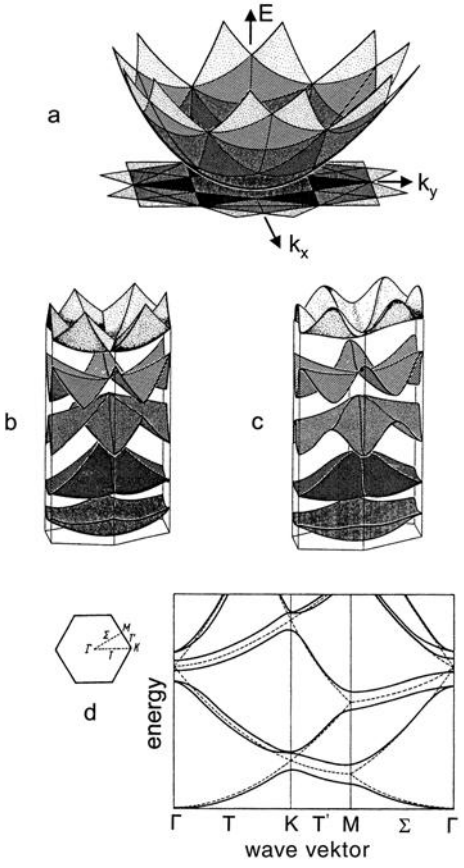


Fig. 8.11. The parabolic dispersion relation of free electrons over a two-dimensional $\boldsymbol{k}$-space, assuming a hexagonal lattice with vanishing potential (the so-called empty lattice) (**a**) the resulting reduced zone scheme for vanishing (**b**) and weak but finite potential (**c**) and the $E(k)$ relation along various lines of high symmetry for both cases (**d**). In **b** and **c** the bands are displaced vertically for clarity ([81M1] of Chap. 1)

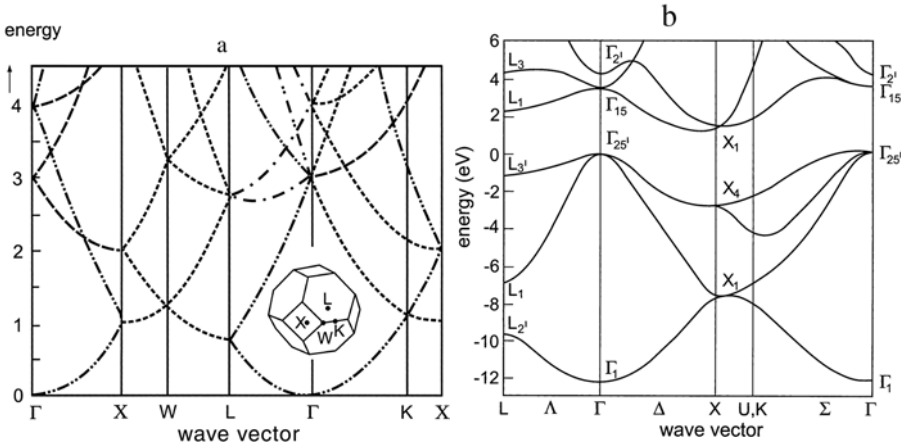


Fig. 8.12. Band structure for the three dimensional diamond structure. For vanishing potential (**a**) and a realistic band structure calculation for Si (**b**). The numbers at Γ give the irreducible representations (see Sect. 26.4 to 6), still neglecting spin: According to [76A1, 81M1, 96Y1] of Chap. 1

maximum of the valence band at the Γ point, set here as energy zero and the minimum of the conduction band close to the X point.

While we stressed, in the nearly free electron approach, the plane wave factor of the Bloch–wave and assumed that the lattice periodic part is essentially a constant, we now start from the opposite side, by placing one or more atomic orbitals at every atom site as shown, e.g., in Fig. 8.3 for simple 1s like atomic orbitals.

Now we translate this concept into an equation assuming, for simplicity, that we have a three-dimensional simple-cubic lattice with one atom per unit cell and a nondegenerate normalized atomic orbital $\varphi_i(\boldsymbol{r}-\boldsymbol{R}_i)$ at every lattice site $\boldsymbol{R}_i$.

In the linear combination of atomic orbitals (LCAO), or tight binding approximation, the total wave function $\psi(\boldsymbol{r})$ is then given by

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_i C_{\boldsymbol{k},i}\varphi_i(\boldsymbol{r}-\boldsymbol{R}_i) \ . \tag{8.23a}$$

The fact that this form has to be of the Bloch–type determines the $C_{\boldsymbol{k},i}$ to be

$$C_{\boldsymbol{k},i} = N^{-\frac{1}{2}}\,\mathrm{e}^{i\boldsymbol{k}\boldsymbol{R}_i} \tag{8.23b}$$

where $N^{\frac{-1}{2}}$ is the normalization factor for a crystal consisting of N atoms. Inserting (8.23b) in (8.23a) results in the LCAO wave function

$$|\boldsymbol{k}\rangle = \psi_{\boldsymbol{k}}(\boldsymbol{r}) = N^{-\frac{1}{2}}\sum_i \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{R}_i}\varphi(\boldsymbol{r}-\boldsymbol{R}_i) \ . \tag{8.24a}$$

The energy belonging to these wave functions can be calculated in first order with the help of the Hamiltonian H which contains, apart from the kinetic energy terms, the sum over the atomic potentials via

$$E(\boldsymbol{k}) = \langle \boldsymbol{k}\,|H|\,\boldsymbol{k}\rangle = N^{-1}\sum_i \sum_j \mathrm{e}^{\mathrm{i}\boldsymbol{k}(\boldsymbol{R}_i-\boldsymbol{R}_j)} \times \int \varphi\cdot(\boldsymbol{r}-\boldsymbol{R}_i) H\varphi(\boldsymbol{r}-\boldsymbol{R}_i)\,\mathrm{d}^3 r\,. \tag{8.24b}$$

Since only relative distances enter, we simplify (8.23b) by introducing $\boldsymbol{R}_m = \boldsymbol{R}_i - \boldsymbol{R}_j$ and obtain

$$E(\boldsymbol{k}) = \langle \boldsymbol{k}\,|H|\,\boldsymbol{k}\rangle = \sum_m \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{R}_m} \int \varphi\cdot(\boldsymbol{r}-\boldsymbol{R}_m) H\varphi(\boldsymbol{r})\,\mathrm{d}^3 r\,. \tag{8.24c}$$

The next simplification is to assume realistically that in the sum only the terms with $\boldsymbol{R}_m = 0$ and contribution for nearest neighbors $\boldsymbol{R}_n$ in the overlap integral give considerable contributions. With the abbreviations

$$\int \varphi^*(\boldsymbol{r}) H\varphi(\boldsymbol{r})\,\mathrm{d}^3 r = -A \tag{8.25}$$

and

$$\int \varphi^*(\boldsymbol{r}-R_n) H\varphi(\boldsymbol{r})\,\mathrm{d}^3 r = -B \tag{8.26}$$

we obtain

$$E(\boldsymbol{k}) = \langle \boldsymbol{k}\,|H|\,\boldsymbol{k}\rangle = -A - B\sum_n \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{R}_n} \tag{8.27}$$

where the sum now runs only over the nearest neighbors.

The term $-A$ is close to but generally slightly lower in energy than the eigenenergy of the parent atomic orbital $\varphi(\boldsymbol{r})$ since the Hamiltonian contains not only the potential of a single atom but the sum over them as stated above.

For the simple-cubic lattice, which we consider, $\boldsymbol{R}_n$ adapts the following six values

$$\boldsymbol{R}_n = (\pm a, 0, 0); (0, \pm a, 0); (0, 0, \pm a)\,. \tag{8.28}$$

This finally results in a simple band structure of width $12B$ centered around $-A$:

$$E(\boldsymbol{k}) = -A - 2B(\cos k_x a + \cos k_y a + \cos k_2 a)\,. \tag{8.29}$$

We see from (8.29) that the overlap between atomic orbitals leads to the formation of bands out of sharp atomic levels while the nearly free electron methods explained the appearance of gaps in the simple parabolic dispersion relation (see again Fig. 8.1).

Physically the formation of bands by the overlap described by B in (8.29) is nothing but the splitting of the eigenenergies of coupled harmonic oscillators as already mentioned in Sect. 8.1.

The band width increases with increasing overlap integral B. Consequently, the width of the forbidden gaps tends to decrease, as does the effective mass, and the lattice periodic part of the Bloch–function $u_k(\boldsymbol{r})$ is smoothed out. See also Sect. 8.5.

An early example of the LCAO method is the Wigner–Seitz approach [47L1] which can already show, for the simple alkali-metals, that the nearly free electron and the LCAO approaches give consistent results.

To finish this subsection on LCAO we want to discuss two further examples. First we consider qualitatively a hypothetic situation in which a valence-band is formed from atomic $2p$ orbitals in a simple-cubic lattice [97U1].

In Fig. 8.13a we show a layer of p_x orbitals. In the x direction the overlap integral B (see above) is small since adjacent wave function lobes of different atoms have opposite sign, i.e., there is a node line. Small overlap means a narrow band in the k_x direction and a heavy hole mass. The same orbitals have a large overlap integral in the y direction since adjacent lobes of neighboring atoms have the same sign. Consequently, B is larger in the y direction the band is broader in the k_y direction and the effective mass is small as shown by the solid line in Fig. 8.13b. For the p_y orbitals we obtain just the opposite behavior as shown by the dashed line in Fig. 8.13b. Including the p_z orbitals gives an analogous behaviour in all three directions of $\boldsymbol{k}$ space. We can reinterpret this finding in the following way. There is one valence-band in all three directions, which has low curvature, i.e., a heavy effective mass and another one, which is degenerate at the Γ-point, with large curvature and low effective mass. Consequently the holes in these bands are called heavy holes (hh) and light holes (lh). We should remember that we were treating a hypothetical material, but we shall encounter in Sect. 8.8 a very similar but slightly more complex situation for the valence bands of many semiconductors crystallizing in diamond- or blende-type structures.

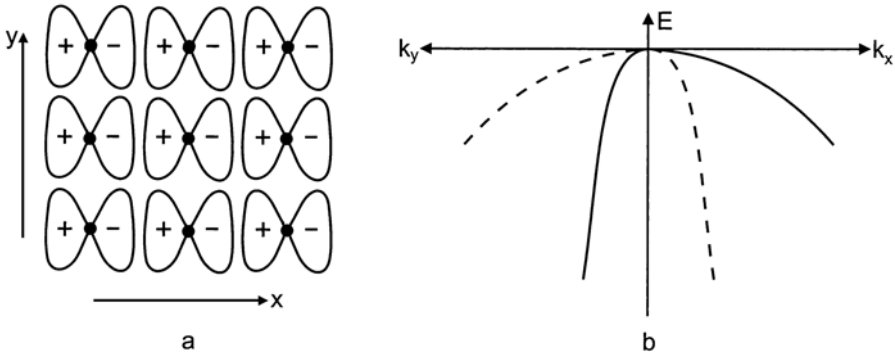


Fig. 8.13. A hypothetic layer of p_x orbitals (**a**) and the resulting $E(\boldsymbol{k})$ relation in the direction k_x and k_y (*solid line*) and the analogon for p_y orbitals (*dashed line*) (**b**). The *heavy dots* in **a** give the positions of the atoms

As a last example we discuss with Fig. 8.14 the tight binding approach for Si (or diamond or Ge) as a function of the lattice constant a. Si has a filled Ne configuration plus $3s^2 3p^2$, i.e., four electrons in the outer $n_B = 3$ shell, which are responsible for the chemical binding, while the lower electron states ($1s^2$, $2s^2$, $2p^6$) form deeper valence bands, which are presently of no interest.

For a large lattice constant the overlap integral vanishes and one starts with the atomic $3s$ and $3p$ levels. With decreasing distance the levels shift to lower energy and broaden since both A and B in (8.29) depend on the lattice constant. For further decreasing distance the two bands start to overlap. This is the distance between the atoms at which the sp^3 hybrid orbitals start to form. For even smaller values the bands split again. In contrast to the two fold degenerate $3s$ and six-fold degenerate $3p$ levels, two four-fold degenerate bands form the binding and the antibinding states of the sp^3 hybrids. Since there are two Si-atoms in the primitive unit cell, the valence band originating from the HOMO (see Sect. 8.2) accommodates eight electrons and the conduction band (LUMO) also has place for eight electrons per unit cell. The binding orbitals of the sp^3 hybrid point towards the corners of a tetrahedron giving rise to the tetrahedral coordination of many semiconductors (see Sect. 7.2).

The minimum of the total free energy gives the equilibrium distance r_0, which is indicated in Fig. 8.14. For even smaller values of a the bands shift to higher energies, again due to the action of Pauli's principle, and the increasing kinetic energy of the wave functions.

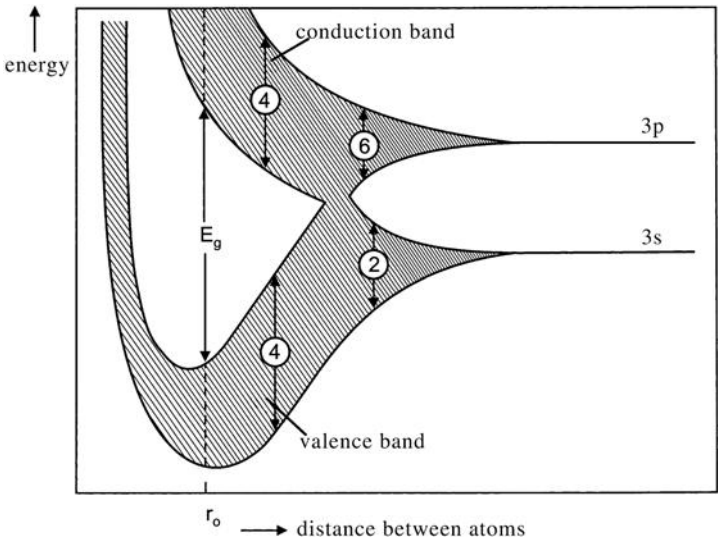


Fig. 8.14. The bands originating from the $3s$ and $3p$ levels of Si as a function of the lattice constant a or the distance between atoms ([95I1] of Chap. 1). The numbers in the circles give the number of electrons per atom in the respective state or band.

To conclude this section we give a few introductory remarks to the widely used $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory, which is very valuable to calculate the band structure in the vicinity of band extrema but by extrapolation allows one to obtain an idea about the band structure over the whole Brillouin zone. For more details see again the textbooks on solid state physics and on semiconductor physics cited in Chap. 1 and in this chapter.

We assume in the following that the band extrema are situated at $\boldsymbol{k}_0$ and introduce the further simplification that $\boldsymbol{k}_0 = 0$. If this is not the case, the following considerations are valid for $\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{k}_0$.

We start again with the one-electron approximation and the periodic potential $V(\boldsymbol{r}) = V(\boldsymbol{r} + \boldsymbol{R})$. The Hamiltonian reads

$$H = \frac{\boldsymbol{p}^2}{2m_0} + V(\boldsymbol{r})\,. \tag{8.30}$$

We insert the Bloch-wave ansatz of Sect. 8.1 into (8.30) and obtain (8.31a) where n is the band index, m_0 the free electron mass and $\boldsymbol{p} = \frac{\hbar}{i}\nabla$:

$$\left(\frac{\boldsymbol{p}^2}{2m_0} + \frac{\hbar^2 \boldsymbol{k} \cdot \boldsymbol{p}}{m_0} + \frac{\hbar^2 \boldsymbol{k}^2}{2m_0} + V(\boldsymbol{r})\right) u_{n\boldsymbol{k}}(\boldsymbol{r}) = E_n(\boldsymbol{k})\, u_{n\boldsymbol{k}}(\boldsymbol{r})\,. \tag{8.31a}$$

For the case $\boldsymbol{k} = \boldsymbol{k}_0 = 0$ (8.31a) simplifies to

$$\left(\frac{\boldsymbol{p}^2}{2m_0} + V(\boldsymbol{r})\right) u_{n0}(\boldsymbol{r}) = E_n(\boldsymbol{k} = 0)\, u_{n0}(\boldsymbol{r}) = E_{n,0} u_{n,0}(\boldsymbol{r})\,. \tag{8.31b}$$

We now assume, that the solution of (8.31b) is known. Then we can use the terms $\hbar^2 \boldsymbol{k}\boldsymbol{p}/m_0$ (this is the reason for the name of this method) and $\hbar^2 \boldsymbol{k}^2/2m_0$ in (8.31a) as perturbations of the first and second order in $\boldsymbol{k}$.

For the simple case of a nondegenerate band (or a two-fold spin degenerate one) we obtain for the lattice periodic part of the eigenfunctions $u_{n\boldsymbol{k}}(\boldsymbol{r})$

$$u_{n\boldsymbol{k}}(\boldsymbol{r}) = u_{n0}(\boldsymbol{r}) + \frac{\hbar}{m_0} \sum_{n' \neq n} \frac{\langle u_{no} | \boldsymbol{k} \cdot \boldsymbol{p} | u_{n'0} \rangle}{E_{n0} - E_{n'0}} u_{n'0}(\boldsymbol{r})\,. \tag{8.32a}$$

The above-mentioned situation of nondegeneracy frequently holds for the lowest conduction band of semiconductors.

The dispersion relation is then given by

$$E_n(\boldsymbol{k}) = E_{n,0} + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{n' \neq n} \frac{|\langle u_{n0} | \boldsymbol{k} \cdot \boldsymbol{p} | u_{n'0} \rangle|^2}{E_{n0} - E_{n'0}} \tag{8.32b}$$

From (8.32b) we can easily deduce the effective mass of e.g. the conduction band n

$$\frac{1}{m_\mathrm{e}} = \frac{1}{m_0} + \frac{2}{m_0^2 \boldsymbol{k}^2} \sum_{n' \neq n} \frac{|\langle u_{no} | \boldsymbol{k} \cdot \boldsymbol{p} | u_{n'0} \rangle|^2}{E_{n0} - E_{n'0}} \tag{8.32c}$$

The coupling between various bands, which influences m_e, depends essentially on the momentum matrix element, which in turn is closely related to the dipole-matrix element (see Sect. 3.2.2) and can at least in principle be deduced from optical absorption spectra.

The knowledge, which of the matrix elements vanish and which are non-zero, can be obtained from group theory (see Chap. 26).

As already stated various times above, the effective electron mass decreases with decreasing width of the gap $E_{n0} - E_{n'0}$ if the dominant term in (8.32c) comes from the coupling between conduction- and valence band.

Since the $\boldsymbol{k} \cdot \boldsymbol{p}$ theory starts for all bands with the free electron mass (see (8.31a)) the appearance of effective masses can be considered again as a consequence of level repulsion, e.g., between conduction and valence band in the way indicated schematically by the arrows in Fig. 8.15. The coupling between the bands arises in this model from the $\boldsymbol{k} \cdot \boldsymbol{p}$ term (also see the Kane-model [57K1]).

The treatment of degenerate bands, especially of the upper valence bands in $\boldsymbol{k} \cdot \boldsymbol{p}$ theory is more complicated, and beyond the scope of this chapter. We refer the reader to the above-mentioned literature for details and formula but mention that we shall meet results in the next section on band structures of real semiconductors.

8.8 Bandstructures of Real Semiconductors

In this section we present band structures of real semiconductors.

As already mentioned the upper valence bands frequently arise, in the case of ionic binding, from the highest occupied atomic p-levels of the anions with

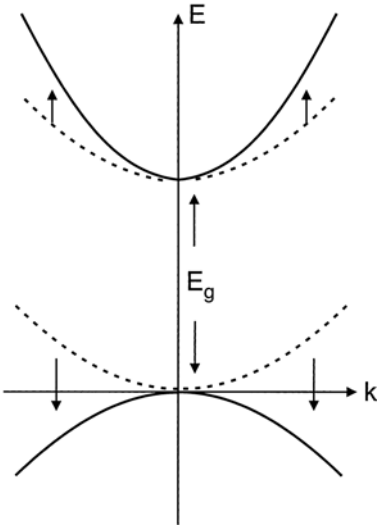


Fig. 8.15. Conduction and valence band with free electron mass m_0 (*dashed lines*) and the modifications (*arrows*) resulting from level repulsion (*solid lines*)

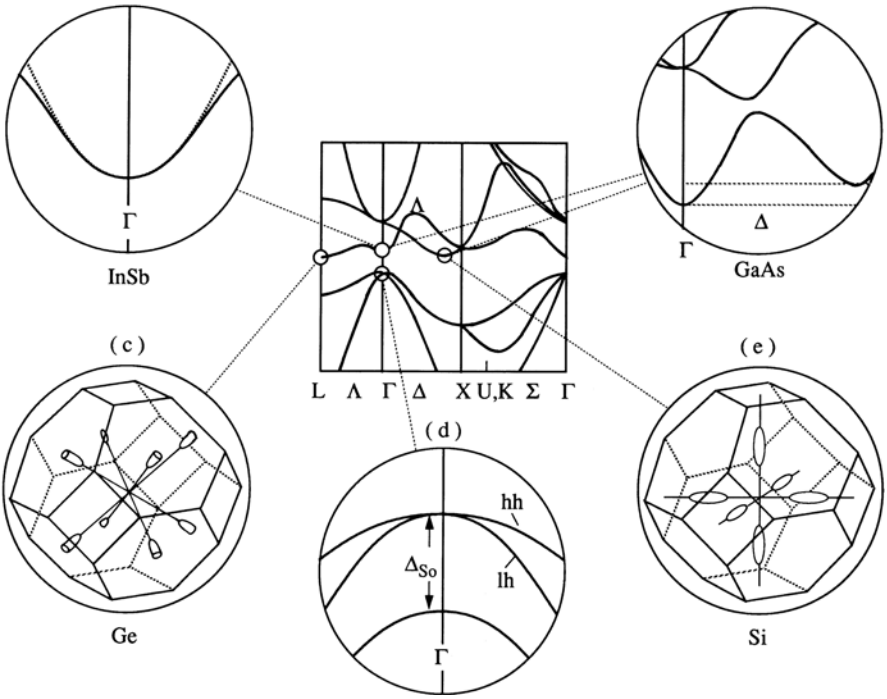


Fig. 8.16. An overview of the band structures of some cubic semiconductors [70M1]

a more-or-less pronounced admixture of d-levels, or from the bonding states of the sp^3 hybrid orbitals for covalent binding. The lowest conduction bands come from the lowest empty s-levels of the cations or the antibonding sp^3 hybrid, respectively. In Fig. 8.16 we give the essentials of the bandstructure for cubic semiconductors like diamond, Si or Ge (point group O_h) or the III–V, II–VI and I–VII compounds crystallizing in zincblende (T_d) structure.

The valence band has its maximum at the Γ point, i.e., at $\boldsymbol{k} = 0$. It is six-fold degenerate including spin corresponding to the parent p orbitals. In Sect. 8.7 we stated that eight levels form the bonding sp^3 hybrid state since there are two atoms in a unit cell. Six of them, which are treated in the following, form the top of the valence band and the other two the rather low lying Γ_1 (with spin Γ_6) level seen, e.g., in Fig. 8.14b. This band is not included in the following. The upper valence band splits due to spin–orbit coupling at $\boldsymbol{k} = 0$ into a two-fold degenerate band (symmetry Γ_7^+ in O_h or Γ_7 in T_d) and a four-fold degenerate band (symmetry Γ_8^+ in O_h or Γ_8 in T_d). Usually the $\Gamma_8^{(+)}$ band is the upper and the $\Gamma_7^{(+)}$ the lower one. The spin–orbit splitting increases in atoms with increasing charge Z of the nucleus and this also applies in the semiconductor. In CuCl the ordering of the valence bands is inverted due to the influence of close-lying d levels.

The $\Gamma_8^{(+)}$ valence band splits for $\boldsymbol{k} \neq 0$ into two bands which have different curvature and are therefore known as heavy- and light-hole bands. All bands have cubic symmetry, which is lower than spherical symmetry. As a consequence the dispersion and thus the hole masses depend on the direction of $\boldsymbol{k}$. This phenomenon is known as band-warping. The valence bands are frequently treated in $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory. The theory has been developed, e.g., by Kane, Luttinger, Cardona, Pollack and many others (see [57K1, 87B1] or [96Y1] of Chap. 1) and is briefly outlined in Sect. 8.7.

Basically one has to start with the 16 bonding and antibonding states resulting from two atoms per primitive unit cell. Often one considers the lowest two-fold degenerate conduction band (symmetry Γ_6) separately, e.g., in the way described in Sect. 8.7 and is then left with eight valence band states. Since the lowest Γ_1 (or Γ_6) state is energetically rather far away, it has only minor influence and may be neglected in the treatment of the upper valence bands. The remaining six states in the Γ_8 and Γ_7 valence bands can be treated in a 6×6 matrix. If the spin-orbit coupling is large, it is, however, often possible to treat the Γ_8 band alone resulting in a 4×4 matrix, which can be solved analytically. The resulting band structure is frequently described by the so-called Luttinger parameters γ_1, γ_2 and and γ_3 according to

$$\begin{aligned} E_{1,2} =& E_0 + \frac{\hbar^2}{2m_0}\Big[\gamma_1 k^2 \pm 2\left[\gamma_2\left(k_x^4 + k_y^4 + k_z^4\right)\right. \\ &\left. +3\left(\gamma_3^2 - \gamma_2^2\right)\left(k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2\right)\right]^{1/2}\Big] \end{aligned} \tag{8.33a}$$

where γ_1^{-1} describes the average effective mass and γ_2 and γ_3 the splitting into heavy- and light-hole bands and the warping via the term $3\left(\gamma_3^2 - \gamma_2^2\right)$. If warping is negligible (i.e. $\gamma_2 = \gamma_3$) the light and heavy hole effective masses are given by

$$m_{\mathrm{hh}} = \frac{m_0}{\gamma_1 - 2\gamma_2}; \quad m_{\mathrm{lh}} = \frac{m_0}{\gamma_1 + 2\gamma_2} \tag{8.33b}$$

The conduction band has a minimum at the Γ point and other minima in the direction Δ close to the X points and at the L points.

If the minimum of the Γ point is the deepest one, the semiconductor is said to have a "direct gap" since transitions between the global maximum of the valence band and the global minimum of the conduction band are directly possible with photons, having $k_{\mathrm{photon}} \simeq 0$. In other cases, the semiconductor is called "indirect" since a momentum-conserving phonon is involved in the transitions between the band extrema.

Examples of indirect semiconductors are diamond, Si and Ge, some of the III–V compounds such as AlAs and GaP, and some of the I–VII compounds like AgBr. Direct gap semiconductors include some of the III–V compounds like GaAs or InP, the II–VI compounds ZnS, ZnSe, ZnTe and CdTe, and I–VII materials like CuCl, CuBr and CuI. See also Table 8.1.

The conduction band minimum at the Γ point is usually, to a very good approximation, isotropic and parabolic. Only some narrow-gap materials like InSb show significant nonparabolicities in the vicinity of $k = 0$ (Fig. 8.16).

The minima of the L points of Ge or in the Δ direction of Si are parabolic but highly anisotropic (Fig. 8.16). The dispersion relation around the minimum at $\boldsymbol{k}_0$ can consequently be expressed by

$$E(\boldsymbol{k}) = E_\mathrm{g} + \frac{\hbar^2}{2}\left(\frac{(k_x - k_{0x})^2}{m_\mathrm{l}} + \frac{(k_y - k_{0y})^2}{m_\mathrm{t}} + \frac{(k_z - k_{0z})^2}{m_\mathrm{t}}\right). \tag{8.34}$$

Here m_l is the effective mass for k-components in the direction from Γ to $\boldsymbol{k}_0$ and m_t that for the two directions perpendicular to it.

In GaAs the minimum at the Γ point is the deepest, but other minima at different points of the Brillouin zone are close in energy, e.g., in the Δ direction giving rise to the Gunn effect, which arises from the transfer of electrons under the influence of a strong electric field from the minimum at $k = 0$ with low effective mass to side minima with higher effective masses.

The direct gap semiconductors are also called "single-valley", and the indirect ones "multi-valley" semiconductors because they have several (6 for Si and $8 \cdot 1/2$ for Ge) equivalent conduction-band minima.

Semiconductors with hexagonal wurtzite structure (point group C_{6v}) are usually "direct". They are found preferentially among the II–VI compounds such as ZnO, ZnS, CdS and CdSe, but also among the III–V materials like GaN.

The upper six valence band states of the C_{6v} semiconductors are split by spin–orbit coupling and by the hexagonal crystal field into three subbands which are usually labelled from higher to lower energies as A, B and C bands with symmetries Γ_9, Γ_7 and Γ_7 (Fig. 8.17a). In ZnO, the symmetries of the two upper bands are most probably inverted as in the case of CuCl. Some references on this presently again ongoing discussion are e.g. [69R1, 82B1, 02W1] or [01G1] of Chap. 4. The effective masses of the valence bands are often strongly anisotropic in these compounds, $m_\perp$ usually being smaller than $m_\parallel$ where the indices refer to $\boldsymbol{k}$-vectors perpendicular and parallel to the polar crystallographic $\boldsymbol{c}$-axis:

$$m_\perp^\mathrm{vb} \ll m_\parallel^\mathrm{vb} \quad \text{for} \quad C_{6v}\,, \tag{8.35a}$$

$$m_\mathrm{DOS} = (m_\perp^2 m_\parallel)^{1/3}\,. \tag{8.35b}$$

If one restricts oneself to the four states in the A and B bands $\boldsymbol{k}\cdot\boldsymbol{p}$ theory predicts for negligible warping, i.e., $\gamma_2 = \gamma_3 = \gamma$ for the effective masses [92E1]

$$m_{\mathrm{A}\parallel} = \frac{m_0}{\gamma_1 - 2\gamma}; m_{\mathrm{A}\perp} = \frac{m_0}{\gamma_1 + \gamma} \tag{8.36a}$$

$$m_{\mathrm{B}\parallel} = \frac{m_0}{\gamma_1 + 2\gamma}; m_{\mathrm{B}\perp} = \frac{m_0}{\gamma_1 - \gamma}\,. \tag{8.36b}$$

The effective mass that enters in the calculation of the density of states m_{DOS} is given in (8.35b). A pecularity of the states of Γ_7 symmetry is that they can have a term linear in k in the dispersion relation for $\boldsymbol{k} \perp \boldsymbol{c}$ as shown in Fig. 8.17b. This term has some influence on the optical properties. In principle, it can also occur for the Γ_7 conduction band, but is much smaller there and usually neglected. The same is true for the hexagonal warping which is, in principle, possible for C_{6v} symmetry in the plane normal to $\boldsymbol{c}$.

As already mentioned, the crystal structures of zincblende and of wurtzite are rather similar and differ only in the arrangements of the next-nearest neighbors. The unit cell of the wurtzite structure is in one direction twice as long as the primitive unit cell of the zinkblende structure. As a consequence the first Brillouin zone is only half as long in one direction. The resulting folding back of the dispersion is shown schematically in Fig. 8.18 neglecting spin [69R1] for the two lowest conduction bands and the highest valence band. Obviously the number of bands doubles in this procedure, as does the number of states at the Γ point. This is in agreement with the fact that diamond and zincblende structures have two atoms in the basis of the primitive unit cell but wurtzite has four.

Due to the p- and s-type character of valence and conduction bands, for the more ionic bound semiconductors and of the binding and antibinding sp^3 states close to the fundamental gap, respectively, the band-to-band transition is dipole allowed – possibly with some additional selection rules for the hexagonal symmetry.

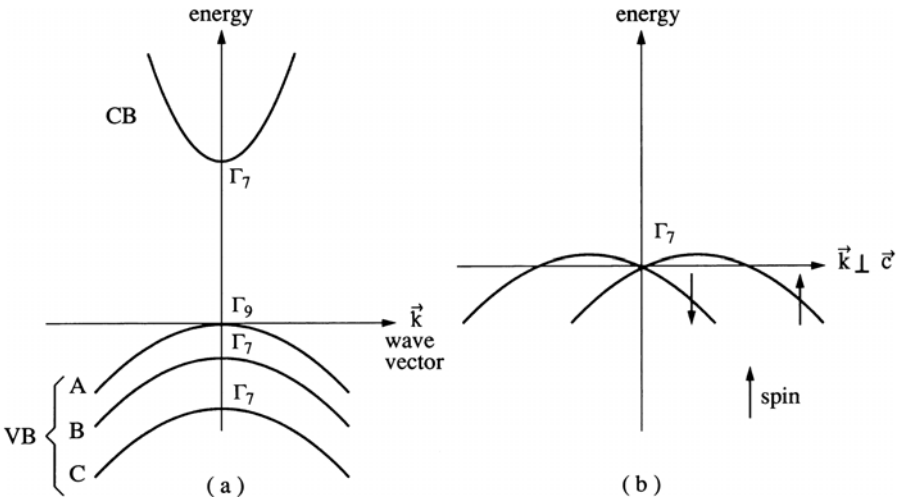


Fig. 8.17. Details of the band structure of hexagonal semiconductors around the Γ point. The splitting into three valence bands (**a**) and the influence of a term linear in k (**b**)

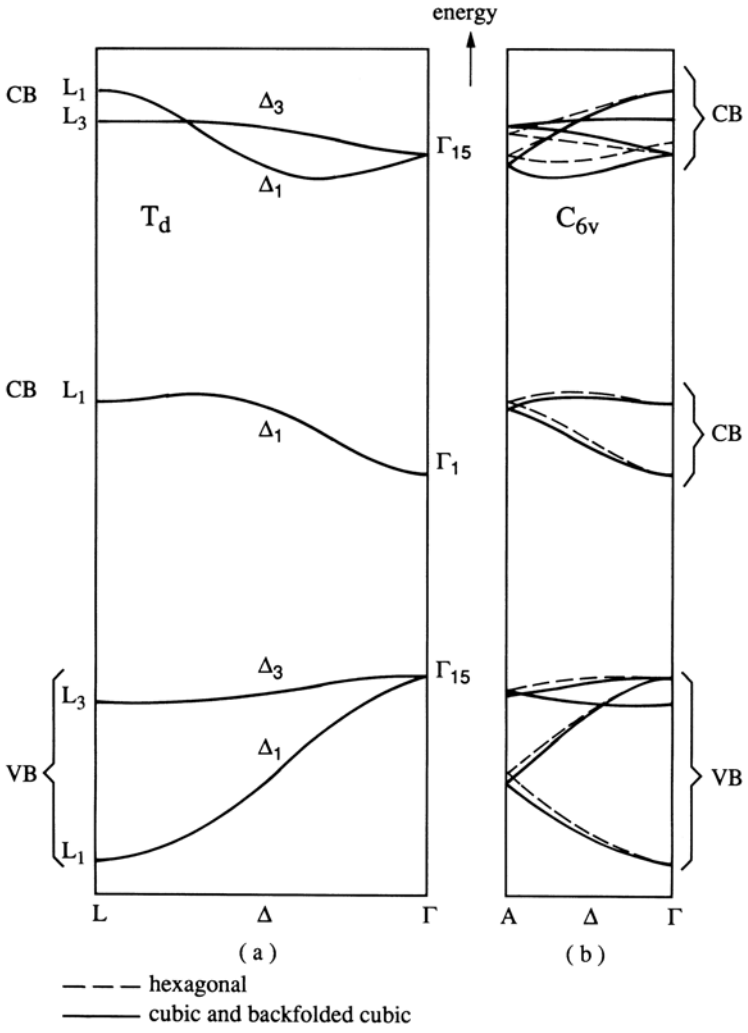


Fig. 8.18. The relation between the band structure of semiconductors with zinkblende and wurtzite type structures along the corresponding direction in $\boldsymbol{k}$-space, neglecting spin [69R1]

In Table 8.4 we summarize some band parameters of semiconductors. An exhaustive list is found in [82L1] of Chap. 1. See as additional information Table 8.2

It is beyond the scope of this book to review all types of semiconductors. Instead we give only some selected examples beyond those mentioned already above.

The lead salts (PbS, PbSe, PbTe) are narrow-gap semiconductors crystallizing in the NaCl structure. They have a direct gap. In contrast to the

Table 8.4. Effective masses of some selected semiconductors. From [82L1] of Chap. 1

SC	Sy	dir/ind	m_e/m_0	m_{hh}/m_0	m_{lh}/m_0	$m_A \perp \parallel /m_0$
C	T_d	i		2.18	0.7	
Si	T_d	i	$\begin{cases}\perp 0.19\\ \parallel 0.92\end{cases}$	0.54	0.15	
Ge	T_d	i	$\begin{cases}\perp 0.081\\ \parallel 1.6\end{cases}$	0.3	0.043	
AlAs	T_d	i	$\begin{cases}\perp 1.56, 5.8\\ \parallel 0.19\end{cases}$	0.76	0.15	
AlSb	T_d	i	$\begin{cases}\perp 0.26\\ \parallel 1.0\end{cases}$	0.94	0.11	
GaN	C_{6v}	d	0.22			≈ 0.8
GaP	T_d	i	$\begin{cases}\perp 0.25\\ \parallel 7.25; 2.2\end{cases}$	0.6	0.17	
GaAs	T_d	d	0.066	0.47	0.07	
GaSb	T_d	d	0.042	0.35	0.05	
ZnO	C_{6v}	d	0.28			$\begin{cases}\perp 0.45\\ \parallel 0.59\end{cases}$
ZnS	C_{6v}	d	0.28			0.5
ZnSe	T_d	d	0.15	0.8	0.145	
ZnTe	T_d	d	0.12	0.6		
CdS	C_{6v}	d	0.2			$\begin{cases}\perp 0.7\\ \parallel 2.5\end{cases}$
CdSe	C_{6v}	d	0.13			$\begin{cases}\perp 0.45\\ \parallel 1.1\end{cases}$
CdTe	T_d	d	0.1	0.4		
CdCl	T_d	d	0.4	2.4		
CuBr	T_d	d	0.25	1.4		
CdI	T_d	d	0.3	≈ 2		

above-mentioned direct semiconductors, the band extrema are situated at the L points. Some semiconductors, such as the IV–VI compounds TiO_2 (rutil) and SnO_2 or the I–VI compound Cu_2O have a direct gap. The transitions between the band extrema are, in contrast to the other materials mentioned above, dipole forbidden because of their equal parity. The symmetry groups of these materials contain inversion as an element and parity is therefore a good quantum number. This feature has significant consequences for the optical properties.

A last example is shown in Fig. 8.19 where we demonstrate the transition of the alloy system $Cd_{1-y}Hg_yTe$ from a normal semiconductor ($y = 0$) to a narrow gap material $0 < y < 0.85$, and finally to a semimetal for $1 \geq y \geq 0.85$.

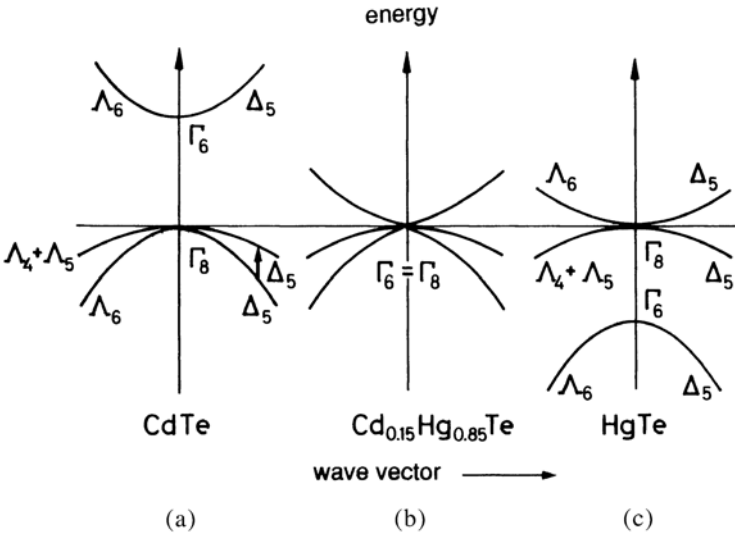


Fig. 8.19. The transition from a semiconductor (**a**) to a semimetal (**c**) for the alloy Cd_{1-y} ,Hg_y,.Te (**b**) as a function of the composition y [70M1]

8.9 Density of States, Occupation Probability and Critical Points

We start now to consider the density of states $D(E)$ for crystal electrons and holes using simple parabolic bands in the effective mass approximation, i.e.,

$$\text{conduction band (CB)} : E(\boldsymbol{k}_\mathrm{e}) = E_\mathrm{g} + \frac{\hbar^2 \boldsymbol{k}_\mathrm{e}^2}{2m_\mathrm{e}}\,, \tag{8.37a}$$

$$\text{valence band (VB)} : E(\boldsymbol{k}_\mathrm{h} = \frac{\hbar^2 \boldsymbol{k}_\mathrm{h}^2}{2m_\mathrm{h}}\,, \tag{8.37b}$$

where we take into account that the energy of a hole increases if it is brought deeper into the valence band.

With the help of (2.79) we find that the density of states depends on the energy in the conduction and valence bands according to

$$\mathrm{CB} : D(E) = (E - E_\mathrm{g})^{d/2-1}\,; \quad E > E_\mathrm{g} \tag{8.38a}$$

$$\mathrm{VB} : D(E) = E^{d/2-1}\,; \quad E > 0\,; \quad d = \text{dimensionality} \tag{8.38b}$$

for three-, two- and one-dimensional systems. This situation is shown schematically in Fig. 8.20 where we also include a set of δ-functions for the density of states in a quasi-zero-dimensional system.

The appearance of the various subbands for $d < 3$ will be explained in the next section.

The statistics with which we describe the occupation probability are the Fermi–Dirac statistics, as already mentioned. It reads for electrons and holes

$$\text{electrons}: \; f_{\mathrm{FD}}(E) = \left(\mathrm{e}^{(E-E_{\mathrm{F}}^{\mathrm{e}})/(k_{\mathrm{B}}T)} + 1\right)^{-1} \tag{8.39a}$$

$$\begin{aligned}\text{holes}: \; f_{\mathrm{FD}}(E) &= 1 - \left(\mathrm{e}^{(E-E_{\mathrm{F}}^{\mathrm{h}})/(k_{\mathrm{B}}T)} + 1\right)^{-1} \\ &= \left(\mathrm{e}^{-(E-E_{\mathrm{F}}^{\mathrm{h}})/(k_{\mathrm{B}}T)} + 1\right)^{-1} . \end{aligned} \tag{8.39b}$$

The Fermi energies (or chemical potentials) for electrons and holes $E_{\mathrm{F}}^{\mathrm{e,h}}$ are non-zero, unlike the situation for phonons or photons, since there is a conservation law for the number of electrons. The Fermi energies depend on the concentrations of electrons (and holes), on the temperature of the electron gas, and on material parameters such as E_{g}, m_{e}, or m_{h} as shown below.

Since electrons can be exchanged between the valence and conduction bands, for example by thermal excitation and recombination, it follows that

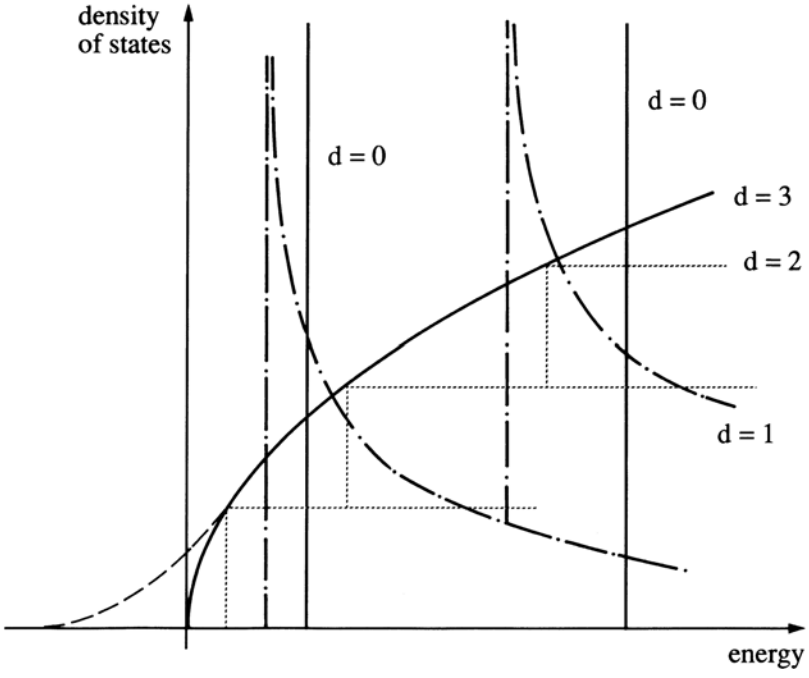


Fig. 8.20. Schematic drawing of the density of states as a function of energy for three-, and quasi-two-, one- and zero-dimensional systems in the effective mass approximation. The *dashed line* corresponds to disorder-induced localized states (Sect. 8.15)

in thermodynamic equilibrium

$$E_{\mathrm{F}}^{\mathrm{e}} = E_{\mathrm{F}}^{\mathrm{h}} = E_{\mathrm{F}} \tag{8.40a}$$

or, in other words, the chemical potential of the electron–hole pair system μ_{eh} is zero

$$\mu_{\mathrm{eh}} = E_{\mathrm{F}}^{\mathrm{e}} - E_{\mathrm{F}}^{\mathrm{h}} = 0\,. \tag{8.40b}$$

We shall come back to this point later.

The density of electrons in a certain energy range, say in the conduction band, is then given by the integral over the density of states weighted by the occupation probability. For parabolic bands with a degeneracy g we find for $d = 3$ and for the total density of electrons in the conduction band n:

$$n = \frac{N}{L^3} = \int_{E_{\mathrm{g}}}^{\infty} g_{\mathrm{e}} \frac{1}{2\pi^2} \left(\frac{2m_{\mathrm{e}}}{\hbar^2}\right)^{3/2} (E - E_{\mathrm{g}})^{1/2} \left(\mathrm{e}^{(E-E_{\mathrm{F}})/(k_{\mathrm{B}}T)} + 1\right)^{-1} \mathrm{d}E\,, \tag{8.41}$$

where we have used our knowledge from Sect. 2.6 and the dispersion relation (8.37). Equation (8.41) leads to the Fermi integral which cannot be solved analytically but can be found in mathematical tables.

Very often E_{F} lies in the gap. In this case the electron gas is said to be non-degenerate and the part of the Fermi function which overlaps with a finite density of states can be approximated quite well by the Boltzmann function, as can be seen in Fig. 2.5 where we compare Fermi–Dirac, Bose–Einstein and Boltzmann statistics.

With this assumption the Fermi–Dirac statistics (8.41) can be replaced by classical Boltzmann statistics and the integral can be solved analytically giving

$$n = g_{\mathrm{e}} \left(\frac{m_{\mathrm{e}} k_{\mathrm{B}} T}{2\pi\hbar^2}\right)^{3/2} \mathrm{e}^{(E-E_{\mathrm{F}})/(k_{\mathrm{B}}T)}\,. \tag{8.42a}$$

The term in front of the exponential is the so-called effective density of states n_{eff} depending on T and on material parameters. So we can write

$$n = n_{\mathrm{eff}}\,(T, m_{\mathrm{e}})\; \mathrm{e}^{(E_{\mathrm{g}}-E_{\mathrm{F}})/(k_{\mathrm{B}}T)}\,. \tag{8.42b}$$

If the electron gas is degenerate, i.e., if E_{F} is situated in the conduction band, we have to use the full (8.41) except for $T \to 0$ where the Fermi function converges to a step function and n is given by

$$n = \frac{2}{6\pi^2} g \left(\frac{2m_{\mathrm{e}}}{\hbar^2}\right)^{3/2} (E_{\mathrm{F}} - E_{\mathrm{g}})^{3/2} \quad \text{for} \quad T = 0\,\mathrm{K}\,. \tag{8.42c}$$

In analogy we find for a three-dimensional non-degenerate gas of holes

$$p = p_{\mathrm{eff}}\, \mathrm{e}^{-E_{\mathrm{F}}/k_{\mathrm{B}}T} \tag{8.42d}$$

with

$$p_{\text{eff}} = g_{\text{h}} \left(\frac{m_{\text{h}} k_{\text{B}} T}{2\pi\hbar^2} \right)^{3/2} . \tag{8.42e}$$

We should mention here that in thermodynamic equilibrium, i.e., when (8.40a) is valid, one finds that the product np is independent of E_{F} in the non-degenerate case, i.e.,

$$np = n_{\text{eff}}\, p_{\text{eff}}\, \mathrm{e}^{-E_{\text{F}}/k_{\text{B}}T} =: n_{\text{i}}^2(T) . \tag{8.43}$$

where $n_{\text{i}}(T)$ is the intrinsic carrier density.

Out of thermal equilibrium, a situation that occurs, e.g., under optical pumping or via carrier injection in a p-n junction, one frequently has a thermal distribution of the carriers within their bands with some effective carrier temperature $T_{\text{e,h}}$, which may be higher than the lattice temperature T_{L}. In this case, the distribution of the carriers in their respective bands can be described by Fermi-functions, which contain separate quasi-Fermi levels for electrons and holes. In this case one has

$$n \cdot p \neq n_{\text{i}}^2\ (T) \tag{8.44a}$$

and in contrast to (8.40b)

$$\mu_{\text{eh}} = E_{\text{F}}^{\text{e}} - E_{\text{F}}^{\text{h}} \neq 0 \tag{8.44b}$$

The occupation of donor or acceptor levels situated in the forbidden gap or of localized tail states can be also covered by the concept of Fermi–Dirac occupation probabilities. For defect states degeneracy factors appear in the Fermi statistics which take care, e.g., of the fact that an unoccupied simple donor in a direct gap semiconductor offers two empty, spin degenerate states. But if one electron sits on the donor, the other electron level is shifted to higher energies due to Coulomb interaction. For details of this topic see, e.g., [87B1].

The last aspect we want to discuss in this section are the so-called critical points or van Hove singularities. One type of these critical points has already been treated. If a parabolic band starts at energy E_0 (e.g., a conduction band) we have above a density of states given by (8.45a) and zero below, i.e., in a d dimensional system

$$D\,(E) \sim \begin{cases} (E - E_0)^{\frac{d}{2}-1} & \text{for } E \geq E_0 \\ 0 & \text{for } E < E_0 \end{cases} . \tag{8.45a}$$

This is a so-called M_0 critical point in three dimensions.

If a parabolic band ends at a certain energy E_0 like a valence band, we have the inverse situation, i.e.,

$$D\,(E) \sim \begin{cases} (E_0 - E)^{\frac{d}{2}-1} & \text{for } E \leq E_0 \\ 0 & \text{for } E > E_0 \end{cases} . \tag{8.45b}$$

The name of this type of critical point is M_d where d is the dimensionality.

In three dimensions we can additionally have a situation where the curvature of $E(\boldsymbol{k})$ is positive in two orthogonal directions and negative in the third one or vice versa. These so-called M_1 and M_2 critical points exhibit a constant density of states D_0 (similar to a two-dimensional system) on the energy side of E_0 where two branches exist and a dependence

$$D(E) \sim \left(1 - A(E - E_0)^{1/2}\right) \tag{8.45c}$$

on the other side.

A saddle point or M_1 critical point in a two dimensional system finally gives a logarithmic singularity on both sides of E_0

$$D(E) \sim -\ln(|E - E_0|) \; . \tag{8.45d}$$

For details see, e.g, [96Y1] of Chap. 1.

Critical points are also important in the optical properties of vertical band-to-band transitions. In this case not the density of states in each of the two bands is important, but rather the joint or combined density of states of conduction and valence band separated in their $E(\boldsymbol{k})$ dispersion vertically by the photon energy $\hbar\omega$.

8.10 Electrons and Holes in Quantum Wells and Superlattices

We have already introduced structures of reduced dimensionality in connection with the discussion of phonons in Sect. 7.10.

This topic is presently of even much wider relevance for the charge carriers in semiconductors, both concerning basic and applied research. Therefore we shall spend some time (or pages) here to explain the main ideas in some detail, trying again to stress the physical and intuitive understanding of the basic concepts and we refer the reader for more theory oriented treatments, e.g., to [86B1,87P1,87W1,88G1,88R1,99B1,99B2,00B1,02K1] or [92L1,93H1, 93S1,95I1,96O1,98D1,98G1,98S1,99B1,01C1,01H1,01L1,02S1] of Chap. 1.

A quasi two-dimensional electron system can be realized, if the motion of the carriers is confined in one dimension by a suitable potential on a length scale comparable to or smaller than the de Broglie wavelength of these particles, or the mean distance between scattering events in the sense of a relaxation time approach. For typical semiconductors this limit is reached for the width of the confining potential below a few tens of nm at least at low temperatures.

In the plane normal to this confining potential the electron is assumed to move freely as a Bloch wave or as a plane wave in the sense of the effective mass approximation.

One of the first realizations was via simple heterostructures, where two different semiconductors where grown on top of each other. If the band-discontinuity at the heterointerface and the position of the Fermi levels are suitable, an arrangement of the bottom of the conduction band may appear as depicted schematically in Fig. 8.21.

In thermodynamic equilibrium the Fermi energy is constant across the interface. The bottom of the conduction band forms a roughly triangular potential (of finite height only) for the electrons. The electrons can move freely in the plane of the interface but are quantized in their motion normal to it. The envelope functions of the first two quantized states are shown schematically.

In effective mass approximation the wave function can be written as

$$\Psi_n\left(\boldsymbol{r}, \boldsymbol{k}_{\parallel}\right) = \exp\left\{i\left(k_x x + k_y y\right)\right\} \chi_n\left(z\right) \tag{8.46a}$$

with

$$E_n\left(\boldsymbol{k}_{||}\right) = E_n + \frac{\hbar^2\left(k_x^2 + k_y^2\right)}{2m_{\mathrm{e}}}\,. \tag{8.46b}$$

The so-called envelope function $\chi^n(z)$ with quantization-energy E_n is obtained as the solution of the given potential. For the solution of the potential

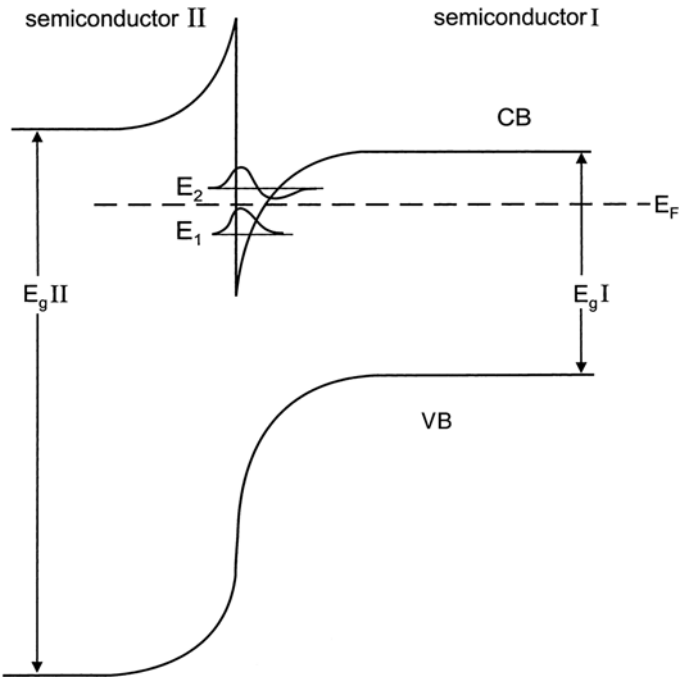


Fig. 8.21. Conduction and valence band edges around a heterojunction between two different semiconductors, forming a quasi-two-dimensional electron gas at the interface. E_{F}: Fermi level

of Fig. 8.21 the first two quantized states are shown schematically. At every value E_n starts a quasi-two-dimensional band with constant density of states as already discussed in Sect. 2.6 and shown in Fig. 8.20.

In the situation of Fig. 8.21 it is assumed that E_F is situated slightly above E_1 so that a quasi-two-dimensional electron gas exists in the first subband already without excitation or carrier injection.

Note that this structure does not provide any confining potential for the holes.

The single heterostructures of Fig. 8.21 play an important role for electronic devices like field-effect transistors. There and especially in (electro) optic devices and structures, however one much more frequently uses so-called quantum wells. They consist of two close-lying heterointerfaces as shown in Fig. 8.22a for the conduction band only. We come back to the valence band in a few pages.

The band discontinuities at the two interfaces form, in the simplest approximation, a one-dimensional square well potential of finite depth. An idealization with infinitely high barriers is shown in Fig. 8.22b. If we define the potential at the bottom of the well as zero and the zero of the z-axis in the middle of the well, we find as the solutions of this textbook problem with increasing energy wavefunctions alternating between even and odd parity of the cosine- and sine-type, i.e., for even n_z

$$\Psi_n(\boldsymbol{r}) = \chi_n(z)\ \mathrm{e}^{\mathrm{i}\boldsymbol{k}_\parallel \boldsymbol{r}_\parallel} = V^{-\frac{1}{2}}\ \mathrm{e}^{\mathrm{i}(k_x x + k_y y)} \cos\left(n_z \frac{\pi}{l_z} z\right) \tag{8.47a}$$

and an analogous expression with sine for odd n_z and

$$E_n\left(k_\parallel\right) = \frac{\hbar^2 \boldsymbol{k}_\parallel^2}{2m_\mathrm{e}} + E_{n_z} = \frac{\hbar^2 \boldsymbol{k}_\parallel^2}{2m_\mathrm{e}} + \frac{\hbar^2 \pi^2 n_z^2}{2m_\mathrm{e} l_z^2} = \frac{\hbar^2}{2m_\mathrm{e}}\left(\boldsymbol{k}_\parallel^2 + n_z^2 \frac{\pi^2}{l_z^2}\right) \tag{8.47b}$$

with $n_z = 1, 2, 3, \ldots$.

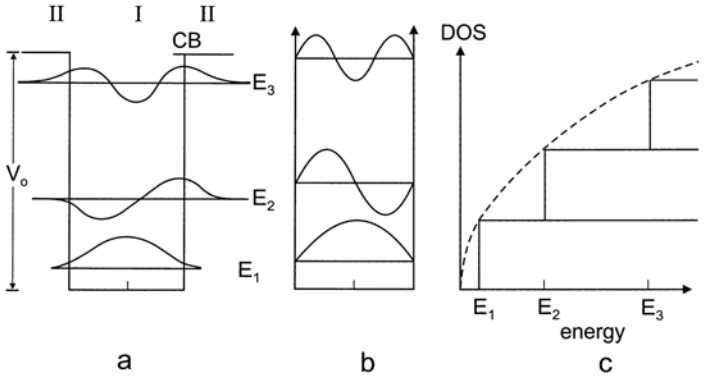


Fig. 8.22. The confining potential for electrons in a quantum well (**a**) and the idealization with infinitely high barriers (**b**) and the density of states (**c**)

This means we have an infinite number of quantized states E_{nz} and at every E_{nz} a new subband starts with constant density of states as shown in Fig. 8.22c.

For the more realistic case of a finite potential V_0 (Fig. 8.22a) the wave functions penetrate exponentially decaying into the barrier.

Since the problem has inversion symmetry, the wave functions again have even or odd parity. From the Schrödinger equation we obtain for the envelope function in z direction:

$$-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\left(\frac{1}{m_\mathrm{e}(z)}\frac{\partial}{\partial z}\right)\chi_{n_z}(z)+V(z)\chi_{n_z}(z)=E_{n_z}\chi(z) \tag{8.48a}$$

with

$$V(z)=\begin{cases}0 & \text{for} \quad |z|\le l_z/2\\ V_0 & \text{for} \quad |z|>l_z/2\end{cases} \tag{8.48b}$$

The term $m_\mathrm{e}(z)$ allows to include different effective masses in well (material I) and barrier (material II).

The ansatz for the wave function $\chi_{n_z}(z)$ is

$$\chi_n(z)=\begin{cases}\pm B\exp(\kappa z) & \text{for} \quad z\le -l_z/2\\ A\begin{Bmatrix}\cos\\ \sin\end{Bmatrix}(kz) & \text{for} \quad -l_z/2\le z\le l_z/2\\ B\exp\{-\kappa z\} & \text{for} \quad z\ge l_z/2\end{cases} \tag{8.48c}$$

with

$$k=\sqrt{\frac{2m_\mathrm{e}^\mathrm{I}E_{nz}}{\hbar^2}};\quad \kappa=\sqrt{\frac{2m_\mathrm{e}^\mathrm{II}(V_0-E)}{\hbar^2}}\,, \tag{8.48d}$$

which fulfils the necessity for normalizability.

The boundary conditions at the interfaces require the steadiness of both the wave function and the current for stationary solutions, i.e., of

$$\chi(z) \quad \text{and} \quad \frac{1}{m_\mathrm{e}}\frac{\partial}{\partial_z}\chi(z)\,. \tag{8.48e}$$

This leads to a transcendental equation, which can be solved graphically or numerically as detailed in the above references.

The main results are that there exists only a finite number of quantized states as shown schematically in Fig. 8.22a and a continuum of states above V_0.

The first few quantized states also follow, to a good approximation, a n_z^2 law, but with a slightly smaller prefactor than in (8.47b) due to the finite barrier height.

We have seen in Sect. 8.8 that the top of the valence band is usually more complex around the Γ-point than the conduction band minimum. Even if we neglect the spin–orbit split-off band we are left with heavy and light hole bands in cubic semiconductors and the crystal field split A and B valence bands in wurtzite type materials.

In the first approximation we expect that the heavy and light hole subbands form two different series of subbands due to their different effective masses and the resulting different quantization energies as can be deduced, e.g., from (8.48). This is shown schematically in Fig. 8.23 where we plot the valence band maximum for a quantum well and the envelope functions of the first four quantized levels.

The next complication is that the wave function of the heavy hole in z direction (which results in the quantization energies $E_{n_z\mathrm{hh}}$) exhibits a light hole dispersion in the plane of the well and vice versa as we have explained schematically for $p_{x,y,z}$ orbitals with Fig. 8.13 (instead of the real sp^3 hybrid wave functions). Consequently, the in-plane dispersion relations tend to cross for finite $\boldsymbol{k}_\parallel$. In fact, this crossing is avoided due to the coupling between the various subbands, resulting in a rather complex in-plane dispersion relation as shown in Fig. 8.24 for a GaAs quantum well embedded between $\mathrm{Al}_{1-y}\mathrm{Ga}_y\mathrm{As}$ barriers. As can be seen, the dispersion relation sometimes deviates strongly from a parabolic one and may in some intervals of $\boldsymbol{k}$ even obtain the opposite curvature. Consequently there will be deviations from the constant two-dimensional density of states expected in the effective mass approximation. Due to these corrections, the $1hh$ dispersion in the plane fortunately shows a dispersion that comes close to m_{hh} while the $2hh$ states obtain an in-plane effective mass, which is even smaller than the light hole mass.

The above consideration is valid if the quantum well is grown on a (100) plane. For high-index planes, things may get even more complicated as detailed in [95I1] of Chap. 1 and [96T1] of Chap. 3.

Until now we have treated only single quantum wells (SQW) and the situation for electrons and holes separately. Now we want to introduce the terms multiple quantum well (MQW), coupled quantum wells and superlattice (SL) and discuss the various types of band alignment.

In Fig. 8.25a we assume, that the band discontinuities of CB and VB are such that electrons and holes are confined in the same material. This is a so-

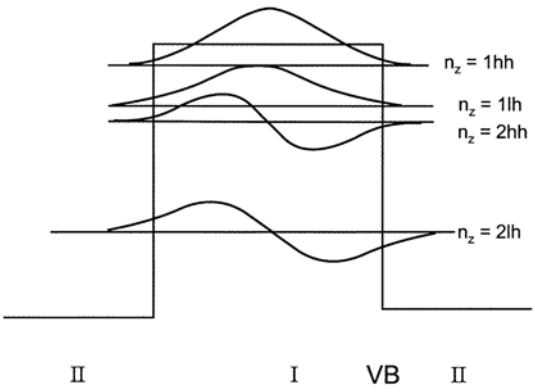


Fig. 8.23. The first two quantized levels for heavy and light holes (schematic)

called type I band alignment. If the width of the barrier is so large, that the wave functions of the first quantized electron states of adjacent wells do not overlap as in Fig. 8.25a we speak of a multiple quantum well (MQW). The density of states still increases as a sum over Heavyside functions at every E_{n_z} only the prefactor grows in proportion to the number of identical QW in the MQW.

If we now start to make the thickness of the barriers thinner and thinner we come to a regime, where the exponential tails of the wave functions in the barriers start to overlap as shown in Fig. 8.25b.

If we couple only a few, e.g., two quantum wells in this way we obtain evidently coupled quantum wells. If they are identical, the ground state splits into a symmetric and an anti-symmetric wave function, which is slightly higher in energy as shown in Fig. 8.26a.

For coupled wells of different width as in Fig. 8.26b parity is no longer a good quantum number and the lowest states of the wide and the narrow well tend to obtain a slight admixture of the other wave function.

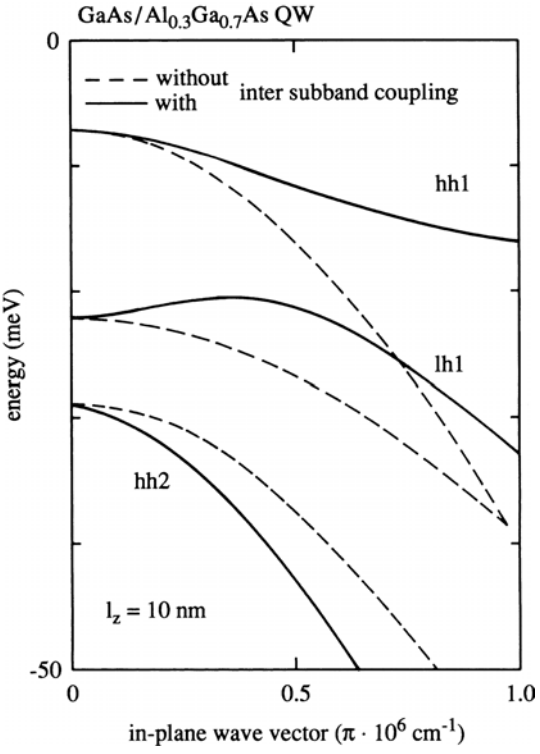


Fig. 8.24. The in-plane dispersion relation of the first three valence subbands in a GaAs/$Al_{1-y}Ga_yAs$ quantum well without (- - -) and with (—) intersubband interaction ([86B1] or [01L1] of Chap. 1)

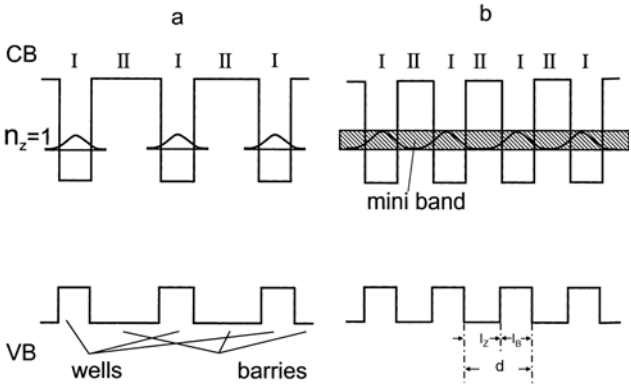


Fig. 8.25. The spatial variation of CB and VB extrema of a multiple quantum well (MQW) (**a**) and of a superlattice (SL) (**b**)

If we couple many identical quantum wells via sufficiently narrow barriers as shown in Fig. 8.25b we obtain in the z direction a new periodicity $d = l_z + l_b$ and the structure is known as a (one-dimensional) superlattice (SL). Ideally d is given by a sum of integer multiple monolayers of materials I and II.

We can now apply what we have learned before e.g. in Sect. 4.1, 5.1 or 7.10. We obtain in the k_z direction a Brillouin zone extending from

$$-\pi/d \leq k_z \leq \pi/d \tag{8.49}$$

much smaller than the extension in the k_x and k_y direction, and a series of mini bands in this mini Brillouin zone with a finite curvature as shown schematically in Fig. 8.27a,b for $k_{\|} = 0$. These bands arise from the quantization levels E_{n_z} in the uncoupled (M)QW. The width of the mini-bands is typically a few meV and increases with increasing band index and decreasing barrier width since the coupling, which determines the bandwidth, increases in both cases (Fig. 8.27c).

The curvature and shape of the bands in the k_x, k_y plane remains essentially the one of the QW. In Fig. 8.26d,e we finally compare the density of states for a SL with the one of a MQW. Instead of a sum of Heavyside functions starting at E_{n_z} we obtain a steady variation of DOS(E), starting with

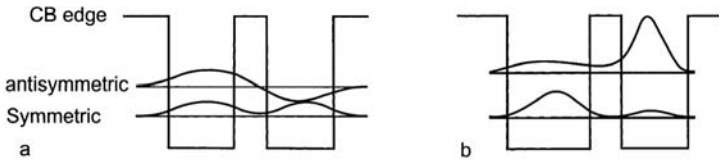


Fig. 8.26. Two coupled identical (**a**) and different (**b**) quantum wells and the envelope functions of the lowest states (schematic)

a square-root behavior of a M_0 critical point and reaching the constant 2d density of states in a M_1 type behavior (see Sect. 8.9).

Similar mini bands can form for the various quantized hole states. If the hole mass is heavier, the penetration of the wave function into the barrier and the width of the bands are smaller.

If the band discontinuities and the band alignment is such that electrons and holes are confined in the same material as shown, e.g., in Fig. 8.28a, we speak of a type I band alignment.

If the arrangements of the band edges are as in Fig. 8.27b we speak of type II staggered. The electrons are confined in material I and the holes are essentially confined to material II. The transition is said to be spatially indirect. In Sect. 9.3 or 15.2 we shall see an example for a type I → type II transition as a function of layer thickness in GaAs/AlAs superlattices.

The arrangement of Fig. 8.28c where the bottom of the conduction band of material I falls below the top of the valence band of material II is known as type II misaligned. The combination of a semiconductor ($E_g > 0$) with a semimetal ($E_g \approx 0$) forms a type III structure (Fig. 8.28d). Examples for the various cases can be found, e.g., in [01L1] of Chap. 1 or in the next sections.

It should be mentioned that it is also possible to produce almost parabolic confinement potentials by a suitable and continuous spatial variation of the

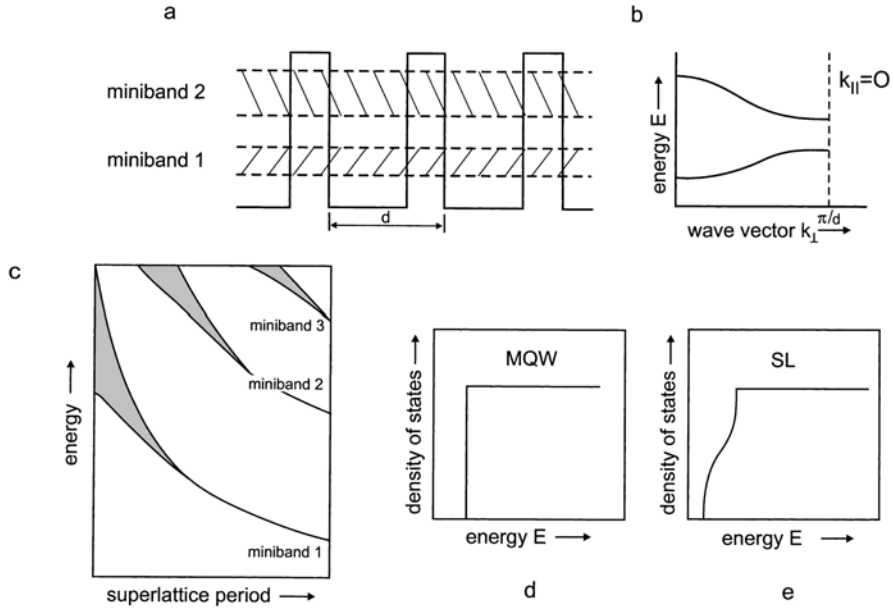


Fig. 8.27. Mini bands in the z direction ($k_x = k_y = 0$) of a SL (**a**, **b**) the width and position of these mini bands as a function of the superlattice period (**c**) and the density of states (**e**) compared to the one of MQW (**d**). According to [99B1, 01L1] of Chap. 1

composition $x(z)$ from the well, e.g., GaAs towards the barrier $Al_{1-x(z)}Ga_{x(z)}As$. In this case the energetic spacing between the first few quantized levels E_{n_z} is almost constant as can be expected for the parabolic harmonic oscillator potential.

An almost parabolic confinement potential is also obtained in so-called nipi structures or doping superlattices. These structures do not contain abrupt or continuous hetero junctions but are made from one material but with alternating n and p doping. Usually the differently doped layers are separated by undoped intrinsic layers explaining the name nipi structure. Electrons from the donors (see Sect. 8.14) and the holes from the acceptors recombine, leaving the space-charge layers of D^+ and A^- behind. By solving the Poisson equation for the electrostatic potential $\phi(\boldsymbol{r})$

$$\boldsymbol{\nabla}^2 \varphi(\boldsymbol{r}) = -\frac{\varphi(\boldsymbol{r})}{\varepsilon\varepsilon_0} \tag{8.50}$$

one obtains in growth direction a periodic modulation of the conduction and valence band edges as shown in Fig. 8.29.

The effective band gap of the spatially indirect transition indicated in Fig. 8.29 is much smaller than that of the parent material. The confinement potential is parabolic in the vicinity of the extrema resulting in energetically equidistant subband levels for electrons and holes.

The spatial overlap between the wave functions in the first confined states of electrons and holes is small. Consequently their lifetime is long. This makes such structures very sensitive, but slow photodetectors.

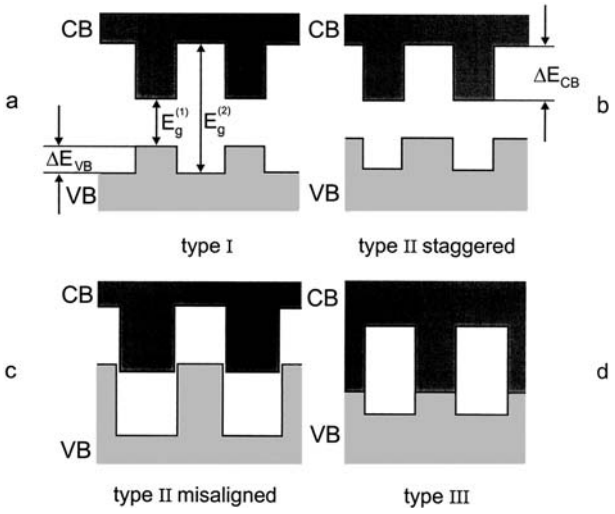


Fig. 8.28. Various band alignments in MQW or SL: type I (**a**), type II staggered (**b**), type II misaligned (**c**), type III (**d**)

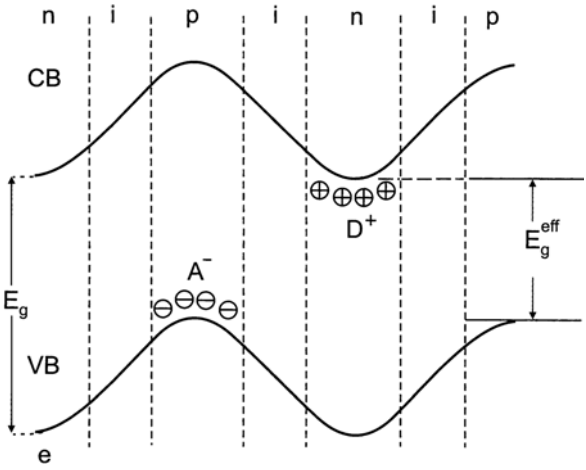


Fig. 8.29. The band alignment in a doping superlattice or nipi structure

In some cases MQW structures are incorporated into nipi structures for special purposes. For more details see, e.g., [86D1, 92M1].

8.11 Growth of Quantum Wells and of Superlattices

The growth of SQWs, MQWs and SLs presently is not a topic of this book but since structures of reduced dimensionality form a large fraction of the research in semiconductor optics (and physics) we give a short outline of the most widely used techniques for the epitaxy of quantum wells and superlattices and later on in this chapter for quantum wires and quantum dots. For more detailed information the reader is referred to [90G1, 01L1] of Chap. 1 and references given therein.

The two most important ways to grow thin (quantum)films and related (nano)structures are molecular beam epitaxy (MBE) and metal-organic chemical vapour deposition (MOCVD) also known as metal-organic vapour phase epitaxy (MOVPE) (see Fig. 8.30).

In MBE one creates, in an ultra-high vacuum chamber, atomic or molecular beams of the constituents of the semiconductor layers to be grown from heated effusion cells with mechanical shutters. By simultaneously or alternatively opening the shutters of, e.g., Ga and As cells it is possible to grow a GaAs layer on a suitable, heated, monocrystalline substrate and with the Al and As cells an AlAs layer, respectively. Doping atoms can be incorporated with further cells, e.g., Si or Be for n and p doping of III–V compounds. In II–VI epitaxy one sometimes also uses compound sources, e.g., for CdS.

Some modifications of the MBE process are known as atomic layer eptiaxy (ALE) or migration enhanced epitaxy (MEE), which allow one to grow layers controlled on a monolayer scale.

In MBE the atomic or molecular beams propagate collision-free from the effusion cells through the ultrahigh vacuum to the substrate where they are physisorbed and then either chemisorbed and integrated in the growing surface or desorbed again. This is a technique where one grows far from thermodynamic equilibrium.

In hot-wall (beam) epitaxy (HWE or HWBE) one grows, in contrast, close to equilibrium. In this technique the substrate is mounted directly on top of the effusion cell or very close to it. The source is kept at a temperature T_{so}, the walls at higher temperature $T_{\mathrm{w}} > T_{\mathrm{so}}$ to avoid condensation of the source material at the walls and the substrate is kept at slightly lower temperature $T_{\mathrm{sub}} < T_{\mathrm{so}}$, to allow a net flow of material from the source to the substrate. In this case, growth occurs close to thermodynamic equilibrium.

In MOCVD so-called precursors like AsH_3 and $Ga(CH_3)_3$ are brought with an inert gas like N_2 as a carrier into the reactor. A chemical reaction ideally takes place directly on the heated substrate in our case, e.g.,

$$AsH_3 + Ga(CH_3)_3 \rightarrow GaAs + 3CH_4 \tag{8.51}$$

resulting in the deposition of GaAs.

It should be noted that many of the materials used in MBE or MOCVD are highly poisonous and some of them cancerogen. Therefore, safety precautions have to be used when opening a MBE chamber or in conditioning the exhaust gas of a MOCVD reactor. It should be noted further that epitaxy is, similarly to crystal growing in general, a highly developed art that demands a lot of experience and skill to be mastered.

The next topic we address in this subsection is which materials can be grown epitaxially on top of each other with good quality? The materials should be similar, in other words, they should have:

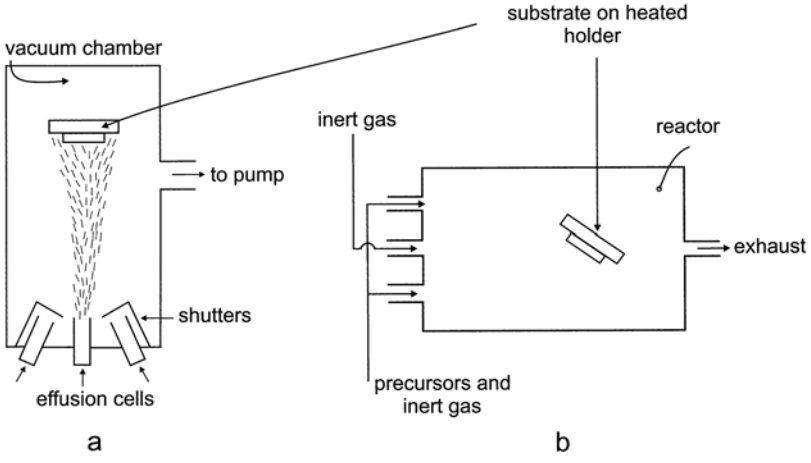


Fig. 8.30. The schematics of a MBE chamber (**a**) and of a MOCVD reactor (**b**)

- Equal or similar crystal structures, e.g., both should crystallize in the zincblende structure or should show at least the same coordination like a tetrahedral one in diamond, zincblende and wurtzite structures.
- Almost identical lattice constants. Below we discuss consequences resulting from a lattice misfit.
- The same valences, which means that III–V semiconductors are more easy to grow on a III–V substrate than on II–VI or group IV substrates and vice versa.
- Similar types of chemical binding, e.g., either preferentially covalent or ionic.

The trivial way to fulfil these conditions is homoepitaxy, i.e., to grow on a substrate the same material. This partly allows one to produce high quality, e.g., chemically very pure layers on a less perfect substrate of the same material.

In contrast, substrate and growing layers are different in heteroepitaxy. As can be seen in Fig. 8.31 the system GaAs, AlAs and their alloys $Ga_{1-y}Al_yAs$, which exist for all values $0 \leq y \leq 1$, fulfil these conditions in an almost ideal manner. The band alignment is of type I where GaAs forms the well with about 60% of the band discontinuity in the conduction band. GaAs is a direct gap semiconductor, AlAs an indirect one. The transition occurs in the alloy for $y \approx 0.43$ (see [96K1] of Chap. 1). To avoid complications by combining a direct gap material with an indirect one, one usually limits y to values $0.3 \leq y \leq 0.4$.

Since GaAs substrates are available in the highest quality compared to all other substrate materials, one tries to use it and to choose the compositions of the layers in a way to match the lattice constants like in $In_{0.51}Al_{0.49}P$, $AlAs_{0.96}P_{0.04}$ or $Ga_{0.51}In_{0.49}P$. InP is another widely used substrate to which III–V alloys can be lattice matched like $Al_{0.48}In_{0.52}As$, $GaAs_{0.5}Sb_{0.5}$ or $Ga_{0.47}In_{0.53}As$. Quaternary alloys like $In_{1-y}Ga_yAs_{1-x}P_x$ allow one to vary the lattice constant and the gap to some extent independently as can be seen from the border lines.

A further III–V system, which has become rather popular in recent years due to the success of building blue- and UV-light emitting and laser diodes, is GaN including its alloys with InN and AlN. Due to the huge lattice mismatch to GaAs, this system is usually grown on SiC, ZnO or Al_2O_3, in contrast to the ZnSe based blue-green emitting structures, which are in most cases grown on GaAs due to the almost perfect lattice match and in spite of the problems resulting from the interface between a III–V and a II–VI compound.

A type-I system, which was and is still intensely investigated because of possible applications of laser diodes in the blue-green to yellow spectral range are ZnSe-based structures including $Zn_{1-y}Cd_ySe$ or $ZnSe_{1-x}Te_x$ wells or $Zn_{1-y}Mg_ySe_{1-x}S_x$ barriers.

In many cases it is unavoidable to grow materials on each other, that have different lattice constants, e.g., for $In_{1-y}Ga_yAs$ on GaAs, $ZnSe_{1-x}Te_x$ wells on

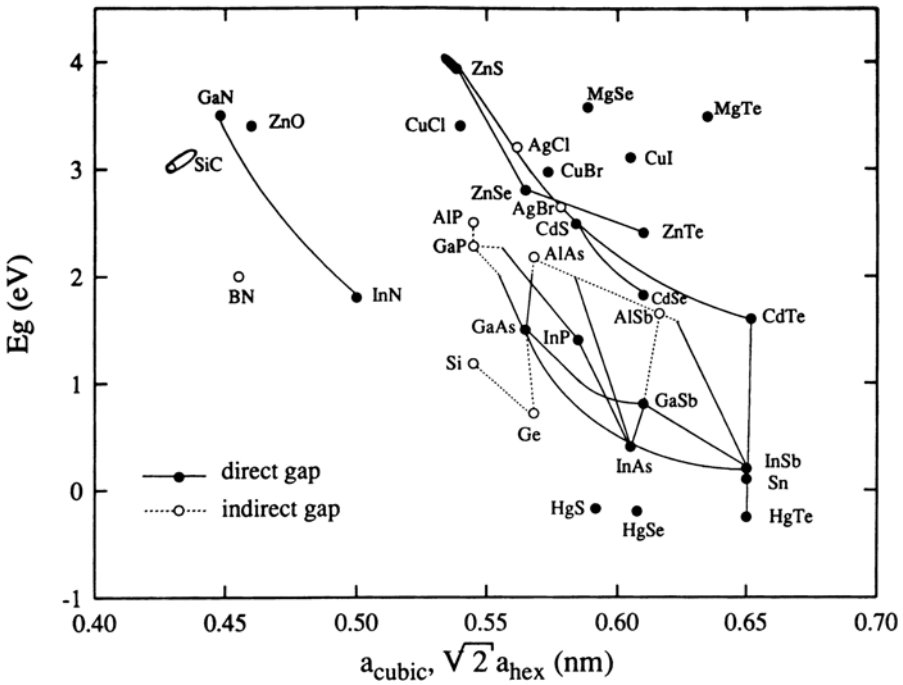


Fig. 8.31. Width of the band gaps as a function of the lattice constant for various semiconductors and their alloys. *Solid lines*: direct gap; *dashed lines*: indirect gap. Note that this Figure gives information about the band gap only, but not about the relative alignment, i.e., whether the combination will be of type I, II or III. Data from [82L1, 01L1] of Chap. 1

ZnSe or for type II CdS/ZnSe or CdSe/CdS superlattices. Another presently widely investigated strained layer superlattice system is formed by the group V elements C, Si and Ge.

Such a lattice misfit has consequences for the growing layer. If the growing layer wets the substrate, it will start to grow as a tensile or compressively strained layer, depending on the sign of the lattice mismatch. In addition tensile in-plane strain usually results in a uniaxial compressive distortion of the lattice in growth direction and vice versa.

The elastic energy stored in the strained layer increases with layer thickness until the lattice relaxes at a critical thickness l_c to its own lattice constant. This relaxation can result in the formation of dislocations like in CdS on GaAs [94G1, 97H1]. The formation of dislocations is highly undesirable because they form scattering and nonradiative recombination centres, degrading the in plane carrier mobility and the luminescence efficiency. The formation (and multiplication) of dislocations in strained ZnSe-based laser diode structures (still) limits their lifetime to values unacceptable for commercial devices [82P1].

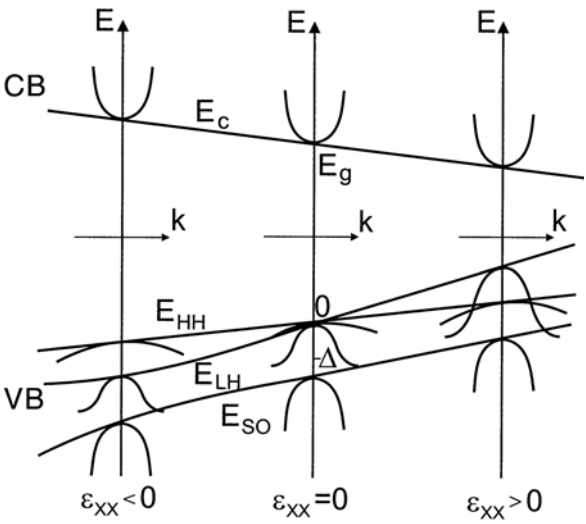


Fig. 8.32. Schematic drawing of the *lhh–lh* splitting, of the spin orbit splitting and of the change of the band gap under two-dimensional compressive ($\varepsilon_{xx} < 0$) or tensile strain ($\varepsilon_{xx} > 0$) [93T1]

The other possibility is that the growing layer starts to form on the wetting layer small islands, in which it can relax to its own lattice constant. This growth mode is known as Stranski-Krastanov growth mode [99B1].

In contrast the ideal two-dimensional layer by layer growth is known as Frank-Van der Merwe mode, and the direct formation of island without wetting layer is called Vollmer-Weber growth mode.

We come back to the Stranski-Kastanov growth mode in Sect. 8.13. As a consequence, strained layers are usually grown only below the critical thickness l_c in (M)QW or SL. Superlattices with many periods (e.g., 50 to 100) may adapt, after a few periods, an intermediate lattice constant which results in two-dimensional compressive and tensile strain in the two different materials, respectively, with a complementary uniaxial distortion normal to the layers. These are so-called free standing superlattices.

The strain also has consequences for the band structure. The width of the gap depends on strain via the deformation potentials, and the heavy and light hole bands split under the influence of strain already at $\boldsymbol{k} = 0$. This effect comes in addition to the splitting of these bands by quantization as already discussed in Sect. 8.10.

The usual behavior under tensile and compressive strain is shown in Fig. 8.32.

For compressive strain the band gap widens and the hh valence band is closer to the conduction band. For tensile strain the situation the opposite.

Depending on the sign of the strain this splitting may either enhance or partly compensate the splitting induced by quantization.

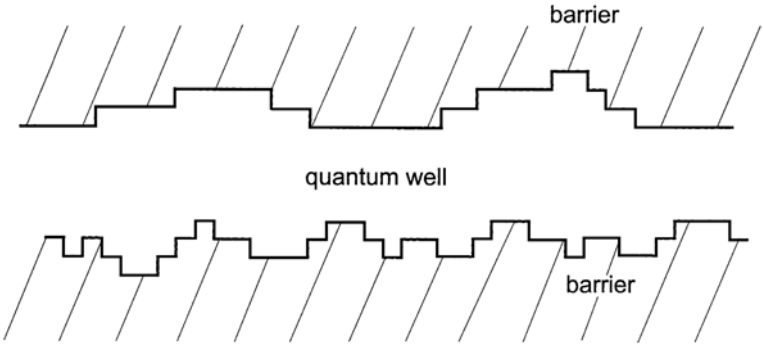


Fig. 8.33. Thickness fluctuations of a quantum well with two different length scales on both sides (schematic)

To conclude this section we briefly mention some complications encountered in real quantum wells and superlattices.

In spite of all progress in epitaxy, real quantum wells exhibit, at both interfaces fluctuations of at least one monolayer height as shown schematically in Fig. 8.33.

The length scale of the fluctuations may be different, e.g., compared to the excitonic Bohr radius and it may even depend on the growth sequence, i.e., if GaAs is grown an AlAs or vice versa.

The next problem is that the interfaces are usually not sharp but washed out by diffusion and segregation. These effects are, e.g., well known for InAs wells in GaAs or for CdSe wells in ZnSe. See, e.g., [97W1] or [01L1] of Chap. 1 and the references given therein. They can be influenced to some extent by the growth parameters.

In the caption of Fig. 8.31 we noted that it gives the band gap of the various compounds, but not the band alignment. The band alignment depends trivially on the two materials but is often not very precisely known (see, e.g., [01L1] of Chap. 1). In the case of alloys in wells or barriers the offset ratio of the conduction and valence band may depend on the composition of the alloy. It depends further on strain, on the crystallographic orientation of the two materials, on well and barrier thicknesses, on the termination of the interfaces especially in case of quaternary systems AB/CD without common anion or cation, on the presence of dipole-layers at the interface, on intermixing (see above) and even on the growth sequence, i.e., if material I is grown on top of material II or vice versa, the so-called non-commutativity.

It should be emphasised that all these types of disorder – like well-width fluctuations, intermixing or alloy formations – result in the formation of localized states, as will be explained in more detail in Sect. 8.15.

These statements show that quantum-structures give, in addition to the confinement effects themselves, ample place for problems, research, well-thought-out experiments and controversial interpretations.

8.12 Quantum Wires

If we confine the motion of the carriers in a second direction the y direction, by a suitable potential, e.g., a discontinuity of the band gap in a heterostructure or the transition to vacuum or air, there remains only free motion in one direction. Consequently, such structures are called quasi-one-dimensional or quantum wires.

The energies of the carriers are given in an effective mass approximation by

$$E\left(k_x\right) = E_{ni,nj} + \frac{\hbar^2 k_x^2}{2m_{\mathrm{e,h}}} \tag{8.52}$$

where we assume that the only remaining direction of free motion is along the x-axis. At every energy level $E_{ni,nj}$ starts a quasi-one-dimensional subband. The singularities of the density of states according to Fig. 8.20 expected at the edge of the energy subband is more or less smeared out by wire-width fluctuations, interdiffusion or segregation as already discussed for quantum wells in the preceding section.

The quantization energy $E_{ni,nj}$ depends on the cross section, of the wire. As we shall see below, these cross sections may be rather complex.

The first attempts to create quantum wires was by lithographical and etching of quantum wells as shown in Fig. 8.34a.

For this rectangular cross section the confinement in two directions is the one of the well E_{nz}. In the y-direction the confinement comes from the transition to vacuum (air) or, if the structure is overgrown by some barrier material to reduce the density of surface states, by the transition to the overgrown barrier material. In this case, the confinement energies are simply additive, i.e.,

$$E_{ni,nj} = E_{nz} + E_{ny}\,. \tag{8.53}$$

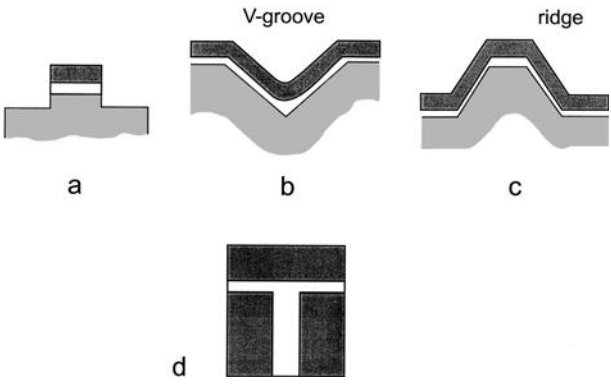


Fig. 8.34. Various possibilities to create quasi-one-dimensional quantum wires. Photolithography and etching of quantum wells (**a**) growth in V-grooves (**b**) on ridges (**c**) or T-shaped quantum wires occurring at the junction of two orthogonal quantum wells. In all cases the wire extends normal to the paper

A problem in these structures is that the width of the quantum well l_z typically has values between 3 and 10 nm while the best values, which can be reached by lithography and dry or wet etching are usually in the range of $20\,\text{nm} \leq l_y \leq 50\,\text{nm}$. Since the quantization energy varies roughly with l_i^{-2} one finds that usually

$$E_{nz} \gg E_{ny} \quad \text{for} \quad nz = ny\,. \tag{8.54}$$

Often it is already considered as a success and an argument for two-dimensional confinement if, e.g., the excitonic transition energies are blue-shifted by a few meV compared to the quantum well. We shall see examples in Sect. 15.3. Therefore many different alternatives to form quantum wires have been invented. We show three of them in Fig. 8.34b–d.

By means of lithography and selective etching it is possible to create pre-patterned substrates with either V grooves (Fig. 8.34b) or ridges (Fig. 8.34c). If a quantum well is now grown on these pre-patterned substrates, generally by MOCVD on the V-grooves and by MBE on the ridges, this quantum well is thicker on the bottom of the V-groove or on the top surface of the ridge, than on the side flanks of the structures or on the flat parts of the substrate. Since the confinement energy increases quadratically with decreasing well width this effect produces a confinement potential along the curved quantum well in addition to the potential step normal to it. Both confining potentials produce together the confinement in two, here usually not orthogonal, directions and leaving a free motion of the carriers only in x-direction.

In the case of V-grooves it is possible to stack several identical quantum wires on top of each other leading to quantum wire superlattices. For more details see, e.g., [99B2, 00B1, 02K1].

The last method to produce wires that we want to discuss in some detail is shown in Fig. 8.34d. One first grows a quantum well. The sample is cleaved in ultra high vacuum normal to this quantum well and another quantum well is grown on this cleaved surface followed by barrier material. This technique is known as cleaved-edge overgrowth. In the region, where the two quantum wells meet, the "average width" of the confining potential is wider than in each of the two quantum wells. This again produces a confining potential in two directionns. If one starts from a multiple quantum well or a superlattice one obtains by the cleaved edge overgrowth an array of uncoupled or coupled quantum wires, respectively.

We mention briefly that on highly indexed planes, like [311], corrugated layers can be grown that have periodically modulated thickness. Actually it has been found that one interface is intermixed and the other is corrugated. For more details of these structures, which were also known as quantum well wire superlattices (QWWSL) and are now generally called corrugated superlattices see, e.g., [98L1] or [96K1, 01L1] of Chap. 1.

To conclude this section we mention the increasing interest in growing thin columnar crystals, so-called nanorods, with several μm length and a few tens

of nm thickness preferentially from ZnO [01H1, 01P1, 03L1, 03Z1, 04Z1] (see Sect 15.3).

8.13 Quantum Dots

The ultimate limit in reducing the quasi-dimension is confinement in all three directions of space leaving no more possibility for free propagation of the carriers. The resulting quasi zero-dimensional structures are known as quantum dots, quantum boxes, artificial atoms and nano crystals.

The energy levels in a single quantum dot are discrete (see Fig. 8.20), i.e.,

$$E = E_{nk,nj,nk} \tag{8.55}$$

like in atoms below the ionization threshold. If a sample contains many quantum dots one usually has some size fluctuations, which lead to an inhomogeneous broadening of the energy spectrum. This statement is generally true, independent of how these quantum dots have been formed.

In the following we give together with Fig. 8.35 some of the many possibilities to produce quantum dots (QD). For more details, including various growth techniques, see, e.g., [88F1, 89H1, 89P1, 90B1, 90P1, 91B1, 91E1, 92H1, 93D1, 93E1, 93M1, 93P1, 93R1, 93W1, 94I1, 95G1, 95W1, 97H3, 99B1, 99G1, 01S1, 03K1] or [93B1, 96K1, 98G1, 98J1, 01L1] of Chap. 1 and references given therein.

The first possibility we mention is again the micro- or nanostructuring of a quantum well by lithography and etching as shown schematically in Fig. 8.35a. As for the wires, the lateral dimensions of the dots are usually considerably larger than the thickness of the well, leading to pancake-shaped dots, with the consequence that the quantization energy from lateral confinement is usually small compared to the one from the parent quantum well itselfs, (see e.g. [93D1, 93P1, 94I1, 95G1]).

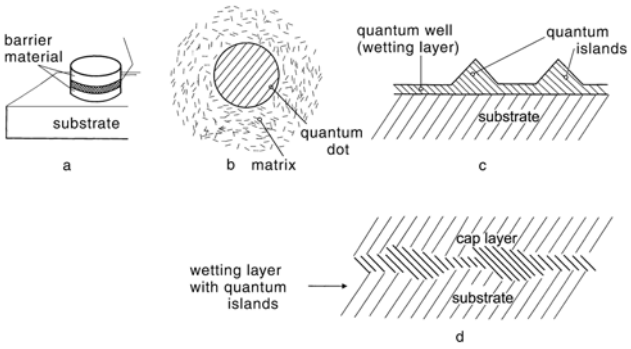


Fig. 8.35. Various possibilities to form quantum dots by microstructuring of quantum wells (**a**), by precipitation in glasses or organic solutions (**b**) and by "self-assembly" (**c**, **d**) (schematic)

The other possibility is to grow QDs by an annealing procedure in glasses doped with the constituents of a semiconductor. This technique is widely used for II–VI compounds such as $CdS_{1-x}Se_x$ including the cases $x = 0$ and $x = 1$, for CdTe or for the Cu halides. The commercially available glass edge filters are made in this way. The average diameter of the dots increases with annealing time and temperature. We show schematically in Fig. 8.35b a high-resolution transmission electron (HRTEM) micrograph of a single dot in a silicate glass matrix. The lattice planes within the almost spherical nano crystallite are already visible [95W1].

Other techniques involve chemical precipitation of, for example, CdS in (organic) liquids or gels [88F1,89H1,89P1,90B1,91B1,91E1,93E1,93M1,93R1]. These techniques have the advantage of giving dots that are almost spherical and have a high barrier, but the disadvantage is that one always gets a distribution of the dot radii with a certain width that is rarely below 10%. Since the quantization energy depends on the radius, this gives one contribution to the inhomogeneous broadening of the ideal δ-function density of states. Still another possibility is to grow the dots in the voids of zeolithe crystals or of synthetic opals [99G1].

A topic that has recently gained interest is "self-assembled" quantum dots or thickness fluctuations occurring, e.g., for $Ga_{1-y}In_y$ as wells under suitable conditions.

As we have seen in Sect. 8.11 there is, in strongly lattice misfit systems (strong means a lattice misfit beyond about 5%), apart from the formation of misfit dislocations, the possibility of so-called Stranski-Krastanov growth mode, where small hillocks form on a wetting layer. In theses hillocks or small pyramids the growing layer relaxes towards its own lattice constant. These pyramids have been observed, e.g., by scanning microscopy (see Fig. 8.35c) [97W1, 99B1, 00K1, 00R1, 01S1] and a lot of effort has been invested in calculating the electron and hole eigenstates and energies in these structures including the spatially varying strain.

The preferential systems in which these so-called self-organized quantum dots or better quantum islands (because of their rather flat island-like shape) are investigated are $In_{1-y}Ga_yAs/GaAs$ including $y = 0$ and CdSe/ZnSe.

Then there come some drawbacks. It has been found by HRTEM investigations in both systems that during overgrowth a substantial intermixing between well and carrier material frequently occurs, with the consequence that the quantum islands consist in most cases of more pronounced thickness or composition fluctuations than well-defined pyramids after overgrowth, where they are no longer directly accessible to scanning microscopy. We give an example in Fig. 8.35d. Furthermore it has been found that in some cases the pyramids even do not have the expected chemical composition. On ZnSe these pyramids formed, e.g., in air even in the absence of CdSe [97H1, 97R1] and in the CdSe/CdS system scanning microscopy revealed also islands, which are more likely some Se precipitates [89G1]. However, recently it became pos-

sible to create quantum islands, which contain, at least in their center, almost pure InAs and CdSe, respectively.

Presently there is no clear criterion from where on a localization site in a quantum well caused by thickness or composition fluctuations should be called a quantum island or a dot. It is more a case of personal taste, of sales-strategy when fundraising, of semantics or of the personality of the scientist, if he called a deep localization site a quantum dot or vice versa.

Maybe one can speak about islands (which, by the way, exist, e.g., in the CdSe/ZnSe system in different sizes [00K1]) if one tries to create them intentionally and about localization sites if they occur unintentionally.

In spite of these draw backs there has also been some progress in the field. Especially it becomes possible to stack these quantum islands on top of each other if the barrier layer in multiple quantum well (or island) systems was thin enough that the $(n+1)$th layer feels the lattice distortion of the islands in the nth layer below. For examples of the InAs/GaAs and the CdSe/ZnSe systems, see, e.g., [99B1, 00K1, 01S1].

To conclude this section we want to mention briefly some further possibilities of dot formation.

If the cleaved edge overgrowth mentioned in Sect. 8.12 is applied a second time in a way that three orthogonal quantum wells meet at one point, one has with the same arguments as above a potential minimum that may confine a carrier in all three dimensions of space.

Another possibility is so-called stressors. These are small areas of a non-lattice matched material grown on top of a thin barrier layer over a quantum well. The strain exerted by the stressors on the quantum well material must be in way to reduce the band gap in the well below the stressor and to create in this way a confining lateral potential.

A further method, which is used less in optical spectroscopy but frequently in transport experiments, is to define a dot by applying suitable electrostatic potentials to micro electrode structures deposited on the barrier material over a quantum well. Keywords are, e.g., Coulomb blockade, single electron tunneling, or electron turn style. For some reviews of these topics, which do not fall into the realm of this book see, e.g., [95H1, 95K1, 95P1, 97P1]. The same holds for quantum dot or antidot arrays [90M1, 93M1].

Still another method uses the growth on prepatterned surfaces containing instead of V-grooves or ridges, corner-cube holes or pyramids [97H3, 02K1].

Presently there is also some effort seen to arrange QDs in three dimensional lattices to create an artificial bulk material [97H1, 97R1, 99A1, 01A1].

This is one of the examples where history repeats itself to some extend. At the beginning of the century one started to understand the discrete energy level scheme of atoms, then of molecules, and finally one assembled the atoms to bulk crystalline solids resulting, e.g., in the band structure and the Ewald–Bloch theorem presented in Sects. 8.1 – 8.9. Then one started to reduce the dimensionality until one arrived at artificial atoms, which have again discrete energy levels caused, however, by a confining potential different from

the (screened) Coulomb potential of atoms. Now one starts again to assemble these artificial atoms to three-dimensional solids. One does not need much of a prophetic gift to predict that one will again find some type of band structure provided that the artificial atoms are sufficiently uniform in size and shape and provide enough electronic coupling between them. A similar approach is used in photonic crystals as will be detailed in Sects. 17.2 – 17.4.

8.14 Defects, Defect States and Doping

In connection with local phonon modes, we have already mentioned the imperfections which are present in all real crystals. Here we outline the electronic states connected with defects and concentrate again on point defect neglecting thus one dimensional defects like dislocations, two dimensional ones like stacking faults, grain boundaries, surfaces or interfaces or three dimensional ones like precipitates or voids. These are present, even in good materials, with densities of up to $10^{13} - 10^{17}\,\mathrm{cm}^{-3}$. Values of $10^{9} - 10^{13}\,\mathrm{cm}^{-3}$ which are given in literature, usually refer only to "electrically active" defects and this means the difference between the concentrations of ionized donors and acceptors (see below) and neglecting defects, which are electrically not active but which may act still as scattering or recombination centers. The density of dislocations ranges from zero to some 10^4 per cm^2 in high quality materials.

In Fig. 7.21 we classified the point defects according to the way in which they are incorporated in the lattice, e.g., as interstitials or -substitutionals. Now we consider their electronic properties and present donors, acceptors, isoelectronic traps and recombination centers in Fig. 8.36. A donor is a shallow center which has an energy level just below the conduction band and can easily give an electron, e.g., by thermal ionization, to this band

$$D^0 \leftrightarrows D^+ + \mathrm{e}\,. \tag{8.56a}$$

Donors are often formed by substitutional atoms situated in the periodic table one column to the right of the atom which they replace, like N or P in Ge or Si, Si on Ga site in GaAs, or Cl on Se sites in ZnSe, Ga or In on Zn site in Zn etc. Furthermore, donors can be formed by interstitials which have a weakly bound electron such as H, Li or Na, in a II–VI compound.

In analogy, acceptors can easily accommodate an electron from the valence band, i.e., they emit a hole into it

$$A^0 \leftrightarrows A^- + \mathrm{h}\,. \tag{8.56b}$$

Acceptors may be formed by substitutional atoms which have one electron less than the one which they replace. Thus these impurities are often found in the periodic table to the left of the atom which they replace, for example, Ga or B in Si and Ge, Li or Na on the cation site and N on the anion site in II–VI compounds, or Si on the As or Be on Ga site in GaAs.

It is clear from some of the above-mentioned examples that the same atom can act as donor or acceptor depending on the way it is introduced into the lattice. This is one possibility for self-compensation. Another arises from the fact that some vacancies or interstitials can act as donors or as acceptors or that the incorporation of an acceptor (donor) in one of the II–VI compounds, which are notoriously n-type (p-type) like ZnSe, CdS or ZnO (ZnTe) triggers in thermodynamic equilibrium rather the formation of another defect, which acts as a donor (acceptor) instead of shifting the Fermi-level close to the valence (conduction) band to produce p-type (n-type) conductivity [88W1, 95F1]. Actually the recent success of p-doping of ZnSe is based entirely on freezing in a non-equilibrium substitutional N occupation on Se sites created during growth. A material which has a higher concentration of donors (acceptors) is called n-type (p-type) and we recall that $np = n_\mathrm{i}^2(T)$ in thermodynamic equilibrium. The ability to choose the type and concentration of carriers over a wide range by doping with donors and/or acceptors is the basis for the widespread and important application of semiconductors in electronic devices like diodes, transistors, thyristors, etc. This topic is beyond the scope of this book and we refer the reader to text books on the physics of semiconductor devices, like [65S1, 81S1, 85P1, 86P1, 92E1] of Chap. 1.

A shallow donor (acceptor) can be considered as a positively (negatively) charged center to which an electron (hole) is bound by Coulomb interaction. So we are faced with a problem similar to that of a hydrogen atom, leading, in the simplest approximation, to a series of states with binding energy

$$E_\mathrm{b}^{\mathrm{D,A}} = \mathrm{Ry}\frac{m_{\mathrm{e,h}}}{m_0}\frac{1}{\varepsilon^2}\frac{1}{n_\mathrm{B}^2} \tag{8.57a}$$

where Ry is the Rydberg energy of the H atom (Ry = 13.6 eV), n_B the main quantum number and ε a dielectric constant. In Fig. 8.36 we show only the

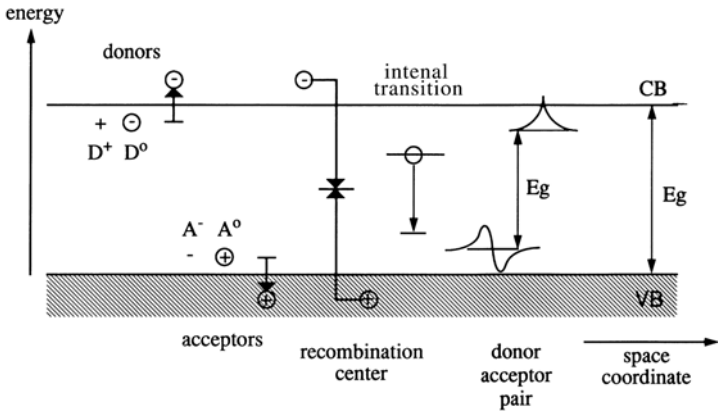


Fig. 8.36. Schematic drawing of various impurity levels in semiconductors

states for $n_\mathrm{B} = 1$. Depending on the material parameters, one usually finds values for the donor and acceptor binding energies of

$$5\,\mathrm{meV} \leq E_\mathrm{b}^\mathrm{D} \leq 50\,\mathrm{meV}\,, \tag{8.57b}$$

$$20\,\mathrm{meV} \leq E_\mathrm{b}^\mathrm{D} \leq 200\,\mathrm{meV},. \tag{8.57c}$$

For a listing of various donor and acceptor binding energies in various semiconductors, center see e.g. [81S1, 82L1] of Chap. 1.

The radius of the $n_\mathrm{B} = 1$ state is given by

$$a_\mathrm{lattice} < a_\mathrm{D,A} = a_\mathrm{B}\varepsilon\frac{m_0}{m_\mathrm{e,h}} \tag{8.57d}$$

ranging from 1 to 20 nm depending again on the material parameters. The fact that these values are larger than the lattice constant justifies the use of the effective mass approximation. Details concerning which value of ε has to be used are similar to those for excitons, as discussed in Sect. 9.2. In addition, there is some smaller influence of the chemical nature of the atom forming the donor or acceptor which is known as the central-cell correction or chemical shift.

Similar to the localized phonon modes, the wave function of a shallow donor (or acceptor) can be described as a superposition of Bloch states

$$\phi(\boldsymbol{r}) = \sum_k a_k \phi_k^\mathrm{CB}(\boldsymbol{r})\,. \tag{8.58}$$

The range of $\boldsymbol{k}$ from which significant contributions can be expected increases with decreasing radius in (8.57d).

The occupation probability of donors and/or acceptors in thermodynamic equilibrium is again governed by Fermi statistics. However, some correction terms appear

$$f_\mathrm{D/A} = \frac{1}{1 + 1/g_i \exp\left\{\left(E_\mathrm{D/A} \mp E_\mathrm{F}\right)/k_\mathrm{B}T\right\}} \tag{8.59}$$

where the g_i describe the degeneracy of the level and take care of the fact, that, e.g., an ionized donor D^+ offers two states for electrons with opposite spin in the ground state. However, if this state is occupied by one electron, the level for the second one is considerably blue shifted due to Coulomb repulsion. For details see, e.g., [87B1].

Pairs of donors and acceptors that are so close in space that their wavefunctions overlap are known as donor–acceptor pairs. As we shall see later in Sect. 14.2, they give rise to a characteristic emission feature.

Donors and acceptors exist in bulk semiconductors, but also in quantum structures, possibly with the complication that the binding energy, e.g., of an electron to a donor also depends on the position of this donor relative to the interfaces to the barrier material.

In connection with (M)QW two special methods of doping should be mentioned:

δ-doping means the introduction of a two-dimensional sheet of doping atoms during epitaxial growth. The concentration in the growth direction then has an almost δ-function-like profile.

Modulation doping means introducing the doping atoms into the barriers of a (M)QW. The electrons (holes) are thermally ionized into the conduction (valence) band of the barrier, reach the well by thermal diffusion, and are captured in it. This allows the production of high two-dimensional carrier densities in the well with high mobility, since the charged impurities are separated spatially from the mobile carriers, partly even by an additional undoped barrier layer between the doped barrier and the quantum well, thus reducing the scattering with them. The space charges of the ionized doping atoms and of the free carriers lead to a characteristic curvature of the bands which can be calculated by solving the Poisson equation

$$-\Delta\phi = \frac{\rho}{\varepsilon\varepsilon_0}\,, \tag{8.60}$$

where ϕ is the electrostatic potential and ρ the space charge as shown schematically in Fig. 8.37. Modulation doping is the basis of devices called high-electron-mobility transistors (HEMT) or modulation-doped field-effect transistors (MODFET). The binding energy of donors and acceptors in SL and (M)QW depends in addition to the parameters (8.57) on the distance from the barrier.

Apart from the shallow donor and acceptors there are deep donors and acceptors and a variety of other deep centers. These are atoms which have one or more energy levels somewhere around the middle of the gap. See e.g. [82L1] of Chap. 1 or [01H1].

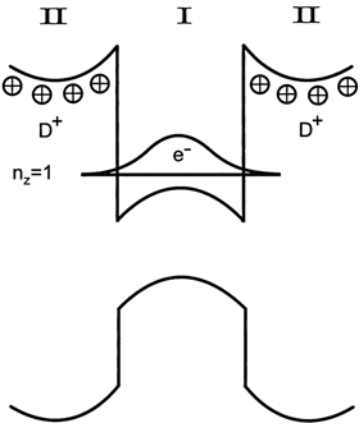


Fig. 8.37. The curvature of the band edges in the case of symmetric modulation doping of a MQW (schematic) including the first quantized electron level

For deep centers an approach as in (8.57) is not adequate. The wave function is better described by the parent atomic orbitals, modified by the influence of the surrounding atoms, i.e., by the symmetry of the arrangements of the neighbors. Copper, nickel, iron, chromium and other elements can give rise to such deep levels.

Some deep centers can exchange carriers with both the conduction and the valence band (in contrast to the donors and acceptors). In this case they are called recombination centers. The recombination can be radiative or non-radiative and some centers provide fast channels of for de-excitation of electron–hole pairs.

A group of (deep) centers is often formed by so-called isoelectronic traps. These are frequently atoms of the same column of the periodic table as the one which they substitute, i.e., they have the same electron configuration in the outer shell. An example of an isoelectronic trap would be Te replacing S or Se in ZnS or ZnSe.

Some centers have various levels in the forbidden gap and so transitions within the center can be investigated as in Cu, Ni or other ions. Since the chance of having various levels in the gap increases with increasing width of the gap, such internal transitions are best investigated in wide gap semiconductors and in insulators. Carriers in deep centers can sometimes couple strongly to bulk and/or localized phonon modes giving rise to very broad emission (and absorption) features with Huang-Rhys factors S significantly larger than 1. This topic however leads beyond the scope of this book and we refer the reader to [81A1]b,c,f of Chap. 1 and references therein.

To summarize this section we show schematically in Fig. 8.38 the density of states including some impurity centers, the occupation probability in thermodynamic equilibrium and the resulting density of electrons and holes per unit energy.

More information on defects in bulk materials and in MQW structures is given e.g. in [88P1, 91E2] and references therein.

8.15 Disordered Systems and Localization

In this section we will first outline briefly what is new in disordered systems and then we present some examples of how disorder can be realized in semiconductors. More exhaustive treatments of these topics, from which the following facts were largely extracted, are [79M1, 79Z1, 83Z1, 84S1, 85K1, 85S1, 86L1, 00A1] or [81A1]d of Chap. 1. Furthermore we mention the pioneering work [58A1].

In Fig. 8.39a we show schematically a periodic potential. It is known as the Kronig–Penney potential. Due to the tunneling of the wavefunction into the barriers, we get finite overlap integrals and consequently a band of Bloch states with a certain width B.

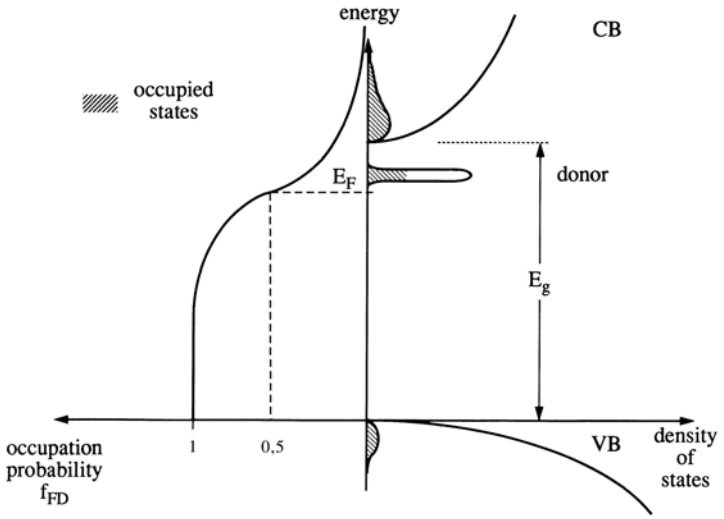


Fig. 8.38. The density of states DOS in the effective-mass approximation of an idealized three-dimensional semiconductor containing some impurities, here a donor state, and the resulting population (*hatched areas*). Usually the relation $k_B T \ll E_g$ holds and the Fermi-function looks more step-like with the consequence that the populations in conduction and valence bands cannot be shown simultaneously on the same linear scale

We now want to introduce disorder into this system. This can be done by varying the depths of the potentials statistically within a width V_0 ("diagonal disorder") or by varying the widths of the potential wells and barriers and thus the coupling ("off-diagonal disorder"). The first case leads to the so-called Anderson model, the second to the Lifshitz model. In practice, both aspects of disorder will occur simultaneously, but from the theoretical point of view it is sufficiently difficult to treat one of them. Figures 8.39 and 8.40 briefly outline the ideas of the Anderson model. If some diagonal disorder is introduced, as in Fig. 8.39c, two things happen – the sharp edges of the density of states are smeared out by exponential tails, and a new type of eigenstate appears, namely localized states. If

$$BV_0^{-1} > 1\,, \tag{8.61a}$$

then there are both localized states at the band tails and extended states in the center, which, however, are different from Bloch states, as we shall see shortly (Figs. 8.39d and 8.40b,c). For

$$BV_0^{-1} < 1\,, \tag{8.61b}$$

there are only localized states in the band (Figs. 8.39c and 8.40c).

To explain the concept of extended and localized states we show in Fig. 8.40a a Bloch wave which is a superposition of atomic orbitals ϕ_n placed

at every lattice site $\boldsymbol{R}_i$ with a proper long-range phase correlation, corresponding to the LCAO method mentioned in Sect. 8.7:

$$\phi_{\boldsymbol{k}}^{\text{Bloch}}(\boldsymbol{r}) = \sum_i \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i}\phi_n(\boldsymbol{r}-\boldsymbol{R}_i)$$
$$= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\sum_i \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{R}_i)}\phi_n(\boldsymbol{r}-\boldsymbol{R}_i) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}u_k(\boldsymbol{r})\,. \qquad (8.62a)$$

In an extended state in a disordered system (Fig. 8.40b) we lose the long-range phase correlation resulting in

$$\phi^{\text{ext}}(\boldsymbol{r}) = \sum_i c_i\phi_n(\boldsymbol{r}-\boldsymbol{R}_i) \qquad (8.62b)$$

and for a localized state we obtain an envelope which is centered around a localization site $\boldsymbol{r}_0$ and which decays exponentially with a localization length ξ (Fig. 8.40c)

$$\phi^{\text{loc}}(\boldsymbol{r}) = \sum_i d_i\phi_n(\boldsymbol{r}-\boldsymbol{R}_i)\,\mathrm{e}^{-|\boldsymbol{r}-\boldsymbol{r}_0|/\xi}\,. \qquad (8.62c)$$

The energy which separates extended from localized states is called the mobility edge. If the carriers were to have an infinite phase-relaxation time T_2 in their states and thus also an infinite lifetime, it is easy to show that the

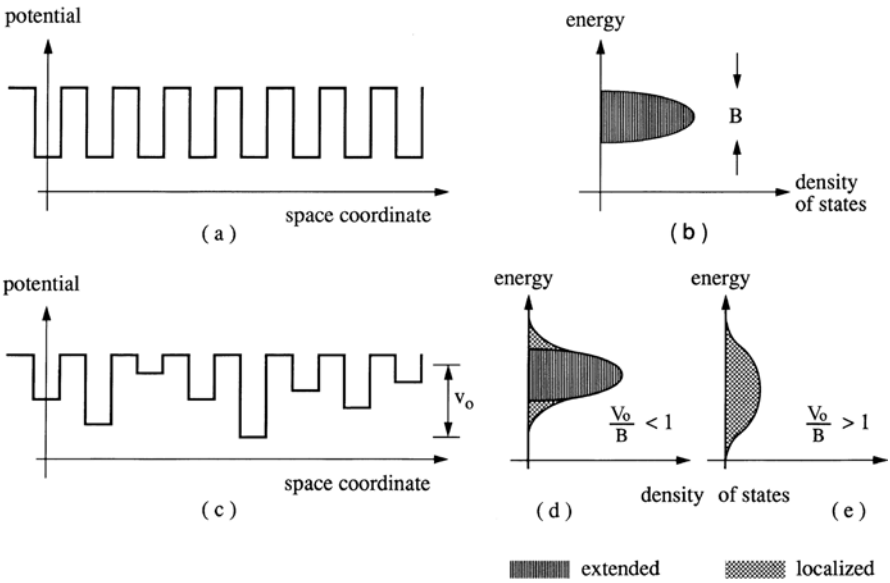


Fig. 8.39. A periodic potential (**a**) and the resulting band, represented by its density of states (**b**), a disordered system (**c**) and bands with localized and delocalized states (**d**) and with localized states only (**e**)

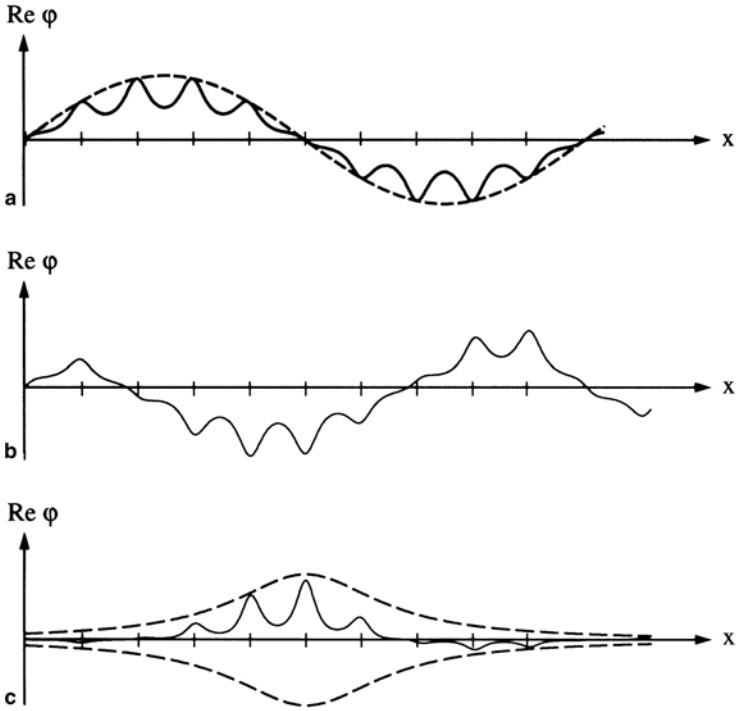


Fig. 8.40. Schematic drawing of the real part of the wave function for a periodic potential (**a**) and for a disordered system an extended state (**b**) and a localized one (**c**)

mobility edge is sharp: if extended and localized states were to coexist in a certain energy range, then the slightest perturbation would mix these states and consequently every localized state would obtain some admixture of an extended wave function (and vice versa) transforming it to an extended state. Since the T_2 times of the carriers in their states are limited, the mobility edge is usually a transition region of a certain width.

Another important consequence of disorder is the following. A disordered system is no longer invariant with respect to translations. As a consequence, the wave vector is no longer a good quantum number and the dispersion relation $E(\boldsymbol{k})$ loses, strictly speaking, its meaning. A disordered system is in principle characterized only by its density of states $D(E)$.

As usual, there are exceptions to this rule: excitations with a wavelength λ long compared to the typical length scale of the disorder fluctuations "average" over the disorder and can be characterized by a wave vector $\boldsymbol{k}$ in the sense of a continuum approximation. This is fulfilled for example for long wavelength (acoustic) phonons or photons. The latter case is easily checked by looking through a glass window or into clear water. See also Sect. 5.3.2.

Weak disorder, e.g., as defined by (8.61a), will produce some localized states at the band edges ($10^{16} - 10^{18}\,\mathrm{cm}^{-3}$) but the extended states in the band ($10^{22} - 10^{23}\,\mathrm{cm}^{-3}$) will be close to Bloch-type waves. See Fig. 8.40.

A crucial property for localization effects is the dimensionality of the system. It has been shown by general arguments that the slightest disorder will in principle localize all states for dimensions $d \leq 2$. However the localization length [ξ in (8.62c)] can be extremely long and in many cases exceeds the length of the sample used. As a consequence these considerations are largely of theoretical interest and only to a limited extent of practical relevance.

After characterizing briefly some properties of disordered systems we now proceed to the inspection of realizations of disorder in semiconductors. We discuss, in roughly increasing magnitude of disorder, heavily doped semiconductors, alloy semiconductors, well-width fluctuations, and amorphous semiconductors.

In Fig. 8.41a, we show the density of states for the conduction band of a semiconductor containing donors of variable concentration. If the concentration of donors is so low that the wavefunctions do not overlap, i.e.,

$$a_{\mathrm{B}}^3 N_{\mathrm{D}} \ll 1\,, \tag{8.63}$$

where N_{D} is the concentration of donors and a_{B} the radius of the wave function according to (8.57d), then their density of states is δ-function-like. We show in Fig. 8.41a, for simplicity, for the donor with one electron only the state with $n_{\mathrm{B}} = 1$. A donor may bind a second electron with opposite spin. As a result of the Coulomb interaction the energy for the two electrons will be higher. This is indicated by the second peak. The states with $n_{\mathrm{B}} > 1$, which the electron could also occupy, are again neglected.

If N_{D} is increased, the wave functions of the donors start to overlap and this again results in the formation of a so-called impurity band. Its width increases with N_{D} and eventually merges with the conduction band, forming an (exponential) tail of the density of states. Depending on the doping and possibly on compensation through acceptors, the Fermi level can be situated for $T \to 0$ above or below the mobility edge. In the first case, the system is called a metal, since it has finite conductivity at $T \to 0$, in the second it is an insulator. The transition from an insulator to a metal with increasing doping is also known as a Mott transition and there was a long discussion as to whether this transition is continuous or not, i.e., whether there is something like minimum metallic conductivity in the second case. Experiments, e.g., with phosphorous-doped silicon (Si:P), indicate that there is no minimum conductivity, [84S1, 85K1, 86L1].

The next case of disorder are alloy semiconductors like $\mathrm{Ga}_{1-y}\mathrm{Al}_y\mathrm{As}$, $\mathrm{CdS}_{1-x}\mathrm{Se}_x$, $\mathrm{ZnSe}_{1-x}\,\mathrm{Te}_x$ or $\mathrm{Zn}_{1-y}\mathrm{Cd}_y\mathrm{S}$. Many, but not all of them can all be grown for every value of x or y between 0 and 1, others like $\mathrm{Ga}_{1-y}\mathrm{In}_y\mathrm{N}$ or $\mathrm{GaAs}_{1-x}\mathrm{N}_x$ have a miscibility gap. They usually have a well-defined crystal structure (e.g., of the zincblende or of the wurtzite type). The disorder is in-

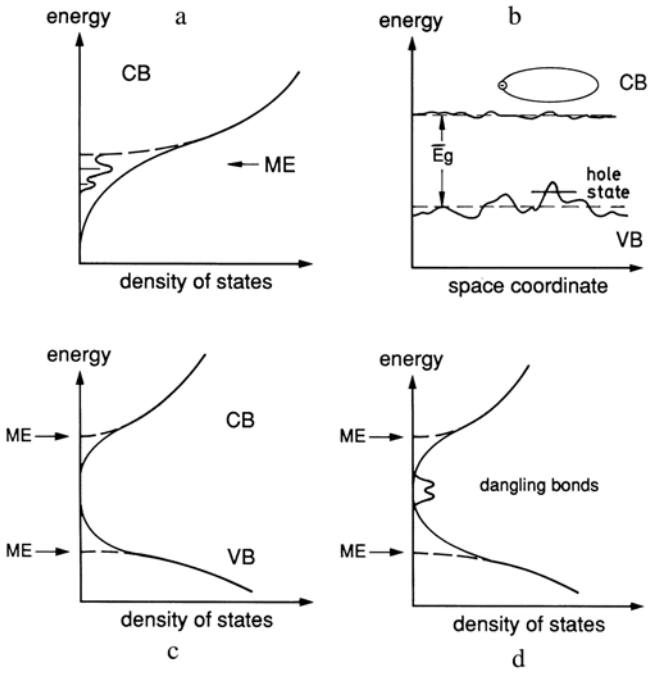


Fig. 8.41. The appearance of tails of localized states due to increasing doping (**a**), potential fluctuations, e.g., in an alloy (**b**, **c**) or in amorphous semiconductors (**d**)

troduced by the statistically distributed occupation of anion (or cation) sites with the two different atoms.

Though an alloy has no translational invariance one can create it artificially in the so-called virtual crystal approach. One assumes that the unit cell contains atoms with a mixture of the properties according to x or y. This approach explains for example the often observed continuous shift of E_g with composition (Fig. 8.31). This shift can be described by

$$E_g(x) = E_g(x=0) + [E_g(x=1) - E_g(x=0)]\left[(1-b)x + bx^2\right] \qquad (8.64)$$

where the usually small "bowing" or "bending" parameter b describes the deviation from a linear relation. In some cases, like in $In_{1-y}Ga_yAs_{1-x}N_x$ the bowing parameter is so large, that the system goes with increasing x through a band gap minimum which is situated considerably below that of $x = 0$ or 1.

The fluctuations now arise from the fact that the composition has fluctuations on a microscopic scale (e.g., the exciton radius, see below). In $CdS_{1-x}Se_x$, for example, there is in some places a little more S, in others more Se. This compositional disorder results frequently in the more ionic systems in preferential fluctuation of the valence band for anion substituted materials like $CdS_{1-x}Se_x$, or of the conduction band for cation-substituted

materials like $Zn_{1-y}Cd_yS$. In the more covalent materials like $Al_{1-y}Ga_yAs$ both bands fluctuate with composition. These fluctuations can again lead to localized states. The "tailing parameter" E_0, which describes the exponential tail in

$$D(E) = \frac{N_0}{E_0} \mathrm{e}^{-E/E_0} , \tag{8.65}$$

where N_0 is the total number of tail states, depends on the material parameters and on the composition. It is obvious that E_0 disappears for $x, y = 1$ and $x, y = 0$, that it increases for increasing $E_g(x = 1) - E_g(x = 0)$ and for decreasing effective mass of the particle which is localized. The latter point depends on the fact that in three dimensions a potential well must have a certain width R and depth V to localize or bind a particle of mass in m with

$$VR^2 \geq \frac{\hbar^2}{m_{\mathrm{e,h}}} \tag{8.66}$$

and explains why holes are more easily localized than electrons.

For the dependence of E_0 on x one finds different but similar formulas in literature. We give here one according to [78B1]. For more recent discussions of this topic concerning also excitons see [78B1,84S1,91K1,92P1,99K1,00P1].

$$E_0 = \frac{1}{178}\beta^4 x^2 (1-x)^2 m_{\mathrm{e,h}}^3 \left(\hbar^6 N^2\right)^{-1} , \tag{8.67}$$

where N is the density of atoms and β the derivative of the position of the edge of the band with respect to x. Often one finds from experiment a slightly asymmetric behavior of $E_0(x)$ caused by the term β and other effects.

We now consider disorder in systems of reduced dimensionality like MQW. Actually there are two origins of disorder. One is the alloy "broadening" or disorder felt by the carriers if the well or the barrier or both are made from an alloy or if an alloy forms at the boundary due to interdiffusion or segregation. Another is due to well-width fluctuations. Though great progress has been made in growing atomically flat interfaces by modern epitaxial techniques, fluctuations of l_z of at least the order of one atomic (or molecular) layer are hardly avoidable. Due to the dependence of the quantization energy on l_z, carriers (and excitons) can be localized or trapped at regions with larger l_z.

The continuous transition from these localization sites to so-called assembled quantum islands has already been mentioned in Sects. 8.10 and 8.13. For examples of localization in two dimensions see, e.g., [85K1, 86L1, 00P1] and references therein.

The strongest disorder occurs in amorphous semiconductors, the prominent example being amorphous or α-Si. In these systems only short-range order remains (e.g., a tetrahedral coordination of most but not of all Si atoms) but no long-range order at all caused by the facts that the coordination of some atoms is different from the one in a perfect crystal and that the bond lengths and the angles between the bonds show some fluctuations. If we start

in a $CdS_{1-x}Se_x$ crystal at a S atom and proceed by a lattice vector $\boldsymbol{R}$ we will hit an atom. We are not sure, however, whether it will be S or Se. In an α-Si there is only an average bond length but nothing like a lattice "constant" and if we move from a Si atom by a vector $\boldsymbol{R}$ of crystalline Si we are not even sure whether we will find an atom at all.

To illustrate this concept, we show in Fig. 8.42 a hypothetic two-dimensional crystal with hexagonal structure and threefold coordinaton and in (b) the corresponding amorphous material in which some five and seven atom rings occur and also some dangling bonds.

This strong disorder leads to substantial exponential tails of the densities of states both for conduction and valence bands, which cover the whole "forbidden gap" (Fig. 8.41d). Unsaturated covalent bonds in Si, so-called "dangling" bonds, form states situated energetically close to the center of the gap (Fig. 8.41d) since it is undetermined whether they would form a bonding or an antibonding orbital with the next Si. The states of the dangling bonds can be saturated by hydrogen doping, opening interesting technological applications for α-Si:H. A short selection of books and conference proceedings for the properties of amorphous semiconductors is [74A1, 84H1, 84T1].

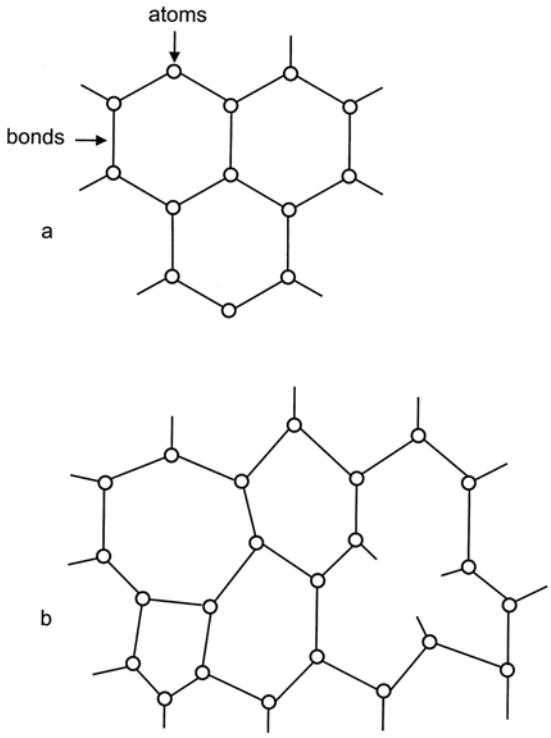


Fig. 8.42. A hypothetical two-dimensional crystalline (**a**) and amorphous (**b**) material

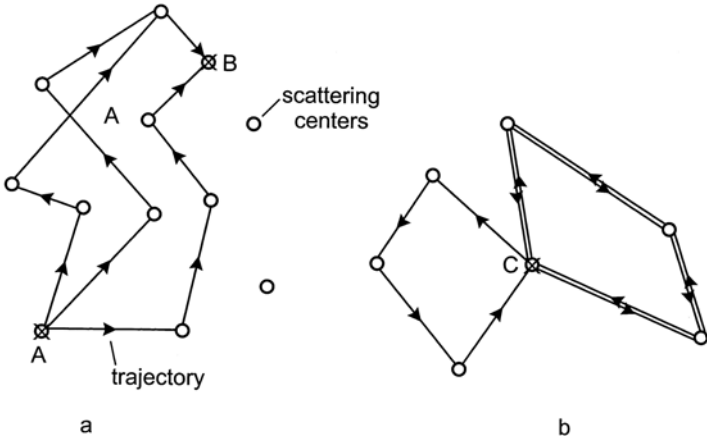


Fig. 8.43. Various possibilities for a particle to come from A to B or to be backscattered from C to C. The *open dots* symbolize scattering centers

To conclude this section, we now outline the concepts of weak localisation by enhanced backscattering and of percolation. We come back to these approaches in Sect. 14.4.

In the first case we consider an otherwise perfect crystal which contains randomly distributed elastic scattering centers. See Fig. 8.43.

To obtain the probability $w_{\mathrm{A}\to\mathrm{B}}$that a particle, e.g., a carrier starting from A reaches B we have to sum over the amplitudes $a_i^{\mathrm{A}\to\mathrm{B}}$ of all possible ways to come from A to B and to square the sum

$$w_{\mathrm{A}\to\mathrm{B}} \sim \left|\sum_i a_i^{\mathrm{A}\to\mathrm{B}}\right|^2 = \sum \left|a_i^{\mathrm{A}\to\mathrm{B}}\right|^2 + \sum_{i,j} a_i^{\mathrm{A}\to\mathrm{B}} a_j^{\mathrm{A}\to\mathrm{B}} \approx \sum \left|a_i^{\mathrm{A}\to\mathrm{B}}\right|^2 . \tag{8.68a}$$

The mixed terms all have different relative phases and cancel essentially as a sum. In the case of backscattering we have to start in the same way, i.e.,

$$w_{\mathrm{C}\to\mathrm{C}} = \left|\sum_i a_i^{\mathrm{C}\to\mathrm{C}}\right|^2 = \sum_i \left|a_i^{\mathrm{C}\to\mathrm{C}}\right|^2 + \sum_{i,j} a_i^{\mathrm{C}\to\mathrm{C}} a_j^{\mathrm{C}\to\mathrm{C}} . \tag{8.68b}$$

What is different from (8.68a) is that there not only remains the term $\left|\sum_i a_i^{\mathrm{C}\to\mathrm{C}}\right|^2$ but also those of the mixed terms, which are made up from $a_i^{\mathrm{C}\to\mathrm{C}}$ and their time-reversed parts shown for one case in Fig. 8.43, since the total phase shift encountered in both directions and assuming, that only *elastic* scattering occurs, is equal. As a consequence we obtain

$$w_{\mathrm{C}\to\mathrm{C}} = 2\left|\sum_i a_i^{\mathrm{C}\to\mathrm{C}} .\right|^2 \tag{8.68c}$$

This result means that backscattering is by a factor of two more probable than scattering from A $\rightarrow$ B $\neq$ A.

It should be noted that similar arguments can also be formulated in $\boldsymbol{k}$-space. Due to elastic scattering only the length of $\boldsymbol{k}$ remains constant and the scattering from an initial $\boldsymbol{k}_i$ to $\boldsymbol{k}_f = -\boldsymbol{k}_i$ is again by a factor of two more probable that scattering to any other $\boldsymbol{k}_f \neq \boldsymbol{k}_i$.

This enhanced backscattering is obviously a beginning of localization. The concept has been verified for scattering of photons from small polysysteme spheres [85A1, 85W1, 86A1, 86S1] and in electric conductivity. In the latter case there are also considerable contributions from inelastic carrier scattering to the resistivity. The part due to enhanced backscattering can be identified by applying a magnetic field $\boldsymbol{B}$, since the simple time reversal argument is no longer valid for $\boldsymbol{B} \neq 0$. For more details see, e.g., [84B1].

The concept of percolation will be explained with a simple-cubic lattice in Fig. 8.44.

Some of the unit cells are randomly painted black. All neighboring cells that are painted black are said to belong to one cluster.

We see in Fig. 8.44 clusters consisting of one, two, three and seven unit cells. With increasing occupation probability (i.e., increasing number of black unit cells) of the average cluster size growths. Eventually a cluster appears that extends from $-\infty \rightarrow +\infty$. This is the percolating cluster that corresponds to an extended state, while finite clusters are "localized". The occupation probability at which the first percolating cluster appears is the critical occupation f_c.

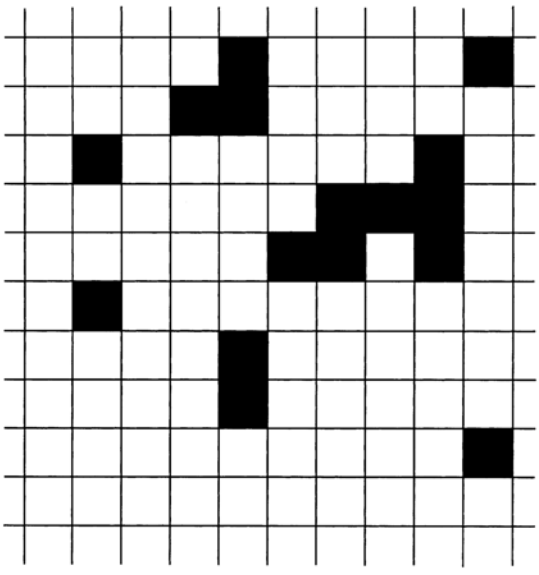


Fig. 8.44. Cluster formation in a simple cubic lattice in two dimensions

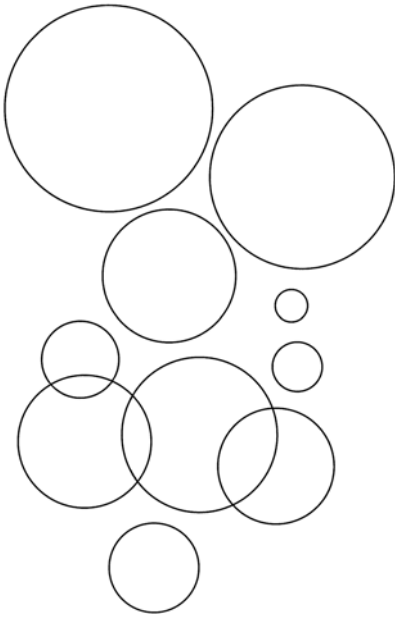


Fig. 8.45. Clusters of overlapping spheres

It is easy to understand that f_c depends on the dimensionality (in a one-dimensional case f_c is trivially equal to unity) on the type of lattice e.g. cubic, hexagonal, or triangular and on the fact if we consider site or bond percolation. For details see, e.g., [85S1].

One possibility to make the percolation model more realistic is e.g. the following one, which has been used with some more sophistication to successfully explain the optical properties of excitons in alloys like $CdS_{1-x}Se_x$ (see [99K1] and Sect. 9.6).

We assume that a localized state like the hole state in Fig. 8.41b is caused by a cluster of Se atoms in $CdS_{1-x}Se_x$ and that the wave function localized at this Se cluster has a spherical shape as indicated in a two-dimensional analogue (which corresponds to quantum wells [00P1]) in Fig. 8.45.

Wave functions that overlap spatially form a cluster. At a certain density of spheres a percolating cluster will appear that corresponds to an extended state while finite clusters again correspond to localized ones. The density and the size of the spheres can be influenced, e.g., by varying the energy in the tail of the density of states. The appearance of a percolating cluster with increasing energy corresponds then to the mobility edge.

An experiment showing localization in a disordered chain of macroscopic, coupled pendula has been described and verified experimentally in [89K1].

We will end the discussion of the properties of electronic states in semiconductors here and concentrate in the next section on electron–hole pairs, which are actually the relevant complexes for optical properties, but which cannot be understood without knowledge of the properties of electrons and holes.

8.16 Problems

1. Try to find the highest and lowest values for effective electron and hole masses in group-IV materials and in binary III–V, II–VI and I–VII semiconductors. (Use a compilation of semiconductor data from the library). What is the trend of the dependence of the effective masses on E_g?
2. Calculate the effective density of states, i.e., the onset of degeneracy, for electrons and holes at 10 K and at room temperature in bulk GaAs and ZnSe.
3. Calculate the effective density of states per unit volume at 300 K for electrons and holes in bulk GaAs and the density of states per unit area in a GaAs QW. Calculate the corresponding density per volume for $l_z = 10\,\text{nm}$.
4. Calculate the density of states and the effective densities of states for a nondegenerate three-, two- or one-dimensional electron gas in the effective mass approximation.
5. What is the minimum power consumption per unit area required to keep a degenerate electron and hole population at room temperature in ZnSe and in GaAs for a lifetime of 0.3 ns and layers with $l_z = 1\,\mu\text{m}$, $0.2\,\mu\text{m}$, and 10 nm thickness? (The first two cases can be considered as bulk material the third one as quasi-two-dimensional).
6. Calculate the binding energy of electrons (holes) to donors (acceptors) for some of the materials mentioned in the first problem. Compare with experimental data such as that given in [82L1] of Chap. 1. Deduce the order of magnitude of the central cell correction. Check the ratio of the binding energy and of the energy of the LO phonon, and try to anticipate the consequence for the choice of ε.
7. Calculate the Bohr radii of acceptors and donors for the extreme values found in the first problem. How many unit cells or atoms are contained in the volume of a donor?
8. Make a sketch of surfaces of constant energy in a two- (or even three-) dimensional $\boldsymbol{k}$-space for spherical, simple cubic, and hexagonal (plane $\boldsymbol{k}\perp\boldsymbol{c}$) symmetries. Is spherical symmetry compatible with cubic and hexagonal symmetry?
9. Show that the wave vectors for which the scattered waves interfere constructively in one- and two-dimensional square lattices in the concept of nearly free electrons are just the borders of the Brillouin zones.
10. Identify the various branches of the 2d electron band structure of Fig. 8.11 in the reduced zone scheme with the corresponding ones in the extended scheme and the $\boldsymbol{G}$-vectors that have been used to shift them into the first Brillouin zone.
11. Verify that (8.23a) fulfils the Bloch criterion, i.e., that the wave function is transformed onto itself under a translation $\boldsymbol{R}_i$ apart from a phase factor.

12. Calculate the positions of the first three quantized electron levels for a GaAs quantum well assuming infinitely high barriers and the realistic band discontinuity to $Al_{0.40}Ga_{0.60}As$ and compare them.
13. Make a sketch and calculate approximately the electrostatic potential assuming a single QW (width 10 nm) with an n-type modulation doping of $10^{17}\,cm^{-3}$ on both sides in a layer of 20 nm thickness and buffer layers of 10 nm ($m_e = 0.1m_0$) and assuming that all donor electrons are collected by the QW. What is the minimum height of the conduction band offset necessary to avoid substantial parallel conductivity of well and barrier?
14. Occupy a one-, two- and three-dimensional simple cubic lattice with six places in every dimension by throwing dies. Indicate the percolation threshold for this finite lattice and compare with data in literature.
15. What is a dislocation? Which types of dislocations do you know or can you find in textbooks?

References to Chap. 8

[47L1] F.C. von der Lage and H.A. Bethe, Phys. Rev. **71**, 612 (1947)
[57K1] E.O. Kane, J. Phys. Chem. Sol. **1**, 249 (1957)
[58A1] P.W. Anderson, Phys. Rev. **109**, 1492 (1958)
[67V1] Y. Varshni, Physica **34**, 149 (1967)
[69R1] U. Rössler, Phys. Rev. **184**, 733 (1969)
[70M1] O. Madelung, Grundlagen der Halbleiterphysik, Heidelberger Taschenbücher **71**, Springer, Berlin (1970)
[74A1] Amorphous and Liquid Semiconductors, J. Stuke and W. Brenig (eds.) Taylor and Trancis, London (1974) and (1990)
[76C1] M.L. Cohen and J.R. Chelikowsky, Phys. Rev. B **14**, 556 (1976)
[77P1] J. Pollmann and H. Büttner, Phys. Rev. B **16**, 4480 (1977)
[78B1] S.D. Baranovskii and A.L. Efros, Sov. Phys. Semicond. **12**, 1328 (1978)
[79M1] N.F. Mott and E.A. Davies, Electronic Processes in Non-Crystalline Materials, 2nd edn., Clarendon, Oxford (1979)
[79T1] H.R. Trebin, phys. stat. sol. (b) **92**, 601 (1979)
[79Z1] J.M. Ziman, Models of Disorder, Cambridge Univ. Press, Cambridge (1979)
[80M1] M. Matsuura and H. Büttner, Solid State Commun. **36**, 81 (1980)
[80P1] B. Pertzsch and U. Rössler, phys. stat. sol. (b) **101**, 197 (1980)
[82B1] G. Blattner, et al., Phys. Rev. B **25**, 7413 (1982)
[82P1] This topic can be nicely followed since 1990 in the Proceeding of the International Conferences on II–VI compounds, published so far in J. Crystal Growth **59** (1982), **72** (1985), **86** (1988), **101** (1990), **117** (1992), **138** (1994), **159** (1996), **184/185** (1998), **214/215** (2000), in phys. stat. sol b **229** (2002), and phys. stat. sol. c **1** Issue 4 (2004)
[83Z1] R. Zallen, Physics of Amorphous Solids, Wiley, New York (1983)
[84B1] G. Bergmann, Physics Reports **107**, 1 (1984)
[84H1] W. Heywang, Amorphe and Polykristalline Halbleiter, Springer, Berlin (1984)

[84S1] B.I. Shklovskii and A.L. Efros, Electronic Properties of Doped Semiconductors, Springer Ser. Solid-State Sci. **45** Springer, Berlin, Heidelberg (1984)
[84T1] The Physics of Hydrogenated Amorphous Silicon, J.D. Jonnopoulos (ed.) Springer, Berlin (1984)
[85A1] M.P. Albada and A. Lagendijk, Phys. Rev. Lett. **55**, 2692 (1985)
[85K1] B. Kramer, G. Bergmann and Y. Bruynserade (eds.), Localization, Interaction and Transport Phenomena, Springer Ser. Solid-State Sci. **61**, Springer, Berlin, Heidelberg (1985)
[85S1] D. Stauffer, Introduction to Percolation Theory, Taylor and Francis, London (1985)
[85W1] P.E. Wolf and G. Maret, Phys. Rev. Lett. **55**, 26969 (1985)
[86A1] E. Akkermans, P.W. Wolf and R. Maynard, Phys. Rev. Lett. **56**, 1471 (1986)
[86B1] G. Bastard, J.A. Brum, IEEE J. QE-**22**, 1625 (1986)
[86D1] G.H. Döhler, IEEE J. QE **22**, 1682 (1986)
[86L1] Localization and Interaction, D.M. Finlayson (ed.), SUSSP, Edinburgh (1986)
[86S1] M. Stephen, Phys. Rev. Lett. **56**, 1809 (1986)
[87B1] J.S. Blakemore, Semiconductor Statistics, Dover Publication, Dover (1987)
[87P1] Physics and Applications of Quantum Wells and Superlattices, E.E. Mendez, K. v. Klitzing (eds.), NATO ASI Ser. **170**, Plenum Press, New York (1987)
[87W1] C. Weisbuch, in Semicond. Semimet. Vol. **24**, Academic Press, London (1987)
[88C1] M.L. Cohen, J.R. Chelikowsky, Electronic Structure and Optical Properties of Semiconductors, Springer Series Solid-State Sci. **75**, Springer, Berlin (1988)
[88F1] Ch. Flytzains, D. Ricard and Ph. Roussignol, NATO ASI Ser. B **194**, 181 (1988)
[88G1] Growth and Optical Properties of Wide-Gap II–VI Low Dimensional Semiconductors, T.C. McGill, C.M. Sotomayor Torres, W. Gebhardt (eds.), NATO ASI Ser. B **200**, Plenum, New York (1988)
[88P1] Properties of Impurity States in Superlattice Semiconductors, C.Y. Fong, I.P. Batra and S. Ciraci (eds.), NATO ASI Series B **183**, Plenum, New York (1988)
[88R1] U. Rössler, F. Malcher and A. Ziegler, in NATO ASI series B **183**, 219, Plenum, New York (1988)
[88W1] W. Walukiewizc, Phys. Rev. B **37**, 4760 (1988)
[89H1] A. Henglein, Chem. Rev. **89**, 1861 (1989)
[89K1] C. Klingshirn in [81A1]d of Chap. 1 p. 111
[89P1] N. Peyghambarian, et al., IEEE J. QE-**25**, 2516 (1989)
[90B1] M.G. Bawendi, M.L. Steigerwald and L.E. Brus, Annual Rev. Phys. Chem. **41**, 477 (1990)
[90M1] U. Merkt, Festkörperprobleme / Advances in Solid State Physics, Vol. **30**, 77 (1990)
[90P1] S.H. Park et al., J. Opt. Soc. Am. B **7**, 2097 (1990)
[91B1] L.E. Brus, Appl. Phys. A **53**, 465 (1991)
[91E1] Electronic Structure and Properties of Semiconductors, W. Schröter (ed.), Materials Science and Technology **4**, VCH, Weinheim (1991)
[91E2] A.I. Ekimov, Phys. Scr. T **39**, 217 (1991)
[91K1] A.A. Klochikhin and S.G. Ogloblin, Sov. Phys. JETP **73**, 1122 (1991)

[92E1] A.L. Efros, Phys. Rev. B **46**, 7448 (1992)
[92H1] F. Henneberger et al., Adv. Solid State Phys./ Festkörperprobleme **32**, 279 (1992)
[92M1] S. Malzer et al., phys. stat. sol. (b) **173**, 459 (1992)
[92P1] S. Permogorov and A. Reznitsky, J. Lumin. **52**, 201 (1992)
[93D1] F. Daminger et al., Proc. 21th ICPS, Beijing (1992), Ping Jiang and Hou-Zhi Zeng (eds.) p 1293 World Scientific, Singapore (1993)
[93E1] A.I. Ekimov et al., J. Opt. Soc. Am. B **10**, 100 (1993)
[93F1] P. Fulde, Electron Correlations in Molecules and Solids, Springer Series in Solid State Science **100**, 2nd, ed. Springer, Berlin (1993)
[93H1] A.C. Hewson, The Kondo Problem to Heavy Fermions, Cambridge Studies in Magnetism, Cambridge University, Press, Cambridge (1993)
[93K1] A.A. Klochikhin et al., Phys. Rev. B **48**, 3100 (1993)
[93M1] C.B. Murray, D.J. Norris and M.G. Bawendi, J. Am. Chem. Soc. **115**, 8706 (1993)
[93M2] U. Merkt et al., Physica B **189**, 165 (1993)
[93P1] K. Ploog, R. Nötzel and O. Brandt, Proc. 21th ICPS, Beijing (1992), Ping Jiang and Hou-Zhi Zeng (eds.) p. 1297 World Scientific, Singapore (1993)
[93R1] M.A. Reed, Scientific American, January issue p. 98 (1993)
[93T1] T. Taguchi, Y. Kawakami and Y. Yamada, Physica B **191**, 23 (1993)
[93W1] U. Woggon et al., Phys. Rev. B **47**, 3684 (1993)
[94A1] A. Abounadi et al., Phys. Rev. B **50**, 11677 (1994)
[94G1] M. Grün et al., J. Crystal Growth **138**, 150 (1994)
[94I1] M. Illing et al., J. Crystal Growth **138**, 638 (1994)
[95C1] A.-B. Chen and A. Sher, Semiconductor Alloys, Plenum Press, New York (1995)
[95F1] W. Faschinger, Material Science Forum **182-184**, 29 (1995)
[95G1] M. Grundmann et al., J. of Nonlinear Optical Physics and Materials **4**, 99 (1995)
[95H1] F. Hofmann and D.A. Wharam, Festkörperprobleme / Advances in Solid State Physics Vol. **35**, 197 (1995)
[95K1] J. König et al., Festkörperprobleme / Advances in Solid State Physics Vol. **35**, 215 (1995)
[95P1] D. Pfannkuche and S.E. Ulloa, Festkörperprobleme / Advances in Solid State Physics, Vol. **35**, 65 (1995)
[95W1] U. Woggon and S.V. Gaponenko, phys. stat. sol. (b) **189**, 285 (1995)
[96B1] H.R. Blick et al., Phys. Rev. B **53**, 78999 (1996)
[97H1] M. Hetterich et al., Phys. Rev. B **56**, 12369 (1997)
[97H2] D. Hommel et al., phys. stat. sol. (b) **202**, 835 (1997)
[97H3] A. Hartmann et al., Appl. Phys. Lett. **71**, 1314 (1997)
[97P1] T. Pohjola et al., Europhys. Lett. **40**, 189 (1997)
[97R1] M. Rabe et al., phys. stat. sol. (b) **202**, 817 (1997)
[97U1] R. Ulbrich, private communication (1997)
[97W1] U. Woggon et al., Appl. Phys. Lett. **71**, 377 (1997)
[98G1] M. Grün et al., Appl. Phys. Lett. **73**, 1343 (1998)
[98L1] D. Lüerßen et al., Phys. Rev. B **57**, 1631 (1998)
[99A1] M.V. Artemyev et al., Phys. Rev. B **60**, 1504 (1999) and references therein
[99B1] D. Bimberg, N. Grundmann and N.N. Ledentsov, Quantum Dot Heterostructures, John Wiley and Sons, Chicester (1999)

[99B2] G. Biasiol and E. Kapon, J. Crystal Growth **201/202**, 62 (1999)
[99G1] S.V. Gaponenko et al., IEEE J. Lightwave Techn. **17**, 2128 (1999)
[99K1] A. Klochikhin et al., Phys. Rev. B **59**, 12947 (1999)
[00A1] D. ben Avraham and S. Havlin, Diffusion and Reactions in Fractals and Disordered Systems, Cambridge University Press, Cambridge (2000)
[00B1] G. Biasiol, K. Leifer and E. Kapon, Phys. Rev. B **61**, 7223 (2000)
[00K1] E. Kurtz et al., Thin Solid Films **367**, 68 (2000) and ibid. **412**, 89 (2002), J. Crystal Growth **214/215**, 712 (2000), phys. stat. sol. (b) **229**, 519 (2001)
[00P1] S. Permogorov et al., J. Crystal Growth **215/215**, 1158 (2000)
[00R1] A. Rosenauer et al., Thin Solid Films **357**, 18 (2000) and Phys. Rev. B **61**, 8276 (2000)
[01A1] M.V. Artemyev et al., phys. stat. sol. (b) **224**, 393 (2001)
[01H1] Q. Huang et al., J. Crystal Growth **227/228**, 117 (2001)
[01H2] M.H. Huang et al., Science **292**, 1897 (2001)
[01P1] Z.W. Pan, Z.R. Dai and Z.L. Wang, Science **291**, 1947 (2001)
[01S1] R.L. Sellin et al., Appl. Phys. Lett. **78**, 1207 (2001)
[02G1] O. Gogolin et al., Solid State Commun. **122**, 511 (2002)
[02K1] E. Kapon in [81A1]k of Chap. 1, p. 243
[02W1] R.L. Walter et al., Phys. Rev. B **65**, 075207 (2002)
[03G1] O. Gogolin et al., J. Lumin. **102/103**, 414 and 451 (2003)
[03K1] S.N. Khanna and A.W. Castleman, Quantum Phenomena in Clusters and Nanostructures, Springer, Berlin, Heidelberg (2003)
[03L1] J.Y. Lao et al., Nano Letters **3**, 235 (2003)
[03Z1] B.P. Zhang et al., Appl. Phys. Lett. **83**, 1635 (2003)
[04Z1] B.P. Zhang et al., Appl. Phys. Lett. **84**, 4098 (2004) and J. Appl. Phys. **96**, 340 (2004)

9

Excitons, Biexcitons and Trions

In Chap. 8 we defined the bandstructure for electrons and holes as the solutions to the $(N \pm 1)$-particle problem and later we saw that the number of electrons in a band can be increased or decreased by donors and acceptors, respectively (Sect. 8.14). In contrast, the number of electrons remains constant in the case of optical excitations with photon energies in the eV or band gap region. What we can do, however, is to excite an electron from the valence to the conduction band by absorption of a photon. In this process, we bring the system of N electrons from the ground state to an excited state. What we need for the understanding of the optical properties of the electronic system of a semiconductor an insulator or even a metal is therefore a description of the excited states of the N particle problem. The quanta of these excitations are called "excitons" in semiconductors and insulators.

We can look at this problem from various points of view.

The ground state of the electronic system of a perfect semiconductor is a completely filled valence band and a completely empty conduction band. We can define this state as the "zero" energy or "vacuum" state. In addition it has total momentum $\boldsymbol{K} = 0$, angular momentum $\boldsymbol{L} = 0$ and spin $\boldsymbol{S} = 0$. From this point $E = 0$, $\boldsymbol{K} = 0$ we will start later on to consider the dispersion relation of the excitons in connection with Fig. 9.1b.

Another point of view is the following. If we start from the above-defined groundstate and excite one electron to the conduction band, we simultaneously create a hole in the valence band (Fig. 9.1a). In this sense an optical excitation is a two-particle transition. The same is true for the recombination process. An electron in the conduction band can return radiatively or non-radiatively into the valence band only if there is a free place, i.e., a hole. Two quasiparticles are annihilated in the recombination process.

This concept of electron–hole pair excitation is also used successfully in other disciplines of physics. If one excites, e.g., an electron in a (larger) atom or a neutron or proton in a nucleus from a deeper lying occupied state into a higher, empty one, a quantitative description concerning the transitions energies is obtained only if one takes into account both the particle in the

excited state and the hole left behind. This is even true if one excites an electron in a metal from a state in the Fermi sea with an energy below the Fermi energy E_F to an empty state above.

Excitons can be described at various levels of sophistication. We present in the next sections the most simple and intuitive picture using the effective mass approximation. Other approaches are described in [57E1, 62N1, 63K1, 63P1, 77B1, 78U1, 79E1, 79R1, 79S1, 81F1, 81K1, 85H1] or [82M1, 86U1, 93P1, 96Y1, 98E1, 98R1, 04O1] of Chap. 1 and references therein.

In Chap. 27 we shall also see how excitons are described in semiconductor Bloch equations.

The concepts of Wannier and Frenkel excitons were introduced in the second half of the 1930s [31F1, 37W1]. There is some controversy concerning the first experimental observations. The author does not wish to act as the referee to settle this point. Instead we give some references to early work [50H1, 52G1, 56G1, 58N1, 62N1] and to Fig. 13.9 and leave the decision to the reader.

9.1 Wannier and Frenkel Excitons

Using the effective mass approximations, Fig. 9.1a suggests that the Coulomb interaction between electron and hole leads to a hydrogen-like problem with a Coulomb potential term $-\,e^2/(4\,\pi\,\epsilon_0\,\epsilon\,|\boldsymbol{r}_e - \boldsymbol{r}_h|)$.

Indeed excitons in semiconductors form, to a good approximation, a hydrogen or positronium like series of states below the gap. For simple parabolic

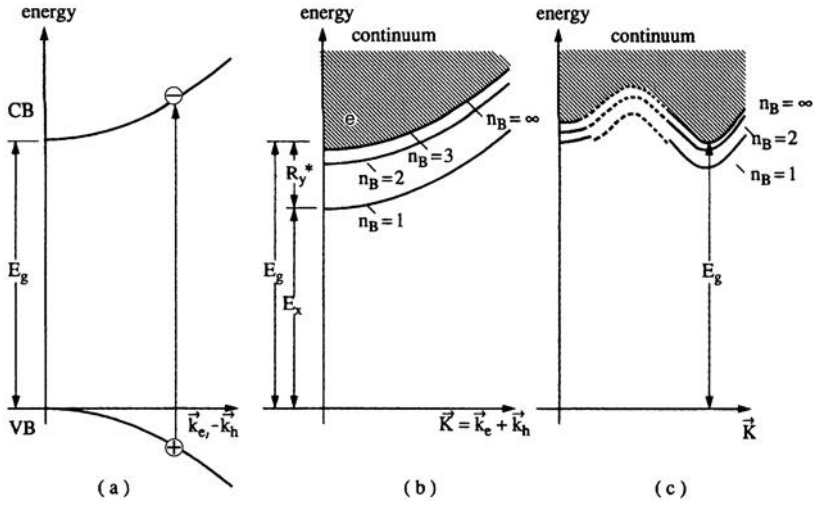


Fig. 9.1. A pair excitation in the scheme of valence and conduction band (**a**) in the exciton picture for a direct (**b**) and for an indirect gap semiconductor (**c**)

bands and a direct-gap semiconductor one can separate the relative motion of electron and hole and the motion of the center of mass. This leads to the dispersion relation of excitons in Fig. 9.1b.

$$E_{\mathrm{ex}}(n_{\mathrm{B}}, \boldsymbol{K}) = E_{\mathrm{g}} - \mathrm{Ry}^* \frac{1}{n_{\mathrm{B}}^2} + \frac{\hbar^2 \boldsymbol{K}^2}{2M} \tag{9.1a}$$

with

$$n_{\mathrm{B}} = 1, 2, 3 \ldots \quad \text{principal quantum number}\,,$$

$$\mathrm{Ry}^* = 13.6\,\mathrm{eV} \frac{\mu}{m_0} \frac{1}{\varepsilon^2} \quad \text{exciton Rydberg energy}\,, \tag{9.1b}$$

$$M = m_{\mathrm{e}} + m_{\mathrm{h}}, \quad \boldsymbol{K} = \boldsymbol{k}_{\mathrm{e}} + \boldsymbol{k}_{\mathrm{h}} \quad \text{translational mass and wave vector of the exciton}\,. \tag{9.1c}$$

For the moment, we use a capital $\boldsymbol{K}$ for the exciton wave vector to distinguish this two-particle state from the one-particle states. When we are more familiar with the exciton as a new quasi-particle we shall return to $\boldsymbol{k}$.

$$\mu = \frac{m_{\mathrm{e}}\, m_{\mathrm{h}}}{m_{\mathrm{e}} + m_{\mathrm{h}}} \quad \text{reduced exciton mass}\,, \tag{9.1d}$$

$$a_{\mathrm{B}}^{\mathrm{ex}} = a_{\mathrm{B}}^{\mathrm{H}} \varepsilon \frac{m_0}{\mu} \quad \text{excitonic Bohr radius}\,. \tag{9.1e}$$

The radii of higher states can be considered on various levels of complexity. If one only takes into account the exponential term $\exp\{Zr/na_{\mathrm{B}}^{\mathrm{H}}\}$ in the radial part of the wave function of the hydrogen problem appearing as the envelope function in (9.4a) and defines the (excitonic) Bohr radius by the decrease of this term to 1/e, one obtains with $Z = 1$ for excitons

$$a_{\mathrm{B}}\,(n_{\mathrm{B}}) = a_{\mathrm{B}}^{\mathrm{H}}\, \varepsilon\, \frac{m_0}{\mu}\, n_{\mathrm{B}}\,, \tag{9.2a}$$

i.e., a linear increase with n_{B}.

If, on the other hand, one takes the full radial function into account, i.e., also including the factor

$$\rho^l\; L_{n_{\mathrm{B}}+l}^{2l+1}\;(\rho) \tag{9.2b}$$

where l is the angular quantum number and $\rho = 2Z\, r/n_{\mathrm{B}} \cdot a_{\mathrm{B}}^{\mathrm{H}}$ and $L_{n_0\mathrm{B}+l}^{2l+1}$ are the Laguerre polynomials and calculates the average distance between electron and proton or hole, respectively, one obtains [55S1, 94L1]

$$\langle r\,(n_{\mathrm{B}})\rangle \;=\; \frac{a_{\mathrm{B}}}{2}\;\left[3\, n_{\mathrm{B}}^2 \;-\; l\;(l\;+\;1)\right] \tag{9.2c}$$

i.e., for the $n_{\mathrm{B}}S$ states (i.e. $l = 0$) a quadratic dependence starting with $3\, a_{\mathrm{B}}/2$ for $n_{\mathrm{B}} = 1$.

The series of exciton states in (9.1a) has an effective Rydberg energy Ry^* modified by the reduced mass of the electron and hole and the dielectric "constant" of the medium in which these particles move; n_B is the principal quantum number. The kinetic energy term in (9.1a) involves the translational mass M and the total wave vector $\boldsymbol{K}$ of the exciton. The radius of the exciton equals the Bohr radius of the H atom again modified by ε and μ. Using the material parameters for typical semiconductors one finds

$$1\,\mathrm{meV} \leq \mathrm{Ry}^* \leq 200\,\mathrm{meV} \ll E_\mathrm{g} \tag{9.3a}$$

and

$$50\,\mathrm{nm} \gtrsim a_\mathrm{B} \geq 1\,\mathrm{nm} > a_\mathrm{lattice}\,. \tag{9.3b}$$

This means that the excitonic Rydberg energy Ry^* is usually much smaller than the width of the forbidden gap and the Bohr radius is larger than the lattice constant. This second point is crucial. It says that the "orbits" of electron and hole around their common center of mass average over many unit cells and this in turn justifies the effective mass approximation in a self-consistent way. These excitons are called Wannier excitons [37W1].

In this limit, excitons can usually also be described in the frame of semiconductor Bloch equations, but finer details or corrections to this simple model, which we shall treat in Sect. 9.2 and which are necessary to understand optical spectra, are usually not incorporated in the semiconductor Bloch equations of Chap. 27.

It should be mentioned that in insulators like NaCl, or in organic crystals like anthracene, excitons also exist with electron–hole pair wavefunctions confined to one unit cell. These so-called Frenkel excitons [31F1] cannot be described in the effective mass approximation. As a rule of thumb, one can state that in all semi-conductors the inequalities (9.3a,b) hold, so that we always deal with Wannier excitons.

A series of conferences devoted both to excitons in anorganic semiconductors and in organic ones as well as in anorganic and organic insulators is EXCON (excitonic process in condensed matter). The proceedings have been published in [94E1]

To get an impression of the wavefunction, we form wave packets for electrons and holes $\phi_\mathrm{e,h}(\boldsymbol{r}_\mathrm{e,h})$ in the sense of the Wannier function of (8.10) and obtain schematically for the exciton wavefunction

$$\phi(\boldsymbol{K}, n_\mathrm{B}, l, m) = \Omega^{-1/2}\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}}\phi_\mathrm{e}(\boldsymbol{r}_\mathrm{e})\phi_\mathrm{h}(\boldsymbol{r}_\mathrm{h})\phi^\mathrm{env}_{n_\mathrm{B},l,m}(\boldsymbol{r}_\mathrm{e}-\boldsymbol{r}_\mathrm{h})\,, \tag{9.4a}$$

with the center of mass $\boldsymbol{R}$

$$\boldsymbol{R} = (m_\mathrm{e}\,\boldsymbol{r}_\mathrm{e} + m_\mathrm{h}\,\boldsymbol{r}_\mathrm{h})/(m_\mathrm{e}+m_\mathrm{h})\,, \tag{9.4b}$$

where $\Omega^{-1/2}$ is the normalization factor. The plane-wave factor describes the free propagation of a Wannier exciton through the periodic crystal similarly as

for the Bloch waves of Sect. 8.1, and the hydrogen-atom-like envelope function $\phi^{\mathtt{env}}$ describe the relative motion of electron and hole.

The quantum numbers l and m, with $l < n_{\mathrm{B}}$ and $-l \leq m \leq l$ have the same meaning as for the hydrogen atom in the limit that the angular momentum is a good quantum number in a solid (see Chap. 26).

As for the H atom, the exciton states converge for $n_{\mathrm{B}} \to \infty$ to the ionization continuum, the onset of which coincides with E_{g}.

Excitons features are especially strong for regions of the electron and hole dispersions, where the group velocities of electrons and holes v_g^e and v_g^h are equal e.g. zero. See e.g. [96Y1] of Chap. 1. This means, that in direct gap semiconductors excitons form preferentially around $\boldsymbol{K} = 0$ if the direct gap occurs at the Γ point.

In indirect gap semiconductors exciton states form preferentially with the hole around $\boldsymbol{k}_h = 0$ and the electron in its respective minimum, as shown in Fig. 9.1c schematically e.g. for Ge where the conduction band has minima at $\boldsymbol{k}_e = 0$ and $\boldsymbol{k}_e \neq 0$. Electrons and holes with different group velocities still "see" their mutual Coulomb interaction, but the dashed continuation of exciton dispersion from the indirect to the direct gap is an oversimplification, among others because the states away from the band extrema are strongly damped and their binding energy varies. See e.g. again [96Y1] of Chap. 1.

The discrete and continuum states of the excitons will be the resonances or oscillators which we have to incorporate into the dielectric function of Chaps. 4 and 5.

For direct semiconductors with dipole allowed band-to-band transitions, one finds an oscillator strength for excitons in discrete states with $S(l = 0)$ envelope function proportional to the band-to-band dipole transition matrix element squared and to the probability of finding the electron and hole in the same unit cell. For the derivation of this relation see [57E1]. This latter condition leads to the n_{B}^{-3} dependence of the oscillator strength for three-dimensional systems.

$$f_{n_{\mathrm{B}}} \propto |H_{\mathrm{cv}}^{\mathrm{D}}|^2 \frac{1}{n_{\mathrm{B}}^3} \,. \tag{9.5}$$

These $f_{n_{\mathrm{B}}}$ result in corresponding longitudinal–transverse splitting as shown in connection with (4.26). Equation ((9.5)) holds for so-called singlet excitons with antiparallel electron and hole spin. Triplet excitons involve a spin flip, in their creation which significantly reduces their oscillator strength (spin flip forbidden transitions).

Since the singlet and triplet pair or exciton states will play some role in later chapters (see, e.g., Chaps. 13 to 16) we give some simplified information on this topic here.

The crystal ground state, i.e., completely filled valence bands and completely empty conduction bands, has, as already mentioned above, angular momentum $\boldsymbol{L}$, spin $\boldsymbol{S}$ and total angular momentum $\boldsymbol{J}$ equal to zero. If we excite optically an electron from the valence band to the conduction band,

e.g., by an electric dipole transition, the spin of the excited electron does not change because the electric field of the light does not act on the spin. Consequently, the simultaneously created hole has a spin opposite to the one of the excited electron and the total spin $\boldsymbol{S}$ of the electron–hole pair state is still zero. Consequently, the electron–hole pair, and likewise the exciton, is said to be in a spin singlet state. The spin S = 1 of the photon is accommodated by the spatial part of the band-to-band matrix element in ((9.5)) or by the envelope function of the exciton (see (9.4a) and [96Y1] of Chap. 1).

If the spin flips in the transition, e.g., by interaction with the magnetic component of the light field, one ends up with a total spin S = 1 corresponding to spin triplet (exciton) state. The triplet state is situated energetically below the singlet state and the splitting is due to a part of the electron–hole exchange interaction, essentially the so-called short-range or analytic (for $\boldsymbol{K} \to 0$) part of the exchange interaction. There are names other than triplet and singlet used for certain materials like para and ortho exciton (e.g., for Cu_2O) or dark and bright excitons (in quantum dots). The last pair of names reflects the fact that triplet excitons have small oscillator strength because they are "spin-flip forbidden." Since spin and angular momentum are, strictly speaking, no good quantum numbers in a crystal (see Chap. 26) it is not obligatory that the triplet exciton is threefold degenerate and the singlet is not degenerate. There are cases where the triplet or para exciton is nondegenerate, as in Cu_2O, and the singlet is threefold degenerate, as in many zinc-blende type crystals like CuCl or ZnSe or GaAs. This is possibly the reason that in some older, especially French, literature the names singlet and triplet exciton states are interchanged.

The oscillator strength of the continuum states is influenced by the so-called Sommerfeld enhancement factor. We come back to this point later in (9.3) and in Chaps. 13 and 15, when we discuss the optical properties.

In the picture of second quantization, we can define creation operators for electrons in the conduction band and for holes in the valence band $\alpha^+_{\boldsymbol{k}_\mathrm{e}}$ and $\beta^+_{\boldsymbol{k}_h}$, respectively. The combination of both gives creation operators for electron hole pairs $\alpha^+_{\boldsymbol{k}_\mathrm{e}}\beta^+_{\boldsymbol{k}_h}$. The exciton creation operator B^+ can be constructed via a sum over electron–hole pair operators (see [81K1] or [93H1] of Chap. 1).

$$B^+_{\boldsymbol{K}} = \sum_{\boldsymbol{k}'_\mathrm{e}\boldsymbol{k}_\mathrm{h}} \delta[\boldsymbol{K} - (\boldsymbol{k}_\mathrm{e} + \boldsymbol{k}_\mathrm{h})]\alpha_{\boldsymbol{k}_\mathrm{e},\boldsymbol{k}_\mathrm{h}}\alpha^+_{\boldsymbol{k}_\mathrm{e}}\beta^+_{\boldsymbol{k}_\mathrm{h}}\,; \tag{9.6}$$

the expansion coefficients $a_{\boldsymbol{k}_\mathrm{e},\boldsymbol{k}_\mathrm{h}}$ correspond, in principle, to those also used in an (slightly old-fashioned) expansion into Slater determinantes of the many particle problem, which contains either all valence band states $\varphi_{\boldsymbol{k}_h}(\boldsymbol{r}_{i,h})$ for the ground state or always has one line being replaced by a conduction band state [63K1].

It can be shown that the $B^+_{\boldsymbol{K}}, B_{\boldsymbol{K}}$, obey Bose commutation relations with a density-dependent correction term which increases with the number

of electrons and holes contained in the volume of one exciton $4\pi(a_{\mathrm{B}}^{\mathrm{ex}})^3/3$ [77H2, 81K1].

This has two consequences: in thermodynamic equilibrium for low densities and not too low temperatures, the excitons can be well described by Boltzmann statistics with a chemical potential ruled by their density and temperature similar to (8.43). For higher densities they deviate more and more from ideal bosons until they end up in an electron–hole plasma made up entirely from fermions (see Sect. 20.5). This makes the creation of a Bose-condensed state of excitons (or of biexcitons) a very complicated problem. Though theory predicts a region in the temperature density plane where an excitonic Bose–Einstein condensed state could occur [77H2], there are currently many hints but no generally accepted, clearcut observation of a spontaneous Bose condensation of excitons. There are some experiments that prove the Bose character of excitons [76L1, 83H1, 83M1, 83P1, 84W1, 87F1]. We come back to this problem in Sect. 20.5 where we shall discuss also more recent results.

9.2 Corrections to the Simple Exciton Model

The simple model outlined in the preceeding section is, as already mentioned, adequate for non-degenerate, parabolic bands. We keep these assumptions for the moment and inspect a first group of corrections which are relevant for the parameters ε and μ entering in (9.1a–e). We already know from Chaps. 4–7 that ε is a function of ω, resulting in the question of which value should be used.

As long as the binding energy of the exciton $E_{\mathrm{ex}}^{\mathrm{b}}$ is small compared to the optical phonon energies and, consequently, the excitonic Bohr radius ((9.1e)) larger than the polaron radius (8.16)

$$E_{\mathrm{ex}}^{\mathrm{b}} < \hbar\omega_{\mathrm{LO}}; \quad a_{\mathrm{B}} > a_{\mathrm{Pol}} \Rightarrow \varepsilon = \varepsilon_{\mathrm{s}}\,, \tag{9.7a}$$

we can use for ϵ the static value ε_{s} below the phonon resonances and the polaron masses and polaron gap. This situation is fulfilled for some semiconductors for all values of n_{B}, e.g., for GaAs where $\mathrm{Ry}^* \simeq 5\,\mathrm{meV}$ and $\hbar\omega_{\mathrm{LO}} \simeq 36\,\mathrm{meV}$.

In many other semiconductors including especially the wide gap semiconductors (see Fig. 9.3) the inequality (9.7a) holds only for the higher states $n_{\mathrm{B}} \geq 2$, while for the ground state exciton ($n_{\mathrm{B}} = 1$) we find

$$E_{\mathrm{ex}}^{\mathrm{b}} \gtrsim \hbar\omega_{\mathrm{LO}}; \quad a_{\mathrm{B}} \simeq a_{\mathrm{Pol}} \longrightarrow \varepsilon_{\mathrm{s}} \geq \varepsilon \geq \varepsilon_{\mathrm{b}}\,. \tag{9.7b}$$

Examples are CdS, ZnO, CuCl or Cu_2O.

In this situation a value for ε between ε_{s} and ε_{b} seems appropriate, because the polarization of the lattice can only partly follow the motion of electron and hole. A useful approach is the so-called Haken potential [55H1] which interpolates between ε_{s} and ε_{b} depending on the distance between electron

and hole, where r_{eh}, $r_{\mathrm{e}}^{\mathrm{p}}$ and $r_{\mathrm{p}}^{\mathrm{h}}$ are the distances between electron and hole, and the polaran radii of electron and hole, respectively:

$$\frac{1}{\varepsilon(\boldsymbol{r}_{\mathrm{e,h}})} = \frac{1}{\varepsilon_{\mathrm{b}}} - \left(\frac{1}{\varepsilon_{\mathrm{b}}} - \frac{1}{\varepsilon_{\mathrm{s}}}\right)\left(1 - \frac{\exp(-\boldsymbol{r}_{\mathrm{eh}}/\boldsymbol{r}_{\mathrm{e}}^{\mathrm{p}}) + \exp(-\boldsymbol{r}_{\mathrm{eh}}/\boldsymbol{r}_{\mathrm{h}}^{\mathrm{p}})}{2}\right) . \quad (9.8)$$

The next correction concerns the effective masses. The polarization clouds of the polarons (Sect. 8.6) have different signs for electron and hole. If both particles are bound together in an exciton state fulfilling (9.7b) the polaron renormalization is partly quenched, with the consequence that values for the effective masses will lie somewhere between the polaron values and the ones for a rigid lattice. The gap "seen" by the exciton in the $1S$ state will likewise be situated between the two above extrema. Fortunately, the above effects tend to partly compensate each other. A transition from the polaron gap to the larger rigid lattice gap shifts the exciton energy to larger photon energies. A transition from ε_{s} to ε_{b} and a increase of the effective masses increases the binding energy and shift the $1S$ exciton to lower photon energies. As a consequence one finds, even for many semiconductors for which inequality (9.7b) holds, that the $1S$ exciton fits together with the higher exciton states reasonably well into the hydrogen-like series of (9.1). We shall use this approach in the future if not stated otherwise. In other cases like Cu_2O or CuCl, the higher states with $n_{\mathrm{B}} \geq 2$ follow a hydrogen-like n_{B}^{-2} series converging to the polaron gap, but the $1S$ exciton shows, with respect to the polaron gap, a binding energy $E_{\mathrm{ex}}^{\mathrm{b}}$, which differs from and is generally larger than the excitonic Rydberg energy deduced from the higher states. We call the experimentally observed energetic distance between the $1S$ exction and the polaron gap the exciton binding energy $E_{\mathrm{ex}}^{\mathrm{b}}$, in contrast to Ry^* in (9.1). This discrepancy introduces some ambiguity when comparing theoretical results with experimental data, since in theory one often normalizes energies with the excitonic Rydberg energy Ry^*, but the $1S$ exciton has a different value for $E_{\mathrm{ex}}^{\mathrm{b}}$.

There is a general trend of the material parameters m_{eff} and ε with E_{g} which results in an increase of the exciton binding energy with increasing E_{g} as shown in Fig. 9.3.

This is an analogous consequence of increasing the band width or in other words decreasing effective mass with decreasing width of the gap as described, e.g., in the band structure model of Sect. 8.5.

The next complication comes from the band structure. If the bands are degenerate, as is the Γ_8 valence band in T_{d} symmetry, it is no longer possible to separate the relative and the center of mass motion – they are coupled together. Similar effects stem from k-linear terms and other sources. We get light- and heavy-hole exciton branches and splittings, e.g., between the 2S and 2P exciton states partly induced by the envelope function. An example for Cu_2O will be given in Sect. 13.2.

Furthermore it should be mentioned, without going into details, that the splitting between singlet and triplet excitons Δ_{st} and the splitting of the

singlet state into a transverse and a longitudinal one Δ_{LT}, are both due to exchange interaction between electron and hole caused by their Coulomb interaction [73D1, 88F1, 03G1] if we consider the N-electron problem in the form of Slater's determinant where the ground state consists only of valence-band states and the excited state of a sum of determinants in each of which one valence-band state is replaced by a conduction-band state. This aspect is treated in detail in [73D1] or in [93H1, 93P1] of Chap. 1.

Usually the following relation holds for Wannier excitons

$$\Delta_{\mathrm{st}} \ll \Delta_{\mathrm{LT}} \quad \text{with} \quad 0.1\,\mathrm{meV} \lesssim \Delta_{\mathrm{LT}} \lesssim 15\,\mathrm{meV}. \tag{9.9}$$

For the AΓ_5-excitons in CdS or ZnO one finds, e.g., $\Delta_{\mathrm{LT}}^{\mathrm{CdS}} = 1.8\,\mathrm{meV}$, $\Delta_{\mathrm{st}}^{\mathrm{CdS}} = 0.2\,\mathrm{meV}$ and $\Delta_{\mathrm{LT}}^{\mathrm{ZnO}} = 1.2\,\mathrm{meV}$, $\Delta_{\mathrm{st}}^{\mathrm{ZnO}} = 0.17\,\mathrm{meV}$. Since the short range interaction increases with decreasing exciton radius as a_{B}^{-3} the situation begins to change for 1S excitons with a value of $a_{\mathrm{B}}^{\mathrm{ex}}$ exceeding only slightly the lattice constant, and leading thus also to the limit of the concept Wannier excitons. One finds, e.g., for CuCl or Cu_2O, $\Delta_{\mathrm{LT}}^{\mathrm{CuCL}} = 5.5\,\mathrm{meV}$, $\Delta_{\mathrm{st}}^{\mathrm{CuCl}} = 2.5\,\mathrm{meV}$, $\Delta_{\mathrm{LT}}^{\mathrm{Cu_2O}} \approx 50\,\mu\mathrm{eV}$, $\Delta_{\mathrm{st}}^{\mathrm{Cu_2O}} = 12\,\mathrm{meV}$ (see [85H1]). The extremely low value of Δ_{LT} in Cu_2O comes from the fact, that the band-to-band transition is parity forbidden and the 1S singlet or ortho exciton is only allowed in quadrupole approximation (see Sect. 13.2).

The singlet-triplet splitting is even enhanced by the compression of the excitonic wave function in quantum wells and especially in quantum dots (see, e.g., [98F1, 03G1]). We come back to this aspect in Sect. 9.3 and Chap. 15.

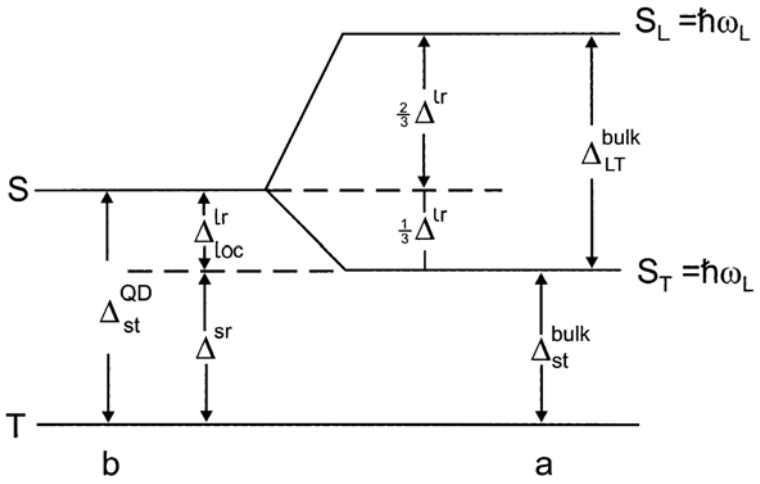


Fig. 9.2. The splitting of the 1S exciton with the notation of the various contributions to exchange splitting for bulk samples (a) and in a quantum dot (b). The abreviations have the following meanings: S = singlet, T = triplet, Δ^{lr} = long range or nanoanalytic (for $\boldsymbol{k} \to 0$) part of the exchange interaction, Δ^{sr} = short range or analytic part of the exchange interaction [03G1]

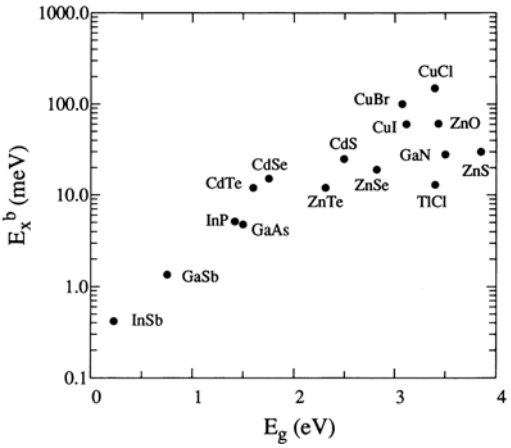


Fig. 9.3. The exciton binding energy $E_{\mathrm{ex}}^{\mathrm{b}}$ as a function of the band-gap for various direct gap semiconductors ([82L1, 93H1, 93P1] of Chap. 1)

It should be mentioned that the excitation of an optically (dipol-) allowed exciton is accompanied by a polarization as detailed in Chaps. 5 and 27 on polaritons and on (semiconductor) Bloch equations, respectively.

The concept of exciton-phonon boundstates has been introduced in [68T1, 72K1, 02B1].

Finally we mention that excitons can also be formed with holes in deeper valence bands. These so-called "core-excitons" are usually situated in the VUV or X-ray region of the spectrum and have a rather short lifetime. An example and further references are given in Sect. 13.1.7 and [87C1, 88K1].

9.3 The Influence of Dimensionality

If we consider the exciton again as an effective mass particle with parabolic dispersion relations, as given by (9.1), we expect a first influence of the dimensionality on the density of states analogous to the situation shown in Fig. 8.20 for every exciton branch $n_{\mathrm{B}} = 1, 2, 3 \ldots$

Another effect of the dimensionality manifests itself in the binding energy, the Rydberg series and the oscillator strength. We consider an exciton, for which the motion of electron and hole is restricted to a two-dimensional plane, but the interaction is still a 3d one, i.e., proportional to $\mathrm{e}^2/|\boldsymbol{r}_{\mathrm{e}} - \boldsymbol{r}_{\mathrm{h}}|$ and find (9.10a), (9.10b) in comparison to the 3d case of (9.1) (see e.g. [93H1] of Chap. 1):

$$\text{3d: } E(\boldsymbol{K}, n_{\mathrm{B}}) = E_{\mathrm{g}} - \mathrm{Ry}^* \frac{1}{n_{\mathrm{B}}^2} + \frac{\hbar^2 (K_x^2 + K_y^2 + K_z^2)}{2M} \tag{9.10a}$$

and for the oscillator strength f for the principal quantum number n_{B} in the limit of (9.2a):

$$f(n_{\mathrm{B}}) \propto n_{\mathrm{B}}^{-3}; \quad a_{\mathrm{B}} \propto a_{\mathrm{B}}^{\mathrm{H}} n_{\mathrm{B}}; \quad n_{\mathrm{B}} = 1, 2, 3 \ldots . \tag{9.10b}$$

$$\text{2d: } E(\boldsymbol{K}, n_\mathrm{B}) = E_\mathrm{g} + E_\mathrm{Q} - \mathrm{Ry}^* \frac{1}{(n_\mathrm{B} - \frac{1}{2})^2} + \frac{\hbar^2 (K_x^2 + K_y^2)}{2M} \tag{9.10c}$$

with E_Q quantization energy

$$f(n_\mathrm{B}) \propto \left(n_\mathrm{B} - \frac{1}{2}\right)^{-3}; \quad a_\mathrm{B} \propto a_\mathrm{B}^\mathrm{H} \left(n_\mathrm{B} - \frac{1}{2}\right), \quad n_\mathrm{B} = 1, 2, 3 \ldots . \tag{9.10d}$$

Essentially n_B has to be replaced by $n_\mathrm{B} - 1/2$ when going from 3d to 2d systems and the quantization energies E_Q of electrons and holes must be considered. Actually E_Q diverges for confinement in a mathematically 2d plane, see Sect. 8.10. The excitonic Rydberg Ry^* is the same in both cases with the consequence that the binding energy of the 1S exciton is Ry^* in three, and $4\mathrm{Ry}^*$ in two dimensions. The oscillator strength increases and the excitonic Bohr radius decreases when going from three- to two-dimensional systems.

The usual realization of quasi-2d excitons is via (M)QW of type I. In this case the motion in the z-direction is quantized, but the width of the quantum well l_z is non-zero.

In Fig. 9.4 we show the exciton binding energy for GaAs as a function of l_z for infinitely high barriers, where the curve reaches $4\mathrm{Ry}^*$ for $l_z = 0$ and for finite barrier height, where the binding energy converges to the value of the barrier material for $l_z = 0$ passing through a maximum of about 2 to 3 times Ry^* depending on the material parameters.

It is possible to describe the binding energy of the exciton in the quasi two-dimensional case of a QW in terms of an "effective" dimensionality d_eff that ranges between three and two and interpolates thus between the limiting cases of (9.10a) and (9.10c) by [91H1, 92M1].

$$E_\mathrm{ex}^\mathrm{b}\ (1\mathrm{S}) \ = \ \frac{\mathrm{Ry}^*}{\left(1 \ + \ \frac{d_\mathrm{eff} - 3}{2}\right)^2} \tag{9.10e}$$

$$\text{with} \quad d_\mathrm{eff} = 3 - \exp\left(\frac{-L_\mathrm{W}}{2a_\beta}\right) \tag{9.10f}$$

where a_β is the three-dimensional excitonic Bohr radius and L_W the width of the quantum well increased by the penetration depth of electron and hole into the barrier.

The increase of the oscillator strength of the 1S exciton comes from the fact that the quantization in the z-direction increases both the overlap between electron and hole and their attraction, which results in turn in a reduction of the two-dimensional Bohr radius.

The Sommerfeld factor F, which describes the enhancement of the oscillator strength of the continuum states and which arises from the residual electron–hole correlation, depends also on the dimensionality [66S1, 91O1]. It reads

$$\text{3d:} \quad F_\mathrm{3d} = \frac{\pi}{W^{1/2}} \frac{\mathrm{e}^{\pi W^{1/2}}}{\sinh(\pi W^{1/2})} \quad \text{with} \quad W = (E - E_\mathrm{g})\mathrm{Ry}^{*-1}; \tag{9.11a}$$

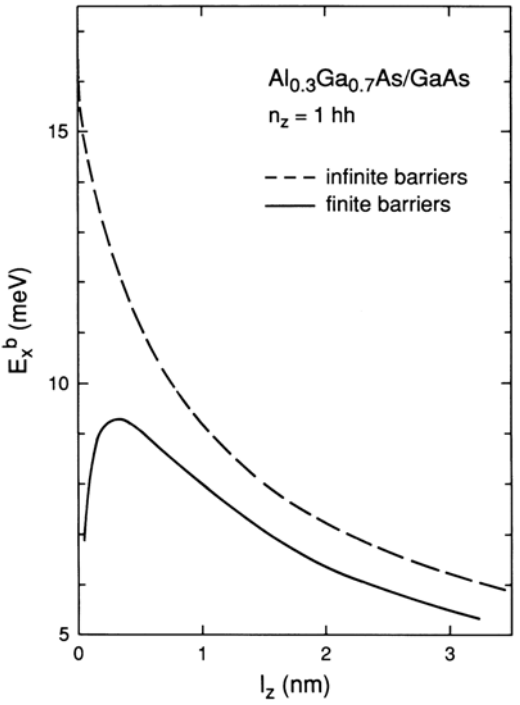


Fig. 9.4. The calculated binding energy of $n_z = 1$hh excitons in AlGaAs/GaAs quantum wells as a function of the well thickness l_z ([81M1, 84G1, 88K2])

$$\text{2d:} \quad F_{2\mathrm{d}} = \frac{\mathrm{e}^{\pi W^{1/2}}}{\cosh(\pi W^{1/2})} \quad \text{with} \quad W = (E - E_\mathrm{g} + E_\mathrm{Q})\mathrm{Ry}^{*-1}\,. \qquad (9.11\mathrm{b})$$

In the three-dimensional case it has a square-root singularity at E_g and decreases gradually to unity for $E > E_\mathrm{g}$. In two dimensions it decays only from two to one with increasing energy and in a one-dimensional system it is even below unity above the gap quenching thus the singularity in the (combined) density of states. We shall see the consequences in Chap. 13.

The corrections which we mentioned in Sect. 9.2 hold partly also in the quasi-2d case. The 2d valence band structure may be even more complex, than the 3d one, comparing e.g. Figs. 8.17 and Fig. 8.24. However, it should be noted that the most widely investigated (M) QW are based on AlGaAs and InGaAs, which fulfill the inequality (9.6).

The most striking feature of excitons in quasi-two-dimensional systems and in systems of even lower dimensionality, however, are the facts, that exciton series ($n_\mathrm{B} = 1 \ldots \infty$) exist for every combination of electron- and hole subband, though partly with small or vanishing oscillator strength and that the exciton splits into light- and heavy-hole excitons which results from the corresponding splitting of the valence-band states (Figs. 8.23 and 8.24). To describe the exciton states we thus need more quantum numbers. Apart from the principal quantum number n_B in (9.1) or (9.10a), we must state which of the quantized conduction- and valence-band states are involved. The sim-

plest optical interband selection rule is $\Delta n_z = 0$, so that we shall see in optical spectra mainly excitons which obey this rule. Finally we must specify whether we are speaking of the light- or the heavy-hole series. Complete information might thus be the $n_z^h = 2\mathrm{hh}, n_z^e = 2, n_\mathrm{B} = 1$ exciton state. Usually one uses the abbreviation $n_z = 2\mathrm{hh}$ exciton involving the above selection rule and the fact that excitons in MQW with $n_\mathrm{B} > 1$ are usually difficult to resolve due to broadening effects. For examples see Sect. 15.1.

As already discussed with one particle states in Sect. 8.10, minibands also form for excitons in superlattices.

In type II structures, the Coulomb interaction is decreased due to the spatial separation of electrons and holes in the two materials with wave function overlap reduced to the interface region. Excitons in such structures are also said to be indirect in real space.

In strictly one- and zero-dimensional cases the binding energy for the exciton diverges. So it is not possible to give general formulas like (9.10c) for these situations. One is always limited to numerical calculations which have to explicitly include the finite dimensions of the quantum wire or quantum dot.

The various possibilities to produce quantum wells, superlattices, quantum wires and dots have been discussed already in Sects. 8.11 to 8.13. In the following, we give some information relevant for approximately spherical quantum dots, as they occur frequently for semiconductor nanocrystals in glass or organic matrices. For details see e.g. [93B1, 97W1, 98G1, 98J1] of Chap. 1.

Three regimes of quantization are usually distinguished in quantum dots in which the crystallite radius R is compared with the Bohr radius of the excitons or related quantities:

$$\text{weak confinement:} \quad R \gtrsim a_\mathrm{B},\ E_\mathrm{Q} < \mathrm{Ry}^* ; \tag{9.12a}$$

$$\text{medium or intermediate confinement:} \quad a_\mathrm{B}^\mathrm{e} \geq R \geq a_\mathrm{B}^\mathrm{h};\ E_\mathrm{Q} \approx \mathrm{Ry}^* , \tag{9.12b}$$

$$\text{with} \quad a_\mathrm{B}^\mathrm{e,h} = a_\mathrm{B}^\mathrm{H} \varepsilon \frac{m_0}{m_\mathrm{e,h}}$$

$$\text{strong confinement:} \quad R \leq a_\mathrm{B}^\mathrm{h};\ E_\mathrm{Q} > \mathrm{Ry}^* . \tag{9.12c}$$

In the first case (9.12a) the quantum dot (QD) is larger than the exciton. As a consequence the center-of-mass motion of the exciton, which is described in (9.1a) and (9.4a) by the term $\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}}$, is quantized while the relative motion of electron and hole given by the envelope function $\phi_{n_\mathrm{B},l,m}(\boldsymbol{r}_\mathrm{e} - \boldsymbol{r}_\mathrm{h})$ is hardly affected. This situation is found, e.g., for the Cu halides where a_B is small, or for CdSe QD with $R \geq 10\,\mathrm{nm}$. In the second case (9.12b) R has a value between the radii of the electron orbit and the hole orbit around their common center of mass. As a consequence, the electron state is quantized and the hole moves in the potential formed by the dot and the space charge of the quantized electron. This case is the most demanding from the theoretical point of view since Coulomb effects and quantization energies are of the same order of magnitude. However,it is often realized for QD of II–VI semiconductors. The regime given by (9.12c) becomes easier again. The Coulomb energy increases

roughly with R^{-1}, and the quantization energy with R^{-2}, so that for sufficiently small values of R one should reach a situation where the Coulomb term can be neglected. However, recent investigation showed that $E_{\mathrm{Q}} \gg \mathrm{Ry}^*$ holds only for R values which are comparable with the lattice constant. In this case the applicability of the effective mass approximation becomes questionable, and the dot may be better considered as a huge molecule. The increase of the electron–hole exchange energy with decreasind dot radius has alredy been addressed in Sect. 9.2 with Fig. 9.2.

In addition to the above-mentioned difficulties, there are some others which lead to an inhomogeneous broadening of the spectra as compared to the δ-like density of states in Fig. 8.20 as well as other complications. The dots usually have a certain spread of R values, as already mentioned, which directly influences the quantization and the Coulomb energies. Though the "gap" of the surrounding amorphous glass matrix is usually much larger than that of the semiconductor.

$$E_{\mathrm{g}}^{\mathrm{SC}} < E_{\mathrm{g}}^{\mathrm{glass}} \tag{9.13}$$

there is no abrupt, infinitely high barrier. Interface states may appear, which depend both on the surrounding matrix and on the growth regime of the QD; image forces have to be considered in quantitative calculations, since the dielectric functions of semiconductor and glass are different; deviations of the QD from an ideal sphere are obvious from TEM investigations but usually neglected; the coupling of excitions to phonons is enhanced, especially in the regime of (9.12b), since the different radial distributions of the electron and hole wavefunctions give rise to a dipole layer. Finally a realistic bandstructure has to be taken into account at least as long as the effective mass approximation is still valid.

Presently, a lot of work is devoted to so-called quantum dots or, their better name, quantum islands, which form in a self-organized way in quantum wells in the form of pronounced thickness or composition fluctuations. Such islands, the diameter of which is usually much larger than their height, form especially in the strained systems InAs/GaAs and CdSe/ZnSe. In addition to the above problems, theses islands show, apart from their more pancake like shape, complications due to alloying and spatially inhomogeneous compositions, due to strain distributions within the island and due to the strong anisotropy of the confinement potential in the plane of the QW and normal to it. For more details of the structures of reduced dimensionality see e.g. [03O1] or [93B1,93H1, 93P1, 93O1, 93S1, 95I1, 96S2, 97W1, 98D1, 98G1, 98J1, 98R1, 98S1, 99B1, 99B2, 99G1, 99M1, 01H1, 01L1, 02S1, 03T1, 04O1] of Chap. 1 and references therein.

9.4 Biexitonns and Trions

When introducing the concept of excitons, we stated that they can be understood on a certain level of sophistication as the analogue to the hydrogen

atom or – even better – to the positronium atom. This analogy can be even pushed further. It is well known that two hydrogen atoms with opposite electron spin can bind to a form a hydrogen molecule. In the same sense it has been calculated, that two positronium atoms can form a positronium molecule as bound state.

So the idea was not far away, that two excitons could bind to form a new quasiparticle, the so-called biexciton or excitonic molecule [58L1, 58M1].

It has been found theoretically that the biexciton should form a bound state for all ratios of effective electron and hole masses and dimensionalities of the sample. The biexciton binding energy expressed in units of the excitonic Rydberg energy (see the discussion in Sect. 9.2) starts for bulk samples for $\sigma = m_e/m_h \Rightarrow 0$ at $E^b_{biex}/Ry^* \approx 0.3$ corresponding to the value of a hydrogen molecule and then drops monotonically until $\sigma = 1$, reaching values of 0.027 or 0.12 for $\sigma = 1$ [72A1, 72B1, 73H1, 81K1]. It approaches for $\sigma \Rightarrow \infty$ again the value of $\sigma = 0$, since the four-particle problem is symmetric against the exchange of electrons and holes, i.e., $\sigma \Rightarrow 1/\sigma$.

The dispersion relation is given in the simplest case by

$$E_{biex}(\boldsymbol{k}) = 2\left(E_g - E^b_{ex}\right) - E^b_{biex} + \frac{\hbar^2 \boldsymbol{k}^2}{4M_{ex}} \tag{9.14}$$

assuming that the effective mass of the biexciton is just twice that of the exciton.

Many of the complications mentioned for excitons in Sect. 9.2 also apply here. As we shall see in Sect. 20.3 biexcitons have indeed been observed experimentally in a wide variety of bulk semiconductors with both direct and indirect band gaps.

Biexcitons also form bound states in quantum wells, wires and dots again in agreement with experiments. The enhancement of the exciton binding energy in structures of reduced dimensionality also shows up in an increasing biexciton binding energy with increasing confinement [83K1, 96S1, 99D1].

In [96S1] it has been calculated, that the biexciton binding energy in quantum wells should be approximately 22% of the corresponding exciton binding energy independent of the well width. Though doubts have been raised concerning the theoretical model, it was found to be in good agreement with experiment.

Furthermore it has been predicted theoretically [76S1, 97S1] and verified experimentally, at least for quantum structures [95F1, 97S2], that trions also form bound states contributing to the luminescence below the free exciton energy. Trions are charged excitons or biexcitons, i.e., quasiparticles consisting of two electrons and one hole or vice versa. As we shall see, in Sect. 20.3 their observation is favored by a moderate n or p (modulation) doping of quantum well structures. Localization effects (see Sect. 9.6) may play a role. See e.g. [04K2]

9.5 Bound Exciton Complexes

Similar to the way that free carriers can be bound to (point-) defects, it is found that excitons can also be bound to defects.

We discuss first shallow impurities. The binding energy of an exciton (X) is highest for a neutral acceptor (A^0X complex), lower for a neutral donor (D^0X) and lower still for an ionized donor (D^+X). An ionized acceptor does not usually bind an exciton since a neutral acceptor and a free electron are energetically more favorable, because the hole mass is usually considerably heavier than the reduced mass of electron and hole. The absorption and emission lines of A^0X, D^0X and D^+X are often labelled I_1, I_2 and I_3 lines, respectively. The binding energy of an exciton to a neutral donor (acceptor) is usually much smaller than the binding energy of an electron (hole) to the donor (acceptor). The ratio of the two energies depends only weakly on the material parameters and amounts approximately to 0.1 . This fact is known as Heynes rule [60H1, 77H1].

The binding energy of the exciton to the complex depends also on the chemical nature of the complex (known as chemical shift or central cell corrections) and on the surroundings, leading in high resolution spectroscopy to a splitting of the I_i lines. Furthermore, bound exciton complexes may have a certain manifold of excited states due to the various mutual arrangements and envelope functions of the two electrons (holes) and the hole (electron) in the $D^0X(A^0X)$ complex [79D1, 81B1]. We shall meet some examples in Sect. 14.1.

The wavefunctions of excitons bound to shallow centers can be described by a superposition of free exciton wavefunctions in a similar way to that shown in (8.14) for free carriers.

To conclude this section on bound exciton complexes we give mention excitons bound to point defects other than single shallow donors or acceptors like deep centers formed partly by isoelectronic traps (see (8.14)).

Donor–acceptor pairs (Fig. 8.36) can be considered as "polycentric" bound excitons. On the other hand, it has been found that one center can, under certain conditions, bind several excitons. The formation of such multi-exciton complexes is especially favored in indirect semiconductors due to the high degeneracy of the multivalley conduction band and the fourfold degenerate Γ_8^+ valence band [78T1, 89G1]. A review of bound exciton complexes is found in [76D1, 79D1, 04M1].

Finally, it should be mentioned that such complexes also exist in quantum wells. In this case, the energy of the bound exciton depends in addition on the spatial position of the impurity relative to the barriers. The binding energy usually decreases if the impurity is located not in the center of the well but closer to one of the barriers because the wavefunction is pushed away from the impurity [88F1]. This phenomenon results in an additional inhomogeneous broadening of the absorption and emission lines, which then often merge with the tail-states caused by disorder (see below) and/or with the free-exciton line.

9.6 Excitons in Disordered Systems

In our discussion of disordered systems in Sect. 8.15 we saw that disorder leads to the appearance of localized electron and/or hole states.

In a similar way the two-particle complex exciton can be localized in a disordered semiconductor.

If we look to the potential wells and barriers in the valence band e.g., of $CdS_{1-x}Se_x$ (Fig. 8.4) we can envisage two different mechanisms of localization. In very deep potential wells for holes, such a quasiparticle can be localized, and the electron is bound to the localized hole by Coulomb interaction. The other possibility is that we have a wide potential well with dimensions larger than the excitonic Bohr radius. In this case the exciton is localized as a whole. As mentional already in Sect. 8.15, approaches based on percolation theory are used to describe carrier or exciton localization.

In Fig. 9.5 we show schematically the density of localized exciton states for $n_B = 1$. At low energies we start with excitons for which one carrier is localized and the other bound to it by Coulomb attraction. With increasing energy there is a continuous transition to excitons which are localized as a whole, which in turn ends at the transition region to extended exciton states known as the mobility "edge". Above this there are the extended exciton states with properties approaching those of free excitons in ordered materials.

In principle, pictures similar to Fig. 9.5 should hold also for $n_B > 1$. Due to the significant inhomogeneous broadening of the $n_B = 1$ state, however, there is not much chance of identifying higher states of the Rydberg series by optical spectroscopy. Therefore usually only the exciton ground states are considered.

As long as the tailing parameters E_0 (8.64) describing the localized excitons is much smaller than E_g, we can use the effective mass approximation for localized excitons to give the following rules of thumb. Since heavier particles are easier to localize according to (8.64), we will find significant features of localized excitons more frequently in those ionic semiconductors with alloy-

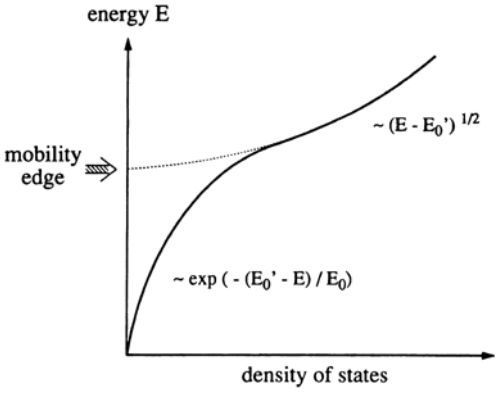


Fig. 9.5. The density of exciton states for $n_B = 1$ in a disordered semiconductor

ing in the anions that form the valence band like $CdS_{1-x}Se_x$ or $ZnSe_{1-x}Te_x$ than in cation-substituted materials such as $Zn_{1-y}Cd_ySe$. In semiconductors with more covalent binding, localization occurs both in the conduction and valence bands, as in α-Si or in $Ga_{1-y}Al_yAs$. However, especially in the latter example, both the electron and hole masses are relatively low and, as a consequence, it is difficult to localize excitons, i.e., the tailing parameter E_0 and the total number of localized exciton states N_0 are small. Another aspect of the same feature is the following: Due to the low effective masses mentioned above, the exciton Bohr radius is much larger in $Al_{1-y}Ga_yAs$ than in $CdS_{1-x}Se_x$. This means that the exciton averages over a larger volume, thus diminishing the effective fluctuation and reducing the tail of localized exciton states.

As a consequence, in $CdS_{1-x}Se_x$ or $ZnSe_{1-x}Te_x$ for x around 0.1 to 0.4, the tail of localized exciton states contains roughly 10^{18} to 10^{19} states per cm^3, while this number may be one or two orders of magnitude smaller for the $Al_{1-y}Ga_yAs$ system (or for cation-substituted materials). More information on localized excitons is given in [85H1, 91C1, 92S1, 92S2, 93C1, 93K1, 01R1, 01T1, 01W1, 04K1, 04K2].

Recently, a very successful model has been developed for the description of exciton localization in bulk semiconductor alloys, which can also be extended for systems of reduced dimensionality. See, e.g. [04K1] and [99K1, 00P1] of Chap. 8. The idea is based on the percolation theory already mentioned in Sect. 8.15 and above. Clusters of atoms of the lower gap material form localization sites. If the wave function of excitons in neighboring clusters overlap, they are assumed to form superclusters. The absorption is governed by ground and excited states of all isolated and superclusters and their phonon side bands. The low temperature and density luminescence is, in contrast, only emitted from the ground states of isolated and/or superclusters. Taking state filling, hopping and relaxation processes into account as well as electrons and holes in the same or in different (super) clusters (so-called distant pairs) this model allows one to describe quantitatively the cw absorption and emission spectra and the dynamics of the latter after pulsed excitation. The model describes qualitatively the S-shaped temperature dependence of the emission maximum observed in many disordered systems [01K1, 01R1, 01T1, 01W1, 04K1].

In quantum wells various types of disorder can contribute to the formation of tails of localized exciton states as already discussed for one-particle states:

- alloy disorder if the well and/or the barrier material is an alloy. The first case obviously has a stronger influence since the probability of finding the electron and hole in the well of a type I structure is larger than that of finding it in the barrier.
- interface roughness, i.e., well-width fluctuations. Usually quantum wells can only be grown with well-width fluctuations of at least one monolayer. These fluctuations of l_z influence the exciton energy via the l_z-dependence of the quantization energy, and via the l_z-dependence of the exciton bind-

ing energy (Fig. 9.4). Usually the first effect is the dominant one. For some recent reviews of this topic see [85H1, 93K1, 04K1] and references therein.

For quantum wells a remarkable evolution has been observed over the last decades concerning the question of exciton localization. In the first euphoric years, many authors claimed that the excitonic features in the emission and absorption spectra are all due to free excitons, because the (MBE-grown) structures are "so pure and the structures so perfect" despite of the facts that it was well-known that the low-temperature luminescence of even the best parent bulk materials was generally dominated by bound exciton complexes and that the low-temperature widths of absorption and emission feature were in the quantum wells frequently in the 5 to 10 meV range, a value far too high to be understood without involving impurity-bound or localized excitons.

Later on, one learned about localization by the various disorder effects in quantum wells mentioned above and presently there is a trend to sell the localization sites as quantum dots. Indeed there is – as already mentioned in (8.15) – no clear and generally accepted criterion to distinguish between a localization site and a quantum island formed by an especially pronounced thickness and/or composition fluctuation of a quantum well. Instead the transition is smooth. But it seems fair to call, e.g., some $Zn_{1-x}Cd_xSe$ or $Ga_xIn_{1-x}As$ regions with x close to unity and considerably larger than of the embedding quantum well (or wetting layer) between ZnSe or GaAs barriers quantum islands, though the lateral in-plane confinement is usually much smaller than the normal confinement caused by the quantum well (see also [01L1] of Chap. 1).

Similar statements and arguments for localization also hold for quantum wires, though localization is less intensively studied in quasi 1d structures. For some examples see, e.g., [91C1, 93C1] or Sect. 15.3.

For systems that are confined in all three dimensions, i.e., quantum dots, the question of localization is irrelevant.

9.7 Problems

1. Calculate the Rydberg energy and the Bohr radius of excitons for some of the semiconductors for which you found the material parameters in the problems of Chap. 8. Compare these with the experimentally determined binding energies and lattice constants, respectively.
2. How many (different) exciton states can be constructed in a semiconductor with zinc-blende (T_d) structure for the principal quantum numbers $n_B = 1, 2$ and 3?
3. Compare the magnitude of the relative splitting between 2s and 2p states in a hydrogen atom (what are the physical reasons?) with the 2s–2p splitting of excitons.
4. Plot the Rydberg series of an idealized three- and two-dimensional exciton and indicate the oscillator strengths.

5. Calculate the (combined) density of states in the continuum of a three- and a two-dimensional exciton in the effective mass approximation. Multiply by the corresponding Sommerfeld enhancement factor.
6. Find in the literature data for the binding energies of the exciton ground state and of the higher states (i.e. $n_{\mathrm{B}}S$ or $n_{\mathrm{B}}P$ states with $n_{\mathrm{B}} \geq 2$), e.g., for GaAs, ZnO, CuCl and Cu_2O and determine for which ones the $1S$ state fits into the hydrogen series with higher states.
7. Show by a semiquantitative guess that the A^-X complex is usually unbound.

References to Chap. 9

[31F1] J. Frenkel: Phys. Rev. **37**, 1276 (1931) and Phys. Z. Sowjetunion **9**, 158 (1936)
[37W1] G.H. Wannier: Phys. Rev. **52**, 191 (1937)
[50H1] M. Hayashi and K. Katsuk: J. Phys. Soc. Japan **5**, 380 (1950) and ibid. **7**, 599 (1952)
[52G1] E.F. Gross and N.A. Karryev: Doklady Academii Nauk USSR **84**, 261 and 471 (1952)
[55H1] H. Haken: Halbleiterprobleme IV, 1 (1955) and Nuovo Cimento **3**, 1230 (1956)
[55S1] L.I. Schiff: Quantum Mechanics, 2nd ed. Mc Graw Hill, New York (1955)
[56G1] E.F. Gross: Nuovo Cimento Supl. **3** (4), 672 (1956) and Sov. Physics Uspekhi **5**, 195 (1962)
[57E1] R.J. Elliot: Phys. Rev. **108**, 1384 (1957)
[58L1] M.A. Lampert: Phys. Rev. Lett. **1**, 450 (1958)
[58M1] S.A. Moskalenko: Opt. and Spectros. **5**, 147 (1958)
[58N1] S. Nikitine: J. Chem. Phys. **55**, 621 (1958) and Progress in Semiconductors **6**, 235 (1962)
[60H1] J.R. Haynes: Phys. Rev. Lett. **4**, 361 (1960)
[62N1] S. Nikitine: Progr. Semicond. **6**, 233, 269 (1962)
[63K1] R.S. Knox: Theory of Excitons, Solid State Phys., Suppl 5, Academic, New York (1963)
[63K2] C.G. Kuper, G.D. Whitfield eds.: Polarons and Excitons, Plenum, New York (1963)
[66S1] M. Shinada, S. Sugano: J. Phys. Soc. Jpn. **21**, 1936 (1966)
[68T1] Y. Toyozawa and J. Hermanson, Phys. Rev. Lett. **21**, 1637 (1968)
[72A1] O. Akimoto, E. Hanamura: J. Phys. Soc. Japan **33**, 1537 (1972) and Solid State Commun. **10**, 253 (1972)
[72B1] W.F. Brinkman, T.M. Rice and B. Bell: Phys. Rev. B **8**, 1570 (1972)
[72K1] L. Kalok and J. Treusch, phys. stat. sol. A **52**, K 125 (1972)
[73D1] M.M. Denisov and V.P. Makarov: phys. stat. solidi (b) **56** 9 (1973)
[73H1] W.T. Huang: phys. stat. sol. (b) **60**, 309 (1973)
[76D1] P.J. Dean, D.C. Herbert, D. Bimberg, W.J. Choyke: Phys. Rev. Lett. **37**, 1635 (1976)
[76L1] R. Lévy et al.: phys. stat. sol. (b) **77**, 381 (1976)
[76S1] B. Stébé, C. Conte: Solid State Commun. **19**, 1237 (1976)

[77B1] D. Bimberg: Festkörperprobleme / Adv. Solid State Phys. **17**, 195 (1977)
[77H1] B. Hönerlage, U. Schröder: Phys. Rev. B **16**, 3608 (1977)
[77H2] H. Haug and E. Hanamura, Phys. Rep. **33C**, 209 (1977)
[78T1] M.L.W. Thewalt: Solid State Commun. **25**, 513 (1978) and references therein
[78U1] R.G. Ulbrich, C. Weisbuch: Festkörperprobleme / Adv. Solid State Phys. **18**, 217 (1978)
[79D1] P.J. Dean, D.C. Herbert: in [79E1] p. 55
[79E1] Excitons, K. Cho ed.: Topics Curr. Phys. Vol. 14 (Springer, Berlin, Heidelberg 1979)
[79R1] U. Rössler: Festkörperprobleme / Adv. Solid State Phys. **19**, 77 (1979)
[79S1] A. Stahl, Ch. Uihlein: Festkörperprobleme / Adv. Solid State Phys. **19**, 159 (1979)
[81B1] G. Blattner et al.: Phys. stat. sol. (b) **107**, 105 (1981)
[81F1] D. Fröhlich: Festkörperprobleme / Adv. Solid State Phys. **21**, 363 (1981)
[81K1] C. Klingshirn, H. Haug: Phys. Rep. **70**, 315 (1981)
[81M1] R.C. Miller et al.: Phys. Rev. B **24**, 1134 (1981)
[83H1] H. Haug and H.H. Kranz: Z. Phys. B **53**, 151 (1983)
[83K1] D.A. Kleinman: Phys. Rev. B **28**, 871 (1983)
[83M1] A. Mysyrowicz et al.: Phys. Rev. B **27**, 2562 (1983)
[83P1] N. Peyghambarian, L.L. Chase and A. Mysyrowicz: Phys. Rev. B **27**, 2325 (1983)
[84G1] R.L. Greene, K.K. Bajaj: Phys. Rev. B **29**, 1807 (1984)
[84W1] J.P. Wolfe and A. Mysyrowicz: Sci. Am. **250**(3) p. 70 (1984)
[85H1] B. Hönerlage et al.: Phys. Rep. **124**, 161 (1985)
[85H2] J. Hegarty, M.D. Sturge: J. Opt. Soc. Am. B **2**, 1143 (1985)
[87C1] R.D. Carson, S.E. Schnatterly: Phys. Rev. Lett. **59**, 319 (1987)
[87F1] D. Fröhlich, K. Reimann and R. Wille: Europhys. Lett. **33**, 853 (1987)
[88F1] C.Y. Fong, I.P. Batra, S. Ciraci (eds.): Properties of Impurity States in Superlattice Semicondcutors, NATO ASI Ser. B, **183**, Plenum, New York (1988)
[88K1] M. Krause, H.-E. Gummlich, U. Becker: Phys. Rev. B **37**, 6336 (1988)
[88K2] E.S. Koteles, J.Y. Chi: Phys. Rev. B. **37**, 6332 (1988)
[89G1] J. Gutowski: NATO ASI Ser. B. **200**, p. 139 (1989) and references therein
[91C1] R. Cingolani et al.: Phys. Rev. Lett. **67**, 891 (1991)
[91H1] X.-F. He, Phys. Rev. B **43**, 2063 (1991)
[91O1] T. Ogawa, T. Takagahara: Phys. Rev. B **44**, 8138 (1991)
[92M1] H. Mathieu, P. Lefebvre, P. Christol: J. Appl. Phys. **72**, 300 (1992) and Phys. Rev. B **46**, 4092 (1992)
[92S1] H. Schwab et al.: phys. stat. sol. (b) **172**, 479 (1992)
[92S2] U. Siegner et al.: Phys. Rev. B **46**, 4564 (1992)
[93C1] R. Cingolani: Phys. Scr. T **49** B, 470 (1993)
[93K1] H. Kalt et al.: Physica B **191**, 90 (1993)
[94E1] EXCON '94, Darwin, Australien SPIE Proc **2362** (1995), EXCON '96, Gohrisch, Germany, M. Schreiber ed., Dresden University Press, Dresden (1996) and Pure and Applied Chemistry **69** (6) (1997), EXCON '98, Boston USA, R.T. Williams and W.M. Yen eds., Electrochemical Society, Pennington (1998), EXCON '00, Osaka, Japan, K. Cho and A. Matsui eds., World Scientific, Singapore (2001), EXCON '02, Darwin, Australia, T. Kobayashi ed., Nonlinear Optics B **29** (4 to 12) (2002), EXCON '04, Crarow, Poland, P. Petelenz ed., J. Lumin., in press

[94L1] A. Lindner: Grundkurs Theoretische Physik, Teubner, Stuttgart (1994)
[94R1] A. Reznitsk et al.: phys. stat. sol. (b) **184**, 159 (1994)
[95F1] G. Finkelstein, H. Shtrikman and I. Bar Joseph: Phys. Rev. Lett. **74**, 976 (1995), Phys. Rev. B **53**, 12593 and R1709 (1996)
[96S1] J. Singh et al: Phys. Rev. B **53**, 15909 (1996)
[97S1] B. Stébé, et al.: Phys. Rev. B **56**, 12454 (1997)
[97S2] A.J. Shields et al.: Phys. Rev. B **55**, R1970 (1997)
[98F1] A. Franceschetti, L.W. Wang and A. Zunger: Phys. Rev. B **58**, 13367 (1998)
[99D1] R. Denschlag, R. v.Baltz: phys. stat. sol. b **215**, 287 (1999)
[01K1] E. Kurtz et al.: Appl. Phys. Lett. **79**, 1118 (2001)
[01R1] A. Reznitsky et al.: Proc. 9th Intern. Symp. On Nanostructures, St. Petersburg, A.F. Ioffe Institute, Russia, (NANO 2001), p. 538 (2001)
[01T1] S.A. Tarasenko et al.: Semiconductor Science and Techn. **16**, 486 (2001)
[01W1] S. Wachter et al., phys. stat. sol. (b) **224**, 437 (2001)
[02B1] M. Betz et al., phys. stat. sol. b **231**, 181 (2002)
[03G1] O. Gogolin et al.: J. Luminesc.**102/103**, 414 and 451 (2003)
[03O1] Optical Properties of 2D Systems with Interacting Electrons, W. Ossau and R. Suris eds. NATO Science Series II **119**, Kluwer, Dordrecht (2003)
[04K1] A. Klochikhin et al.: Phys. Rev. B **69**, 085308 (2004)
[04K2] A. Klochikhin et al., 27^th ICPS, Flagstaff, USA, to be published
[04M1] B.K. Meyer et al., Phys. Stat. Sol. b **241**, 231 (2004)

10

Plasmons, Magnons and some Further Elementary Excitations

Here we will briefly address some other collective excitations in semiconductors and the quasi-particles which result from the quantization of these excitations like plasmons or magnons.

10.1 Plasmons, Pair Excitations and Plasmon-Phonon Mixed States

The excitons presented in Chap. 9 are the energetically lowest elementary excitations of the electronic system of an ideal semiconductor (or insulator). However, if we produce in a semiconductor a large density of free electrons and/or holes, e.g., by doping (Sect. 8.1) or by high (photo-) excitation (Chap. 21), other elementary excitations appear in the electronic system which are partly well known in metals like plasmons.

We consider in the following a semiconductor which contains a large number of electrons (say $10^{17} - 10^{19}\,\mathrm{cm}^{-3}$). Analogous results are found for holes.

The gas of free electrons can perform collective oscillations relative to the positive background of ionized donors. We consider in Fig. 10.1a the three-dimensional situation. A displacement of the electron gas of density n by an amount Δx produces a surface charge density ρ_s

$$\rho_\mathrm{s} = ne\Delta x \tag{10.1a}$$

and an electric field, according to (2.1a)

$$E_x = \frac{ne\Delta x}{\varepsilon\varepsilon_0} \tag{10.1b}$$

provided we can neglect boundary effects in the y and z directions.

This electric field acts back on the electrons, leading to an equation of motion

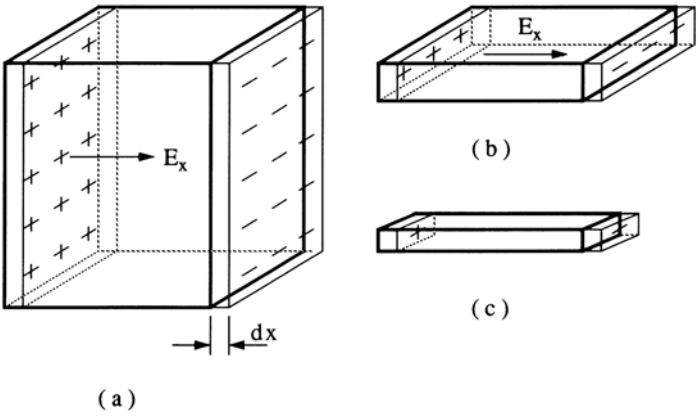


Fig. 10.1. Schematic drawing of a displacement of the electron system in a plasma oscillation in a three- (**a**), two- (**b**) and one-dimensional semiconductor (**c**)

$$eE_x = \frac{e^2 n \Delta x}{\varepsilon \varepsilon_0} = m_\mathrm{e} \frac{\partial^2}{\partial t^2}(\Delta x)\,. \tag{10.1c}$$

The solution of ((10.1c)) is a harmonic oscillation with frequency

$$\omega_\mathrm{PL}^0 = \left(\frac{e^2 n}{m_\mathrm{e} \varepsilon_0 \varepsilon}\right)^{1/2} = \omega_\mathrm{L}, \quad \omega_\mathrm{T} = 0\,. \tag{10.1d}$$

The dielectric "constant" ε which enters (10.1a–d) is ε_s provided $\omega_\mathrm{PL}^0 \ll \omega_\mathrm{L0}$. What happens if this condition does not hold will be discussed in a few moments.

The oscillations described by (10.1a–d) are known as plasma oscillations. Their quantization leads to new quasiparticles which obey Bose statistics and which are known as plasmons. In gases, including the electron gas considered here, only longitudinal oscillations can propagate since gases have a non-vanishing compression but no shear stiffness. See also Sects. 4.1 and 4.3 or 5.6.

Consequently the value of ω_PL^0 in (10.1d) gives the longitudinal eigenfrequency for $\lambda \Rightarrow \infty$ or $\boldsymbol{k} \Rightarrow 0$. The transverse one is zero. If we go to shorter wavelengths we find a weak parabolic dispersion for the plasmons. See [97B1, 80R1, 88R1] or [02D1] in Chap. 1:

$$\omega_\mathrm{PL}(\boldsymbol{k}) = \omega_\mathrm{PL}^0 (1 + ak^2 + \cdots) \tag{10.2}$$

shown in Fig. 10.2a.

For the densities mentioned above, $\hbar\omega_\mathrm{PL}^0$, is situated in the range $10 - 100$ meV for typical semiconductors, i.e., in the (F)IR. This situation is different in metals, where the plasma frequency is usually situated in the VIS or UV part of the spectrum and causes the high reflectivity of this class of materials which extends from the IR up to ω_PL.

For large $\boldsymbol{k}$-vectors, the plasmon modes are strongly damped because they fall in the continuum of one-particle (or, as shown below more precisely, two-particle) intraband excitations. These excitations are shown in Fig. 10.3 where we give the dispersion relation of the conduction band filled up to the Fermi energy E_F by a degenerate electron gas and, for simplicity, $T = 0$. We can produce excitations in this "Fermi sea" of electrons by lifting an electron from a state below E_F into a state above, actually by simultaneously creating a hole in the Fermi sea, below E_F, again resulting in a two particle excitation. In contrast to the two particle excitations in semiconductors and insulators across the gap, this two-particle state occurs in one band. Additionally the Coulomb interaction is substantially screened by the free carriers. We come back to this aspect in connection with electron–hole plasmas in Chap. 21. The excitation energies range from zero to values given by the width of the band, i.e., several eV. Small excitation energies can be created for all wave vectors between zero and $2k_\mathrm{F}$ if the $\boldsymbol{k}$-space is at least two dimensional. For finite excitation energies the shaded range in Fig. 10.2 is accessible.

For strictly two- and one-dimensional systems, the restoring electric field $\boldsymbol{E}$ is not constant but decreases for long wavelengths as λ^{-1} or λ^{-2}, respectively. As a consequence the dispersion of plasmons starts for $k = 0$ at zero energy as shown in Fig. 10.2b, but plasmons in MQW or SL show rather the 3d dispersion relation, if the electrons are displaced in phase for many adjacent wells. For details see [90M1, 98P1, 99A1, 99F1, 00J1, 01O1].

The dielectric function of plasmons can be deduced in different ways, here for bulk materials. We can either use the dielectric function of a Lorentz

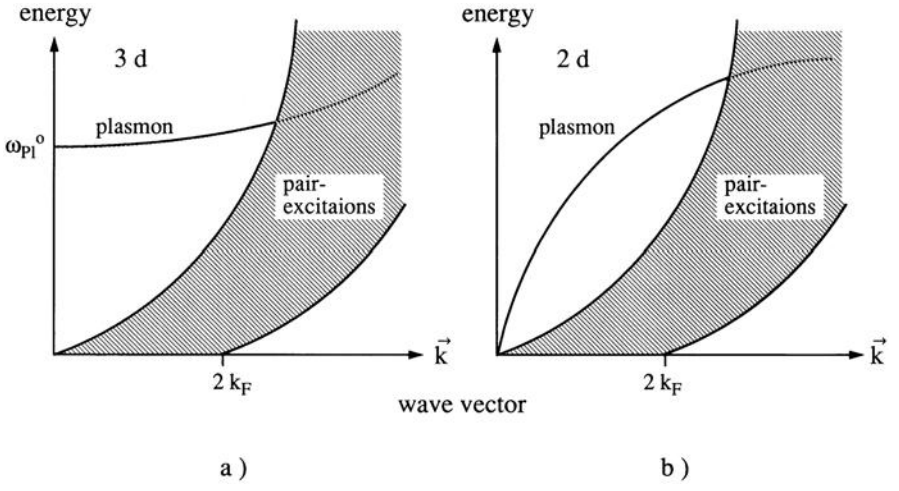


Fig. 10.2. The plasmon dispersion and the range of two-particle excitations in a three- (**a**) and a two-dimensional system (**b**) (schematic)

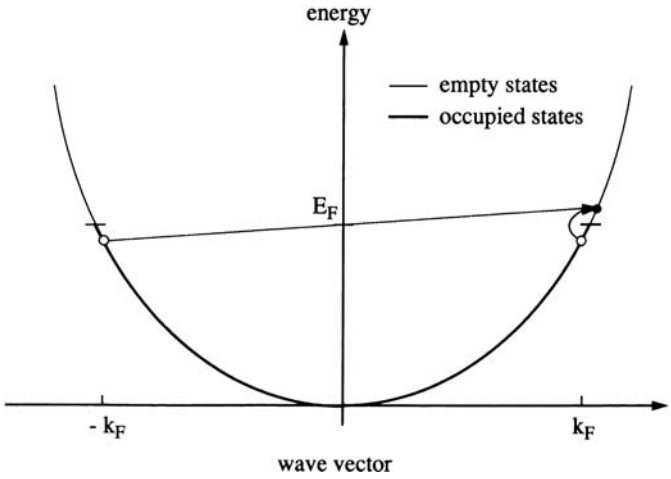


Fig. 10.3. Two-particle excitations within one band with degenerate population

oscillator but setting $\omega_0 = \omega_\mathrm{T} = 0$ resulting in

$$\varepsilon(\omega) = \varepsilon_\mathrm{b} + \frac{\frac{ne^2}{\varepsilon_0 m}}{-\omega^2 - \mathrm{i}\omega\gamma} \tag{10.3}$$

where n is now the density of free carriers and $\gamma = 1/\tau$.

Equation (10.3) again reproduces for vanishing damping the longitudinal plasma frequency of (10.1d) by setting $\varepsilon(\omega_\mathrm{L}) = 0$.

The so-called Drude–Lorentz model considers bound and unbound charges separately. For bound charges the Lorentz-oscillator model of Chap. 4 is used while the equation of motion reads for free carriers for an electric field in the x-direction

$$m\ddot{x} + \frac{m}{\tau}\dot{x} = F_x = eE_x = eE_x^0 e^{-\mathrm{i}\omega t} \tag{10.4a}$$

where $\tau = 1/\gamma$ is the relaxation time, i.e., the time between collisions resulting in a frequency-dependent electrical conductivity $\sigma(\omega)$

$$\sigma(\omega) = \sigma_0 \frac{1}{1 - \mathrm{i}\omega\tau} \tag{10.4b}$$

and to a total dielectric function

$$\varepsilon_\mathrm{total}(\omega) = \varepsilon_\mathrm{Lorentz}(\omega) + \chi_\mathrm{free\,carrier}(\omega)\ . \tag{10.4c}$$

With the help of Maxwells equations.

$$\nabla \times \boldsymbol{H} = \boldsymbol{j} + \dot{\boldsymbol{D}} = \boldsymbol{\sigma E} + \varepsilon_0 \varepsilon_\mathrm{Lorentz} \dot{\boldsymbol{E}}\,, \tag{10.4d}$$

we obtain for $\chi_\mathrm{free\,carrier}(\omega)$

$$\chi_{\text{free carrier}}(\omega) = \frac{\mathrm{i}}{\varepsilon_0 \omega} \sigma(\omega) \tag{10.4e}$$

and for the total dielectric function

$$\varepsilon_{\text{total}}(\omega) = \varepsilon_{\text{Lorentz}}(\omega) + \frac{\mathrm{i}}{\varepsilon_0 \omega} \sigma(\omega) = \varepsilon_{\text{Lorentz}}(\omega) + \frac{\omega_{\text{PL}}^{0^2}}{-\omega^2 - \mathrm{i}\omega/\tau}. \tag{10.4f}$$

The $\boldsymbol{k}$-dependence of ω_{PL}, or in other words the spatial dispersion of this resonance (see Sect. 5.4.1), can be neglected in a good approximation for bulk materials but is crucial for systems of lower dimensionality as seen from Fig. 10.2b.

If all Lorentz oscillators are situated at frequencies much higher (lower) than ω_{PL}, their influence on the plasmon resonance can be approximated by their static (background) dielectric constant ε_{s} (ε_{b}).

We show in Fig. 10.4 the dielectric function of a plasmon resonance for such a situation for vanishing damping and the resulting reflection spectrum. Compare with Fig. 4.3 and 4.6 for $\omega_0 = 0$.

In this way we can incorporate a finite conductivity into a total dielectric function. Evidently, the real (dissipative) part of the conductivity contributes to the imaginary part of $\varepsilon(\omega)$ and vice versa. In some disciplines it is also common to include all contributions in the total conductivity, but the author has strong semantic problems to speak about the complex conductivity of phonons or excitons and recommends speaking of conductivity only when carriers can move freely through the sample.

Surface plasmon modes exist in agreement with the statements in Sect. 5.6 at the frequency for which $\varepsilon(\omega) = -1$.

Quantized (surface) plasmon modes exist in small metal spheres or colloids. They are, e.g., responsible for the beautiful color of gold-ruby glasses.

The plasmon energy increases with the square-root of the carrier density. What happens when ω_{PL}^0 approaches the energies of the optical phonons is

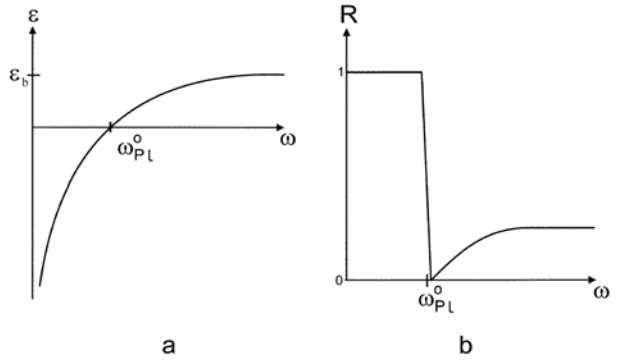


Fig. 10.4. The dielectric function of a plasmon resonance (**a**) and the resulting reflection spectrum (**b**) for vanishing damping

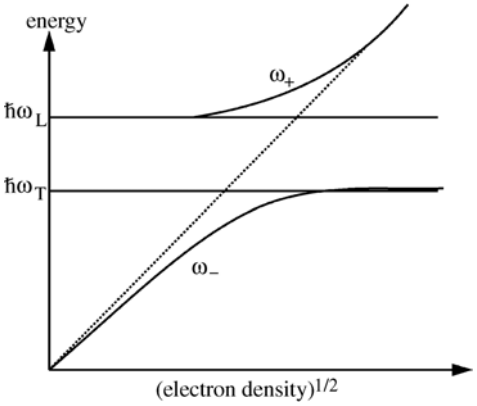


Fig. 10.5. The eigenenergies of the plasmon–phonon mixed state as a function of the square-root of the carrier density n

shown in Fig. 10.5. Plasmons and phonons interact with each other due to their electric fields. This interaction results in another example of the quantum-mechanical non-crossing rule (Chap. 5) here as a function of the carrier density n. The plasmon frequency bends over to the transverse optical phonon branch and reappears above the longitudinal one. These two new branches are usually known as ω_- and ω_+ modes of this plasmon phonon mixed state, respectively. For examples see e.g. [00N1, 02N1] or Sect. 12.3.

This behavior can also be understood classically in the frame of linear optics by just adding the susceptibilities of a plasmon resonance and (for simplicity) of one single Lorentz oscillator, again neglecting spatial dispersion

$$\varepsilon(\omega) = \varepsilon_b + \frac{f}{\omega_0^2 - \omega^2 - \mathrm{i}\omega\gamma} + \frac{\omega_{\mathrm{PL}}^{0^2}}{-\omega^2 - \mathrm{i}\omega/\tau}\,. \tag{10.5}$$

We show in Fig. 10.6 the contribution of both parts separately and their sum for vanishing damping, which explains the appearance of the ω_-mode for situations when $\omega_{\mathrm{PL}}^0 < \omega_0$. Care has to be taken to introduce ε_{b} in (10.5) only once.

For plasmon frequencies considerably above those of longitudinal optical phonons, the background dielectric constant ε_{b} has to be used in (10.1d). For more details see. e.g. [97B1].

Finally it should be mentioned that there are also valenceband plasmons connected with collective excitations of the electron system of a filled valence band. Their eigenfrequencies are situated at energies much larger than E_{g}. Therefore they are not further considered. For details of this aspect see [85E1].

10.2 Magnons and Magnetic Polarons

We discuss next the collective excitations of the spin system, the so-called magnons, existing in ferro-, ferri- and antiferromagentic materials. Semicon-

ductor examples are NiO and MnSe. See, e.g., ref [81M1] of Chapt. 1 and other textbooks on solid state physics.

The material examples show, that magnetic semiconductors are not the most widely used ones, but might become interesting in the future in "spin-electronics" or, i.e., "spintronics." See, e.g., [90M1, 98P1, 99A1, 99F1, 00J1, 01O1] and Chap. 24.

The trivial excitation of a ferromagnetic spin system would be to tilt one spin by an amount $\hbar$ off the magnetization direction. Due to the (exchange) interaction with the other (surrounding) spins and magnetic moments, this excitation has a finite energy. The nontrivial possibility is to tilt every spin by a small amount in a correlated way so that the total angular momentum is also reduced by an amount $\hbar$. This results in a collective excitation, the quanta of which are called magnons. The dispersions relation starts with zero energy for infinite wavelength, i.e., $E(\boldsymbol{k} = 0) = 0$ and shows a parabolic dispersion see Fig. 10.7a. In antiferromagnetic or ferrimagnetic materials there are various magnon branches due to in-phase or antiphase excitation of the spin subsystems. In analogy to phonons, the lowest branch is called the "acoustic" magnon branch, the higher ones "optic" magnon branches though magnons exhibit generally weak or vanishing coupling to the light field. The dispersion relation of magnons is usually investigated by inelastic neutron scattering.

To complete the zoo of quasiparticles relevant to semiconductors, we introduce the concept of magnetic polarons. These are analogous to the usual or phonon polarons of Sect.8.6. Free carriers, i.e., electrons or holes that are accompanied by a magnetization cloud of paramagnetic ions (mostly Mn or Fe) incorporated with a concentration of up to several percent in a nonmagnetic semiconductor, generally II–VI materials like $Cd_{1-x}Mn_xTe$ or $Zn_{1-x}Mn_xSe$,

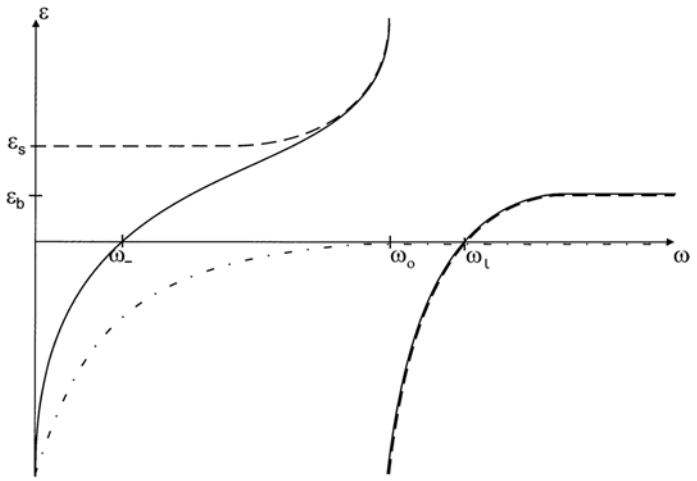


Fig. 10.6. The (undamped) dielectric functions of a Lorentz oscillator (- - -) of a plasma resonance (– · – · –·) and of the sum of both (——) (schematic).

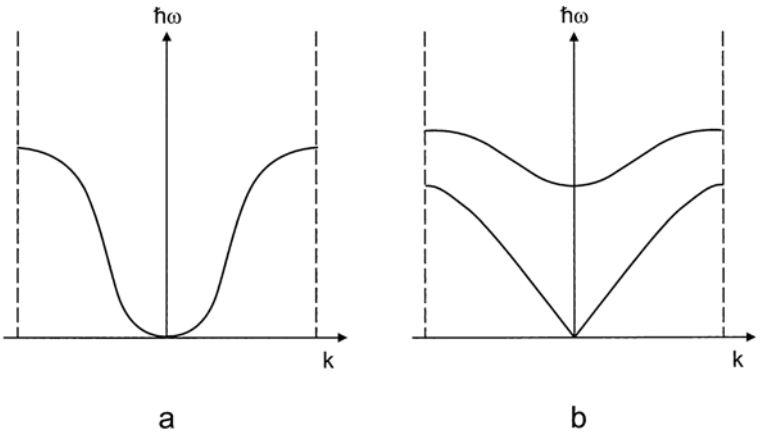


Fig. 10.7. The schematic dispersion relation of magnons in a ferromagnet (**a**) and an antiferromagnet (**b**)

but also III–V compounds like $Ga_{1-x}Mn_xAs$. In the latter group the incorporation of Mn^{2+} results simultaneously in a (usually unwanted) p-type doping. Consequently $G_{1-y}Mn_yAs$ tends to be a metal i.e. it has finite electrical conductivity at $T \Rightarrow 0\,\mathrm{K}$. See Sect. 8.2. The p-doping can be compensated by a codoping with donors. This group of materials is known as diluted or semimagnetic semiconductors (DMS). The idea is the following. While MnSe shows a antiferromagnetic ordering, an alloy with a small percentage of Mn^{2+} or Fe^{3+} does not develop a spontaneous magnetic ordering due to the larger distance of the magnetic centers, but behaves as a paramagnet with very high susceptibility resulting, e.g., in excitonic Zeeman splitting several orders of magnitude larger than in nonmagnetic materials. In some cases the magnetic field resulting from the magnetic moment of a free carrier, e.g., an electron, is able to align the magnetic moments of the surrounding Mn^{2+} ions. The electron can still move freely through the lattice, but it carries with it a magnetization cloud. Therefore these charge carriers are called magnetic polarons.

Some further quasiparticles in solids are, e.g., fluxons, skirmions and composite fermions, but they are of minor importance for the optical properties of semiconductors and therefore not treated here.

10.3 Problems

1. Calculate the plasmon energy $\hbar\omega_{\mathrm{PL}}^0$ for a typical three-dimensional semiconductor ($m_e = 0.1\,m_0$) and $n = 10^{16}$, 10^{17} and $10^{18}\,\mathrm{cm}^{-3}$. Compare with the eigenenergies of optical phonons.
2. Calculate $\hbar\omega_{\mathrm{PL}}^0$ for a metal ($n \approx 10^{22} - 10^{23}\,\mathrm{cm}^3$). Using the knowledge of Chap. 4, consider which value should be taken for the dielectric "constant" ε?

3. What is the origin of the color of some metals like gold or copper? Remember that there are, apart from plasmons, interband-transitions in metals.
4. Why are radio waves in the short wave range (KW) reflected by the upper layers of the atmosphere but not ultrashort waves (UKW)?
5. Would you expect differences in the diffraction pattern of a crystal using X-rays or (thermal) neutrons if these materials are dia-, para-, ferro- or antiferromagnetic?
6. Qualitatively, what would be the temperature dependence of the saturation magnetisation of a ferromagnet if magnons do exist or do not exist? Compare your hypothesis with the information in a textbook.
7. Why are surface plasmons important for the spectral efficiency of a metal-covered diffraction grating?
8. Make a sketch of the dielectric function of plasmon-phonon mixed states for $\omega_{\mathrm{PL}}^{0} > \omega_0$.

References to Chap. 10

[82H1] R. Höpfel et al., Surf. Science **113**, 118 (1982)
[83P1] A. Pinczuk and J.M. Worlock, Physica **117/118** B, 637 (1983)
[85E1] E. Egri, Phys. Rep. **119**, 363 (1985)
[86H1] D. Heitmann, Surface Science **170**, 332 (1986)
[89H1] M. Helm et al., Phys. Rev. Lett. **63**, 74 (1989)
[90M1] U. Merkt, Festkörperprobleme/Advances in Solid State Physics **30**, 70 (1990)
[91E1] Th. Egeler, Festkörperprobleme/Advances in Solid State Physics **31**, 315 (1991)
[97B1] R.v. Baltz in [81A1]h of Chap. 1 p. 303 and textbooks on solid state physics listed also in Chap. 1
[98P1] G.A. Prinz, Science **282**, 1660 (1998)
[99A1] D.D. Awschalom and J.M. Kikkawa, Physics Today, June 1999, 33
[99F1] R. Fiederling et al., Nature **402**, 6783 (1999)
[00J1] C. Joachim, J.K. Gimzewski and A. Aviram, Nature **408**, 541 (2000)
[00N1] M. Nagai and M. Kuwata-Gonokami, Phys. Stat. Sol. b **221**, 261(2000) and J. Luminesc. 233 (2002)
[02N1] M. Nagai, K. Ohkawa and M. Kuwata-Gonokami, Appl. Phys. Lett. **81**, 484 (2002)
[01O1] M. Oestreich et al., Festkörperprobleme/Advances in Solid State Physics **41**, 173 (2001)

11

Optical Properties of Phonons

Having presented the optical properties of a system of model oscillators and the elementary excitations in semiconductors, we shall now start to put these two parts together. As a first example we investigate the optical properties of phonons. We start with the properties of bulk materials.

11.1 Phonons in Bulk Semiconductors

As already mentioned, optical phonons can couple strongly to the electromagnetic field if the solid has at least partly ionic binding. The $\boldsymbol{k}$ dependence of the eigenfrequency is rather weak and covers generally only some ten meV or even less over the whole Brillouin zone. It is completely negligible if we concentrate on the region of $\boldsymbol{k}$ vectors reached in the IR or VIS part of the spectrum, i.e., $k < 10^6\,\mathrm{cm}^{-1}$ as seen from Figs. 7.15, 7.16, bearing in mind that the first Brillouin zone extends up to $\boldsymbol{k}$ values of around $10^8\,\mathrm{cm}^{-1}$.

We can therefore treat the phonon polariton according to Chaps. 4 and 5, neglecting spatial disperison remembering, however, that due to the small but finite spatial dispersion, a transverse branch extends over the hole Brillouin zone and, if appropriate (see Fig. 5.3), also a longitudinal one.

11.1.1 Reflection Spectra

In Fig. 11.1a we show the reflection spectrum of CdS in the IR around the phonon resonance. Spectra of other more or less ionic materials look very similar. For further examples see e.g. [81M1,82L1,90K1,96Y1,98R1] of Chap. 1 and the references given therein. We can clearly see the reststrahlbande i.e. the stop band between ω_T and ω_L as we predicted in Fig. 4.6. The reflectivity reaches values above 0.9 at room temperature and for the polarization $\boldsymbol{E} \parallel \boldsymbol{c}$ where the Γ_1 phonons are dipole allowed (see Fig. 11.1a). The spectrum of the reststrahlenbande around $250\,\mathrm{cm}^{-1}$ looks very similar to the one calculated for the model of Lorentz oscillators without spatial dispersion in Fig. 4.6 (the

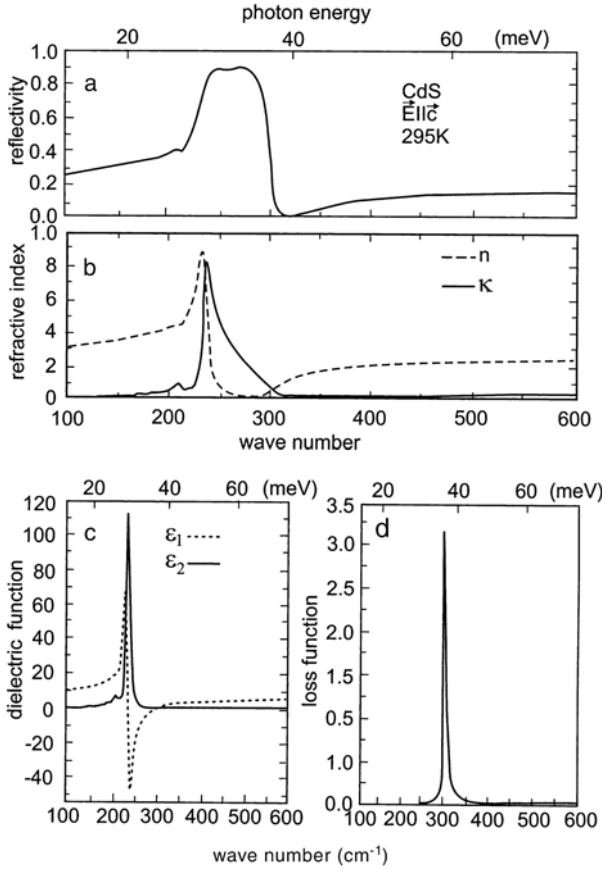


Fig. 11.1. (**a**) The reflection spectrum of the phonon stop band (reststrahlbande) in CdS at room temperature for $\boldsymbol{E} \parallel \boldsymbol{c}$ (**b**) Real and imaginary parts of the complex index of refraction (**c**) Real and imaginary parts of the dielectric function (**c**) The loss function $\mathrm{Im}\{-1/\varepsilon(\omega)\}$ [97G1]

small feature around $203\,\mathrm{cm}^{-1}$ is a multi-phonon resonance). In Fig. 11.1b we show also the real and imaginary parts of the complex index of refraction $\tilde{n}\,(\omega) = n\,(\omega) + i\kappa\,(\omega)$ (b), the real and imaginary parts of the dielectric function $\varepsilon\,(\omega) = \varepsilon_1\,(\omega) + i\varepsilon_2\,(\omega)$ (c) and the so-called loss function $\mathrm{Im}{-1/\varepsilon(\omega)}$ (d), which describes the energy loss of electrons transmitted through a thin sample. The transverse and longitudinal eigenfrequencies coincide very well with the maxima of $\varepsilon_2\,(\omega)$ and $\mathrm{Im}\{-1/\varepsilon(\omega)\}$, respectively.

The spectra in Fig. 11.1b to d have been deduced from the reflection spectrum with the help of the Kramers–Kronig relations and other formulae given in Chap. 6.

Also, these curves coincide nicely with the results of our model calculations in Figs. 4.4 and 4.5. Transmission spectra of thin samples also agree with

Fig. 11.1as shown, e.g., in [97G1]. This agreement allows us to claim good understanding of the optical properties of phonons. It should be remembered here, that optical phonons in semiconductors with purely covalent binding like Si or Ge have, at $k = 0$, zero longitudinal-transverse splitting and zero oscillator strength, as shown in Fig. 7.16. Furthermore, it should be noted that in samples with (partly) ionic binding, optical phonon modes may exist, which do not carry a dipole moment and therefore do not couple to the radiation field, i.e., they are not optically active.

Coherent optical phonon packets have been reported in [98B1].

11.1.2 Raman Scattering

Raman scattering with phonons is the inelastic scattering of light under emission or absorption of an optical phonon or a phonon polariton for the case, that the phonon mode under consideration is both Raman and IR active. Raman active phonons modulate the polarizability of the crystal while infrared active phonons modulate the electric dipole moment. In crystals with inversion symmetry, optical phonons are either Raman or IR active. The selection rules thus complement each other. In samples without inversion symmetry, optical phonons may be both Raman and IR active.

In the weak coupling picture one can also say that a photon creates virtually an excition or an electron–hole band-to-band transition and is scattered by emission or absorption of an optical phonon, but this model fails in some respects as we shall see below.

In the Raman-scattering process energy and (quasi-)momentum have to be conserved, i.e.,

$$\hbar\omega_{\mathrm{R}} = \hbar\omega_{\mathrm{i}} \pm \hbar\Omega_{\mathrm{Ph}}\,, \boldsymbol{k}_{\mathrm{R}} = \boldsymbol{k}_{\mathrm{i}} \pm \boldsymbol{k}_{\mathrm{Ph}}\,, \tag{11.1}$$

where the index i stands for the initial or incident light quantum, e.g., a (phonon- or excitonlike) exciton polariton, R for the Raman signal and Ph for the created or annihilated phonon (polariton). Usually one chooses ω_{i} and ω_{R} in the transparent spectral region of the semiconductor well above the phonon resonance and below the exciton resonances. Often $\hbar\omega_{\mathrm{i}}$ is determined by the emission lines of a readily available laser such as an Ar^+ laser.

Equation (11.1) results in values around $10^5\ \mathrm{cm}^{-1}$ for $\boldsymbol{k}_{\mathrm{i}}$ and $\boldsymbol{k}_{\mathrm{R}}$. In a backward or 90° scattering geometry $\boldsymbol{k}_{\mathrm{Ph}}$ is also around $10^5\ \mathrm{cm}^{-1}$ and thus clearly lies on the phonon-like part of the dispersion relation (see e.g., Fig. 11.2) or on the longitudinal branch. In Fig. 11.2a we plot the dispersion of acoustic phonons, of transverse optical phonon polaritons and of a longitudinal optical phonon from the position of the incident (exciton-) polariton $\hbar\omega_{\mathrm{i}}$; $\boldsymbol{k}_{\mathrm{i}}$ to the left (backward scattering) and into the negative energy direction (representing the "–" sign in Stokes emission). In Fig. 11.2 we also give typical Raman spectra.

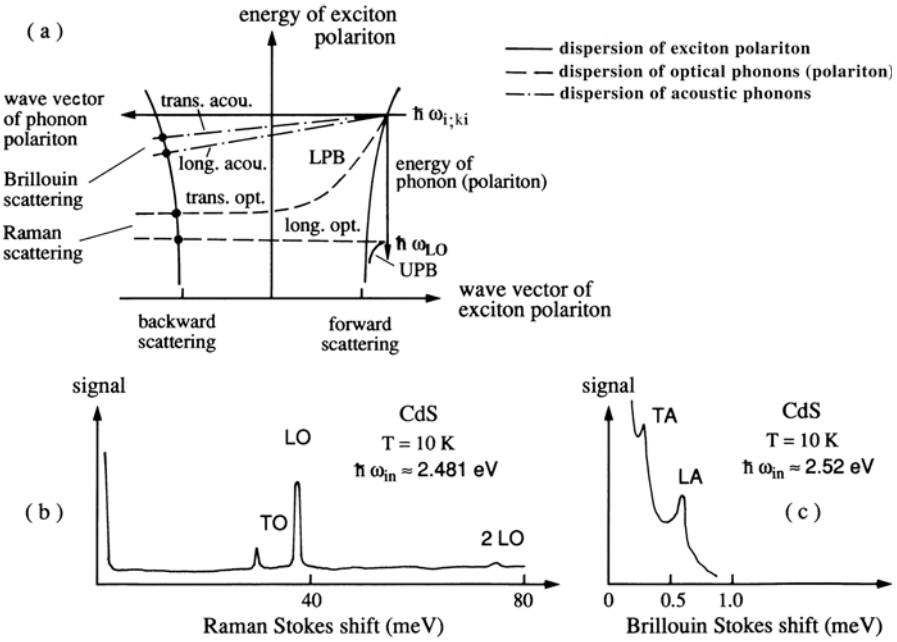


Fig. 11.2. Schematic drawing of energy and momentum conservation in (11.1) for a Raman scattering process (Stokes emission) in a backscattering configuration (**a**), and typical spectra for Raman (**b**) and Brillouin scattering (**c**). Note the different scales on the x-axes in (**b**) and (**c**) [71M1, 80B1, 94C1]

This case is covered by the weak-coupling picture. The selection rules are given in a classical description by the Raman tensor (see, e.g., [90K1, 96Y1] of Chap. 1) and in quantum mechanics by the transition matrix elements. The group theory presented in Chap. 26 gives in this case preliminary information about which transitions are allowed and which are not. We do not want to go into details on this topic which is treated at length, e.g., in the above cited references but cite here only the easily intelligible formula for the intensity ratio of Stokes to anti-Stokes emission, i.e., scattering under emission or absorption of a phonon and assuming that the density of final states for Stokes and anti-Stokes emissions are practically identical

$$\frac{I_\mathrm{s}}{I_\mathrm{a}} = \left(\frac{\omega_\mathrm{i} - \Omega_\mathrm{Ph}}{\omega_\mathrm{i} + \Omega_\mathrm{Ph}}\right)^4 \cdot \frac{N_\mathrm{Ph} + 1}{N_\mathrm{Ph}} = \left(\frac{\omega_\mathrm{i} - \Omega_\mathrm{Ph}}{\omega_\mathrm{i} + \Omega_\mathrm{Ph}}\right)^4 \cdot \exp\left\{\frac{\hbar\Omega_\mathrm{Ph}}{k_\mathrm{B}T}\right\}, \qquad (11.2)$$

which just reflects the fact that the phonon emission is proportional to $(N_\mathrm{ph} + 1)$ and the absorption proportional to N_Ph, where N_Ph is the phonon occupation number of the respective mode.

The first term incorporates the emission probability proportional to ω^4. This factor is usually close to unity because $\omega_\mathrm{i} \gg \Omega_\mathrm{Ph}$.

11.1.3 Phonon Polaritons

If we go now from a backward (or 90°) Raman scattering geometry to forward scattering we see that $\boldsymbol{k}_{\mathrm{Ph}}$ can become small and can fall for IR and Raman active modes into the transition region from phonon-like to photon-like polaritoris. Measurements in this geometry allow one to measure the dispersion of the phonon polariton rather directly. Beautiful experiments of this type, described, e.g., in [65H1, 75C1, 80C1], are performed by collecting the scattered light on the entrance slit of a spectrometer so that the height on the slit is a measure of $\boldsymbol{k}_{\mathrm{i}} - \boldsymbol{k}_{\mathrm{R}}$ (Fig. 11.3a). The wavelength dispersion of the spetrometer then gives the Ω axis. We show an example schematically in Fig. 11.3b. Figure 11.4 gives the dispersion relation for the phonon polariton in GaP reconstructed from this type of experiments, which of course, has to take into account the refraction of the beams at the surface of the sample in contrast to the simplifications in Fig. 11.3a. These experiments can obviously only be understood in the strong-coupling or polariton picture.

In another group of experiments, the dispersion of phonon polaritons has been investigated or verified by an inelastic scattering of two light quanta in which the differences of the wave vectors and of the quantum energies match a point on the dispersion curve of the phonon polariton [98B1, 99R1, 01W1, 02C1].

To conclude our discussion of Raman scattering in bulk semiconductors, we want to mention that Raman processes are also possible in which two phonons are created (or annihilated), e.g., due to the strong anharmonicity of the potential between atoms resulting in phonon–phonon interaction mentioned in connection with (7.47a–c). If this is a two-step process, in the sense that the Raman polariton ω_{R} in (13.1) undergoes a second Raman scattering process, then both phonons have small wavevector compared to the size of the Brillouin zone. If, on the other hand, both phonons are emitted simultaneously, only the sum of the phonon wavevectors has to fulfill (13.1b). The individual phonon can come from any part of the Brillouin zone. Consequently

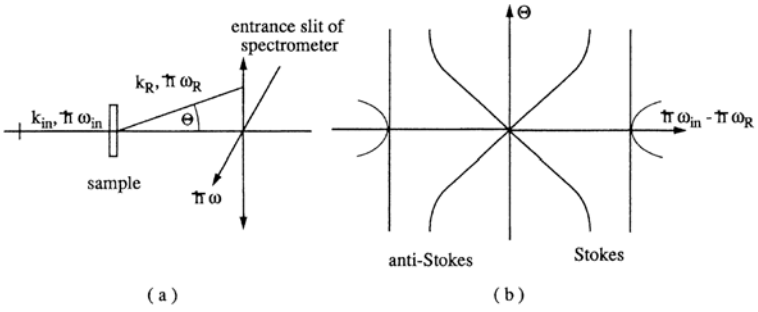


Fig. 11.3. Schematics of a Raman scattering experiment in the forward direction with a spectrometer (a) and the visualization of the relation between Θ and $\hbar\Omega$ of **a** in the output plane of the spectrometer (**b**)

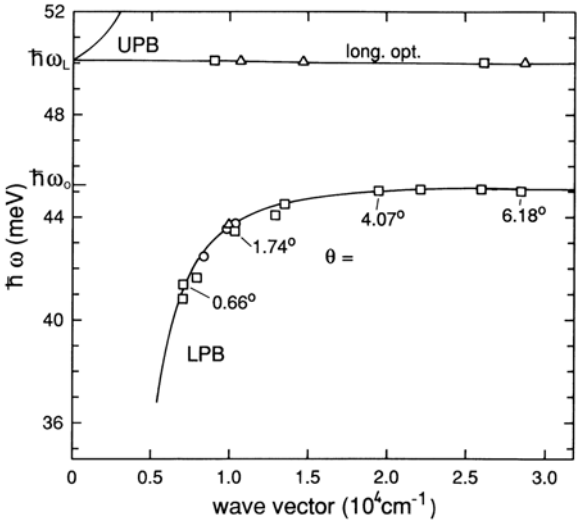


Fig. 11.4. The dispersion of the phonon polariton in GaP, measured by the method shown in Fig. 11.3 [65H1]

the Raman spectrum then reflects to a certain extent the density of states of the phonons. Details of such processes are given in [79S1, 94C1, 03H1].

For the direct observation of the THz emission of optical phonons see [92K1] and references therein and for the influence of isotopes [93C1].

11.1.4 Brillouin Scattering

Brillouin scattering is the analog of Raman scattering for acoustic phonons. Because of the rather flat dispersion relation of acoustic phonons, with a slope given by the velocity of sound (instead of c/n), one finds even in a back cattering configuration only much smaller (anti-) Stokes shifts, which are usually $\leq 1\,\mathrm{meV}$ (Fig. 11.2c). Since the coupling of acoustic phonons to photon-like exciton polaritons is also much weaker than for optical phonons, some high-resolution techniques and an efficient suppression of stray light are necessary to detect Brillouin scattering, as can be seen by comparing the abscissa of Figs. 11.2b and c. Since the dispersion relation of acoustic phonons is linear in the range of interest, (11.1) can be rewritten for Brillouin scattering in a backscattering configuration as

$$\frac{\Omega_{\mathrm{Ph}}^{\mathrm{LA,TA}}}{v_{\mathrm{s}}^{\mathrm{LA,TA}}} = \frac{\omega_{\mathrm{i}} - \omega_{\mathrm{R}}^{\mathrm{LA,TA}}}{v_{\mathrm{s}}^{\mathrm{LA,TA}}} = 2\frac{\omega_{\mathrm{i}}}{c} n(\omega_{\mathrm{i}}) \tag{11.3}$$

if we assume that we are so far away from the exciton resonance that the real part of the refractive index $n(\omega)$ does not change significantly over the Brillouin shift. In this case either $n(\omega)$ or v_{s} can be determined from Brillouin

scattering if the other quantity is known. If ω_i approaches the resonance of the exciton polariton, Brillouin scattering is an efficient means of $\boldsymbol{k}$-space spectroscopy of these resonances. We come back to this aspect in Sect. 13.1.4

11.1.5 Surface Phonon Polaritons

We already mentioned in Sect. 5.6 that surface polariton modes exist in the range between the transverse and longitudinal eigenmodes. These modes can be observed by attenuated total reflection (ATR), i.e., by coupling the evanescent wave of Fig. 3.3 to the material under investigation. More details of this technique are given in Chap. 25.

Here we show in Fig. 11.5a ATR spectra and in b the dispersion relation of the surface phonon polariton in GaP. The agreement with the schematic drawing of Fig. 5.14a is obvious. For more details see [74B1] and references given therein.

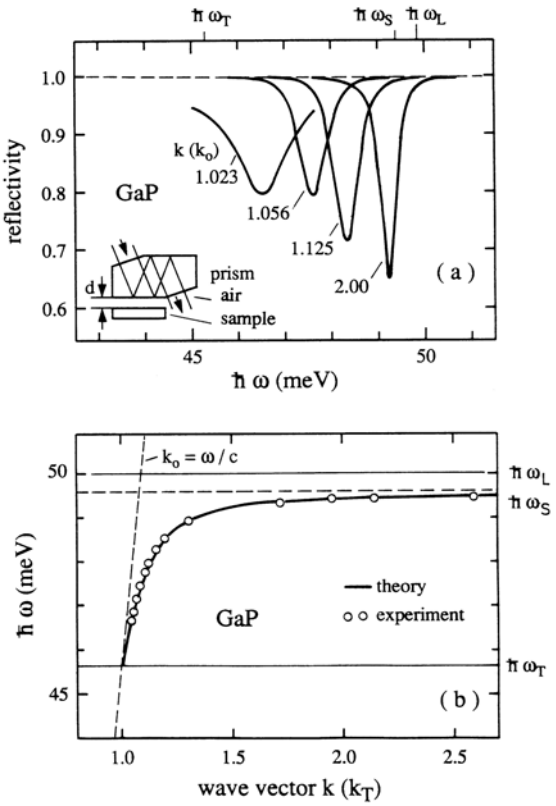


Fig. 11.5. Spectra of attenuated total reflection (**a**) in the region of the optical phonon resonances of GaP and the resulting dispersion relation of the surface phonon polariton (**b**). Compare with Fig. 11.4 for the bulk modes. According to [72M1]

11.1.6 Phonons in Alloys

Many semiconductors form alloys like $Ga_{1-y}Al_yAs$, $CdS_{1-x}Se_x$ or $Zn_{1-y}Mg_yS_{1-x}Se_x$. See, e.g., [95C1] of Chap. 1 or Sects. 7.8 and 8.15. In some cases the alloys exist for the whole range $0 \leq x,\ y \leq 1$. In other cases, there is a miscibility gap. We concentrate here on ternary alloys of two binary compounds. There is a simple rule of thumb for how the phonon modes behave in an alloy: If the equivalent branches of the two parent binary compounds, e.g., the LA or the TO, overlap energetically, there is a smooth transition of the phonon properties from one limiting case (e.g., $x = 0$) to the other ($x = 1$). This is the so-called amalgamation type. If, on the other hand, the two equivalent branches do not overlap energetically, the phonon modes remain essentially as they are in the two parent binary compounds, e.g., CdS and CdSe, but their weights, e.g., their oscillator strengths, vary continuously with x or y. This is the so-called persistent mode type. The reason for this rule is essentially the coupling between oscillators with approximately equal or clearly different resonance frequencies.

The acoustic phonons start in all materials at $\boldsymbol{k} = 0$ and $E = 0$ so they are always of the amalgamation type, i.e., the longitudinal and transverse sound velocities vary in alloys continuously with composition.

For optical phonons both cases can appear. $CdS_{1-x}Se_x$ belongs to the persistent mode type concerning the optical phonons. We show in Fig. 11.6a the composition dependence of the TO and LO eigenenergies. One sees clearly that the frequencies remain essentially unchanged but the longitudinal-transversal splitting Δ_{LT} which is proportional to the oscillator strength varies continuously and complementary with composition. Figure 11.6 shows data for

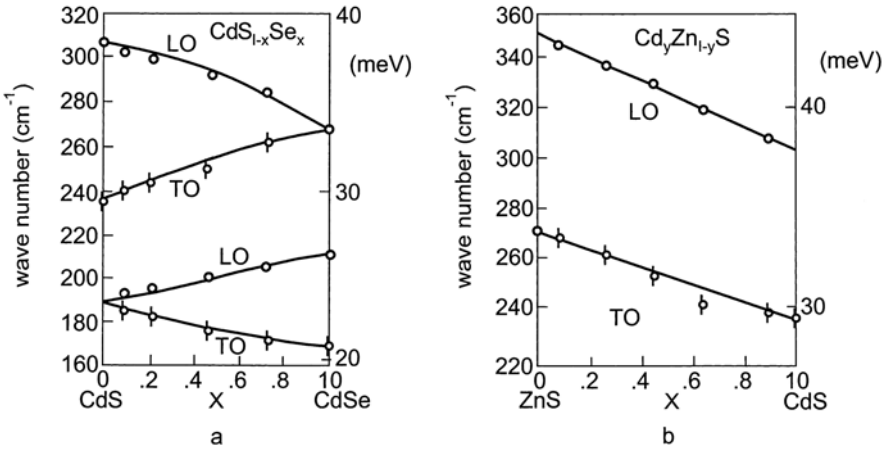


Fig. 11.6. The composition dependence of the optical phonon frequencies for a persistent mode, type alloy, here $CdS_{1-x}Se$ (**a**) and an amalgamation, type alloy, here $Zn_{1-y}Cd_yS$ (**b**) [68C1]

$Cd_{1-y}Zn_y$. The TO and LO phonon energies shift continuously with composition corresponding to the amalgamation type. Data have been deduced from Raman and IR spectroscopy.

Isotopes of one element have per definition different masses. For phonons, the eigenfrequencies depend on the mass of the atoms (see Sect. 7.1). Thus isotope mixtures in a solid already present some (weak) alloying or disorder, contributing to a broadening of phonon resonances. The narrowing of these resonances by materials containing only one isotope has been beautifully demonstrated in [93C1, 94C2].

11.1.7 Defects and Localized Phonon Modes

As a last example of the optical properties of phonons in bulk materials we consider the observation of localized phonon modes e.g. in absorption. The example is a GaAs crystal doped with Si which can occupy Ga or As sites as Si_{Ga} and Si_{As} and thereby act as donor or acceptor, respectively. Additionally the samples have been codoped with Li to keep the concentration of free carriers low since the Si acceptors and donors do not compensate completely. The absorption spectrum of this system in the energy range above the optical bulk phonon modes is shown in Fig. 11.7. Most of the peaks can be identified as local phonon modes of the centers or complexes indicated by arrows. Since Si and Li have considerably lower atomic masses than Ga or As and since the force constants are not too much lower, it is not surprising that the localized phonon modes appear at frequencies above the bulk modes. The isotope shift introduced by ^{7}Li and ^{8}Li is clearly visible in further data of [71S1]. More information on this topic may be found in [89J1, 94R1, 02K1] and references given therein. Localized phonon modes can also be observed in Raman scattering or in luminescence or absorption in the visible as side-bands. See, e.g., [81A1]a of Chap. 1.

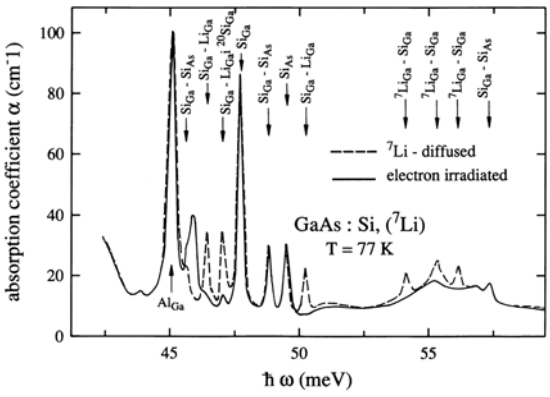


Fig. 11.7. Absorption spectrum of a GaAs: Si, Li, doped sample, showing localized phonon modes [71S1]

11.2 Phonons in Superlattices

Phonons in heterostructures like superlattices have been investigated for many systems including AlAs/GaAs or CdS/CdSe. See, e.g., [01L1] of Chap. 1 or [72C1, 79S1, 97C1, 98G1, 99D1, 01D1] and references given therein. We select here as an example the quaternary system CdS/ZnSe since backfolded acoustic, confined optic and interface phonons have been observed.

11.2.1 Backfolded Acoustic Phonons

The rules for backfolding or confinement of phonons in a superlattice are essentially the same as for amalgamation and persistent mode type in alloys (see Sect. 11.1.6 above and Sect. 7.9).

Consequently, backfolding always applies for acoustic phonons. For the above-mentioned CdS/ZnSe system we show in Fig. 11.8 the calculated dispersion of backfolded longitudinal acoustic phonons, which can be observed in the geometry used in [99D1, 01D1] for two different samples and the first three backfolded LA modes (so-called FLAPS) in the corresponding Raman spectra. The experimental data are compared in Fig. 11.8 with calculated dispersion relations. Excellent agreement has been found.

Under favorable conditions it is even possible to observe backfolded acoustic phonons directly in the IR. An example of a CdS/CdSe SL is given in [98G1]. References for backfolded acoustic phonons in other systems like GaAs/AlAs or CdS/ZnS superlattices are found, e.g., in [72C1, 93R1, 99D1].

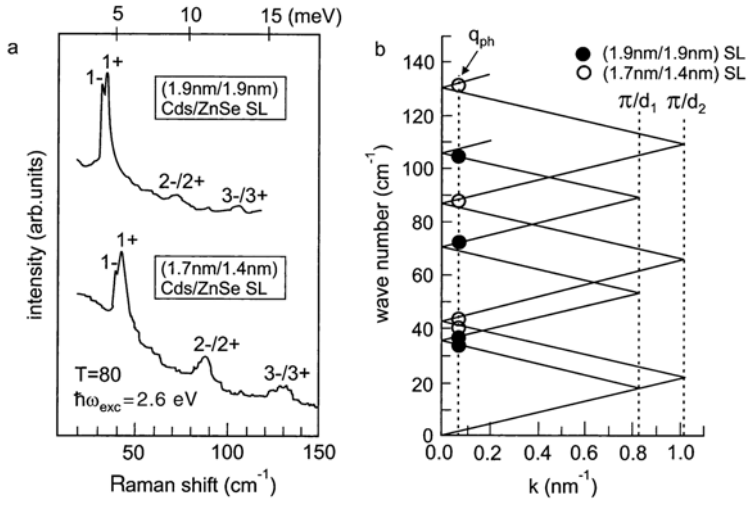


Fig. 11.8. Raman spectra of the first three backfolded longitudinal acoustic phonons (FLAPS) in two different CdS/ZnSe superlattices (**a**) and their calculated dispersion relation together with the data points from (**a**) (**b**) [01D1]

11.2.2 Confined Optic Phonons

The CdS/ZnSe system is a persistent mode type. Consequently there are confined optical phonon modes in the two different sublattices. In Fig 11.9a the displacement patterns of the first two confined LO phonon modes are shown for the CdS and for the ZnSe layers. Due to the mismatch of the eigenfrequencies, the amplitudes decay extremely fast beyond the boundary. In the Raman spectrum are well-resolved the first two confined LO phonon modes in ZnSe, while those in CdS are more broadened and the corresponding structure is washed out. The polarization-dependent selection rules are obvious, i.e. the odd confined phonon modes are seen preferentially in off-diagonal polarisa-

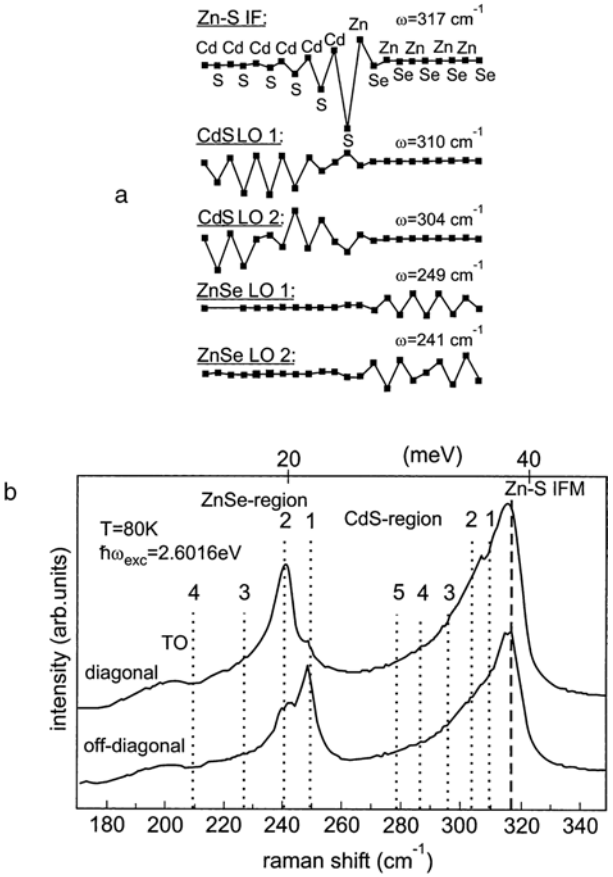


Fig. 11.9. Calculated atomic displacement patterns for the first two confined longitudinal optical phonon modes in a $(CdS)_6/(ZnSe)_5$ superlattice and of the ZnS interface mode (**a**) and the Raman spectra for parallel (diagonal) and orthogonal (off-diagonal) polarization with the calculated frequencies indicated by dashed vertical lines (**b**) [01D1]

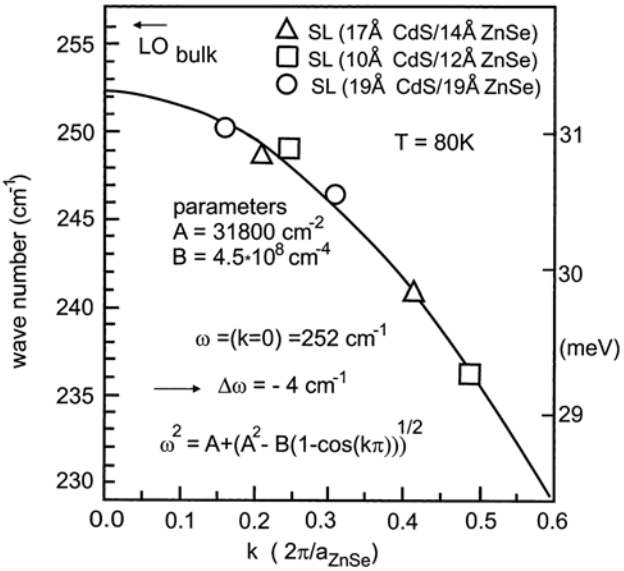


Fig. 11.10. The dispersion relation of LO phonons of ZnSe over one half of the first Brillouin zone deduced from confined optic phonons in a CdS/ZnSe superlattice [01D1].

tion and the even ones in diagonaql polarisation. The calculated energies are marked by the dashed lines.

By measuring the eigenfrequencies of confined phonons of different order and in superlattice layers of different thicknesses it is possible to investigate optically the phonon dispersion over a substantial part of the first Brillouin zone, including possibly some minor shifts or modifications in strained layer superlattices due to the phononic deformation potentials. An example is shown for ZnSe in Fig. 11.10. The dispersion relation is shifted by $4\,\mathrm{cm}^{-1}$ due to strain as compared to bulk ZnSe. This technique is much cheaper than inelastic neutron scattering and can be performed in almost every well-equipped laboratory for optical spectroscopy and not only at one of the few thermal neutron sources.

More examples of this type of experiment and references are found, e.g, in [93R1, 94C1, 97V1] or [96Y1, 01L1] of Chap. 1.

11.2.3 Interface Phonons

In some quaternary superlattices, like CdS/ZnSe or BeTe/ZnSe, so-called interface phonons may appear. These are phonon modes that occur, as the name suggest, at the interface between two materials with amplitudes decaying rapidly on both sides. In this sense they correspond to the surface (polariton) modes of Sect. 5.6.

In the CdS/ZnSe system, two bonds can appear at the interface, which are not present in either of the two compounds, namely CdSe and ZnS bonds. While CdSe interface phonon modes are difficult to detect since their frequency is rather low due to the heavy masses of Cd and Se and they might consequently merge in some other confined or backfolded modes, the situation is different for the ZnS mode, which involves the light cation and anion masses. The peak with the highest energy in the Raman spectrum of Fig. 11.9b is attributed to this interface mode. The displacement pattern is shown in Fig. 11.7a and the excellent coincidence of the calculated and measured energies supports this interpretation.

11.3 Phonons in Quantum Dots

In quantum dots, phonons may be confined in all three directions of space similarly to charge carriers or excitons. One peculiar consequence is that acoustic phonons show a discrete energy spectrum starting at finite energy. There are modes where, e.g., one half of a spherical quantum dot rotates against the other and others where a sphere oscillates adopting the shape of an American football. The appearance of the various modes is obviously closely linked to the boundary conditions of the dot. For details see, e.g., [92N1, 93W1, 96W1]. These modes can be observed, e.g., as satellites in resonant photoluminescence or photoluminescence excitation spectroscopy.

The confined acoustic phonons shift in energy according to R^{-1} where R is the dot radius and can thus be distinguished from other splitting mechanisms with a comparable order of magnitude but a different size dependence. The singlet-triplet splitting also observed in [96W1] varies, e.g., with powers $R^{-2.5}$ to R^{-3}.

11.4 Problems

1. Try to find more reflection spectra like that in Fig. 11.1b in the literature. Deduce $\hbar\omega_{\mathrm{T}}$, Δ_{LT}, ε_{s} and ε_{b} from these spectra and compare with values in the literature.
2. Show that the (eventually only weak) dependence of ω on $\boldsymbol{k}$ is important to explain the experimental fact that the TO and LO phonon modes can be followed through the whole Brillouin zone by neutron scattering. Compare for the explanation with Fig. 5.1 and (6.2) for vanishing and finite damping.
3. Why is the phonon spectrum of high T_{c} superconductors so complex?
4. Which trend would you expect for the zone boundary LA and TA phonons when going from ZnO via ZnS and ZnSe to ZnTe. Compare with data in the literature. What do you expect for zone center optic phonons?

5. Can you give qualitative arguments for why the optical phonons in $CdS_{1-x}Se_x$ and $Zn_{1-y}Cd_yS$ are of the persistent and amalgamation type, respectively? Consider the atomic masses of the oscillating atoms.

References to Chap. 11

[65H1] Ch. Henry and J.J. Hopfield, Phys. Rev. Lett. **15**, 964 (1965)
[68C1] I.F. Chang and S.S. Mitra, Phys. Rev. B **172**, 924 (1968)
[71M1] R.M. Martin and T.C. Damen, Phys. Rev. Lett. **26**, 86 (1971)
[71S1] W.G. Spitzer, Festkörperprobleme / Adv. Solid State Phys. **11**, 1 (1971)
[72C1] C. Colvard et al., Rev. Lett. **28**, 811 (1972)
[72M1] N. Marshall and B. Fischer, Phys. Rev. Lett. **28**, 811 (1972)
[74B1] G. Borstl, H.J. Falge and A. Otto, Surface and Bulk Phonon Modes Observed by Attenuated Total Reflection, Springer Tracts in Modern Physics **74**, Springer, Berlin, Heidelberg (1974)
[75C1] R. Claus, L. Merten and J. Brandmüller, Light Scattering by Phonon-Polaritons, Springer Tracts Modern Phys. **75**, Springer, Berlin, Heidelberg (1975)
[79S1] R.L. Schmidt et al., Phys. Rev. B **20**, 3345 (1979)
[80B1] I. Broser et al., J. Phys. Soc. **49**, Suppl. A, 401 (1980)
[80C1] R. Claus, phys. stat. sol. (b) **100**, 9 (1980)
[89J1] B. Jusserand and M. Cardona, Light Scattering in Solids **V**, Topics in Applied Physics **66**, 49, Springer Heidelberg (1989)
[92K1] W. Kütt, Festkörperprobleme / Adv. Solid State Phys. **32**, 133 (1992)
[92N1] S. Nomura and T. Kobayashi, Solid State Commun. **82**, 335 (1992)
[93C1] M. Cardona et al., J. Phys. Condensed Matter **5A**, 61 (1993)
[93R1] T. Ruf et al., Phys. Rev. Lett. **71**, 3035 (1993)
[93W1] P.D. Wang and C.M. Sotomayor Torres, Phonons in Semiconductor Nanostructures, Kluwer Academic Publishers, Dordrecht (1993), Solid State Commun. **88**, 63 (1993), and J. Appl. Phys. **74**, 5047 (1993)
[94C1] M.P. Chamberlain, C. Trallero-Giner and M. Cardona, Phys. Rev. B **50**, 1611 (1994)
[94C2] M. Cardona, Festkörperprobleme / Adv. Solid State Phys., **34,** 35 (1994)
[94R1] T. Ruf et al., Festkörperprobleme / Adv. Solid State Phys. **34**, 237 (1994)
[96W1] U. Woggon et al., Phys. Rev. B **54**, 1506 (1996)
[97C1] F. Comas, C. Trallero-Giner and M. Cardona, Phys. Rev. B **56**, 4115 (1997)
[97G1] M. Göppert, Diplom Thesis, Karlsruhe (1997)
[97V1] R.V. Velasco and F. Garcia-Moliner, Surf. Science Rep. **28**, 123 (1997)
[98B1] H.J. Bakker, S. Hunsche and H. Kurz, Rev. Mod. Phys. **70**, 523 (1998)
[98G1] M. Göppert et al., Phys. Rev. B **57**, 13068 (1998)
[99D1] A. Dinger et al., phys. stat. sol. (b) **215**, 413 (1999)
[99R1] V. Romero-Rochin et al., J. Chem. Phys. **111**, 3559 (1999)
[01D1] A. Dinger et al., Phys. Rev. B **64**, 245310 (2001)
[01W1] T. Watanuki et al., J. Phys. Soc. Japan **70**, 2784 (2001)
[02C1] T.F. Crimmins, N.S. Stoyanov and K.A. Nelson, J. Chem. Phys. **117**, 2882 (2002)
[02K1] A. Kaschner et al., Appl. Phys. Lett. **80**, 1909 (2002)
[03H1] U. Haboeck et al., Phys. Stat. Sol. c **0** (6), 1710 (2003)

12

Optical Properties of Plasmons, Plasmon-Phonon Mixed States and of Magnons

Plasmons as collective excitations of the carriers in a partly filled band do not usually exist in pure semiconductors under thermodynamic equilibrium, since the probability of thermal excitation across the forbidden gap is almost negligible, except for narrow gap semiconductors or for semimetals.

There are, however, two conditions under which plasmons can be observed, and these are either highly doped semiconductors or highly excited ones. Under high doping and in thermodynamic equilibrium one has (see Sects. 8.9,14)

$$np = n_{\mathrm{i}}^2(T) = N_{\mathrm{eff}}^{\mathrm{e}} N_{\mathrm{eff}}^{\mathrm{h}} \mathrm{e}^{-E_{\mathrm{g}}/k_{\mathrm{B}}T} \tag{12.1}$$

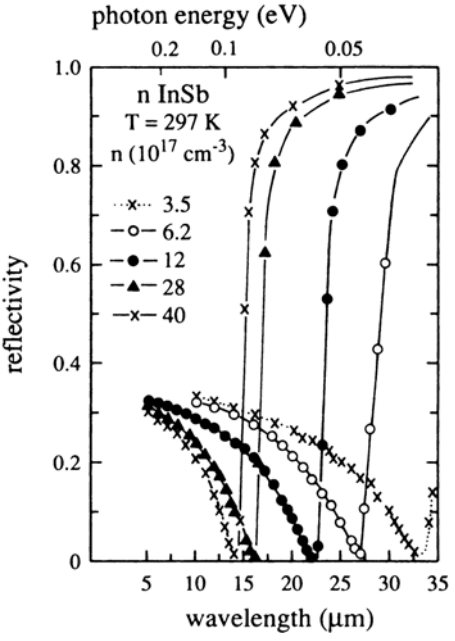


Fig. 12.1. The reflection spectrum in the vicinity of the plasmon resonance in doped InSb samples [57S1]

i.e., one has a high density either of electrons or of holes, and the data of the majority carriers have to be used in the calculation of ω_{P1} according to (10.1).

Under high excitation an electron–hole plasma can be formed, which consists of electrons and holes (for details see Chap. 21). In this case the reduced mass of electron and hole enters in (10.1).

In Fig. 12.1 we show the IR reflection spectra of InSb samples with different n-doping. In agreement with our statements that the transverse eigenfrequency of plasmons is zero, we see reflectivity close to one from zero up to the plasma frequency, which is, as we remember, the longitudinal eigenfrequency. The reflection minimum corresponds to the frequency at which the refractive index of the upper polariton branch is unity. Above the minimum R reaches a value determined by the background dielectric constant of the plasma resonance. The shift of the reflection minimum with increasing doping reflects the $n^{1/2}$ dependence of the plasmon frequency.

12.1 Surface Plasmons

Adapting Fig. 5.14 to the situation $\omega_T = 0$ we expect, for the case of plasmons, that surface plasmon modes exist in the whole frequency range between zero and ω_{PL} (Fig. 12.2). These surface plasmons can be investigated either by attenuated total reflection (see Sects. 11.1 or 25.1) or by another method which we will now outline briefly, since it has been used to obtain the data in Fig. 12.2. If a grating with spacing Λ is engraved in the surface, then the parallel component of the wave vector is conserved only modulo reciprocal vectors of this surface grating (see Sect. 3.1.3), i.e., modulo

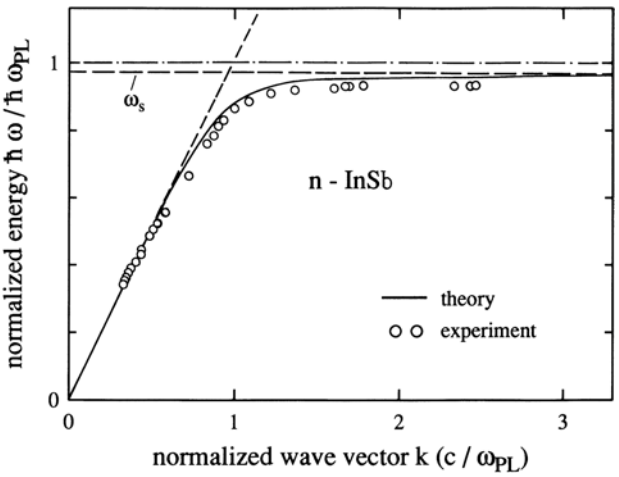


Fig. 12.2. The dispersion of the surface plasmon mode in n-doped InSb [71M1]

$$G_{\parallel} = \frac{2\pi}{\Lambda} m;\ m = 0, \pm 1, \pm 2 \ldots . \tag{12.2}$$

Adding such $G_{\parallel}$ values to the wave vectors of the incident light beam allows coupling to the surface polariton modes for $m \neq 0$. By varying the frequency and the angle of the incident light beam and thus

$$\omega, \boldsymbol{k}_{\parallel} + \boldsymbol{G}_{\parallel} \tag{12.3}$$

independently, it is possible to measure the dispersion of the surface plasmon polariton.

By changing the doping concentration (or the pump power in the case of an electron–hole plasma or the current in forward direction in a p-n junction) it is possible to deliberately vary the carrier concentration in a semiconductor and thus the plasmon frequency.

12.2 Plasmon-Phonon Mixed States

Eventually we encounter conditions where $\omega_{\mathrm{PL}} \approx \omega_{\mathrm{LO}}$ and plasmon-phonon mixed states develop (see Sect. 10.1). Depending on the selection rule, they may be detected in Raman scattering and/or in IR reflection spectroscopy. We give examples for both cases. In Fig. 12.3. the longitudinal modes $\hbar\omega_{-}$ and $\hbar\omega_{+}$ of Fig. 10.5 have been detected in an electron–hole plasma as a function of electron–hole pair density confirming nicely the concept developed in Sect. 10.1.

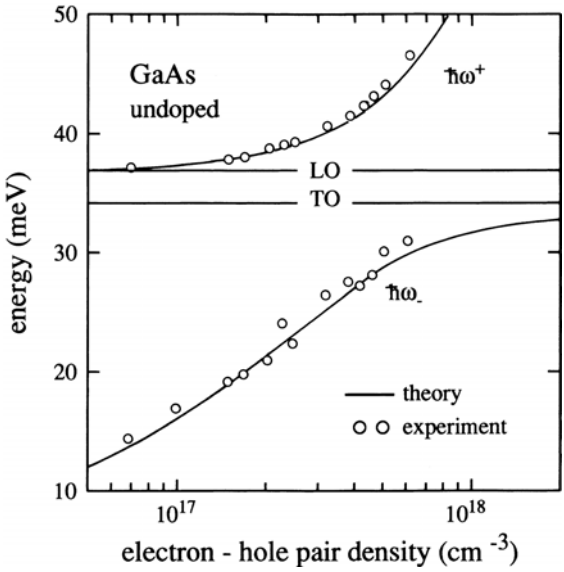


Fig. 12.3. The dependence of the plasmon-phonon mixed mode on the electron–hole pair density in highly photon-excited GaAs [84N1]

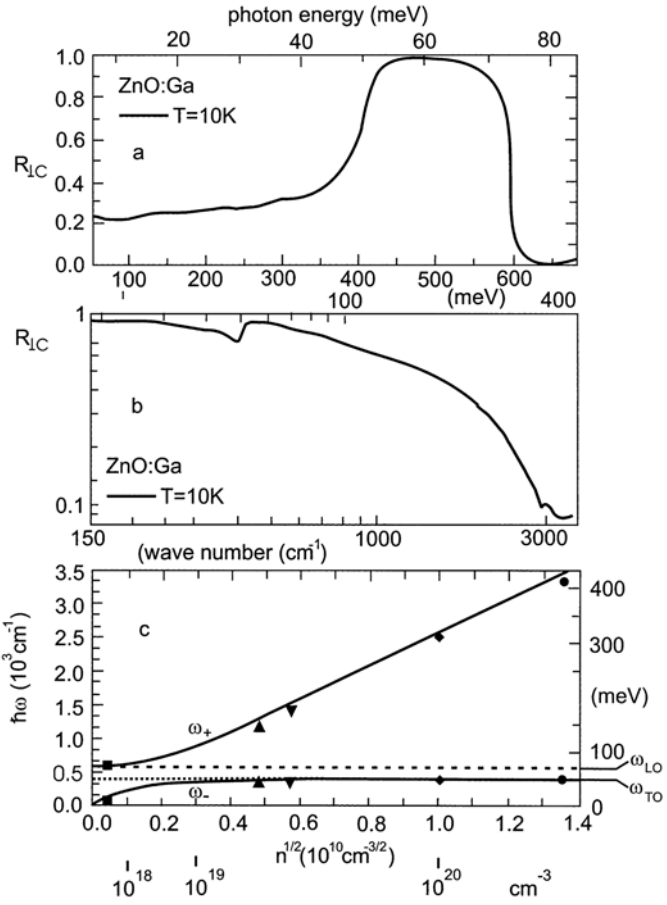


Fig. 12.4. Reflection spectra of n-doped ZnO:Ga samples with two different doping levels (note the different abscissas) (**a**,**b**) and the various transverse and longitudinal eigenenergeis of the plasmon-phonon mixed states as a function of carrier density (**c**) [97G1]

In [00N1, 02N1] of Chapt. 10 the plasmon-phonon mixed states have been observed in transient time-resolved reflection spectroscopy in various semiconductors after pulsed excitation.

In Fig. 12.4a,b we show reflection spectra of ZnO:Ga with various n-type doping levels and in Fig. 12.4c the resulting transverse and longitudinal eigenenergies. One can nicely see that in Fig. 12.4a, the plasma resonance occurs below 15 meV, but in Fig. 12.4b the density of $2.3 \times 10^{19}\,\mathrm{cm}^{-3}$ gives $\hbar\omega_{\mathrm{PL}}$ around 400 meV.

The alternating stopbands starting at the various transverse eigenfrequencies ($\hbar\omega_{\mathrm{T,PL}} = 0$; $\hbar\omega_{\mathrm{TO}}$) and the reflection minima, occurring at the longitudinal eigenfrequencies $\hbar\omega_-$ and $\hbar\omega_+$, respectively, can be nicely seen.

In simple metals the plasma frequency occurs in the 5 to 15 eV range, i.e., they show $R \lesssim 1$ over the whole IR and visible spectral range up to the UV. The colour of some metals like Cu, Au or the dip in the reflectivity of Ag in the near UV are due to a superposition of the plasmon resonance and band-to-band transitions in a similar way as the plasmon-phonon interaction in semiconductors in the IR [81N1].

12.3 Plasmons in Systems of Reduced Dimensionality

As already explained in Sect. 10.1 the plasma frequency starts at $\boldsymbol{k} = 0$, i.e., $\lambda = \infty$ in two- and one-dimensional systems at $\hbar\omega_{\mathrm{PL}} = 0$. This holds, e.g., for single quantum wells and multiple quantum wells as long as the coupling between them is negligible.

In Fig. 12.5 we show such a situation for two different modulation-doped $\mathrm{GaAs/Al_{1-y}Ga_yAs}$ samples. The data points have been deduced from angle-resolved Raman scattering, which allows variation of $\boldsymbol{k}_{\parallel}$.

If the quantum wells couple in a MQW sample one can also define a transverse wavelength or $\boldsymbol{k}_{\perp}$.

A dispersion for $\omega_{\mathrm{PL}}\left(\boldsymbol{k}_{\parallel}, \boldsymbol{k}_{\perp}\right)$ taking such effects into account is given, e.g., by [82O1]

$$\omega_{\mathrm{PL}} = \left\{ \frac{2\pi n e^2}{\varepsilon_{\mathrm{M}} m^*_{\mathrm{eff}}} k_{\parallel} \frac{\sin k_{\parallel} d}{\cosh k_{\parallel} d - \cos k_{\perp} d} \right\}^{1/2} \tag{12.4}$$

with n being the two-dimensional carrier density in every quantum well, m^* the effective mass of the carriers, d the distance between quantum wells and ε_{M} the dielectric constant of the barriers.

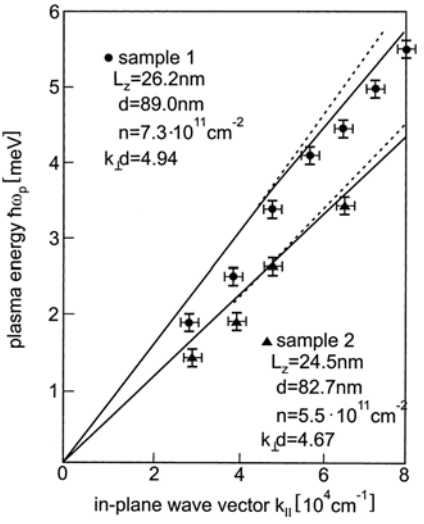


Fig. 12.5. The in-plane dispersion relation of plasmons in two different modulation doped $\mathrm{GaAs/Al_{1-y}Ga_yAs}$ samples [82O1]

The solid lines have been calculated with (12.4) and the dashed lines are linear approximations.

Obviously, various branches can be expected for various values of $k_\perp$. Additionally, a coupling to phonon resonances or to intersubband transitions may complicate things. Examples are found, e.g., in [86F1, 89Z1, 93S1] or in [01L1] of Chapt. 1.

Just as an addendum we mention again that the beautiful red color of gold ruby glasses results from surface plasmon excitation of gold colloids (or quantum dots) in the glass matrix [95K1].

12.4 Optical Properties of Magnons

As already mentioned in Sect. 10.2 only few semiconductors show magnetic ordering like the antiferromagnet MnSe.

One of the very few examples for the observation of magnon polaritons has been reported for antiferromagnetic FeF_2 in [97J1]. The optical magnons have been observed in reflection and the corresponding surface polaritons by attenuated total reflection, both in a narrow spectral range around $52.5\,\mathrm{cm}^{-1}$ (6.5 meV).

An example the difficulty to assign spectral features to plasmons see the references in [96G1, 96R1].

The optical properties of diluted (or semimagnetic) semiconductors (DMS) will be treated in Sect. 16.1.2.

12.5 Problems

1. Calculate the plasmon energy for n- (or p-) doped GaAs and carrier concentrations of 10^{10}, 10^{15} and $10^{20}\,\mathrm{cm}^{-3}$.
2. Calculate the density of electrons at which $\hbar\omega_{\mathrm{PL}} = \hbar\Omega_{\mathrm{LO}}$ for ZnO and InAs. Up to which temperatures are the electron gases degenerate.
3. Show that a treatment of the optical properties of free carriers like the Drude model outlined in (10.3 to 5) results in similar optical properties to those discussed above.
4. Which interband transition are responsible, together with the plasmon reststrahlbande for the color of Cu and Au?

References to Chap. 12

[57S1] G.W. Spitzer and H.Y. Fan, Phys. Rev. **106**, 882 (1957)
[71M1] N. Marshall, B. Fischer and H.-J. Queisser, Phys. Rev. Lett. **27**, 95 (1971)
[81N1] U. Nowak, W. Richter and G. Sachs, phys. stat. sol. (b) **108**, 131 (1981)
[82O1] D. Olego et al., Phys. Rev. B **25**, 7867 (1982)

[84N1] H. Nather and L.G. Quagliano, Solid State Commun. **50**, 75 (1984)
[86F1] G. Fasol et al., Phys. Rev. Lett. **56**, 2517 (1986)
[89Z1] T. Zettler et al., Phys. Rev. B **39**, 3931 (1989)
[93S1] L.C.Ó Súilleabháin et al., Solid State Comun. **87**, 517 (1993)
[95K1] U. Kreibig and R. Vollmer, Optical Properties of Metal Clusters, Springer Series in Material Sciences **25**, Springer, Berlin (1995)
[96G1] M. Grüninger et al., Europhys. Lett. **35**, 55 (1996)
[96R1] D. Reznik et al., Phys. Rev. B **53**, R 14741 (1996)
[97G1] M. Göppert et al., J. Luminesc. **72-74**, 430 (1997)
[97J1] M.F.R. Jensen et al., Phys. Rev. B **55**, 2745 (1997)

13

Optical Properties of Intrinsic Excitons in Bulk Semiconductors

Having treated the optical properties of phonons, plasmons and magnons, we come in this and the following chapters to the essence of semiconductor optics, namely the optical properties of excitons.

Phonons are necessary to describe the optical properties of semiconductors and of insulators in the IR; plasmons determine the optical properties of metals from the IR through the visible to the near UV, and in semiconductors, if present at all, they contribute along with the phonons to the IR spectra. Excitons, on the other hand, determine together with their continuum stats or the band-to-band transition the optical properties around the band gap, i.e., in the visible including the near UV and IR in the case of semiconductors and in the (V)UV for insulators. Although inorganic insulators like the alkali halides and organic ones such as anthracene have specific optical properties, many of the aspects presented in the following for excitons in semiconductors also apply to them.

13.1 Excitons with strong Oscillator Strength

We concentrate in this chapter on the intrinsic linear optical properties of excitons in bulk semiconductors starting from semiconductors with a dipole-allowed, direct band-to-band transition because they exhibit dipole-allowed excitons with the highest oscillator strength. Values of their longitudinal–transverse splitting Δ_{LT} range from 0.1 to beyond 10 meV.

It should be mentioned, however, that not all excitons in this group of semiconductors have high oscillator strength and that some excitons in semiconductors with dipole-forbidden band-to-band transitions may be dipole allowed, but with considerably lower oscillator strength. We come back to these cases in Sect. 13.2, ending with indirect gap materials.

13.1.1 Exciton–Photon Coupling

In semiconductors with dipole-allowed direct band-to-band transitions, excitons occur, which couple strongly to the radiation field. As a consequence

many optical properties can be understood quantitatively only in the strong-coupling or polariton picture. Thus we use this occasion to elucidate once more for this particular case the concept of weak and of strong coupling to the radiation field, heeding the classical dogma "repetitio est mater studiorum".

In Sects. 2.1–2.4 we introduced the electromagnetic radiation field, and in Sect. 2.5 the photons as its quanta. In Chaps. 9–11 we presented the properties of various elementary excitations. The interaction between the two can be treated in perturbation theory. This is the so-called weak coupling approach. The one-photon absorption coefficient $\alpha(\omega)$ is then at resonance proportional to the dipole matrix element squared in (13.1a), i.e., by first order-perturbation theory with the initial and final state:

$$\alpha(\omega) \propto |\langle f \mid H_{\mathrm{D}} \mid i\rangle|^2 \, \delta\left(E_{\mathrm{i}} - E_{\mathrm{f}} + \hbar\omega\right). \tag{13.1a}$$

The refractive index is obtained at this level of approximation either by a Kramers–Kronig transformation of $\alpha(\omega)$ or, away from the resonance, by second-order perturbation theory according to

$$n^2(\omega) - 1 \propto \sum_z \frac{\langle i|H_{\mathrm{D}}|z\rangle \, \langle z|H_{\mathrm{D}}|i\rangle}{E_z - E_{\mathrm{i}} - \hbar\omega} \tag{13.1b}$$

A photon $\hbar\omega$ creates virtually an excited intermediate state $|z\rangle$ under momentum conservation, which, after a time Δt limited by

$$\Delta E \, \Delta t = (E_z - E_{\mathrm{i}} - \hbar\omega) \, \Delta t \lesssim \hbar\,, \tag{13.2}$$

emits again a photon which is identical to the incident one, while the electronic system returns to the initial state $|i\rangle$. The time Δt during which the energy is "stored" in the virtually excited state reduces the phase velocity of the light and thus evidently describes an $n(\omega)$ which increases when ω approaches the resonance energy $E_z - E_{\mathrm{i}}$ from below, since ΔE goes to zero and Δt can be very long, in agreement with Fig. 4.4.

In the polariton concept, on the other hand, one quantizes the mixed state of the electromagnetic radiation and the excitation of the medium, i.e., the polarization wave. We already introduced this concept in Chap. 5. Since it is a very important one we want to demonstrate it here again for the exciton polariton.

For readers who are not satisfied with the simple statement that the polaritons are the quanta of the mixed states of electromagnetic radiation and excitation (or polarization), we give two other approaches. The first is just a diagrammatic representation of what was said before.

In Fig. 13.1 an incident photon creates an electron–hole pair, which recombines again to give a photon, and so on. The Coulomb interaction between electron and hole, which is responsible for the formation of the exciton, is represented by a virtual exchange of photons between electron and hole, i.e.,

by the vertical lines. Consequently the whole diagram of Fig. 13.1 can be considered as a representation of the exciton polariton.

In the other approach, which follows [93H1] of Chap. 1, we start with the electron and hole operators, construct from them the exciton and finally the exciton polariton. (see also Sect. 9.1)

We start with the creation and annihilation operators for excitons:

$$B^{+}_{v,\boldsymbol{k}}\,;\, B_{v,\boldsymbol{k}}\,. \tag{13.3}$$

The index v stands for the quantum numbers n_{B}, l, m.

It can be shown that the $B_{v,\boldsymbol{k}}$ deviate from the commutator relations of ideal bosons by a term proportional to the mean number of electron–hole pairs, n, contained in the volume of an exciton a^{d}_{B} (see e.g. [93H1, 93P1] of Chap. 1)

$$\langle [B_{0\,,\,0}, B^{+}_{0\,,\,0}]^{-} \rangle = 1 - \mathrm{O}\,(n a^{d}_{\mathrm{B}})\,, \tag{13.4}$$

where d is the dimensionality of the system. The Hamiltonian of a non-interacting exciton gas is then

$$H = \sum_{v,\boldsymbol{k}} E\,(v,\boldsymbol{k})\, B^{+}_{v,\boldsymbol{k}}\, B_{v,\boldsymbol{k}}\,. \tag{13.5}$$

Using an analogous expression for the photons with the number operator $c^{+}_{\boldsymbol{k}} c_{\boldsymbol{k}}$, for the interacting system of excitons and photons considering the leading, i.e., resonant terms around a specific resonance only, we obtain

$$H = \sum_{\boldsymbol{k}} \left[\sum_{v} E_{v\boldsymbol{k}}\, B^{+}_{v\boldsymbol{k}}\, B_{v\boldsymbol{k}} + \hbar\,\omega_{\boldsymbol{k}}\, c^{+}_{\boldsymbol{k}} c_{\boldsymbol{k}} - \mathrm{i}\,\hbar \sum_{v} g_{v\boldsymbol{k}}\, (B^{+}_{vk}\, c_{\boldsymbol{k}} - h.c.) \right]\,. \tag{13.6}$$

The coupling coefficients $g_{v\boldsymbol{k}}$ contain the transition matrix elements as in (13.1).

If we consider the third term on the right-hand side of (13.6) as a perturbation, we are back once more to the weak coupling limit.

The polariton concept is obtained if we diagonalize the whole Hamiltonian (13.6) by a suitable linear combination of the $B_{v\boldsymbol{k}}$ and the $c_{\boldsymbol{k}}$, leading to the polariton operator $P_{\boldsymbol{k}}$

$$P_k = u_{v\boldsymbol{k}}\, B_{v,\boldsymbol{k}} - v_{\boldsymbol{k}}\, c_{\boldsymbol{k}} \tag{13.7}$$

$$\text{with} \quad |u_{v\boldsymbol{k}}|^2 + |v_{\boldsymbol{k}}|^2 = 1\,. \tag{13.8}$$

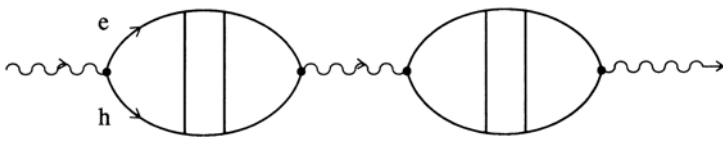


Fig. 13.1. Diagrammatic representation of an exciton polariton

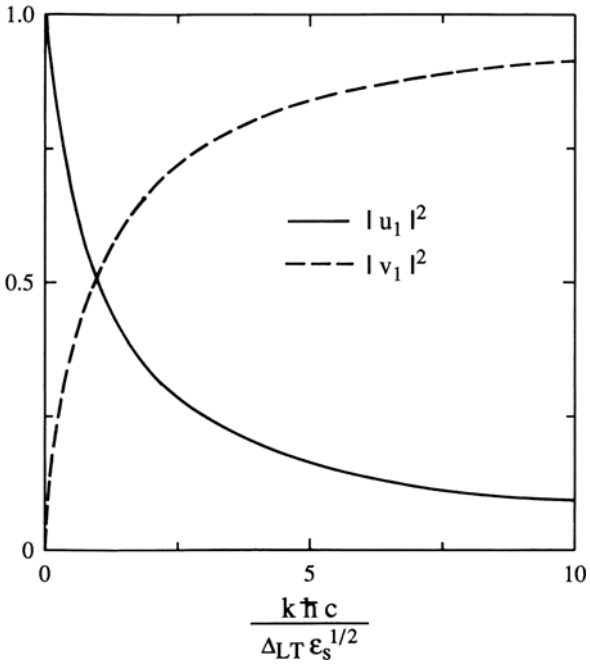


Fig. 13.2. The energy dependence of the $u_{v,\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$ of (13.7) for the upper polariton branch. According to [93H1] of Chap. 1

The $u_{v\boldsymbol{k}}$ give the exciton-like character of the polaritons. They are close to one around the exciton energy $E_{v,0}$ and decrease on the upper and lower polariton branches with decreasing energetic distance from the resonance. The $v_{\boldsymbol{k}}$ give the photon-like part and, according to (13.8), show the opposite behavior. In Fig. 13.2 we give an example for the upper polariton branch. As a rule of thumb we can state that the polariton wavefunction contains considerable exciton-like parts over energies

$$|\hbar\omega - E_{\rm ex}| \lesssim 10\,\Delta_{\rm LT}\,. \tag{13.9}$$

It is interesting to note that the dispersion relation that we obtain from this approach is identical to the one obtained from the set of classical coupled oscillators treated in Chaps. 4 and 5.

13.1.2 Consequences of Spatial Dispersion

In contrast to that of optical phonons, the $\boldsymbol{k}$ dependence of the exciton energy is significant. For $\boldsymbol{k}$ vectors in the transition region from the photon-like to the exciton-like part of the dispersion relation, the so-called bottle-neck, the kinetic energy term in (9.1a) becomes comparable to the longitudinal transverse splitting $\Delta_{\rm LT}$. The consequences of spatial dispersion have already been outlined in connection with Figs. 5.3–5.5, so that we can restrict ourselves to just recalling them here. For all frequencies there is at least one propagating

mode. This fact reduces the reflectivity in the reststrahlbande to values below 1, even in the case of negligible damping. For frequencies above ω_L there are several propagating modes, and below it there is at least one propagating mode and one or more evanescent ones. This situation is not covered by the boundary conditions deduced from Maxwell's equations and additional boundary conditions (abc) have to be introduced containing the information about what fraction of the energy transmitted through the interface travels on which polariton branch. Since this "branching" ratio is ω-dependent and since the imaginary parts of the various branches differ, the decay of the intensity into the depth of the sample can be nonexponential. This means the "effective" absorption coefficient can be thickness dependent. Furthermore it looks more complex (Fig. 5.5) than Fig. 4.4.

The abc which have been introduced by Pekar and by Hopfield [58H1, 62H1, 62P1, 63H1, 64M1] assume that the excitonic part of the polarization at the surface vanishes, or its derivative normal to the surface, or a linear combinations of both. See Sect. 5.4 and e.g. [74A1, 75L1, 78B1, 78H1, 78S1, 79B1, 79S1, 80B1, 81B1, 81L1, 81S1, 82O1, 82R1, 82S1, 83M1, 84H1, 84S1, 85H1] of Chap. 5. In [98H1] a way out of this abc problem has been shown. For even more recent approaches see [00T1, 01S1]. Furthermore one can assume that excitons do not "leak out" of the semiconductor into vacuum and that there should consequently be an exciton-free surface layer (dead layer), the

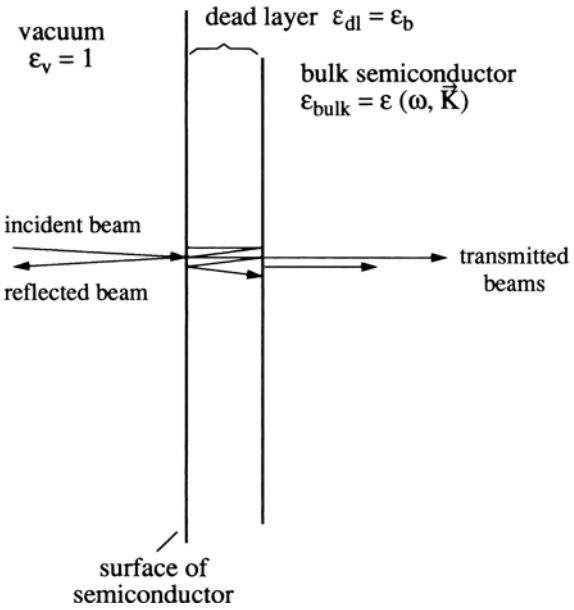


Fig. 13.3. The problem of reflection for a semiconductor in the vicinity of an exciton resonance for normal incidence including multiple reflection in a dead layer and two propagating modes due to spatial dispersion

optical properties of which are described by ε_b and which has a minimum thickness of the excitonic Bohr radius. Electric fields, which occur often at surfaces normal to them, can ionize, i.e., destroy the exciton and lead to an increase in the thickness of the dead layer. The problem which thus has to be solved to calculate a reflection spectrum is shown in Fig. 13.3. An incident beam passes first the dead layer in which multiple reflection occurs and then enters the semiconductor in which several modes can be excited. Sometimes scientists apply even more complex models assuming, e.g., that the damping and/or the eigenfrequencies are depth dependent [79E1, 81L1, 82S1].

The formulas to calculate the spectra are rather complex and we do not give them here but refer the reader to [75L1, 78S1, 79E1, 80B1, 81L1, 82B1, 82S1, 84R1, 85H1, 93K1, 95B1] or the references from Chap. 5 given above. Instead we give in the next section examples of reflection, transmission, and luminescence spectra of the exciton polariton in bulk semiconductors with direct, dipole-allowed band-to-band transitions.

13.1.3 Spectra of Reflection, Transmission and Luminescence

In Fig. 13.4a we show the bandstructure of CdS in part b the dispersion of the exciton polariton for the orientation $\boldsymbol{k} \perp \boldsymbol{c}, \boldsymbol{E} \perp \boldsymbol{c}$ and including the $n_B = 1$ excitons involving a hole either in the A or in the B valence band. The A exciton is a rather simple resonance for this orientation, comparable to our model system in Sect. 5.2. The k-linear term of the B valence band (Sect. 8.8) mixes the singlet and triplet states for $\boldsymbol{k}_\perp \neq 0$ (see Sect. 13.2.1.1) and gives rise to an additional polariton branch. Figure 13.4c finally gives the reflection spectra of the two resonances for $\boldsymbol{E} \perp \boldsymbol{c}$ and $\boldsymbol{E} \parallel \boldsymbol{c}$ and of some higher states ($n_B \geq 2$). The combination of the Γ_1 (S-) envelope function for $n_B = 1$ with the symmetries of the electron Γ_7 and the holes (AΓ_9, BΓ_7) gives excitons of symmetries AΓ_5 and AΓ_6 and BΓ_1, BΓ_2, BΓ_5 as explained in more detail in Chap. 26. The Γ_6 and Γ_2 states are triplets which couple only weakly to the radiation field since they are spin-flip and dipole-forbidden and do not show up in reflection (see Sect. 13.2.1.1). Γ_5 and Γ_1, states couple to the radiation field for the orientations $\boldsymbol{E} \perp \boldsymbol{c}$ and $\boldsymbol{E} \parallel \boldsymbol{c}$, respectively. These selection rules show up clearly in the reflection spectra, the $n_B = 1$ A excitons being seen only in $\boldsymbol{E} \perp \boldsymbol{c}$.

A fit to the reflection spectra (not shown here) using spatial dispersion, an exciton-free layer and some abc coincides with experiments within a few percent.

The AΓ_5 resonance is quite simple, as already mentioned. R remains significantly below 1 due to spatial dispersion as predicted above. For smaller oscillator strength or longitudinal-transverse splitting, the maximum almost disappears and only a narrow dip close to the longitudinal eigenfrequency remains, as shown in Fig. 13.5 for ZnTe.

The small spike around AΓ_5^L in Fig. 13.4c is partly caused by the onset of the UPB but mainly by the dead layer. An increase of its thickness increases the importance of the spike due to multiple reflections (Fig. 13.3) and may

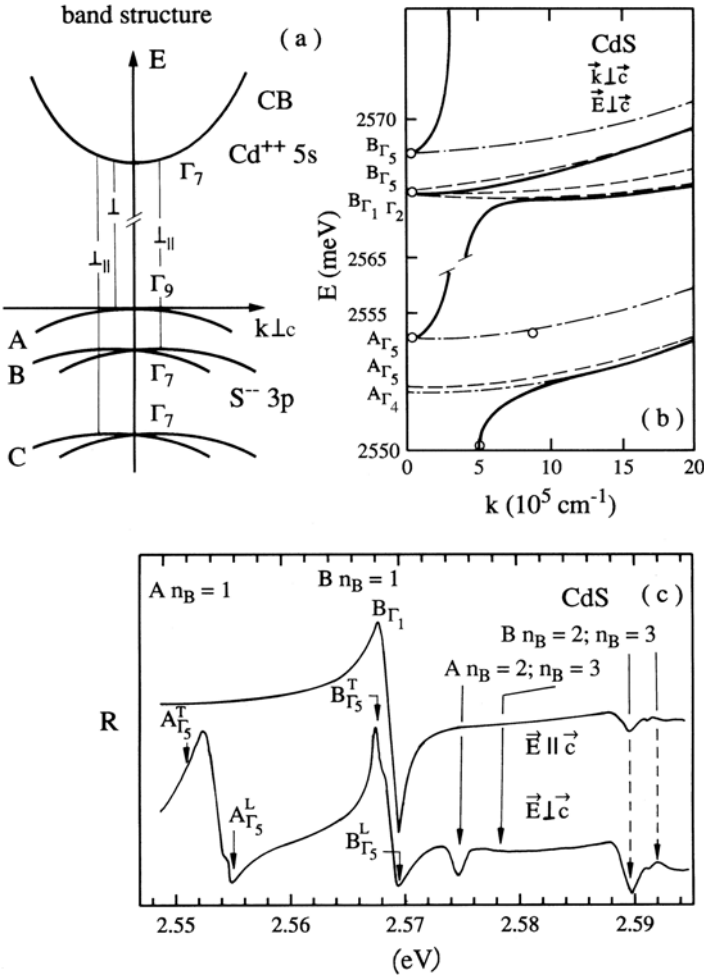


Fig. 13.4. The bandstructure of CdS around the Γ-point (**a**); the dispersion of the $n_B = 1$ A and B exciton polariton resonances (**b**); and the reflection spectra for the polarizations $\boldsymbol{E} \perp \boldsymbol{c}$ and $\boldsymbol{E} \parallel \boldsymbol{c}$ (**c**). According to [82B1, 85H1, 93K1]

even lead to an "inversion" of the usual reflection spectrum, i.e., to a dip at low energies and a maximum above. A set of calculated spectra showing this phenomenon is given in Fig. 13.6.

The BΓ_1 exciton resonance is again a simple one, but the BΓ_5 has a small dip stemming from the additional polariton branch shown in Fig. 13.4b, which at this energy reaches exactly $n = 1$.

While the A and BΓ_5 excitons have in CdS for $\boldsymbol{E} \perp \boldsymbol{c}$ roughly equal oscillator strength and Δ_{LT} [85H1], the situation changes for close lying resonances as occurs for ZnO [85H1]. In this case the longitudinal transverse splitting

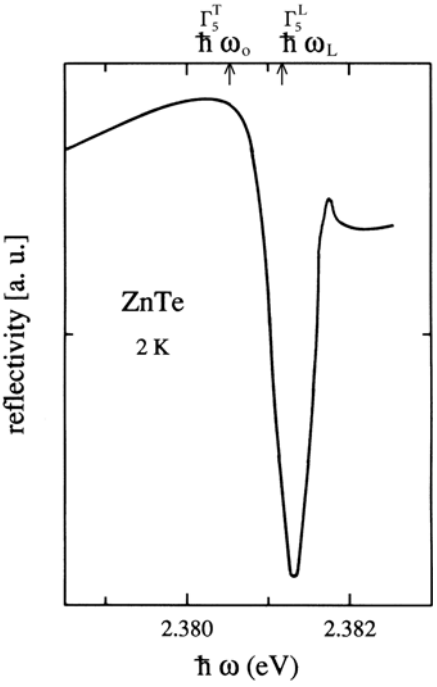


Fig. 13.5. A reflection spectrum for ZnTe. According to [83M1]. For the dispersion of the exciton polariton see Sect. 16.1.1

Δ_{LT} of the AΓ_5 exciton is reduced and that of the BΓ_5 increased as discussed already in Sect. 4.5. Only the sum of $\Delta_{\mathrm{LTA}\Gamma_5}$ and $\Delta_{\mathrm{LTB}\Gamma_5}$ has to be constant and equal to $\Delta_{\mathrm{LTC}\Gamma_1}$ [60H1]. See also Fig. 13.21

At higher energies we see $n_{\mathrm{B}} \geq 2$ exciton states which split into several sublevels due to the various L and/or L_z values of the envelope function. The reflection signal of these higher states decreases due to the n_{B}^{-3} dependence of the oscillator strength. In the band-to-band transition region the reflection spectra are usually flat and structureless. The C-excitons expected from Fig. 13.4a are situated around 2.61 eV and are off the scale of Figs. 13.4b, c. These resonances are washed out even at low temperature because the C-exciton is situated in the continuum of the A- and B-excitons and thus has a rather short phase relaxation time T_2, i.e., strong damping.

With increasing lattice temperature the exciton resonances are broadened due to increasing scattering with phonons. Sometimes they are hardly visible at RT as shown in Fig. 13.7 for CdS. A similar washing out of the exciton resonance can occur even at low temperatures in samples with high impurity content and/or lower crystalline quality or for increasing excitation desenty.

To summarize, we can state that the reflection spectra of semiconductors are determined around the gap by exciton polaritons. The longitudinal eigenenergy can be reasonably well determined from the reflection minimum which corresponds to $n = 1$ on the UPB and is therefore situated only slightly

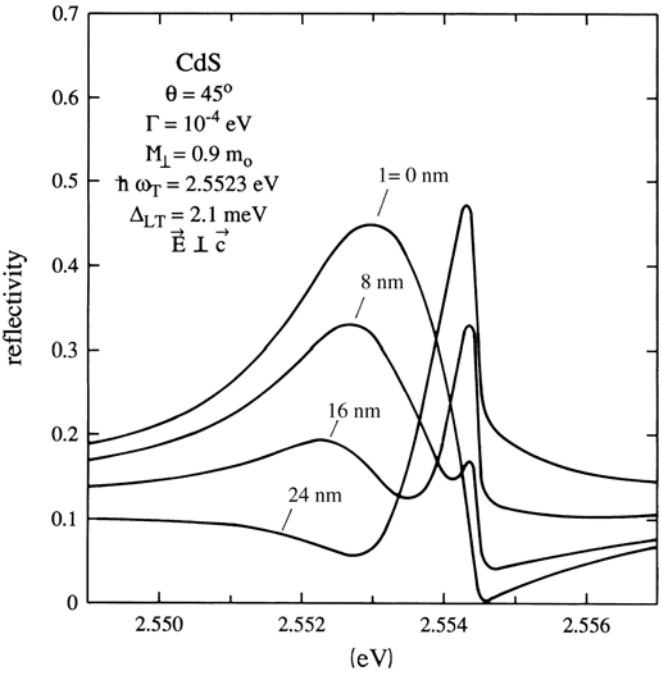


Fig. 13.6. A set of calculated exciton reflection spectra of CdS for 45° incidence and for various thicknesses of the exciton-free layer with all other parameters kept constant. According to [82R1]

(< 1 meV) above $\hbar\omega_{\mathrm{L}}$. The transverse eigenenergy, the oscillator strength, the damping, and the effective mass of the exciton can be extracted only with the help of a rather complicated line-shape analysis.

For uniaxial, hexagonal materials like CdS or ZnO it can be shown from group-theoretical considerations that the Γ_5-excitons are the resonances for the ordinary beam (Sect. 3.1.7) because the polarization $\boldsymbol{E} \perp \boldsymbol{c}$ can be realized for all angles between $\boldsymbol{k}$ and $\boldsymbol{c}$. The Γ_1 resonances have dipoles oriented parallel to $\boldsymbol{c}$. As a consequence they have maximum coupling to the radiation field for $\boldsymbol{E} \parallel \boldsymbol{c}$ and $\boldsymbol{k} \perp \boldsymbol{c}$, but develop as extraordinary or mixed-mode polaritons to the longitudinal state if the angle $\angle(\boldsymbol{k}, \boldsymbol{c})$ is continuously changed from $\boldsymbol{k} \perp \boldsymbol{c}$ to $\boldsymbol{k} \parallel \boldsymbol{c}$. In Fig. 13.8 we show the dispersion of the CΓ_1 exciton polariton in ZnO, which has an oscillator strength varying according to

$$\Delta_{\mathrm{LT}} = \Delta^0_{\mathrm{LT}} \sin^2 \angle(\boldsymbol{k}, \boldsymbol{c})\,. \tag{13.10}$$

Experimental data for mixed mode polaritons can be found in [72W1, 75L1]. The small refection spike seen at B$\Gamma_{5\mathrm{L}}$ in [03C1] might similarly well be due to a mixed mode polariton due to the finite angle of aperture of the incident beam.

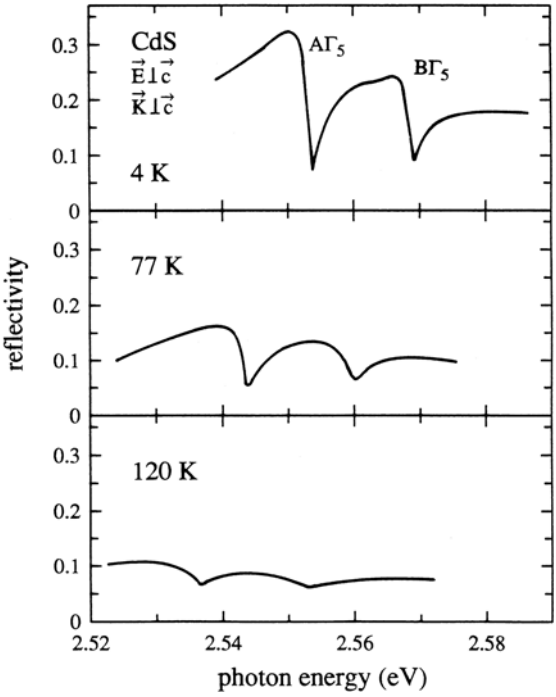


Fig. 13.7. A set of reflection spectra of the A- and B-exciton resonances of CdS for various lattice temperatures. According to [80B2]

The absorption spectra of $n_{\mathrm{B}} = 1$ polariton resonances are usually difficult to measure quantitatively, since in the resonance region the effective absorption, or rather extinction, coefficient reaches values in the range

$$10^4\,\mathrm{cm}^{-1} \lesssim \alpha_{\mathrm{eff}} \lesssim 10^6\,\mathrm{cm}^{-1}\,. \tag{13.11}$$

As a consequence for samples with $d \gtrsim 1\,\mu\mathrm{m}$, the transmitted light intensity goes to "zero", i.e., to values comparable to or smaller than the stray light of the spectrometer. We show in Fig. 13.9a an absorption spectrum of a ZnO layer at RT. The layer is of polycrystalline nature, produced by evaporating Zn on a substrate of quartz glass and subsequent oxidation [43M1]. The crystallites have their $\boldsymbol{c}$ axis oriented perpendicular to the substrate. The peak at 3.3 eV is due to the close lying and thermally broadened AΓ_5 and BΓ_5 exciton resonances ([75L1,78H1,80B1,81L1,82B1,85H1,03C1] and references therein). This is possibly one of the first experimental observations of a Wannier exciton in absorption, though the author of [43M1] was most probably not aware of this fact. See also Sect. 9.1. In Fig. 13.9b experimental transmission spectra are shown A and BΓ_5 resonance of a thin CdS platelet type sample. There are Fabry–Perot modes (see Sect. 3.1.6) in the transparent region due to the natural reflectivity of the surfaces of the as-grown platelet type sample. A de-

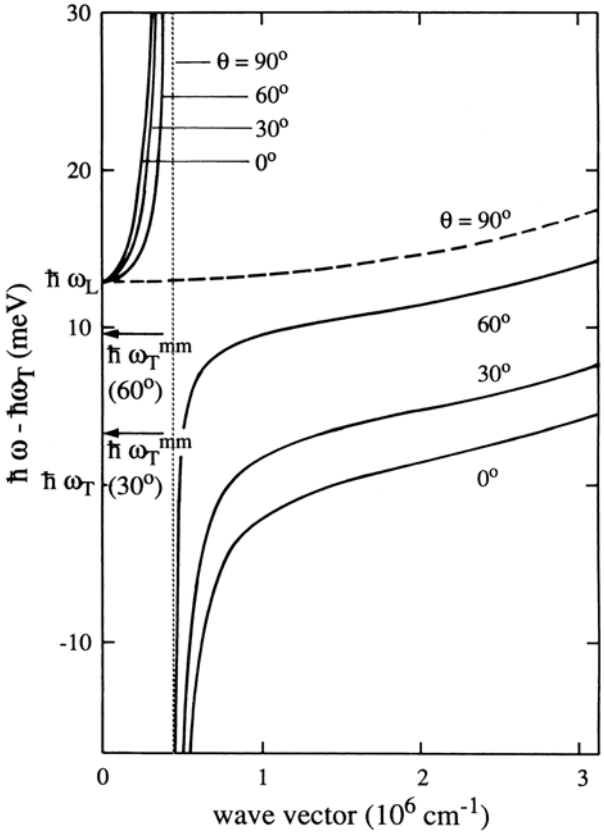


Fig. 13.8. Dispersion of the exciton polariton in a uniaxial material for various angles between $\boldsymbol{k}$ and $\boldsymbol{c}$. Data for the CΓ_1 exciton polariton in ZnO. According to [78H1]

tailed analysis of $\alpha_{\text{eff}}(\omega)$ in the resonance region is possible only for sample thicknesses of the order of 1 µm, such that $\alpha d \lesssim 4$. Such samples cannot usually be produced by grinding and polishing of thicker samples, since these processes introduce so much damage to the lattice that the damping γ of the resonance becomes too large (i.e., $\hbar\gamma > \Delta_{\text{LT}}$).

Sometimes epitaxial layers can be used for this type of investigation, but care has to be taken that these layers do not contain (inhomogeneous) strain since transmission measurements integrate over the whole sample thickness. Fortunately some semiconductors, such as CdS or CdSe, tend to grow as thin, single crystal platelets. The absorption spectrum of Fig. 13.9b stems from such a sample. Figure 13.9c shows a calculated tranmission spectrum of the AΓ_5 $n_{\text{B}} = 1$ resonance. The similarity to the A_1 resonance in Fig. 13.9b and the deviations from the simple case without spatial dispersion of Figs. 4.4 or 11.1 are evident. In Fig. 13.9d we give an overview of the absorption

spectrum of a thin GaAs sample. GaAs has, in comparison to CdS or ZnO, a much lower oscillator strength due to the larger value of a_B (see Sect. 9.1); (Δ_{LT} GaAs ≈ 0.1 meV). Therefore it is easier to measure the absorption of the $n_B = 1$ exciton. We can see in this figure similarly as for CdS in Fig. 13.9b also the $n_B = 2$ and 3 levels with S-envelope function. Even higher states ($n_B > 3$) merge with the continuum. The decrease of the oscillator strength with n_B^{-3} [57E1] is at least qualitatively confirmed by this spectrum.

The rather constant value of $\alpha(\omega)$ in the region of the continuum states comes from the product of the square root combined density of states (the dashed line gives the calculated absorption spectrum for a simple band-to-band transition without Coulomb effects for comparison) and the Sommerfeld factor, already discussed in Sect. 9.1,3 for $d = 3$.

At higher temperatures, the excitons develop an absorption tail to lower photon energies, which is described by the so-called Urbach or Urbach–Martienssen rule [53U1, 57M1, 58D1, 71K1, 71S1, 72D1, 85L1]

$$\alpha(\hbar\omega) = \alpha_0 \exp\left[-\sigma(T)\,(E_0 - \hbar\omega)/\,k_B\,T\right], \qquad \hbar\omega < E_0 \tag{13.12}$$

where α_0 and E_0 are material parameters. E_0 is an energy situated several 10 meV above the energy of the lowest free exciton at $T_L = 0$ K. σ is a function varying only weakly with temperature. An example is given in Fig. 13.10 for CdS.

The temperature dependence of the band gap, which can also be nicely seen in Figs. 13.7 or 13.10, is already incorporated in (13.12) since the exciton peak shifts in parallel to the gap.

The reason for the behavior described by (13.12), which is of rather universal nature in semiconductors and insulators, is the interaction of excitons with optical phonons. Two effects are usually discussed in theory: a momentary localization of the excitons in the randomly fluctuating field of optical phonons, or an ionization in their electric field. These two effects seem to contribute with a weight that depends on the material parameters. Details about the theory can be found in [71S1, 72D1, 85L1] and references therein.

The absorption of direct excitons in indirect materials will be addressed in Sect. 13.2.2.

The investigation of the luminescence from excitons or more precisely exciton polaritons is a rather difficult task. For early investigations see [72B2, 72W2]. The emission from these states is generally very weak even at low temperatures even in high quality sampels. This has various reasons: the total luminescence yield η_{lum} of semiconductors is often very low. Frequently one finds, even for direct gap materials,

$$10^{-1} \gtrsim \eta_{lum} \gtrsim 10^{-3}, \tag{13.13}$$

i.e., the main recombination channel is non-radiative involving defect centers. A large part of the emission stems from phonon replica and, especially at

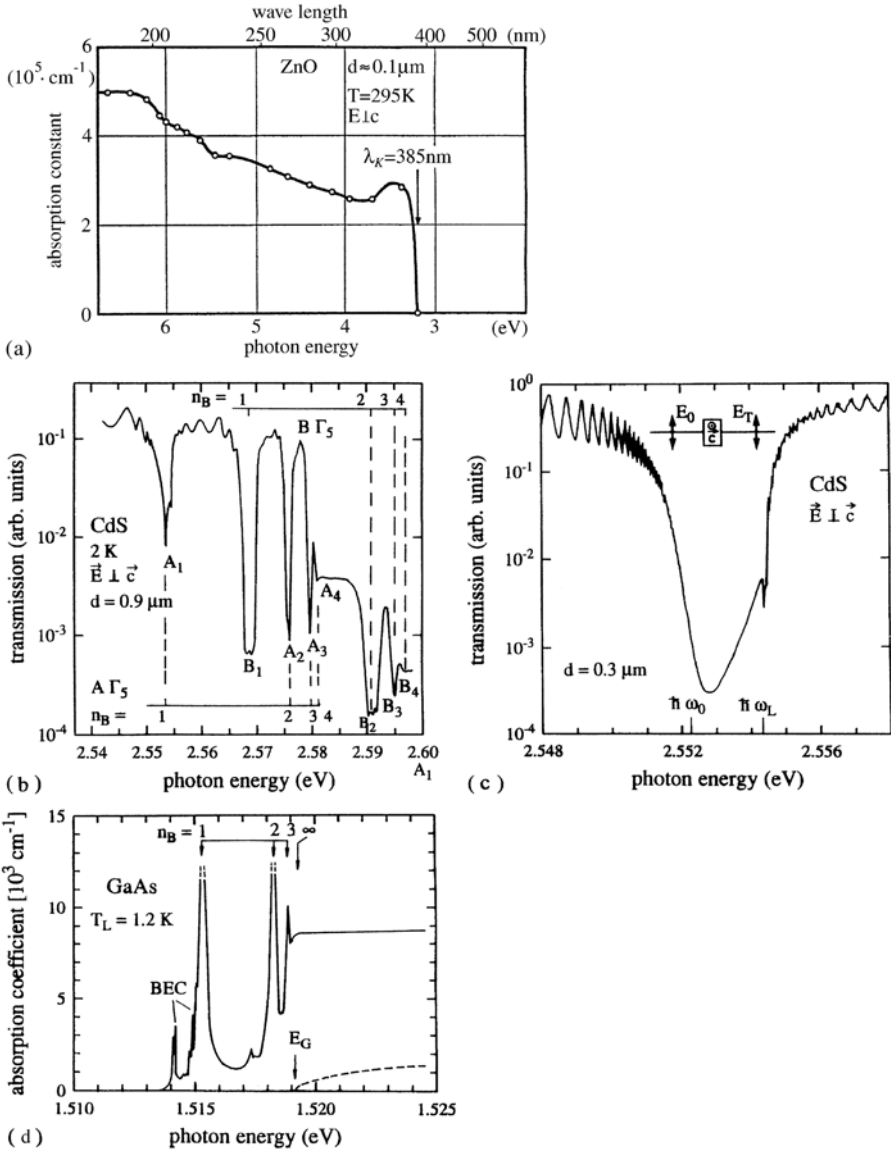


Fig. 13.9. An experimental absorbtion spectrum of ZnO at RT (**a**), an experimental transmission spectrum of CdS for the orientation $\boldsymbol{E} \perp \boldsymbol{c}$ in the region of the A and B exciton resonances (**b**); a calculated one for the AΓ_5 resonance for $n_\mathrm{B} = 1$ (**c**); and an absorption spectrum for a thin GaAs sample (**d**). According to [43M1] to [79V1, 82R1] and to [91U1], respectively

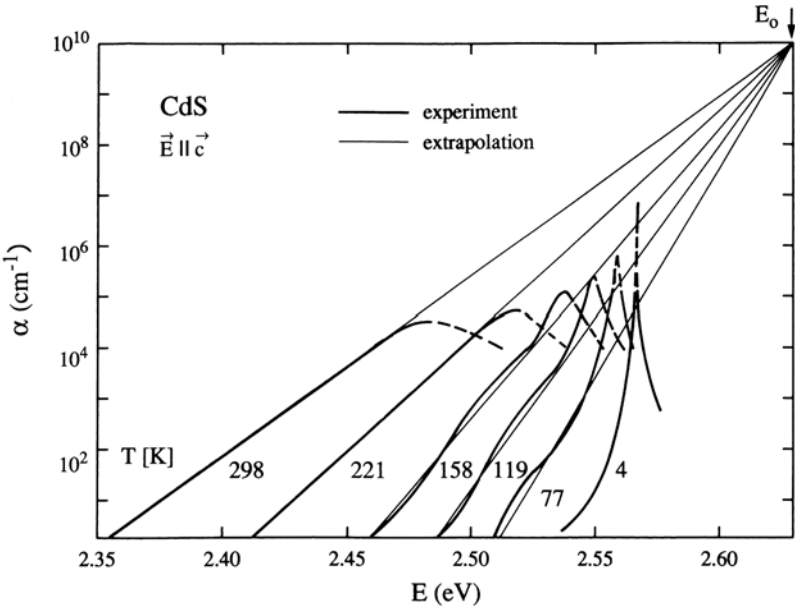


Fig. 13.10. The absorption edge of CdS for various temperatures and the polarization $\boldsymbol{E} \parallel \boldsymbol{c}$. According to [76S1]

low temperatures, from bound-exciton complexes, donor–acceptor pairs, or other defect centers which are considered in Chap. 14. Furthermore, the direct emission from free exciton polaritons is limited by various effects: One is the internal reflection, another the small escape depth. These points will be further clarified below.

If we excite an electron–hole pair, e.g., in the continuum states, it will relax to lower energies and thermalize by emission of phonons, as described in more detail in Chap. 23. At very low temperatures ($k_{\mathrm{B}}T < \Delta_{\mathrm{LT}}$) the excitons end up on the LPB where they further relax by acoustic phonon emission. Since the scattering matrix element and the density of final states both decrease in the transition region between exciton- and photon-like dispersion, the excitons accumulate there. This is the reason why this region is called a bottle-neck. At higher temperatures the excitons reach essentially a Boltzmann-like distribution on the exciton-like part of the LPB, on the longitudinal branch, and on the UPB. In the polariton picture the luminescence from these states cannot be described as a "recombination" of the exciton polariton with emission of a photon, since the photon will be immediately reabsorbed to form an exciton, or in other words, since we are considering the quanta of the mixed state of exciton and photon.

The proper description is the following: The exciton polariton moves with its group velocity through the sample. It can be scattered by impurities or phonons or be trapped. Eventually it reaches the surface. In most cases it

will be reflected back into the sample. The limiting angle for total internal reflections α_{TR} is, e.g., for $n = 5$ – a typical value on the LPB in the bottle neck region as shown in Sect. 13.1.4, only about 13°. Of the excitons impinging under an angle smaller than α_{TR}, a considerable fraction are also reflected back into the sample, as becomes clear if one considers the formulas for the reflection under normal incidence or if one integrates over Fresnel formula of Sect. 3.1.4.

Furthermore, the luminescence yield of free-exciton polaritons is limited by the small escape depth l_{esc}, i.e., the depth from which they can reach the surface. If one excites in the band-to-band transition region, the exciting light penetrates about 0.1–1 μm into the sample corresponding to α-values of 10^4–$10^5\,\mathrm{cm}^{-1}$ in this spectral range. By diffusion, the excitons spread out over a region of 1–2 μm. The depth from which they can reach the surface is much less than this. A rough estimate can be obtained either from the inverse effective absorption coefficient in the exciton resonance or from the product of phase-relaxation time T_2 (Sect. 23.1) and the group velocity in the exciton resonance (Sect. 13.4):

$$\alpha_{\mathrm{eff}}^{-1} = (10^4 - 10^6\,\mathrm{cm}^{-1})^{-1} = 0.01 - 1\,\mu\mathrm{m} \tag{13.14a}$$

$$l_{\mathrm{esc}} = v_g T_2 = (10^{-3} - 10^{-5} \cdot c)\,(10 - 40\,\mathrm{ps}) = 10 - 0.03\,\mu\mathrm{m}\,. \tag{13.14b}$$

In spite of all these difficulties it was possible to observe the emission from the exciton polariton in many semiconductors and we give an example for ZnO in Fig. 13.11. On the left the dispersion relation of the A-exciton is shown for $\boldsymbol{k} \perp \boldsymbol{c}$ and $\boldsymbol{k} \parallel \boldsymbol{c}$. Since the Γ_7 and Γ_9 valence bands are most probably inverted in ZnO as compared to CdS, see Sect. 8.8, the k-linear term appears in the A-exciton for the orientation $\boldsymbol{k} \perp \boldsymbol{c}$. The influences of the resulting additional polariton branch, of the longitudinal branch, and of the UPB are seen by comparison with the orientation $\boldsymbol{k} \parallel \boldsymbol{c}$ where the longitudinal branch and the k-linear term are missing. A luminescence polarized $\boldsymbol{E} \parallel \boldsymbol{c}$ in the A exciton resonance has been observed in [03C1] possibly indicative for the weak $\mathrm{A}\Gamma_1$ contribution according to [60H1]. Some other examples for the polariton luminescence including also other materials, both as bulk samples and epitaxial layers are found e.g. in [68V1, 77V1, 79P1, 80D1, 95S1, 97N1, 97N2, 98M1, 99R1, 00T1, 01R1, 01S1].

A luminescence channel of the exciton polaritons with higher luminescence yield in semiconductors with strong exciton-LO phonon coupling are the LO-phonon replicas. In this case a polariton on the exciton-like part of the dispersion relation or in the bottleneck is scattered onto the photon-like branch by emission of one or more longitudinal optical phonons. The coupling with this type of phonon is stronger than with transverse optical or acoustic phonons since the lattice distortion of the polaron (Sect. 8.6) can be described largely as a superposition of longitudinal optical phonons. Once the polariton is on the photon-like branch, it travels over long distances with almost negligible damping and is transmitted through the surface into vac-

uum with quite high probability, since $n \lesssim 2$ in the corresponding spectral range.

In Fig. 13.12 we show the appearance of the LO-phonon satellites schematically. If we neglect the bottleneck region for the moment, we can deduce with the Boltzmann occupation probability the distribution of the excitons as a function of their kinetic energy E_{kin}

$$N(E_{\mathrm{kin}}) \propto \begin{cases} E_{\mathrm{kin}}^{1/2} \exp\{-E_{\mathrm{kin}}/k_{\mathrm{B}}T\} & \text{for } E_{\mathrm{kin}} \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{with } E_{\mathrm{kin}} = \frac{\hbar^2 \boldsymbol{k}^2}{2M} \,. \tag{13.15a}$$

The lineshape of the luminescence of the m-th LO-phonon replica is then given by [82P1]:

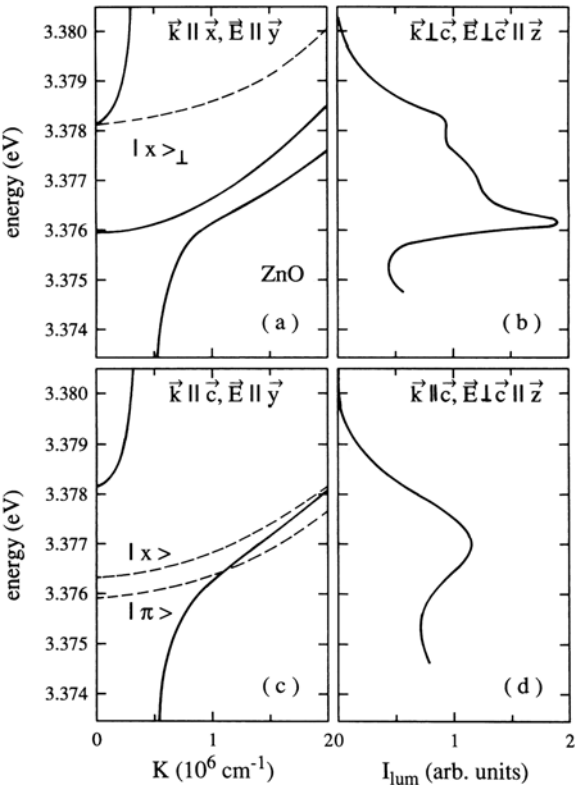


Fig. 13.11. The dispersion of the $n_{\mathrm{B}} = 1$ A-exciton resonance in ZnO for the polarizations $\boldsymbol{E} \perp \boldsymbol{c}\,(\Gamma_5)$ and $\boldsymbol{E} \parallel \boldsymbol{c}\,(\Gamma_1)$ (**a**, **c**) and the observed polariton luminescence (**b**, **d**). According to [78H1, 81K1]

$$I_m^{\mathrm{lum}}(\hbar\omega) \propto \begin{cases} E_{\mathrm{kin}}^{1/2} \exp\left(-E_{\mathrm{kin}}/k_{\mathrm{B}}T\right) W_m(E_{\mathrm{kin}}) & \text{for } E_{\mathrm{kin}} \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{with } \hbar\omega = E_0 - m\hbar\omega_{\mathrm{LO}} + E_{\mathrm{kin}} \,. \tag{13.15b}$$

where E_0 is the energy of the dipole allowed, transverse exciton at $\boldsymbol{k} = 0$.

The transition probability $W_m(E_{\mathrm{kin}})$ can be often expressed by a power law, i.e.,

$$W_m(E_{\mathrm{kin}}) \propto E_{\mathrm{kin}}^{l_m} \,. \tag{13.16}$$

For $m = 1$ one finds $l_1 = 1$ since the density of final states for the LO phonons increases with $E_{\mathrm{kin}} \propto k^2$ assuming that the wave vector of the photon-like exciton polariton in the final state is negligible.

For $m = 2$ many different combinations of the two-phonon wave vectors are possible for a given $\boldsymbol{k}$ of the exciton-like polariton. As a consequence l_2 is zero and the lineshape of the second LO phonon replica directly reflects the distribution of exciton polaritons in the initial state. In Fig. 13.13 we show the emission of ZnO at 55 K. The free exciton polariton is not seen in emission for the reasons given above. There is a little bound exciton emission (Sect. 14.1) around 3.34 eV and the LO phonon replicas for $m = 1, 2, 3$. The theoretical curves are calculated according to (13.15), (13.16) assuming that the lattice temperature and the temperature of the gas of exciton-like polaritons are equal. The fit coincides very nicely with experiment, thus confirming concepts developed above and proving especially that excitons are good quasi particles, the distribution of which can be described in many cases by Boltzmann statistics. A further example will be given for Cu_2O in Sect. 13.2.1.2.

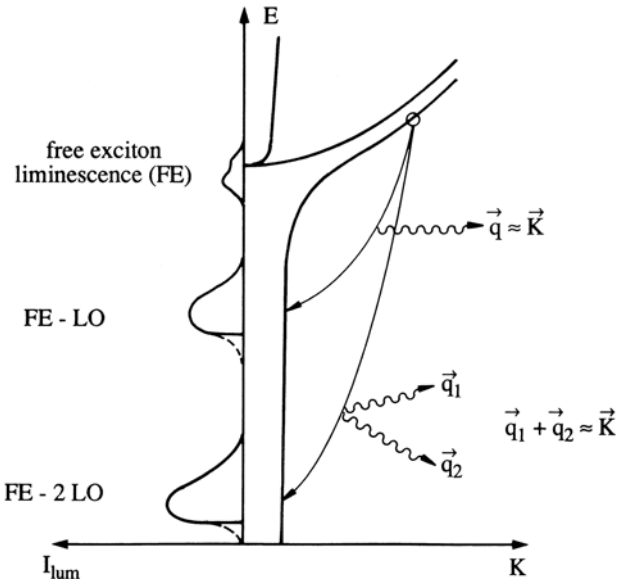


Fig. 13.12. Schematic drawing of the decay mechanisms of the exciton-mLO phonon emission processes

The small tail on the low energy side comes from the population in the bottleneck which is not considered in (13.15), (13.16) and from a homogeneous broadening of the excitons increasing with temperature. From (13.15), (13.16) one can deduce that the ratio of the integrated intensities of the first and the second LO-phonon replica is proportional to T

$$Q_{1,2} = \frac{\int I_1^{\mathrm{lum}}(\omega)\,\mathrm{d}\omega}{\int I_2^{\mathrm{lum}}(\omega)\,\mathrm{d}\omega} \propto T\,. \tag{13.17}$$

In Fig. 13.14 we give experimental data for $Q_{1,2}$ in ZnO for volume excitation and surface excitation. The first case has been realized by two-photon excitation with a ruby laser, which allows relatively homogeneous excitation of samples up to thicknesses in the mm range, and the second by UV excitation in the continuum states where the excitation depth is limited mainly by diffusion to values of the order of μm as discussed above.

Up to temperatures of 100 K the points follow nicely the predictions of (13.17), then they drop. This deviation is due to reabsorption effects caused by the absorption tail described by (13.12), which starts to influence the escape depth of the polaritons also in the $m = 1$ range at higher temperatures. This effect is evidently more pronounced for volume excitation than for surface excitation. In semiconductors with less polar coupling, such as GaAs, the LO-phonon replicas are less pronounced.

For those who work on luminescent ions in insulators and are therefore familiar with the concept of the Huang–Rhys factor S, it should be mentioned that for free excitons in most semiconductors S is below one. The large ratio of first to zero phonon intensity seen in Fig. 13.13 is not connected with a large value of S but with other processes that reduce the zero phonon emission, as outlined above (See e.g. [68V1, 79P1]).

LO-phonon replicas do not only appear in the excitonic luminescence spectra on the low energy side, but may also appear in case of sufficiently strong exciton LO-phonon coupling in the absorption spectra on the high energy side, extending into the continuum. Examples for one- and two-

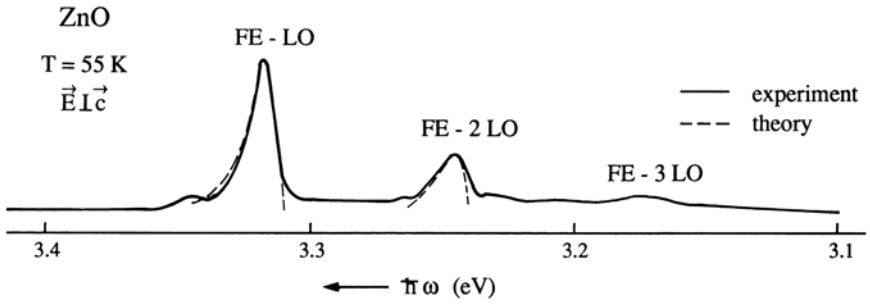


Fig. 13.13. LO-phonon replica in the luminescence spectrum of ZnO at 55 K. According to [75K1]

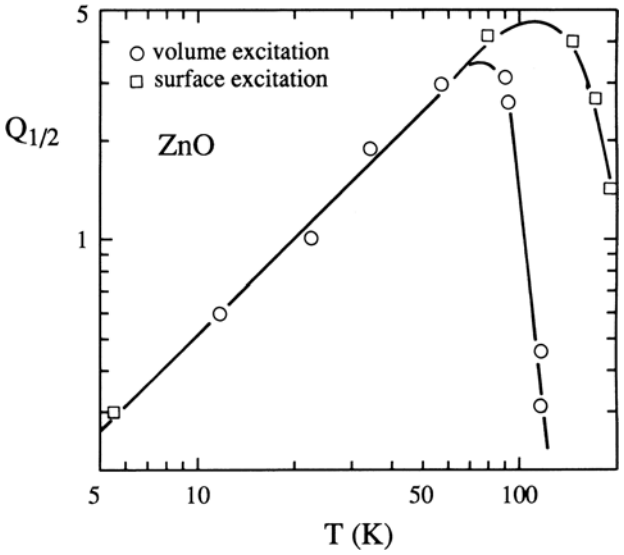


Fig. 13.14. The ratio of the integrated intensities of first and second LO-phonon replica in ZnO as a function of temperature for two different excitation conditions. According to [75K1]

photon absorption in Cu halides are found, e.g., in [66R1, 72B1] and for PbI_2 in [94W1]. The term exciton-phonon bound state was introduced in [68T1, 72K1].

The interaction of excitons and photons with acoustic phonons in Cu_2O leads to the introduction of the concept of phonoritons [00H1].

It should be mentioned that excitons can also be scattered also by emission of quasiparticles other than phonons onto the photon-like part of the polariton branch. Examples can be found, e.g., in Sect. 20.2 and we mention here scattering by magnons, which may occur in (anti) ferromagnetic semiconductors like $RbMnF_3$ or Cr_2O_3 [76S2, 99D1].

To conclude this discussion of free exciton luminescence, it should be mentioned, that a recent calculation [98K1] predicts, that a luminescence feature may occur at the position of the free exciton resonance, resulting, however, from the recombination of an electron–hole plasma (see Chap. 21) going under emission of a photon from a state containing n electron–hole pairs to one with $(n-1)$ pairs. Though these calculations do by no means rule out the existence of excitons as good quasi-particles, a luminescence feature at the exciton energy alone is not necessarily a proof of their existence. The observation of additional features, like intra excitonic transitions (see Sect. 13.3 and [03K1]) can help to clarify possible discrepancies.

As a rule of thumb it can be stated that the formation of excitons is likely for resonant and low or intermediate exciton densities (see Chap. 20). If the thermal energy $k_{\mathbf{B}}T$ is comparable to or larger than the exciton binding

energy $E_{\mathrm{x}}^{\mathrm{b}}$ the excitons will be thermally ionized with time. Under band-to-band excitation, exciton formation is likely for $k_{\mathrm{b}}T \ll E_{\mathrm{x}}^{\mathbf{b}}$ and for exciton formation times shorter than the exciton lifetime. The formation time is short, if the exess energy can be dissipated by LO-phonon emission especially in polar materials, while acoustic phonon emission is slower, especially in less polar materials. We come back to this topic in Chaps. 15 and 23.

13.1.4 Spectroscopy in Momentum Space

It is clear from the above discussion that the concept of exciton polaritons allows one to understand the spectra of reflection, transmission, and luminescence, but that a quantitative interpretation of the data usually involves a rather elaborate theoretical fit. Therefore various techniques have been developed which allow more or less directly the spectroscopy of exciton polaritons in momentum space, i.e., they provide the possibility of measuring the dispersion relation $E(\boldsymbol{k})$ more directly.

In this section we therefore recall briefly the consequences of "spatial dispersion", i.e., of the dependence of ω_0 on $\boldsymbol{k}$, and then present various methods of $\boldsymbol{k}$- or momentum-space spectroscopy.

As mentioned earlier in Chap. 5, the combination of the dielectric function $\varepsilon(\omega, \boldsymbol{k})$ (6.1) and the polariton equation forms an implicit representation of the polariton dispersion. For $\omega > \omega_{0,\mathrm{L}}$ we get two propagating modes in the sample or even more if the dispersion relation is complex (Fig. 13.4b or Fig. 13.11) or if the longitudinal branch couples to the radiation field as may occur for $\boldsymbol{k} \neq 0$ or for oblique incidence in uniaxial crystals. Since the $\boldsymbol{k}$ vectors of the various modes in the sample are different, the diffracted beams propagate in different directions, thus giving some meaning to the term "spatial dispersion". Below $\omega_{0,\mathrm{L}}$ we have at least one propagating and one evanescent mode, which however, for finite damping also acquires a small real part. The consequences of this fact for the reflection spectra have already been discussed above.

The first method of $\boldsymbol{k}$-space spectroscopy uses the analysis of the Fabry–Perot modes introduced in Sect. 3.1.6, which appear for example in as-grown, thin, platelet type samples with plane-parallel surface. For an example see Fig. 13.9b. As already pointed out there, transmission maxima occur when an integer number of half-waves fit into the resonator, i.e.,

$$k_m(\omega_m) = m\frac{\pi}{d}\,,\ m = 1, 2, 3 \ldots\,, \tag{13.18}$$

where d is the geometrical thickness of the sample and k is the real part of the wave vector in the medium. If ω_m and k_m are known for one m_1 then the dispersion relation can be reconstructed from (13.18) by reading the $\hbar\omega_m$ from the transmission spectrum and by progressing in steps of πd^{-1} on the k-axis. In Fig. 13.15 we show the dispersion of the $n_{\mathrm{B}} = 1$ A-Γ_5-exciton polariton in CdSe, together with a measured and a calculated reflection spectrum. The

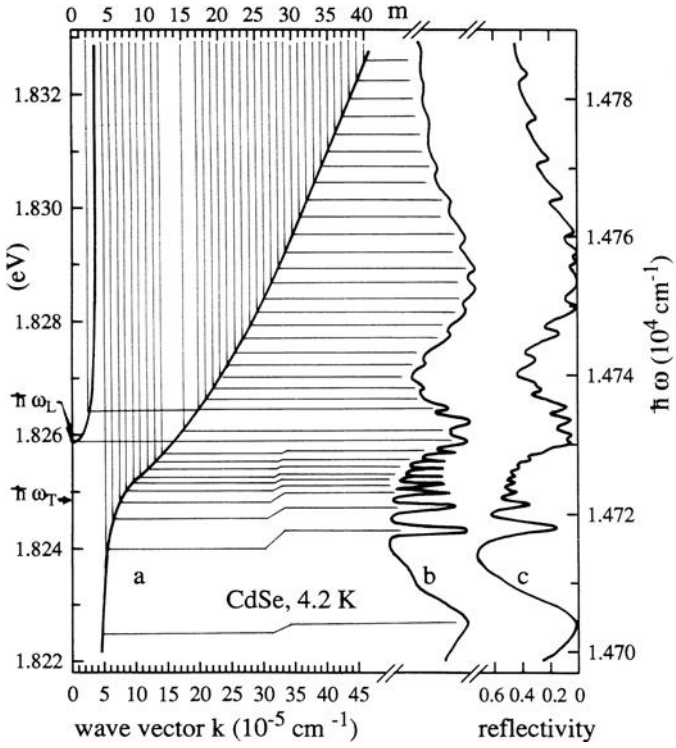


Fig. 13.15. The polariton dispersion in CdSe with an equidistant ruling of the $\boldsymbol{k}$-axis (**a**) and the observed (**b**) and calculated structure (**c**) of the resulting Fabry–Perot modes. According to [75K2]

condition for the reflection minima coincides with that for the transmission maxima in (13.18).

The equidistance on the k-axis is clearly visible and the good agreement of theory and experiment is obvious. Above $\omega_{0\mathrm{L}}$ one can clearly see closely spaced modes of the LPB and superimposed widely spaced ones from the part of the light travelling on the UPB. The fact that the small modulation decreases with increasing photon energy indicates that the fraction of light travelling on the LPB through the sample decreases with increasing energy above $\omega_{0\mathrm{L}}$.

Similar data have also been found in CdS and CuCl [80M1, 84M1, 93K1].

The next method is resonant Brillouin scattering. We introduced this scattering process with acoustic phonons in Sect. 11.1.4 and know that the Brillouin shift is directly proportional to the transfer of momentum due to the linear dispersion relation of acoustic phonons around the Γ point.

Figure 13.16 shows schematically the Brillouin scattering in the resonance region of an exciton polariton in a backward scattering configuration. The effect has been observed in GaAs, CdS, ZnSe and CuBr [78U1, 79S1, 80B3,

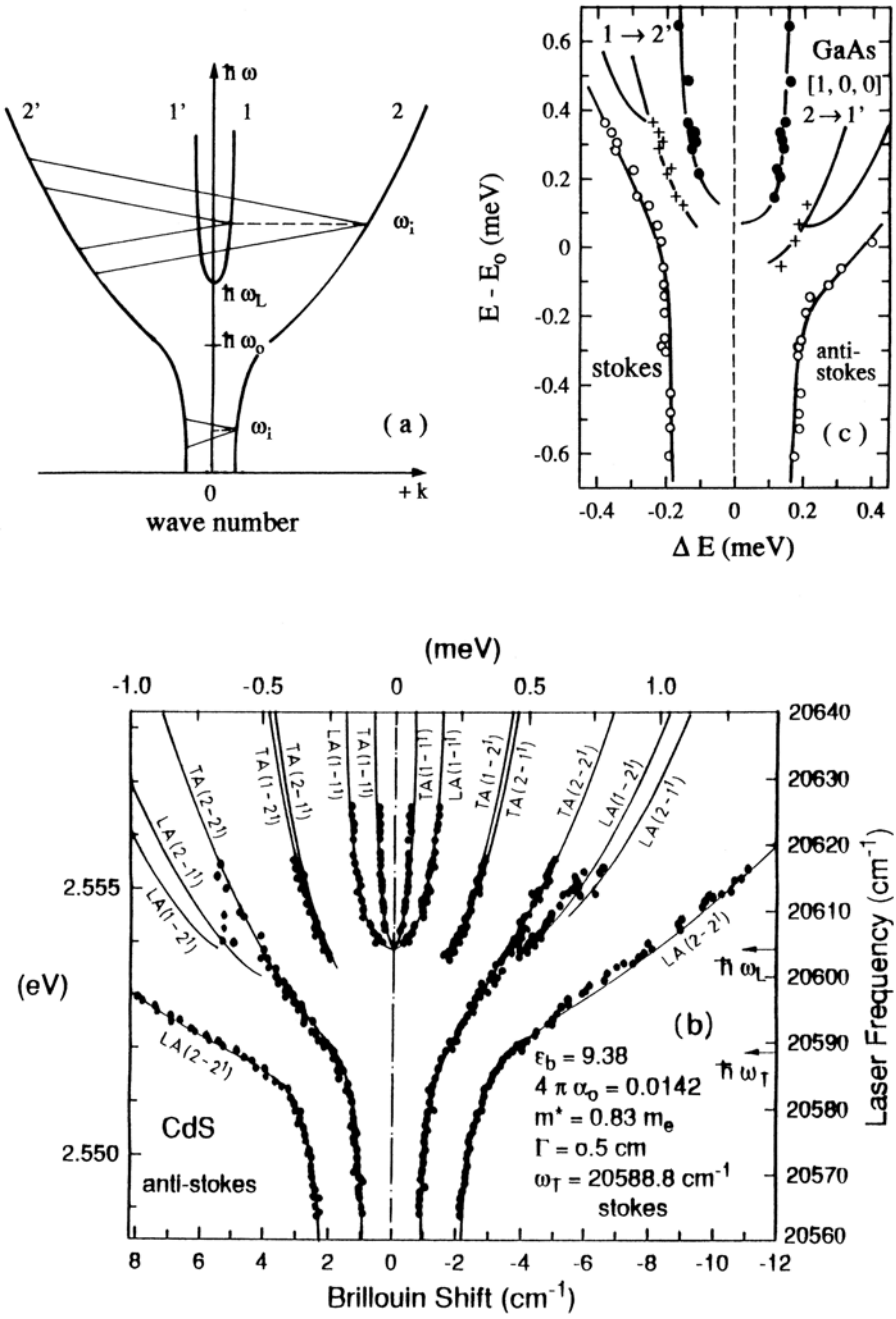


Fig. 13.16. Schematic drawing of Brillouin scattering in backward geometry in the resonance of an exciton polariton [78U1] (**a**) and experimental data for GaAs [78U1] (**c**) and for CdS [84M1, 84W1] (**b**)

80K1, 81C1, 84M1, 84S1, 84W1]. In Fig. 13.16b and c we give data for CdS and GaAs partly involving TA and LA phonons. The soild lines have been calculated in a fit procedure with the parameters in the inset; these of course, also allow the calculation of the polariton dispersion itself.

A close inspection of the Brillouin line shape shows that the damping constant is energy or $\boldsymbol{k}$-dependent [84S1]. This result has also been obtained in CuCl from nondegenerate four-wave mixing [85H1]. See Sect. 20.2 and 25.2 for this technique. Similarly one finds that Δ_{LT} is also a function of $\boldsymbol{k}$ [82O1, 85H1].

It is well known that refraction from a prism can be used to determine the real part of $\tilde{n}$ and thus of $\boldsymbol{k}$. If sufficiently thin ($d \lesssim 1\,\mu\mathrm{m}$), prism-shaped samples are available, it is also possible to extend this technique to the resonance region of the exciton polariton. Fortunately some CdS platelets grow in the desired form, presumably involving a small-angle grain boundary.

Figure 13.17 gives two examples for CdS, where n is given and not k, but where both quantities are simply connected with each other according to $\mathrm{Re}\,\{k\} = \omega/c\,\mathrm{Re}\,\{\tilde{n}\}$, see (2.36). In Fig. 13.17a one sees again the birefringence and dichroism for the polarizations $\boldsymbol{E} \perp \boldsymbol{c}$ (i.e., Γ_5) and $\boldsymbol{E} \parallel \boldsymbol{c}$ (i.e., Γ_1) of the $n_{\mathrm{B}} = 1$ A-exciton resonance known already from the reflection spectra of Fig. 13.4. In Fig. 13.17b the dispersion of the LPB can be followed up to $n \approx 20$ corresponding to $\varepsilon_1 = 400$ or $k = 2.5 \times 10^6\,\mathrm{cm}^{-1}$.

If one compares the time of flight of a picosecond pulse through a sample with its propagation through vacuum one can deduce the group velocity v_{g} and, with (2.13)b, the slope of the dispersion relation. In Fig. 13.18b experimental values are given for v_{g} at the lowest exciton polariton resonance in CuCl, together with a curve deduced from the dispersion relation of Fig. 13.18a. The excellent agreement between experiment and theory again proves the validity of the polariton concept. One can see from Fig. 13.18b that in CuCl v_{g} can be as low as $5 \times 10^{-5} c$. In the region above $\omega_{0\mathrm{L}}$ two pulses are created in the medium from one incident pulse due to the different group velocities on the LPB and the UPB. The spatial distance between them increases with sample thickness and this is another reason to call the $\boldsymbol{k}$-dependence of ω_0 "spatial dispersion". Similar experiments have been performed in GaAs, CdSe and CdS [79U1, 81I1, 83S1].

We conclude this subsection with two nonlinear methods of $\boldsymbol{k}$-space spectroscopy, namely two-photon spectroscopy or absorption (TPA) and two-photon (or hyper-) Raman scattering (TPRS, HRS), anticipating some of the results of Sect. 20.3. In TPA two laser beams are directed onto the sample, sometimes with a finite angle between them. They have energies $\hbar\omega_{1,2}$ and momenta $\hbar\boldsymbol{k}_{1,2}$. A TPA signal, i.e., a signal which occurs only if both beams are on, is observed only if the sum of energies and momenta coincide in the sample with those of an excited state, i.e., (see Fig 13.19a)

$$\begin{aligned} \hbar\omega_1 + \hbar\omega_2 &= E_{\mathrm{f}}\,, \\ \hbar\boldsymbol{k}_1 + \hbar\boldsymbol{k}_2 &= \hbar\boldsymbol{k}_{\mathrm{f}}\,. \end{aligned} \tag{13.19}$$

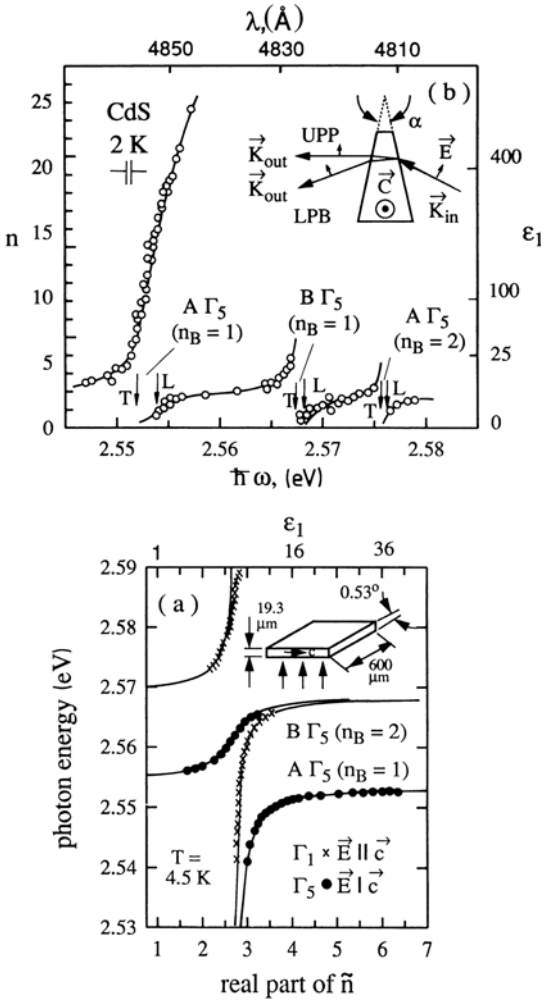


Fig. 13.17. Dispersion relations of the A- and B-exciton resonances determined with the thin-prism method [81B1, 84L1]

Due to the curvature of the LPB it is only possible with TPA to reach states on the longitudinal exciton branch or on the UPB. The selection rules are usually different from those of one-photon absorption and depend in addition on the polarizations of the beams relative to each other and to the crystallo-graphic axes. Examples will be given in Sect. 20.1. More recently, even three-photon spectroscopy has been used to determine the exciton polariton branches in various semiconductors and insulators [93F1, 94F1].

In HRS one or two incident laser beams create in the sample two (generally photon-like) polaritons. This two polariton state decays under energy and momentum conservation into a photon-like polariton $\hbar\omega_{\mathrm{R}}$, which is observed as a Raman-like emission, and another final state particle $\hbar\omega_{\mathrm{f}}$ which can be photon- or exciton-like. If a single incident beam is used to deliver

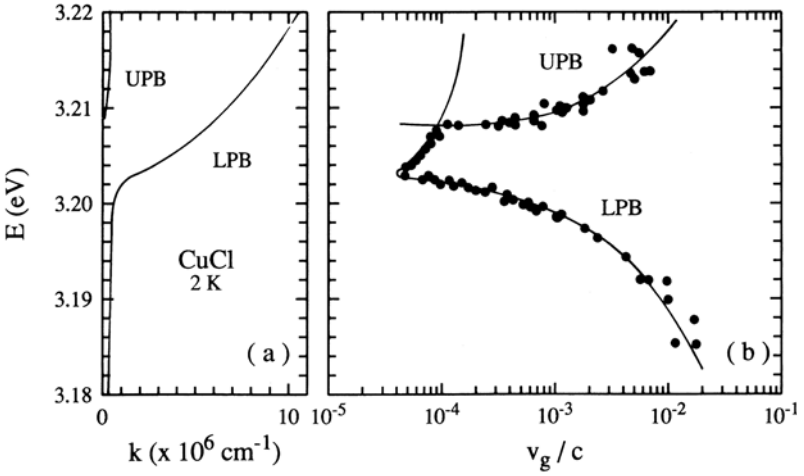


Fig. 13.18. The dispersion relation of the lowest free exciton resonance in CuCl (**a**) and the group velocity determined from the time-of-flight of picosecond laser pulses and from the dispersion relation (**b**) [79M1]

both incident quanta, i.e., $\hbar\omega_1 = \hbar\omega_2$, this process reads

$$\begin{aligned} 2\hbar\omega_1 &= \hbar\omega_{\mathrm{R}} + \hbar\omega_{\mathrm{f}} \\ 2\hbar\boldsymbol{k}_1 &= \hbar\boldsymbol{k}_{\mathrm{R}} + \hbar\boldsymbol{k}_{\mathrm{f}}\,. \end{aligned} \tag{13.20}$$

Usually one aims to have one or both of the incident quanta almost in resonance with the exciton and/or the biexciton state (see below). This choice enhances the transition probability due to the small resonance denominators appearing in perturbation theory. In Fig. 13.19 we show schematically TPA and the HRS processes where in both cases the longitudinal state has been reached. Figure 13.20 shows the polariton dispersion in CuCl around the lowest $n_{\mathrm{B}} = 1$, Γ_5-resonance and the states reached by two photon absorption and by HRS.

A large number of semiconductors have been investigated by TPA and HRS. TPA has also been used to detect higher states ($n_{\mathrm{B}} > 1$). For recent reviews see, e.g. [71F1, 81K2, 85H1]. Three-photon absorption has proved the validity of the polariton concept also in insulators like KCl [85B1, 94F1].

A slightly more involved observation of the polariton dispersion with short and thus spectrally broad pulses ($\Delta_{\mathrm{LT}} < \Delta_{\mathrm{Pulse}}$) has been already addressed in Sect. 5.3 in the context of propagation quantum beats between lower and upper branch polaritons. The example of the dipole-forbidden direct gap semiconductor for Cu_2O can be complemented by data for dipole-allowed ones, like ZnSe or InSe [96N1, 97N1].

The investigation of the properties of exciton polaritons came to a temporary stop, when structures of reduced dimensionality moved into the focus of interest, mainly quantum wells. At the beginning, these structures

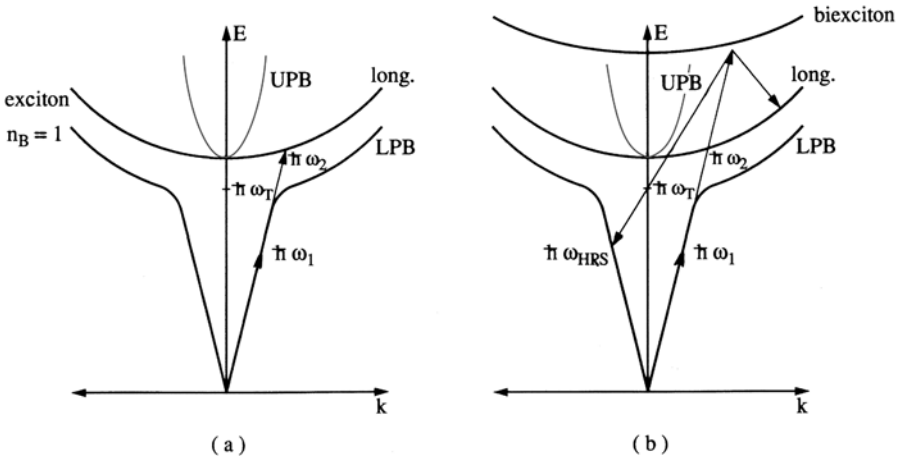


Fig. 13.19. Two-photon transition from the crystal ground state to the longitudinal exciton branch (**a**) and the hyper-Raman process which is almost resonant with the exciton and the biexciton state (**b**)

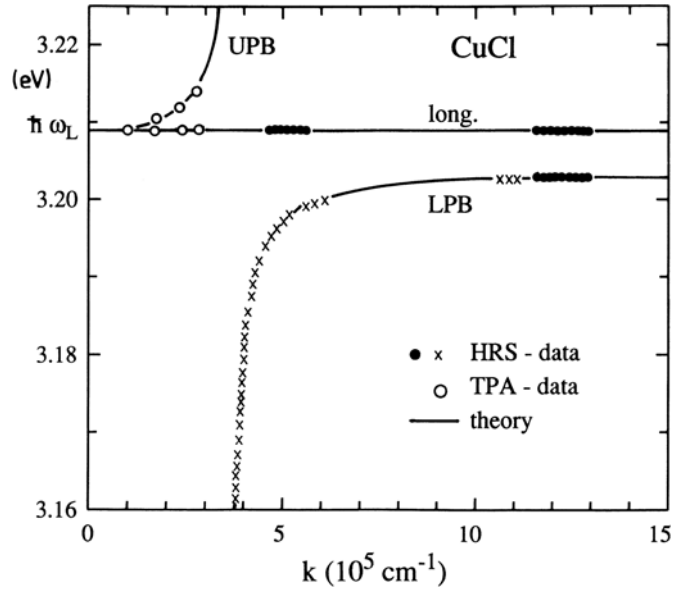


Fig. 13.20. The polariton dispersion of the lowest Γ_5-exciton resonance in CuCl measured by two-photon absorption and by hyper-Raman scattering [71F1, 81K2, 85H1]

exhibited such strong inhomogeneous broadening of the exciton resonances, that details of the polariton concept could not be observed. In the meantime, the concept of exciton polaritons has been developed for quantum

wells and wires (see Chap. 15) and it has been extended to cavity polaritons. See Chap 17. Even more recently, exciton polaritons were discussed again for materials like ZnO (see the references in Sect. 8.8 and Figs. in this chapter including [03C1]), GaN [99R1, 01R1, 02L1, 02P1] or GaAs [00T1, 01S1].

13.1.5 Surface-Exciton Polaritons

As already mentioned in Sect. 5.6, one also finds surface-exciton polaritons in the spectral region between the LPB and the longitudinal exciton. Their dispersion can be measured by attenuated (or frustrated) total internal reflection (ATR). A light beam is sent into a prism in such a way that total internal reflection occurs at its base (Fig. 13.21a). Under these conditions an evanescent wave leaks into the vacuum as is indicated schematically (see also Fig. 3.3). The frequency of this wave can be tuned trivally, and its wave vector $\boldsymbol{k}_{\|}$ by varying the angle of incidence. If a semiconductor is brought close enough to the base of the prism, the evanescent wave can couple to the surface-exciton polariton modes in the semiconductors if both $\hbar\omega$ and $\boldsymbol{k}_{\|}$ coincide. "Close enough" means distances of the order of λ, i.e., fractions of a μm and the realization of this condition involves some experimental skill. The coupling to the surface-exciton polariton mode attenuates the total reflection. In Fig. 13.21b measured and calculated ATR spectra are shown for the $n_{\mathrm{B}} = 1C\Gamma_1$-exciton resonance in ZnO, which has a rather large Δ_{LT} splitting of about 12 meV. The calculated dispersion relation of the surface-exciton polariton is given in Fig. 13.21c together with the states reached by ATR. Good agreement between experiment and theory can be claimed in both Figs. 13.21b and c. More details about surface-exciton polaritons can be found in [78L1] or in [82M1]a,d of Chap. 1.

13.1.6 Excitons in Organic Semiconductors and in Insulators

The number of well-characterized organic semiconductors is much smaller than that of inorganic ones. In [82L1] of Chap. 1 roughly 600 inorganic semiconductors are listed but only a dozen or so organic ones, with anthracene being a model substance.

On the other hand, research on organic semiconductors has increased greatly in recent years. It is hoped that applications as cheap, large area luminescence devices and as electronic devices like field-effect transistors will be developed.

Crystalline organic semiconductors usually contain several molecules per unit cell, anthracene, e.g., four molecules of $C_{14}H_{10}$. Correspondingly complex is the spectrum of lattice vibrations. Sometimes it is possible to distinguish between phonons or intermolecular modes, i.e., modes where the molecules in the unit cell oscillate relative to each other in the sense of acoustic and optic

modes and vibrons or intramolecular modes, where the atoms oscillate within one molecule. Obviously there may be some coupling between these two types of modes.

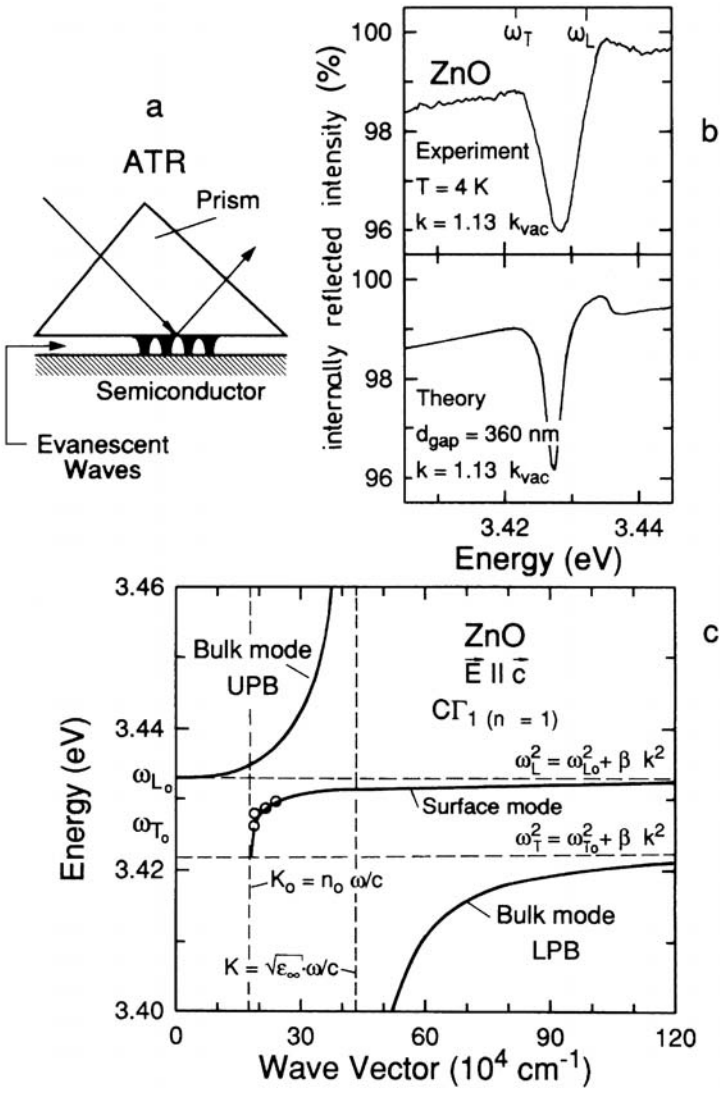


Fig. 13.21. The principle of the experimental technique of attenuated total reflection (ATR) (**a**) experimental and calculated ATR spectra for the $n_{\mathrm{B}} = 1C\Gamma_1$-exciton in ZnO (**b**) and the resulting dispersion of the surface and volume polaritons (**c**). According to [78L1]

Concerning the electrons, there are valence bands (highest occupied molecular orbitals, HOMO) and conduction bands (lowest unoccupied molecular orbitals, LUMO) with a gap between them.

The bands are usually rather narrow, due to the small electronic overlap between molecules, resulting in large effective masses of the charge carriers (electrons and holes) and low mobilities ranging from 1 to $10\,\mathrm{cm}^2/\mathrm{V\,s}$.

The interband excitations in the electronic system form excitons, however with some characteristic differences from the (Wannier) excitons in most inorganic semiconductors.

One usually deals in organic semiconductors and especially in the wide gap ones with Frenkel excitons, i.e., electrons and holes reside in the same molecule. This close spatial confinement results in singlet-triplet splittings which can easily exceed 1 eV. See again, e.g., [82L1] of Chap. 1 and below. The triplet states have low oscillator strength, low luminescence yield and long lifetime in the ms range. The singlet states have high oscillator strength and a short life time in the ns-range and a tendency to relax into the lower lying triplet states.

Depending on the nature of the chemical bonding, coupling of excitons to (Raman- and/or IR-active) phonons may be strong, possibly leading to self-trapping.

Furthermore there are charge transfer (CT) excitons, in which the two carriers (i.e., electron and hole) reside in adjacent molecules forming, e.g., a pair of a positively and a negatively charged anthracene molecules.

Excitons with large oscillator strength form polaritons in sufficiently pure materials in a similar way as in inorganic semiconductors.

Below we give two examples of the optical properties of excitons in organic semiconductors.

In Fig. 13.22a,b we give an overview of the absorption spectra of anthracene and tetracene. One can see the transitions from the crystal ground states (S_0) to triplet (T_1) and singlet (S_1) excitons. The oscillator strength of the first transition is too low to observe it directly in absorption.

Therefore, the photoluminescence excitation spectra (see Chap. 25) of the delayed triplet fluorescence are plotted. The singlet states have, as already mentioned, a high oscillator strength resulting in values of the absorption coefficients exceeding $3 \times 10^{10^5}\,\mathrm{cm}^{-1}$. CT stands for charge transfer exciton, IPC for the intrinsic photoconductivity threshold and EPE for the external photoemission threshold. The crystals are anisotropic, therefore the spectra are given for the polarizations parallel and perpendicular to the crystallographic b axis.

Figure 13.22c shows a close-up of the absorption spectrum of the singulet exciton polariton absorption, measured here with the technique of photocurrent excitation spectroscopy (see Chap. 25).

In recent years oligothiophenes and especially α-hexathiophene are under intensive investigation concerning both their optical properties and their (con-

troversial) applications in optoelectronic devices like field effect transistors or luminescence diodes. See, e.g., [00F1, 01F1].

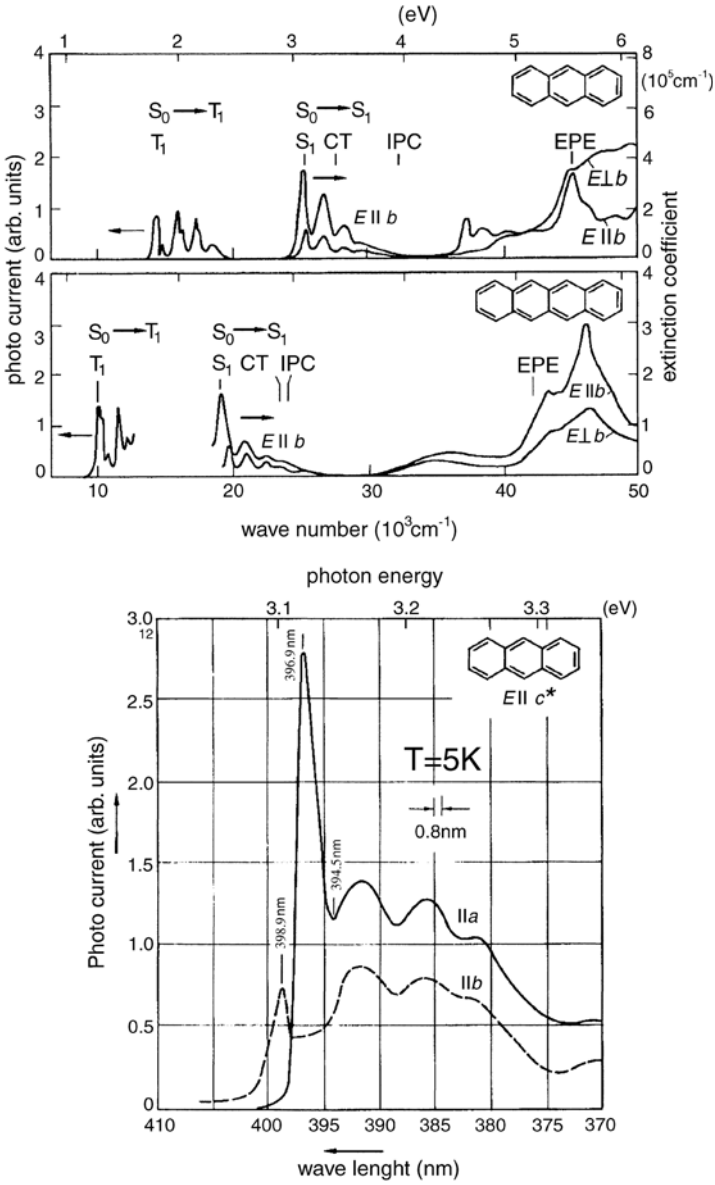


Fig. 13.22. Overview of the absorption spectra of anthracene (**a**) and tetracene (**b**) and photocurrent excitation spectra (PCE) of the singulet exciton in anthracene. According to [77C1, 79S2] and [82L1] of Chap. 1

These two papers also contain together with [82L1] of Chap. 1 references to various (text)books on the properties of organic semiconductors. For further recent publications of optical properties of organic and inorganic insulators and organic semiconductors see e.g. [04P1] or [94E1] of Chap. 9.

As an example of the exciton polariton dispersion in an inorganic insulator, namely KI, we show the dispersion relation obtained from three-photon sum and difference frequency spectroscopy in Fig. 13.23. Note the high value of $\Delta_{LT} \approx 100\,\text{meV}$. Further information including similar compounds, is found, e.g., in [85B1]. More information on excitons in insulators, including self trapping or solid rare gases can be found, e.g., in [64K1, 67K1, 67O1, 78G1, 83G1] or in [94E1] of Chap.9.

13.1.7 Optical Transitions Above the Fundamental Gap and Core Excitons

So far, this chapter has concentrated on electronic transitions in semiconductors close to the fundamental gap. But transitions from deeper valence bands and/or into higher conduction bands are also possible. The structures connected with these transitions are obviously found at $\hbar\omega > E_g$, i.e., in the (V)UV part of the spectrum.

The absorption in the exciton continuum is influenced by the Sommerfeld enhancement as discussed in [91O1] and shown in Fig. 13.9d. Its importance, i.e., the dependence on the transition energy of this Sommerfeld factor de-

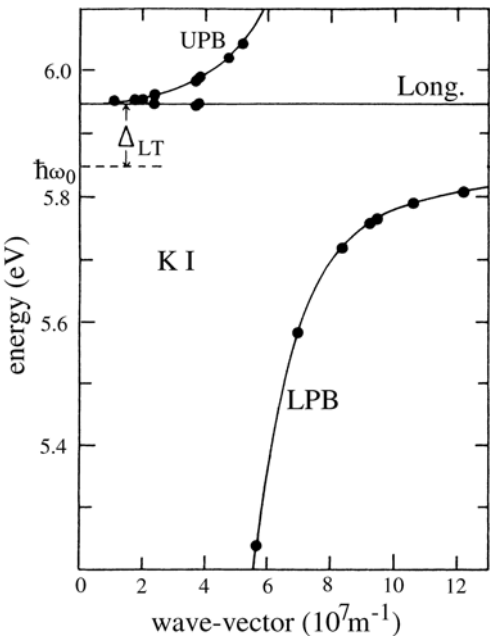


Fig. 13.23. The exciton polariton dispersion in KI obtained from three-photon spectroscopy including the lower and upper polariton branches and the longitudinal exciton. According to [94F1]

creases with increasing energy above the onset of the excitonic ionization continuum, (the band gap) and we are left more or less with band-to-band transitions. We therefore discuss their properties here and show various examples, but return later to the influence of electron–hole interaction.

We introduced in Sect. 8.9 the concept of critical points in the band structure and their consequences for the density of states. For band to band transitions it is not the density of states in a single band that determines the spectral shape of the optical properties, i.e., of the reflectivity $R(\omega)$, of the complex dielectric function $\varepsilon(\omega) = \varepsilon_1(\omega) + \varepsilon_2(\omega)$ or of the complex index of refraction $\tilde{n}(\omega) = n(\omega) + \mathrm{i}\kappa(\omega)$, but the so-called combined density of states. This concept will be explained with the schematic band structure of Fig. 13.24.

We show vertical transitions between a filled valence band and an empty conduction band for a one-dimensional $\boldsymbol{k}$-space. We assume as in the dipole approximation, that the wave vector of the photon is negligible. At an arbitrary photon energy (indicated by arrow 'a') there is only one transition possibility for a given photon energy, i.e., only one point in $\boldsymbol{k}$-space for which transitions between the bands are possible (in a 2-dimensional $\boldsymbol{k}$-space actually along a line etc.). If the two bands are parallel it is possible for transitions with the same energy to occur over a small interval in $\boldsymbol{k}$-space and we can

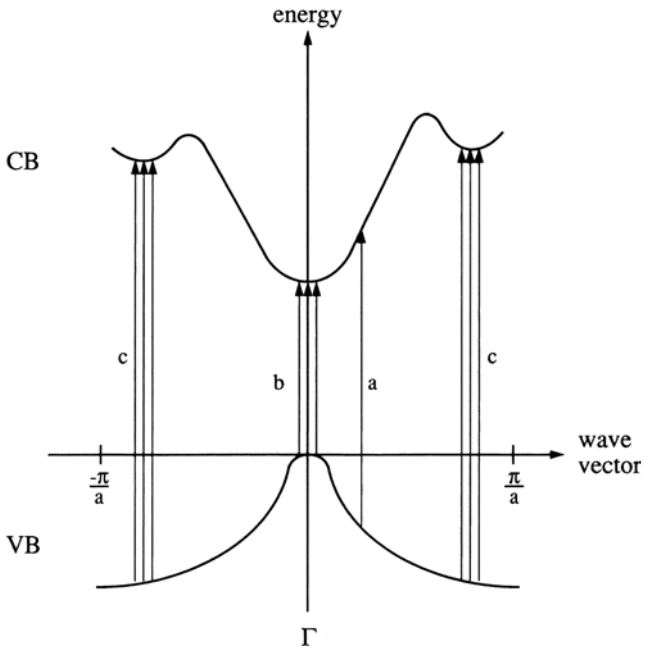


Fig. 13.24. Band-to-band transitions at an arbitrary point of the Brillouin zone (**a**) and at critical points of the combined density of states where the two bands are parallel (**b**, **c**)

expect a distinct structure in the absorption spectrum at this photon energy. Such a situation occurs if either the $\mathrm{grad}_{\boldsymbol{k}} E_j(\boldsymbol{k})$ of both bands is zero as shown in Fig. 13.24 for the Γ-point or if the distance between the bands is constant, i.e., they are parallel, having the same value of $\mathrm{grad}_{\boldsymbol{k}} E_j(\boldsymbol{k})$.

If we go from the one-dimensional $\boldsymbol{k}$-space of Fig. 13.24 to the usual three-dimensional one, the situation remains qualitatively similar. This means, that we can expect structures, whenever

$$\mathrm{grad}_{\boldsymbol{k}}[E_{\mathrm{i}}(\boldsymbol{k}) - E_j(\boldsymbol{k})] = 0\,, \tag{13.21}$$

where the indices i and j stand for valence and conduction bands, respectively. If we insert (13.21) into (2.77) we get the so-called combined density of states, which is the one relevant for the description of optical band-to-band transitions. If we further make the generally realistic assumption that the transition matrix element varies only smoothly with the transition energy, we expect in the band-to-band absorption spectra features, that look similar to

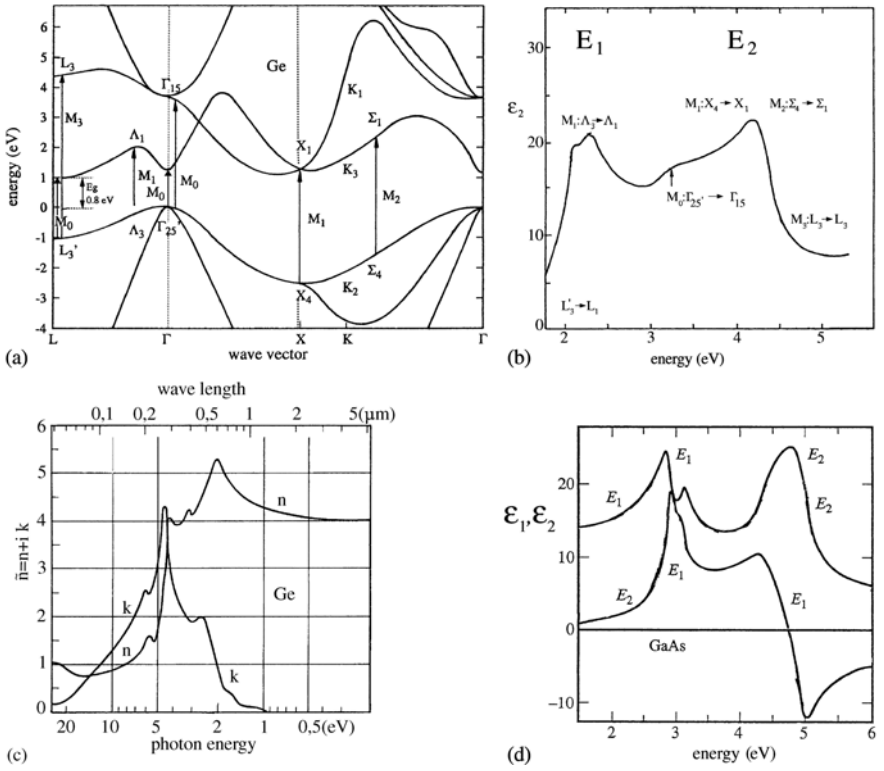


Fig. 13.25. The band structure of Ge (**a**), the spectra of ε_2 for Ge (**b**), of n and κ of Ge (**c**) and of ε_1 and ε_2 of GaAs (**d**). According to [90K1] of Chap. 1 (**a**), [90K1,96Y1] of Chap. 1 (**b**), [76P1] of Chap. 3 (**c**), and [96Y1] of Chap. 1 and [83A1] (**d**)

the (combined) density of states in Sect. 8.9 for various critical points (and various quasi dimensions of the system).

In experiments, one usually measures the reflection spectrum from $\hbar\omega \approx E_g$ up to higher energies and applies Kramers–Kronig relations (Chap. 6) to deduce from $R(\omega)$ either $\varepsilon_1(\omega)$ and $\varepsilon_2(\omega)$ or $n(\omega)$ and $\kappa(\omega)$.

Another possibility is to use ellipsometry to determine these quantities. See Chap. 25

In Fig. 13.25 we show various examples for higher energy transitions starting from the band-structure of Ge. Other examples may be found in [80H1, 83M3, 93M1, 96T1] or in [82L1, 90K1, 96Y1] of Chap. 1.

The various transitions at critical points and the type of the critical point in the combined three-dimensional density of states ($M_i = 0, 1, 2, 3$) are indicated. In Fig. 13.25b we show the spectra of ε_2, where the various transitions are labeled. Theory is able to reproduce these experimental data reasonably well, and tend, however, to vary slightly from author to author (compare e.g. data for Ge in [90K1, 96Y1] of Chap. 1). In Fig. 13.25c we show rather old data for n and κ, again for Ge and in Fig. 13.25d finally ε_1 and ε_2 for GaAs. In c and d the absorptive and dispersive behavior of the real and imaginary parts is obvious. It is remarkable that at these energy scales neither the low-

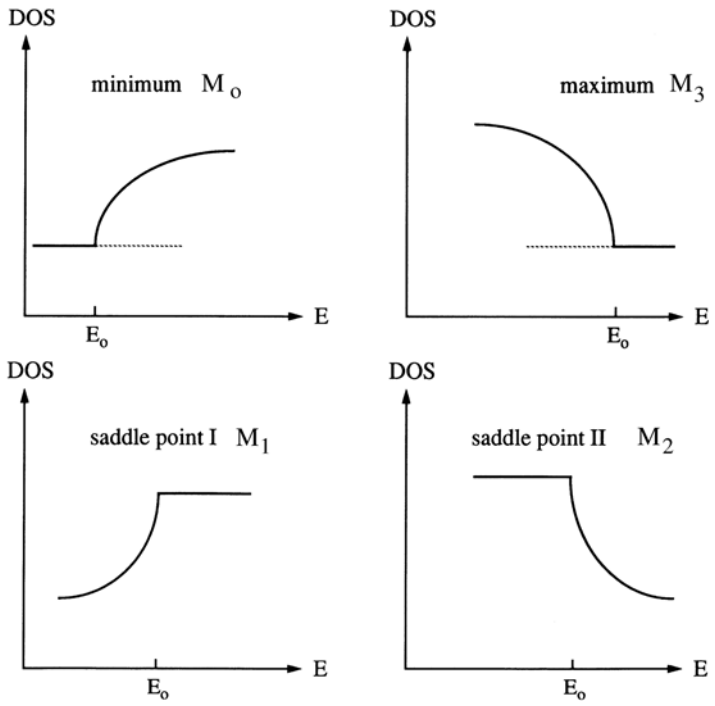


Fig. 13.26. The (combined) density of states for the three-dimensional critical points M_0 to M_3. See also Sect. 8.9.

est M_0 critical point, which results in the beautiful exciton spectra shown for GaAs in Fig. 13.9c, nor the indirect transition (e.g., at the indirect gap of Ge) contribute significantly to the overall spectra. The two peaks occurring around 2 eV and 5 eV are usually called E_1 and E_2. They are dominated by M_1 and M_2 critical points.

The fact that the structures in Fig. 13.25 look generally more like peaks and less like Fig. 13.26 is to a large extent due to the neglect of the electron–hole Coulomb interaction in the latter. The resulting excitons, which are responsible for the rich variety of phenomena discussed in the preceding part of this chapter, also modify the optical properties of the higher energy transitions (core excitons, saddle-point excitons). However, these excitons are often not resolved as individual structures, even at low temperatures, because the damping of these states is too high (or the phase-relaxation time too short) since they can decay rapidly into lower energy states.

To illustrate this statement, we give in Fig. 13.27 an example for CdTe. The direct gap is at 1.4 eV. The E_1 feature is located around 3.5 eV. The dashed line shows the expected spectrum for ε_2 bearing in mind the M_1 critical point in the combined density of states. The measured spectrum deviates significantly, but coincides nicely with a model calculation taking, in addition to the M_1 critical point, the Coulomb interaction between electron and hole, i.e. excitonic effects into account.

If we got to even higher photon energies than those shown in Fig. 13.25 we come to the X-ray regime. There we expect the K, L, ... absorption edges of the elements forming the semiconductor [93M1, 96T1]. These deep core levels are influenced only slightly by the chemical binding, but the absorption spectrum is still influenced by the position and proper-

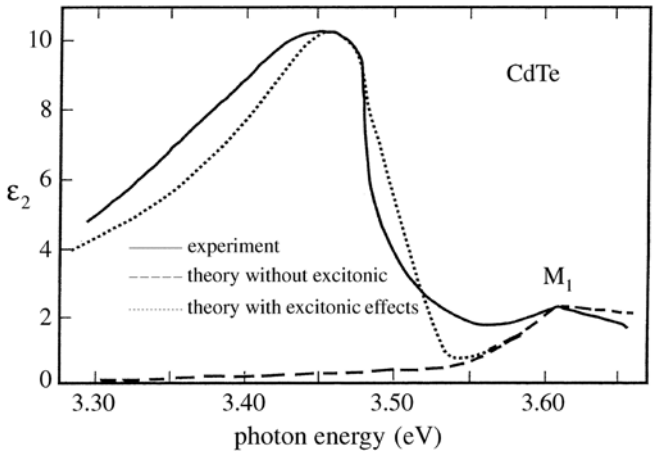


Fig. 13.27. The imaginary part of ε_1 in the region of the E_1 feature in CdTe compared with two calculations taking electron–hole Coulomb correlations into account or not. According to [69K1, 70W1] and [96Y1] of Chap. 1

ties of the conduction band since the electron has to be brought to an empty state, or in other words, by the interference of excited electrons scattered from neighboring atoms. These facts are exploited in EXAFS (extended X-ray absorption fine structure) and XANES (X-ray absorption near edge structure). The absorption coefficient may still reach considerable values in the X-ray regime, but has an overall tendency to decrease with increasing photon energy as shown in Fig. 13.28. The imaginary part of $\tilde{n}$ is in contrast rather small in this region and the real part is very close to unity, approaching this value above the highest resonance from below.

In Fig. 13.28 we show these trends for NaCl. This is an insulator but this does not matter to much in the present context. The data compiled in Fig. 13.28 are remarkable, because they cover eight orders of magnitude in photon energy and show nicely the optical phonon resonance in the IR, the exciton and band edge in the UV slightly below 10 eV leading to the element specific absorption edges in the X-ray regime.

More information about core excitons and related topics can be found in [65C1, 65C2, 93M1, 94J1, 96T1] or [90K1] of Chap. 1 and references given therein.

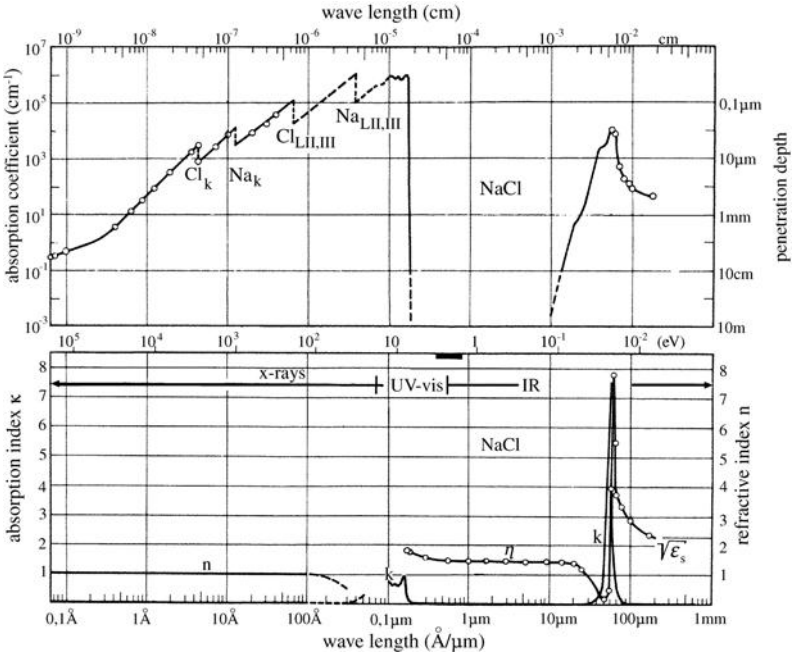


Fig. 13.28. The spectra of the absorption coefficient of NaCl (**a**) and of the real and imaginary parts of $\tilde{n}$ (**b**). According to [76P1] of Chap. 3

13.2 Forbidden Exciton Transitions

In the preceding section, we treated the optical properties of exciton resonances with large oscillator strength, requiring the polariton concept for a quantitative description. There are also intrinsic excitons which, for various reasons, couple only weakly to the radiation field. In this case the weak coupling approach is usually but not always sufficient for their description.

Reasons why excitons can have low oscillator strength include spin-flips, dipole-forbidden transitions, or the necessity to involve a third particle, e.g., a phonon in their creation. We give some examples in the following.

13.2.1 Direct Gap Semiconductors

We start with direct gap semiconductors.

13.2.1.1 Triplet States and Related Transitions

Even in direct gap semiconductors with dipole-allowed band-to-band transitions there are exciton states with very small oscillator strength, so-called forbidden transitions. Among them there are spin triplet states (see Sect. 9.2), which involve a spin-flip, longitudinal excitons, and states whose symmetry forbids the transition in the dipole approximation (Sect. 26.5). The latter situation can occur e.g. for $n_\mathrm{B} \geq 2$ and L or $L_z = 1, 2 \ldots$. Such exciton states can sometimes be observed in one-photon transitions in higher-order perturbation theory, e.g. as electric quadrupole or magnetic dipole transitions. This, however, is only possible if no strong one-photon transition occurs in the same spectral range, which masks the weak transition. An example is again CdS. We have already stated that the $n_\mathrm{B} = 1$ A-excitons are dipole allowed only for the polarization $\boldsymbol{E} \perp \boldsymbol{c}$, i.e., Γ_5. For $\boldsymbol{E} \parallel \boldsymbol{c}$ one can weakly see the AΓ_6 triplet or the AΓ_5^L. In the latter case the finite angle of aperture of every real light beam plays a role because small deviations from $\boldsymbol{k} \perp \boldsymbol{c}$ lead to a small oscillator strength of mixed-mode states according to (13.10). In Fig. 13.29 we show a transmission spectrum of CdS around the $n_\mathrm{B} = 1$ A-exciton states for the orientation $\boldsymbol{E} \parallel \boldsymbol{c}$; $\boldsymbol{k} \perp \boldsymbol{c}$, where the A$\Gamma_6$ triplet and the AΓ_5^L states are seen. A further example for ZnO will be given in Sect. 16.1.

The other way of reaching some of these states is via two-photon absorption (TPA). The selection rules for TPA differ from those for one-photon absorption as will be outlined in Sect. 26.5.

A first example has already been presented in Fig. 13.20, where the longitudinal branch of the Γ_5 exciton is seen in TPA. Another example for states with $n_\mathrm{B} \geq 2$ will be given in Sect. 16.1, where we discuss the influence of magnetic fields on excitons.

13.2.1.2 Parity Forbidden Band-to-Band Transitions

There is a group of semiconductors that have a direct gap at the Γ point. But the transition between the uppermost valence band and the lowest conduction band is dipole forbidden because both bands have the same parity. The best investigated material of this group is Cu_2O, but there are many others such as SnO_2, TiO_2, and GeO_2. In these materials exciton states with an S envelope (i.e., $L = 0$) are all dipole forbidden.

For $n_B \geq 2$ there are also envelope functions with $L = 1$, which have odd parity. Via their envelope, these states acquire a weak oscillator strength and can be seen in absorption.

Figure 13.30 gives an example of the "yellow" exciton series in Cu_2O, where the one- and two-photon absorption spectra are compared. The $n_B = 1$ states of the yellow and green exciton series have been found by two-photon absorption. Their positions are indicated. These states do not fit into the n_B^{-2} series, for the reasons discussed already in Sect. 9.2. However, even for the tiny quadrupole oscillator strength of this $n_B = 1$ S-exciton, it has been shown that the polariton effect exists and can be measured. See [91F1] and the polariton propagation quantum beats discussed in Sect. 5.3.

The 1S ortho exciton of the yellow series, to which we concentrate in the following, has a binding energy of 140 meV while the nP (n > 1) states follow a hydrogen series with $Ry^* = 93.26$ meV. See, e.g., [98J1] and references therein.

Due to the small exciton Bohr radius in Cu_2O of about 0.53 nm the singlet-triplet (or ortho-para) splitting is much larger than in usual semiconductors and amounts to 12 meV. The lower lying para-exciton is optically forbid-

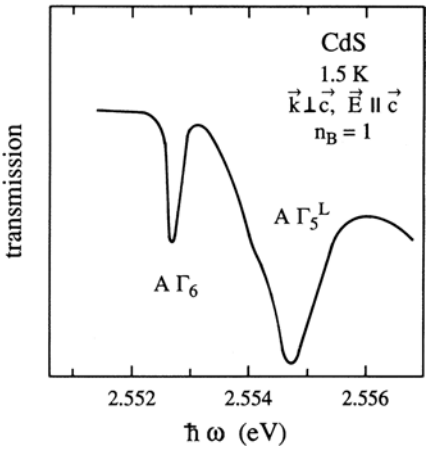


Fig. 13.29. A transmission spectrum of CdS in the resonance region of the $n_B = 1A$-exciton for the orientation $\boldsymbol{k} \perp \boldsymbol{c}$ showing the dipole-forbidden transitions $A\Gamma_6$ and $A\Gamma_5^L$. According to [82B1]

den to all orders of perturbation theory. Its position has been deduced by applying magnetic (or strain) fields, which mix ortho and para states, and by extrapolating to zero field [79M1]. Additionally, it has been claimed that a weak Γ_5^- phonon replica of the para-exciton occurs in luminescence around 2.01 eV [78B1].

Since Cu_2O is a kind of model substance and textbook example for excitons and brings after several decades of intense research still new and exciting results (also see Sect. 20.5) we dwell a little bit longer and give in Fig. 13.31 some further examples of the optical properties of its yellow exciton series.

The 1s ortho exciton states around $\boldsymbol{k} = 0$ show up only weakly in absorption or luminescence as seen in Figs. 13.30a and 13.31a,c, respectively. But the LO phonon replica reproduce nicely the square root density of states of this effective mass particle. The onset of this Γ_{12}^- LO-phonon absorption band is shown together with a calculated square-root function in Fig. 13.31a. This is one of the few cases, where this square root DOS shows up directly in experiment.

Superimposed on this absorption band (also including higher bands) is the absorption of the nP ortho-excitons as shown in Fig. 13.31b. Their line shape is asymmetric resulting from the constructive and destructive interference of the transition amplitudes on both sides of the resonance, respectively, with the LO-phonon continuum in the sense of a Fano type resonance.

The exciton luminescence is similarly dominated by the phonon replicas as seen in Fig. 13.31c showing both Stokes and anti-Stokes emissions. The long exciton lifetime results in thermalization of the excitons with the lattice. The theoretical curves in this figure are calculated essentially as a product of the

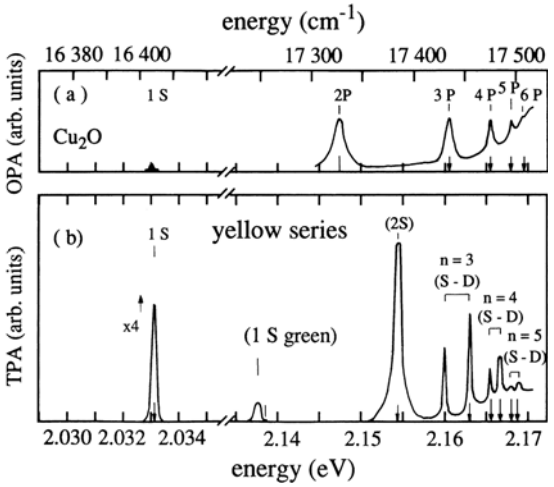


Fig. 13.30. The absorption of the forbidden "yellow" exciton series in CuO_2 in one and two absorption (OPA and TPA, respectively). According to [81U1]

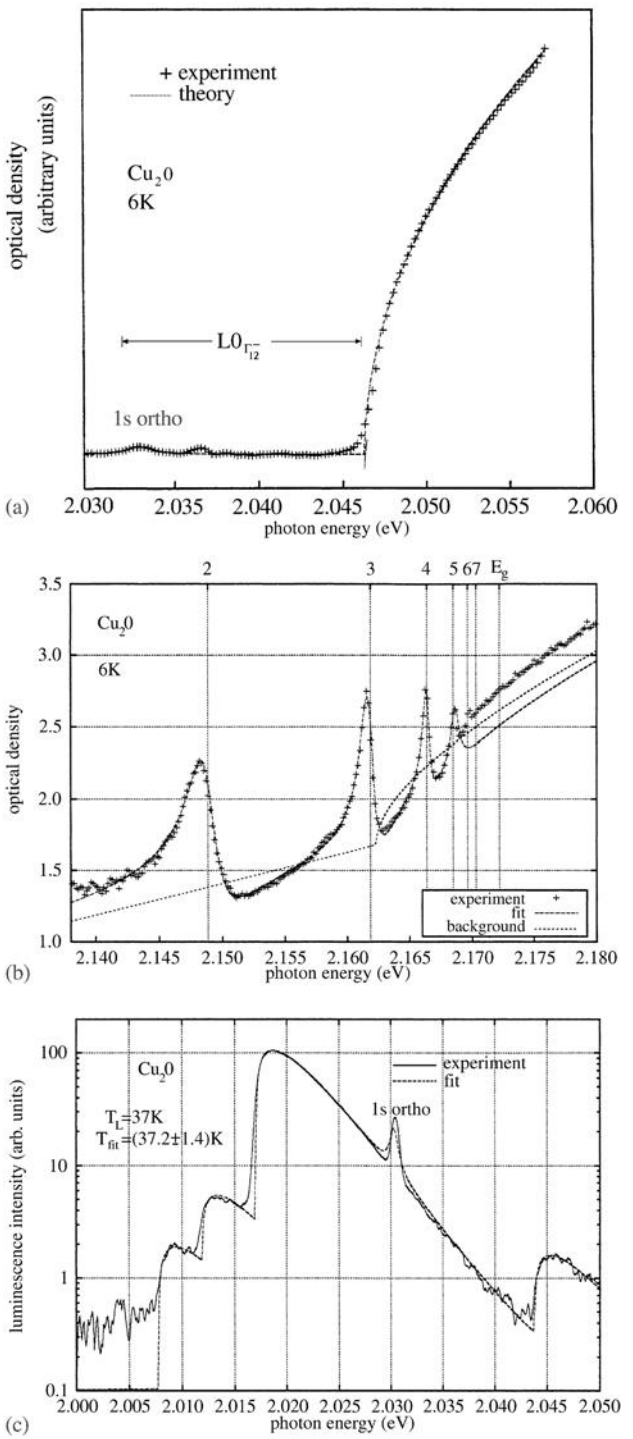


Fig. 13.31. Absorption (**a**, **b**) and luminescence spectra (**c**) of Cu_2O. According to [98J1]

square root DOS and a Bose–Einstein or Boltzmann occupation probability. The exciton temperatures deduced from this fit and from the ratio of Stokes and anti-Stokes emission agree very well with the lattice temperature. This also holds for temperatures down to or even below 5 K [98J1].

Further recent topics in Cu_2O exciton spectroscopy include the determination of the absolute luminescence yield as a function of temperature [99J1,01J1], the observation of the excitonic Lyman series (see Sect. 13.3), the discussion about Bose–Einstein condensation or of superfluid exciton transport (see Sect. 20.5), the observation of the coupling of exciton polaritons with acoustic phonons, the so-called phonoritons [00H1], the observation of extremely narrow 1s absorption peaks [03D1], the investigation of the effective masses of ortho- and para-excitons, which become different and direction dependent du to exchange interaction [04D1], or of extremely weak LO-phonon assisted absorption below the exciton resonance deduced from luminescence and generalized Planck's law [99J1].

13.2.2 Indirect Gap Semiconductors

Many semiconductors, some of great technological importance such as Si or Ge, but also GaP or AgBr, have an indirect band structure, as explained in Sect. 8.8. Consequently, the lowest free excitons occur at $\boldsymbol{k} \neq 0$, as outlined in Sect. 9.2. Due to momentum-conservation, these states cannot couple directly to the radiation field. In both absorption and emission processes a third, momentum conserving particle has to be involved, usually a phonon. At low temperature only photon absorption with the creation (i.e., emission) of one or more phonons is possible. At higher temperatures, when the relevant phonon states are populated thermally with finite probability, absorption of light quanta is possible with both phonon emission and absorption.

Figure 13.32a shows schematically the process of absorption accompanied by the creation or emission of a phonon, and in Fig. 13.32b, as an example, the absorption spectrum of GaP at low temperature. The exciton states do not appear as peaks as in direct gap materials (Fig. 13.9) but instead as onset energies, since the participation of phonons allows one to reach the whole density of states. Thus the absorption spectra are given in the simplest approximation by sums of expressions like

$$\alpha_{\mathrm{em}}(\hbar\omega) \propto \alpha_0[\hbar\omega - E_{\mathrm{ex}}(\boldsymbol{k}_0) + \hbar\omega_{\mathrm{ph}}]^{1/2}(1 + N_{\mathrm{ph}})\,, \tag{13.22a}$$

$$\alpha_{\mathrm{abs}}(\hbar\omega) \propto \alpha_0[\hbar\omega - E_{\mathrm{ex}}(\boldsymbol{k}_0) - \hbar\omega_{\mathrm{ph}}]^{1/2}N_{\mathrm{ph}}\,, \tag{13.22b}$$

where α_0 contains the transition matrix element squared, which is assumed not to depend (or only weakly) on the momentum of the phonon. $E_{\mathrm{ex}}(\boldsymbol{k}_0)$ is the exciton energy at the indirect minimum, $\hbar\omega_{\mathrm{ph}}$ the phonon energy for $\boldsymbol{k} \approx \boldsymbol{k}_0$, and N_{ph} is the number for the phonons in a given mode.

The photon energy in the square root term, which describes the density of states of the excitons above $E_{\mathrm{ex}}(\boldsymbol{k}_0)$, has to be chosen so that the argument is positive. For other values of $\hbar\omega_{\mathrm{ph}}$, α_{em} and α_{abs} are zero.

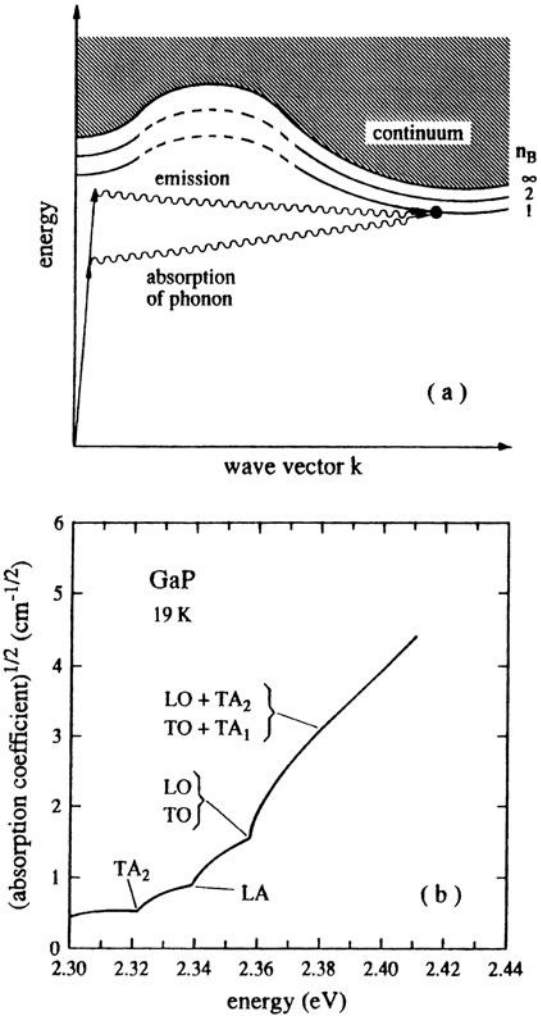


Fig. 13.32. The creation of an exciton in an indirect gap semiconductor accompanied by reation or annihilation of a momentum-conserving phonon (**a**). The absorption spectrum of GaP (**b**). According to [70M1]

The participation of the phonon makes the transition probability for absorption in indirect-gap semiconductors several orders of magnitude smaller than that for direct, dipole-allowed transitions. Typical values of the "indirect" absorption coefficient are in the range

$$1\,\mathrm{cm}^{-1} \leq \alpha_{\mathrm{ind}} \leq 10^2\,\mathrm{cm}^{-1}\,. \tag{13.22c}$$

These values are so low that no significant structures appear in the reflection spectra.

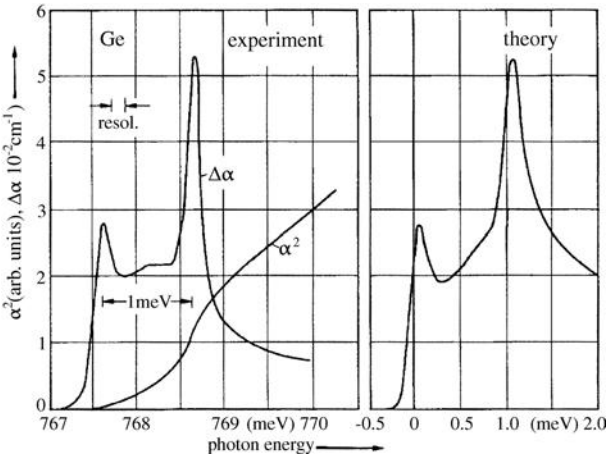


Fig. 13.33. The spectra of the square of the absorption coefficient of the phonon-assisted absorption onset in Ge and its wavelength modulated signal. According to [75F1]

In Fig. 13.33 we show the onset of the phonon assisted exciton absorption band of Ge (actually $\alpha^2\,(\hbar\omega)$) and the wavelength modulated spectrum $\Delta\,\alpha\;(\hbar\omega)$, exhibiting some structures due to the mass reversal of the exciton, i.e. a change of the curvature, and to the anisotropy of its dispersion close to the L-point (see Sect. 8.8).

The absorption process under absorption of phonons is even weaker at low temperatures, i.e., at $k_\mathrm{B}\,T < \hbar\omega_\mathrm{Phonon}$. Nevertheless, it was possible to deduce the absorption spectra, e.g., in Si down to $\alpha = 10^{-17}\,\mathrm{cm}^{-1}$ at 90 K from the luminescence spectra, from Kirchhoff's law and the generalized Planck equation [95D1, 95W1].

Above the indirect exciton, there can also be direct excitons which show absorption structures of the type already discussed in the preceding section.

We give an example in Fig. 13.34 for Ge. Since the direct exciton can decay rapidly into lower states, it is strongly damped, preventing the observation of finestructure such as states with $n_\mathrm{B} > 1$.

The luminescence of free excitons in indirect-gap materials involves – as does the absorption – one or more phonons (Fig. 13.32a). Consequently the excitonic emission is red-shifted with respect to $E_\mathrm{ex}(\boldsymbol{k}_0)$. In Fig. 13.35 we show an example for Ge where one sees the contributions of the various phonons.

Due to the participation of a third particle, the luminescence yield of indirect-gap materials is much smaller than that of direct ones. This makes the application of bulk Si in light emitting devices impossible. On the other hand, the lifetime of excitons (or carriers) is rather long in indirect materials and ranges from less than 1 µs to many µs depending on the purity of the material, while the lifetime in direct-gap materials is in the (sub-) ns range as will be demonstrated later.

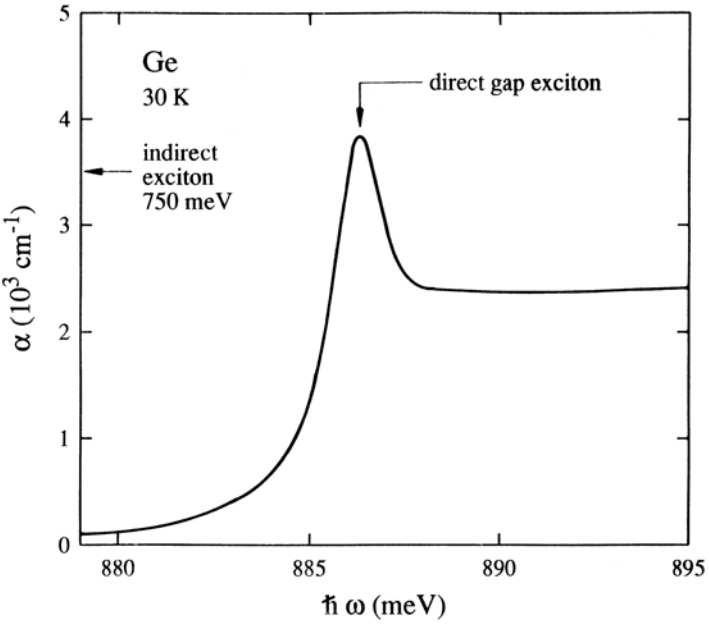


Fig. 13.34. Absorption around the direct exciton in Ge. According to [83S2]

13.3 Intraexcitonic Transitions

Excitons are frequently compared to hydrogen or positronium atoms, essentially with modifications of the masses of the two charge carriers and of the dielectric constant of the embedding medium. We have also used this approach, e.g., in Sect. 9.1 and 9.2.

Concerning the spectroscopy of excitons described, e.g., in this chapter so far and of H atoms, there is a crucial difference. If we do, e.g., absorption spectroscopy on H atoms, the H atoms are there already before we send a light beam on them, and with the light beam we measure, e.g., the transitions from the 1s state to the n_{B}p states ($n_{\mathrm{B}} = 2, 3, \ldots$), i.e., the Lyman series.

In a semiconductor, in contrast, there is no exciton before the light is switched on, i.e., a "vacuum" state and the incident light beam creates the excitons, e.g., in the n_{B}s states ($n_{\mathrm{B}} = 1, 2, 3, \ldots$).

A further difference resulting from the a priori presence of H atoms is the fact that the 1s $\rightarrow$ 2p transitions are usually Doppler broadened, while the excitons can be resonantly created on a well-defined point on their dispersion relation.

The question is now if it is possible to observe the Lyman series also for excitons? The answer is yes! What one needs is a certain density of 1s excitons in the semiconductor created, e.g., by optical pumping from the vacuum into some exciton states including their continuum.

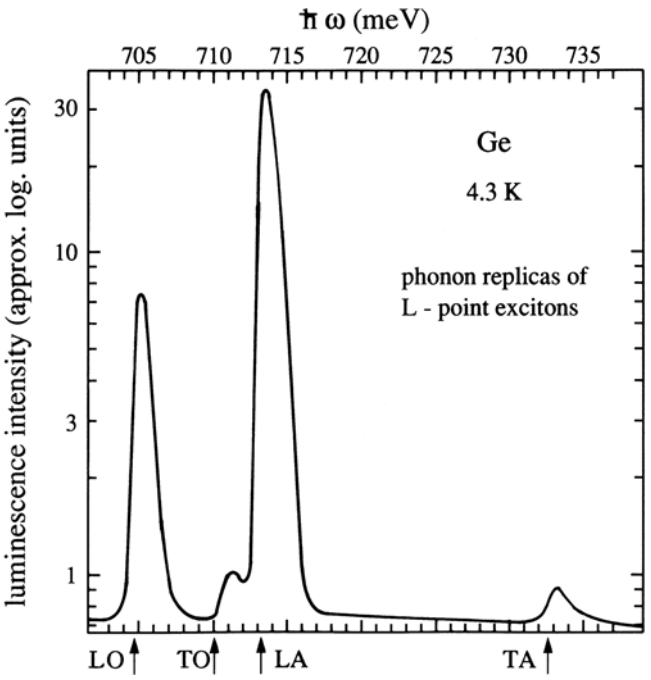


Fig. 13.35. A luminescence spectrum of Ge showing various phonon satellites. According to [76T1]

With a probe beam in the suitable spectral range one tries to detect the changes in the transmission induced by the created excitons. This so-called pump-and-probe or differential transmission spectroscopy is, after two-photon absorption and hyper-Raman scattering, in Sect. 13.1.4, a further example of nonlinear optical spectroscopy. We will see more examples for these techniques in Chaps. 19 to 24.

Early examples of such transitions between excitonic sublevels, also known as intra excitonic transitions, have been given for the indirect gap materials Si and Ge [76T2, 78T1, 88L1].

We give here an example for Si in Fig. 13.36. The sample is cw pumped by a laser and the induced absorption is plotted. The three main features ranging from 10 to 12 meV are attributed to the transitions from the 1s exciton state to various 2p states split by the anisotropy and degeneracy of the band structure of Si at the Δ-point, including the interaction of the carrier spins and the angular momentum of the envelope function.

Though there is a clear signature of these transitions, there are not too many reports of bulk materials for the following reason: The transition energies fall in the range of a few to over 100 meV depending on the exciton binding energy necessitating Fourier spectroscopy or other means of THz spectroscopy. The transition energy should not coincide with some phononic absorption fea-

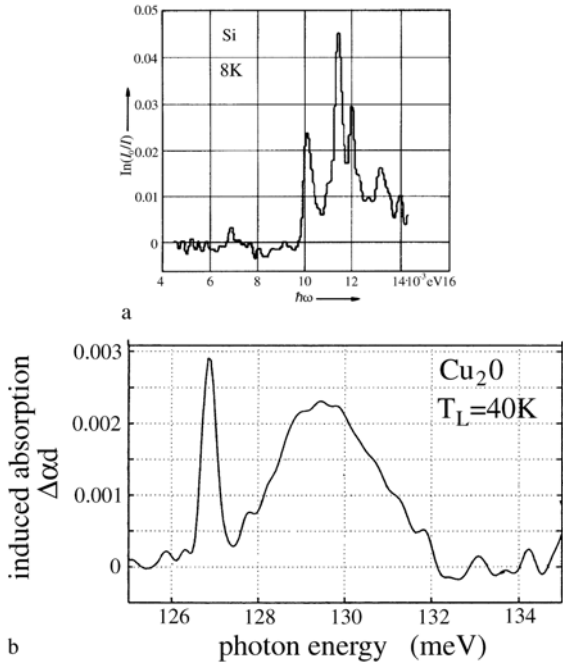


Fig. 13.36. The 1s → 2p transitions in an optically-pumped Si crystal at 8 K. (a) and in Cu_2O (b). According to [78T1, 04J1].

tures and, finally, the lifetime of the excitons should be long, so that it is possible to accumulate a sufficiently high density n of excitons under cw pumping, which should exceed e.g. the density of bound exciton sites (see Sect. 14.1).

Apart from the indirect gap semiconductors like Si or Ge mentioned above, Cu_2O fulfils the requirements because it has a direct but dipole forbidden band-to-band transition, as detailed in Sect. 13.2, which results in rather long lifetimes. Indeed signatures of 1s → 2p transitions have been reported. See [03J1] and references therein. However, contributions from defect states, from biphonons or from phonon overtones can not be ruled out in these data [04J1, 04K2]. The transition from the 1s para exciton to 2p states has been identified [04J1, 04K1, 04K3].

We show an example in Fig. 13.36b. Thought the total feature looks temptingly similar to the calculated induced absorption spectrum of a Bose condened gas of para excitons [01J1] (see Sect. 20.5) a detailed analysis shows, that the peak at 129 eV is due to a 1s para → 2s para transition in a classical Boltzmann exciton gas while the narrow line at 127 meV is probably due to a similar transition in a bound exciton complex (see Sect. 14.1)

The fast development of time resolved spectroscopy with THz probe pulses in recent years (see Chap. 23 and 25) allows one to observe the 1s → 2p transition in GaAs under resonant exciton creation and to follow the rather

slow relaxation of excitons from the continuum states into their 1s ground state in this material [03K1].

13.4 Problems

1. Make a plot of the longitudinal-transverse splitting of $n_B = 1$ exciton resonances as a function of the exciton binding energy. Include only semiconductors with direct, dipole-allowed band-to-band transitions. Compare with similar Figs. in Chap. 9.
2. Consider a band-to-band transition in a direct-gap semiconductor neglecting the Coulomb interaction between electron and hole and calculate the absorption spectrum for a dipole-allowed and a dipole-forbidden transition, i.e. for a transition with a matrix element varying linearly with $\boldsymbol{k}$.
3. Consider the $n_B = 1\,A\Gamma_5$-polariton resonances in CdS (Fig. 13.4b) and determine for a light beam incident at 45° to the surface ($\boldsymbol{E} \perp \boldsymbol{c}$, $\boldsymbol{k} \perp \boldsymbol{c}$) the length and direction of the wave vectors of the propagating modes in the sample and their phase and group velocities. Select a few characteristic photon energies. Explain the term "spatial dispersion".
4. Explain the differences between the concepts of polaritons, of spatial dispersion and of birefringence.
5. Can the transition energy of a dipole-allowed intraexcitonic transition 1s $\rightarrow$ 2p coincide with the energy of a dipole-allowed TO phonon?

References to Chap. 13

[43M1] E. Mollwo, Reichsber. der Physik **1**, 1 (1943) see also G. Heiland, E. Mollwo and F. Stöckmann, Solid State Physics **8**, 191 (1959)
[53U1] F. Urbach, Phys. Rev. **92**, 1324 (1953)
[57E1] R.J. Elliot, Phys. Rev. **108**, 1384 (1957) and in Ref [63P1] p 269
[57M1] W. Martienssen, J. Phys. Chem. Solids **2**, 257 (1957)
[58D1] D. Dutton, Phys. Rev. **112**, 785 (1958)
[58H1] J.J. Hopfield, Phys. Rev. **112**, 1555 (1958)
[59M1] W. Martienssen, J. Phys. Chem. Solids **8**, 294 (1959)
[60H1] J.J. Hopfield, J. Phys. Chem. Sol. **15**, 97 (1960)
[62H1] J.J. Hopfield and D.G. Thomas, Phys. Rev. **128**, 2135 (1962)
[62P1] S.J. Pekar, Sov. Phys. Sol. State **4**, 953 (1962)
[63H1] J.J. Hopfield and D.G. Thomas, Phys. Rev. **132**, 563 (1963)
[63P1] Polares and Excitons, C.G. Kupper and G.D. Whitfield eds., Oliver and Boyd, Edinburg (1963)
[64K1] M.N. Kabler, Phys. Rev. A **136**, 1296 (1964)
[64M1] G.D. Mahan and J.J. Hopfield, Phys. Rev. A **135**, 428 (1964)
[65C1] M. Cardona and G. Harbeke, Phys. Rev. A **137**, 1467 (1965)
[65C2] M. Cardona, M. Weinstein and G.A. Wolff, Phys. Rev. A **140**, 633 (1965)
[66R1] J. Ringeissen, Physics Lett. **22**, 571 (1996)

[67K1] M.N. Kabler and D.S. Patterson, Phys. Rev. Lett. **19**, 652 (1967)
[67O1] Y. Onodera and Y. Toyozawa, J. Phys. Soc. Japan. **22**, 833 (1967)
[68T1] Y. Toyozawa and J. Hermanson, Phys. Rev. Lett. **21**, 1637 (1968)
[68V1] J. Voigt and F. Spiegelberg, phys. stat. sol. (b) **30**, 659 (1968)
[69K1] E.O. Kane, Phys. Rev. **180**, 852 (1969)
[70M1] O. Madelung, Grundlagen der Halbleiterphysik, Heidelberger Taschenbücher, **71**, Springer, Berlin, Heidelberg (1970)
[70W1] J.P. Walter et al., Phys. Rev. B **1**, 2661 (1970)
[71F1] D. Fröhlich, E. Mohler and P. Wiesner, Phys. Rev. Lett. **26**, 554 (1971)
[71K1] M.V. Kurik, phys. stat. sol. (a) **8**, 9 (1971)
[71S1] H. Sumi and Y. Toyozawa, J. Phys. Soc. Jpn. **31**, 342 (1971)
[72B1] A. Bivas et al., Optics Commun. **6**, 142 (1972)
[72B2] H.B. Bebb and E.W. Williams, Semicond. and Semimetals **8**, 182 (1972)
[72D1] J.D. Dow and D. Redfield, Phys. Rev. **85**, 94 (1972)
[72K1] L. Kalok and J. Treusch, Phys. Stat. Sol. A **52**, K125 (1972)
[72W1] R.L. Weiher and W.C. Tait, Phys. Rev. B **5**, 623 (1972)
[72W2] E.W. Williams and H.B. Bebb, Semicond. and Semimetals **8**, 321 (1972)
[75F1] A. Frova et al., Phys. Rev. Lett. **34**, 1572 (1975)
[75K1] C. Klingshirn, phys. stat. sol. (b) **71**, 547 (1975)
[75K2] V.A. Kiselev, B.S. Razbirin and I.N. Uraltsev, phys. stat. sol. (b) **72**, 161 (1975) and references therein
[75L1] J. Lagois and K. Hümmer, phys. stat. sol. (b) **72**, 393 (1975)
[76S1] F. Spiegelberg, E. Gutsche, and J. Voigt, phys. stat. sol. (b) 77, 233 (1976)
[76S2] E. Strauss, V. Gerhardt and H. Riederer, J. Luminesc. **12/13**, 239 (1976)
[76T1] G.A. Thomas and M. Capizzi, in Proc. 13th Int. Conf. Phys. Semicond., ed. by F.G. Fumi, Rome, p. 915 (1976)
[76T2] T. Timusk et al., Phys. Rev. B **13**, 3511 (1976)
[77C1] A. Coret and A. Fort, Il Nuovo Cimento **39B**, 544 (1977)
[77V1] H. Venghaus, S. Suga and K. Cho, Phys. Rev. B **16**, 4419 (1977)
[77W1] G. Winterling, E.S. Koteles and M. Cardona, Phys. Rev. Lett. **39**, 1286 (1977)
[78B1] P.D. Bloch and C. Schwab, Phys. Rev. Lett. **41**, 514 (1978)
[78G1] G.G. Grosso et al., Solid State Commun. **25**, 435 (1978)
[78H1] K. Hümmer, Habilitation Thesis, Erlangen (1978)
[78L1] J. Lagois and B. Fischer, Festkörperprobleme / Adv. Solid State Phys. **18**, 197 (1978)
[78S1] W. Stößel and H.J. Wagner, phys. stat. sol. (b) **89**, 403 (1978)
[78T1] T. Timusk et al., Solid State Commun. **25**, 217 (1978)
[78U1] R.G. Ulbrich and C. Weisbuch, Festkörperprobleme / Adv. Solid State Phys. **28**, 217 (1978)
[79E1] W. Ekardt, K. Lösch, and D. Bimberg, Phys. Rev. B **20**, 3303 (1979)
[79M1] Y. Masumoto et al., J. Phys. Soc. Jpn. **47**, 1844 (1979)
[79M2] A. Mysyrowicz, D. Hulin and A. Antonetti, Phys. Rev. Lett. **43**, 1123 (1979)
[79P1] J. Puls and J. Voigt, phys. stat. sol. (b) **94**, 199 (1979)
[79S1] B. Sermage and G. Fishman, Phys. Rev. Lett. **43**, 1043 (1979)
[79S2] E. Swenberg and N. Geacintov, Organic Molecular Photophysics **1**, 489 (1979)
[79U1] R.G. Ulbrich and G.W. Fehrenbach, Phys. Rev. Lett. **43**, 963 (1979)

[79V1] J. Voigt, F. Spiegelberg and M. Senoner, phys. stat. sol. (b) **91**, 189 (1979)
[80B1] I. Broser and M. Rosenzweig, Phys. Rev. B **22**, 2000 (1980)
[80B2] K. Bohnert, G. Schmieder and C. Klingshirn, phys. stat. sol. (b) **98**, 175 (1980)
[80B3] I. Broser and M. Rosenzweig, Solid State Commun. **36**, 1027 (1980)
[80D1] W. Dreybrodt et al., Phys. Rev. B **21**, 4692 (1980) and references therein
[80H1] W. Hanke and L.J. Sham, Phys. Rev. B **21**, 4656 (1980)
[80K1] E.S. Koteles and G. Winterling, Phys. Rev. Lett. **44**, 948 (1980)
[80M1] I.V. Makarenko, I.N. Uraltsev and V.A. Kiselev, phys. stat. sol. (b) **98**, 773 (1980)
[81B1] I. Broser et al., Solid State Commun. **39**, 1209 (1981)
[81C1] K. Cho and Y. Yamane, Solid State Commun. **40**, 121 (1981)
[81I1] T. Itho et al., Solid State Commun. **37**, 925 (1981)
[81K1] R. Kuhnert, R. Helbig and K. Hümmer, phys. stat. sol. (b) **107**, 83 (1981)
[81K2] C. Klingshirn and H. Haug, Phys. Rep. **70**, 315 (1981)
[81L1] J. Lagois, Phys. Rev. B **23**, 5511 (1981)
[81U1] Ch. Uihlein, D. Fröhlich and R. Kenklies, Phys. Rev. B **23**, 2731 (1981)
[82B1] G. Blattner et al., Phys. Rev. B **25**, 7413 (1982)
[82O1] Y. Onodera and T. Oshikiri, J. Phys. Soc. Japan **49**, 1845 (1980) and ibid. **51**, 2194 (1982)
[82P1] S. Permogorov, in Excitons, ed. by E.I. Rasha, M.D. Sturge, Mod. Probl. Cond. Mat. Sci., Vol. **2**, p. 177, North Holland, Amsterdam (1982)
[82R1] M. Rosenzweig, PhD Thesis TU Berlin (1982)
[82S1] T. Skettrup, phys. stat. sol. (b) **109**, 663 (1982)
[83A1] D.E. Aspnes and A.A. Studna, Phys. Rev. B **27**, 985 (1983)
[83G1] R. Grasser and R. Scharmann, NATO ASI Series B **88**, 317 (1983)
[83M1] W. Maier, G. Schmieder and C. Klingshirn, Z. Phys. B **50**, 193 (1983)
[83M2] Y. Masumoto, S. Shionya and T. Takagahara, Phys. Rev. Lett. **51**, 923 (1983)
[83M3] N. Meskini, H.J. Mattausch and W. Hanke, Solid State Commun. **48**, 807 (1983)
[83S1] Y. Segawa, Y. Aoyagi and S. Namba, J. Phys. Soc. Jpn. **52**, 3664 (1983)
[83S2] H. Schweizer et al., Phys. Rev. Lett. **51**, 698 (1983)
[84L1] M.V. Lebedev et al., JETP Lett. **39**, 366 (1984)
[84M1] M. Matsushita, J. Wicksted and H.Z. Cummins, Phys. Rev. B **29**, 3362 (1984)
[84M2] T. Mita and N. Nagasawa, Solid State Commun. **44**, 1003 (1984)
[84R1] R. Ruppin, Phys. Rev. B **29**, 2232 (1984)
[84S1] T. Shigenari, X.Z. Lu and H.Z. Cummins, Phys. Rev. B **30**, 1962 (1984)
[84W1] J. Wicksted et al., Phys. Rev. B **29**, 3350 (1984)
[85B1] F. Beerwerth and D. Fröhlich, Phys. Rev. Lett. **55**, 2603 (1985)
[85H1] B. Hönerlage et al., Phys. Rep. **124**, 161 (1985)
[85L1] J.G. Liebler, S. Schmitt-Rink and H. Haug, J. Lumin. **34**, 1 (1985)
[86B1] F. Beerwerth and D. Fröhlich, Phys. Rev. Lett. **57,** 1344 (1986)
[88L1] D. Labrie et al., Phys. Rev. Lett. **61**, 1882 (1988)
[91F1] D. Fröhlich et al., Phys. Rev. Lett. **67**, 2343 (1991)
[91O1] T. Ogawa and T. Takagahara, Phys. Rev. B **43**, 14325 (1991)
[91U1] R. Ulbrich, in Materials Science and Technology, ed. by W. Schröter, VCH, Weinheim, Vol. **4**, p. 65 (1991)

[93F1] M. Fiebig, D. Fröhlich and Ch. Pahlke-Lerch, phys. stat. sol. (b) **177**, 187 (1993)
[93K1] C. Klingshirn, NATO ASI Ser. B **301**, 119 (1993)
[93M1] Y. Ma et al., Phys. Rev. Lett. **71**, 3725 (1993)
[94F1] D. Fröhlich in [81A1] [81A1] g of Chater 1, p. 289 and references therein
[94J1] C. Janowitz et al., Phys. Rev. B **50**, 2181 (1994)
[94W1] M. Watanabe and T. Hayashi, J. Phys. Soc. Japan **63**, 785 (1994)
[95B1] F. Bassani, G. Czajkowski and A. Terdicucci, Z. Physik B **98**, 39 (1995)
[95D1] E. Daub and P. Würfel, Phys. Rev. Lett. **74**, 1020 (1995)
[95S1] H. Schneider and K. Köhler, Phys. Rev. B **52**, R14364 (1995)
[95W1] P. Würfel, S. Finkbeiner and E. Daub, Appl. Phys. A **60**, 67 (1995)
[96N1] U. Neukirch et al., phys. stat. sol. (b) **196**, 473 (1996)
[96T1] S. Tanaka and Y. Kayanuma, Solid State Commun. **100**, 77 (1996)
[97N1] S. Nüsse et al., Phys. Rev. B **55**, 4620 (1997)
[97N2] U. Neukirch et al., Phys. Rev. B **55**, 15408 (1997) and ibid. **57**, 9208 (1998)
[98H1] K. Henneberger, Phys. Rev. Lett. **80**, 2889 (1998)
[98J1] A. Jolk and C. Klingshirn, phys. stat. sol. (b) **206**, 841 (1998)
[98K1] M. Kira, F. Jahnke and S.W. Koch, Phys. Rev. Lett. **81**, 3263 (1998)
[99D1] J.S. Dodge et al., Phys. Rev. Lett. **83**, 4650 (1999)
[99J1] A. Jolk et al., Proc. 24th Intern. Conf. Phys. Semiconductors, Jerusalem (1998), CD, File II C4 (1999)
[99R1] A.V. Rodina et al., Phys. Stat. Sol. B **216**, 21 (1999)
[00F1] V.V. Frolov et al., Chem. Phys. Lett. **326**, 558 (2000)
[00H1] L. Hanke et al., phys. stat. sol. (b) **221**, 287 (2000)
[00T1] J. Tignon et al., Phys. Rev. Lett. **84**, 3382 (2000)
[01F1] V.V. Frolov et al., Phys. Rev. B **63**, 205203 (2001)
[01J1] M. Jörger et al., Phys. Rev. B **64**, 113204 (2001)
[01J2] K. Johnsen and G.M. Kavoulakis, Phys. Rev. Lett. **86**, 858 (2001)
[01R1] A.V. Rodina et al., Phys. Rev. B **64**, 115204 (2001)
[01S1] H.C. Schneider et al., Phys. Rev. B **63**, 045202 (2001)
[02L1] Low-Dimensional Nitride Semiconductors, B. Gil, ed., Oxford University Press, Oxford (2002)
[02P1] Proc. Intern. Workshop on Nitride Semiconductors, A. Hoffmann and A. Rizzi, eds., phys. stat. sol. C **0**, Issue 1 (2002)
[03C1] S.F. Chichibu et al., J. Appl. Phys. **93**, 756 (2003)
[03D1] G. Dasbach et al., phys. stat. sol. (b) **238**, 541 (2003)
[03J1] M. Jörger et al., phys. stat. sol. (b) **238**, 470 (2003)
[03K1] R.A. Kaindl et al., phys. stat. sol. (b) **238**, 451 (2003)
[04D1] G. Dasbach et al., Phys. Rev. B (2004) in press
[04J1] M. Jörger, PhD Thesis, Karlsruhe (2004) and Proc. EXCON 04, Cracow, J. Luminesc. in press
[04K1] M. Kuwata-Gonokami et al., J. Phys. Soc. Japan **73**, 1065 (2004) and Proc. ICSCES, Pittsburgh (2004) Solid State Commun. in press
[04K2] K. Karpinska, P.H.M. van Loosdrecht and I.P. Handayani, Proc. ICSCES, Pittsburgh (2004), Solid State Commun. in press and EXCON 04, Cracow, J. Luminesc. in press
[04K3] C. Klingshirn et al., Proc. ICSCES, Pittsburgh (2004) Solid State Commun. in press
[04P1] Physics of Organic Semicond., W. Brütting ed., phys. stat. sol. (a) **201** (6), (2004)

14

Optical Properties of Bound and Localized Excitons and of Defect States

In the previous chapter we discussed mainly the optical properties of intrinsic excitons. Here we consider the optical properties of defect and localized states in bulk materials, but mention that many of these aspects are also relevant for the structures of reduced dimensionality presented in the next chapter.

14.1 Bound-Exciton and Multi-exciton Complexes

We have already introduced point defects in Sect. 8.14. Some of these defects can bind an exciton resulting in a bound exciton complex (BEC), see also Sect. 9.5. In Fig. 14.1 we visualize excitons bound to an ionized donor (D^+X), a neutral donor (D^0X), and a neutral acceptor (A^0X). Excitons are usually not bound to ionized acceptors, as explained in Sect. 9.5 The binding energy E^b of the exciton to the complex usually increases according to

$$E^b_{D^+X} < E^b_{D^0X} < E^b_{A^0X}\,. \tag{14.1}$$

The binding energy is defined as the energetic distance from the lowest free exciton state at $\boldsymbol{k} = 0$ to the energy of the complex.

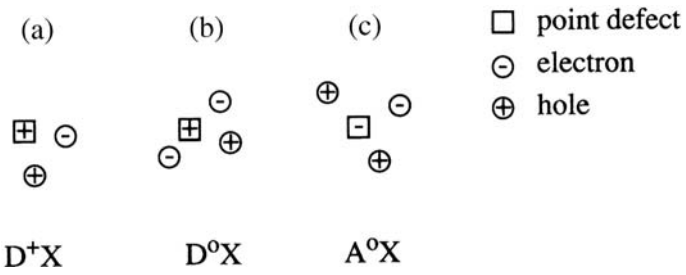


Fig. 14.1. Visualization of an exciton bound to an ionized donor (**a**), a neutral donor (**b**), and a neutral acceptor (**c**)

There is a rule of thumb, known as Hayne's rule, which relates the binding energy of the exciton to the neutral complex with the binding of the additional carrier to the point defect. For the D^0X complex, for example, this says that the ratio of $E^{\mathrm{b}}_{\mathrm{D^0X}}$ to the binding energy of the electron to the donor $E^{\mathrm{b}}_{\mathrm{D}}$ is a constant, depending only on material parameters such as the effective masses:

$$E^{\mathrm{b}}_{\mathrm{D^0X}}/E^{\mathrm{b}}_{\mathrm{D}} = \mathrm{const}\,. \tag{14.2}$$

The constant is often found to have a value around 0.1 to 0.2 (Hayne's rule, see [60H1, 77H1, 04M1]). Isoelectronic traps (such as a Te on a Se site in ZnSe) sometimes form deep centers, with binding energies for excitons exceeding those of neutral acceptors. A long-standing and controversially discussed problem is that of the EL2 luminescence in GaAs. For details including related defects see, e.g., [93S1, 97S1, 98S1] and references therein.

The BEC do not have any degree of freedom for translational motion. As a consequence BEC often show up in luminescence and absorption spectra as extremely narrow peaks. In Fig. 14.2 we give as examples luminescence spectra of ZnO, ZnSe and GaAs. See also Fig. 13.9d.

Figure 14.2a gives an overview of the low-temperature luminescence of a high-quality ZnO single crystal from [76H1, 76T1]. The interpretation of the various groups of emission bands is as follows. FE is the recombination of free exciton polaritons (see also Chap. 13). Then comes a group of bound exciton complexes, the assignment of which is presently reconsidered. According to the above references the complexes were attributed with decreasing energy, i.e., with increasing binding energy to the defect, to D^+X, D^0X and A^0X, respectively, according to Fig. 14.1. The neutral acceptors were assumed to be, e.g., Na or Li on Zn sites. Later on, a large fraction of the lines D^0X and A^0X has been attributed to excitons bound to neutral acceptors [97G1]. Presently, all of the lines labeled D^0X in the lower line in Fig. 14.2 are assigned to excitons bound to neutral donors based on the analysis of the two electron transitions, while hardly any shallow acceptors have been identified [04M1]. For the nitrogen acceptor in ZnO:N see [02Z1, 04W1].

Independently of the detailed interpretation of the various lines, we see within one group a splitting of the lines caused by the chemical nature of the binding atom. This effect is known as chemical shift or central cell correction. Similar shifts or splittings can be caused by the presence of other defects in the vicinity of the center binding the exciton.

Towards lower energies in Fig. 14.2a we see the luminescence of excitons bound to deep centers, possibly deep acceptors. Then follows a range of two electron transitions (see below) and the LO-phonon replica of free and bound excitons, denoted as X-LO and BEC-LO, respectively.

In ZnSe or GaAs we find a similar behavior, as can be seen in Fig. 14.2b,c.

To conclude this part it can be stated that the low temperature luminescence of high quality samples is generally dominated by defect-(or localiation-) related recombination processes.

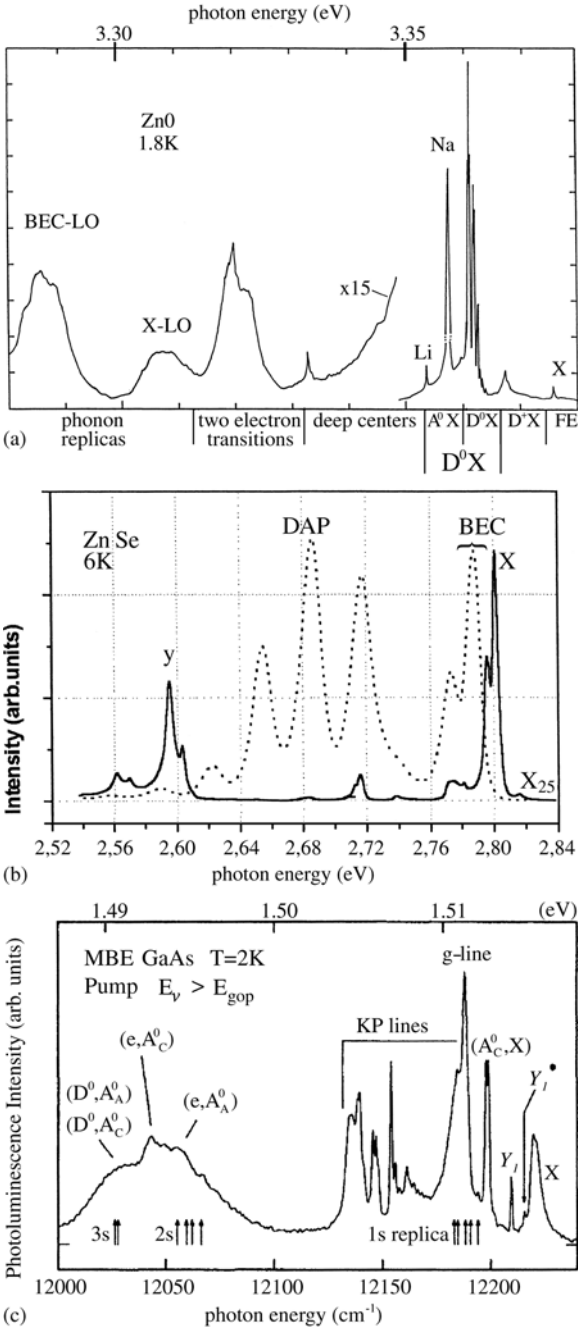


Fig. 14.2. Low-temperature and low-excitation luminescence spectra of ZnO (**a**) of two differently grown ZnSe samples (*solid line*: MBE growth with element sources, *dashed line*: with a compound source) (**b**) and of GaAs (**c**) [76H1,76T1,91V1,03D1]

Since $\boldsymbol{k}$-conservation is relaxed for BEC due to the lack of translational invariance, BEC can also couple to acoustic phonons, for example via the deformation potential or piezoacoustic coupling. This coupling leads to a so-called acoustic wing in the emission or absorption spectra of a BEC. Figure 14.4 shows an example for the I_1, line in CdS. Since only few phonon states are thermally populated at low temperature, the emission manifests itself essentially as a wing on the low energy side according to

$$E_{\mathrm{A}^0X} \longrightarrow E_{\mathrm{A}}^0 + \hbar\omega_{1\mathrm{um}} + \hbar\Omega_{\mathrm{A}}\,, \tag{14.3a}$$

where $\hbar\Omega_{\mathrm{A}}$ is the energy of the emitted acoustic phonon. The shape of the phonon wing is determined by the energy dependence of the coupling of the BEC to the phonons and by their population, especially the ratio of Stokes to anti-Stokes emission. In Fig. 14.3 the influence of the lattice temperature on the emission line shape is illustrated.

Figure 14.4 shows the acoustic phonon side bands of the I_1 and I_2 BEC in CdS and in addition an absorption spectrum of the same sample. In this special case actually the reabsorption of a rather broad luminescence band appearing under higher excitation was used. The acoustic wing appears in absorption on the high energy side of the zero-phonon line according to

$$E_{\mathrm{A}^0} + \hbar\omega_{\mathrm{abs}} \longrightarrow E_{\mathrm{A}^0\mathrm{X}} + \hbar\Omega_{\mathrm{A}}\,. \tag{14.3b}$$

Another recombination process leaves the point defect in an excited state, this means that the donor electron of a $\mathrm{D}^0\mathrm{X}$ complex is transferred in the

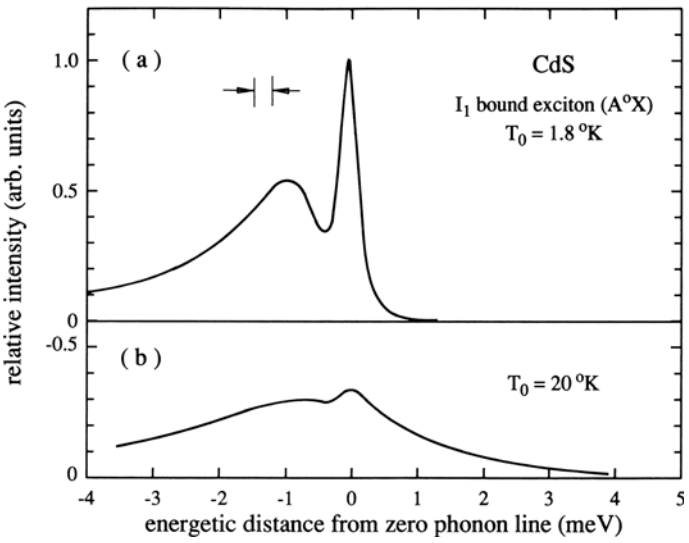


Fig. 14.3. The Stokes and anti-Stokes emission in the vicinity of the I_1 BEC in CdS for two lattice temperatures [74D1, 74S1]

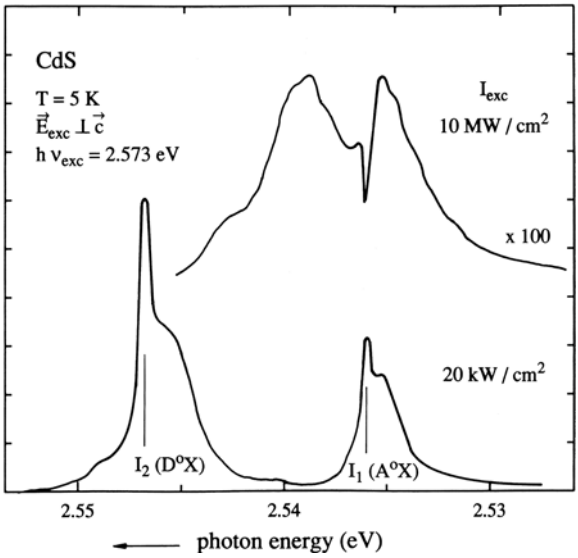


Fig. 14.4. The emission (lower trace) and (re)absorption spectrum of a CdS sample around the I_1 BEC, showing the acoustic wing [79S1]

recombination process from its ground state into an excited state. The energy conservation for these "two-electron transitions" reads in the simplest approximation for the D^0X complex, and assuming that the binding energy of the donor forms a hydrogen-like series of states described by the main quantum number n_B,

$$\hbar\omega_{1um} = E_{D^0X} - E_{D^0}\left(1 - n_B^{-2}\right) \tag{14.4}$$

with $E_{D^0} = \mathrm{Ry}\, m_e(m_0\varepsilon^2)^{-1}$.

In Fig. 14.5 we show an example for ZnO. See also [04M1]

A similar "two-hole transition" may occur for acceptor-bound excitons (A^0X). Corresponding transitions to various 1s, 2s and 3s states are indicated in Fig. 14.2c.

The BEC have a very rich and complex spectrum of excited states. Apart from the quantum numbers known from free excitons n_B, l, m, (see Sects. 9.1,2), excited states can be created by the relative orientations of the spins of identical carriers (with corresponding parities of the envelope function), or in the A^0X complex, for example, by the participation of holes from deeper valence bands. As an example we give in Fig. 14.6 the photoluminescence excitation spectrum of a A^0X or D^0X complex in ZnO. See the discussion above with Fig. 14.2a. The peaks labelled R_2 and R_5 are thought to be due to complexes which involve one or two holes from the B valence band instead of the A valence band. The former is separated from the uppermost A valence band by about 5.4 meV. The resonances R_1, R_3 and R_6 could then be due to the excitation of internal degrees of freedom of the complexes with various

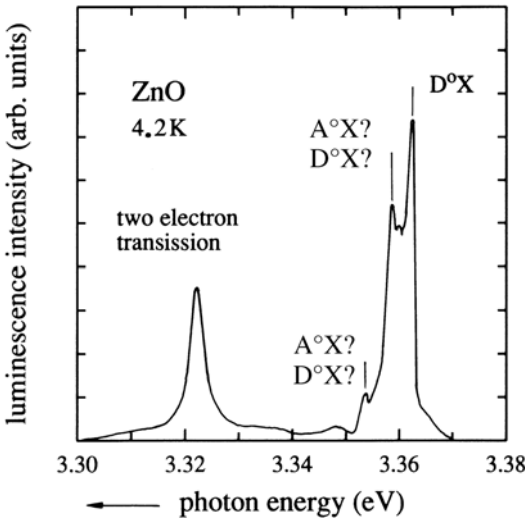


Fig. 14.5. An emission spectrum of a ZnO sample with an especially pronounced two-electron transition. According to [75K1]. Concerning the identification of the lower BEC see the discussion above

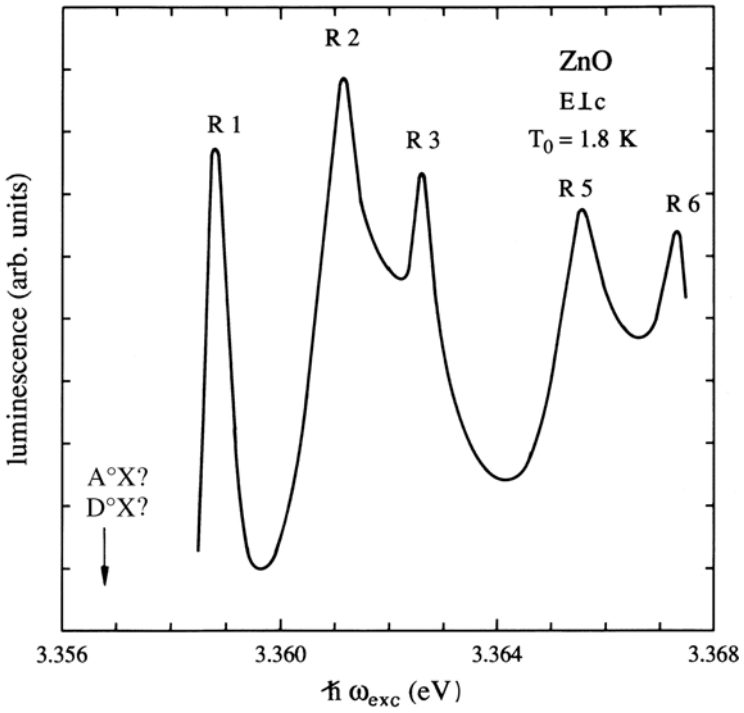


Fig. 14.6. Photoluminescence excitation spectrum of an BEC complex in ZnO [81B1]

hole configurations. Note that the interpretation for R_5 does not work for a D^0X complex, which continues only one hole.

A more elaborate treatment of this topic is described in [81B1,86G1,97G1, 04M1].

Many defects act as fast nonradiative recombination centers or luminescence killers like dislocations [99L1, 00F1]. In some cases they give rise, however, to characteristic luminescence features. The so-called Y-emission line in ZnSe is, e.g., due to the recombination of an exciton bound to the lattice disortion field of a dislocation. An example is given in Fig. 14.2b.

The use of A^0X luminescence in a ZnSe quantum well as a single photon source has been described in [02S1]

Although we first treat quantum structures in the next section, we want to mention here that BEC also exists in quantum wells. Partly their luminescence merges with that of the so-called free, i.e., generally disorder broadened intrinsic emission lines as is often the case in $GaAs/Al_{1-y}Ga_yAs$ structures, partly it appears as a spectrally separated line. An example is shown in Fig. 14.7 for structures with different quantum well widths l_z, centre-doped with Be acceptors. For further references see, e.g., [01L1] of Chap. 1. It should be especially noted that for quantum structures, the binding energy of the excitons to the center additionally depends on its position relative to the barriers.

The bandstructure of some indirect semiconductors allows multiple occupation of the conduction band due to the many-valley structure, and of the valence band due to the four-fold degeneracy of the Γ_8 valence-band maximum. In these materials multiple bound-exciton complexes can be formed, i.e., BEC which contain one, two, three or even more electron–hole pairs. The emission lines are due to transitions of a complex containing m electron–hole pairs to one with $m-1$ pairs. In Fig. 14.8 we give as an example a luminescence spectrum of Si:P. The usual ($m = 1$) emission line of the phosphorous D^0X complex dominates. It is followed at lower energies by a series of lines with indices m up to 6. The decay of the higher members of the series is obviously faster than of the lower ones, as expected already from simple reaction kinetics. These multiple BEC can be considered as a precursors or nucleation centers for the electron–hole plasma droplet discussed in Chap. 21. Another noteworthy point is that Si is an indirect semiconductor and optical transitions usually involve a momentum conserving phonon. Since the $\boldsymbol{k}$-selection rule is relaxed for BEC as mentioned above, we can see these lines even without phonon participation.

BECs can be best observed at low lattice temperatures. With increasing T_L the BEC disappear depending on the material, often at temperatures between 20 and 100 K due to thermal dissociation of the exciton from the complex. For an example see Fig. 14.4. Below this temperature the halfwidth of the emission lines increase partly linearly and partly quadratically with T_L.

For optimal observation conditions the concentration of the point defects should be low enough ($\lesssim 10^{16}\,\mathrm{cm}^{-3}$) to avoid broadening due to interaction between the BEC.

Due to this low concentration BEC are usually best observed in luminescence or photoluminescence excitation spectroscopy and only partly in absorption see e.g. Fig. 13.9d, and only very rarely do they give rise to reflection structures, simply because the modulation of n and/or κ is too small to result in a significant signal in $R(\omega)$. For an exception to this rule see [78B1].

To conclude this section it should be mentioned that BEC are sometimes said to have a "giant" oscillator strength. This expression is partly the result of a misconception because the maximum absorption coefficient $\alpha^{\max}$ for the free excitons per unit cell was compared with the corresponding quantity of the BEC per defect center. However, in both cases the (Wannier-) excitons cover many unit cells.

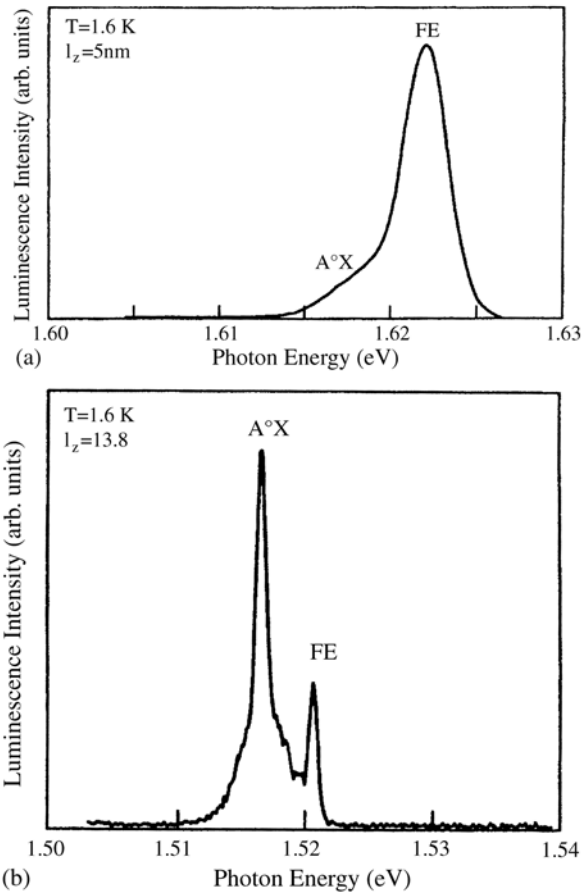


Fig. 14.7. The luminescence of GaAs/$Al_{1-y}Ga_yAs$ structures, of two different well width l_z center doped with Be acceptors at low temperature showing the emission from localized intrinsic excitons labelled FE and of acceptor bound excitons (A^0X) [89H1]

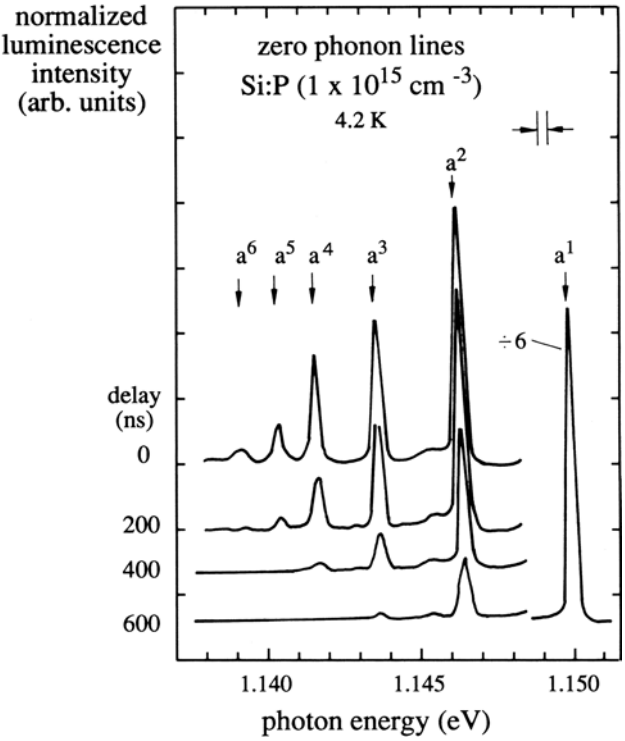


Fig. 14.8. The zero-phonon luminescence spectra of a multiple-bound-exciton complex in Si:P after pulsed excitation, normalized to the a_1 line [78T1]

For some reviews on bound exciton complexes in bulk materials and quantum wells see e.g. [79D1, 97G1, 04M1] or Ref. [82L1, 01L1] in Chap. 1 and the references given therein or in Sect. 9.5.

14.2 Donor–Acceptor Pairs and Related Transitions

Until now we have assumed that the BEC involves only one defect center. Actually one can also imagine poly-centric bound-exciton complexes, which involve two or more close-lying centers.

The simplest such defect combination is the donor–acceptor pair. In Fig. 14.9 we show schematically an electron on a donor and a hole on an acceptor. If their wavefunctions overlap, they can recombine. The energy of the photon resulting from radiative recombination is given by

$$\hbar\omega_{\mathrm{DA}} = E_{\mathrm{g}} - E^{\mathrm{b}}_{\mathrm{D}^0} - E^{\mathrm{b}}_{\mathrm{A}^0} + \frac{e^2}{4\pi\varepsilon\varepsilon_0 r_{\mathrm{DA}}} - m\hbar\omega_{\mathrm{LO}} \qquad (14.5)$$

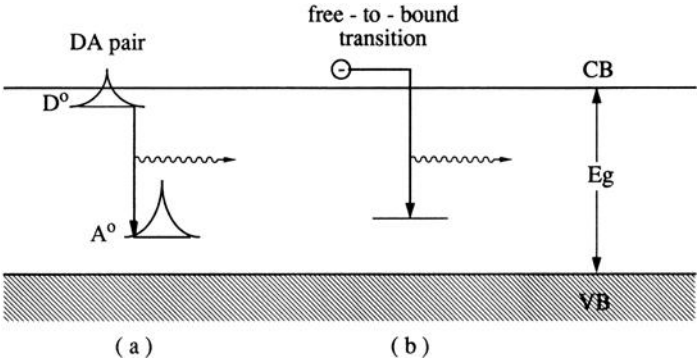


Fig. 14.9. A donor–acceptor pair (**a**) and a free-to-bound recombination process (**b**)

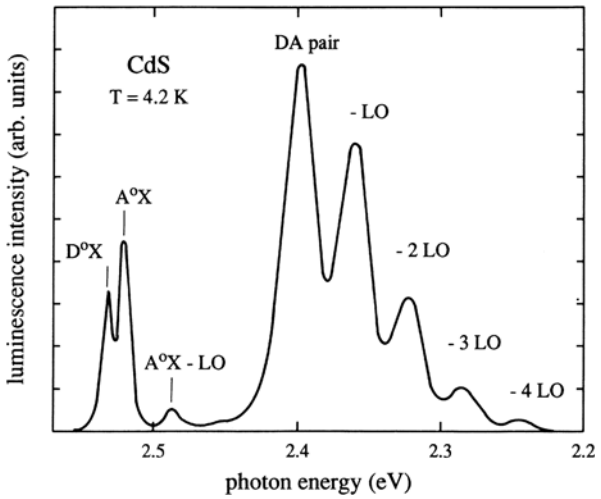


Fig. 14.10. Donor–acceptor pair luminescence in CdS including some phonon replica [79B1]

for singly charged centers if we neglect the energetic shift caused by the overlap of the wavefunctions of the neutral donor and acceptor. $E^{\mathrm{b}}_{\mathrm{D^0A^0}}$ are the binding energies of electron and hole to their respective centers.

The forth term on the rhs of (14.5), which depends on the distance r_{DA} of the centers, reflects the Coulomb energy of the ionized centers after the recombination. The last term describes LO phonon replica. We give examples in Figs. 14.2b and 14.10.

If the donors and acceptors are introduced substitutionally in a rather simple (e.g., cubic) lattice, then only discrete, well-defined values for r_{DA} are possible, to nearest neighbors and next-nearest neighbors, etc. Consequently under high resolution one sometimes observes that the zero-phonon band consists of discrete lines corresponding to the discrete values of r_{DA} [79C1,79D1].

With increasing pump intensity, the number of occupied donor and acceptor centers increases and their average distance r_{DA} necessarily decreases. As a consequence one finds that the emission maximum of the pair-band shifts to the blue with increasing excitation due to the Coulomb term in (14.5). This is a very characteristic feature of donor–acceptor pair recombination. For an example in the less commonly investigated semiconductor SnO_2 see [80B1].

Another recombination process connected with neutral donors or acceptors is the so-called free-to-bound transition. In this case a free electron or hole recombines with a neutral acceptor or donor, respectively, as shown schematically in Fig. 14.9. The corresponding emission peak is at

$$\hbar\omega_{\mathrm{FB}} = E_{\mathrm{g}} - E^{\mathrm{b}}_{\mathrm{D}^0/\mathrm{A}^0} - m\hbar\omega_{\mathrm{LO}}\,. \tag{14.6}$$

Obviously $\hbar\omega_{\mathrm{FB}}$ is blue-shifted as compared to $\hbar\omega_{\mathrm{DA}}$ and often both processes overlap together with their phonon replica in one complex or broadened luminescence spectrum. Exampels are shown in Fig. 14.2c both for donor–acceptor pairs $(\mathrm{D}^0, \mathrm{A}^0)$ (the index C means a carbon acceptor) and free electron to bound hole transitions (e, A^0).

14.3 Internal Transitions and Deep Centers

Some defects have not only one (or a few) levels close to one band, but they have several of them, partly around the middle of the gap. The chance of encountering such a situation evidently increases with increasing E_{g} and is thus of higher relevance for insulators, e.g, solid state laser materials, like the (Cr, Ti or Nd) doped ones, than for semiconductors.

Such deep centers can sometimes interact with both the conduction and valence band and serve then as recombination centers. If this recombination is fast and non-radiative, these centers are known as "luminescence killers". They are to a large extent responsible for the low luminescence yield of many semiconductors. Iron is an example of such a center.

Other deep centers show either also free-to-bound or internal transitions that reflect the atomic orbitals. These, however, are perturbed with respect to the free atom by the environment of neighboring atoms. Examples of this type of center are copper, sodium, the rare earths, and the transition metals. They give rise to the green, orange and red emission bands of wide-gap semiconductors such as CdS, ZnO and ZnS. We give an example in Fig. 14.11 but do not go into details, but mention, that there is also another green emission band in ZnO, which is spectrally smooth without the pronounced phonon structure and which is related to an oxygen vacancy [96V1] in addition to ZnO:Cu and citing for ZnO:OH [78G1]. Recently an emission band has been observed in ZnO in the orange spectral range, with the unusual property that the luminescence cloud expands after the switching on of cw band-to-band excitation over distances of several mm in some tenths of a second. The effect can be

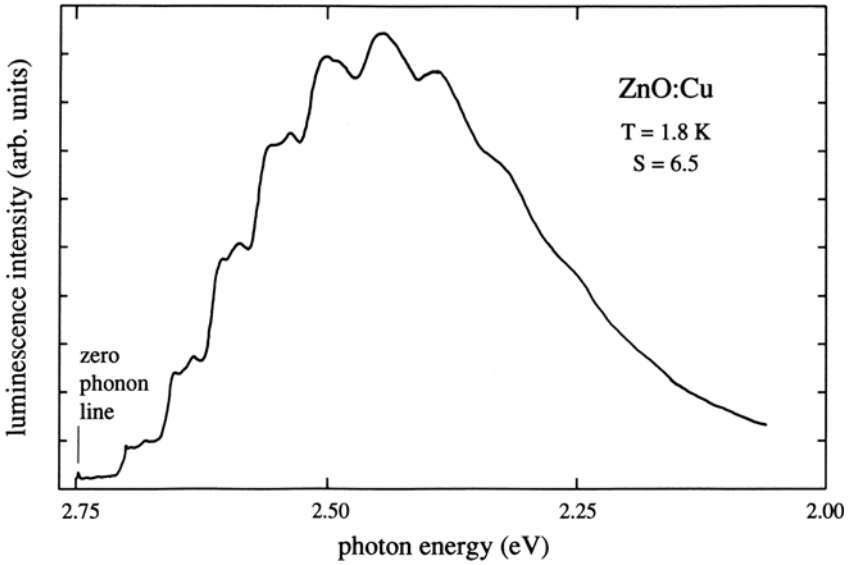


Fig. 14.11. Deep center emission of ZnO:Cu [81K1]

observed at temperatures from 70 K to 150 K. It has been attributed tentatively to a (possibly thermally activated) Forster or Foster–Dexter energy transfer [04D1].

As already mentioned above, the topic of deep centers is much more important for insulators. Ample information for this group of materials and for semiconductors can be found in [75O1, 81H1, 90K1] or in [81A1] b,c,e of Chap. 1.

14.4 Excitons in Disordered Systems

In Sect. 8.15 we outlined the appearance of localized tail states with increasing doping, for alloy semiconductors and in amorphous materials. In quantum wells the fluctuations of l_z may also result in the formation of localized tail states because of the dependence of the quantization energy on l_z as mentioned earlier.

If we create electron–hole pairs in such systems, they may be localized, too, in the energetic regime below the mobility "edge" which is usually a transition region of finite width.

Whole excitons can be localized in potential fluctuations of sufficient depth and with diameters larger than their Bohr radius. In other cases, only one of the carriers is localized–usually the hole, because of its heavier mass–and binds the other carrier to it by Coulomb attraction. For a more detailed discussion of localization, including a percolation approach, see Sects. 8.15 and 9.6, [03E1]

and the references therein. For localization in various quantum structures see [02R1] and for organic alloys see [76K1] and for amorphous semiconductors [03S2].

The typical shape of the absorption spectrum of a strongly disordered (e.g., amorphous) system is shown in Fig. 14.12.

In region I the absorption is weak and is caused by impurities in the alloy or amorphous semiconductor, i.e., atoms of a different chemical nature.

Then region II follows where the absorption coefficient increases exponentially with photon energy. This regime is also known as the "Urbach tail" in analogy to Fig. 13.10 and (13.12). Its origin in disordered semiconductors is, however, not the interaction with phonons, but simply the exponential tail of the density of states of localized excitons which can be reached without $\boldsymbol{k}$-selection. Above follows region III comprising the absorption via

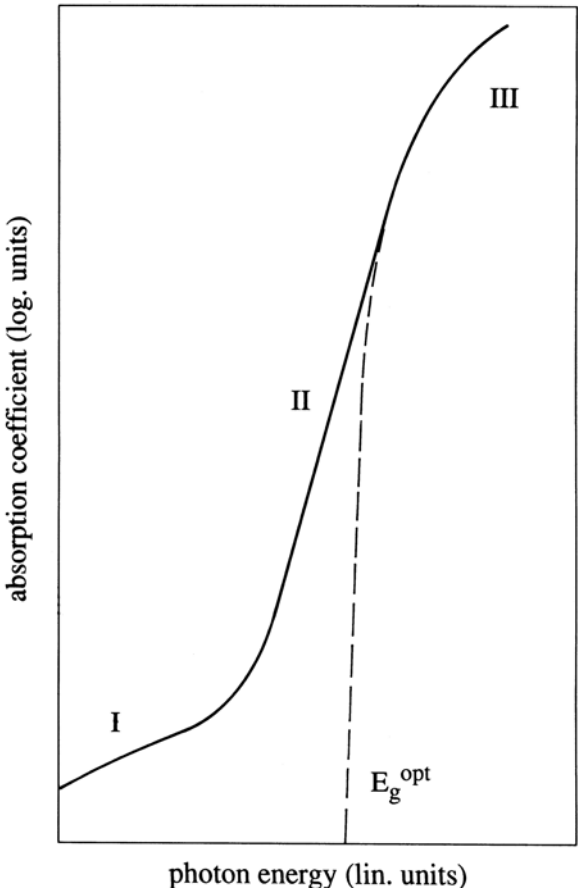


Fig. 14.12. The various regimes of the absorption coefficient in disordered semiconductors [72W1]

transitions into extended states. An extrapolation of this part to lower energies, e.g., with a $(E' - E_0)^2$ law, allows one to define an optical or mobility gap. For this so-called Tauc-regime see, e.g., [89S1, 89T1]. The latter quantity can also be found by excitation-spectroscopy of the photo-current. Absorption spectra corresponding to the schematic drawing of Fig. 14.12 have been found in amorphous Si, in chalcogenide glasses, and in alloy semiconductors such as $CdS_{1-x}Se_x$ and $ZnSe_{1-x}Te_x$. Examples are presented, e.g., in [72W1, 89T1, 99R1, 04K1].

The gradual increase of the absorption, and via Kramers–Kronig analysis also of the real part of the reflective index, in many cases prevents the observation of discrete exciton resonances in reflection.

The analysis of the "Urbach" part of the absorption spectrum gives the tailing parameter $\varepsilon_{\rm loc}$ describing the density of localized states below the mobility edge $E_{\rm ME}$.

$$N(E) = \frac{N_0}{\varepsilon_{\rm loc}} e^{-{\rm E}/\varepsilon_{\rm loc}} \quad E < E_{\rm ME} \tag{14.7}$$

In Fig. 14.13 we show $\varepsilon_{\rm loc}$ for $CdS_{1-x}Se_x$ as a function of x. This quantity necessarily vanishes for the ordered binary compounds $x = 0$ and $x = 1$, see also Sects. 8.15 and 9.6. In between it goes through a maximum with a shape that is not symmetric with respect to $x = 0.5$, partly since the quantity $\mathrm{d}E_{\rm g}/\mathrm{d}x$ changes with x, which enters in (8.64,67).

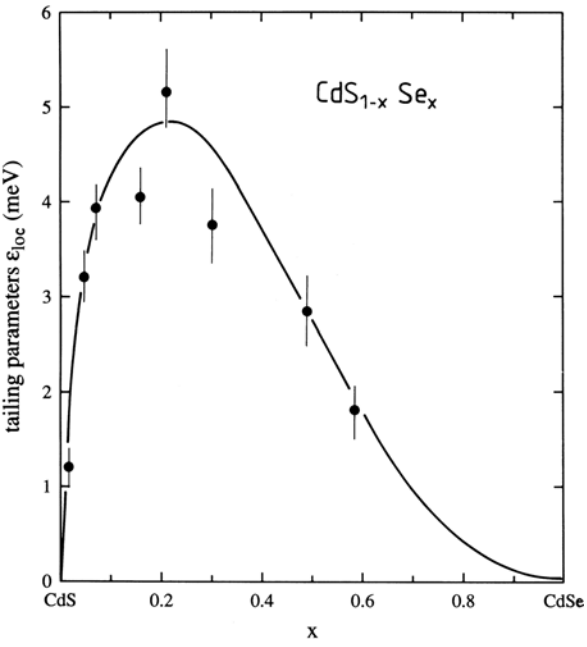


Fig. 14.13. The dependence of the tailing parameter [92P1]

Further information on the localized states can be deduced from the luminescence spectra. In Fig. 14.14 we show the luminescence of a $CdS_{1-x}Se_x$ crystal under excitation of the extended states and under resonant excitation of the localized ones; see also the spectra in [92P1]. One sees here a rather broad zero-phonon line peaking at 2.21 eV and the first LO-phonon replica around 2.18 eV. The phonon replicas are resolved under resonant excitation into the vibrations of CdS and of CdSe, since in this alloy the phonons are of the "persistent-mode" type, i.e., they keep approximately the energies corresponding to $x = 0$ and $x = 1$ and contribute to the emission with a weight changing with x, while the gap in this case varies continuously with x. See Sects. 7.8, 11.1.6 and 8.15 and 9.6, respectively.

The high energy edge of the luminescence band can be identified with the transition region from extended to localized states [92P1, 99R1, 04K1]. This statement can be supported by the following argument: As long as excitons are in extended states, they "scan" the sample and have a good chance of hitting one of the fast nonradiative recombination centers mentioned above in Sect. 13.1.3. Once it is sitting in a localized state which does not have such a center within the localization length, the exciton can either recombine radiatively or reach a deeper localized state, e.g., by phonon-assisted tunneling. The probability for the latter process decreases with decreasing energy and density of the localized states, resulting in increasing luminescence yield. This

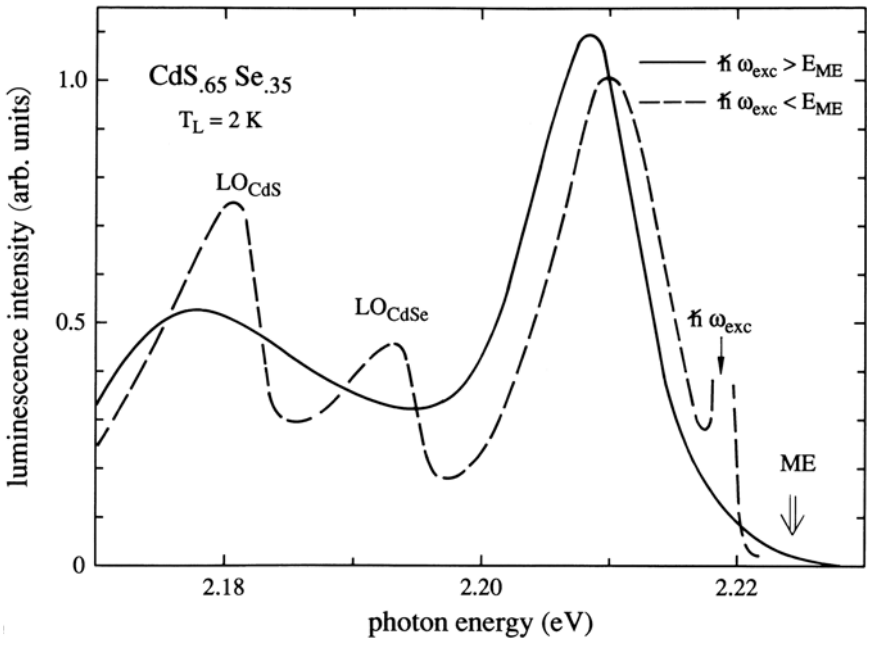


Fig. 14.14. Luminescence spectrum of $CdS_{1-x}Se_x$ under resonant and non-resonant excitation [93K1]

idea has been confirmed by investigations of the phase-relaxation times and of the polarization memory as will be outlined in Chap. 23. For a more recent discussion of the absorption and luminescence properties of bulk semiconductor alloys at low temperature involving the ideas of clusters, superclusters and of metastable states introduced in Sect. 8.15 and 9.6 see [99R1, 04K1] and the references given therein.

At higher lattice temperatures defined by

$$k_{\mathrm{B}} T_{\mathrm{L}} \gtrsim \varepsilon_{1\mathrm{oc}} \tag{14.8}$$

the majority of the excitons are thermally (re-) excited from the localized into extended states where they behave similarly to free excitons in ordered crystals except for a rather short phase-relaxation time due to the alloy scattering. This transition is partly connected to a non-monotonous dependence on the emission maximum and also partly on the FWHM of the band as shown for a $CdS_{1-x}Se_x$ sample in Fig. 14.15, which includes the CdS LO phonon replica. For an earlier example see, e.g., [78K1].

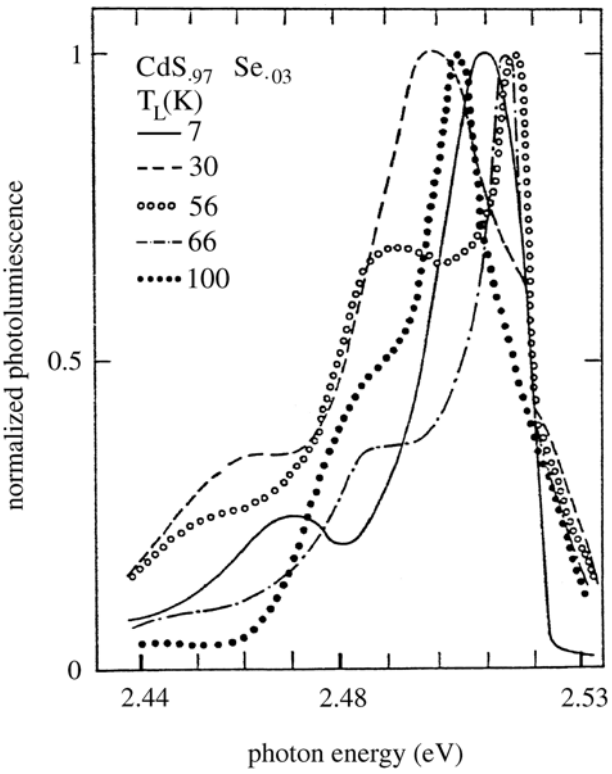


Fig. 14.15. The low excitation photoluminescence spectra of $CdS_{.97}Se_{.03}$ for various lattice temperatures T_l [88S1]

Such behavior was rediscovered recently in quantum well structures. We come back to this topic in the next chapter.

At low temperatures the luminescence yield of localized excitons can come close to unity in selected samples [97W1].

Similar results as presented for $CdS_{1-x}Se_x$ have also been observed in $ZnSe_{1-x}Te_x$, [92P1] and in $Ga_{1-y}A1_yAs$ [92S1]. In the first case the localization effects are even more pronounced than in $CdS_{1-x}Se_x$. Furthermore, self-trapping of excitons seems to occur for $x \approx 0.01$ at single Te atoms or at small Te clusters. In the second case, $\varepsilon_{\mathrm{loc}}$ is much smaller, partly due to the smaller translational exciton mass which makes localization less probable and partly because of the larger Bohr radius of the exciton which averages over a larger area of the lattice. More recently ZnO based alloys like $Zn_{1-y}Mg_yO$ or GaN based ones like $Ga_{1-y}In_yN$ have been investigated with similar results concerning the optical properties of excitons [01P1, 03S1]. Localisation effects occur also in group IV alloys like $Si_{1-x}Ge_x$. For some references see, e.g., [92L1, 92N1, 94T1]. The importance of disorder for Rayleigh scattering and the appearance of speckles is investigated in [04M1, 04Z1] and will be treated in Sect. 23.2.

For recent reviews on localized excitons in disordered semiconductors see [86L1, 89T1, 91K1, 92S1, 92S2, 99K1, 04K1] and references given in Sects. 8.15 and 9.6. As already mentioned localization of excitons in quantum wells and wires will be treated in the next chapter. But we mention here that (quantum-mechanical) level repulsion occurs for states, which are spatially and energetically close lying [03Z1] because this effect should also occur in bulk material, but is there much more difficult to observe experimentally.

14.5 Problems

1. For some standard semiconductors, such as Si, Ge, or GaAs, calculate the binding energy of electrons and holes to donors and acceptors, respectively. Find some data for the binding energies of excitons to these complexes and compare the results with Haynes' rule.
2. Calculate the shift in energy for the donor–acceptor pair recombination when donor and acceptor are nearest possible neighbors in a zinc-blende lattice, and when they have a separation of three, and of ten lattice constants. Can one observe emission from pairs with a separation $d \gg a_{\mathrm{B}}\varepsilon\frac{m_0}{m_c}$

References to Chap. 14

[60H1] J.R. Haynes, Phys. Rev. Lett. **4**, 361 (1960)
[72W1] D.L. Wood and J. Tauc, Phys. Rev. B **5**, 3144 (1972)
[74D1] A.F. Dite, V.I. Revenko and V.B. Timofeev, Sov. Phys. Solid State **16**, 1273 (1974)

[74S1] J. Shah, R.F. Leheny and W.F. Brinkmann, Phys. Rev. B **10**, 659 (1974)
[75K1] C. Klingshirn, phys. stat. sol. (b) **71**, 547 (1975)
[75O1] Optical Properties of Ions in Solids, B. Di Bartolo (ed.), Plenum, New York (1975)
[76H1] R. Helbig, Habilitation Thesis, Erlangen (1976)
[76K1] R. Kapelman, J. Lumin. **12/13**, 775 (1976)
[76T1] E. Tomzig and R. Helbig, J. Lumines. **14**, 403 (1976)
[77H1] B. Hönerlage and U. Schröder, Phys. Rev. B **16**, 3608 (1977)
[78B1] K. Bohnert et al., Solid State Commun. **27**, 295 (1978)
[78G1] F. Gärtner and E. Mollwo, phys. stat. sol. (b) **89**, 381 and ibid **90**, 22 (1978)
[78K1] A. Kozanecki et al., phys. stat. sol. (b) **89**, 313 (1978)
[78T1] M.L.W. Thewalt, Solid State Commun. **25**, 513 (1978)
[79B1] G. Blattner, Diploma Thesis, Karlsruhe (1979)
[79C1] K. Cho (ed.), Excitons, Topics Curr. Phys. Vol. **14**, Springer, Berlin, Heidelberg (1979)
[79D1] P. Dean, in Ref. [79C1], p 11
[79S1] H. Schrey, PhD, Karlsruhe (1979)
[80B1] G. Blattner, R. Helbig and C. Klingshirn, Solid State Commun. **33**, 341 (1980)
[81B1] G. Blattner et al., phys. stat. sol. (b) **107**, 105 (1981)
[81H1] K.H. Hellwege, Einführung in die Festkörperphysik, 2^{nd} edn., Springer, Berlin, Heidelberg (1981)
[81K1] R. Kuhnert and R. Helbig, J. Lumin. **26**, 203 (1981)
[84S1] B.I. Shkolovskii and A.L. Efros, Electronic Properties of Doped Semiconductors, Springer Ser. Solid-State Sci. **45**, Springer, Berlin, Heidelberg (1984)
[86G1] J. Gutowski, Solid State Commun. **58**, 523 (1986) and references therein
[86L1] Localization and Interaction, D.M. Finlayson (ed.), 31^{st} Scottish Universities Summer School in Physics, SUSSP, Edinburgh (1986)
[88S1] H.-E. Swoboda et al., phys. stat. sol. (b) **150**, 749 (1988)
[89H1] P.O. Holtz et al., Phys. Rev. B **40**, 12338 (1989)
[89S1] E.A. Shiff in Ref. [81A1]d, p 153 of Chap. 1
[89T1] P.C. Taylor in Ref. [89L1], p 257 of Chap. 1
[90K1] M. Kunz et al., Materials Chemistry and Physics **25**, 27 (1990)
[91K1] A.A. Klochikhin and S.G. Ogloblin, Sov. Phys. JETP **73**, 1122 (1991)
[91V1] A. Villemaire, T. Steiner and M.L.W. Thewalt, Phys. Rev. B **44**, 13426 (1991)
[92L1] L.C. Lenchshyn et al., Appl. Phys. Lett. **60**, 3174 (1992)
[92N1] J.-P. Noël et al., Appl. Phys. Lett. **61**, 690 (1992)
[92P1] S. Permogorov and A. Reznitsky, J. Lumin. **52**, 201 (1992) and references therein
[92S1] U. Siegner et al., Phys. Rev. B **46**, 4564 (1992)
[92S2] H. Schwab et al., phys. stat. sol. (b) **172**, 479 (1992)
[93K1] C. Klingshirn, NATO ASI Series B **301**, 119 (1993)
[93S1] T.W. Steiner et al., Phys. Rev. B **47**, 1265 (1993)
[94T1] K. Terashima et al., Appl. Phys. Lett. **65**, 601 (1994)
[96V1] K.A. Vanheusden et al., Appl. Phys. Lett. **68**, 403 (1996)
[97G1] J. Gutowski, P. Bäume and K. Hauke in Properties of Wide Gap II–VI Semiconductors, R. Bhargava ed., INSPEC, London (1997)

[97S1] M.I.N. da Silva et al., J. Appl. Phys. **82**, 3346 (1997)
[97W1] R. Westphäling et al., J. Luminesc. **72-74**, 980 (1997)
[98S1] T. Shinagawa and T. Okamure, Jpn. J. Appl. Phys. **37**, 1939 (1998)
[99K1] A. Klochikhin et al., Phys. Rev. B **59**, 12947 (1999)
[99L1] D. Lüerßen et al., Appl. Phys. Lett. **75**, 3944 (1999)
[99R1] A. Rezmitsky et al., Phys. Rev. B **59**, 10268 (1999)
[00F1] G. von Freymann et al., Appl. Phys. Lett. **76**, 203 (2000)
[01P1] W.I. Park, G.-C. Yi and H.M. Jang, Appl. Phys. Lett. **79**, 2022 (2001)
[02R1] E. Runge, Solid State Physics **57**, H. Ehrenreich and F. Spaepen (eds.), Academic Press, San Diege (2002)
[02S1] S. Stauf et al., Phys. Rev. Lett. **89**, 177403 (2002)
[02Z1] A. Zeuner et al., phys. stat. sol. b **234**, R7 (2002)
[03D1] M. Dremel et al., J. Appl. Phys. **93**, 6142 (2003)
[03E1] Electronic Structure of Alloys, Surfaces and Clusters, A. Mookerjee (ed.), Taylor and Francis, London (2003)
[03S1] M. Strassburg et al., phys. stat. sol. c **0** (6), 1835 (2003)
[03S2] J. Singh and K. Simakawa, Advances in Amorphous Semiconductors, Taylor and Francis, London (2003)
[03Z1] R. Zimmermann, E. Runge and V. Savona, phys. stat. sol. (b) **238**, 478 (2003)
[04D1] M. Decker, H. Priller and C. Klingshirn, to be published
[04K1] A. Klochikhin et al., Phys. Rev. B **69**, 085308 (2004)
[04M1] B.K. Meyer et al., phys. stat. sol. (b) **241**, 231 (2004)
[04M2] G. Mannarini, W. Langbein and R. Zimmermann, phys. stat. sol. c **1** (3), 489 (2004) and Phys. Rev. B **69**, 085326 (2004)
[04W1] D. Wang et al., J. Physics-Cond. Matter **16**, 4635 (2004)
[04Z1] R. Zimmermann and E. Runge, Phys. Rev. B **69**, 155307 (2004)

15

Optical Properties of Excitons in Structures of Reduced Dimensionality

As in the presentation of elementary excitations in Chaps. 7–12, we now proceed from bulk materials to the optical properties of excitons in systems of reduced dimensionality. For some recent reviews see [89S1, 90C1, 90G1, 93C1, 01H1, 03R1, 03S1] or, e.g., the following References [81A1 h,i,k,l, 93B1, 93H1, 93O1, 93P1, 93S1, 95I1, 96S2, 97W1, 98D1, 98G1, 98S1, 99B1, 01C1, 01H1, 01L1, 02S1] of Chap. 1.

In Sects. 8.10, 8.13 and 9.3 we outlined the influence of the dimensionality on the eigenstates of carriers and of excitons. Now we describe the optical properties of excitons in such structures, starting with quantum wells in Sect. 15.1 and proceeding to coupled wells, superlattices and to structures of even lower (quasi-) dimensionality in the following sections.

15.1 Qantum Wells

Here we treat transitions between valence– and conduction band (interband transitions) and intraexatonic transitions. For intersubband transitions see Sect. 21.5 We start with type–I structures, where the electron and hole are quantized in the same material.

The description will remain almost exclusively in the weak coupling limit, mainly because the quality of presently available samples in relation to interface roughness or alloy disorder is not sufficient for the fine details stemming from the polariton concept to be observable. The most widely investigated groups of MQW are based on GaAs/$Al_{1-y}Ga_yAs$ and on InP/InGaAs/InAlAs. See Sect. 8.11. In the meantime, considerable interest has also arisen in II–VI systems, especially for wide gap materials, like structures based on ZnCdMg/SSeTe, ZnMgCd/O and in new III–V systems like GaAlIn/N and GaInNAs for the blue and IR regions, respectively. The II–VI compounds are covered, e.g., in the corresponding proceedings like [82P1] in Chap. 8, while data on the nitrides can be found in [97N1, 00I1] of Chap. 1.

In Fig. 15.1 we show how the spectra of a quantum well system develop from the bulk material.

The layer with $l_z = 400\,\text{nm}$ is almost bulk-like and shows the 1 s exciton absorption peak followed by its ionization continuum. Compare this with Sect. 13.1.3.

With decreasing quantum well width, the excitons of the various subband transitions appear (labelled by $n_z = 1, 2, \ldots$) and the splitting into heavy hole and light hole excitons caused by the different, mass dependent quantization energies. Some weaker structures appear in between, which are dipole forbidden, but obtain some oscillator strength, e.g., via internal band bending. For further examples see, e.g., [01L1] of Chap. 1 and the references given therein.

In Fig. 15.2a,b we compare the spectra for $GaAs/Al_{1-y}Ga_yAs$ and $In_{1-y}Ga_yAs/InP$ multiple quantum well structures MQW.

We see the exciton and the continuum transitions between the first quantized heavy- and light-hole levels and the first quantized electron state, $n_z = 1hh$ and $n_z = 1lh$, respectively. The peaks are due to the $n_B = 1$ state. The higher states $n_B \geqslant 2$ are usually not visible, due to their rela-

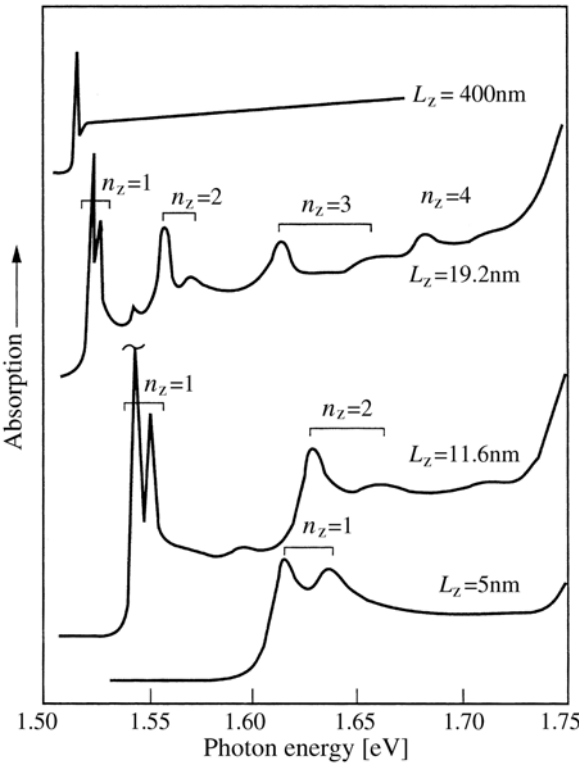


Fig. 15.1. Low temperature absorption spectra of GaAs (quantum) films of various thicknesses ([83G1, 90G1] or [01L1] of Chap. 1)

tively small binding energies and oscillator strengths (see Sect. 9.3) and due to the broadening mechanisms mentioned above. The absorption coefficient in the continuum states essentially reflects the constant density of states for two-dimensional effective-mass particles. It is not influenced much by the Sommerfeld enhancement because it varies only by a factor two. See Sect. 9.3 unlike the three-dimensional systems mentioned in Sects. 9.2 and 13.1. The next prominent structures are the $n_z = 2hh$ and lh excitons. Higher states are usually less clearly seen, among other reasons because the electron states are often no longer quantized for $n_z > 3$, depending on l_z, m_e and the band offset.

In contrast to the InP system, the GaAs substrate of the AlGaAs MQW is opaque for the exciton resonances. So it has to be removed by selective

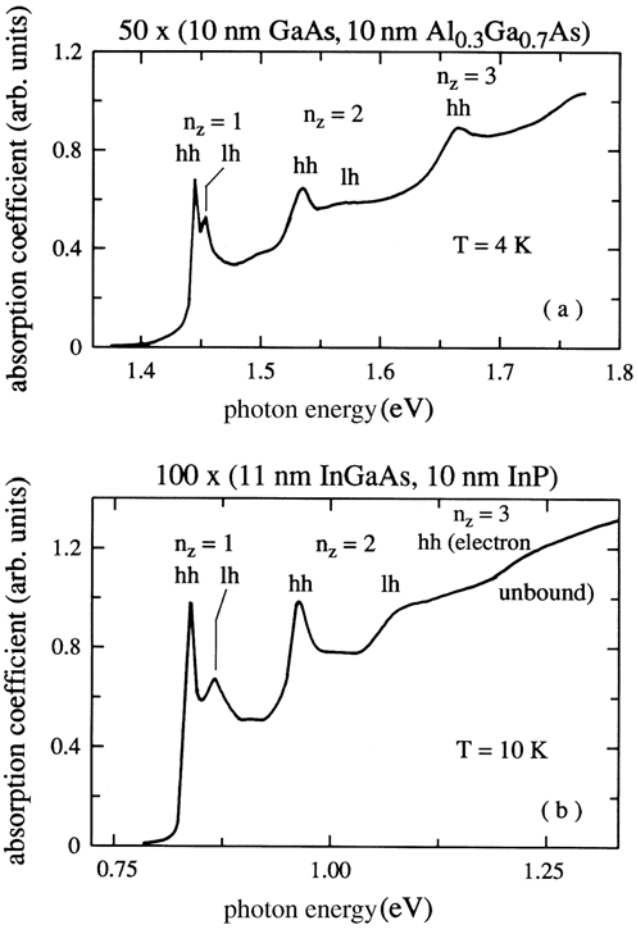


Fig. 15.2. The low temperature absorption spectra of multiple quantum well samples of $GaAs/Al_{1-y}Ga_yAs$ (**a**) and of $In_{1-y}Ga_yAs/InP$ (**b**) [89S1] and [01L1] of Chap. 1, respectively

etching [80P1] for the observation of absorption spectra. Before this technique was known, scientists often measured the luminescence–excitation spectra, i.e., they detected the exciton luminescence as a function of the photon energy of excitation $\hbar\omega_{\mathrm{exc}}$ for constant incident intensity (or rather constant photon flux density). Similarly, the dependence of the photo current on $\hbar\omega_{\mathrm{exc}}$ can be used in a pin diode biased in the blocking direction, which contains the (M)QW in the intrinsic region. The quantities

$$\left.\begin{aligned} I^{\mathrm{lum}} &= f(\hbar\omega_{\mathrm{exc}}) \\ j^{\mathrm{photo}} &= f(\hbar\omega_{\mathrm{exc}}) \end{aligned}\right\} I_{\mathrm{exc}} = \mathrm{const} \tag{15.1}$$

are related to, but not identical with, the absorption spectrum $\alpha(\hbar\omega)$. See Chap. 25. The above-mentioned techniques are still useful if a single quantum well (SQW) is investigated. The absorption coefficients of two-dimensional excitons in the above-mentioned systems are of the order of $10^4\,\mathrm{cm}^{-1}$. A SQW with $l_z = 10\,\mathrm{nm}$ therefore gives an optical density of only

$$l_z \cdot \alpha \approx 10\,\mathrm{nm} \cdot 10^4\,\mathrm{cm}^{-1} = 10^{-2}\,, \tag{15.2}$$

which is hardly detectable in a simple transmission experiment. The observation of such a small variation needs either highly developed modulation techniques or photo-luminescence excitation spectroscopy.

The advantage of studying only a SQW is that fluctuations of l_z from one QW to the next can be avoided. Furthermore some (opto-) electronic devices contain only a SQW, e.g., some field-effect transitors (MOD–FET, HEMT). The additional optical selection rules which arise when the excitons are confined to a QW are rather simple:

$$\Delta n_z = 0\,. \tag{15.3}$$

Transitions with odd Δn_z are forbidden by parity and those for even non-zero Δn_z are forbidden for rectangular shaped QW in the limit of infinetly high barriers. Transitions which violate (15.3) can be observed if there are symmetry breaking perturbations such as external or internal electric fields. Examples will be given in Sect. 16.2.

The luminescence spectra of excitons in QW are often rather broad and partly Stokes-shifted with respect to the absorption. We give two-rather different examples in Fig. 15.3. The effects can be explained by assuming that the luminescence is a superposition of the recombination of excitons in extended states at $\boldsymbol{k}_{\parallel} \approx 0$ (excitons with larger $\boldsymbol{k}_{\parallel}$ cannot decay radiatively because of momentum conservation), of excitons localized in tail states due to well-width fluctuations and/or alloy disorder, and finally of excitons bound to impurities such as neutral acceptors or donors. As we saw in Chap. 14, the latter complexes lead to spectrally very narrow emission bands in 3d systems, but in QW, their energies are strongly inhomogeneously broadened because the energy of the complex depends on its positions relative to the barriers.

The latter effect can be reduced by deliberately doping only a narrow layer, e.g., in the center of the well, as shown, e.g., earlier in Fig. 14.7.

In samples of very high quality the mechanisms mentioned above are less pronounced, resulting simultaneously in a reduction of the spectral widths of absorption and emission bands and of the Stokes shift. Both quantities are related to each other as shown, e.g., in [93Y2, 94G1]. An example is given in Fig. 15.3. The shoulders in the emission on the low energy side can have various origins: fluctuations of l_z by one monolayer, recombination of a bound exciton complex, or recombination of a trion or a biexciton (see Sects. 9.4 and 20.3 for the latter two processes). Due to the weaker coupling of carriers and excitons to optical phonons in the III–V compounds compared to the more

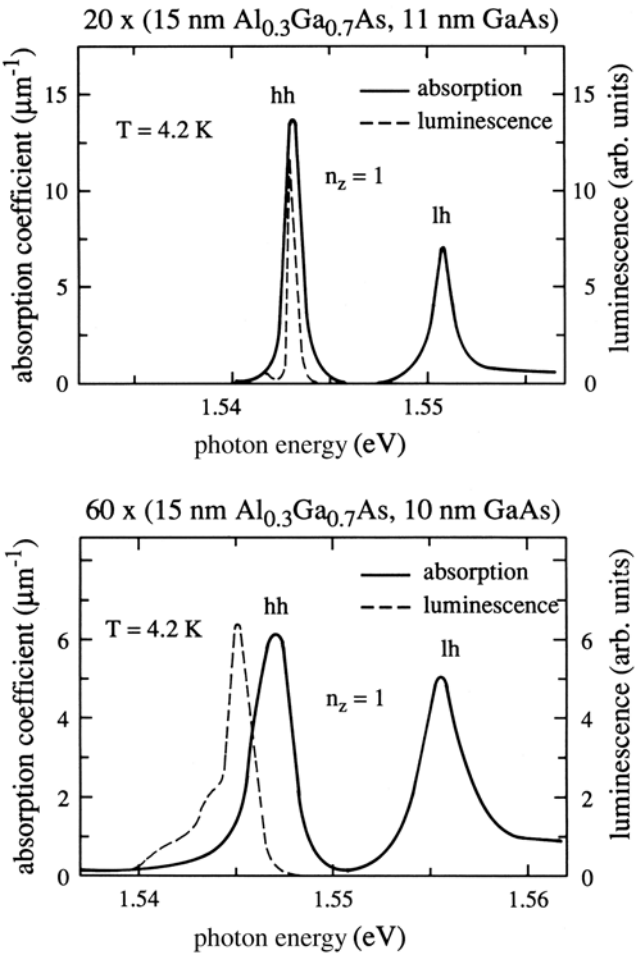


Fig. 15.3. A comparison of absorption and luminescence spectra in two samples of different quality [92O2]

ionic II–VI semiconductors, LO-phonon replica are hardly detectable in the former materials.

As an example of a II–VI system we show in Fig. 15.4 the photoluminescence (PL) and photoluminescence excitation (PLE) spectra of a ZnSe/$ZnS_{1-x}Se_x$ MQW sample.

One observes in PL the light and heavy hole exciton, a bound exciton complex (BEC) and their LO phonon replica. The PLE spectrum taken at the position of the hh luminescence shows the lh exciton resonance and above peaks spaced by integer multiples of LO phonons. They indicate that electron-hole pairs first form excitons and then relax by LO phonon emission in contrast to, e.g., GaAs systems where the carriers relax individually and an exciton is formed only at the end of the LO-phonon emission cascade [93K2, 95B2, 96W3, 96W4]. Concerning the discussion to which extend luminescence at the exciton energy is due to excitons or to a Coulomb correlated electron–hole gas or plasma see [02H1, 04C1] or [98K1] of Chap. 13. More details on relaxation processes will be given in Chap. 23.

In this context it is remarkable that in the early and euphoric times of quantum well research, most excitonic low temperature absorption and luminescence bands have been attributed to free excitons, even if their width reached or even exceeded 10 meV. Absorption and emission of free, quasi-two-dimensional excitons within the "light cone" $\boldsymbol{k}_{\parallel} \leq \omega/c$ would result in a linewidth of only one or a few hundred μeV, depending on the effective mass. Even if a homogeneous broadening corresponding to dephasing

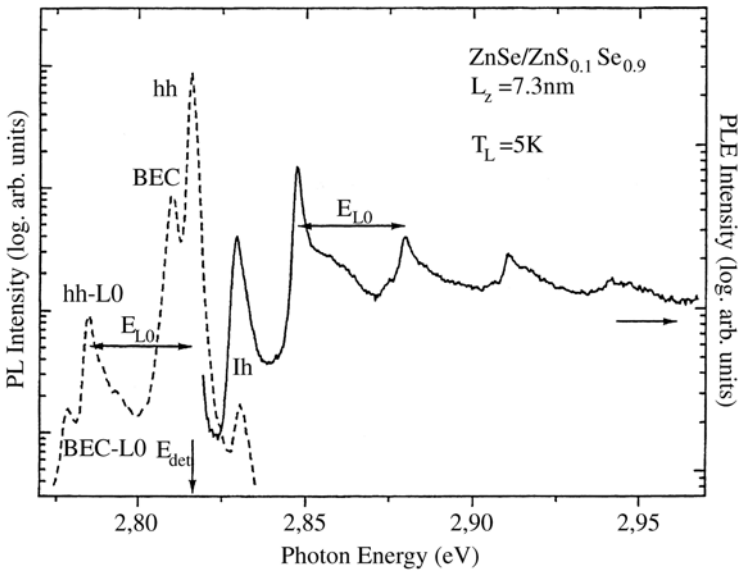


Fig. 15.4. Photoluminescence (PL) (- - - -) and PL-excitation spectra (——) of a ZnSe/$ZnS_{1-x}Se_x$ MQW sample [98K1]

times on the order of a few ps is added (see Chap. 23) one has to consider low temperature exciton line width exceeding 1 meV as substantially inhomogeneously broadened due to the formation of localized states caused by well width and alloy fluctuations and especially in luminescence by contributions of bound exciton complexes. Presently we see the opposite trend that every localization site is considered a quasi zero-dimensional quantum dot. As already mentioned in Sects. 8.13 and 9.3 and discussed in [01L1] of Chap. 1, there is presently no clear and generally accepted set of criteria to decide when a localization site should be called a quantum dot or island and vice versa.

The homogeneous width of excitons localized by well width and alloy fluctuations are at low temperatures often rather narrow ($< 100\,\mu$eV). If only a small ensemble of such localized states is probed, e.g., in luminescence, the localized excitons show up in emission as individual narrow lines while they merge to an inhomogeneously broadened band if their number increases. Examples are found, e.g., in [98P1,00L1] or [01L1] of Chap. 1. We give an example of a CdSe/ZnSe quantum well in Fig. 15.5, showing a macro- and microphotoluminescence spectrum (resolution 1 μm in the latter case) of a single CdSe quantum well in ZnSe.

The nominally deposited amount of CdSe was two monolayers. Due to segregation and diffusion it is spread over about 10 monolayers and islands of higher CdSe are formed resulting in the localization sites.

The relaxation and thermalization of the excitons in localized and extended states give rise to a frequently observed so-called S-shape dependence of the spectral position of the emission maximum of the macro-PL with temperature. Actually, it is really an n-shaped dependence, rather than an S-shaped one, but the term is established and is difficult to change.

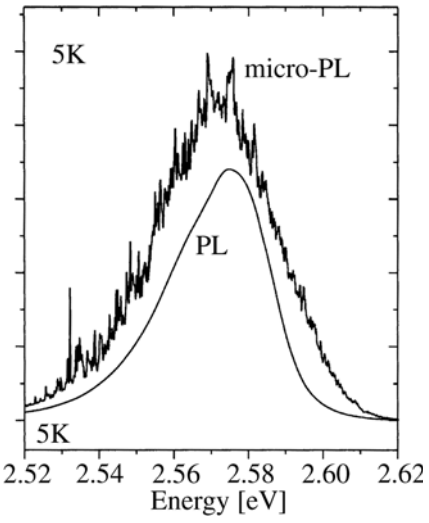


Fig. 15.5. Macro- and microphotoluminescence spectra of a CdSe/ZnSe single quantum well [01K3]

Examples of the non-monotonous variation of the emission maximum and of the half-width with temperature are given in Fig. 15.6 and in [98P1, 01T1, 01W1, 02R1, 03K2]. The behavior here is very similar to the behavior of localized excitons in bulk alloys, as already shown in Sect. 14.3. There are various approaches to theoretically model this behavior [01T1, 04K1].

If two localization sites of excitons with almost identical energies are spatially so close to each other in one quantum well that they interact slightly, a phenomenon occurs that we have already seen several times, namely, level repulsion. This level repulsion can be observed if high spatial and spectral resolutions are used simultaneously. For theoretical calculations and experimental evidence see, e.g., [03K3, 03N1, 03Z1] and references therein. A further effect, connected with the disorder inherent to quantum wells is (resonant) Rayleigh scattering. See, e.g., [03M1, 03N1] and references therein. We come back to this topic in Chap. 23.

Presently there is a trend to investigate excitons in individual localization sites or quantum islands by spatially resolved spectroscopy. See also Sect. 15.4.

At excitations below about 40 K the excitons are trapped, after excitation into the continuum states, randomly by various localization sites. The individual lines follow the temperature dependence of the band gap. The maximum of the macro-PL tends to show a stronger red shift for T between 50 and 100 K. In this regime the excitons undergo thermally activated hopping processes to neighboring, deeper localization sites. At even higher temperatures $T > 100$ K scattering into extended states dominates, which have a much higher density of states. Above 150 K the temperature dependence again roughly follows the band gap. The transition from localized to extended states is frequently ac-

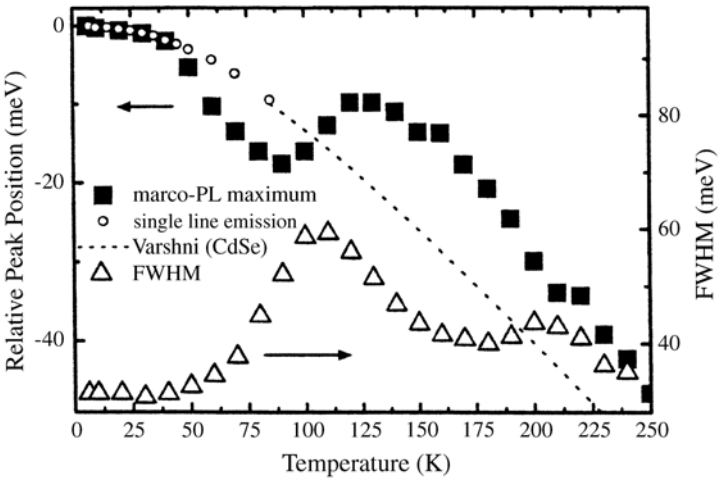


Fig. 15.6. The temperature dependences of the band gap of CdSe according to Varshni's formula, of the narrow emission lines of individual localization sites (o) of the maximum of the macrophotoluminescence (■) and of the FWHM (Δ) [01W1]

companied by a non-monotonous behavior of the width of the emission band as shown in Fig. 15.6.

The investigation of the reflection spectra of quantum wells is difficult due to inhomogeneous broadening and due to the fact that a single layer typically 10 nm in width contributes much less to a resonance signal than a bulk sample.

Nevertheless, it became possible to observe excitonic reflection spectra with improving sample quality. In Fig. 15.7 we give an example of a single GaAs well embedded between superlattices as barriers.

Under normal incidence, the typical reflectivity is around 30% for the AlGaAs system away from the resonance, and the heavy and light hole excitons in the well (A, B) and in the superlattice (C, D) produce only weak structures. To enhance them, the authors of [88U1] used a nice trick. They illuminated the sample with linearly polarized light under oblique incidence close to Brewster's angle (Fig. 15.7c). This procedure reduced the background reflectivity and thus enhanced the exciton resonances. Note the different scales in Fig. 15.7b and c on the ordinate.

In the meantime good reflection features have also been obtained under normal incidence in III–V and II–VI quantum wells. For examples see [96S1] or [01L1] of Chap. 1.

Exciton polaritons also exist in quantum well structures, however, the usually strong inhomogeneous broadening prevents the observation of finer

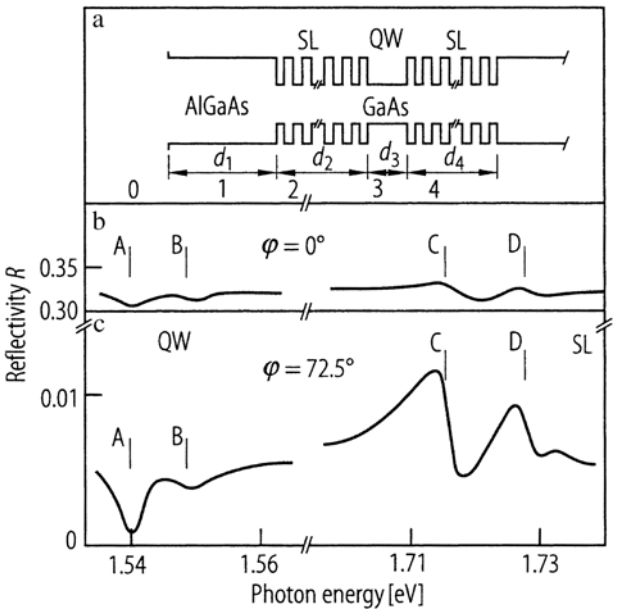


Fig. 15.7. Reflection spectra of a single GaAs quantum well with an $Al_{.32}Ga_{.68}As$/GaAs superlattice as barrier (**a**) and the reflection spectra for normal (**b**) and oblique (**c**) incidence [88U1]

details. In Fig. 15.8 we show calculated in-plane polariton dispersions for the heavy and light hole exciton resonances of an ideal single QW without any disorder using the assumption that the background dielectric constant of the well is equal to the dielectric constant of the barrier.

The polariton dispersion in the barrier material is shown by the dotted, almost vertical line. There exists a lower polariton branch that looks similar to the bulk material, with one difference being that the longitudinal-transverse splitting Δ_{LT} depends on the well width l_z and on $k_{\|}$ according to

$$\Delta_{\mathrm{LT}} \sim l_z \cdot k_{\|} \tag{15.4}$$

since there is no coupling of light to excitons in a single well with vanishing thickness, nor for infinite wavelength, i.e., $k_{\|} = 0$. The lower polariton branch corresponds to a guided mode since the refractive index below the resonance is higher in the well than in the barrier. There is a longitudinal mode but no upper polariton branch in the usual sense, since the refractive index above the resonance is smaller than that of the barrier. Consequently, one has a radiative or antiguided mode, which looses its amplitude rapidly (typical in 10 ps) by radiating into the barrier. If the background constant off the well exceeds that of the barrier, an upper polariton branch may appear as well [94J1]. The light hole resonance shows a splitting in the Γ_5 and Γ_4 state. The Γ_5 hh and lh states are dipole allowed for $\boldsymbol{E} \perp z$, where z is the growth direction, and show up in usual transmission spectroscopy, while Γ_4 is dipole allowed for $\boldsymbol{E} \parallel z$ only, and is therefore not seen in the usual geometry $\boldsymbol{k} \parallel z$; $\boldsymbol{E} \perp z$.

From the experimental side, time of flight measurements have been performed, which show a clear decrease of the group velocity around the resonances [90O1], but much less pronounced than for bulk material. See Sect. 13.1.4.

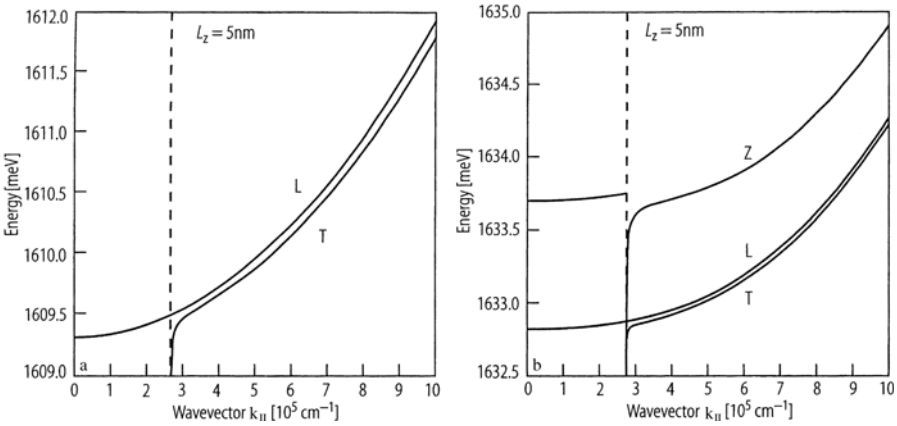


Fig. 15.8. The calculated in-plane polariton dispersion for an ideal 5 nm-wide single quantum well between $Al_{1-y}Ga_yAs$ barriers for the heavy (**a**) and light (**b**) hole exciton. Only dipole allowed states are shown [93J1]

Further information may be found in [90A1, 90T1, 95A1, 95B1, 95C1, 00M2, 01K1]. Some further suggestions to observe quantum well polaritons in experiment have been presented in [98K2].

For the geometry $\boldsymbol{k} \parallel z$ and $\boldsymbol{E} \perp z$, a stack of quantum wells acts as an ensemble of localized oscillators with infinite translational mass and with a very heavy one in the case of a superlattice, in the sense that it is an effective medium, as long as the distances between the wells and the well thickness itself are considerably lower than the wavelength of light. For radiative coupling between adjacent quantum wells see, e.g., [90I1, 94C1, 96S2, 96W1].

The topic of polaritons obtained new interest when quantum wells were incorporated in a microcavity. We come back to this aspect in Chap. 17.

We mention the inversion between the longitudinal transverse splitting Δ_{LT} and the singlet triplet splitting Δ_{st}. For bulk semiconductor samples with direct, dipole-allowed gaps one finds for the 1 s Wannier exciton (polariton)

$$\Delta_{\mathrm{LT}} > \Delta_{\mathrm{st}} \tag{15.5}$$

while Δ_{LT} decreases with decreasing volume fraction of the material showing the resonance Δ_{st} increases with increasing confinement. This is already true for (multiple) quantum wells and even more so for quantum dots. See also Fig. 9.2.

Intraexcitonic transitions have been observed, e.g., in $\mathrm{GaAs/Al_{1-y}Ga_yAs}$ quantum wells in [89O1, 96C1, 96S3]. For the observation of this transition as a proof for the presence of excitons see [04C1].

15.2 Coupled Quantum Wells and Superlattices

Symmetric and asymmetric coupled double quantum wells show interesting features like level splitting. We show in Fig. 15.9 luminescence spectra of an asymmetric, coupled quantum well consisting of two GaAs wells, which are 9 and 18 nm wide and are separated by 3 nm of $\mathrm{Al_{.32}G_{.68}As}$.

With increasing pump power, the exciton gas is heated (the effective exciton temperature is given in the inset as a function of pump power) resulting in wider emission bands and increasing importance of the higher states proceeding from the *hh* exciton of the wide well (ww) to the *lh* exciton of the narrow well (nw). Coupled quantum wells are an attractive topic for intersubband spectroscopy. We come back to this aspect in Chap. 21.

Now we concentrate on superlattices, which are characterized by miniband formation in growth direction, as outlined in Sect. 9.3. We first give some general examples and then concentrate on some peculiarities.

The optical properties of type-I superlattices (SL) do not deviate very much from those of MQW except for some broadening due to the formation of mini-bands in the direction of periodicity of the SL and the increase of the importance of monolayer fluctuations with decreasing l_z. In superlattices

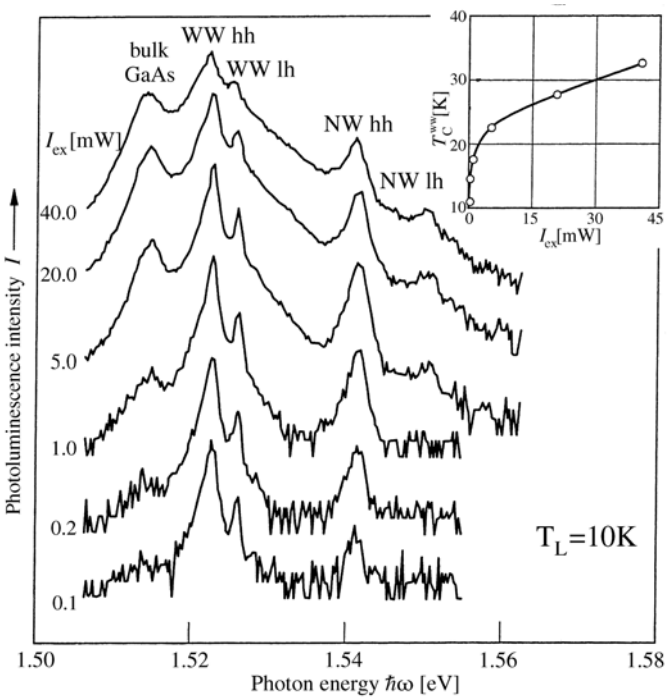


Fig. 15.9. Luminescence spectra of an asymmetric coupled double quantum well sample at $T = 10$ K for various cw excitation conditions [94L1]

obviously also fluctuations in the barrier width are contributing to disorder. For an investigation of the concept of polaritions in superlattices see, e.g., [95B1, 97P1].

Figure 15.10 shows the absorption spectrum of a superlattice consisting of $200 \times (2.2\,\text{nm GaAs}/2.5\,\text{nm Al}_{.3}\text{Ga}_{.7}\text{As})$.

Very similarly to a multiple quantum well sample one observes the $n_B = 1hh$ and lh exciton resonances and possibly, as a small shoulder, the $n_B = 2hh$ exciton. The features S_{hh} and S_{lh} are attributed to saddle point excitons at the van Hoove singularities at the boarder of the first mini-Brillouin zone.

In Fig. 15.11 we show absorption spectra for a series of ZnSe/ZnS superlattices. The increasing blue shift with decreasing well width is obvious. Also, the increasing spectral broadening of the light and heavy hole features are due to the increasing relative importance of well and barrier width fluctuations and alloying as already mentioned above.

The CdS/ZnSe system forms superlattices and quantum wells of type II with electron confinement in the CdS layers and a remarkably large conduction band offset by 750 ± 50 meV. The second remarkable finding both in quantum wells and superlattices is the wide tunability of the emission peak position from about 2.0 to 2.6 eV [99D1, 00D2, 01S3, 04P1].

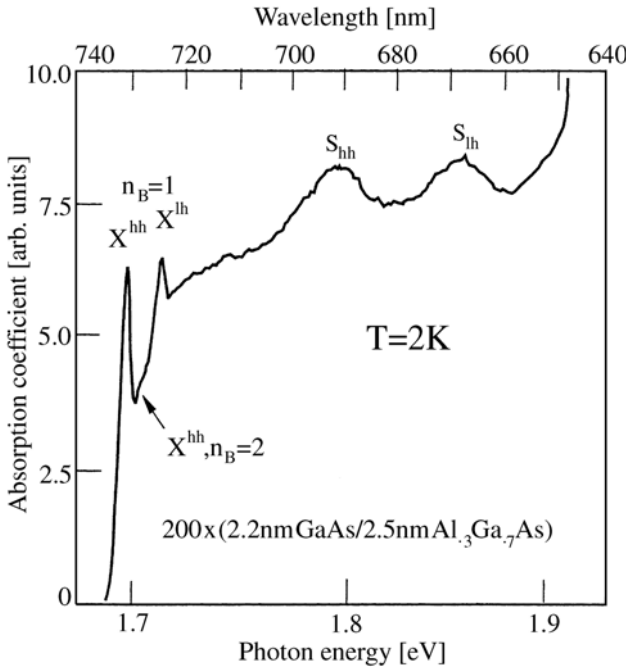


Fig. 15.10. The absorption spectrum of a GaAs/$Al_{1-y}Ga_yAs$ superlattice [90C2]

In the superlattice, the electron and hole wave functions have a stronger overlap than in single and multiple quantum wells of the same CdS well width due to the confinement of the holes in ZnSe in the first case. Additionally, the overlap increases in both cases with decreasing width of the CdS well. Consequences are a more strongly pronounced hh exciton-feature in the superlattice compared to quantum wells, a higher oscillator-strength and a faster luminescence decay time. Peculiarities of single CdS quantum wells will be presented in Chap. 23. In quaternary type II systems like BeTe/ZnSe or CdS/ZnSe the overlap of electron and hole wave functions is limited to the interface region. Depending on the type of interface the bands can be oriented in (110) or (1-10) directions. Depending on this orientation the luminescence can be polarized preferentially in one of these two directions, although both constituents have cubic symmetry. Examples for the BeTe/ZnSe system are found, e.g., in [00Y1] and for CdS/ZnSe in [00S1].

The II–VI compounds usually have a rather strong lattice mismatch, as can be seen from data of Sec. 8.11. As a consequence only rather thin films (a few mono-layers) can be grown if one is to avoid the formation of misfit dislocations. The results are so-called strained-layer superlattices where the two different materials forming the wells and the barriers adapt to a common, intermediate lattice constant.

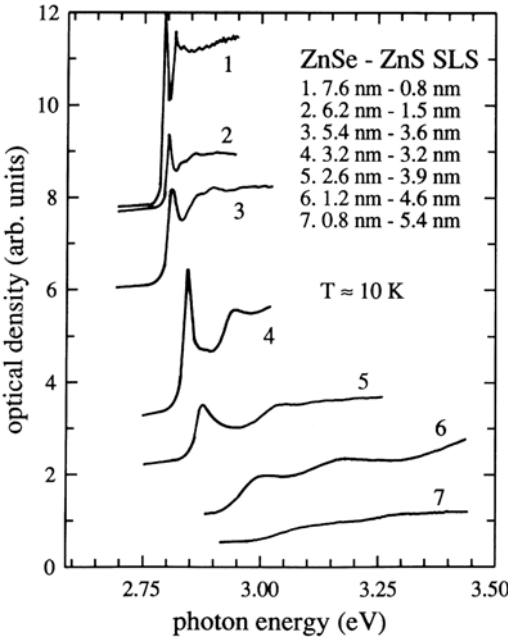


Fig. 15.11. Absorption spectra of a series of ZnSe/ZnS superlattices with various well (ZnSe) and barrier (ZnS) widths [92O1, 92S1]

The strain in superlattices mentioned above produces, via the piezoeffect in the layers, electric fields in the growth direction. The piezoeffect requires at least partially ionic binding and the absence of inversion symmetry in the crystal. Both conditions are fulfilled, e.g., for the zincblende- and wurtzite-type III–V and II–VI compounds. Since these built-in fields tilt the bands and shift the exciton energies, such superlattices are also known as Stark-superlattices.

The piezoeffect is generally stronger in the more ionic II–VI compounds compared to III–V materials and in wurtzite structure compared to the zincblende one. We present here example data for wurtzite CdSe/CdS superlattices from [94G1, 94K1, 94L2, 95L2]. References for other materials are found, e.g., in [93C2, 94K1, 95L1].

CdSe/CdS has a slightly type II band alignment with a conduction band offset of $100 \pm 100\,\mathrm{meV}$. In Fig. 15.12a we show the conduction and valence band edges including the envelope wave functions of electrons and holes in the first two minibands at $k_z = 0$, the temporal evolution of the emission spectra after ps excitation with $2.46\,\mathrm{eV}$ photons, a fluence of $0.3\,\mathrm{mJ/cm^2}$ and the calculated shift of the various transitions with time together with experimental data of the emission maximum.

The conduction and valence band edges are tilted by alternating piezoelectric fields with a field strength of $2 \times 10^8\,\mathrm{V/m}$. Figure 15.12b gives the

same results calculated self-consistently for a carrier density per layer of $1.2 \times 10^9\,\mathrm{cm}^{-2}$. Obviously the carrier separation into the different layers produces a strong depolarisation field that almost compensates the piezo fields under the conditions of Fig. 15.12b. As can be seen by comparing Figs. 15.12a and b both the transition energies and the electron hole overlap, i.e., the transition matrix element depends strongly on the carrier density. This effect can be nicely seen under pulsed (Fig. 15.12c) or cw excitation [94K1].

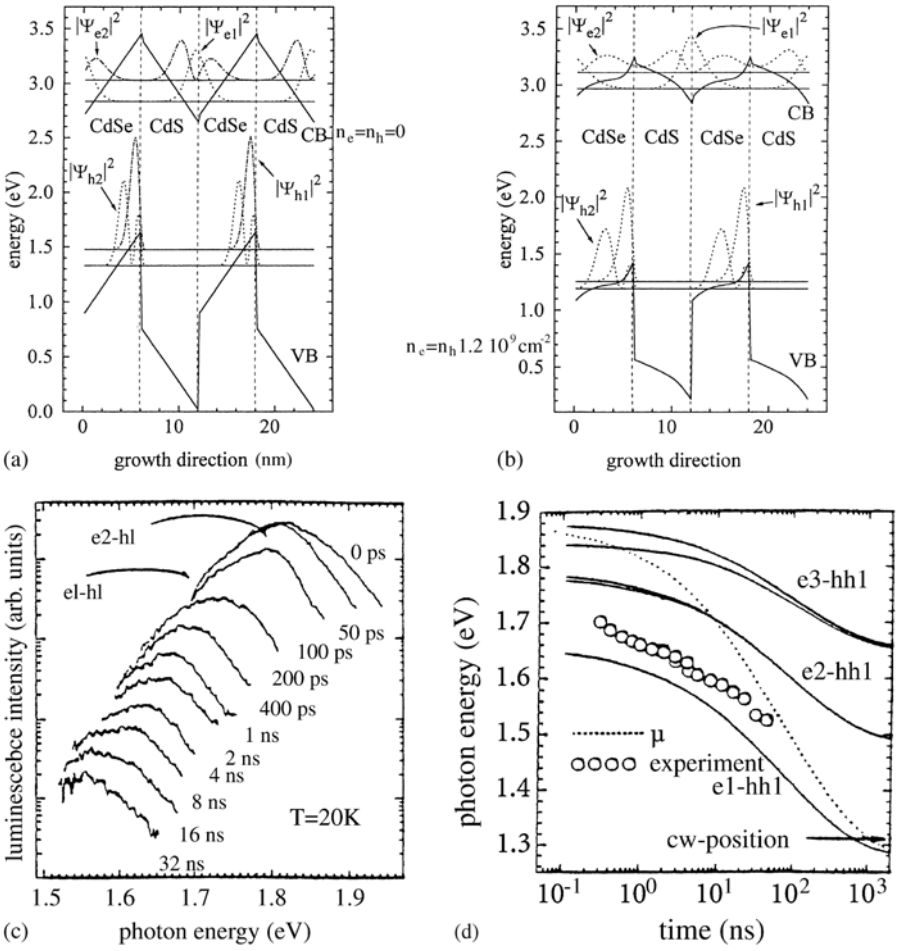


Fig. 15.12. The conduction and valence band edges of CdSe/CdS superlattices influenced by the piezofield (**a**) for zero carrier density and for a density of $n_\mathrm{e} = n_\mathrm{h} = 1.2 \times 10^9\,\mathrm{cm}^{-2}$ (**b**) the temporal evolution of the luminescence spectra after ps excitation (**c**) and the experimental and calculated shifts of emission maxima with time (**d**) [94K1, 95L2]

Both the decay rate and the position of the emission maxima shift to lower values with increasing time after the excitation pulse, leading to an almost logarithmic dependence, as seen in Fig. 15.12d. The temporal evolution of various calculated inter-band transitions are compared to the one of the emission maximum.

In addition, the position of the chemical potential μ of the electron hole pair system is given, i.e., the energetic distance between the quasi-Fermi levels of electrons and holes. At the beginning, it reaches values in the miniband, i.e., one has population inversion. Indeed stimulated emission has been observed from the transition e2 $\rightarrow$ *hhl*, which has a larger oscillator strength than the e1 $\rightarrow$ *hhl* process [95L2].

A good review of the present state of the art of II–VI epitaxy and of their optical properties is found in the proceedings of recent conferences on this topic ([82P1] of Chap. 8 or [91T1]).

Since the luminescence tail of excitons localized by well-width fluctuations in II–VI superlattices is sometimes spectrally broader than the energy of the optical phonons, and since the coupling of the excitons to the phonons is stronger than in the III–V system AlGaAs, one can observe interesting phenomena related to the relaxation or "freezing" of excitons in these tail states. More information about this topic and about the magnetic polaron effects which appear for II–VI SL containing Mn can be found in [93H1, 93J2, 93K1].

We come back to this topic in the sections of magnetic field effects (Sect. 16.1) and on time-resolved spectroscopy Sect. 23.3.

Corrugated superlattices have been grown on highly indexed substrate surfaces like (113). For some time the aspect of quantum well-wire superlattices has been discussed for this type of structure. However, more recently it has been found that the quantum wire aspect in these structures is marginal or completely absent [96L1].

The transition from a type-I to type-II band alignment can be observed in short period GaAs/AlAs SL.

The generally used nomenclature is $(\mathrm{GaAs})_m(\mathrm{AlAs})_n$ where m and n give the number of respective monolayers. GaAs is a direct gap material and AlAs an indirect one with a larger gap. If the width of the GaAs barrier is made smaller and smaller, then the $k = 0$ $n_z = 1$ conduction band state shifts to higher and higher energies until it is situated above the indirect conduction band state in AlAs. This situation is shown schematically in Fig. 15.13.

In contrast to the type-II GaSb/AlAs SL, the short period GaAs/AlAs SL are indirect in real space and in $\boldsymbol{k}$-space. The peculiarities of absorption in the well, relaxation of the electrons into the "barrier", and subsequent recombination are discussed in detail in [90G1]. An example is shown in Fig. 15.13, where the photoluminescence excitation spectra and the luminescence spectra are given for a type-I and a type-II GaAs/AlAs SL.

In the case of a type-II arrangement, the optical matrix element from the direct maximum in the valence band of GaAs to the indirect minimum in the conduction band of AlAs is so small that the transition hardly shows up in

the absorption spectrum. The thermal population factor is responsible for the fact that the low-temperature luminescence nevertheless originates essentially from this transition (Fig. 15.13c).

We mention that efforts are also being made to grow strained Ge/Si SL, though there are considerable technical problems in growing good quality samples due to the different lattice constants. The idea is that a similar folding-back mechanism as outlined in connection with phonons, in Sects. 7.10 and 11.2 should yield a direct electronic bandstructure which would allow the highly developed Si technology to be used also for light-emitting devices.

For recent publications and reviews of this topic see, e.g., [92N1, 92V1, 92W1,93A1,93M1,93V1,94H1,95F1,95L1,95Z1,96L2,96O1,97K1,00G1,00P1]

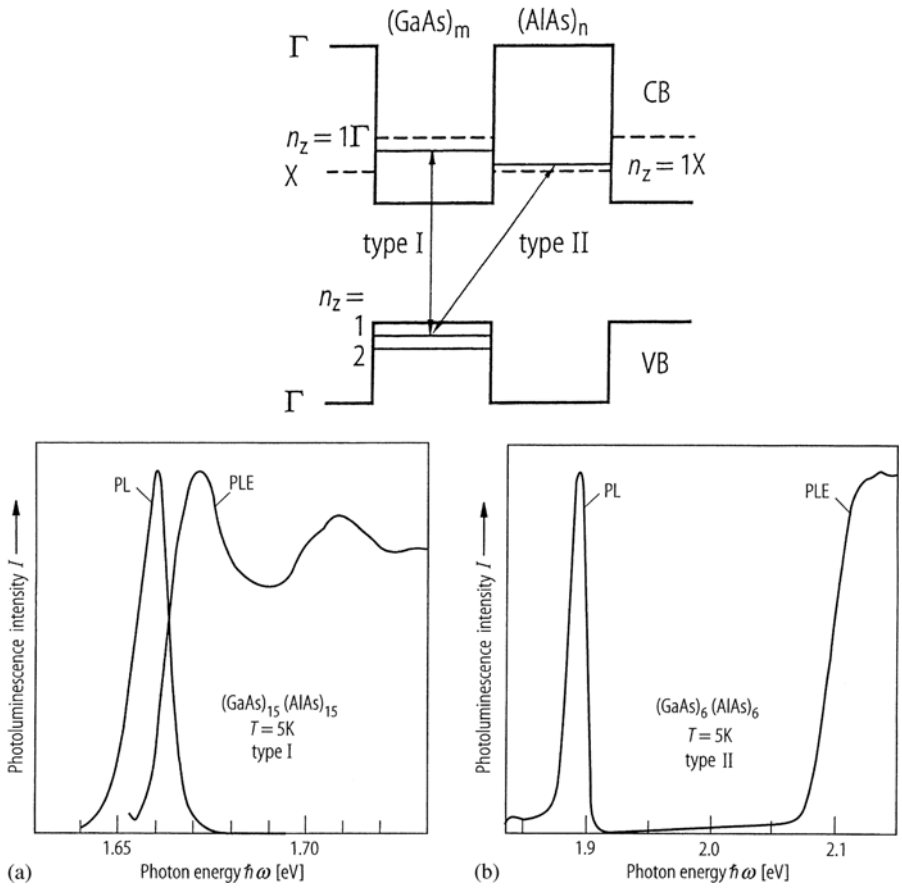


Fig. 15.13. The band alignment of a $(GaAs)_m(AlAs)_n$ short-period superlattice of type II showing the conduction and valence band edges at Γ (——), the conduction band edge at X (- - - -) and the quantized electron and hole states (**a**), PL and PLE for a type I (**b**) and a type II superlattice (**c**) [90G1]

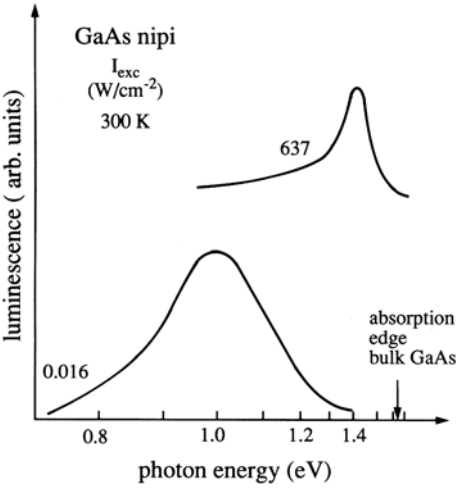


Fig. 15.14. Luminescence spectra of a nipi structure compared to the RT absorption edge of bulk GaAs [86D1, 92M1]. Compare with Fig. 15.12c

and references therein, which partly indicate that even this approach is unlikely to result in extremely high luminescence yields.

To conclude this section we briefly return to the concept of nipi structures or doping superlattices introduced in Sect. 8.11. The probability of transitions between electron states in the conduction-band minima and holes in the valence band maxima is low because of the "spatially indirect" bandstructure. Consequently strong absorption sets in only around the energy of the "direct" gap, i.e., at an energy corresponding to that between the bands at the same place, modified by the Coulomb interaction. The carriers are spatially separated after their creation resulting in a very long lifetime. The radiative recombination which can nevertheless occur due to the small overlap between electron and hole wavefunctions in their respective spatial band minima is strongly Stokes-shifted with respect to the spatially direct absorption edge. We give an example in Fig. 15.14, which shows the emission spectra for two different pump intensities and the position of the absorption edge of bulk GaAs. An analogous shift of the emission with excitation is also seen in piezo-superlattices see Fig. 15.12 or [93C2, 94K1, 95L2]. The nipi structures have quite amazing nonlinear optical properties [86D1, 92M1].

15.3 Quantum Wires

As already mentioned in Sects. 8.12 and 9.3 there are various possibilities to produce quantum wires, e.g., by laterally structuring quantum wells, by growing in V-groves or on ridges, by cleaved edge overgrowth producing T-shaped structures or by growing free standing, whisker-like needles, also known as nanorods. See, e.g., [90C1, 90G1, 93C1, 93D1, 93P1, 02K1, 03Y1] and references therein.

In the following, we present a few examples.

The problem encountered in lateral structuring is that the quantization length l_z in the growth direction of the QW is usually much smaller than the length l_y of the subsequent lateral structuring. Typical values are

$$\begin{aligned} l_z &\approx 10\,\mathrm{nm}\,, \\ l_y &\approx 50 - 100\,\mathrm{nm}\,. \end{aligned} \tag{15.6}$$

Consequently, the confinement energies resulting from confinement in the two directions are very different. Usually one observes only a narrow modulation of the luminescence or photoluminescence excitation spectra due to quantization in the y-direction.

We provide examples for the confinement in the second direction in Figs. 15.15 and 15.16.

Figure 15.15 shows PLE spectra of a 8 nm-wide GaAs quantum well between $Al_{1-y}Ga_yAs$ barriers in the upper trace, showing the well-known heavy and light hole excitons.

In the PLE spectrum of the quantum wires these two resonances are slightly blue shifted according to the lateral confinement and for the heavy hole exciton even the first two laterally confined states could be resolved, labeled hh_{11} and hh_{12}. The PLE spectra have been taken at the spectral positions indicated by the vertical arrows.

The progress in nanolithography is obvious when comparing the data in Fig. 15.16 for InAs wells.

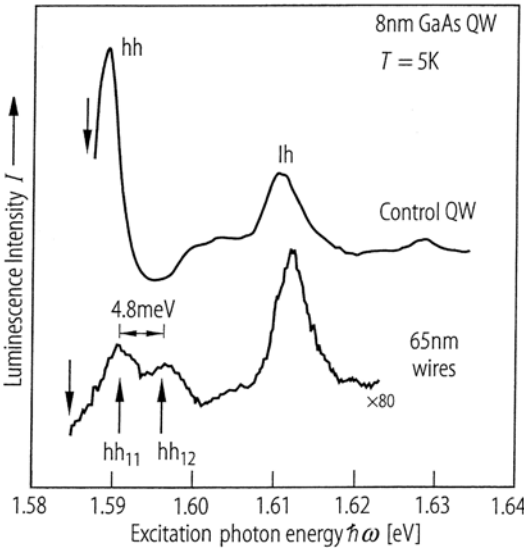


Fig. 15.15. PLE spectra of a 8 nm-wide GaAs/$Al_{1-y}Ga_yAs$ multiple quantum well sample (*upper trace*) and of 65 nm-wide quantum wires patterned from this sample (*lower trace*) [95S2]

Samples have been prepared by either etching through the upper barrier and the well into the substrate (deep etched), or only the upper barrier has been thinned down close to the well (modulated barrier). In both cases, a significant increase of the exciton binding energy from the value of the well around 8 meV by a factor of 1.5 or 2 has been obtained by additional lateral confinement.

In Fig. 15.17 we show the cathodo-luminescence spectra of a single GaAs V-groove wire.

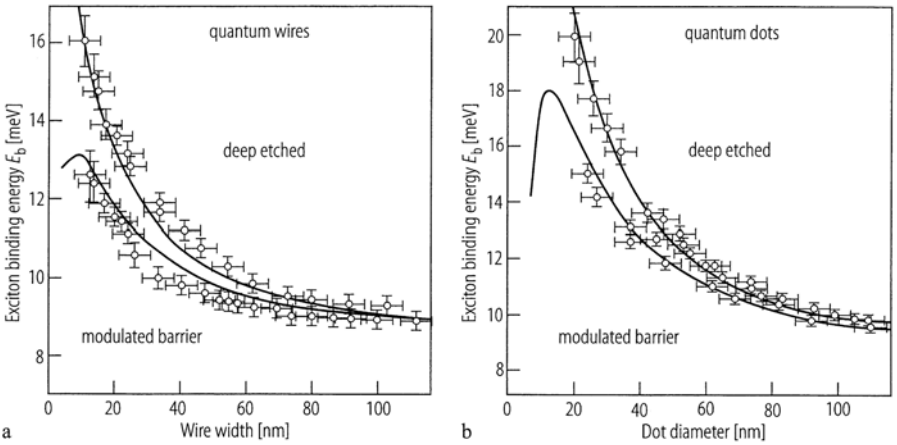


Fig. 15.16. Exciton binding energies as a function of wire width (**a**) or of dot diameter (**b**) for InAs samples [98B2]

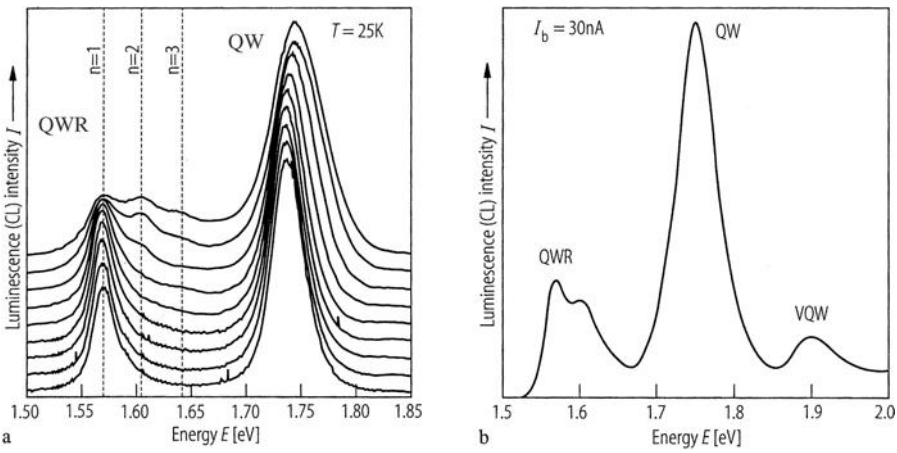


Fig. 15.17. Normalized cathode-luminescence spectra of a single GaAs V-groove quantum wire for increasing electron beam density on the sample at an acceleration voltage of 12.5 kV (**a**) and an overview over the whole spectrum (**b**) [95G1]

The overview spectrum shows, at the lower energy side, the emission from the quantum wire (QWR) at the bottom of the V-groove. At the higher energy side it shows the emission from the quantum wells (QW) at the side flanks of V-groove, which are considerably narrower, and therefore the form lateral confining potential of the wire-states though have higher quasi dimensionality $d = 2$. The emission VQW finally results from a vertical quantum well which forms unintentionally in the middle of the V-groove.

The lateral confining potential of the V-groove is not as abrupt as, e.g., quantum wires formed by nanolithography, but rather smooth. Therefore the higher states, which are populated at higher excitation current and are seen in luminescence, are in first approximation as equally spaced as in a harmonic oscillator. They are labeled in Fig. 15.17a by $n = 1, 2, 3$.

As a last example we show in Fig. 15.18 the luminescence spectra of two different T-shaped quantum wires formed by cleaved edge overgrowth. Their emission is again red shifted with respect to the emission from the quantum wells on the (001) and (110) surfaces.

For recent data both for GaAs and $In_{1-y}Ga_yAs$ based structures see, e.g., [88H1, 96L4, 98L1, 99F1, 00D1, 00K1] and references therein.

For localization in quantum wires see, e.g., [97R1, 98V1].

The dispersion relations of quantum wire polaritons are calculated for various cross-sections, e.g., [92T1,93C1,93K1,93Y1,99C1] yielding partly different results. For experimental data see, e.g., [88K1].

For the optical properties of nanorods see the above references and [97S1, 01H2, 02K3, 04P2, 04L1].

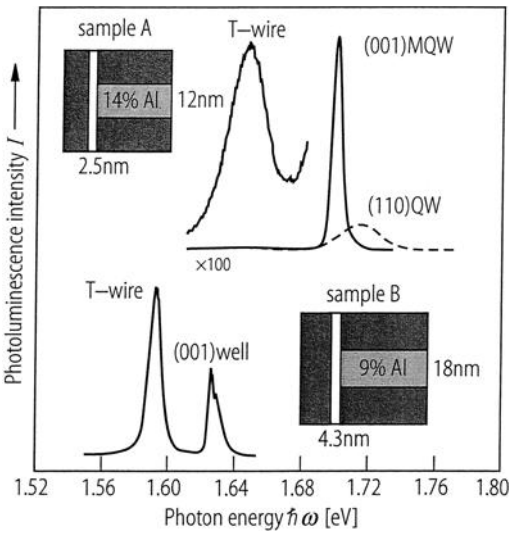


Fig. 15.18. Low-temperature photoluminescence spectra of two different T-shaped quantum wires [96L3]

15.4 Quantum Dots

As already explained in Sects. 8.13 and 9.3 there are many different ways to produce quantum dots or islands.

We present here some selected examples of "real" dots or nanocrystals that are embedded in glass or polymer matrices with high barrier energies in the eV range and almost spherical shape and for quantum islands formed in quantum wells by generally more oblate insertions of low-band-gap material.

The most frequently used groups of semiconductors for the first type of quantum dots are Cu-halides and II–VI compounds and to a much lower degree III–V materials. For the latter group, examples are found in [96G1, 97B1] and for the two former groups in [93B1, 97W1, 98G1, 01L1] of Chap. 1.

The research on epitaxially-grown quantum islands started essentially with the III–V system In(Ga)As/GaAs followed by the II–VI material combination CdSe/ZnSe.

A limited selection of textbooks, monographs and review articles on this topic is: [81A1 g,k,l, 93B1, 97W1, 98G1, 01H1] of Chap. 1 and [86B1, 87S1, 90M1, 92P1, 93W1, 95P1, 95W1, 00K2, 00L2, 01K1, 01K2, 01K3]. In these references some comments can also be found on the history of this field. Actually, II–VI semiconductor doped glasses have been used as color- and edge filters from the IR to the near UV from the 1930s. It was also known that these glasses obtain their color only during an annealing process and that the absorption edges shifts to the red with increasing annealing time and temperature. The fact that this is a consequence of quantum confinement in nanocrystallites was mentioned explicitly only in the beginning of the 1980s, e.g., by [82E1, 82H1, 83J1].

A rather recent development in this field is nanostructures grown in inverted tetrahedral pyramids [98H1, 00K2, 00L2] and references therein.

One distinguishes three quantization regimes. Below we give some formulae for the quantization energies assuming spherical dots, simple isotropic and parabolic conduction and valence bands and infinite barriers. As outlined already in Sects. 8.13 and 9.3 the regime I is obtained for dots larger than the excitonic Bohr radius;

$$\text{regime I: } \bar{R} \gtrsim a_{\mathrm{B}} \,. \tag{15.7a}$$

In this case, only the center of mass motion of the exciton is quantized, while the relative motion of electron and hole are hardly affected. The quantization energy is given, in the above approximations of a spherical dot with infinitely high barriers and neglecting the difference in the dielectric properties of the microcrystallite and the surrounding matrix, by

$$\Delta E_{\mathrm{Q}}^{\mathrm{I}} = \frac{\hbar^2 \pi^2}{2M\bar{R}^2} \,, \tag{15.7b}$$

where $M = m_\mathrm{e} + m_\mathrm{h}$ is the translationed mass of the exciton. Regime II is given by

$$\text{regime II: } a_\mathrm{B}^\mathrm{e} > \bar{R} > a_\mathrm{B}^\mathrm{h}\,, \tag{15.7c}$$

where $a_\mathrm{B}^\mathrm{e,h}$ are the respective radii of the electron and hole orbits in the exciton around their common center of mass.

In this case the motion of the electron is quantized and the hole moves in the potential of the QD and of the space-charge of the electron for the usual situation $m_\mathrm{e} \ll m_\mathrm{h}$.

Continuing to use the above-mentioned approximations, and neglecting the Coulomb energy between electron and hole, one finds

$$\Delta E_\mathrm{Q}^\mathrm{II} = C\mathrm{Ry}^* \left(\frac{a_\mathrm{B}\pi}{\bar{R}}\right)^2 ; \quad \text{with} \quad C = 0.67\,. \tag{15.7d}$$

Regime III is defined by

$$\text{regime III: } \bar{R} \ll a_\mathrm{B}^\mathrm{e,h}\,. \tag{15.7e}$$

Since the quantization energy scales with $\bar{R}^{-2}$ and the Coulomb energy with $\bar{R}^{-1}$, for this regime one can ww.mozilla.org/start/1.7/ expect as a first guess

$$\frac{\hbar^2\pi^2}{2\mu\bar{R}^2} = E_\mathrm{Q}^\mathrm{III} \gg E_\mathrm{Coulomb}^\mathrm{III} \tag{15.7f}$$

where μ is the reduced mass of electron and hole. The influence of the Coulomb energy is still partly incorporated, e.g., in a formula used in [97M1]

$$\Delta E(R) = \frac{\hbar^2\pi^2}{2\mu\bar{R}^2} - \frac{1.78e^2}{\varepsilon\bar{R}} + 0.752E_\mathrm{Ry}\,. \tag{15.7g}$$

Slightly different formulae can be found, e.g., in [86S1, 91E1] or in [01L1] of Chap. 1.

The values of $\bar{R}$ where (15.7e,f) hold are comparable to the lattice constant. In this situation, the whole effective mass concept breaks down and a description of the dot in terms of molecular orbitals becomes more appropriate. An exhaustive list of references for calculations using the full valence band structure of cubic and hexagonal quantum dots and of calculations for very small clusters where tight binding methods are more appropriate is given in [01L1] of Chap. 1.

Recently it became obvious that the size of the dots is not the only relevant quantity affecting their linear and especially nonlinear optical properties. The three different growth regimes of the dots which occur during the annealing process of the glasses, namely nucleation, normal growth of the dots and growth through coalescence, result is significantly different properties of, among other things, the interface states between glass matrix and semiconductor dot. Furthermore, the capping of dots plays a crucial role. In glass or organic matrices one generally has a lot of surface and interface states, the

number of which can be significantly decreased by capping with hydrogen or organic ligands, which saturate the dangling bonds, or with similar semiconductors with a higher band gap of type-I band alignment. For recent reviews see [88H2, 90K1, 91E1, 93G1, 93W1] and the above mentioned references.

From our considerations of the density of states in Sects. 7.7 and 8.9 we would expect QD to have an absorption spectrum consisting of a series of δ-functions or of Lorentians, if some homogeneous broadening is included. Such behavior is not observed in reality for ensembles of dots since various processes contribute to a broadening of the resonances, e.g., the size distribution of the crystallites considered in Fig. 15.19, the coupling to optical phonons, which may be enhanced in the quantum dots as compared to the bulk material, and the influence of impurity, surface, or interface states; see [91E1, 93G1, 93W1, 95W1, 96W1] and again references of Sects. 8.13 and 9.3 and references therein.

The range of the size distribution for nanocrystallites in glass matrices is typically in the range of 10 to 30%. In organic solvents it can be reduced by size selective precipitation to a few percent [95W1]. The size distribution of quantum dots produced by nanolithography is, especially for the smallest dots, which show substantial lateral confinement effects, also in the range

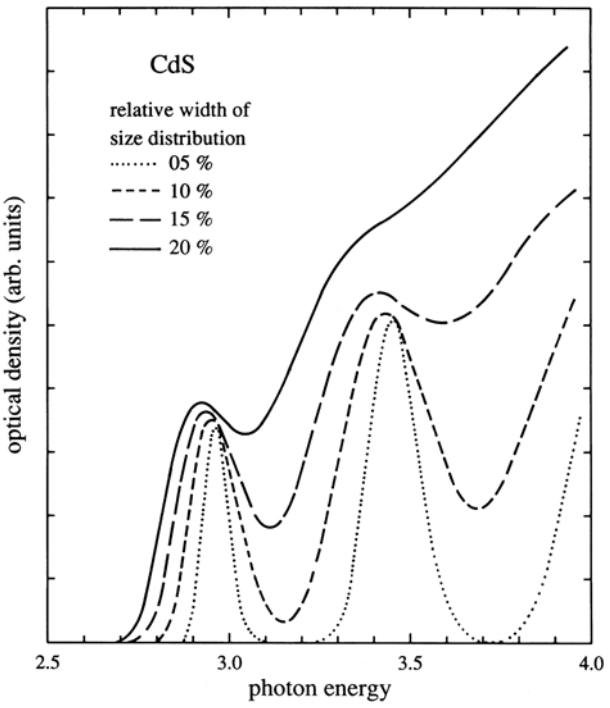


Fig. 15.19. Calculated spectra of the optical density of quantum dots for various relative widths of thier size distribution [90K1]

of $(20 \pm 10)\%$. Quantum islands in MBE-grown layers usually show strong size and composition fluctuations. Ways out of the resulting inhomogeneous broadening are single dot spectroscopy and site- or size-selective spectroscopy, including spectral hole burning.

In Fig. 15.19 we present calculated absorption spectra, with a simplified energy level scheme, showing the influence of inhomogeneous broadening. In Figs.15.16, 15.20 and 15.21 we show experimental data.

Regime I is usually realized for QD made from the copper halides CuCl, CuBr and CuI since the values of a_B for these materials are in the range of 1 nm. Figure 15.20a shows an example for CuBr which clearly illustrates the blue-shift of the $n_B = 1$ exciton resonances from the two uppermost valence bands with decreasing $\bar{R}$.

In Fig. 15.20b we show the blue-shift of the absorption edge of glasses containing CdSe QD with decreasing values of $\bar{R}$. The calculated transition energies and their oscillator strengths are given by the positions and heights of the vertical lines. For more data see also [92K1].

The luminescence is usually considerably Stokes-shifted with respect to the absorption, in Fig. 15.21 by roughly 200 meV. This shift can be partly attributed to a relaxation of the carriers into defect or surface states, and partly to stronger phonon coupling resulting in Huang-Rhys factors S^{QD} in the QD of around 1, which is larger than for the corresponding bulk materials. For $CdS_{1-x}Se_x$QD, for example, one finds

$$0.2 \approx S^{bulk} < S^{QD} \approx 1\,. \tag{15.8}$$

In Fig. 15.22a,b we show for CdSe dots in an organic solvent, the energy of the lowest electron-hole pair state as a function of the dot radius, and the energy spacing to higher states as a function of the energy of the lowest state. Good agreement between experiment and theory is observed.

Further examples for other materials are found in [97W1,01L1] of Chap. 1.

Transitions between the various electron-hole pair states in the IR under optical excitation across the band gap have been reported for CdSe in [98G1] and for ZnO dots in [00H1].

To demonstrate the relation of inhomogeneous broadening of the spectrum of a dot ensemble and to the homogeneous broadening of a single dot, we show corresponding luminescence spectra of CdSe dots capped with ZnS in Fig. 15.23. Capping with larger gap semiconductors or with organic ligands tends to reduce the homogeneous line width compared to QD in simple (boro–silikate) glass matrix. The satellite shifted by 25.6 meV is the LO phonon replica.

The absolute photo-luminescence quantum efficiency can be rather high, reaching values of 0.1 to 0.9 [95N1].

The homogeneous width $\Gamma(T)$ increases with temperature typically with a function like [99R1]

$$\Gamma(T) = \Gamma_0 + \frac{\Gamma_{\text{el–phon}}}{\exp\{\hbar\Omega_{LO}/k_B T\} - 1} \tag{15.9}$$

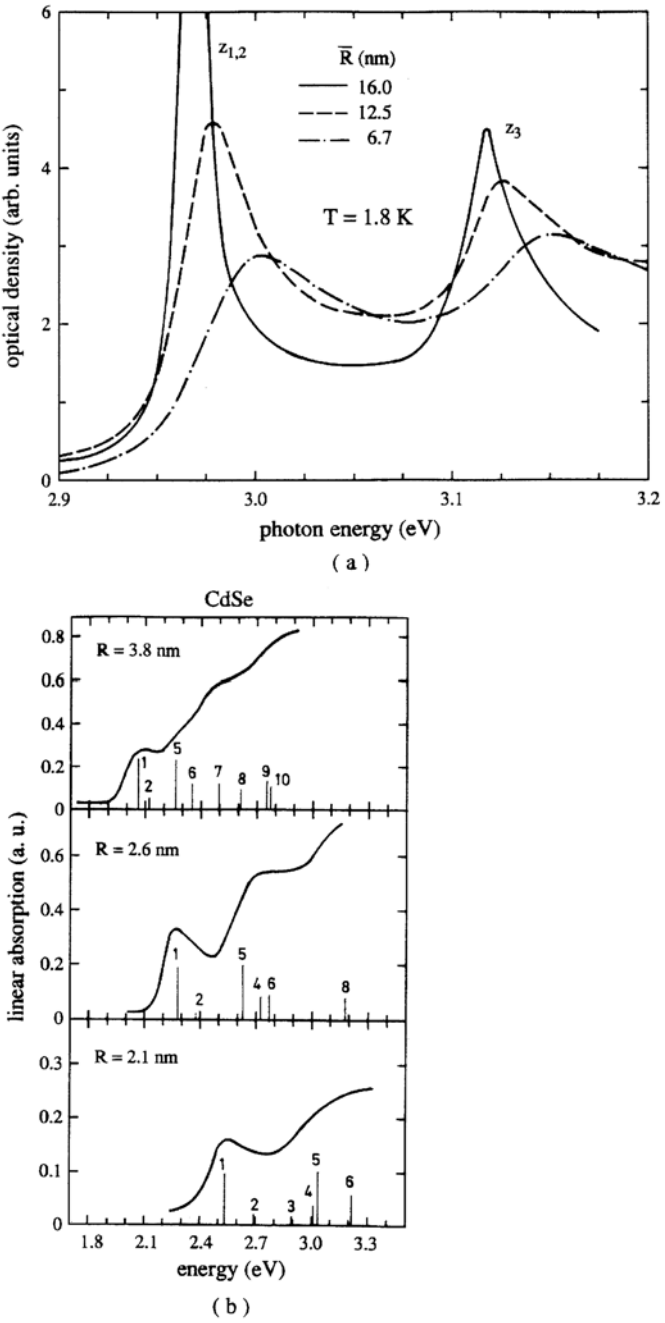


Fig. 15.20. A set of absorption spectra of CuBr quantum dots with different average radii $\bar{R}$ (**a**) and a set of measured absorption spectra together with the calculated positions of the optical transitions in CdSe QD (**b**) [88H1, 93E1]

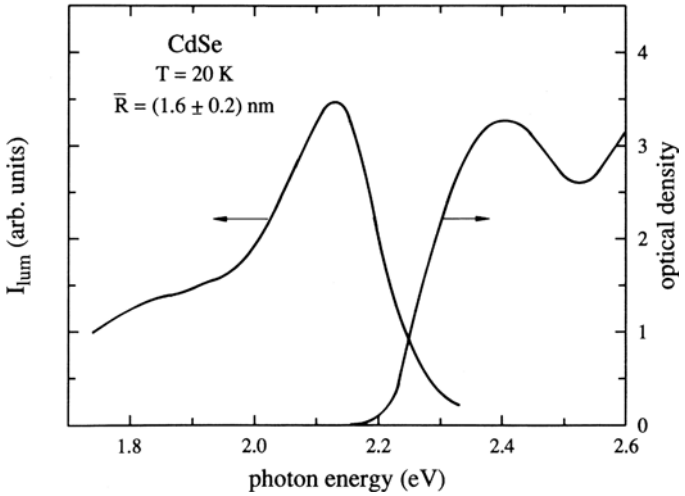


Fig. 15.21. The Stokes-shift between the absorbtion and emission spectra of $Cd_{1-x}Se_x$ quantum dots [93W2]

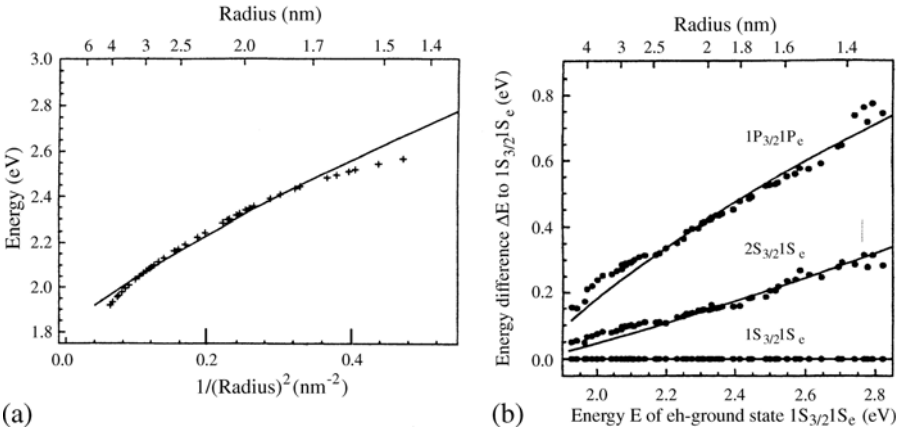


Fig. 15.22. Measured and calculated energies of the lowest electron-hole pair state in CdSe as a function of the dot radius (**a**) and the splitting of higher states as a function of this energy (**b**) [96N1]

where Γ_0 describes the low temperature limit. The second term gives the interaction with LO phonons including their thermal population probability. A term linear in T could be added, describing the interaction with (weakly confined) acoustic phonons. See also [01B1]. An alternative formula is given in [01P1]. Values of Γ_0 down to a few tens of μeV have been found for excitons in quantum dots and islands. See [02K5, 03S2, 04S1] and references therein.

Various interesting features can be observed in single dot spectroscopy: One is spectral diffusion and fluorescence intermittency or blinking [96N2,

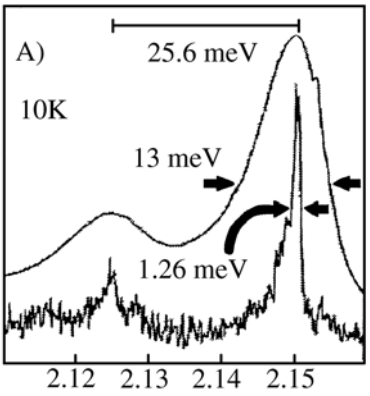


Fig. 15.23. Luminescence spectra of ZnS-capped CdSe quantum dots for an ensemble of average radius $\bar{R} = 4.5\,\mathrm{nm}$ and of a single dot [96E1]

00N1, 01S4]. Both effects are attributed to the ionization of carriers from the dot and their capture at the interface or in the surroundings of the dot. These carriers either shift the emission energy by the quantum confined Stark effect, caused by the electric field of the extra carrier, or even bring them into a non-luminescent state. Removal of the carrier by optical or thermal activation brings the QD back to the original state.

For the mapping of the excitonic wave function in quantum dots by optical near field microscopy see, e.g., [01C1, 03M3, 04K1].

Another interesting phenomenon is the fact that a single dot can emit only one photon at a time. It then has to be re-excited again before it can emit another photon. As a consequence, the photons emitted from a single dot are anti-correlated. This photon anti-bunching shows up in a dip in the correlation function of the photons for zero time delay. Examples are found, e.g., in [00L3, 00M3, 01M1, 02S1, 03U1].

A related topic is quantum entanglement of excitons in single or coupled quantum dots [98B1, 99Q1, 00C1, 00R1, 01B1, 03Z2].

The exciton or electron–hole pair ground state usually shows a fine-structure, due to the exchange splitting. The lower state is the triplet, in quantum dots usually called the dark state, since it is spin flip forbidden. This ordering of the states is the same since it already exists in bulk materials, with the difference that in bulk materials the splitting between the triplet and the transverse dipole allowed singlet states of Wannier excitons is much smaller than the transverse-longitudinal splitting of the singlet. In quantum dots the situation is inverted due to a strong enhancement of the (short-range) exchange interaction caused by three-dimensional confinement, while the oscillator strength is reduced by the dilution of the oscillators, i.e., the dots in their matrix. See also Fig. 9.2.

In Fig. 15.24 we show data from various authors for the exchange splitting of the excitons in CdSe quantum dots as a function of the dot radius.

Similar data are found for Cu and Ag halogenid dots [00G2, 03G1] where bulk band parameters have also been deduced from dot spectroscopy.

The interaction of excitons (i.e., electron–hole pair states) in quantum dots with phonons has already been addressed above in connection with the homogeneous width. It also shows up in luminescence. An example is given in Fig. 15.23. Typically one finds that the spacing of the zero phonon line and its one LO phonon replica is slightly smaller than the LO phonon energy in bulk at $\boldsymbol{k} = 0$. This effect is caused by the facts that the LO phonon mode confined in a quantum dot has an effective $\boldsymbol{k} > 0$ and that the dispersion of LO phonons is generally negative, i.e., $E_{\mathrm{LO}}(\boldsymbol{k})$ decreases with increasing $\boldsymbol{k}$. See, e.g., [96E2]. Furthermore one observes surface phonon modes [99H1] and confined acoustic phonons [96W1, 01P1]. For exciton-phonon complexes or exciton-polarons see [04K2] and references therein.

The Huang-Rhys factor, which is a measure for the coupling between an electronic resonance and LO phonons can be deduced from luminescence spectra, in contrast to bulk materials, because problems with reabsorption and $\boldsymbol{k}$ selection rules in the zero phonon line are relaxed, which are relevant in the latter case. The values of the Huang–Rhys factor depend partly on the individual properties of the dots like capping or surface states and have values between 0.2 and 0.9. For data see [01L1] of Chap. 1 or [93S1] or the discussion with (15.8).

The interaction of excitons with single carriers or other excitons will be considered in connection with trions and biexcitons in Sect. 20.3.

Quantum dots, especially from the group of II–VI semiconductors, have been used as markers attached to organic molecules [98B3, 01D3, 02G1] to impregnate the voids in photonic crystals [99R2, 02P1] or for incorporation in microcavities [03M3].

It is well known, that the vapor pressure of small droplets increases with decreasing droplet radius. Consequently their boiling point or temperature decreases. A similar phenomenon also occurs for the solid to liquid phase

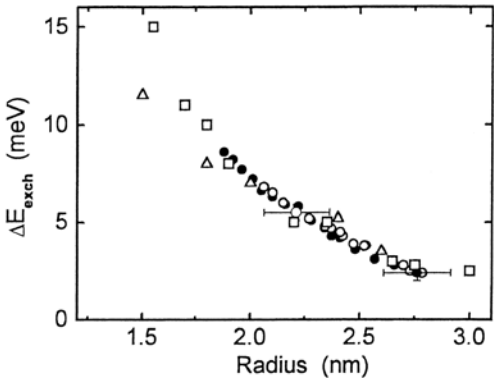


Fig. 15.24. Exchange splitting as a function of the exciton ground state in CdSe dots as a function of the dot radius (data from various authors, [96W1])

transition of small crystallites. This effect has been known and used since ancient times for metals like Au, e.g., in granulation techniques. Recently this phenomenon has also been observed for CdS(e), Si or Cu-halide semiconductor quantum dots [92G1, 96G1, 99V1, 01Z1, 03G1].

In Fig. 15.25 we show examples for two glass samples doped with CuCl crystals of different average radii. The optical density in the lowest exciton resonance (compare with Fig. 15.20) is plotted as a function of sample temperature. The steep decrease (increase) is attributed to the melting (solidification) of the CuCl dots. Note that these temperatures are below the glass transition temperature of the host. One observes two prominent features: a hysteresis in the temperature cycle, as indicated by the arrows, which is characteristic of a first order phase transition, and a shift of this hysteresis loop to temperatures below the bulk melting temperature T_M of CuCl, which is indicated by the vertical arrow. This shift increases with decreasing dot radius as predicted above.

Quantum dots or islands have also been made from indirect gap semiconductors, with an emphasis on Si, which is the dominant material for electronic devices. The hope is that the relaxation of the $\boldsymbol{k}$-conservation under three-dimensional confinement might increase the luminescence yield of this material, allowing one to integrate light emitters with other (opto) electronic devices based on the same material (Si) on one chip. However, at the end of 2004 this aim had only been achieved with limited success for Si based quantum wells or superlattices.

Various approaches have been used to prepare Si nanocrystals, which are reviewed in, e.g., Chap. 9 and [97W1] of Chap. 1. For more recent data see, e.g., [94K2, 94P1, 96G3, 96M1, 98P2, 99K1, 01Y1, 02Y1, 02Z1] and references therein.

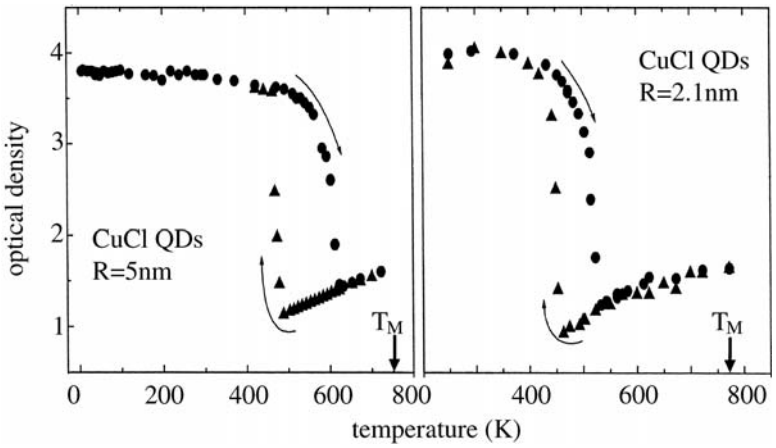


Fig. 15.25. Optical density in the exciton resonance of CuCl quantum dots in glass matrices as function of temperature for two different average dot radii [03G1]

A special type of Si nanocrystals is porous silicon (p-Si). In 1984 a visible, low temperature emission was observed in p-Si [84P1, 90C3, 97C1]. Since then more than 1000 papers have appeared on this topic, which is still being controversially discussed. Therefore we give here only a very short overview and a few references.

Porous silicon is prepared by anodic oxidation of doped Si, e.g., in HF-ethanol electrolytes. The result is a porous layer on the Si chip, which contains nanocrystals and nanorods with dimensions of a few to several nm. These samples show (photo) luminescence in the visible spectrum from the red-orange to the blue region depending on the preparation and excitation conditions.

One idea to explain the blue shift and the enhanced intensity of the emission compared to bulk Si is by the above mentioned quantum confinement effect. However, alternative explanations have been put forward, attributing the origin of the emission to siloxene $(Si_6O_3H_6)_n$ or to Si-oxihydrides or suboxides and other Si-related compounds, which may form during the etching process in the porous layer, partly even in the form of surface states or traps at the nanocrystals. On the other hand, the siloxene usually shows luminescence bands comparable to those of p-Si only after a certain heat treatment, during which Si nanocrystals might form.

There seems to be a trend to assume that the blue-shifted absorption of p-Si and parts of its red-orange luminescence including the polarization anisotropy might be related to Si nanocrystals or nanorods, while the shorter wavelength emission components might have their origin in the alternative compounds mentioned above. We cannot currently provide a definite answer to this problem. Instead we give some references for both approaches and leave the final decision to the reader or to future research [92S2, 95S2, 97G1, 97R2, 97R3, 99B1, 99J1, 01D1, 02K4].

As already mentioned in the introduction we presently see increasing work to pack quantum dots so closely that the wave functions start to overlap. According to what we have learned for periodic potentials in general and more specifically for superlattices in Sects. 4.1, 5.4, and in Chap. 8 this interaction will result in the formation of (mini)bands from the discrete energy levels in isolated dots. If the dots are periodically arranged and if their size distribution and thus their inhomogeneous level broadening are sufficiently small, these new bands can be described by a dispersion relation $E(\boldsymbol{k})$. In the opposite case, there will be no good quantum number $\boldsymbol{k}$, but there will be bands in the density of states, similarly to disordered systems. The first examples of this development can be found in the following small selection of references: [92T2, 95L3, 96K1, 98M1, 01D2, 00L4].

A similar phenomenon is found if fullerene molecules C_{60}, are put together to form a solid. Solid C_{60} crystals are semiconductors with a band gap of about 2.2 eV [99M1].

To finish this section on quantum dots we want to give some references for quantum islands or self-assembled quantum dots occurring in thin MBE layers for lattice-mismatched systems. A calculation of the wave functions for

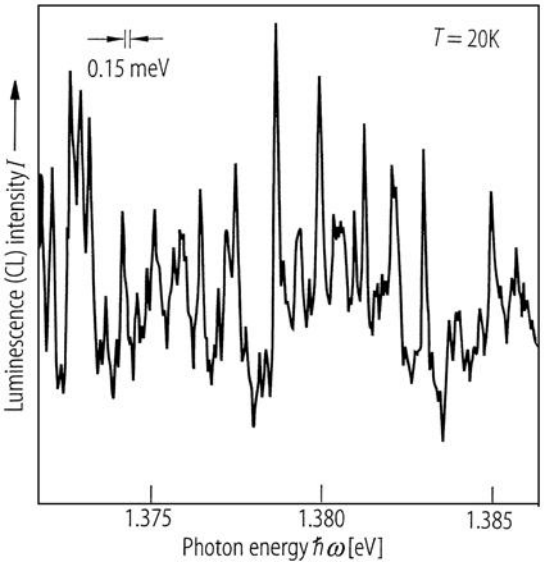


Fig. 15.26. Cathodo-luminescence spectrum of an InAs/GaAs sample containing self-assembled islands [95G3]

idealized, pyramid shaped islands can be found in [95G2]. Some recent reviews are, e.g., [01K2, 01K3, 02K2, 02K5, 03S2, 04S1] or [01L1] of Chap. 1. We give two examples of the optical spectra of these quantum islands. In Fig. 15.26 we show a cathodo-luminescence spectrum of an ensemble of In(Ga)As islands embedded in a InGaAs well sandwiched between GaAs barriers. The nominal deposited InAs layer thickness was 0.6 nm.

The narrow emission lines arise from various islands or localization sites. The width of individual lines is ≤ 0.15 meV and does not change significantly with temperature up to 70 K, giving thus some evidence for the immobility of the state and its isolation from the surroundings. Compare, e.g., Figs. 15.5 and 15.23. For more details on InAs-based islands see [99B2].

As a second example we show in Fig. 15.27 the photoluminescent spectra of the ground state luminescence of three different CdSe/ZnSe quantum islands.

The energy of the emission line is set to zero. The sharp features in the PLE spectra around 40 meV are considered as excited states while the wing above zero and the broad peak at 25 meV are considered as acoustic and LO phonon sidebands, respectively. See also Sects. 14.1 and 14.4 or Figs. 14.3 and 14.5.

The exciton states in the quantum islands thus show analogies both to localized excitons in quantum wells and alloys and even to bound exciton complexes, but also to the strongly confined states in quantum dots in glass and organic matrices. One has narrow homogeneous widths of the states,

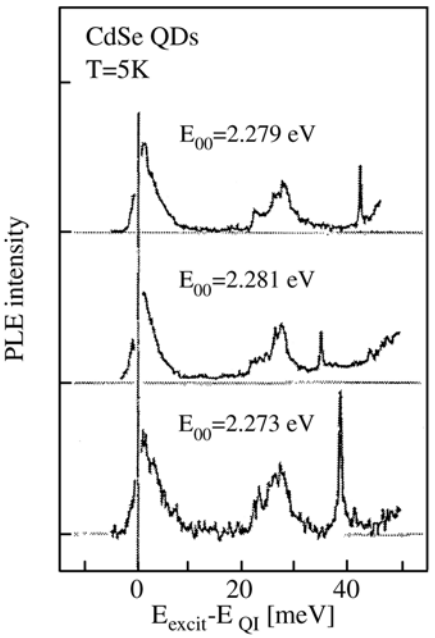


Fig. 15.27. PLE spectra of three different Cd(Zn)Se/ZnSe islands [01H3]

i.e., long dephasing times (see Chap. 23) and a confinement-enhanced singlet-triplet splitting as in QD doped glasses and organic matrices (Fig. 15.24). See, e.g., [98G2, 01H1] for the system CdS/ZnSe. One can also observe spectral diffusion and intermittency [98Z1, 00S2] and one tries to build up higher dimensional structures by coupling quantum islands, e.g., in vertical stacks or even in three dimensional ordered structures [99Y1, 01H2, 01K3, 01P1, 03K2].

For examples of level repulsion between spatially and energetically close lying states see, e.g., [03G1, 03W1] and references therein.

With these examples and the following problems we finish the chapter on excitonic optical properties of semiconductors quantum structures. We shall meet reduced dimensionalities again including "photonic dots" in the next chapters.

15.5 Problems

1. Try to make a simple estimate of the radius of an ideal quantum dot at which the quantization energy and the Coulomb energy are equal.
2. In a quantum well with infinitely high barriers one expects an energy level scheme that depends quadratically on the quantum number n_z. Check if a similar relation holds for the (few) bound exciton states in wells with finite barriers. Inspect, e.g., data for the system GaAs/AlGaAs.

3. Make a qualitative sketch of the conduction and valence band edges of an ideal, undoped, type I quantum well, of a barrier or a well-doped sample or of a type II MQW with a slight homogeneous n-doping.
4. Make a rough sketch of the absorption spectrum for the band-to-band transition with and without Coulomb interaction of a direct gap, dipole allowed bulk semiconductor and for quantum wells, wires and dots.
5. Calculate the first few energy levels and wave functions in a spherical and a cubic dot with infinitely high barriers.

References to Chap. 15

[65C1] M. Cardona and G. Harbeke, Phys. Rev. A **137**, 1467 (1965)
[65C2] M. Cardona, M. Weinstein and G.A. Wolff, Phys. Rev. A **140**, 633 (1965)
[80P1] J.J. Le Pore, J. Appl. Phys. **51**, 6441 (1980)
[82E1] Al.L. Efros and A.L. Efros, Sov. Phys. Semicond. **16**, 772 (1982)
[82H1] A. Henglein et al., Ber. Bunsen, Phys. Chem. **86**, 301 (1982)
[83G1] A.C. Gossard, Thin Films, Preparation and Properties, Academic Press, New York (1983)
[83J1] R. Jain and R. Lind, J. Opt. Soc. Amer. **73**, 647 (1983)
[84P1] C. Pickering et al., J. Phys. C **17**, 6535 (1984)
[86B1] L. Brus, J. Chem. Phys. **90**, 2555 (1986) and IEEE J. QE **22**, 1909 (1986)
[86D1] G.H. Doehler, IEEE J. QE-**22**, 1682 (1986)
[86S1] H.M. Schmidt and H. Weller, Chem. Phys. Lett. **129**, 615 (1986)
[87S1] S. Schmitt-Rink, D.A.B. Miller and D.S. Miller, Phys. Rev. B **35**, 8113 (1987)
[88H1] Y. Hirugama et al., Phys. Rev. B **37**, 2774 (1988)
[88H2] F. Henneberger et al., Appl. Phys. B **46**, 19 (1988)
[88K1] M. Kohl et al., Phys. Rev. B **37**, 10927 (1998)
[88U1] I.N. Uraltsev et al., phys. stat. sol. (b) **150**, 673 (1988)
[89K1] M. Kohl et al., Superlattices and Microstructures **5**, 235 (1989)
[89O1] M. Olszabier et al., Phys. Rev. Lett. **62**, 2997 (1989), Surf. Science **228**, 123 (1990) and J. Luminesc. **45**, 186 (1990)
[89S1] S. Schmitt-Rink, D.S. Chemla and D.A.B. Miller, Adv. Phys. **38**, 89 (1989)
[90A1] L.C. Andreani and F. Bassani, Phys. Rev. B **41**, 7536 (1990)
[90C1] R. Cingolani and K. Ploog, Adv. Phys. **40**, 535 (1990)
[90C2] R. Cingolani, L. Tapfer and K. Ploog, Appl. Phys. Lett. **56**, 1233 (1990)
[90C3] L.T. Canham, Appl. Phys. Lett. **57**, 1046 (1990)
[90G1] E.O. Göbel and K. Ploog, Prog. Quantum Electron. **14**, 289 (1990)
[90I1] E.L. Ivchenko et al., phys. stat. sol. (b) **161**, 219 (1990) and Phys. Sol. State **36**, 1156 (1994)
[90K1] S.W. Koch, Phys. Bl. **46**, 167 (1990)
[90M1] U. Merkt, Festkörperprobleme/Adv. Solid State Phys. **30**, 77 (1990)
[90O1] K. Ogawa, T. Katsuyama and H. Nakamura, Phys. Rev. Lett. **64**, 796 (1990)
[90T1] F. Tassone, F. Bassani and L. Andreani, Il Nuovo Cimento **12D**, 1673 (1990)

[91E1] A.I. Ekimov and Al.L. Efros, Acta Phys. Polnica A **79**, 5 (1991) and A.L. Ekimov, Physica Scripta **39**, 217 (1991)
[91L1] V. Lehmann and U. Goesele, Appl. Phys. Lett. **58**, 856 (1991)
[91T1] H. Tuffigo et al., Phys. Rev. B **43**, 14629 (1991)
[92A1] L.C. Andreani and F. Bassani, Phys. Rev. B **45**, 6023 (1992)
[92G1] A.N. Goldstein et al., Science **256**, 1425 (1992)
[92K1] S.W. Koch et al., J. Crystal Growth **117**, 592 (1992)
[92M1] S. Malzer et al., phys. stat. sol. (b) **173**, 459 (1992)
[92N1] G.A. Northrop et al., Appl. Phys. Lett. **60**, 865 (1992)
[92O1] D. Oberhauser, Dissertation, Kaiserslautern (1992)
[92O2] K.P. O'Donnel et al., J. Cryst. Growth **117**, 497 (1992)
[92P1] Phyiscs of Nanostructures Proc. 38th Scottish Universities Summer School in Physics, St. Andrews (1991), SUSSP, Edinburg (1992)
[92S1] W. Sack et al., J. Lumin. **53**, 409 (1992)
[92S2] M. Stutzmann et al., Festkörperprobleme/Advances in Solid State Physics **32**, 179 (1992)
[92T1] F. Tassone and F. Bassani, Il Nuovo Cimento **14D**, 1241 (1992)
[92T2] F. Tassone, F. Bassani and L. Andreani, Phys. Rev. B **45**, 6023 (1992)
[92T3] T. Takagahra, Surf. Science **267**, 310 (1992)
[92V1] L. Vescan et al., Thin Solid Films **222**, 5 (1992)
[92W1] M. Wachter et al., Thin Solid Films **222**, 10 (1992)
[93A1] G. Abstreiter, Physica Scripta T **49**, A + B, 42 (1993)
[93C1] J. Cibert et al., Phys. Scr. T **49**, B, 487 (1993)
[93C2] D.S. Citrin, Phys. Rev. B **48**, 2535 (1993)
[93C3] R. Cingolani, Phys. Scripta T **494**, 470 (1993)
[93D1] F. Daiminger et al., in Proc. 21st Int'l Conf. Phys. Semicond. p 1293, ed. by P. Jiang, H.Z. Zheng, World Scientific, Singapore (1993)
[93E1] A.I. Ekimov et al., J. Opt. Soc. Am. B **10**, 100 (1993)
[93G1] S.V. Gaponenko et al., J. Opt. Soc. Am. B **10**, 1947 (1993)
[93H1] R. Hellmann et al., in Proc. 21st Int'l Conf. Phys. Semicond. p 1008, ed. by P. Jiang, H.Z. Zheng, World Scientific, Singapore (1993)
[93J1] S. Jorda, U. Rössler and D. Broido, Phys. Rev. B **48**, 1669 (1993)
[93J2] S. Jorda, Solid State Commun. **87**, 439 (1993)
[93K1] H. Kalt et al., Physica B **191**, 90 (1993)
[93K2] T. Katsuyama et al., Semicond. Science and Tech. **8**, 1226 (1993)
[93L1] N. Linder et al., Appl. Phys. Lett. **62**, 1916 (1993)
[93M1] U. Menczigar et al., Phys. Rev. B **47**, 4099 (1993)
[93P1] K.G. Ploog, R. Nötzel and O. Brandt, in Proc. 21st Int'l Conf. Phys. Semicond. p 1297, ed. by P. Jiang, H.Z. Zheng, World Scientific, Singapore (1993)
[93S1] J. J. Shiang, S. H. Risbud and A. P. Alivisatos, J. Chem. Phys. **98**, 8432 (1993)
[93V1] P. Vogl, Physica Scripta T **49**, A + B, 476 (1993)
[93W1] U. Woggon et al., Phys. Rev. B **47**, 3684 (1993)
[93W2] O. Wind et al., Adv. Mat. Opt. El. **3**, 89 (1993)
[93Y1] D.R. Yakovlev et al., in Proc. 21st Int'l Conf. Phys. Semicond. p 1136, ed. by P. Jiang, H.Z. Zheng, World Scientific, Singapore (1993)
[93Y2] F. Yang et al., Phys. Rev. Lett. **70**, 323 (1993)
[94C1] D.S. Citrin, Solid State Commun. **89**, 139 (1994) and Phys. Rev. B **51**, 14361 (1995)

[94G1] M. Gurioli et al., Solid State Commun. **91**, 931 (1994)
[94G2] M. Grün et al., Superlattices and Microstruct. **15**, 463 (1994)
[94H1] V. Higgs et al., Appl. Phys. Lett. **64**, 607 (1994)
[94J1] C. Janowitz et al., Phys. Rev. B **50**, 2181 (1994)
[94K1] C. Klingshirn et al., J. Crystal Grwoth **138**, 191 (1994)
[94K2] Y. Kanemitsu et al., J. Luminesc. **60/61**, 337 (1994)
[94L1] C. Luo et al., Superlattices and Microstructures **15**, 221 (1994)
[94L2] W. Langbein et al., Appl. Phys. Lett. **65**, 2465 (1994)
[94P1] Porous Silicon, Z.C. Feng and R. Tsu (eds.), World Scientific, Singapor (1994)
[95A1] L.C. Andreani, phys. stat. sol. (b) **188**, 29 (1995)
[95B1] F. Bassani, G. Cajkowski and Trediucci, Z. Physik B **98**, 41 (1995) and references therein
[95B2] L. Banyai et al., Phys. Rev. Lett. **75**, 2188 (1995)
[95C1] D.S. Citrin, phys. stat. sol. (b) **188**, 43 (1995)
[95F1] S. Fakatsu, J. Cryst. Growth **157**, 1 (1995)
[95G1] A. Gustafsson et al., Appl. Phys. Lett. **67**, 3673 (1995)
[95G2] M. Grundmann, Festkörperprobleme/Advances in Solid State Physics **35**, 123 (1995)
[95G3] M. Grundmann et al., Phys. Rev. Lett. **74**, 4043 (1995)
[95L1] H. Lee et al., J. Appl. Phys. **78**, 6327 (1995)
[95L2] W. Langbein, M. Hetterich and C. Klingshirn, Phys. Rev. B **51**, 9922 (1995)
[95L3] R. Leon et al., Science **267**, 1966 (1995)
[95N1] M. Nirmal et al., Phys. Rev. Lett. **75**, 3728 (1995)
[95P1] Proc. of the Intern. Conf. on Physics, Chemistry and Application of Nanostructures, V.E. Borisenko, A.B. Filanov, S.V. Gaponenko and V.S. Gurin (eds.), Minsk, Belarus (1995)
[95S1] Y.P. Song et al., Semicond. Science and Technol. **10**, 1404 (1995)
[95S2] M. Stutzmann, phys. stat. sol. (b) **192**, 273 (1995)
[95W1] U. Woggon and S.V. Gaponenko, phys. stat. sol. (b) **189**, 285 (1995)
[95Z1] A. Zrenner et al., Phys. Rev. B **52**, 16608 (1995)
[96A1] A.P. Alivisatos, Science **271**, 933 (1996) and J. Phys. Chem. **100**, 13226 (1996)
[96C1] J. Cerne, Phys, Rev. Lett. **77**, 1131 (1996)
[96E1] S.A. Empedodes, D.J. Norris and M.G. Bawendi, Phys. Rev. Lett. **77**, 3873 (1996)
[96E2] A. Ekimov, J. Luminesc. **70**, 1 (1996)
[96G1] A.A. Guzelian et al., Appl. Phys. Lett. **69**, 1432 (1996)
[96G2] A.N. Goldstein, Appl. Phys. A **62**, 33 (1996)
[96G3] S.V. Gaponenko et al., J. Lumines. **70**, 364 (1996)
[96K1] C.R. Kargan, Phys. Rev. Lett. **76**, 1517 (1996) and Phys. Rev. B **54**, 8633 (1996)
[96L1] W. Langbein et al., Phys. Rev. B **54**, 10784 (1996)
[96L2] X. Lin et al., J. Phys. Cond. Matter **8**, 3947 (1996)
[96L3] W. Langbein, H. Gislason and J.M. Hvam, Phys. Rev. B **54**, 14595 (1996)
[96L4] Ch. Lienau, A. Richter and T. Elsässer, Appl. Phys. Lett. **69**, 325 (1996)
[96M1] K.S. Min et al., Appl. Phys. Lett. **68**, 2511 (1996)
[96N1] D.J. Norris and M.G. Bawendi, Phys. Rev. B **53**, 16338 (1996)

[96N2] M. Nirmal et al., Nature **383**, 802 (1996)
[96O1] L.K. Orlov, J. Appl. Phys. **80**, 415 (1996)
[96S1] V. Sirinivas, Y.J. Chen and C.E.C. Wood, J. Opt. Soc. Am. **13**, 989 (1996)
[96S2] T. Stroucken et al., Phys. Rev. B **53**, 2026 (1996)
[96S3] M.S. Salib et al., Phys. Rev. Lett. **77**, 1135 (1996)
[96W1] U. Woggon et al., Phys. Rev. B **54**, 1506 (1996)
[96W2] D. Weber et al., J. Opt. Soc. Am. B **13**, 1241 (1996)
[96W3] M.U. Wehner et al., J. Opt. Soc. Am. B **13**, 977 (1996)
[96W4] M.U. Wehner, D. Steinbach and M. Wegener, Phys. Rev. B **54**, R5211 (1996)
[97B1] U. Banin et al., Phys. Rev. B **55**, 7059 (1997)
[97C1] R.T. Collins, P.M. Fauchet and M.A. Tischler, Physics Today, p **24**, Jan. (1997)
[97G1] J.L. Gole et al., Phys. Rev. B **56**, 2137 (1997)
[97K1] S. Kim et al., Phys. Rev. B **55**, 7130 (1997)
[97K2] C. Klingshirn, in [81A1]h of Chap. 1, p 85
[97M1] Y. Masumoto and K. Sonobe, Phys. Rev. B **56**, 9734 (1997)
[97P1] M.F. Pereira et al., phys. stat. sol. (a) **164**, 199 (1997)
[97R1] A. Richter et al., Phys. Rev. Lett. **79**, 2145 (1997)
[97R2] M. Rosenbauer and M. Stutzmann, J. Appl. Phys. **82**, 4520 (1997)
[97R3] M. Rosenbauer et al., Phys. Rev. B **55**, 10117 (1997)
[97S1] V. Sperling et al., J. Luminesc. **72-74**, 395 (1997)
[98B1] H. Bonadeo et al., Science **282**, 1473 (1998)
[98B2] U. Banin et al., J. Chem. Phys. **109**, 2306 (1998)
[98B3] M. Bayer, S. Walck and T.L. Reinecke, Phys. Rev. B **57**, 6584 (1998)
[98B4] M. Bruches, Science, **281**, 2013 (1998)
[98C1] C.P. Collier, T. Vossmeyer and J.R. Heath, Annu. Rev. Phys. Chem. **49**, 371 (1998)
[98G1] P. Guyot-Sinnoest and M.A. Hines, Appl. Phys. Lett. **72**, 680 (1998)
[98G2] F. Gindele et al., Solid State Commun. **106**, 653 (1998)
[98H1] A. Hartmann et al., Appl. Phys. Lett. **73**, 2322 (1998)
[98K1] H. Kalt et al., J. Crystal Growth **184/185**, 795 (1998)
[98K2] C. Klingshirn, NATO ASI Series B **372**, 143 (1998)
[98L1] Ch. Lienau et al., Phys. Rev. B **58**, 2045 (1998)
[98M1] O.I. Micic et al., J. Chem. Phys. B 102, 9791 (1998)
[98P1] W. Petri et al., J. Crystal Growth **184/185**, 320 (1998)
[98P2] I. Pelant et al., phys. stat. sol. (a) **165**, 25 (1998)
[98R1] A. Richter et al., Appl. Phys. Lett. **73**, 2176 (1998)
[98V1] F. Vouilloz et al., Physica E **2**, 862 (1998)
[98Z1] B.P. Zhang et al., Appl. Phys. Lett. **73**, 1266 (1998)
[99B1] M.S. Brandt et al., phys. stat. sol. (b) **215**, 409 (1999)
[99B2] D. Bimberg, M. Grundmann and N.N. Ledentsov, Quantum Dot Heterostructures, John Wiley and Sons, Chichester (1999)
[99C1] D. Crisinel, M.A. Dupertius and E. Kapon, Optical and Quantum Electronics **31**, 797 (1999)
[99D1] A. Dinger et al., Semicond. Science and Technol. **14**, 595 (1999)
[99F1] C.J. Fall, M.A. Dupertius and E. Kapon, Optical and Quantum Electr. **31**, 201 (1999)
[99H1] Y. N. Huang, Phys. Rev. B **59** 7285 (1999)

[99J1] R. Janssen, Phys. Rev. B **60**, 13561 (1999)
[99K1] D. Kovalev et al., phys. stat. sol. (b) **215**, 871 (1999)
[99M1] K.P. Medletov and V.D. Negrii, phys. stat. sol. (b) **211**, 217 (1999)
[99O1] A. Ohtomo et al., Appl. Phys. Lett. **75**, 980 (1999)
[99Q1] L. Quiroga and N.F. Johnson, Phys. Rev. Lett. **83**, 2270 (1999)
[99R1] L.M. Robinson et al., Phys. Rev. Lett. **73**, 2797 (1999)
[99R2] A. L. Rochachev et al., J. Phys. Cem. B **103**, 3065 (1999)
[99V1] P.M. Valov and V.I. Leiman, Phys. of the Solid State **41**, 278 (1999)
[99Y1] P. Yu et al., Phys. Rev. B **60**, 16680 (1999)
[00B1] M. Bayer et al., Phys. Rev. B **61**, 7273 (2000)
[00C1] G. Chen et al., Science **289**, 906 (2000)
[00D1] M.A. Dupertius et al., Europhys. Lett. **52**, 420 (2000)
[00D2] A. Dinger et al., J. Crystal Growth **214/215**, 660 (2000)
[00F1] P.W. Fry et al., Phys. Rev. Lett. **84**, 733 (2000)
[00F2] C.E. Finlayson, D.S. Ginger and N.C. Greenham, Appl. Phys. Lett. **77**, 2500 (2000)
[00G1] S. Ghlosh, J. Weber and H. Presting, Phys. Rev. B **61**, 15625 (2000)
[00G2] O. Gogolin, Phys. Rev. B, **62**, 13053 (2000)
[00H1] H. Hartmann et al., Phys. Rev. Lett. **84**, 5648 (2000)
[00K1] C. Konstantin et al., J. Appl. Phys. **88**, 141 (2000)
[00K2] E. Kapon, in Frontiers of Nano-Optoelectronic Systems, e.d. by L. Pavesi and E. Buzaneva, p 41, Kluver, Dordrecht (2000)
[00L1] K. Leosson et al., Phys. Rev. B **61**, 10322 (2000)
[00L2] K. Leifer et al., Appl. Phys. Lett. **77**, 3923 (2000)
[00L3] B. Lounis et al., Chem. Phys. Lett. **329**, 399 (2000)
[00L4] C.A. Leatherdale et al., Phys. Rev. B **62**, 2669 (2000)
[00M1] T. Makino et al., Appl. Phys. Lett. **77**, 975 (2000)
[00M2] G. Malpuech et al., Phys. Rev. Lett. **85**, 650 (2000)
[00M3] P. Miller et al., Science **290**, 2282 (2000) and Nature **406**, 968 (2000)
[00N1] R.G. Neuhauser et al., Phys. Rev. Lett. **85**, 3301 (2000)
[00P1] C. Penn et al., Thin Solid Films **369**, 394 (2000)
[00R1] J.H. Reina, L. Qurioga and N.F. Johnson, Phys. Rev. A **62**, 012305 (2000)
[00S1] M. Schmidt et al., Appl. Phys. Lett. **77**, 85 (2000)
[00S2] J. Seufert et al., Appl. Phys. Lett. **76**, 1872 (2000)
[00Y1] D.R. Yakovlev et al., Phys. Rev. B **61**, 2421 (2000)
[01A1] M.V. Artemyev et al., Appl. Phys. Lett. **78**, 1032 (2001) and Nano Lett. **1**, 309 (2001)
[01B1] M. Bayer et al., Science **291**, 451 (2001)
[01B2] L. Besombes et al., Phys. Rev. B **63**, 155307 (2001)
[01C1] A. Crottini et al., Phys. Rev. B **63**, 121313 (2001)
[01D1] J. Diener, SPIE Proc. **4355**, 137 (2001)
[01D2] H. Döllefeld, H. Weller and A. Eychmüller, Nano Letters **1**, 267 (2001)
[01D3] M. Dahan et al., Optics Lett. **26**, 825 (2001)
[01H1] K. Hild et al., phys. stat. sol. (b) **224**, 379 (2001)
[01H2] J. Hu et al., Science, **292**, 2060 (2001)
[01H3] A. Hundt et al., phys. stat. sol. (b) **224**, 159 (2001)
[01H4] S.-K. Hong et al., Appl. Phys. Lett. **78**, 3349 (2001)
[01K1] A.V. Kakovin et al., Solid State Commun. **120**, 259 (2001)
[01K2] I.L. Krestnikov et al., phys. stat. sol. (a) **183**, 207 (2001)

[01K3] E. Kurtz et al., J. Crystal Growth **227/228**, 630 (2001)
[01M1] E. Moreau et al., Appl. Phys. Lett. **79**, 2865 (2001)
[01M2] G. Messin et al., Optics Lett. **26**, 1891 (2001)
[01P1] P. Palinginis and H. Wang, Appl. Phys. Lett. **78**, 1541 (2001)
[01P2] T. Passow, K. Leonardi and D. Hommel, phys. stat. sol. (b) **224**, 143 (2001)
[01S1] H.D. Sun et al., Appl. Phys. Lett. **78**, 2464 (2001)
[01S2] M.H. Szymanska, P.B. Littlewood and R.J. Needs, Phys. Rev. B **63**, 205317 (2001)
[01S3] M. Schmidt et al., phys. stat. sol. (b) **229**, 643 (2001)
[01S4] K.T. Shimizu et al., Phys. Rev. B **63**, 205316 (2001)
[01T1] S.A. Tarasenko et al., Semicond. Science and Technol. **16**, 486 (2001)
[01W1] S. Wachter et al., phys. stat. sol. (b) **224**, 437 (2001)
[01Y1] A.I. Yakimov et al., Phys. Rev. B **63**, 045312 (2001)
[01Z1] Z. Zhang, M. Zhao and Q. Jiang, Semicond. Science and Technol. **16**, L33 (2001)
[02E1] H.J. Eisler, Appl. Phys. Lett. **80**, 4614 (2002)
[02G1] D. Gerion, J. Am. Chem. Soc., **124**, 7070 (2002)
[02H1] G.R. Hayes and B. Devaud, phys. stat. sol. (a) **190**, 637 (2002)
[02K1] E. Kapon, NATO Sciences Series II **90**, 256 (2002)
[02K2] E. Kurtz et al., NATO Science Series II **90** 633 (2002)
[02K3] D. Katz et al., Phys. Rev. Lett. **89**, 086801 (2002)
[02K4] D. Kovalev et al., J. Appl. Phys. **91**, 4131 (2002)
[02K5] I.L. Krestnikov et al., Phys. Rev. B **66**, 155310 (2002)
[02P1] W. Park, phys. stat. sol. (b) **229**, 949 (2002)
[02R1] A. Reznitsky et al., phys. stat. sol. (c) **0**, No 1, 280 (2002)
[02S1] Ch. Santori et al., Nature **419**, 594 (2002) and Phys. Rev. B **66**, 045308 (2002)
[02Y1] L.X. Yi et al., Appl. Phys. Lett. **81**, 4248 (2002)
[02Z1] M. Zacharias et al., Appl. Phys. Lett. **80**, 661 (2002)
[03G1] O. Gogolin et al., J. Luminesc. **102-103**, 414 and 451 (2003)
[03K1] A.M. Kapitanov et al., phys. stat. sol. (b) **238**, 317 (2003)
[03K2] E. Kurtz et al., in [81A1]k of Chap. 1, p 633
[03K3] G. Kocherscheidt, W. Langbein and V. Savona, phys. stat. sol. (b) **238**, 486 (2003)
[03M1] G. Mannarini et al., phys. stat. sol. (b) **238**, 494 (2003)
[03M2] K. Matsuda et al., Phys. Rev. Lett. **91**, 177401 (2003)
[03M3] B. Möller et al., Appl. Phys. Lett. **83**, 2686 (2003)
[03N1] U. Neuberth et al., phys. stat. sol. (b) **238**, 494 (2003)
[03R1] V. Rotello, Nanostructure Science and Technology, Kluwer, Dordrecht (2003)
[03S1] Semiconductor Nanocrystals, A.L. Efros, D.J. Lockwood and L. Tsybeskov (eds.), Kluwer, Dordrecht (2003)
[03S2] M. Strassburg et al., phys. stat. sol. c **0** (6), 1835 (2003)
[03U1] S.M. Ulrich et al., Appl. Phys. Lett. **83**, 1848 (2003)
[03W1] U. Woggon et al., Phys. Rev. B **67**, 045204 (2003)
[03Y1] M. Yan et al., Appl. Phys. **94**, 5240 (2003)
[03Z1] R. Zimmermann, E. Runge and V. Savona, phys. stat. sol. (b) **238**, 478 (2003)
[03Z2] P. Zhang et al., Phys. Rev. A **67**, 012312 (2003)

[04C1] S. Chatterjee et al., Phys. Rev. Lett. **92**, 067402 (2004)
[04K1] A. Klochikhin et al., Phys. Rev. B **69**, 085308 (2004)
[04K2] A.M. Kapitonov et al., Phys. Rev. B, submitted (2004)
[04L1] N. Le Thomas et al., Phys. Rev. Lett., submitted (2004)
[04P1] H. Priller et al., phys. stat. sol. (c) **1** (4), 747 (2004)
[04P2] H. Priller et al., Proc. EXCON '04, Cracow, J. Lumin., in press
[04S1] R. Seguin et al., Appl. Phys. Lett. **84**, 4023 (2004)

16

Excitons Under the Influence of External Fields

A technique which reveals fascinating new phenomena as well as providing a powerful tool to detect and probe the properties of excitons is the application of external fields. For a rather early treatment see [08V1].

An exhaustive review of this field including the cases of temporally constant fields [79E1, 85H1] and of modulation techniques [69C1, 73S1, 04G1] would itself fill a whole book. Therefore, we restrict ourselves here to the presentation of general features and of some selected topics and examples.

We consider the influence of magnetic, electric and strain fields on the optical properties of excitons including their continuum states and we proceed again from bulk materials to structures of lower quasi-dimensionality.

16.1 Magnetic Fields

First we consider magneto-optics, i.e., the influence of a magnetic field $\boldsymbol{B}$ on excitons. There are two "natural" energy units which can be compared. One is the excitonic Rydberg energy Ry^* the other the cyclotron energy $\hbar\omega_\mathrm{c} = \hbar(e\boldsymbol{B}/\mu)$, where μ is the reduced mass of the exciton. The regime

$$\mathrm{Ry}^* \gg \hbar\omega_\mathrm{c}; \Rightarrow \quad \gamma = \hbar\frac{eB}{\mu \mathrm{Ry}^*} \ll 1 \tag{16.1a}$$

characterizes the weak field limit. The Coulomb energy dominates and the magnetic field can be treated as a perturbation.

In the strong-field limit

$$\mathrm{Ry}^* \ll \hbar\omega_\mathrm{c}; \Rightarrow \quad \gamma \gg 1\,, \tag{16.1b}$$

we have to consider first the Landau levels resulting from the free particle states by the quantization of the motion in the plane perpendicular to $\boldsymbol{B}$, and then the Coulomb energy. The intermediate regime $\gamma \approx 1$, which pertains in

many semiconductors in typical dc-fields of superconducting or Bitter-type magnets ($\boldsymbol{B}_{\max} \approx 40\,\mathrm{T}$) is more complicated to describe quantitatively.

It should be mentioned that a sufficiently strong magnetic field also produces quasi one-dimensional subbands similar to the case of quantum wires, i.e., of quasi one-dimensional systems (Sects. 8.9 and 8.12) according to the following argument. The carriers can move freely only parallel to $\boldsymbol{B}$, i.e., in one dimension. In the plane perpendicular to $\boldsymbol{B}$ the motion is quantized to circles in the classical picture resulting in a quantum mechanical description in a harmonic-oscillator-like term level scheme, the so-called Landau levels,

$$E\left(n_{\mathrm{l}}, \boldsymbol{k}_{\parallel}\right) - E_0 = \left(n_l + \frac{1}{2}\right)\hbar\omega_{\mathrm{c}} + \frac{\hbar^2 \boldsymbol{k}_{\parallel}^2}{2m_{\mathrm{eff}}}\,, \tag{16.1c}$$

where n_{l} is the quantum number of the Landau level and E_0 the band extremum without magnetic field. The density of states is again as in quantum wires for every subband proportional to $(E - E_{n_{\mathrm{l}}, \boldsymbol{k}_{\parallel}})^{-1/2}$, i.e., one obtains a similar picture as in Fig. 8.20 for vanishing damping with the main difference that the Landau levels are equidistant in energy. Damping washes the singularities out (see below). The selection rules for inter-Landau level transition are within one fan of Landau-levels and for simple, parabolic bands $\Delta n_{\mathrm{l}} = \pm 1$ and for interband transitions $n_{\mathrm{l,CB}} - n_{\mathrm{l,VB}} = 0$.

In accordance with the philosophy of this book, we present here the main effects, namely the diamagnetic shift, the Zeeman splitting, and the appearance of Landau levels, in a very elementary way and give some examples. References leading deeper into this field are [69C1, 73S1, 79E1, 82A1, 85H1, 97P1, 01K1, 04G1].

If we apply a magnetic field to an exciton, the relative motion of electron and hole is deformed by Lorentz forces. In perturbation theory, this deformation can be described by a weak admixture of other states. If we consider the ground-state $n_{\mathrm{B}} = 1$, $l = 0$, which has an S-like envelope function without angular momentum, we can describe this deformation as a small admixture of $l = 1$ (or P-) envelope states. The angular momentum resulting from this admixture is proportional to $\boldsymbol{B}$ and is oriented according to Lenz' rule (i.e., the minus sign in (2.1c)) antiparallel to $\boldsymbol{B}$. Since the energy of a magnetic dipole in a magnetic field also increases linearly with $\boldsymbol{B}$ we obtain in total a quadratic, so-called diamagnetic shift ΔE_{dia} to higher energies

$$\Delta E_{\mathrm{dia}} = a\boldsymbol{B}^2\,. \tag{16.2a}$$

The constant a is a material parameter which is proportional to the square of the Bohr radius of the exciton or more general, to its cross-section normal to $\boldsymbol{B}$. The dependence on a_{B} or n_{B} and on the material parameters explains that the diamagnetic shift increases with n_{B}. A typical value for the A-excitons in CdS ($a_{\mathrm{B}} \approx 2.8\,\mathrm{nm}$) is $a(n_{\mathrm{B}} = 1) \approx 2 \times 10^{-6}\,\mathrm{eV\,T^{-2}}$, i.e., for 10 T one finds a diamagnetic blue-shift of about 0.2 meV, which is just at the detection limit.

Exciton states that at $B = 0$ already have a non-vanishing magnetic moment, which can be aligned relative to $\boldsymbol{B}$, exhibit in addition the linear Zeeman splitting.

This magnetic moment can come from the spins of electron and hole. For singlet and triplet excitons with S envelope the difference or the sum of electron and hole g-factors enters, in simpliest approximation,

$$\Delta E_z = \pm\frac{1}{2} \mid g_\mathrm{e} \pm g_\mathrm{h} \mid \mu_\mathrm{B} B \tag{16.2b}$$

due to the relative alignment of electron and hole spin. The g-factors of electrons and holes can deviate significantly from two, due to influences of the bandstruckture or crystal symmetry. Tables are found e.g. in [82B1, 85H1, 02L1] or in Ref. [82L1, 01L1] of Chap. 1.

For states with $n_\mathrm{B} \geqslant 2$ there is an additional contribution from the magnetic moment of the envelope function for $l \geqslant 1$, depending on the orientational quantum number m in Sects. 9.1, 9.2.

If the states, which are subject to Zeeman splitting are split for some other reasons already for $\boldsymbol{B} = 0$ by an amount Δ, the Zeeman splitting is suppressed until ΔE_z becomes according to (16.2b) comparable to Δ. This fact can be easily seen by solving (or diagonalizing) a 2×2 Hamiltonian of the following type

$$\begin{pmatrix} E_0 + aB^2 + \frac{\hbar^2 \boldsymbol{k}_\parallel{}^2}{2m} - E & -\frac{1}{2}\mu_\mathrm{B} g B \\ -\frac{1}{2}\mu_\mathrm{B} g B & E_0 + \Delta + aB^2 + \frac{\hbar^2 \boldsymbol{k}_\parallel{}^2}{2m} - E \end{pmatrix} \tag{16.3}$$

For a quantitative calculation, perturbation theory for (almost) degenerate states has to be used, e.g., in the eight-fold space of $n_\mathrm{B} = 1$ excitons which can be constructed from the four-fold degenerate Γ_8 valence band and the twofold degenerate Γ_6 conduction band in T_d symmetry. The terms describing the diamagnetic shift appear together with the kinetic energy terms, the singlet-triplet and the longitudinal-transverse splitting in the main diagonal of the resulting 8×8 matrix, while the Zeeman terms, k-linear terms and others contribute to the off-diagonal elements. A detailed description is – as already mentioned – beyond the scope of this book; see e.g., [79E1, 82B1, 85H1]. Instead we give in the following some results starting with bulk materials.

16.1.1 Nonmagnetic Bulk Semiconductors

Figure 16.1 shows the behavior of the $n_\mathrm{B} = 1$ AΓ_5-exciton resonance of CdS with increasing magnetic field. The splitting is obviously connected with the Zeeman effect while the diamagnetic shift is hardly visible (see above).

In Fig. 16.2 we give transmission spectra of ZnO in the spectral region of the $n_\mathrm{B} = 1$ A- and B-exciton resonances, for the polarization $\boldsymbol{E} \parallel \boldsymbol{c}$. In this orientation all transitions have only very weak oscillator strength. Again

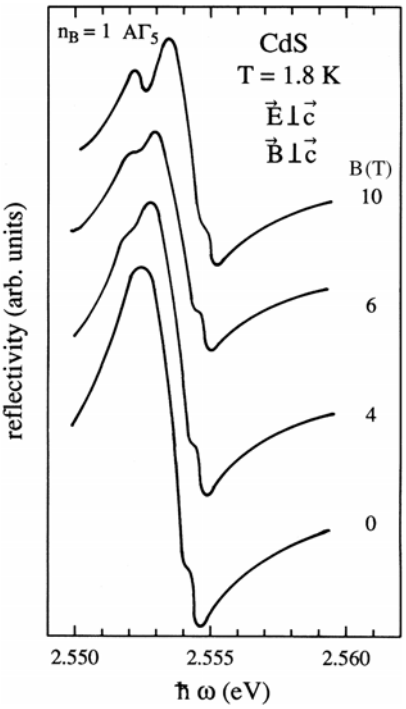


Fig. 16.1. Reflection spectra of the $n_B = 1A\Gamma_5$-exciton resonance for various magnetic fields [82B1]

we observe the Zeeman splitting of the A$\Gamma_{1,2}$ triplet states with increasing $\boldsymbol{B}$, which show at $\boldsymbol{B} = 0$ an accidental degeneracy, and the AΓ_5^L state. The narrow spike S is due to an isotropic point, i.e., a photon energy for which the exciton polariton dispersion relations for the polarizations $\boldsymbol{E} \parallel \boldsymbol{c}$ and $\boldsymbol{E} \perp \boldsymbol{c}$ cross or in other words where the ordinary and extraordinary refractive indices are equal. In this situation energy can be transferred from one polarization (here $\boldsymbol{E} \parallel \boldsymbol{c}$) to the other under energy and momentum conservation, leading to the dip in transmission. For a more recent discussion of the g-factors resulting from these data see also [60H1, 60T1, 02L1, 03C2, 04A1, 04M1] and Sect. 14.1.

Figure 16.3 depicts the splitting pattern of the $n_B = 1$A- and B-exciton resonances in CdS as a function of the magnetic field for the orientation $\boldsymbol{B} \perp \boldsymbol{c}$, $\boldsymbol{E} \perp \boldsymbol{c}$, showing nicely the combined influences of terms linear and quadratic in $\boldsymbol{B}$.

Figure 16.4a gives the dispersion relation of the $n_B = 1\Gamma_5$-exciton-polariton resonance in cubic (T_d) ZnTe. The influence of the light and heavy holes on the dispersion is clearly visible. Figure 16.4b shows the situation for finite $\boldsymbol{B}$ in a fairly arbitrary direction. In this situation all degeneracies are lifted and all exciton branches acquire some admixture of oscillator strength resulting in a total of eight different exciton and nine polariton branches. This is a rather frightening example for the dispersion of so-called magneto-

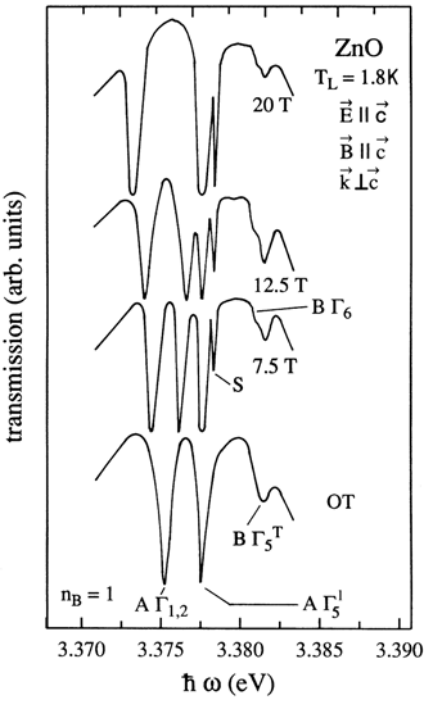


Fig. 16.2. Transmission spectra in the region of the $n_B = 1$A and B exciton resonances of ZnO for the orientation $\boldsymbol{E} \parallel \boldsymbol{c}$ [82B1]

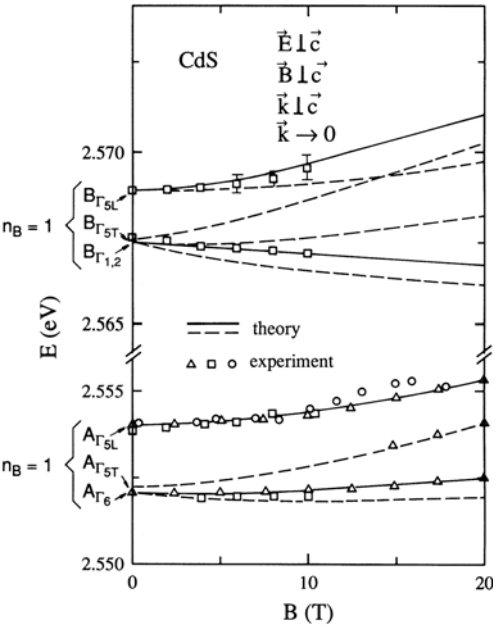


Fig. 16.3. The splitting pattern of the $n_B = 1$A- and B-exciton resonances in CdS with increasing magnetic field [82B1, 85H1]

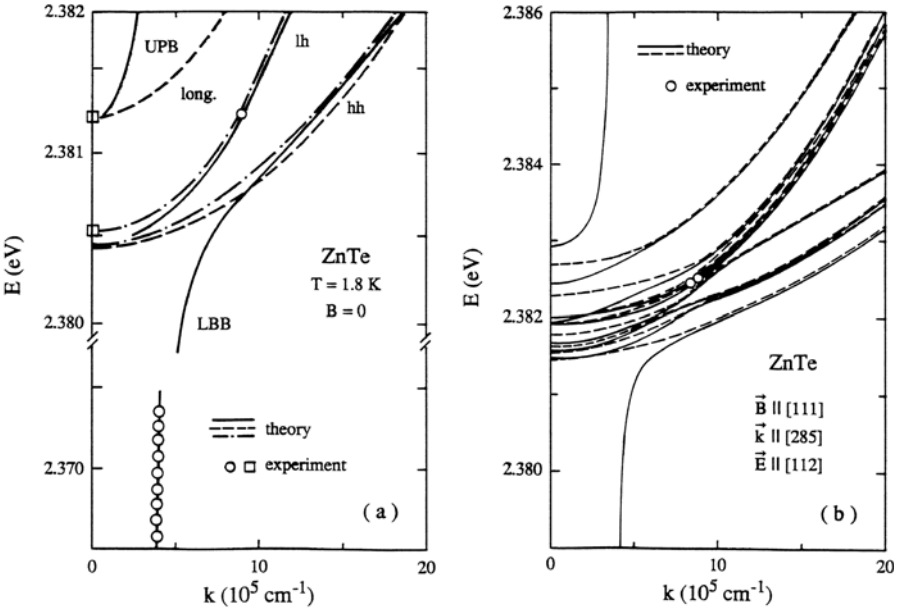


Fig. 16.4. Dispersion relation of the $n_B = 1$ exciton poloriton resonances in ZnTe without (**a**) and with (**b**) an applied magnetic field [85H1]

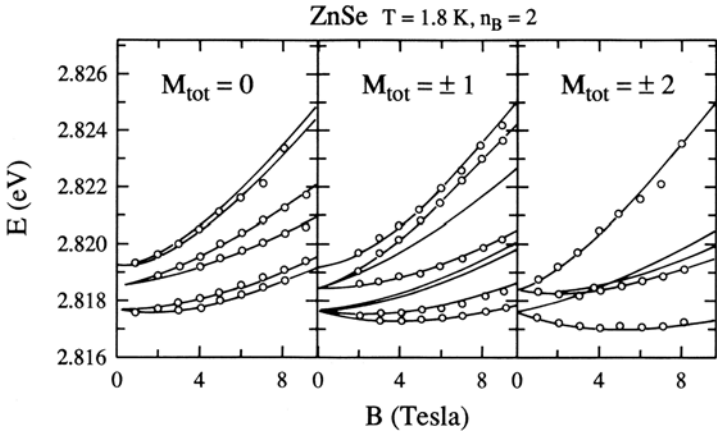


Fig. 16.5. The influence of a magnetic field on the $n_B = 2$ exciton resonances in ZnSe, determined in two-photon absorption spectroscopy [82H1]

polaritons. From Fig. 16.4b it is clear that it is wise to use "simple" geometries for the measurements, so that the effects remain relatively simple.

Until now we have considered in the examples exclusively the excitonic ground state $n_B = 1$. Therefore we now give in Fig. 16.5 the magnetic field behavior of the $n_B = 2$, $l = 1$ exciton resonances in ZnSe. There is already

a finite splitting for $B = 0$. For higher fields this splitting increases due to Zeeman terms. In addition, there is a pronounced diamagnetic shift due to the large radius. The total angular momentum M, consisting of spin and envelope contributions, could be chosen in the two-photon absorption experiments of Fig. 16.5 by using circular and/or linear polarizations of the two beams (see also Sects. 19.1 and 25.5).

Similar splitting patterns as for free excitons can also be observed for bound exciton complexes (BEC), where one has to take into account that the $\boldsymbol{B}$-field influences both the initial and the final state, e.g., in a recombination process from A^0X to A^0. Examples of the magnetic field behavior of BEC are given in [79E1, 81B1, 82K1, 86G1, 89G1, 04A1, 04M1].

At the beginning of this section we noted that the observation of Landau levels might be difficult, since for most semiconductors the limit $\gamma \gg 1$ can be reached only for fields $\boldsymbol{B} \gtrsim 10^2$ T, which, if available at all, are usually only in pulsed form.

There is, however, a way to overcome this difficulty by investigating the continuum states, where the Coulomb interaction still influences the oscillator

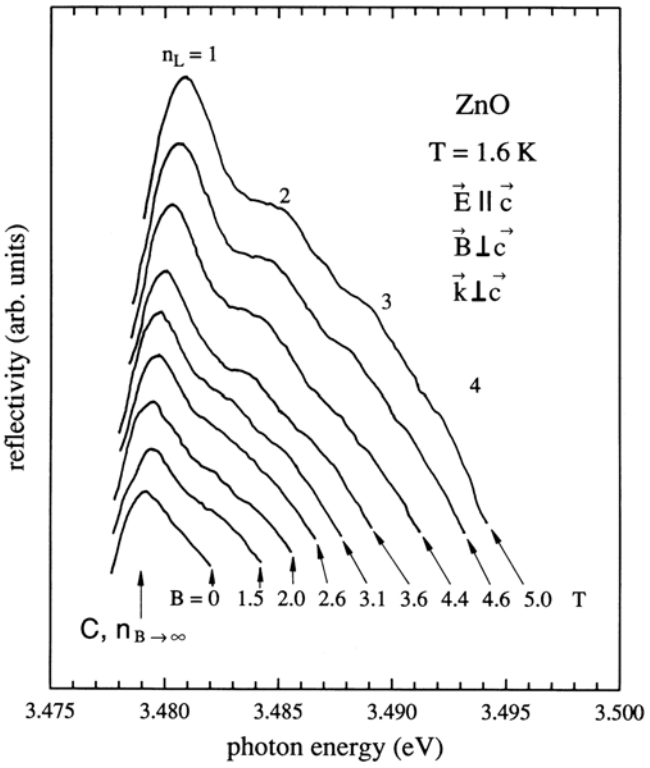


Fig. 16.6. The reflectivity in the exciton continuum in ZnO for various magnetic fields [73H1]

strength but where the motion of electron and hole is almost free. To observe the Landau levels, their broadening must be smaller than the cyclotron energy as in micro wave experiments, in other words

$$\omega_c T_2 \gtrsim 1 \tag{16.4}$$

where T_2 is the phase-relaxation time. Bulk samples with direct, dipole-allowed band-to-band transitions are usually opaque in this region so that reflection spectroscopy is the appropriate tool.

In Fig. 16.6 we give as an example the reflectivity of ZnO at the beginning of the continuum states of the C-exciton series. The Landau-level structure becomes obvious for the highest B-fields. The observation of the Landau levels allows the sum of electron and hole masses to be determined. If the electron mass is known, e.g., from cyclotron absorption by n-type materials, the hole mass can then be determined.

In semiconductors with relatively small exciton binding energies like GaAs, the magnetic field may shift states into the continuum, resulting in Fano-interference-type lineshapes. The quantization of the continuum leads to a reduction of this effect. For some examples see, e.g., [95S1, 97B1].

16.1.2 Diluted Magnetic Bulk Semiconductors

An interesting group of semiconductors are the so-called diluted or semimagnetic semiconductors (DMS or SMSC).

Some review books and papers on their properties are [88S1, 91O1, 92G2, 92S1,92Y1,94G1,95M1,96Y1,01D1,01O1,02S1,03D1,03D2] and the references therein. While the older articles concentrate more on bulk materials, the more recent ones give general results for quantum structures.

Most of the information given below is taken from the above references.

The DMS are generally II–VI and more recently also III–V semiconductors doped with magnetic ions up to several tens of percent, i.e., they can also be considered as alloys for higher concentrations, with the corresponding consequences for composition dependencies for phonons, the band gap, exciton binding energies and localization, as outlined in Sects. 7.8, 8.15, 9.6, 11.1.6 and 14.4. The doping atoms must carry, by definition, spin and magnetic moment. The most widely used one is Mn^{2+} with a half-filled 3d shell but also the influences of Fe, Co, Ni, V, Cr, Eu or Gd are investigated. These atoms sit ideally on the metal or cation site of the host. The most widely used host materials are narrow and wide gap II–VI materials like $Hg_{1-x}Mn_xTe$, $Cd_{1-x}Mn_xTe$ or $Zn_{1-x}Mn_xSe$. More recently ZnO:V and ZnO:Co have also been investigated in addition to III–V compounds like GaAs:Mn. Since Mn^{2+} acts on a Ga site and acts as an acceptor, GaAs:Mn samples are metallic without compensating co-doping.

Most of the binary DMS compounds show antiferromagnetic ordering like MnTe and $Cd_{1-x}Mn_xTe$ for concentrations x down to about 0.35. Some

others are ferromagnetic like $Ga_{1-x}Mn_xAs$ or $Ga_{1-x}Mn_xN$ or ZnO:V and ZnO:Co [98O1,00D1,02D1,02S2,02S3]. The ferromagnetism can be influenced by the concentration of free carriers e.g. holes in GaAs:Mn. Their density can be decreased by codoping with a donor. Similarly codoping with a donor like Ga or Al increases the electron density in ZnO:V or ZnO:Co.

At low concentration they are paramagnetic with very high saturation magnetization and can exhibit spin glass behavior at low T.

We now start to list more of the specific semimagnetic properties for bulk DMS.

The main point is that the exchange interaction occurs between the s- and/or p-type conduction and valence bands of the host and the 3d orbitals of the magnetic ions, known as sp-d exchange. A free or bound carrier can polarize the magnetic moments of the surroundings, e.g., Mn^{2+} ions. This fact results in extremely high Zeeman splitting and effective g-factors for both the free or bound carriers and for the excitons. The effective g-factors can reach values up to 100 and the Zeemann splitting saturates at values between 10 and 100 meV for fields of a few T. Both the splitting values and the g-factors are thus by one or two orders of magnitude larger than in nonmagnetic semiconductors. We give an example in Fig. 16.7 that can be compared, e.g., to Figs. 16.1–16.3. These large values of the Zeeman splitting are also held responsible for the large values of Faraday and Kerr rotation ob-

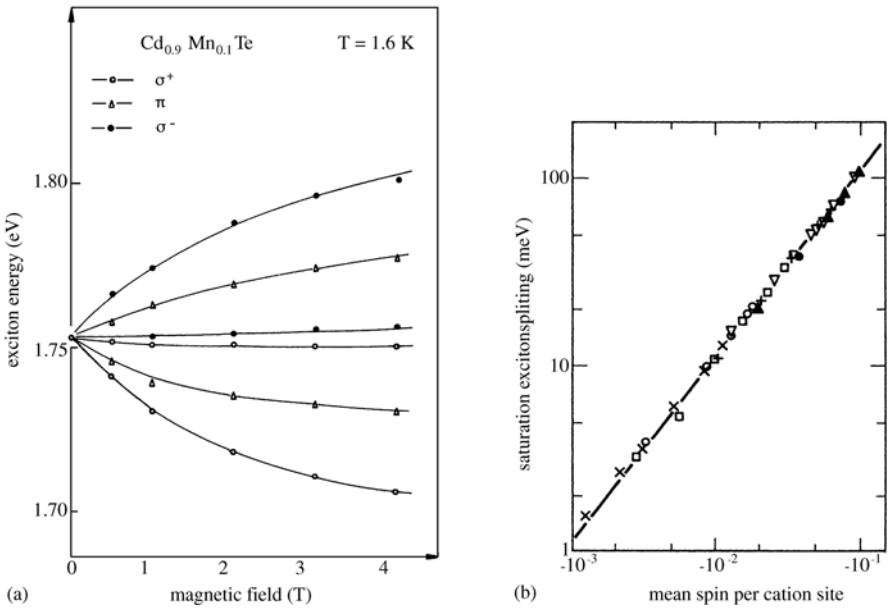


Fig. 16.7. (**a**) Zeeman splitting of the exciton of $Cd_{0.9}Mn_{0.1}Te$ as a function of magnetic field in the Faraday configuration, i.e., $\boldsymbol{k} \parallel \boldsymbol{B}$. (**b**) Saturation Zeeman splitting per average Mn^{2+} spin per cation site [88G1]

served in transmission and reflection geometries, respectively. For examples see, e.g., [78G1].

Free and even more bound carriers can form so-called magnetic polarons. This is a carrier surrounded by a magnetization cloud of oriented Mn^{2+} moments in a similar way as a usual carrier can be surrounded by a phonon cloud forming a polaron (see Sect. 8.6). The formation time of a magnetic polaron after creation of a carrier (e.g., by optical excitation) is finite and in the range of 100 ps or longer. The formation of the magnetic polaron cloud lowers the energy of the carrier. This energy decreases with increasing temperature and increasing external magnetic field [92Y1, 96Y1].

In narrow gap DMS one observes with increasing magnetic field well-resolved Landau-level fans [91O1]. Mn^{2+} ions in II–VI DMS show an internal transition resulting in a luminescence band around 2.2 eV, i.e., in the yellow spectral range and absorption features starting at or above this energy when the band gap, or better exciton energies, are situated above theses transitions. This is the case for $Cd_{1-xy}Mn_xTe$ approximately for $x > 0.4$ [84L1].

For a collection of data see, e.g., [82L1] of Chap. 1.

16.1.3 Semiconductor Structures of Reduced Dimensionality

In quasi two-dimensional systems the transmission spectra are easily measured if the substrate is transparent or has been removed by selective etching. If both conditions are not fulfilled, there remain the techniques applicable to bulk materials, i.e., photoluminescence, photoexcitation or reflection spectroscopy.

Many properties are similar to bulk samples, e.g., there is a diamagnetic shift proportional to the area of the exciton in the quantum well if $\boldsymbol{B}$ is normal to it, there is a Zeeman splitting, and there are Landau levels. See, e.g., [92B2, 92O1, 92R1, 92S1, 95B1, 01K1, 03S1] or [01L1] of Chap. 1.

It should be noted that the electron and hole states are fully quantized in all three directions for a magnetic field normal to the quantum well.

In Fig. 16.8 we give an example for the development of the Landau-levels of a GaAs quantum well sample with 8.5 nm-wide wells for magnetic fields up to 12 T.

The two features at lowest energy, the heavy and light hole excitons, are only weakly affected by $\boldsymbol{B}$ but energetical above them beautifully develops the fan of Landau levels.

Apart from similarities to bulk materials there are also new phenomena:

Due to the strong anisotropies of quantum wells and wires, the magnetic field effects depend not only on its orientation to the crystallographic axes but also on its orientation relative to the well or wire, e.g., [96O1].

A further effect comes from the well-width dependence of the lateral part of the wavefunctions. With decreasing well width the wave function penetrates deeper into the barrier. Consequently the influence of the magnetic properties of the barrier, e.g., of its g-factors, grows. In Fig. 16.9 we give two examples. The electron g-factor is negative in GaAs and positive in AlAs.

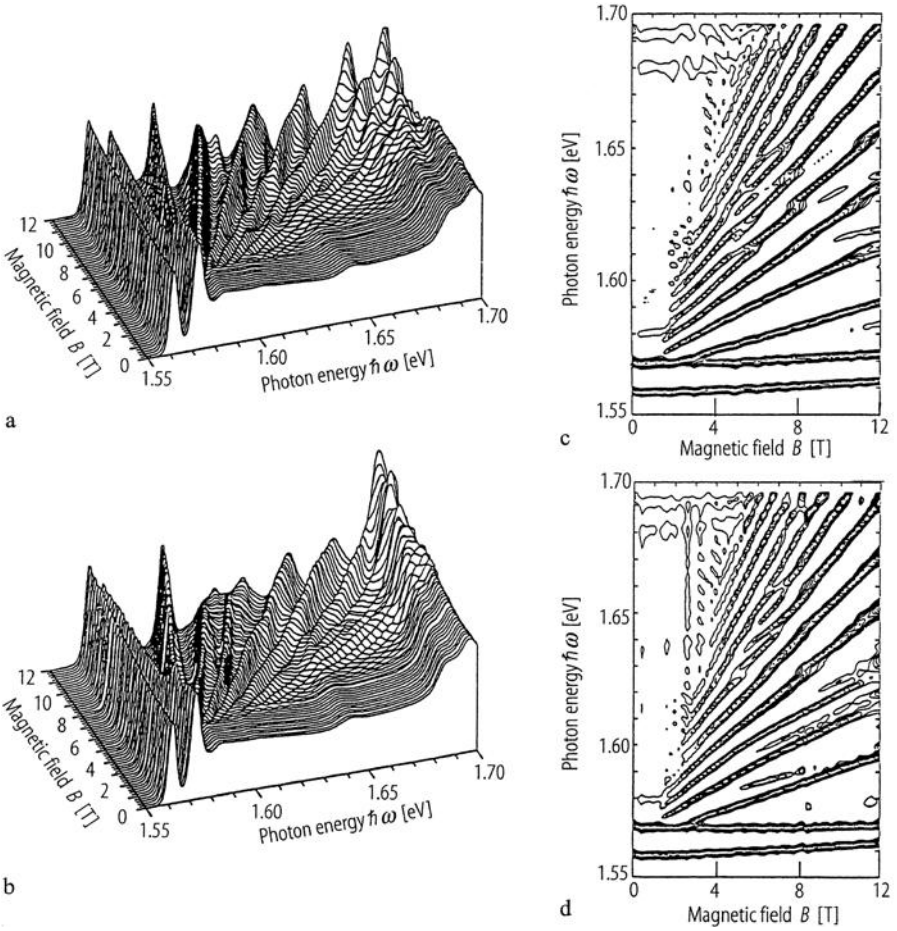


Fig. 16.8. The linear magneto absorption spectra for two different circular polarizations of the light (**a**, **b**) The resulting contour plots (**c**, **d**) [91S1, 92S1]

Consequently g_e increases with decreasing well width of GaAs/$Al_{1-y}Ga_yAs$ (a) or GaAs/AlAs structures and even exhibits a change of sign. Furthermore one observes a variation of g_h with l_z and a dependence on the orientation of the magnetic field being either perpendicular or parallel to the quantum well.

For an example of Si/$Si_{0.76}Ge_{0.24}$ quantum wells see [00P1].

In diluted magnetic semiconductors the magnetic ions can be incorporated into one or both barriers or into the well itself, allowing information on the interface to be obtained [94G1]. The huge Zeeman splitting and the formation of bound or free magnetic polarons have been investigated, e.g., in [92G1, 92Y1, 96Y1]. Due to the huge Zeeman splitting it might even be possible to change the band alignment from type I to type II by a magnetic

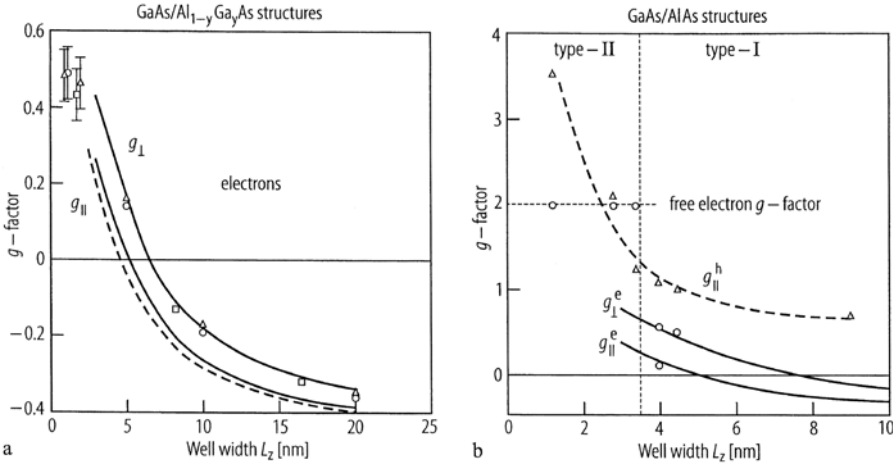


Fig. 16.9. Dependence of electron (and hole) g-factors on the well width and on the orientation of $\boldsymbol{B}$ in $GaAs/Al_{1-y}Ga_yAs$ (**a**) and GaAs/AlAs structures (**b**). In (**b**) the transition from a type I to a type II band alignment is also indicated (see Sect. 15.2) [92I1, 95H1]

field if the respective band offset is small [86B1, 91J1, 94F1, 03C1] or [01L1]of Chap. 1.

For biexciton states in diluted magnetic semiconductor quantum wells in high magnetic fields see [01M1].

To conclude this section we give some references on excitonic magneto optics in quantum wires and dots, which allow the reader to become familiar with this field of research.

For microstructured, V-groove and T-shaped quantum wires see, e.g., [94P1, 95O1, 96W1, 98B1, 98L1, 00L1, 01K1, 02M1]. For DMS wire structures see [00O1,00T1,01I1,01K2], for self-organized quantum islands including DMS see [96R1, 99B1, 99B2, 00K1, 01K1, 02B1, 02M1, 03C1, 03P1] and for quantum dots in insulating matrices see [94N1, 97O1, 98K1, 99G1].

16.2 Electric Fields

In the same way that we know the quadratic (diamagnetic) and linear (Zeeman) effects of a $\boldsymbol{B}$-field from atomic physics and now also for excitons in semiconductors we can guess what will happen if we apply a static electric field $\boldsymbol{E}_\mathrm{s}$ to the exciton resonances in semiconductors. For the ac or optical Stark effect see Sect. 20.4. There will be a Stark effect, i.e., a shift (and splitting) of the exciton resonances, usually quadratic but possibly also linear in $\boldsymbol{E}_\mathrm{s}$. The influence of $\boldsymbol{E}_\mathrm{s}$ on the band-to-band transitions, i.e., on the continuum states is also known as the Franz–Keldysh effect. The tilting of the

bands by an applied static electric field $\boldsymbol{E}_\mathrm{s}$ allows, in this case, the Bloch-type electron and hole wave functions to tunnel into the forbidden gap thus resulting in a red-shift of the band-to-band absorption tail and the appearance of an oscillatory structure above, i.e., in the continuum states. The modifications of the absorption spectra due to $\boldsymbol{E}_\mathrm{s}$ also influence the refractive index in the transparent spectral region below the exciton resonances according to the Kramers–Kronig relations (Chap. 6), resulting in Pockels- or Kerr-effect-like phenomena, which describe the modification of an existing or the introduction of a new birefringence, depending linearly or quadratically on $\boldsymbol{E}_\mathrm{s}$, respectively. We shall concentrate here on the influence of $\boldsymbol{E}_\mathrm{s}$ on the eigenfrequencies.

16.2.1 Bulk Semiconductors

The observation of the Stark effect for excitons in three-dimensional (or bulk) samples is not easy, for the following reason: To get an observable shift of the eigenstates, the "electric field energy" $e\boldsymbol{E}_\mathrm{s}a_\mathrm{B}$ should be comparable with the spectral width of the absorption bands, i.e. at low temperatures,

$$e\boldsymbol{E}_\mathrm{s}a_\mathrm{B} \gtrsim \Delta_\mathrm{LT}\,. \tag{16.5}$$

Equation (16.5) requires fields of the order of $10^6\,\mathrm{Vm}^{-1}$ depending on the material parameters. On the other hand, such fields broaden or even destroy the exciton resonances due to two effects. One is field ionization of the exciton, by tunneling through the finite Coulomb barrier at finite fields, as illustrated in Fig. 16.10. The other is impact ionization; this means that carriers which are always present in a semiconductor at finite temperature can gain such high energies in the electric field, that they can ionize an exciton if they hit it, resulting in two more carriers and a collision-broadening of the exciton

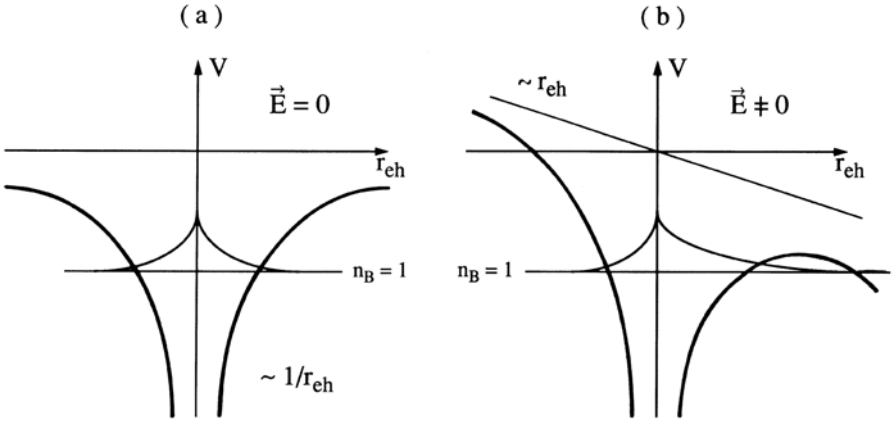


Fig. 16.10. The Coulomb potential of a three-dimensional exciton without (**a**) and with (**b**) an applied constant electric field

resonance. For the above reason, not too many successful attempts have been reported to observe the Stark effect for excitons in bulk materials directly.

Instead, small modifications occurring at low fields have usually been used in modulation spectroscopy [67G1, 69C1, 70H1, 71M1, 73S1, 76H1, 82L1, 84S1, 85S1].

Some (early) references for the observation of the Stark effect of excitons in inorganic semiconductors like CdS, CdSe or CuCl are [69H1,70M1,71L1,73R1] and for organic ones [73B1].

Examples for the Franz–Keldysh effect in bulk semiconductors can be found, e.g., in [71S1, 73B2, 76B1] and for the Pockels effect, e.g., in ZnO in [70M1, 71M1].

In Fig. 16.11 we show an example for the quadratic Stark effect of the $n_{\mathrm{B}} = 1$ exciton in CdS. The shift and a slight broadening of the excitonic absorption band are clearly visible.

As mentioned above, the Franz–Keldysh oscillations appear in the band-to-band transition region. The energies of the alternating maxima and minima E_n follow the relation [04G1]

$$(E_n - E_g)^{3/2} \sim n \tag{16.6}$$

In Fig 16.12 we show modulation spectra of GaAs samples at 100 K exhibiting the Franz-Keldysh effect and the verification of (16.6). The Franz–Keldysh effect has been induced by a surface electric field.

In Fig. 16.13 we give results of a rather unusual type of excitation spectroscopy. A GaSe sample is illuminated with monochromatic light through

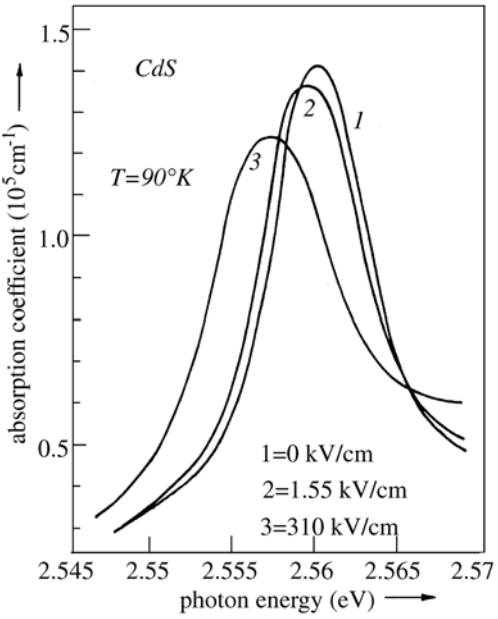


Fig. 16.11. Influence of an external field on the absorption spectrum of the B, $n_{\mathrm{B}} = 1$ exciton of CdS for various applied electric fields [71L1]

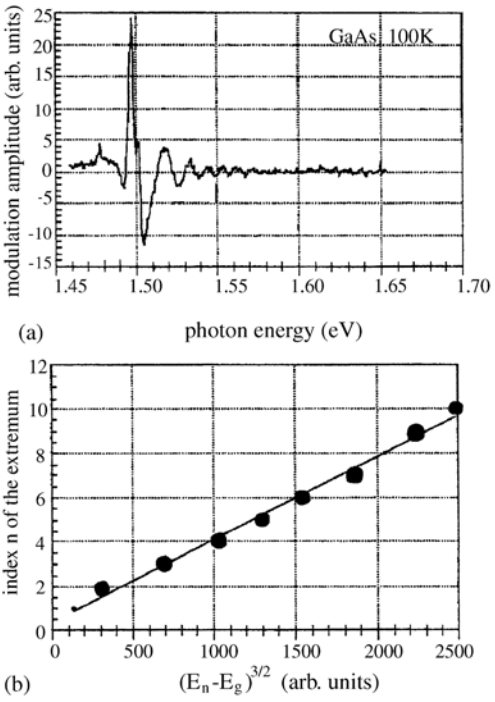


Fig. 16.12. Franz–Keldysh oscillations observed in a GaAs sample (**a**) and the relation (16.6) (**b**) [95L1, 04G1]

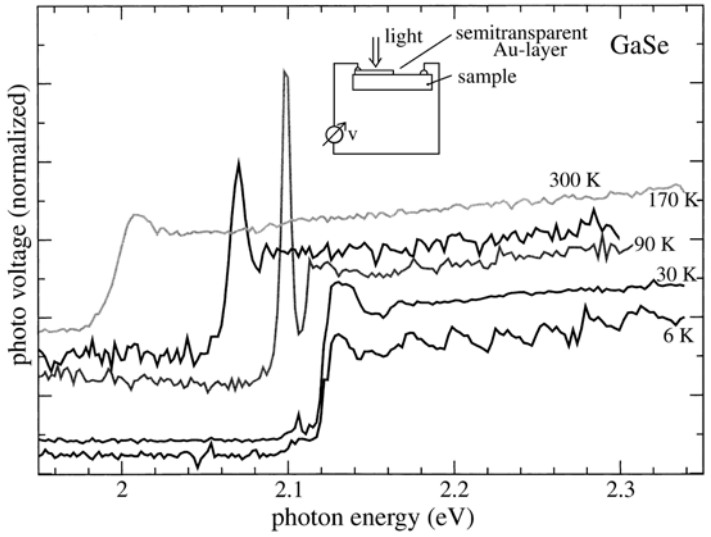


Fig. 16.13. Photovoltage excitation spectra of GaSe. The spectra are normalized in the continuum regime and shifted for clarity [02G1]

a semitransparent Au electrode and the photovoltage that develops between the contacts is measured as a function of the photon energy. This type of photovoltage excitation spectroscopy gave different spectra depending on sample temperature.

At low temperature one only sees the continuum transitions superimposed by some oscillatory structure, possibly Franz–Keldysh oscillations, which may result from the built-in field of the Schottky contact. With increasing temperatures these oscillations disappear and the exciton resonance shows up, shifts with the band gap and broadens. The exciton resonance is probably suppressed at low temperatures by field ionization in the Schottky barrier. Which features dominate depends on the various length scales like the extension of and field strength in the Schottky barrier or the penetration depth of the light.

16.2.2 Semiconductor Structures of Reduced Dimensionality

The problem of field ionisation in bulk semiconductors can be overcome by confining the electron and hole between barriers, e.g., in quantum wells, wires or dots, as shown schematically in Fig. 16.14. With increasing $\boldsymbol{E}_\mathrm{s}$ oriented perpendicular to the layers the electron and hole shift into their respective corners, reducing their energetic separation. This results in a roughly quadratic redshift of the gap and thus of the exciton resonance.

The overlap between electron and hole wavefunction is thereby reduced, resulting in a decrease of the oscillator strength and a decrease of the exciton binding energy. This latter effect, however, is only a small correction to the reduction of the energetic separation of the first quantized levels. In addition the selection rule $\Delta n_z = 0$ is relaxed because $\boldsymbol{E}_\mathrm{s}$ mixes states with odd and even parity, inducing some transitions which are forbidden for $\boldsymbol{E}_\mathrm{s} = 0$. For a detailed elaboration of this so-called quantum-confined Stark effect (QCSE) see [85M1, 86M1]. All three above-mentioned effects, namely the redshift of the exciton, the decrease of its oscillator strength, and the appearance of forbidden transitions, are illustrated in Fig. 16.15, where this quantum-confined Stark effect is shown for an $\mathrm{InP}_{1-y}\mathrm{Ga}_y\mathrm{As}$ based MQW located in the intrinsic region of a reverse-biased pin diode as shown in the inset.

For the quantum confined Stark effect in $\mathrm{Si}_{1-y}\mathrm{C}_y/\mathrm{Si}$ quantum wells see, e.g., [99F1] and the references therein.

If an electron moves in a quantum well in an electric field normal to the well, the spin of the electron precesses due to spin-orbit coupling. This effect allows the manipulation of the electron spin as shown in [03R1].

The application of an electric field normal to the planes of superlattices shifts adjacent wells energetically with respect to each other by an amount

$$\Delta E = e\,|\boldsymbol{E}_\mathrm{s}| \cdot d \tag{16.7}$$

where $\boldsymbol{E}_\mathrm{s}$ is the applied field and d is the superlattice period. This fact destroys the origin of the miniband formation namely the overlap of the levels of adjacent wells with equal energy. Instead one obtains a so-called Wannier–Stark ladder. An example will be shown later in Sect. 23.2 in connection with the Bloch oscillations. Further examples can be found, e.g., in [93M1, 94Y1] or Ref. [01L1] of Chap. 1

For the observation of the quantum confined Stark effect in quantum wires see, e.g., [99W1] and the references therein.

The quantum confined Stark and Franz–Keldysh effects are also known for quantum dots in insulating matrices or for (self-assembled) quantum islands in quantum wells. For references see, e.g., [88M1, 90E1, 98H1, 02W1].

The electric field of charges trapped in the vicinity of quantum dots and islands is also (at least partly) responsible for the abrupt steps in the emission energy of single dots. For examples see [89H1, 90R1, 92N1, 93W1, 98R1, 99E1, 01F1, 01K3, 01K4].

Electroluminescence of quantum dots has been reported in [97A1, 02C1].

Phenomena like that of Fig. 16.15 lead into the regime of opto-electronics or electro-optics, to which we shall return briefly in Sect. 24.1.5 and 24.2.

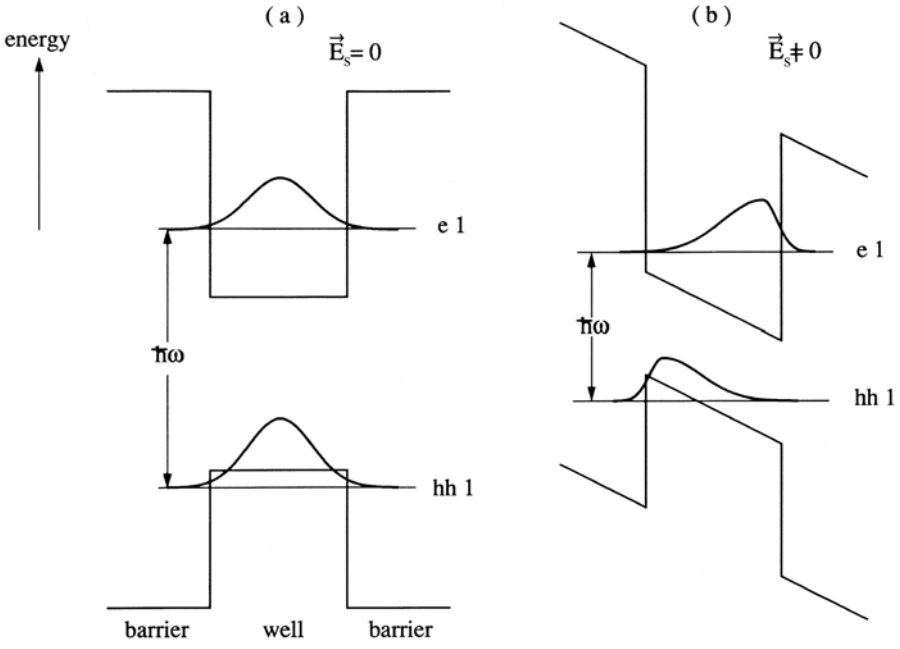


Fig. 16.14. The bandstructure of a quantum well without (**a**) and with (**b**) an applied constant electric field perpendicular to the wells, but neglecting the Coulomb interaction between electrons and holes

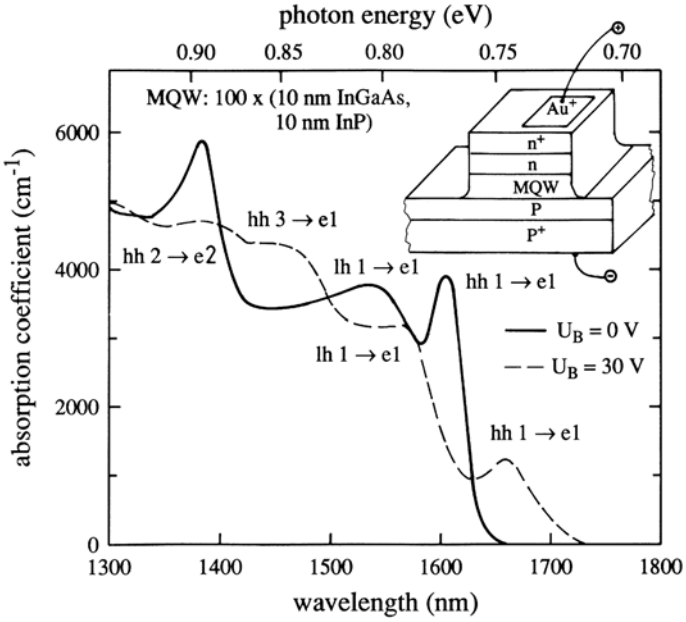


Fig. 16.15. The quantum-confined Stark effect in InGaAs MQW samples [87B1]

16.3 Strain Fields

The third external perturbation which we shall consider here are mechanical strain fields. The crucial quantity which enters here is the dependence of the energy of the band extrema on the strain, the so-called deformation potential. It is defined for conduction and valance band by (see (8.21))

$$\Xi = a\frac{\mathrm{d}E_{\mathrm{c,v}}}{\mathrm{d}x} \tag{16.8}$$

where a is the lattice constant, $\mathrm{d}E_{\mathrm{c,v}}$ is the shift of the band edge caused by a length variation of the unit cell by an amount $\mathrm{d}x$. Deformation potentials are defined for various types of strain like hydrostatic, uniaxial or shear strain. For details see, e.g., [82L1, 96Y1] of Chap. 1 or Sect. 8.6.

For excitons the relevant quantity is the sum of the deformation potentials of conduction and valence bands. As in the case of magnetic and electric fields, an applied stress changes the eigenenergies of the exciton states and may also lead to a splitting of degenerate states, if the resulting strain is oriented such that it reduces the symmetry of the crystal. The latter situation is e.g. realized when stress is applied perpendicular to the crystallographic $\boldsymbol{c}$ axis of a uniaxial material (e.g., the Wurtzite structure) where it lifts the degeneracy in the plane perpendicular to $\boldsymbol{c}$, but not if the stress acts in the direction of $\boldsymbol{c}$. This orientation leads only to a shift of eigenenergies but not to a splitting since the symmetry is not changed. For details see Chap. 26 on group theory.

16.3.1 Bulk Semiconductors

We now give various examples: Figure 16.16 shows reflection spectra of CdS for zero and finite stress perpendicular to the $\boldsymbol{c}$ axis. The shift and the splitting of the $n_{\mathrm{B}} = 1\mathrm{A}\Gamma_5$ and $\mathrm{B}\Gamma_5$ exciton resources and the variation of their oscillator strengths are obvious. Measurements of this type can be used to determine the deformation potentials. For materials with zinc-blende structure see [70L1, 79T1, 79R1].

Since bound exciton complexes (BEC) often have very narrow absorption and emission bands (Sect. 14.1) one can easily study the influence of magnetic or strain fields, which affect both the ground and the excited state as already mentioned above. In Fig. 16.17 we give the shift of the luminescence of a BEC in ZnO which shows clearly the influence of the strain. According to the statement given above, a shift but no splitting occurs for the orientation $\boldsymbol{S} \parallel \boldsymbol{c}$ in contrast to the situation of Fig. 16.16. For more details see [68S1].

Recently the influence of strain acquired new importance in the field of epitaxial growth of heterolayers.

There are two contributions to the "biaxial" strain, which occurs in the plane of epitaxial growth, and which can be considered as a superposition of a hydrostatic and a uniaxial strain. One contribution to the strain already arises during the growth process if the lattice constants of substrate and layer do not coincide at growth temperature. The first epitaxial lay-

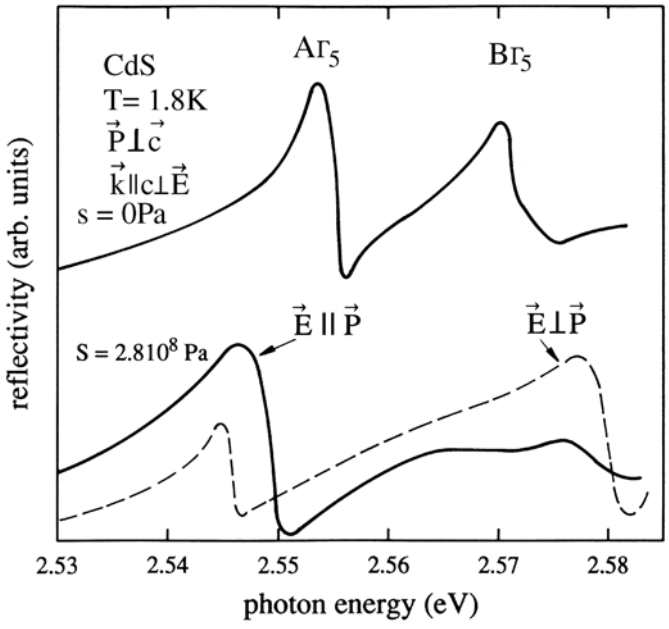


Fig. 16.16. The excitonic reflection spectra of CdS under uniaxial stress [70L1]

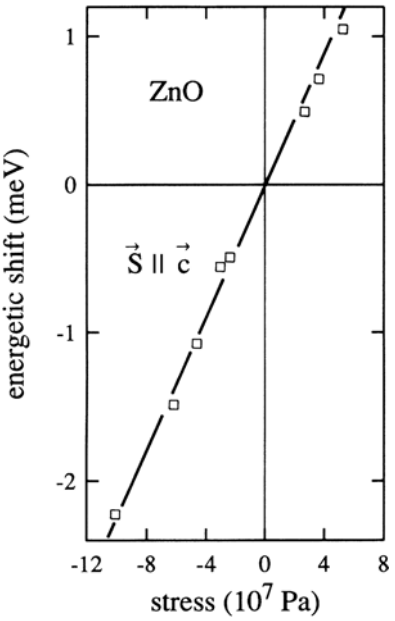


Fig. 16.17. The shift of a BEC luminescence line in ZnO under the influence of strain [68S1]

ers will often grow with some strain, trying to match the lattice constant of the substrate. With increasing layer thickness it is energetically more favorable for the layer to create a dislocation network which relaxes the strain [85P1, 87D1], or to form islands. This latter growth mode is known as Stronski–Krastanov mode. See Sect. 8.13. Sometimes a small residual strain remains, independent of the layer thickness [92W1, 92W2, 92G1]. The critical thickness l_c for the onset of dislocation formation depends on the lattice misfit, on the elastic constants of the epilayer, on the energy necessary to produce the dislocations, on the surface energies etc. It can be as small as a few atomic layers only. The dislocation network explains the function of a buffer layer or a short period superlattice, which are often used if two materials of different lattice constant are grown on top of each other. In this case the crystalline quality of the interface is considerably poorer than that of the surface of the layer after a thickness of about 0.1 μm.

The other part of the biaxial strain arises during the cooling from the growth temperature to the temperature at which the (optical) measurements are performed if the coefficients of thermal expansion are different for substrate and layer, as is usually the case. A part of this thermal strain can sometimes be accommodated by modifications in the dislocation network. In Fig. 16.18 we show reflection spectra in the region of the $n_B = A\Gamma_5$ and $B\Gamma_5$ excitons in CdS grown on SrF_2.

Since the cubic SrF_2 substrate was oriented (111), the hexagonal CdS grows with the $\boldsymbol{c}$ axis perpendicular to the interface. Consequently the C_{6v}

symmetry is not affected by the biaxial strain, but the eigenenergies are shifted. The bulk values for the longitudinal eigenenergies (or the reflection minima) are given by the vertical lines. In agreement with the discussion above, we see that the excitons are shifted and that this shift is different at the interface and the surface. Furthermore, the damping is larger at the interface than at the surface due to the dislocations in the former region.

In contrast to Fig. 16.18, we show in Fig. 16.19 a luminescence spectrum of the cubic material ZnTe grown on GaAs(00l). In this case, the biaxial strain reduces the cubic T_d symmetry of bulk ZnTe and the Γ_5 exciton of Sect. 13.1 splits into its light and heavy hole components.

Splittings as in Fig. 16.19 also play a role in strained-layer superlattices in addition to the mass-dependent quantization energies.

The biaxial in-plane tensile strain results essentially from the different thermal expansion coefficients of the GaAs and ZnTe layers since the ZnTe layer thickness of 2 μm is well above l_c and the ZnTe layer grows with its own lattice constant.

Inhomogeneous strain can be used to create potential traps for excitons. See, e.g., Sect. 21.3 for Si or Ge and [03N1] for Cu_2O. However, strain resulting from the fixation of the sample on the holder in a cryostat is usually unwanted and should be avoided.

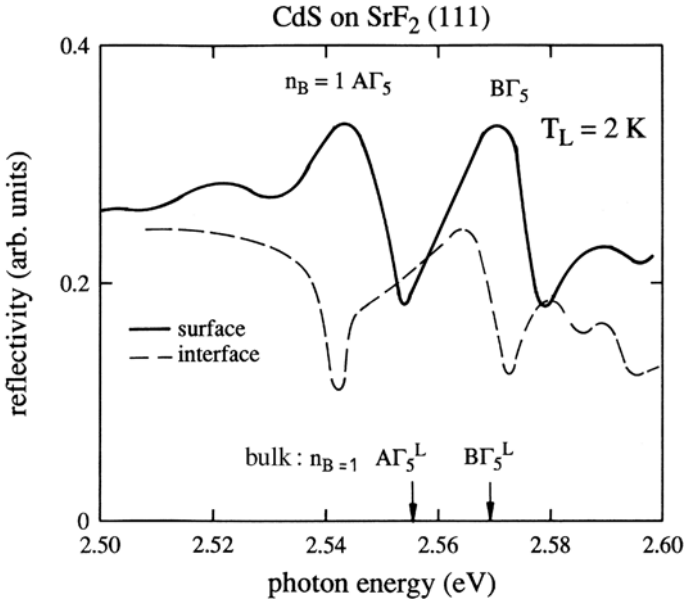


Fig. 16.18. Reflection spectra of the $n_\mathrm{B} = 1A\Gamma_5$ and $B\Gamma_5$ excitons of a CdS layer grown on SrF_2 and measured at the surface and at the interface. The positions of the longitudinal eigenenergies in the bulk are indicated by *vertical arrows* [92B1]

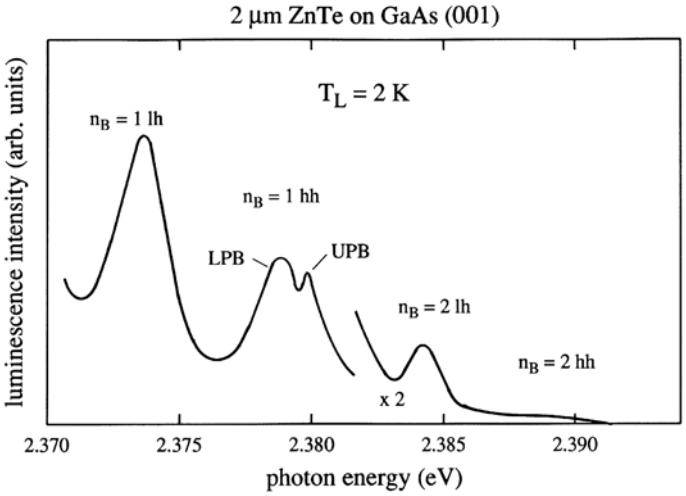


Fig. 16.19. Luminescence spectrum of a ZnTe layer grown on GaAs (001), showing the strain-induced splitting of the light and heavy hole $n_B = 1$ exciton resonances [92G1, 92W1, 92W2]. Compare with Fig. 16.4a

16.3.2 Structures of Reduced Dimensionality

The above-mentioned strain-induced splitting of the Γ_8 valence band is of special importance for strained quantum wells and superlattices whenever a layer is grown with a lattice constant different from that of the substrate.

A two-dimensional tensile strain in the plain of the layer is usually accompanied by a uniaxial compressive strain normal to it and vice versa. The results of this strain on the width of the gap and on the arrangement of the *hh* and *lh* valence bands is shown schematically in Fig. 16.20 and verified for tensile strain, e.g., in Fig. 16.19.

In quantum wells and superlattices (and in structures of even lower quasi-dimensionalities) the strain induced effects come in addition to the quantization energies, which are always larger for the *lh* exciton compared to the *hh*. Consequently, strain effects may enhance or reduce the splitting caused by confinement. For strained superlattices one frequently uses a substrate with a lattice constant situated between the values of the two different layer materials of the superlattice resulting in so-called strain symmetrized superlattices because the layers show alternating compressive and tensile strain. A similar situation can be reached in "free standing" superlattices, which occur when the thickness of the superlattice is so thick that the coherence of its lattice constant to the substrate is lost, e.g., by the formation of dislocations.

While many III–V systems can be grown lattice-matched to GaAs, InAs or InP substrates, the II–VI and IV–IV systems are usually strained, as seen in Fig. 8.31. Typical examples are, e.g., $Si_{1-y}Ge_y$/Si superlattices (e.g. [97K1] and references above), CdS/ZnS [97H1] or CdS/ZnSe [99P2, 01D1].

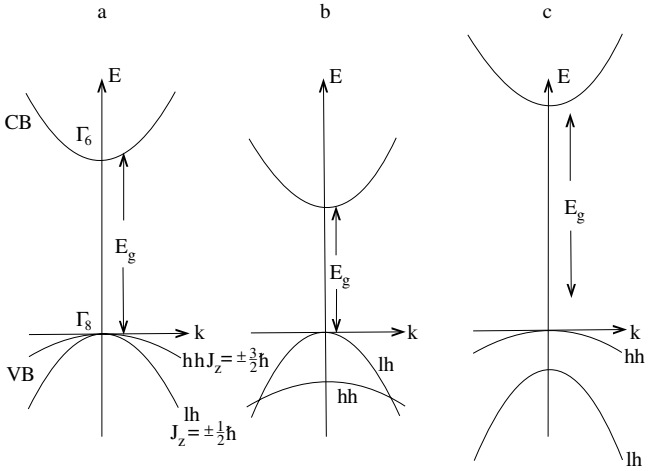


Fig. 16.20. Schematic drawing of the influence of strain on the alignment of the *hh* and *lh* valence bands and on the value of the gap (**a**) zero strain (**b**) two-dimensional tensile strain and normal uniaxial compressive strain (**c**) and two-dimensional compressive strain and normal uniaxial tensile strain. See also Fig. 8.32

For the creation of an inhomogeneous strain, acting as a potential trap for excitons in quantum wells see, e.g., [99N1].

Phase transitions from zincblende or wurtzite to the rock salt structure can be induced by high hydrostatic pressures. For examples see [93T1, 94T1].

16.4 Problems

1. Typical values of the deformation potential are around 10 eV. What is the shift of the band edges for $\Delta a/a = 10^{-3}$?
2. Do you expect the band-gap shifts due to the lattice deformation by acoustic and by optical phonons to be identical or not? Why?
3. Calculate the diamagnetic shift at $\boldsymbol{B} = 10$ T for excitons characterized by $E_{\mathrm{ex}}^{\mathrm{b}} = 5\,\mathrm{meV}$, $\varepsilon = 15$ and by $E_{\mathrm{ex}}^{\mathrm{b}} = 100\,\mathrm{meV}$, $\varepsilon = 6$, respectively.
4. Calculate the Zeeman splitting for a spin singlet and a triplet exciton for $n_{\mathrm{B}} = 1$, $g_{\mathrm{e}} = 1.6$ and $g_{\mathrm{h}} = 2.2$. What is the difference if at $\boldsymbol{B} = 0$ the two states already show a finite energy splitting δ?
5. Do you expect that the linear Stark effect can occur for $n_{\mathrm{B}} = 1$ and/or for $n_{\mathrm{B}} = 2$ excitons? Why?
6. Calculate in the simplest approximation the Stark shift for excitons with the data given in connection with Problem 3 and for an electric field strength E of $10^3\,\mathrm{Vm}^{-1}$, $10^6\,\mathrm{Vm}^{-1}$, and $10^{-2}\,E_{\mathrm{ex}}^{\mathrm{b}}(ea_{\mathrm{B}})^{-1}$.
7. Calculate the quantum confined Stark effect by perturbation theory for a quantum well with infinitely high barriers and the other data as for the *hh* exciton in GaAs or InAs. Compare with experimental data.

References to Chap. 16

[08V1] W. Voigt, Magneto- und Elektrooptik, Teubner, Leipzig (1908)
[60H1] J.J. Hopfield, J. Phys. Chem. Solids **15**, 97 (1960)
[60T1] D.G. Thomas, J. Phys. Chem. Solids **15**, 86 (1960)
[67G1] E. Gutsche and H. Lange, phys. stat. sol. (b) **22**, 229 (1967)
[68S1] Chr. Solbrig, Z. Phys. **211**, 429 (1968)
[69C1] M. Cardona, Modulation Spectroscopy, Academic, New York (1969)
[69H1] C. Gahwiller and G. Harbeke, Phys. Rev. **185**, 1147 (1969)
[70H1] N. Hase and M. Onuka, J. Phys. Soc. Jpn. **28**, 965 (1970)
[70L1] D.W. Langer et al., Phys. Rev. B **2**, 4005 (1970)
[70M1] E. Mohler, phys. stat. sol. **38**, 81 (1970)
[71L1] H. Lange, phys. stat. sol. (b) **48**, 791 (1971)
[71M1] O.W. Madelung and E. Mollow, Z. Physik A **249**, 12 (1971)
[71S1] B. Stone and Y. Brada, Bull Isr. Phys. Soc. 21 (1971)
[73B1] L.M. Blinov and N.A. Kirchenov, Sov. Phys. Sol. State **8**, 2163 (1973)
[73B2] B.A. Bobylev, A.F. Kravchenko and A.S. Terekhov, Sov. Phys. Semicond. **6**, 1635 (1973)
[73H1] K. Hümmer, phys. stat. sol. (b) **56**, 249 (1973)
[73R1] B.S. Razbirin, I.N. Uraltsev and A.A. Bogdanov, Sov. Phys. Sol. State **15**, 604 (1973)
[73S1] B.O. Seraphin, ed., Modulation Spectroscopy, North-Holland, Amsterdam (1973)
[76B1] B. Buckel, U. Birkholz and K. Ziegler, phys. stat. sol. (b) **78**, K23 (1976)
[76H1] Y. Hamakawa and T. Nishino, in Optical Properties of Solids, New Developments, ed. by B.O. Seraphim, North-Holland, Amsterdam, p 255 (1976)
[78G1] J.A. Gaj et al., Solid State Commun. **28**, 193 (1978) and ibid **29**, 435 (1979)
[79E1] Excitons, K. Cho, ed., Topics Curr. Phys. Vol. **14**, Springer, Berlin, Heidelberg (1979)
[79R1] R. Ranvaud et al., Phys. Rev. B **20**, 701 (1979)
[79T1] H.-R. Trebin, U. Rössler and R. Ranvaud, Phys. Rev. B **20**, 686 (1979)
[81B1] G. Blattner et al., phys. stat. sol. (b) **107**, 105 (1981)
[82A1] Application of High Magnetic Fields in Semiconductor Physics, G. Landwehr, ed., Lecture Notes Phys. Vol. **177**, Springer, Berlin, Heidelberg (1982)
[82B1] G. Blattner et al., Phys. Rev. B **25**, 7413 (1982)
[82H1] H.W. Hölscher, A. Nöthe and Ch. Uihlein, Physica B **117/118**, 2379 (1982)
[82K1] C. Klingshirn in [82A1], p 214
[82L1] M.V. Lebedev and V.G. Lyssenko, Sov. Phys. Sol. State **24**, 1721 (1982)
[84L1] Y.R. Lee and A.K. Romdas, Solid State Commun. **51**, 861 (1984)
[84S1] L. Schultheis and J. Lagois, Phys. Rev. B **29**, 6784 (1984)
[85H1] B. Hönerlage et al., Phys. Rep. **124**, 161 (1985)
[85M1] D.A.B. Miller et al., Phys. Rev. B **32**, 1043 (1985)
[85P1] R. People and J.C. Bean, Appl. Phys. Lett. **47**, 322 (1985)
[85S1] L. Schultheis and C.W. Tu, Phys. Rev. B **32**, 6978 (1985)
[86B1] J.A. Brum, G. Bastard and M. Voos, Solid State Commun. **59**, 561 (1986)
[86G1] J. Gutowski, Solid State Commun. **58**, 523 (1986)
[86M1] D.A.B. Miller, D.S. Chemla and S. Schmitt-Rink, Phys. Rev. B **33**, 6976 (1986)

[87B1] I. Bar-Joseph et al., Appl. Phys. Lett. **50**, 1010 (1987)
[87D1] B.W. Dodson and J.Y. Tsao, Appl, Phys. Lett. **51**, 1325 (1987)
[88G1] J.A. Gaj in [88S1], p 275
[88M1] D.A.B. Miller, D.S. Chemla and S. Schmitt-Rink, Appl. Phys. Lett. **52**, 2154 (1988)
[88S1] Semiconductors and Semimetals **25**, J.K. Furdyna and J. Kossut, eds., Academic Press, New York (1988)
[89G1] J. Gutowski, NATO ASI Ser. B **200**, 139, Plenum, New York (1989)
[89H1] F. Haché, D. Ricard and C. Flytzanis, Appl. Phys. Lett. **55**, 1504 (1989)
[90E1] A.I. Ekimov et al., Journ. of Luminesc. **46**, 83 (1990)
[90L1] M.V. Lebedev and V.G. Lyssenko, phys. stat. sol. (b) **161**, 395 (1990)
[90R1] H. Rossmann et al., phys. stat. sol. (b) **159**, 287 (1990)
[91J1] B.T. Jonker et al., J. Appl. Phys. 69, 6097 (1991)
[91O1] M. v. Ortenberg, Festkörperprobleme/Adv. in Solid State Physics **31**, 261 (1991)
[91S1] S. Schmitt-Rink et al., Appl. Phys. A **53**, 491 (1991)
[92B1] U. Becker et al., J. Crystal Growth **125**, 384 (1992)
[92B2] L.V. Butov, Phys. Rev. B **46**, 12765, 13627 and 15156 (1992)
[92G1] W. Gebhardt, Mater. Sci. Eng. B **11**, 1 (1992)
[92G2] O. Goede and W. Heimbrodt, Festkörperprobleme/Adv. in Solid State Physics **32**, 237 (1992)
[92I1] E.L. Ivchenko and A.A. Kiselev, Sov. Phys. Semicond. **26**, 827 (1992)
[92N1] S. Nomura and T. Kobayashi, Phys. Rev. B **45**, 1305 (1992)
[92O1] D. Oberhauser et al., phys. stat. sol. (b) **173**, 53 (1992)
[92R1] T. Rappen, G. Mohs and M. Wegener, phys. stat. sol. (b) **173**, 77 (1992)
[92S1] S. Schmitt-Rink, Festkörperprobleme/Adv. Solid State Phys. **31**, 243 (1992)
[92W1] H.P. Wagner and H. Leiderer, Festkörperprobleme/Adv. Solid State Phys. **32**, 221 (1992)
[92W2] H.P. Wagner et al., J. Lumin. **52**, 41 (1992)
[92Y1] D.R. Yakovlev, Festkörperprobleme/Adv. in Solid State Physics **32**, 237 (1992)
[93M1] E.E. Mendez in [93O1] of Chap. 1, p 181
[93T1] R. Tommasi, M. Lepore and I.M. Catalano, Solid State Commun. **85**, 539 (1993)
[93W1] U. Woggon et al., Phys. Rev. B **48**, 1979 (1993)
[94F1] J.K. Furdyna, Solid State Electronics **37**, 1065 (1994)
[94G1] J.A. Gaj et al., Phys. Rev. B **50**, 5512 (1994)
[94N1] S. Nomura et al., Phys. Rev. B **49**, 13571 (1994)
[94P1] A.S. Plan et al., Surf. Science **305**, 576 (1994)
[94T1] S. Tolbert et al., Phys. Rev. Lett. **73**, 3266 (1994)
[94Y1] M. Yamaguchi et al., Sol. State Electronics **37**, 839 (1994)
[95B1] L.V. Butov et al., Phys. Rev. B **52**, 12153 (1995)
[95H1] R.M. Hannak et al., Solid State Commun. **93**, 313 (1995)
[95L1] D. Liu, MS Thesis, Boston College (1995)
[95M1] Materials Science Forum **182-184**, H. Heinrich and J.B. Mullin, eds., Trans Tech. Publ. (1995)
[95M2] H. Mariette et al., Annales de Physique, Colloq. C2, Suppl. au n^{o} 3, **20**, C2-143 (1995)

[95O1] M. Oestreich et al., Europhys. Lett. **31**, 399 (1995)
[95S1] U. Siegner et al., phys. stat. sol. (b) **188**, 361 (1995) and Phys. Rev. B **51**, 495 (1995)
[96O1] M. Oestreich, S. Hallstein and S. Rühle, IEEE J. Selected Topics Quant. Electr. **2**, 747 (1996)
[96R1] R. Rinaldi et al., Phys. Rev. Lett. **77**, 342 (1996)
[96W1] H. Weman et al., Phys. Rev. B **53**, 6959 (1996)
[96Y1] D.R. Yakovlev, Comments Cond. Mat. Phys. **18**, 51 (1996)
[97A1] M.V. Artemyev, V. Sperling and U. Woggon, J. Appl. Phys. **81**, 6975 (1997)
[97B1] S. Bar-Ad, Phys. Rev. Lett. **78**, 1363 (1997)
[97H1] M. Hetterich et al., Phys. Rev. B **56**, 12396 (1997)
[97K1] S. Kim et al., Phys. Rev. B **55**, 7130 (1997)
[97O1] Y. Oka, phys. stat. sol. (b) **202**, 795 (1997)
[97P1] Proc. Intern. Conf. on High Magnetic Fields in Semiconductors, G. Landwehr and W. Ossau, eds., Vols. 1 and 2, World Scientific, Singapor (1997)
[97S1] A.A. Sivenko in [97P1], Vol. 2, p 561
[98B1] M. Bayer et al., Phys. Rev. B **57**, 6584 (1998)
[98H1] W. Heller, U. Bockelmann and G. Abstreiter, Phys. Rev. B **57**, 6270 (1998)
[98K1] M Kuno et al., J. Phys. Chem. B **102**, 9245 (1998)
[98L1] W. Langbein et al., J. Opt. Soc. Am. **256**, 86 (1998)
[98O1] H. Ohno, Science **281**, 951 (1998)
[98R1] S. Raymond et al., Phys. Rev. B **58**, R13415 (1998)
[99B1] M. Bayer et al., Phys. Rev. Lett. **82**, 1748 (1999)
[99B2] M. Bayer et al., Phys. Rev. B **60**, R8481 (1999)
[99E1] S.A. Empedocles et al., Adv. Matr. **11**, 1243 (1999)
[99F1] S. Fukatsu et al., J. Luminesc. **80**, 429 (1999)
[99G1] J.A. Gupta, Phys. Rev. B **59**, 10421 (1999)
[99N1] V. Negoita, D.W. Snoke and K. Eberl, Appl. Phys. Lett. **75**, 2059 (1999) and Phys. Rev. B **60**, 2661 (1999)
[99P1] J. Puls et al., Phys. Rev. B **60**, R16303 (1999)
[99P2] S. Petillon et al., J. Crystal Growth **201/202**, 453 (1999)
[99W1] H. Weman et al., Appl. Phys. Lett. **74**, 2334 (1999)
[00D1] T. Dietl et al., Science **287**, 1019 (2000)
[00F1] P.W. Fry et al., Phys. Rev. Lett. **84**, 733 (2000)
[00K1] M. Keim et al., J. Appl. Phys. **88**, 7051 (2000)
[00L1] A. Lorenzoni et al., phys. stat. sol. (a) **178**, 239 and 243 (2000)
[00O1] Y. Oka et al., Proc. SPIE **4086**, 62 (2000)
[00P1] C. Penn et al., Phys. Rev. B **61**, 13055 (2000)
[00T1] N. Takahashi et al., J. Appl. Phys. **87**, 6569 (2000)
[01D1] A. Dinger et al., Phys. Rev. B **64**, 225310 (2001)
[01F1] J.J. Finley et al., Phys. Rev. B **63**, 073307 (2001)
[01I1] H. Ikada et al., Physica E **10**, 373 (2001)
[01K1] R. Kotlyar et al., Phys. Rev. B **63**, 085310 (2001)
[01K2] F.V. Kyrychenko and J. Kossut, Physica E **10**, 378 (2001)
[01K3] T.D. Krauss et al., J. Phys. Chem. B **105**, 1725 (2001)
[01K4] M. Kuno et al., J. Chem. Phys. **115**, 1028 (2001)
[01M1] H. Mino et al., Physica B **289**, 421 (2001)

[01O1] M. Oestreich et al., Adv. in Solid State Physics **41**, 173 (2001)
[02B1] F.J. Brieler et al., Chem. Eur. J. **8**, 185 (2002)
[02C1] S. Coe et al., Nature **420**, 800 (2002)
[02D1] T. Dietl, Semicond. Science and Technol. **17**, 377 (2002)
[02G1] O. Gogolin, M. Schmidt and M. Dremel, private communication (2002)
[02L1] W.L. Lambrecht et al., Phys. Rev. B **65**, 075207 (2002)
[02M1] J.R. Madureira et al., Phys. Rev. B **66**, 075332 (2002)
[02M2] N. Miura et al., Physica E **13**, 263 (2002)
[02S1] Semiconductor Spintronics, H. Ohno, ed., Semicond. Science and Technol. **17**, Issue 4 (2002)
[02S2] S. Sonda, IEEE Trans. Mag. **38**, 2859 (2002)
[02S3] K. Sato and H. Katayama-Yoshida, Semicond. Science and Technol. **17**, 367 (2002)
[02W1] O. Wolst et al., Physica E **13**, 283 (2002) and ibib. E **17**, 554 (2003)
[03C1] G.P. Collins, Scientific American **288**, March, p 13 (2003)
[03C2] S.F. Chichibu et al., J. Appl. Phys. **93**, 756 (2003)
[03D1] T. Dietl, Europhys. News **34**, 216 (2003) and phys. stat. sol. b **240**, 433 (2003)
[03D2] T. Dietl and H. Ohno, MRS Bull **28**, 714 (2003)
[03N1] N. Naka and N. Nagasawa, phys. stat. sol. (b) **238**, 397 (2003)
[03P1] R.T. Phillips et al., phys. stat. sol. (b) **238**, 601 (2003)
[03R1] E.I. Rashba and A.L. Efros, Appl. Phys. Lett. **83**, 5295 (2003)
[03S1] R.T. Senger et al., Appl. Phys. Lett. **83**, 5425 (2003)
[04A1] S. Adachi, EXCON 04, Cracow, to be published in J. Lumin.
[04G1] G.J. Goldsmith in [81A1]l of Chap. 1, in press
[04M1] B. Meyer et al., phys. stat. sol. (b) **241**, 231 (2004)

17

From Cavity Polaritons to Photonic Crystals

In this chapter we return briefly to the concept of a Fabry–Perot resonator in the form of a (micro) cavity and then proceed to the cavity polaritons as a mixed state between a resonance in a solid (these are generally exciton resonances in quantum wells, wires or dots) and a cavity resonance. From there we reach, via different paths, the presently very active and potentially application-relevant field of photonic crystals with a subspecies known as photonic band gap materials.

17.1 Cavity Polaritons

We start from an empty cavity and introduce then the concept of a cavity polaritons.

17.1.1 The Empty Resonator

A one-dimensional Fabry–Perot resonator, etalon or cavity is, in the simplest case, an arrangement of two plane-parallel, lossless mirrors with a reflectivity close to unity, i.e., $R \leq \underline{1}$ at a distance d (see Fig. 17.1a). Concerning the nomenclature, it can be stated that all the above names mean physically essentially the same thing. Sometimes the following distinctions are made.

A Fabry–Perot resonator (also known as Perot–Fabry) often has a length of many wavelengths like, e.g., the resonators of many gas, dye or solid-state lasers with a consequently narrow spectral longitudinal mode spacing. A cavity often has a width of only one or a few multiples of $\lambda/2$. An etalon is often made of a plane-parallel slab of transparent matter like Al_2O_3 with dielectric mirrors on both sides to enhance the reflectivity.

We consider here only the case of plane-parallel surfaces. But Fabry–Perot resonators may also have one or two curved, generally spherical mirrors. This topic is generally treated in books or articles on laser physics or laser design.

The cavities may be also (micro)spheres ([01A1, 02A1, 02M1, 03M2] or Ref. [03W1] of Chap. 15) or even cylinders [02K1].

An incident plane, monochromatic wave will be partly reflected and partly transmitted at each of the two mirrors. If the coherence length is $\gg d$, then the partial waves can interfere. In most cases this interference will be destructive, resulting in a reflectivity of the whole structure $R_{\rm tot} \approx 1$ and a transmission $T_{\rm tot} = 1 - R_{\rm tot} \approx 0$ and a very low field amplitude in the resonator. See Fig. 17.1b. There are, as already outlined in Sect. 3.1.6, special situations when an integer number of half waves fits into the resonator, i.e., if

$$m\frac{\lambda}{2} = d \quad \text{or} \quad k_\perp = m\frac{\pi}{d} \quad m = 1, 2, 3, \ldots . \tag{17.1}$$

In this case all partial waves in the resonator interfere constructively and the field amplitude in the resonator can considerably exceed the incident one and the transmission of the whole arrangement is close to unity, although the reflectivity of every single mirror is also close to unity. If the finesse (see Sect. 3.1.6) is sufficiently high, this constructive interference results in the narrow spikes (dips) in the spectra of $T_{\rm tot}$ ($R_{\rm tot}$) as shown in Fig. 17.1b. This situation is shown for $m = 2$ in Fig. 17.1a.

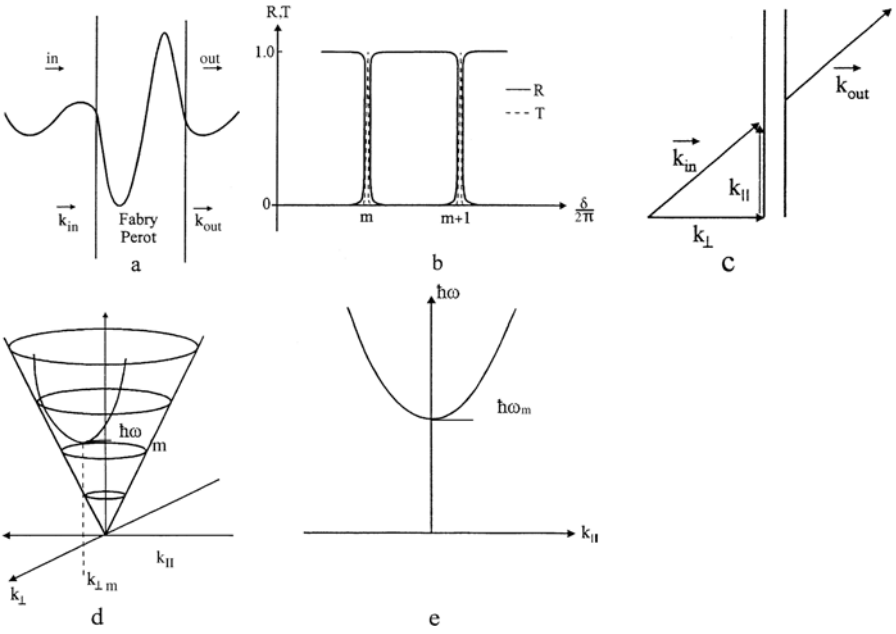


Fig. 17.1. (**a**) Schematic drawings of the incident, internal and transmitted fields in a Fabry–Perot resonator for normal incidence. (**b**) The spectra of total transmission T and reflection R as a function of the phase shift δ per roundtrip. (**c**) The decomposition of the incident and transmitted wave vectors for oblique incidence. (**d,e**) The construction of the dispersion relation from the light cone [01K1]

We can reformulate ((17.1)) by saying that these so-called eigenmodes of a Fabry–Perot resonator or cavity or etalon occur for the normal component of the incident $\boldsymbol{k}$-vector

$$k_{\perp,\,m} \;=\; m\,\frac{\pi}{d}; \quad m \;=\; 1,\; 2,\; 3,\; \ldots \tag{17.2}$$

or that the modes are equally spaced on the $\boldsymbol{k}_\perp$ axis with a distance

$$\Delta k_\perp = k_{\perp,m+1} - k_{\perp,m} = \frac{\pi}{d}\,. \tag{17.3}$$

For oblique incidence, the wave vector of the incident beam can be decomposed in components parallel and normal to the resonator as shown in Fig. 17.1c:

$$\boldsymbol{k} = \boldsymbol{k}_\parallel + \boldsymbol{k}_\perp\,. \tag{17.4}$$

The normal component has to fulfill (17.2) for the resonator modes, while a conservation law holds for $\boldsymbol{k}_\parallel$ in reflection and transmission, due to the translational invariance of the problem in the plane of the resonator (see Noether's theorem in Sect. 3.1.3).

For photons, i.e., for the quanta of light in vacuum, the dispersion is a cone when we plot $\hbar\omega$ over a two-dimensional $\boldsymbol{k}$-plane. The cross section of this cone with a fixed $k_{\perp\,m}$ is a parabola as shown in Fig. 17.1d. The photon energy $\hbar\omega_m$ belonging to $\mathrm{k}_{\perp\,m}$ is

$$E_m = \hbar\omega_m = \hbar c k_{\perp m}\,. \tag{17.5}$$

This energy corresponds to the minimum of the dispersion relation of a resonator as shown in Fig. 17.1e.

If one tilts a Fabry–Perot resonator in a parallel light beam away from normal incidence, i.e., away from $k_\parallel = 0$ the photon energy of the transmitted mode shifts of higher values, i.e., to shorter wavelength. This effect is slightly counterintuitive, but can be easily verified when viewing an incandescent lamp through a line interference filter and tilting this filter.

If a semiconductor physicist sees a parabolic dispersion relation, the effective mass concept comes immediately to mind. The effective mass connected with the curvature is given by (see Sect. 8.5):

$$1/m_{\mathrm{eff}} = \frac{\partial^2}{\hbar\partial k^2}\omega\,(k) \tag{17.6}$$

and it turns out that m_{eff} is something like the "rest mass" of the light quanta of the resonator modes in the sense of (17.7):

$$m_{\mathrm{eff}}c^2 = \hbar\omega_{m,k_\parallel=0}\,. \tag{17.7}$$

For a more detailed treatment of this topic, also considering wave guides and resonators with confinement in more than one direction, see [03B1] and the references therein.

17.1.2 Cavity Polaritons

In the next step we bring some matter into the resonator, which has an eigenmode at a certain frequency $\hbar\omega_{\mathrm{M}}$. Since such a resonance is generally connected with some absorption, and since absorption deteriorates the finesse of a resonator (see, e.g., Sect. 3.1.6), only a small overall oscillator strength is recommended, e.g., only a thin layer. Therefore one places generally one or a few quantum wells in the resonator at the position(s) of an antinode of the electric field with an eigenfrequency of the lowest free exciton resonance close to a resonance frequency of the resonator.

Some early examples for such situations are, e.g., [92W1, 94H1]. More recent work can be found in [95K1, 98Q1, 98S1, 99K1, 03K1, 04G1] and the references given therein. Recently quantum wires and quantum islands have also been incorporated into resonators [99A1, 99C1]. For microsphere resonators see, e.g., [01A1, 01A2, 02K1, 02M1, 03M1, 03M2]. After this short excursion to the literature, we return to the physics of the planar resonator. In Fig. 17.2a we show schematically the dispersion relation of a resonator mode $\hbar\omega_{FP}\left(\boldsymbol{k}_{\parallel}\right)$ and of an excitation in matter, e.g., of the hh exciton resonance in a quantum well $\hbar\omega_{M}\left(\boldsymbol{k}_{\parallel}\right)$.

The curvature of the latter is negligible compared to the one of $\hbar\omega_{\mathrm{FP}}$. If both modes do not couple to each other, their dispersion relations may cross as indicated by the dashed lines. If there is a finite coupling, we observe once again the non- (or avoided) crossing behavior (see Chap. 5) and the dispersion describes mixed states of the resonator or cavity mode with an exciton resonance of the quantum well, the so-called cavity polaritons.

In Fig. 17.2b we show luminescence spectra obtained under band-to-band excitation as a function of the angle of observation Θ relative to the normal of the cavity. The variation of this angle corresponds to a variation of $\boldsymbol{k}_{\parallel}$. Both branches of the cavity polariton show up in luminescence. This fact shows, that we really observe the quanta of a mixed polariton state. A resonator alone would not luminesce at $\hbar\omega_{\mathrm{FP}}$ if illuminated with light of a considerably higher photon energy.

In Fig. 17.2d the dispersion relations deduced from the luminescence data are shown together with the calculation as a dashe-dotted line for three samples with slightly different detuning, together with calculations without coupling, beautifully verifying the above statements. Figure 17.2c finally shows schematically the layout of the sample. The cavity has a width of $3\lambda/2$ and three quantum wells are placed at the positions of the two antinodes. The mirrors are dielectric Bragg mirrors (see below). The calculated buildup of the electric field in these Bragg reflectors is also shown.

The above concept is valid if a resonator mode with low index m coincides with an exciton (or other) resonance.

If the resonator is made wider and wider, the modes start at decreasing photon energy and their spacing decreases until one approaches a bulk-like situation. See, e.g., [95Ü1, 98N1].

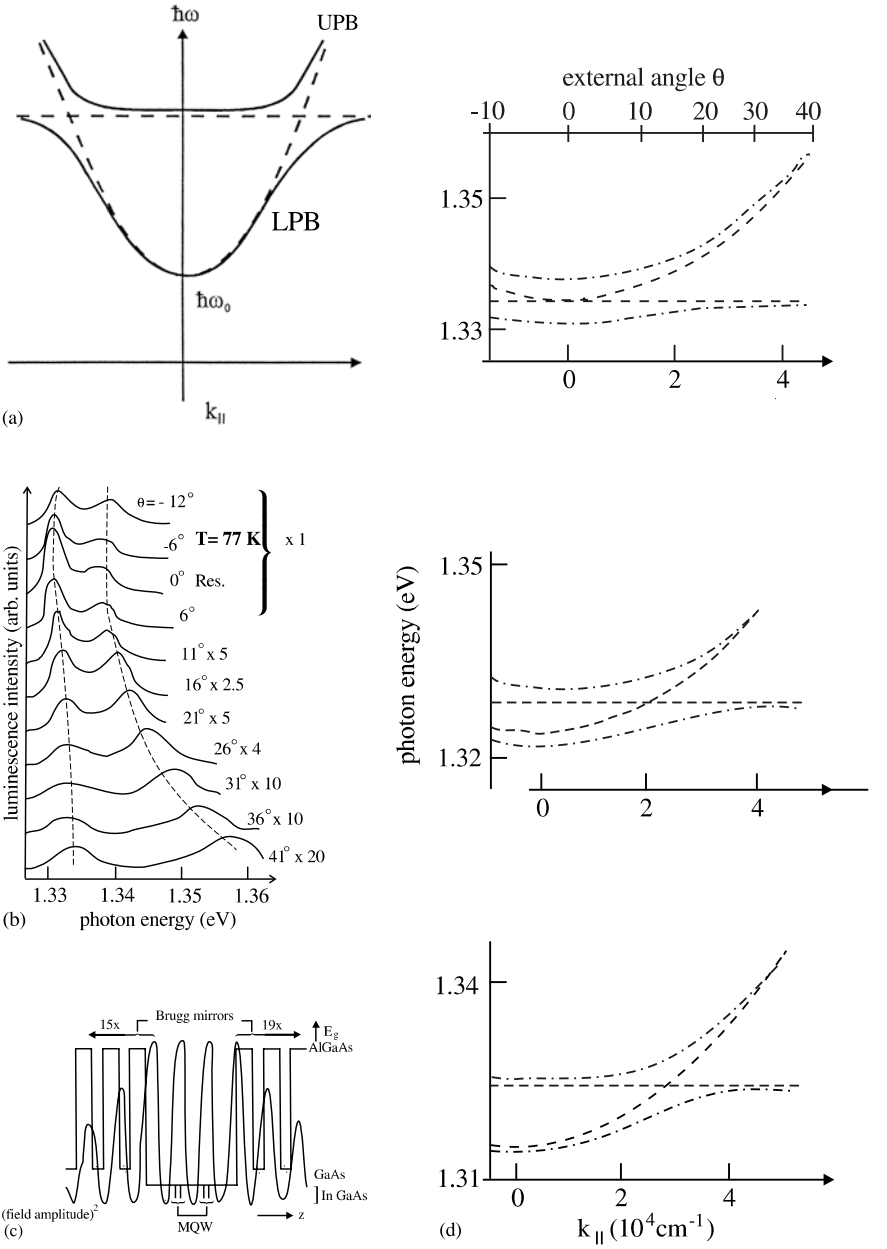


Fig. 17.2. (**a**) Schematic drawing of the dispersion relation of a microcavity and of an exciton resonance without (*dashed line*) and with coupling (*solid line*) LPB and UPB stand for lower and upper polariton branch. Compare to Chaps. 5, 13 and 15. (**b**) Experimental luminescence spectra for various external angles of incidence of the structure shown in (**c**) and the measured and calculated data with coupling (-·-·-) and the calculated dispersion witout coupling (- - -). According to [94H1]

If the resonator is, on the other hand, too narrow, its eigenmodes shift energetically to above the exciton resonances.

A further necessary prerequisite for the appearance of cavity polaritons is the following. The dephasing time of the excitation in matter T_2 must be longer than the roundtrip time of the resonator. Otherwise, the coupled modes can not develop. Such a situation may occur for increasing excitation (or temperature), since T_2 becomes shorter, e.g., by excitation induced dephasing via exciton-exciton or exciton-electron collisions (see [99K1, 03H1, 04G1] or Chap. 23) or the transition to an electron–hole plasma (see Chap. 21). It should be mentioned, however, that lasing from microcavity modes can be observed. We come back to this point in Chap. 22.

17.2 Photonic Crystals and Photonic Band Gap Structures

We introduce first the basic concepts of photonic crystals using a certain analogy between the Schrödinger equation and Maxwell's equations, then we continue with various examples and approaches.

17.2.1 Introduction to the Basic Concepts

The concept of photonic crystals has been introduced by the pioneering papers of Yablonovitch and John [87J1, 87Y1] although the first experimental realizations date many years earlier, as it occurs frequently in science and as we shall see below.

For some recent reviews of the topic see [95J1, 96M1, 96P1, 99B1, 01B1, 01S1, 03B1, 03P1] and the references given therein. Recently scientists started to use also metals instead of semiconductors in photonic structures. Though this is not a topic of this book we give at least one Ref. [03C1].

A first approach to a one-dimensional photonic crystal or photonic band gap material can be found via the understanding of dielectric or Bragg mirrors. A photonic band gap material is, in analogy to the band gap for electrons in crystals a material that has in a certain energy interval no propagating light modes.

It has been long known that a high reflectivity very close to unity is equivalent to a gap caused by the absence of a propagating mode and can be achieved over a certain frequency interval (the photonic band gap) if one produces a stack of layers of different refractive indices n_{I} and n_{II}, where each layer has a thickness d_i equal to a quarter wavelength, i.e.,

$$d_i = \frac{\lambda_i}{4} = \frac{c\pi}{2n_i\omega_0} \quad i = \mathrm{I}, \mathrm{II}\,. \tag{17.8}$$

See also Fig. 17.2d. The partial waves reflected at every interface interfere under the condition (17.8) in a certain interval $\Delta\omega$ of frequencies around ω_0 in

a way that high reflectivity occurs and that there is no propagating mode, i.e., there is a photonic energy gap. The spectral width $\Delta\omega$ of this photonic band gap increases with increasing difference of the refractive indices $|n_{\mathrm{I}} - n_{\mathrm{II}}|$ and the reflectivity converges to unity with an increasing number of layer pairs, provided that these layers and the interfaces are lossless.

An example of such a structure and its reflectivity are shown schematically in Fig. 17.3c,f. These things have been known for decades and can be found in textbooks like [89T1].

Now, we approach the problem from another point of view, recalling what we have learned for the properties of electrons in a periodic potential in Chap. 8.

The Schrödinger equation reads in its general form

$$-\frac{\hbar^2}{2m}\Delta\Phi(\boldsymbol{r},t) + V(r)\,\Phi(\boldsymbol{r},t) = -\frac{\hbar}{i}\frac{\partial}{\partial t}\Phi(\boldsymbol{r},t)\,. \tag{17.9a}$$

We deduced in Sect. 2.3 from Maxwell's equations for the electric field strength $\boldsymbol{E}(r,t)$

$$\Delta\boldsymbol{E} - \frac{1}{c^2}\varepsilon(\omega)\,\mu(\omega)\,\ddot{\boldsymbol{E}} = 0\,. \tag{17.9b}$$

If we consider stationary states with a $e^{-i\omega t}$ time dependence, real V and ε, nonmagnetic material ($\mu(\omega) \equiv 1$) and use the approximation of a dispersionless dielectric function $\varepsilon \neq \varepsilon(\omega)$ and $n^2 = \varepsilon$, which always applies over a sufficiently small frequency interval, we obtain

$$\Delta\Phi(\boldsymbol{r}) + \frac{2m}{\hbar^2}(E - V)\,\Phi(\boldsymbol{r}) = 0 \tag{17.9c}$$

and

$$\Delta\boldsymbol{E}(\boldsymbol{r}) + \frac{\omega^2}{c^2}n^2\boldsymbol{E}(\boldsymbol{r}) = 0\,. \tag{17.9d}$$

Evidently there is a close analogy between (17.9c) and (17.9d). There is a second derivative with respect to space. The term $E - V$ corresponds to n^2 and since the eigenenergy E is simply a constant for the stationary states, in which we are presently only interested in, we can identify in this analogy as essentially the potential V with n^2 or ε.

There are, however, differences: $\Phi(\boldsymbol{r})$ is a scalar quantity $\boldsymbol{E}(\boldsymbol{r})$ a vector field, the eigenenergy $E = \hbar\omega$ appears linearly in (17.9c) while (17.9d) is quadratic in ω. These similarities and differences between the two equations allow us to predict similar but not identical properties for electrons and light. We consider first the situation of spatially constant $V(\boldsymbol{r}) = V_0$ and $n(\boldsymbol{r}) = n_0$.

Then the solutions are plane waves (for $E \geq V_0$ or $n^2 > 0$) with the dispersion relations typical for massive and massless particles, as shown in Fig. 17.3a,b, i.e.,

$$E = \frac{\hbar^2 k^2}{2m} \quad \text{for} \quad E \geq V_0 \tag{17.10a}$$

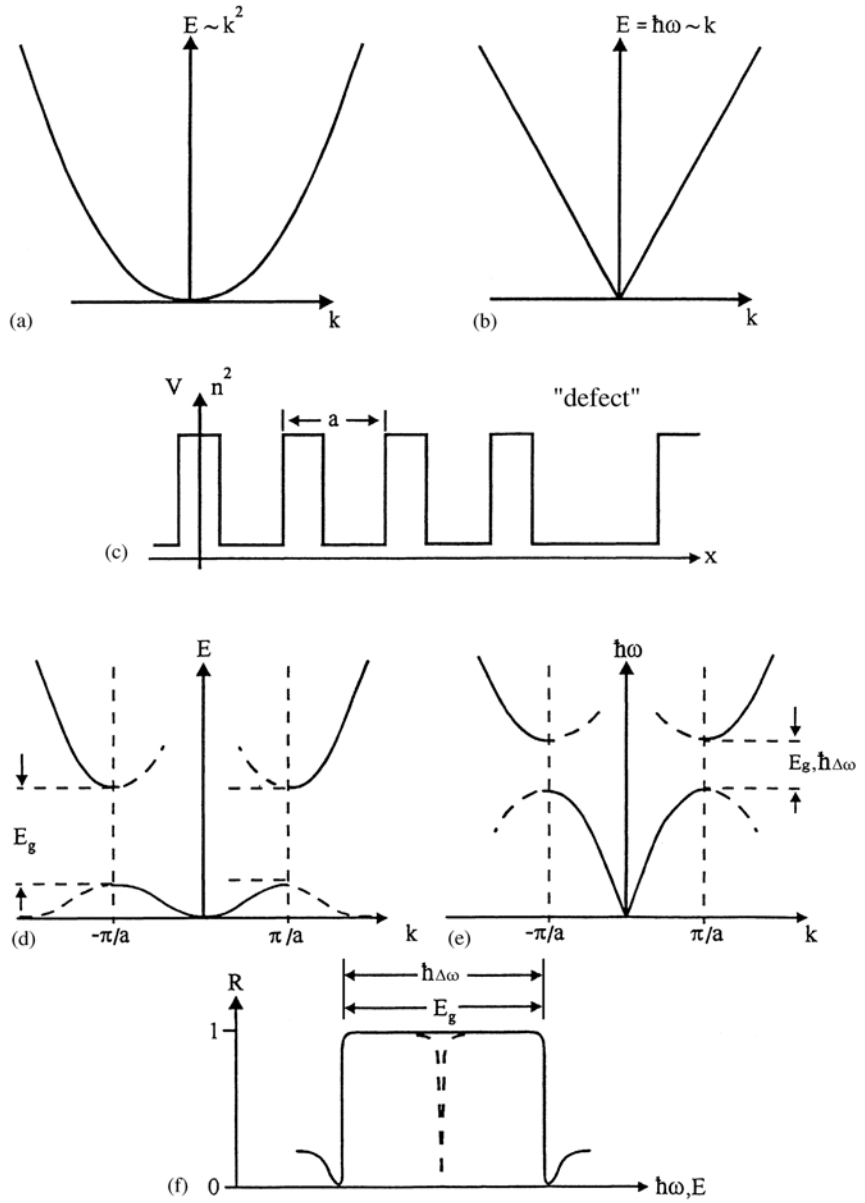


Fig. 17.3. (**a**) The parabolic dispersion of electrons in a constant potential V_0. (**b**) The linear dispersion of photon-like polaritons in a material with constant refractive index n_0. (**c**) The periodic potential of the Kronig-Penney model with a "defect". (**d,e**) The resulting dispersion relations and band gaps for electrons and light without defects. (**f**) The reflectivity of a Bragg mirror and of a Fabry–Perot cavity formed by two Bragg mirrors with a cavity (*solid* and *dashed lines*, respectively). All drawings are schematic according to [01K1]

and

$$E = \hbar\omega = \hbar\frac{c}{n_0}k_{vac} = \hbar ck\,. \tag{17.10b}$$

The difference comes from the linear and quadratic E or ω dependencies mentioned above, which in turn go back to the first and second time derivatives in (17.9a,b).

Now we allow for a periodic variation of $V(\boldsymbol{r})$, e.g., in the sense of a periodic square potential (the so-called Kronig–Penney potential) of Fig. 17.3c. The results are as known from Chap. 8. The appearance of band gaps, the concepts of the periodicity of $E(\boldsymbol{k}) = E(\boldsymbol{k} + \boldsymbol{G})$, the quasi momentum $\hbar\boldsymbol{k}$ are shown schematically in Fig. 17.3d.

We now consider the case of a periodic variation of $n(\boldsymbol{r})$. If we want to use (17.9d) we are limited to small spatially periodic variations of n (e.g., in the sense of nearly free electrons in Sect. 8.7 for the following reason.

In the derivation of (17.9d) one comes across the term

$$\nabla \times (\nabla \times \boldsymbol{E}) = \nabla\,(\nabla \cdot \boldsymbol{E}) - \Delta\boldsymbol{E}\,. \tag{17.11a}$$

If and only if ε is spatially constant, we can use the chain of arguments

$$0 = \rho = \nabla\boldsymbol{D} = \nabla\,(\varepsilon\boldsymbol{E}) \stackrel{!}{=} \varepsilon\nabla \cdot \boldsymbol{E} = 0\,. \tag{17.11b}$$

If $\varepsilon = \varepsilon\,(\boldsymbol{r})$ we are either left with an additional term $\boldsymbol{E}\nabla \cdot \varepsilon\,(\boldsymbol{r})$, which is only small for small variations of $\varepsilon\,(\boldsymbol{r})$ (see above), but is in any case also periodic in space. Furthermore it is possible in one- and twodimensional arrays to select orientations with $\boldsymbol{E} \perp \boldsymbol{\nabla}\varepsilon(\boldsymbol{r})$. Otherwise, we have to use the general from (17.11c)

$$\nabla \times (\nabla \times \boldsymbol{E}\,(\boldsymbol{r})) - \frac{\omega^2}{c^2}\varepsilon\,(\boldsymbol{r})\,E\,(\boldsymbol{r}) = 0 \tag{17.11c}$$

with $\nabla\,(\varepsilon\,(\boldsymbol{r})\,\boldsymbol{E}) = 0$, which is correct for every $\varepsilon\,(\boldsymbol{r})$ but is not Hermitian with respect to the usual scalar product. Alternatively one can use the equation for the magnetic field $\boldsymbol{H}$

$$\nabla \times \left(\frac{1}{\varepsilon\,(\boldsymbol{r})}\nabla \times \boldsymbol{H}\right) - \frac{\omega^2}{c^2}\boldsymbol{H} = 0 \tag{17.11d}$$

with $\nabla\boldsymbol{B} = \nabla\boldsymbol{H} = 0$ since we assumed nonmagnetic materials, i.e., $\mu = 1$. Equation (17.11d) is Hermitian.

If we introduce $\bar{\varepsilon}$ as the spatial average value of $\varepsilon\,(\boldsymbol{r})$ then the variations around this average value are given by $\varepsilon\,(\boldsymbol{r}) - \bar{\varepsilon}$ and we end up with

$$\nabla \times (\nabla \times \boldsymbol{E}) - \frac{\omega^2}{c^2}\,(\hat{\varepsilon}\,(\boldsymbol{r}) - \bar{\varepsilon})\,\boldsymbol{E} = -\frac{\omega^2}{c^2}\boldsymbol{E}\,, \tag{17.11e}$$

which makes the analog to the eigenvalue (17.9c) even more plausible:

$$\nabla \cdot \nabla\Phi\,(\boldsymbol{r}) - \frac{2m}{\hbar}V\,(\boldsymbol{r})\,\Phi\,(\boldsymbol{r}) = E\Phi\,(\boldsymbol{r})\ . \tag{17.11f}$$

The vector character of $\boldsymbol{E}(\boldsymbol{r})$ compared to the scalar field $\Phi(\boldsymbol{r})$ causes the $\nabla \times \nabla \times$ term in (17.11e) instead of the $\nabla \cdot \nabla$ in (17.11c,i).

It should be noted that $\boldsymbol{E} \perp \boldsymbol{H}$ is still valid, but due to the variations or steps in $\varepsilon(\boldsymbol{r})$, $\boldsymbol{E}$ is not necessarily perpendicular to $\boldsymbol{k}$.

Having now played, to some extent, with the equations for $\boldsymbol{E}$ and $\boldsymbol{H}$, we now present some analogies and differences assuming a periodic variation of $n^2(\boldsymbol{r}) = \varepsilon(\boldsymbol{r})$, e.g., as in Fig. 17.3c. The results include again the concept of quasi momentum $\hbar\boldsymbol{k}$ and the periodicity $\omega(\boldsymbol{k}) = \omega(\boldsymbol{k} + \boldsymbol{G})$ with $\boldsymbol{G}$ a vector the reciprocal lattice. The differences to the electronic case are a linear (photon-like) increase of the dispersion close to the origin. Furthermore, we have to calculate the band structure for two orthogonal orientations of $\boldsymbol{E}$ as soon as we deviate from normal incidence on a one-dimensional structure.

As a last aspect of this introductory treatment, we consider the following point.

A defect in the periodic potential as indicated on the right hand side of the Kronig–Penney model of Fig. 17.3c may result in an energetically narrow, localized state in the gap if the defect has a suitable shape, e.g., concerning width and/or depth.

In a similar way, a "defect" in the periodicity of $n(\omega)$ may result in an energetically narrow defect state in the photonic band gap of a Bragg-mirror if its parameters are adequately chosen. This defect state is nothing but the Fabry–Perot mode, which allows propagation of light in a narrow frequency interval (Fig. 17.3f).

The light quanta in a photonic crystal with periodicity

$$n^2(\boldsymbol{r}) = \varepsilon(\boldsymbol{r}) = \varepsilon(\boldsymbol{r} + \boldsymbol{R}) \tag{17.12a}$$

with

$$\boldsymbol{R} = a_1 n_1 \boldsymbol{a}_1 + n_2 \boldsymbol{a}_2 + n_3 \boldsymbol{a}_3\,, \tag{17.12b}$$

as in Sects. 7.2 and 8.1, are actually also a superposition of a polarization and an electromagnetic wave. Therefore the terms polaritonic crystal and polaritonic band structure would be more adequate. However, the terms photonic crystals and band structures have been introduced and are established. So there is no sense in trying to change them and we continue to use them here.

17.2.2 Realization of Photonic Crystals and Applications

While dielectric or Bragg mirrors have been known for decades and are now considered an example of a one-dimensional photonic structure, artificial two- and three-dimensional systems have been developed only recently and are based on the pioneering work mentioned at the beginning of this chapter. This means that we see here the opposite trend as compared to semiconductors where the development started with three-dimensional or bulk material and proceeded to quasi two-, one- and zero-dimensional structures. However, we shall also see a similar approach in Sect. 17.3, and should mention that

there is a precious gemstone namely (noble) opal that consists of a three-dimensional periodic array of small SiO_2 spheres forming a natural photonic crystal. We come back to this below but start first with two-dimensional structures.

The technology of processing Si is highly developed. Using patterning and selective etching, among other methods, it is possible to produce periodic (hexagonal) arrays of holes in Si that have a spacing or lattice constant on the order of 1 μm and some are 10 μm deep (see Fig. 17.4). Such structures form evidently a two-dimensional photonic crystal. This structure has a true two-dimensional photonic band gap, i.e., a band gap that exists for all values and directions of $\boldsymbol{k}_{\parallel}$ centered around a frequency ω_0 given by

$$\omega_0 \approx 0.4c\frac{2\pi}{a} \tag{17.13}$$

where a and c are the lattice constant and the vacuum speed of light, respectively. Obviously, the band structure depends on the polarization of the light field and it turns out that the photonic gap is narrower for the electric field polarized parallel to the tubes than for the orthogonal polarization.

Obviously a in (17.13) has to be chosen such that $\hbar\omega_0$ falls in the transparent region of Si, i.e., $\hbar\omega_0 < E_g^{\mathrm{Si}}$.

Details of such structures are found, e.g., in [97F1, 98B1, 99B1, 99B2] and the references therein.

Here we concentrate on only a few aspects. A chain of defects, missing holes, act for frequencies in the photonic gap as a wave guide. The insert in Fig. 17.4a shows a simple linear wave guide. Since the light field cannot radiate from this wave guide into the two-dimensional photonic structure in the spectral region of the photonic gap, curvatures of wave guides can be realized (Fig. 17.4a) that have much narrower curvatures than classical optical fibers can have. A realization of a beam splitter is shown in Fig. 17.4b and a resonator in Fig. 17.4c. The central defect "localizes" light field. It is weakly contacted by the two wave guides.

In a similar way, a ridge wave guide of Si on SiO_2 has been converted in a Fabry–Perot resonator by a periodic arrangement of holes, which act as

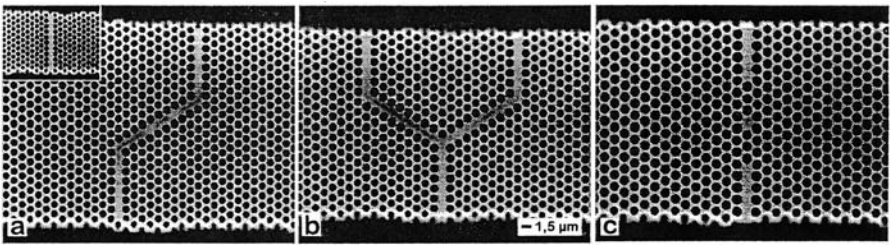


Fig. 17.4. (**a**) A hexagonal periodic array of cylindrical holes in Si with a wave-guide (**b**) a beam splitter (**c**) a resonator [98B1, 99B2]

Bragg mirrors and the central part as a cavity. See [97F1, 98B1, 98B2, 99B1] and references therein.

The application potential of such structures is obvious, allowing the integration of wave guides with dimensions of only a few tens of μm on a Si chip. The drawback is presently still that these structures have very high insertion losses. Furthermore, they are much larger than the building blocs in highly integrated electronic chips, despite the fact, that they are smaller than classical waveguide structures.

Now we discuss three-dimensional structures that have a true photonic band gap for all polarizations and directions $\boldsymbol{k}$-vectors within a certain energy interval.

Different lattices and ways to create them have been proposed and are considered. An early example is the so-called Yablonovite. Three sets of parallel holes are drilled in a solid. The angle between the axes of the holes is 120 ° and their inclination to the normal of the surface is 35.26 °. If the refractive index of the material is sufficiently large, one obtains a photonic gap. For details see, e.g., [91Y1].

Drilling of holes under oblique, well-defined directions is not yet possible for distances comparable to the wavelength of light, but it has been shown for longer wavelengths that the concept works.

A modified version uses the interference of three laser beams falling obliquely on a photoresistant layer together with a reference beam impinging normal on the surface. The interference creates a three-dimensional intensity pattern that results after development of the photoresistant layer in a photonic crystal. The present status of this technique is found, e.g., in [03M1].

Still another idea is so-called wood piles. These are semiconductor "rods" (with diameters on the order of 1 μm) placed in subsequent layers orthogonally on top of each other. Again, one expects a photonic gap for sufficiently high index contrast $|n_{\mathrm{I}} - n_{\mathrm{II}}|$.

A system that is presently under intense theoretical and experimental investigation is inverted opals [98B2, 01B2, 01R1, 02G1, 02H1, 03B2, 03W1]. An opal itself is a periodic arrangement of SiO_2 spheres forming a diamond lattice. The spheres are surrounded by air (or water). The difference of the refractive indices in opal is too small to obtain a complete gap, as can be seen, e.g., in the calculated density of states in Fig. 17.5c.

Natural opal, however, has gaps for certain directions in $\boldsymbol{k}$-space. The position and width depends on the orientation of $\boldsymbol{k}$ relative to the lattice. This fact, together with differently oriented domains, explains the changing colors of opal in different places and for varying viewing angles. The idea of an inverted opal is now to fill or impregnate the voids between the SiO_2 spheres by Si, which below E_{g} has a rather high refractive index above 3, and to remove then the SiO_2 spheres by selective etching, resulting in a periodic arrangement of air-spheres surrounded by Si. The band-structure calculation in Fig. 17.5a shows the gap around $\omega a/2\pi c = 0.8$. The complexity of the band-structure, similarly to the electron case, is to a large extent due to the

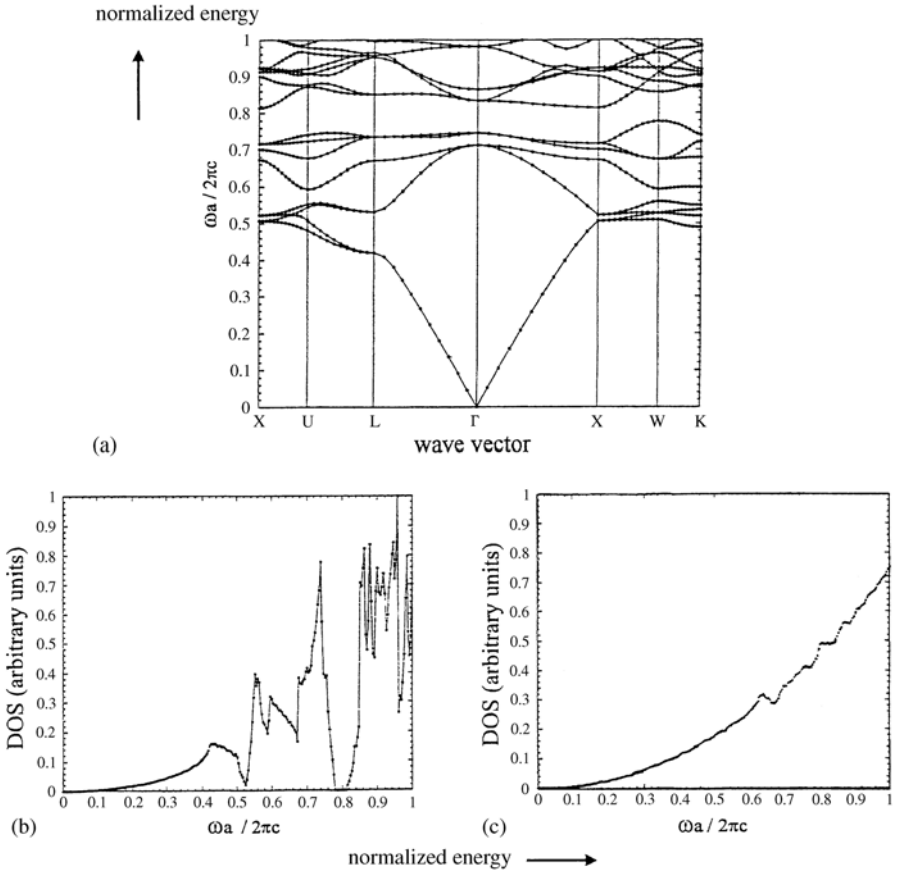


Fig. 17.5. (**a**) The calculated photonic band-structure of close-packed air spheres in Si, a so-called inverted opal. (**b**) The resulting density of states (DOS), showing clearly a gap around $\omega a/2\pi c \approx 0.8$ for this system and for (**c**) SiO_2 spheres in air corresponding to natural opal [99J1]

back folding of the extended scheme into the first Brillouin zone (see Chap. 8). The calculated density of states starts quadratically as expected for a linear dispersion relation in three dimensions, but then develops a rather complex structure, which is much more "peaky" compared to the one of natural opal in Fig. 17.5c.

17.3 Photonic Atoms, Molecules and Crystals

In this section we present still another approach to photonic crystals, starting in this case with "photonic atoms" i.e., three-dimensional microcavities and coupling them to create molecules or extended periodic structures. For reviews

and some original work see, e.g., [97R1, 98B3, 98G1, 98G2, 99B3, 99G1, 01B1, 01B3, 01G1, 02G2, 03S1].

The concept of a photonic-atom is explained in Fig. 17.6a. A quantum well is grown in a cavity consisting of two stacks of dielectric Bragg mirrors as explained above for cavity polaritons. The luminescence is shown in the lowest trace of Fig. 17.6b.

We now see the well-known effect that the emission maximum is shifted away from the exciton resonance of the quantum well without a cavity.

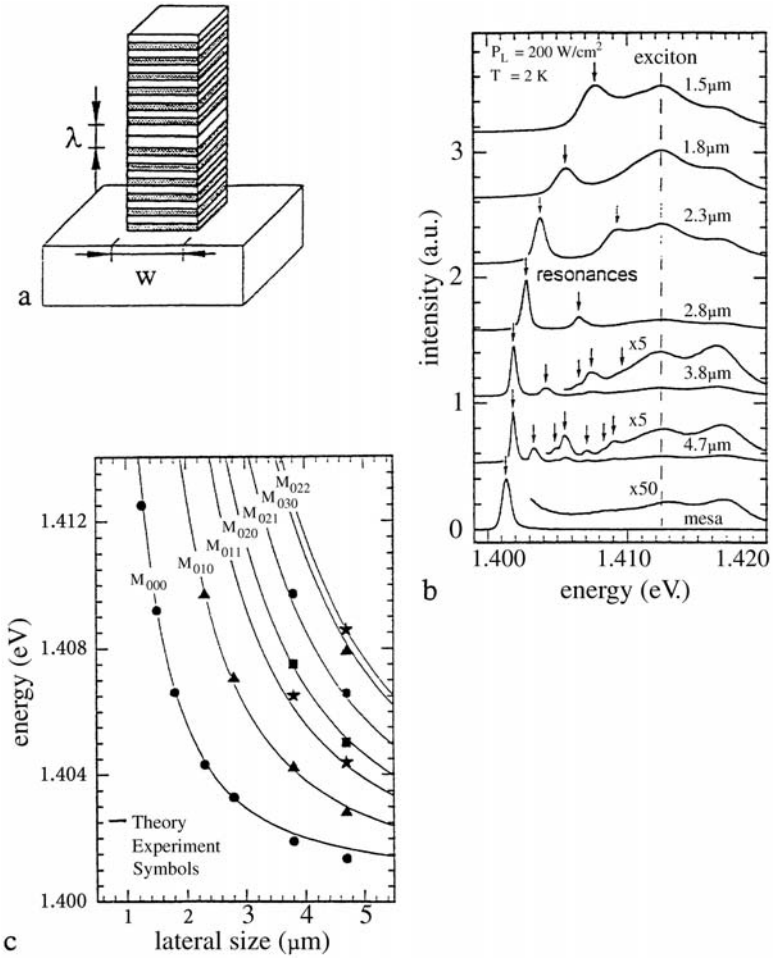


Fig. 17.6. (**a**) A photonic atom consisting of a square mesa prepared from a structure containing a quantum well in a cavity with stacks of Bragg mirrors on both sides. (**b**) The luminescence of the mesa with $w \Rightarrow \infty$ and for various finite values of w. (**c**) The coincidence of the measured and calculated emission peaks as a function of w [97R1]

In the next step a three-dimensional resonator is formed by etching a small square (or cylindrical) mesa of width w (see Fig. 17.6a). The natural reflectivity of the side walls resulting from refractive indices around three is sufficient to form a resonator with three-dimensional confinement for light as seen from the increasing spacing of the emission peaks, which correspond to the various cavity modes, with decreasing w in Fig. 17.6b,c.

The name "photonic atom" is coined in analogy to the artificial atoms or quantum dots of Sects. 8.13 and 15.4 in which there is no translational degree of freedom and the energy level spectrum is discrete as in an atom.

The spectral positions of the emission peaks are given by

$$\hbar\omega = \frac{\hbar c}{n}\left(k_0^2 + k_x^2 + k_y^2\right)^{1/2} \tag{17.14a}$$

where k_0 is the wave vector defined by the Bragg mirror cavity (see Sect. 17.1) and the discrete values of k_x and k_y are given by

$$k_i = (m_i + 1)\,\frac{\pi}{w}\,, \quad m_i = 0, 1, 2, \ldots\,, \quad i = x, y\,. \tag{17.14b}$$

Figure 17.6c shows excellent agreement between the experimental data and the results of (17.14a,b) for various modes.

The next step is to couple two of the photonic atoms to a photonic molecule by a bridge between them as shown in Fig. 17.7a.

The two lowest modes have even and odd parity in close analogy to the binding and anti-binding states in an H_2 molecule. Increasing coupling, as seen in Fig. 17.7c, by increasing the width of the bridge at constant length of 1 μm leads to a splitting of the otherwise degenerate levels with one node line in each photonic dot or atom.

The obvious extension is to couple more and more identical photonic atoms, for simplicity, in a linear array. One observes a splitting in more and more discrete levels as expected from coupled identical harmonic oscillators. These discrete levels arrange themselves into bands in a similar way as for the periodic Kronig–Penney model.

In Fig. 17.8 we show the one-dimensional band-structure of a linear array of 50 coupled photonic atoms. The formation of bands and of band gaps at the border of the Brillouin zones (indicated by the vertical dotted lines) is obvious. The absence of a gap at the border of the third Brillouin zone is an effect of the specific choice of parameters of the sample used in [99B3, 01B1].

Now it is possible to introduce a defect in the linear chain of otherwise identical photonic atoms, e.g., by incorporating one atom with a different size. Depending on this size, the defect may form a state in the bands, which is not so interesting, or one in the gap. This latter case corresponds again to the Fabry–Perot resonator with a spectral narrow transmitting mode in the high-reflectivity photonic gap [02G2].

Sometimes these states are called acceptor- or donator-like, if they are situated energetically just above a (valence) band or just below a (conduction)

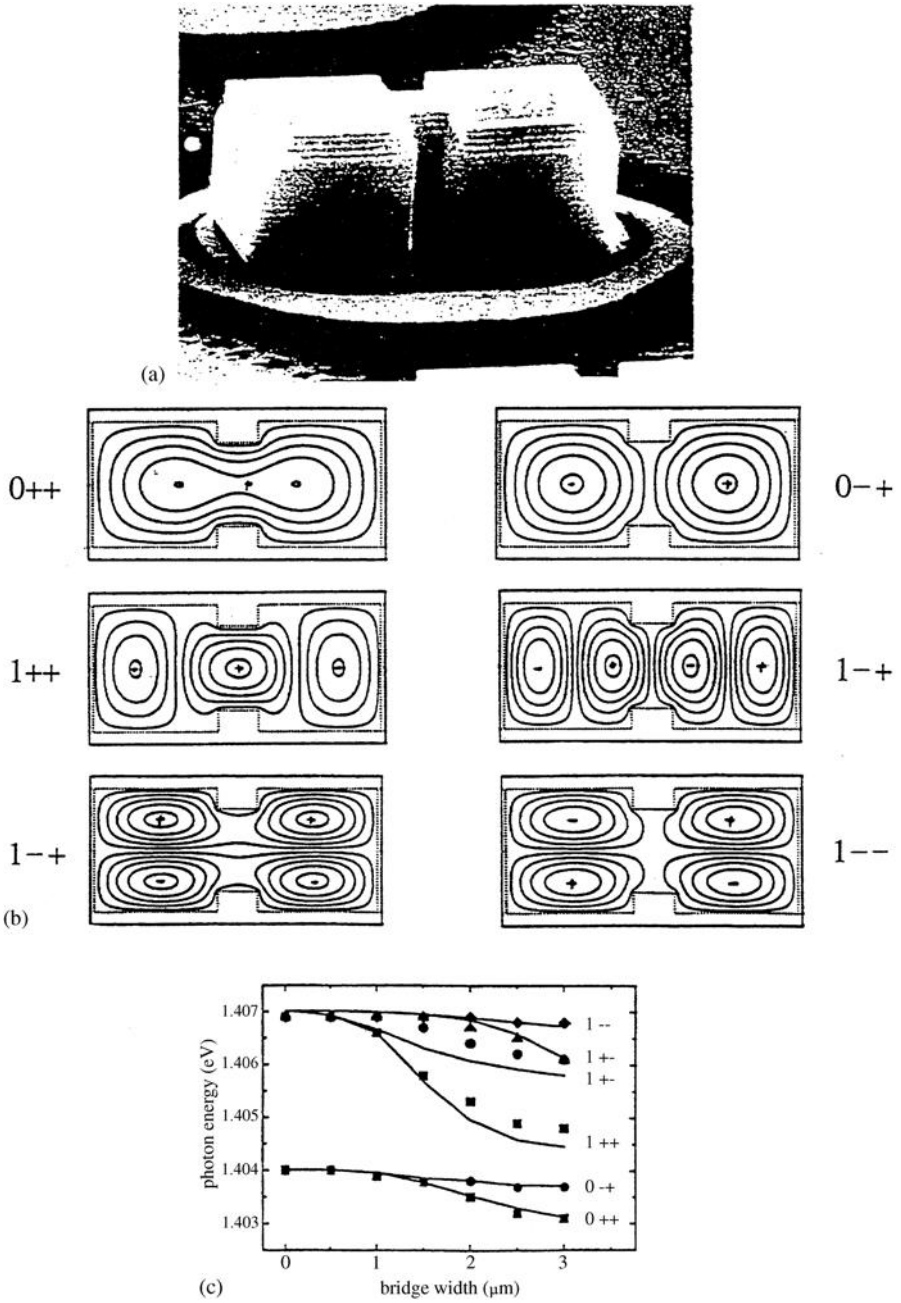


Fig. 17.7. (**a**) A photonic molecule made by coupling of two atoms via a bridge. (**b**) Calculations of the field distributions of the six lowest coupled states. (**c**) Their spectral positions as a function of the bridge width [98B3, 01B1]

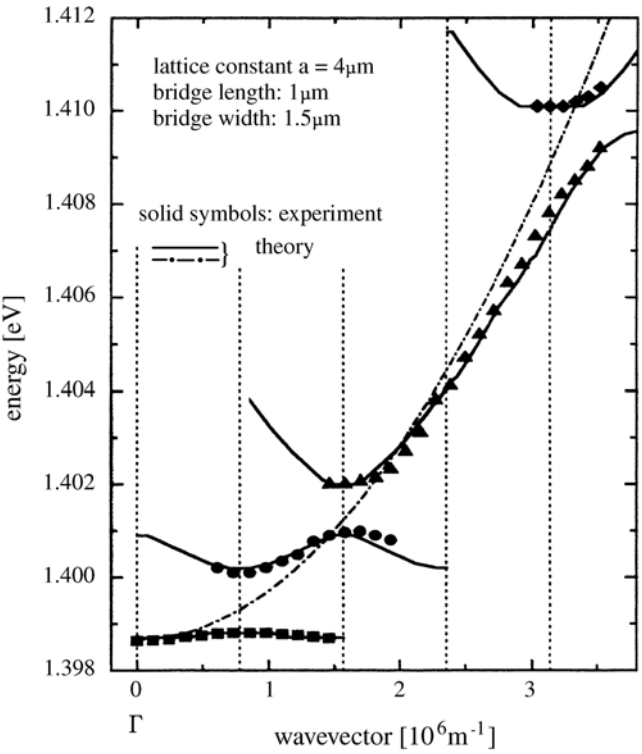


Fig. 17.8. The dispersion relation of a linear chain of 50 coupled microresonators. The *vertical dotted lines* give the borders of the one-dimensional Brillouin zones, $n\pi/a; n = 1, 2, 3, \ldots$ [99B3]

band and a photonic band gap material is referred to as a "semiconductor for light". The author feels that this nomenclature is stressing the analogy to electronic band-structure somewhat too far, since electrons are fermions and it consequently makes sense to speak about an up most, completely filled valence band and an empty conduction band. But since light quanta are bosons there are no such things as a completely filled photonic valence band or an ionized donor.

17.4 Further Developments of Photonic Crystals

The investigation of photonic crystals and especially of photonic band gap materials is presently a very active field of research into which more and more international research groups are joining.

One aim is of course to improve the quality of the periodic arrays, to verify the existence of band gaps and to compare to theoretical models.

This field, apart from being a beautiful physical playground, holds promises for both fundamental research and applications. We give a few examples.

Imagine that a luminescent atom or an electronic quantum dot is incorporated into a three-dimensional photonic band gap material such that the emission wavelength of this atom falls into the forbidden photonic gap. Then this atom can not radiate. An emitted light quantum cannot propagate away from the radiator but remains confined in the vicinity and will be reabsorbed. This is a chance to test basic questions of quantum electrodynamics. The first examples showing at least an increase in the radiative decay time have been reported, e.g., in [00V1].

An application of this fact could result from the possibility of influencing the band-structure and the position and/or width of a photonic gap by an external electric or magnetic field, e.g., by incorporating liquid crystal material [99B2] or semimagnetic semiconductors in the photonic structure. The emission of the above mentioned atom can then be switched, e.g., by an electric field if it shifts the band gap away from the emission wavelength.

The application of photonic structures as wave guides, mirrors, cavities, etc., has already been mentioned and one hopes to build, with the help of such structures, microlasers with extremely low threshold and pumping power.

Presently there are (theoretical) attempts to calculate the band-structure for the case of spectrally strong varying $\varepsilon(\omega, \boldsymbol{r})$ as it occurs in the vicinity of resonances or to treat complex dielectric functions that involve absorption with the obvious consequence that there are no longer stationary eigenstates.

As has been seen many times in the past, the availability of a new class of materials allows one to repeat experiments done before on simpler systems. In this sense, one can expect to see, in the near future, experiments of time-resolved and/or nonlinear spectroscopy, like time of flight measurements for light pulses through photonic crystals, similar to the data presented in Sect. 13.1.4, or experiments on nonlinear optics in the simplest case second harmonic generation or of $\boldsymbol{k}$-space spectroscopy (see Sect. 13.1.4 and Chap. 19).

17.5 Problems

1. Calculate the parabolic dispersion relation of Fig. 17.1e and the effective mass of (17.7).
2. In connection with disorder, we introduced the concept of enhanced backscattering. Could this concept also work for elastic scattering of light by small polystyrol or H_2O spheres?
3. Recall the way we calculated the density of states in $\boldsymbol{k}$- and E-space. Are there similarities to the mode counting in a Fabry–Perot resonator?
4. Why is a dielectric mirror also called a Bragg mirror?
5. Formulate the equations for the magnetic and electric fields for a material with spatially periodic, but frequency independent, refractive index n.

What is different if $n = n(\omega)$? What would you expect qualitatively for a complex index of refraction (or dielectric function) in analogy to a complex potential?

References to Chap. 17

[87J1] S. John, Phys. Rev. Lett. **58**, 2486 (1987)
[87Y1] E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987)
[89T1] A. Thelen, The Design of Optical Interference Coatings, Mc Graw Hill, New York (1989)
[91Y1] E. Yablonovitch, T.J. Gnütter, and K.M. Leung, Phys. Rev. Lett. **67**, 2295 (1991)
[92W1] C. Weisbuch et al., Phys. Rev. Lett. **69**, 3314 (1992)
[94H1] R. Houdré et al., Phys. Rev. Lett. **73**, 2043 (1994)
[95J1] J.D. Joannopoulos, R.D. Meade, and J.N. Winn, Photonic Crystals, Princeton University Press, Princeton (1995)
[95K1] P. Kelkar et al., Phys. Rev. B **52**, R5491 (1995)
[95Ü1] H. Übbing et al., Il Nuovo Cimento **17D**, 1753 (1995)
[96M1] Microcavities and Photonic Band gaps: Physics and Applications, J. Rarity and C. Weinsbuch, eds., NATO ASI Series E **324** Kluver, Amsterdam (1996)
[96P1] Photonic Bandgap Materials, C.M. Soukoulis, ed. NATO ASI Series E **315** Kluver, Amsterdam (1996)
[97F1] J.S. Faresi et al., Nature **390**, 143 (1997)
[97R1] J.P. Reithmaier et al., Phys. Rev. Lett. **78**, 378 (1997)
[98B1] K. Busch and S. John, Phys. Rev. E **58**, 3896 (1998)
[98B2] A. Birner et al., phys. stat. sol. (a) **165**, 111 (1998)
[98B3] M. Bayer et al., Phys. Rev. Lett. **81**, 2582 (1998)
[98G1] T. Gutbrod et al., Phys. Rev. B **57**, 9950 (1998)
[98G2] J.M. Gerard et al., Phys. Rev. Lett. **81**, 1110 (1998)
[98N1] U. Neukirch et al., Phys. Rev. B **57**, 9208 (1998)
[98Q1] F. Quochi et al., J. Cryst. Growth **184/185**, 754 (1998)
[98S1] V. Savona, J. Cryst. Growth **184/185**, 737 (1998)
[99A1] L.C. Andreani and G. Panzarini, Phys. Rev. B **60**, 13276 (1999)
[99B1] A. Birner, K. Busch and F. Müller, Physik. Blätter **55**, Issue 1, p. 27 (1999)
[99B2] K. Busch, Physikal. Blätter **55**, Issue 4, p. 27 (1999)
[99B3] M. Bayer et al., Phys. Rev. Lett. **83**, 5374 (1999)
[99C1] C. Constantin et al., Phys. Rev. B **59**, R7809 (1999) and Mat. and Science Engineering B **74**, 158 (2000)
[99G1] T. Gutbrod et al., Phys. Rev. B **59**, 2223 (1999)
[99J1] S. John and K. Busch, J. Lightwave Technol. IEEE **17**, 1931 (1999)
[99K1] G. Khitrova et al., Rev. Mod. Phys. **71**, 1591 (1999)
[00C1] C. Constantin et al., Mat. and Science Engineering B **74**, 158 (2000)
[00V1] Yu. Vlasov, J. Appl. Phys. **76**, 1627 (2000)
[01A1] S. Arnold, American Scientist **89**, 414 (2001)
[01A2] M.V. Artemyev, U. Woggon, and R. Wannemacher, Appl. Phys. Lett. **78**, 1032 (2001)
[01B1] M. Bayer, Physik. Blätter **57**, Issue 7/8, p. 75 (2001)

[01B2] A. Blanco et al., Appl. Phys. Lett. **78**, 3181 (2001)
[01B3] M. Bayer et al., Phys. Rev. Lett. **86**, 3168 (2001)
[01G1] G. Guttroff et al., Phys. Rev. E **63**, 36611 (2001)
[01K1] C. Klingshirn, in Advances in Energy Transfer Processes, B. Di Bartolo and X. Chen, eds., World Scientific, Singapore (2001) p. 165
[01R1] S.G. Ramanov et al., Appl. Phys. Lett. **79**, 731 (2001)
[01S1] K. Sakoda, Optical Properties of Phontic Crystals, Springer, Berlin (2001)
[02A1] M. Artemyev, U. Woggon and W. Langbein, phys. stat. solidi B **229**, 423 (2002)
[02G1] V.G. Golubev et al., J. Non-Cryst. Solids **299-302**, 1062 (2002)
[02G2] G. Guttroff et al., Phys. Rev. B **64**, 155313 (2002)
[02H1] C. Herrmann and O. Hess, JOSA B **19**, 3013 (2002)
[02K1] M. Kazes et al., Adv. Materials, **14**, 317 (2002)
[02M1] B. Möller et al., Appl. Phys. Lett. **80**, 3253 (2002) and to be published (2004)
[03B1] R. v. Baltz, NATO Science Series II **90**, 91 (2003)
[03B2] V. Babin et al., Appl. Phys. Lett. **82**, 1553 (2003)
[03C1] A. Chist et al., Phys. Rev. Lett. **91**, 183901 (2003)
[03H1] A. Huynh et al., Phys. Rev. B **68**, 165340 (2003)
[03K1] A. Kakovin and G. Malpuech, Cavity Polaritons, Elsevier, Amsterdam (2003)
[03M1] Yu.V. Miklayev et al., Appl. Phys. Lett. **82**, 12846 (2003)
[03M2] B. Möller et al., Appl. Phys. Lett. **83**, 2686 (2003)
[03P1] Photonic Crystals: Optical Materials for the 21^{th} Century, K. Busch and R. Wehrspohn eds., phys. Stat. solidi A **197**, Issue 3 (2003)
[03S1] R.E. Slusher and B.J. Eggleton, Nonlinear Photonic Crystals, Springer, Berlin, Heidelberg, (2003)
[03W1] Z.L. Wang et al., Phys. Rev. E **67**, 16612 (2003)
[04G1] H.M. Gibbs, Spinger Series in Solid-State Science **146**, 189 (2004)

18

Review of the Linear Optical Properties

In this brief chapter, we shall review and summarize some of the aspects of the linear optical properties of semiconductors that were presented in the preceeding chapters in some detail.

18.1 Review of the Linear Optical Properties

In Fig. 18.1 we give a schematic overview of the spectra of the complex dielectric function, of the complex index of refraction and of the reflectivity for a typical direct-gap semiconductor over the whole spectral range from the IR to the UV. For simplicity, we consider one optically active phonon mode, one exciton resonance, and one further resonance which represents all continuum and band-to-band transitions. We include some damping for every resonance but neglect details including spatial dispersion. The additive structure of the resonances in the dielectric function is clearly visible. The back ground dielectric constant ε_b of one resonance is simultaneously the "static" one ε_s for the next higher resonance.

Figure 18.1a includes contributions which might come from two other effects. The dashed-dotted line gives the modifications introduced by a plasma. [This plasma can be caused either by high doping levels (Sects. 10.1, 12.1) or by high excitation (Chap. 21)]. In the first case it consists of either electrons or of holes, in the second case it is a bipolar plasma. If we remember that the transverse eigenfrequency of a plasma is zero, we immediately obtain the dashed-dotted line in Fig. 18.1. We should mention that the presence of a plasma also influences the other resonances, e.g., the phonon resonance, due to plasmon–phonon mixed states (Sect. 12.3), or the exciton resonances by increasing damping, by screening of the Coulomb attraction of electron and hole and by band gap renormalization and band filling (Chap. 21). These effects are not shown in Fig. 18.1 for sake of clarity.

The other contribution to the dielectric function is the so-called orientational polarization shown by the dashed line. It describes the contribution

of freely rotatable permanent electric dipoles to $\varepsilon(\omega)$, i.e., a para-electric contribution. Freely rotatable means that the dipoles do not have a restoring force, but that an orientation of the dipoles which has been established, e.g., by an external field, decays after switching off of this field with a relaxation time τ by (thermal) collisions with the surroundings. The contribution of the orientational polarization to ε_1 is constant for low frequencies as long as the dipoles can follow the driving field and decays for $\omega\tau \geqslant 1$

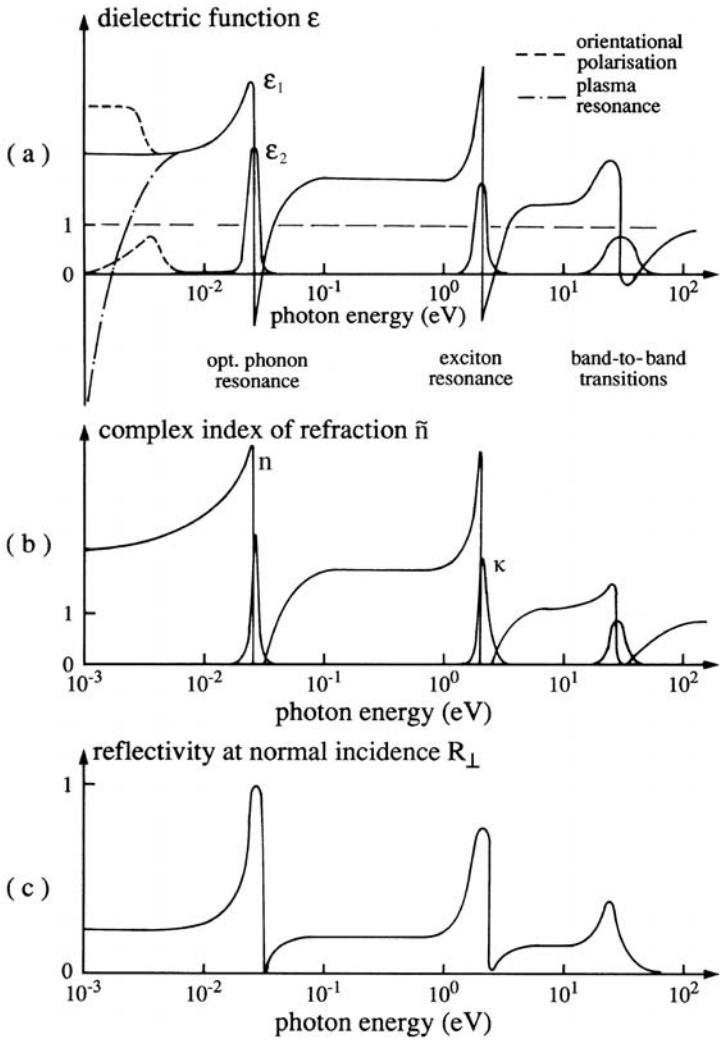


Fig. 18.1. A schematic overview from the IR to the UV of the spectra of the real and imaginary parts of $\varepsilon(\omega)$ (**a**), of $\tilde{n}(\omega)$ (**b**), and of the reflectivity (**c**) for a semiconductor

as is usual for relaxation phenomena. The imaginary part has its maximum at $\omega\tau = 1$. This is reasonable if we assume that the "friction losses" of the rotating dipoles are proportional to the amplitude and the speed ω of the rotation. The orientational polarization is a very important effect for polar liquids (e.g., water) and organic solids containing some rotatable radicals. In semiconductors this effect is less common but is mentioned here for sake of completeness.

Another "summary" of what we have learned up to now is the dispersion relation in Fig. 18.2, where we again include one phononic and two electronic resonances. For simplicity we consider zero damping and give only the real part of $\boldsymbol{k}$, but properly including the spatial dispersion, i.e., the $\boldsymbol{k}$-dependence of the eigenfrequencies. The "global" polariton dispersion of Fig. 18.2 shows nicely the various resonances and how, above the highest eigenfrequency, the dispersion asymptotically approaches that of photons in vacuum.

Finally, Fig. 18.3 summarizes the spectra of the dispersion, the absorption coefficient, the reflectivity, and the luminescence of a direct-gap semiconductor at low temperature including free and bound excitons and donor–acceptor pairs, all with their LO-phonon replica.

To end this summary, we refer the reader to a small collection of more-or-less recent textbooks covering some aspects of semiconductor optics [69O1, 72O1,72W1,75B1,90H1,91K1,91L1,93P1,93O1,95K1,96S1,96Y1,01H1,02D1, 02S1,03Q1,04O1]. Further titels are given in the references of Chap. 1. These books also demonstrate nicely the progress that has been made in this field

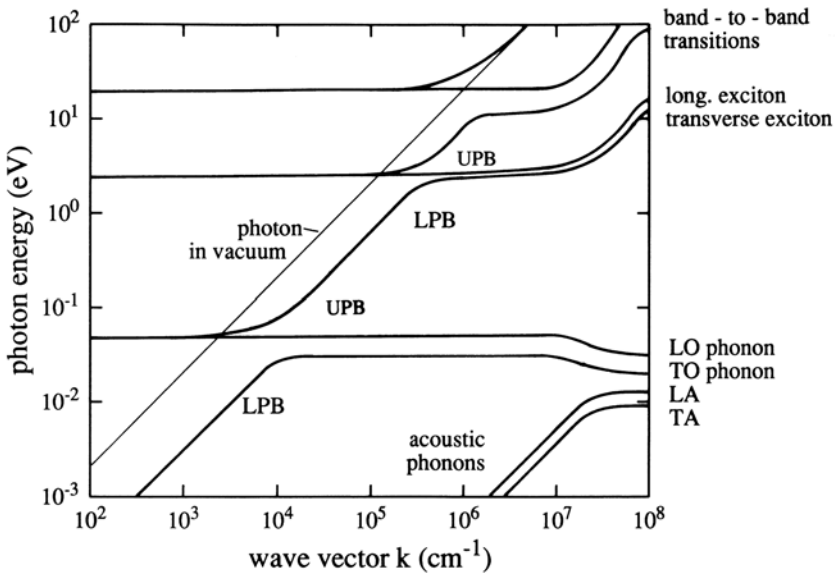


Fig. 18.2. A schematic overview of the real part of the dispersion relation of light in a semiconductor from the IR to the UV, neglecting damping

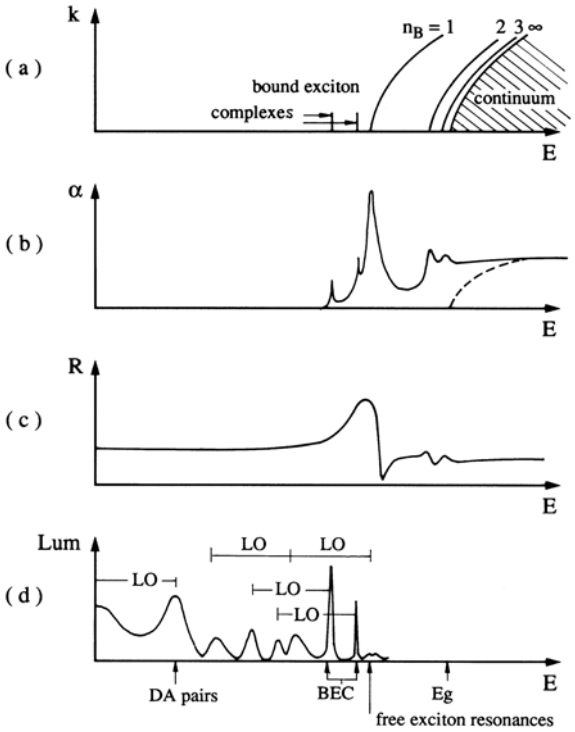


Fig. 18.3. A schematic drawing of the dispersion (**a**), and the spectra of the absorption (**b**), reflection (**c**) and luminescence (**d**) of a high-quality, direct-gap semiconductor at low temperatures in the region of the exciton resonances

over the last four decades. The author considers all these book of high scientific and didactic value, but enjoys especially [04O1] as becomes clear from its preface.

18.2 Problem

Try to identify what is sketched on the front cover. Which features are wrong?

References to Chap. 18

[69O1] *Optical Properties of Solids*, S. Nudelman, S.S. Mitra (eds.) (Plenum, New York 1969)

[72O1] *Optical Properties of Solids*, F. Abeles (ed.) (North-Holland, Amsterdam 1972)

[72W1] F. Wooten: *Optical Properties of Solids* (Academic, New York 1972)

[75B1] F. Bassani, G.P. Parravicini: *Electronic States and Optical Transitions in Solids* (Pergamon, Oxford 1975)

[90H1] H. Haug, S.W. Koch: *Quantum Theory of the Optical and Electronic Properties of Semiconductors* (World Sientific, Singapore 1990)

[91K1] H. Kuzmany: *Festkörperspektroskopie* (Springer, Berlin, Heidelberg 1991)
[91L1] P.T. Landsberg: *Recombination of Semiconductors* (Cambridge Univ. Press, Cambridge 1991)
[93P1] N. Peyghambarian, S.W. Koch, A. Mysyrowicz: *Introduction to Semiconductor Optics* (Prentice Hall, Engelwood Cliffs, NJ 1993)
[93O1] *Optics of Semiconductor Nanostructures*,F. Henneberger, S. Schmitt-Rick, E.O. Göbel (eds) (Academic, Berlin 1993)
[95K1] C.F. Klingshirn (1^{st} ed.): *Semiconductor Optics* (Springer, Berlin 1995)
[96S1] J. Shah: *Ultrafast Spectroscopy of Semiconductors and of Semiconductor Nanostructures* (Springer, Series in Solid State Sciences **115** (1996))
[96Y1] P.Y. Yu, M. Cardona: *Fundamentals of Semiconductors* (Springer, Berlin 1996)
[01H1] C. Hamaguchi: *Basic Semiconductor Physics* (Springer, Berlin 2001)
[02D1] M. Dressel, G. Grüner: *Eletrodynamics of Solids* (Cambridge Univ. Press, Cambridge 2002)
[02S1] W. Schäfer, M. Wegener: *Semiconductor Optics and Transport Phenomena* (Springer, Berlin 2002)
[03Q1] *Quantum Coherence, Correlation and Decoherence in Semiconductor Nanostructures*, T. Takagahara (ed.) (Academic Press, Amsterdam 2003)
[04O1] *Optics of Semiconductors and Their Nanostructures*, H. Kalt and M. Hetterich (eds.), Springer Series in Solid-state Science **146** (2004)

19

High Excitation Effects and Nonlinear Optics

In this and in some of the following chapters we shall leave the regime of linear optics introduced in Sect. 2.3 and proceed to the field of nonlinear optics. Nonlinear optics including high excitation phenomena, laser emission and electro-optics, forms together with the investigation of semiconductors of reduced dimensionality, presently the most active fields in semiconductor science.

In the next two sections we give the definition and the general scenario. In Sect. 19.3 we then present a short outlook to extreme nonlinear optics, which leads, however, beyond the field of semiconductor optics and thus beyond the scope of this book. Then we continue in Chaps. 20–24 with discussions of the most important effects and phenomena contributing to nonlinear optics in semiconductors including time-resolved spectroscopy.

For a rather early theoretical work on nonlinear optics see [31G1] and for early experiments and more recent reviews [61K1, 62B1, 62M1, 63F1, 64B1, 64B2, 65B1, 65B2, 66K1, 71K1] and [75A1, 84S1, 86E1, 89N1, 90B1, 91M1, 95M1, 99N1, 00N1, 02S1, 03W1, 04O1], respectively, the references given therein and some further textbooks given in the references of Chaps. 1 and 20–24. In [00N1] an overview is given among others of various definitions of nonlinear susceptibilities and a conversion tables between various units like SI and esu. As can be seen from [31G1], in which Göppert-Mayer predicted two photon absorption on the basis of perturbation theory, there were early predictions of nonlinear optical effects, but the field started to develop only after the invention of the laser in 1960.

19.1 Introduction and Definition

In linear optics we learned how the optical properties of matter depend on the frequency of the incident radiation field and on the direction of polarization or propagation relative to the crystallographic axes. But we explicitly assumed that the optical properties do not depend on the field amplitude(s) $\boldsymbol{E}_\mathrm{i}$ (or the

intensities, I) of the incident light beam(s). As a consequence, the polarization of matter oscillates with the same frequency as the incident field(s) and two light beams can cross each other in matter without mutual interaction.

The regime of nonlinear optics comprises all effects for which the above assumptions are no longer valid. As a definition of the term "nonlinear optics" we can thus say that these are all phenomena in which the optical properties like $\varepsilon(\omega)$ or $\tilde{n}(\omega)$ depend in a reversible way on the illumination. We stress here the term "reversible". This means that the system returns to its initial state, after the illumination has been switched off, possibly after some time delay. We therefore exclude from our considerations phenomena like the photographic process or simply drilling a hole into a sample by intense laser excitation, a process known also as "laser ablation" "evaporation" or optical breakdown. Some references to this field are found, e.g., in [84S1] and more recently [92B1, 97G1, 00H1], which show that these effects may even have some relevance towards application. An example from every-day life for an optical nonlinearity are phototropic sun-glasses, which are transparent under weak illumination, become dark under the influence of UV radiation, e.g. from sunlight, and become transparent again after a few seconds if the illumination is reduced.

There are two limiting approaches to nonlinear optics. In the first we assume that the response of the medium to the incident field(s) depends only on the instantaneous field amplitudes. This condition is fulfilled if the electronic excitations, on which we shall concentrate in the following, are created only virtually using the language of weak coupling (see Sect. 3.2) or, in the language of strong coupling, introduced in Chap. 5 and Sect. 13.1, if we treat only interactions of coherent polaritons. This approach may also be used, if the duration of the exciting laser pulse is shorter than the dephasing time T_2 of the excitation.

In fact, the weak coupling approach is used in most cases including the optical- and semiconductor-Block equations treated in Chap. 27 to describe optical nonlinearities, and we follow this trend here. But we also give some examples and hints for the proper description of the phenomena in the polariton picture.

As already stated, we now assume that the dielectric susceptibility depends on the instantaneous field amplitudes, i.e.,

$$\chi(\omega_i, \boldsymbol{E}_i) = \varepsilon(\omega_i, \boldsymbol{E}_i) - 1\,. \tag{19.1}$$

Since the dependence on $\boldsymbol{E}_i$ is usually not known explicitly we expand $\chi(\omega_i, \boldsymbol{E}_i)$ into a power series of the incident field amplitudes [63F1, 65B1, 65B2, 84S1]:

$$\frac{1}{\varepsilon_0}\boldsymbol{P}_i = \sum_j \chi^{(1)}_{ij}\boldsymbol{E}_j + \sum_{j,k} \chi^{(2)}_{ijk}\boldsymbol{E}_j\boldsymbol{E}_k + \sum_{j,k,l} \chi^{(3)}_{ijkl}\boldsymbol{E}_j\boldsymbol{E}_k\boldsymbol{E}_l + \ldots\,. \tag{19.2}$$

If all frequencies are equal, a similar expansion can be formulated for the refractive index using intensities, i.e.,

$$\tilde{n}(\omega, I) = \tilde{n}_0(\omega) + \tilde{n}_2 I + \ldots \tag{19.3}$$

where $\tilde{n}_0$ is again the linear refractive index and $\tilde{n}_2$ describes changes induced by I.

The first term on the right-hand side of (19.2),(19.3) describes the linear optical properties discussed in Chaps. 2–18. The second term of (19.2) containing the phenomenologically introduced parameter $\chi^{(2)}$ describes effects like second-harmonic, sum- and difference-frequency generation or the dc effect, i.e., rectification of the electric field of a light beam.

It immediately becomes clear what is meant by these terms if we direct a (laser) light beam with frequency ω onto the sample. The second term then reads

$$\chi^{(2)} \boldsymbol{E}_0^2 \sin^2 \omega t = \frac{1}{2} \chi^{(2)} \boldsymbol{E}_0^2 (1 - \cos 2\omega t) \,. \tag{19.4}$$

The $\cos 2\omega t$ term in (19.4) tells us that a contribution to the polarization is created, which oscillates at 2ω, and is radiated by the sample, the so-called second harmonic generation. The first term in the bracket describes a temporally constant polarization, which results in a voltage across the sample as shown schematically in Fig 19.1. This corresponds to a partial rectification of the ac field of the light beam. The corresponding effect is therefore also called the "dc effect". The dc effect is, to a good approximation, an inversion of the linear electro-optic effect as mentioned, e.g., in [71K1, 84S1]. If two fields with different frequencies ω_1 and ω_2 interact in the sample via $\chi^{(2)}$ effects, a similar approach as in (19.4) gives contributions to the polarization which oscillate with frequencies $\omega_1 \pm \omega_2$, i.e.,

$$\begin{aligned} \text{sum-frequency generation:} \quad & (\omega_1 + \omega_2) \,, \\ \text{difference-frequency generation:} \quad & (\omega_1 - \omega_2) \,, \end{aligned} \tag{19.5}$$

which are also radiated.

$\chi^{(2)}$ and all other even terms in the expansion (19.2) vanish for crystals whose symmetry elements contain the inversion. This is immediately clear by letting $\boldsymbol{E} \rightarrow -\boldsymbol{E}$ and consequently $\boldsymbol{P} \rightarrow -\boldsymbol{P}$.

For non-centrosymmetric crystals, the $\chi^{(2)}$ tensor contains one or more non-zero elements. Which elements vanish and which are equal depends on the point group of the crystal. Details are given in [63F1, 65B1].

Another aspect of the $\chi^{(2)}$ and of higher order phenomena concerns interference. When the fundamental beam propagates in the sample, it creates everywhere the second harmonic with the same relative phase between fundamental wave and second harmonic. On the other hand, the second harmonic, created at one place propagates and interferes with the second harmonic generated deeper in the sample. In order to get a maximum output of the second harmonic, this interference should be always constructive. The requires that

$$n(\omega) = n(2\omega) \tag{19.6a}$$

or

$$\boldsymbol{k}(2\omega) = 2\boldsymbol{k}(\omega)\,. \tag{19.6b}$$

The second version is nothing but momentum conservation i.e., two light quanta of frequency $\hbar\omega$ and (quasi-) momentum $\hbar\boldsymbol{k}(\omega)$ are annihilated to give one with $2\hbar\omega$ and $\hbar\boldsymbol{k}(2\omega)$. The condition (19.6) is also known as phase-matching and can usually be fulfilled only in some birefringent materials, when the polarizations of the waves with ω and 2ω are different, because $n(\omega)$ tends to increase in the transparent window of a crystal with increasing ω, generally leading to

$$n(\omega) < n(2\omega)\,. \tag{19.6c}$$

Additionally one needs a special orientation for a given frequency. Similar arguments to (19.6b) hold also for sum- and difference-frequency generation.

If phase matching is perfectly realized, the intensity of the second harmonic starts to grow quadratically with the propagation distance in the sample with a simultaneous decrease in the fundamental wave. The increase of the second harmonic intensity tends, however, to saturation, if the intensity of the fundamental wave drops substantially. Ideally one could expect 100% transfer of energy into the second harmonic for an infinitely thick crystal. In reality, an efficiency of 10 to 30% is realistic, depending on the quality of the laser beam, on the sample thickness or the value of $\chi^{(2)}$, etc.

If phase matching is not fulfilled, the second harmonic also starts to increase with sample thickness. For thicker samples the second harmonic contributions created at various places start, however, to interfere destructively,

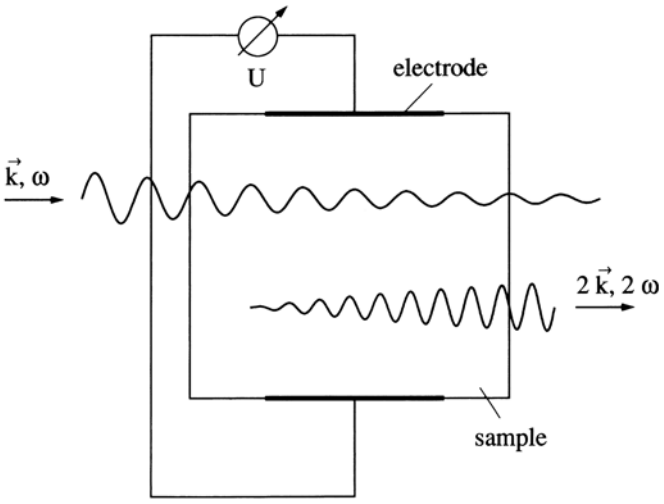


Fig. 19.1. Schematic drawing of an experimental situation in which second harmonic generation and the dc effect can be observed

again feeding the fundamental wave. This effect results in an oscillation of the second harmonic intensity with sample thickness (the so-called Maker fringes) [62M1, 65B2, 84S1]. Consequently, sufficiently thin samples are adequate for non-phase-matching conditions.

The $\chi^{(3)}$ effects describe four-wave mixing (FWM) and hyper-Raman scattering (HRS), coherent anti-Stokes Raman scattering (CARS) etc. Some of these effects are described with the examples below.

The linear and nonlinear optical effects discussed here can be described in perturbation theory of increasing order either with the dipole operator $H_1 \Rightarrow H^{\mathrm{D}}$ or the second-order term $H_2^{(2)}$ of (3.46).

We give in Fig. 19.2 schematic drawings of various nonlinear processes in the picture of two- and three-level systems and in (19.7) to (19.16) a schematic or intuitive translation into equations. A detailed treatment of the formalism including Feynman diagrams can be found, e.g., in [65B2, 84S1].

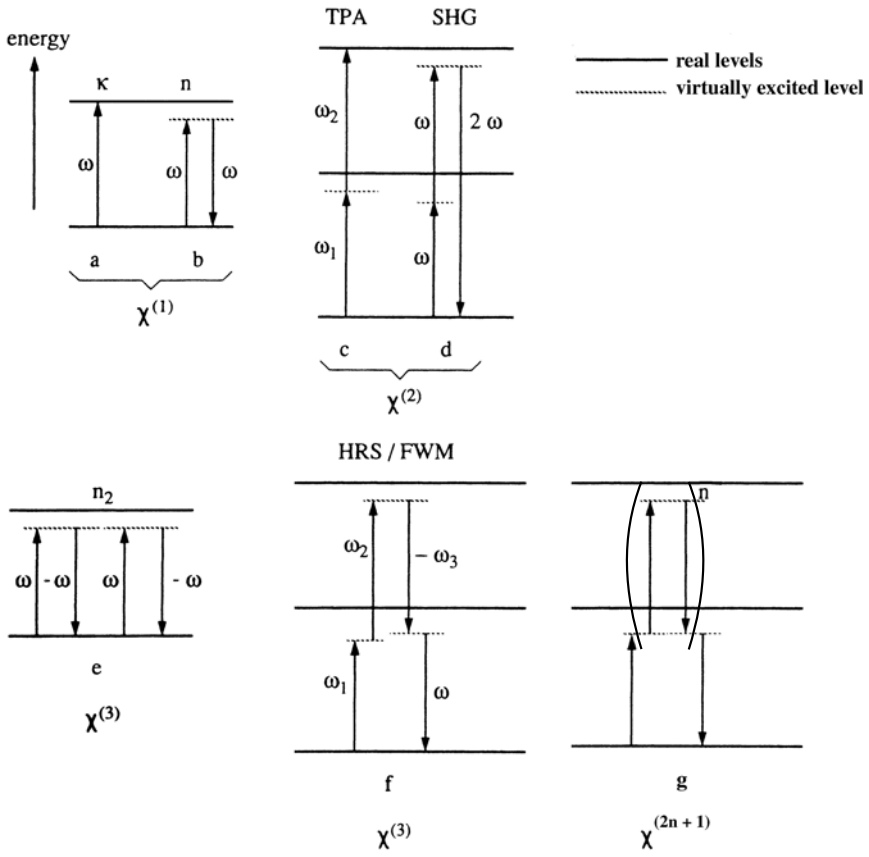


Fig. 19.2. Schematic presentation of various linear (**a**,**b**) and nonlinear optical processes (**c**-**f**)

The linear one photon absorption $\alpha(\omega)$ is then at resonance proportional to

$$\alpha(\omega) \propto w_{if} \propto |\langle f| H^D(\omega) |i\rangle|^2 \,. \tag{19.7a}$$

It should be noted that the absorption process is only completed when the final state undergoes some scattering process, which destroys the coherence of the polarization of the transition with the incident light field. Otherwise, an intensity dependent coherent oscillation between the ground and excited states initiates, in the sense of Rabi oscillations [37R1] or flopping discussed in Sect. 20.4 and in Chap. 27.

The corresponding linear refractive index can be visualized off resonance by a process involving one virtually excited intermediate states $|z\rangle$ (Fig 19.2b) which after some phase delay spontaneously emits a photon, which is otherwise identical to the incident one. We use the following notation: Absorption processes leading to real or virtual excited states are described by positive frequencies in $\chi^{(n)}$ and stimulated or spontaneous emission processes induced by an incident field or by the zero-point fluctuations, respectively, by negative frequencies.

We recall the dipole operator H^{D} from (3.58) and write for small imaginary parts $\kappa(\omega)$ of $\tilde{n}(\omega)$, i.e., away from the resonance energy E_{zj} for the real part of the linear susceptibility (see also (4.22c))

$$\chi^{(1)}(\omega) = n^2(\omega) - 1 \propto \sum_{z_j} \frac{\langle i \left| H^{\mathrm{D}}(-\omega) \right| z_j \rangle \langle z_j \left| H^{\mathrm{D}}(\omega) \right| i \rangle}{\hbar\omega - (E_{zj} - E_i)} \tag{19.7b}$$

Alternatively $n(\omega)$ can be obtained from $\kappa(\omega)$ via the Kramers–Kronig relations (Chap. 6) or vice versa.

The two-photon absorption coefficient β introduced phenomenologically for a one-beam experiment in (19.8) and Fig 18.2c,

$$-\frac{\mathrm{d}I}{\mathrm{d}x} = \beta I^2 \,, \tag{19.8}$$

is then at resonace with the final state $|f\rangle$ proportional to

$$\beta \propto \left| \sum_{z_j} \frac{\langle f \left| H^{\mathrm{D}}(\omega_2) \right| z_j \rangle \langle z_j \left| H^{\mathrm{D}}(\omega_1) \right| i \rangle}{\hbar\omega_1 - (E_{zj} - E_i)} + \text{c.p.} \right|^2 \,, \tag{19.9}$$

where we allow for two different frequencies ω_1 and ω_2, represented by their respective dipole operators, and where c.p. stands for cyclic permutations. Again a decoherence process must occur in the final states for the two photon absorbtion process to be completed.

Energy conservation has to be fulfilled between initial state $|i\rangle$ and final state $|f\rangle$, momentum conservation in every step, i.e.,

$$\left.\begin{array}{l} E_f = E_i + \hbar\omega_1 + \hbar\omega_2 \,, \\ \boldsymbol{k}_i + \boldsymbol{k}(\omega_1) = \boldsymbol{k}_{z_j} \\ \boldsymbol{k}_{z_i} + \boldsymbol{k}(\omega_2) = \boldsymbol{k}_{tf} \,. \end{array} \right\} \text{ and} \tag{19.10}$$

The dipole approximation actually means $\boldsymbol{k}(\omega_{1,2}) = 0$; see (3.55). For early examples of two-photon absorption see, e.g., [62B1, 64B2] and for two step transitions with a really and not virtually excited intermediate state [67B1].

After the short discussion of n, κ and β we turn now to the $\chi^{(n)}$. For the linear susceptibility we obtain (see Chaps. 4, 27 and (19.7b) above)

$$\chi^{(1)}(\omega) \propto \left[\sum_{zj} \frac{\langle i \left| H^{\mathrm{D}}(\omega) \right| z_j \rangle \langle z_j \left| H^{\mathrm{D}}(\omega) \right| i \rangle}{\hbar\omega - (E_i - E_{zj}) + i\gamma} - \sum_{zj} \frac{\langle i \left| H^{D}(\omega) \right| z_j \rangle \langle z_j \left| H^{\mathrm{D}}(\omega) \right| i \rangle}{\hbar\omega + (E_i - E_{zj}) + i\gamma} \right] \tag{19.11}$$

where the first term on the r.h.s. is the resonant one and where we have introduced a small damping γ to avoid a singularity.

From the imaginary part of $\chi^{(1)}$ we get $\kappa(\omega)$ and $\alpha(\omega)$ and from the real part $n^2(\omega) - 1$ for weak absorption as indicated above and in more detail in Chaps. 4 and 27.

Second-harmonic and sum- or differece-frequency generation involve two virtually excited intermediate states $|z_i\rangle$ and $|z_j\rangle$ and are thus described by a contribution to $\chi^{(2)}$ which reads, according to Fig 19.2d for off-resonance conditions:

$$\chi^{(2)} \propto \sum_{z_j, z_k} \frac{\langle i \left| H^{\mathrm{D}}[-(\omega_1 + \omega_2)] \right| z_k \rangle \langle z_k \left| H^{\mathrm{D}}(\omega_2) \right| z_j \rangle \langle z_j \left| H^{\mathrm{D}}(\omega_1) \right| i \rangle}{[\hbar(\omega_1 + \omega_2) - (E_{z_k} - E_{\mathrm{i}})][\hbar\omega_1 - E_{z_j} - E_{\mathrm{i}}]} + \text{c.p.} \tag{19.12}$$

Another off-resonance contribution to the second-harmonic generation comes from first-order perturbation theory, using the term $H^{(2)}$ in (3.46) which was considered to be small of second order.

$$\chi^{(2)} \propto \sum_{z_k} \frac{\langle i \left| H^{\mathrm{D}}(-2\omega) \right| z_k \rangle \langle z_k \left| H^{\mathrm{D}}(2\omega) \right| i \rangle}{2\hbar\omega - E_{z_k}} \tag{19.13}$$

If the decay of the second virtually excited intermediate state is stimulated by a third photon field, this induces a $\chi^{(3)}$ effect of the type shown in Figs. 19.2e,f. We denote the frequencies in these contributions to $\chi^{(3)}$ in the following way:

$$\chi^{(3)}(\omega : \pm\omega_j, \pm\omega_k, \pm\omega_l), \quad \text{with} \quad \omega = \pm\omega_l \pm \omega_k \pm \omega_j, \tag{19.14}$$

where the incident, absorbing, and/or stimulating frequencies are given with their respective sign after the colon, and the resulting frequency of the polarization that is radiated in the process under consideration before the colon.

One of the many contributions to $\chi^{(3)}$ (Fig. 19.2f) is then given by (19.15) again for off-resonance conditions

$$\chi^{(3)}(\omega : -\omega_3, \omega_2, \omega_1) \propto$$
$$\sum_{z_j, z_k, z_l} \frac{\langle i \left| H^{\mathrm{D}}(\omega) \right| z_l \rangle \langle z_l \left| H^{\mathrm{D}}(-\omega_3) \right| z_k \rangle \langle z_k \left| H^{\mathrm{D}}(\omega_2) \right| z_j \rangle \langle z_j \left| H^{\mathrm{D}}(\omega_1) \right| i \rangle}{[\hbar(\omega_1 + \omega_2 - \omega_3) - (E_{z_l} - E_{\mathrm{i}})][\hbar(\omega_1 + \omega_2) - (E_{z_k} - E_{\mathrm{i}})][\hbar\omega_1 - (E_{z_j} - E_{\mathrm{i}})]}$$
$$+ \text{c.p.} \tag{19.15}$$

If we inspect the terms describing one- and two-photon absorption (19.7a) and (19.9), or for three photon absorption not given here, or those which describe processes in which a photon is emitted like (19.7b), (19.10) and (19.13), the rules become intuitively clear. For every next higher order of perturbation theory there is one more dipole matrix element and one more resonance denominator. A product of two dipole matrix elements and of two resonance denominators can be replaced by one H^2 term as can be seen by comparing (19.10) and (19.11).

With increasing order of the perturbation there is an increasing number of cyclic permutations and of $\pm$ as in (19.14). An exhaustive presentation of all phenomena up to order three is given, e.g., in [63F1, 65B1, 65B2, 84S1] and an investigation of higher orders, e.g., in [64B1, 65B2, 66K1, 95A1, 00A1, 01A1] and references therein. A group of possible contributions to higher orders of $\chi^{(n)}$ is shown schematically in Fig. 19.2g.

As a further example, we give in (19.16) a combination which contributes to an excitation-induced change of the real part of the refractive index, i.e., to $\mathrm{Re}\{n_2\}$ in (19.3) in the way sketched in Fig 19.2f:

$$\chi^{(3)}(\omega : \omega, -\omega, \omega) =$$
$$\sum_{z_j, z_k, z_l} \frac{\langle i \left| H^{\mathrm{D}}(-\omega) \right| z_l \rangle \langle z_l \left| H^{\mathrm{D}}(\omega) \right| z_k \rangle \langle z_k \left| H^{\mathrm{D}}(-\omega) \right| z_j \rangle \langle z_j \left| H^{\mathrm{D}}(\omega) \right| i \rangle}{[\hbar\omega - (E_{z_l} - E_{\mathrm{i}})][\hbar(\omega - \omega) - (E_{z_k} - E_{\mathrm{i}})][\hbar\omega - (E_{z_j} - E_{\mathrm{i}})]}$$
$$+ \text{c.p.} \tag{19.16}$$

This process describes evidently an intensity dependent contribution to the refractive index at the frequency ω as in (19.3). The sum of all analogous $\chi^{(3)}(\omega : \pm\omega, \pm\omega, \pm\omega)$ is proportional to $\mathrm{Re}\{\tilde{n}_2(\omega)\}$. These effects lead, e.g., to a self-focussing or defocussing of a light beam with a Gaussian beam profile, depending on the sign of $\mathrm{Re}\{\tilde{n}_2(\omega)\}$ or to self-phase modulation of a short (laser-) pulse. The dynamic blue and red shift of the light frequency resulting from this effect on the leading and trailing edges of the pulse is known as chirp.

In all of the above processes, momentum has to be conserved in every step and energy must be conserved between initial and final states, as already mentioned. Apart from the above mentioned references, more details can be found, e.g., in [81K1, 89S1, 90G1, 91C1, 94P1, 97C1, 99N1, 01Z1, 02S3] and the references given therein.

The second group of nonlinear optical phenomena involves modifications of the optical properties by really or incoherently excited species, e.g., electron–hole pairs, excitons, or phonons. These species have finite lifetimes T_1 which can be of the order of sub-ns to ms. Due to this finite lifetime, their density N does not instantaneously follow the incident light field, but depends on the generation rate in the past weighted by some decay function such as an exponential.

In this case we have:

$$\frac{1}{\varepsilon_0}\boldsymbol{P} = \chi(\omega, N)\boldsymbol{E} \tag{19.17}$$

with

$$N(t) = \int_{-\infty}^{t} G(t') \exp[-(t-t')/T_1]\,\mathrm{d}t' \,. \tag{19.18}$$

The generation rate $G(t')$ is connected with the intensity at time t', e.g., in the presence of one- and two-photon excitation by

$$G(t', x) = \frac{\alpha I(t', x)}{\hbar\omega_{\mathrm{exc}}} + \frac{\beta I^2(t', x)}{2\hbar\omega_{\mathrm{exc}}} \,, \tag{19.19}$$

where α and β are the one- and two-photon absorption coefficients for the excitation light with photon energy $\hbar\omega_{\mathrm{exc}}$ and intensity $I(t', x)$.

A further complication is introduced if we assume that the parameters T_1, α, β in (19.18),(19.19) are not constants but depend themselves on N or I. In this case we are left with rather complex systems of coupled integro-differential equations, which generally can be solved only numerically, often using some further approximations as shown, e.g., in [90H1, 02S1]. Alternatively these problems are treat in the framework of the optical or semiconductor Bloch equations of Chap. 27.

It is obvious that a power expansion as in (19.2), (19.3) is not adequate to describe optical nonlinearities which depend on an incoherent population $N(t)$. This population is created by a laser pulse which ends, e.g., at $t = 0$. For $t > 0$, the electric fields in (19.2) are zero; nevertheless there are changes of the optical properties for as long as the incoherent population lives.

In order to be able to compare the magnitude of the coherent and incoherent optical nonlinearities, some authors prefer to describe, e.g., the diffraction efficiency of a laser-induced population grating (Chap. 25) as a $\chi^{(3)}$ process also in the case of incoherent population gratings. Though this approach is basically wrong, it may be of some practical use for the above-mentioned purpose especially if quasi-stationary conditions are used. We shall see later an example for such an effective $\chi^{(3)}_{\mathrm{eff}}$.

On the other hand we should note, that one can follow with extremely short laserpulses the temporal build-up and decay of coherent polarisation fields going thus beyond (19.2). Examples are given in Chap. 23 or in [04W1].

At the end of this section, we want to introduce some terms that are closely related to nonlinear optics and which may even be used as synonyms.

At the beginning of the investigation of optical nonlinearities in semiconductors in the 1960s, these physical effects were and sometimes still are known as high-excitation, high-density or many-particle effects.

The reason is that optical nonlinearities can often be observed under high excitation, usually with laser light. Typical values of the light intensities in the field of nonlinear optics range from 10^3 to far above 10^6 W/cm^2. Under these conditions many particles, i.e., electron–hole pairs, are created, really or virtually. This means that they exist in high density. The nonlinearities are due to the interactions between these many particles.

Changes of the optical properties under strong illumination are also known as renormalization effects of the optical properties. When they are induced by a laser of frequency ω_{exc} at this very frequency one thus speaks of self-renormalization of the optical properties.

To conclude this excursion into semantics we may state that optical nonlinearities are the consequence of the many-particle or renormalization effects that occur under high excitation.

19.2 General Scenario for High Excitation Effects

In the following chapters we shall discuss optical nonlinearities in detail. We concentrate on those due to many-particle effects in the electron–hole pair system of semiconductors. This means that we again concentrate on the spectral region around the exciton resonances and the band gap. We later consider interaction processes between excitons and phonons, but we shall not treat nonlinearities that occur in the spectral region of the (optical) phonon resonance, e.g., due to anharmonic phonon–phonon interaction.

A general scenario for the many-particle effects and the resulting optical nonlinearities has been developed over the last four decades. It is shown schematically in Fig 19.3. At low light levels (i.e., in the low-density regime) the optical properties are determined by single electron–hole pairs, either in the exciton states or in the continuum. At low temperatures these excitons may also be bound to some defects to form BEC.

With increasing excitation intensity we reach the so-called intermediate density regime. In this regime, excitons are still good quasi–particles, but their density is so high, that they start to interact with each other. There are, e.g., elastic and inelastic scattering processes between excitons and (at higher temperatures) between excitons and free carriers. These scattering processes may lead to a collision-broadening of the exciton resonances and to the appearance of new luminescence bands, to an excitation-induced increase of absorption, to bleaching or to optical amplification, i.e., to gain or negative absorption depending on the excitation conditions. Another group of coherent and incoherent interaction processes in this interme-

diate density regime involves transitions to the excitonic molecule or biexciton. The biexciton, which was briefly mentioned in Sect. 9.4, is a new quasiparticle which consists of two electrons and two holes. If the exciton can be seen in analogy to the hydrogen or rather the positronium atom, then the biexciton corresponds to the H_2 or the positronium molecule. Similarly, there are so-called trions, which consist of two electrons and one hole ore vice versa. They are also known as X^- and X_2^+, respectively, and correspond to H^- or H_2^+. As we shall see in Sect. 20.3, transitions involving biexcitons or trions give rise to a large variety of optical nonlinearities.

If we pump the sample even harder, we leave the intermediate and arrive at the high density regime, where the excitons lose their identity as individual quasiparticles and where a new collective phase is formed which is known as the electron–hole plasma (EHP). In this regime, the density of

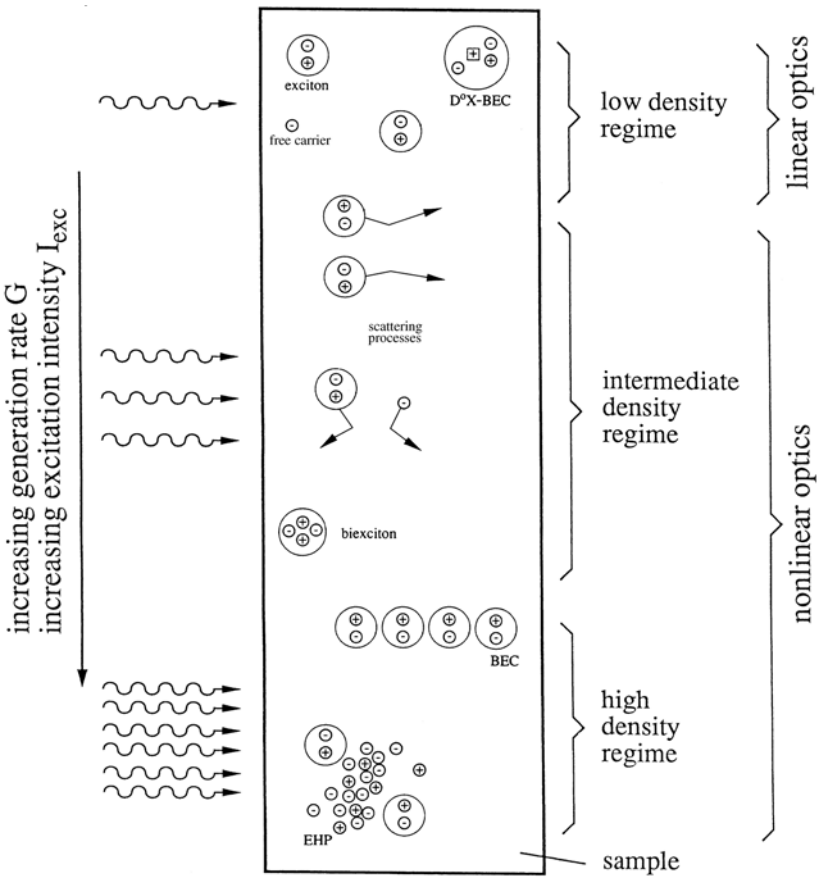


Fig. 19.3. The general scenario for many-particle effects in semiconductors

electron–hole pairs n_p is at least in parts of the excited volume so high that their average distance is comparable to or smaller than their Bohr radius a_B, i.e., we reach a "critical density" n_p^c in an EHP, given to a first approximation by

$$a_\mathrm{B}^3 n_\mathrm{p}^\mathrm{c} \approx 1\,. \tag{19.20}$$

We shall meet some more elaborate formulas than (19.20) in Chap. 21.

In this high-density regime, we can no longer say that a certain electron is bound to a certain hole; instead, we have the new collective EHP phase. As we shall see in more detail in Chap. 21, the transition to an EHP is connected with very strong changes of the electronic excitations and the optical properties of semiconductors.

The scenario outlined in Fig 19.3 has been observed in every group of semiconductors investigated so far including structures of reduced dimensionality. However, the ranges of observability of the various groups of phenomena can be very different. In the indirect gap material Ge it is possible in favorable cases to observe an EHP even under illumination with an incandescent lamp, due to the long lifetime of the carriers which can reach ms in the bulk of high quality samples. In contrast, for CuCl pump powers in the range of GW/cm^2 have to be used to fulfill the condition of (19.20) and to produce an EHP. Such high power densities can be applied only for a few ps, otherwise the sample will be evaporated.

As a rule of thumb, one can say that plasma phenomena are most easily observed in indirect gap materials due to the long carrier lifetime or in direct gap materials with large excitonic Bohr radius, so that (19.20) can also be readily satisfied as in GaAs. The Cu halides, on the other hand, are model substances for biexciton phenomena, due to the large exciton and biexciton binding energies. The II–VI compounds have an intermediate position allowing observation of all of the effects shown in Fig 19.3 (and some more). This is very satisfying but poses the challenge of separating the various contributions, which partly overlap spectrally. Recent reviews which cover high excitation effects are [81K1, 89S1, 90G1, 90H1, 91C1, 94P1, 99N1, 02S1, 04O1] and references therein.

In the following chapters we describe the various effects in some detail, always beginning with three-dimensional materials but also giving examples of systems of reduced dimensionality. We start in Chap. 20 with the intermediate density regime and continue in Chap. 21 with the electron–hole plasma.

It should be added that at low temperatures theory predicts between the intermediate and high density regime the possibilities of excitonic Bose–Einstein condensation and superfluidity. We come back to this aspect in Sect. 20.5.

A Bose–Einstein condensation (also generally abbreviated by BEC) is difficult to depict in a figure like Fig. 19.3 since it is a condensation in $\boldsymbol{k}$-space. We tried very tentatively to do it as shown in Fig. 19.3

19.3 Beyond the $\chi^{(n)}$ Approximations

An expansion of the susceptibility in a power series of the electric field like in (19.2) only makes sense if it converges fast. Though terms up to $\chi^{(7)}$ are discussed in the literature, e.g., in [95A1, 00A1, 01A1], one usually stops with $\chi^{(3)}$ terms.

On the other hand, presently available pulsed lasers bring focus intensities up to $10^{20}\,\mathrm{Wcm}^{-2}$. These intensities, the resulting generation rates, and the related electrical field strengths clearly go far beyond the applicability of the approaches used in Sects. 19.1 and 19.2.

Expected effects have been discussed in [51S1, 64B1, 66K1, 97C1, 98M1, 01C1, 01Z1, 02M1, 02S2, 02S3, 03W1, 04W1] and the references given therein. We give here only a short and limited listing since most of these effects are no longer specific for semiconductors and thus lead beyond the scope of this book.

- The Rabi frequency mentional already, discussed in more detail in Sect. 20.4 and in Chap. 23 and 27, can become comparable to the light frequency, allowing the population to flip coherently from the ground to the excited state and back during a laser pulse with only a few cycles of light. This so-called (carrier-wave) Rabi flopping has been predicted in [37R1] and observed, in semiconductors in [01M1, 02S1].
- Higher harmonics can be generated up to the soft X-ray regime from atoms [66K1, 93H1, 01P1] or from a laser induced plasma on a solid surface [01C1, 02S2].
- Since laser pulses with extreme intensity are often very short and partly comprise only a few cycles of light, the nonlinear properties start to depend on the relative phase of the carrier wave of the light and the maximum of the pulse envelope. If this phase shifts in a controlled manner from pulse to pulse, the resulting carrier envelope offset frequency can be measured and possibly used as a frequency normal (see, e.g., [02M2, 03C1, 04W1]).
- Solids can be melted or evaporated [92B1, 97G1, 00H1, 01C1].
- A plasma can be produced at the surface or in the bulk of solids [97G1,02S2].
- A fusion can be ignited [98M1, 02M1].
- The so-called ponderomotive energy of electrons can reach relativistic energies. The ponderomotive energy is defined as the energy a free electron can reached during acceleration in the light field of a laser pulse. If we assume that the electron is free, initially at rest, and consider only the electric field over one half a period of light, neglecting the $\boldsymbol{B}$ field or relativistic effects, we can easily show that [96J1]

$$E_{\mathrm{pond}} = \frac{p^2}{2m} = \frac{2e^2E_0^2}{m\omega_0^2} \tag{19.21a}$$

with

$$p = \int_0^\pi \dot{p}\,\mathrm{d}t = \int_0^\pi E_0 \sin\omega_0 t\,\mathrm{d}t = 2eE_0\omega_0^{-1} \tag{19.21b}$$

Since the therm $e\boldsymbol{v}\times\boldsymbol{B}$ of the Lorenz-force depends via (19.21b) quadratically on the field amplitude it can give rise to photon-drag effects [03H1]. E_{pond} can easily exceed the width of the forbidden gap of a semiconductor or the work function of electrons by orders of magnitude, thus giving only limited importance to whether an electron is initially free or bound in an atom or a solid. Expected effects that are partly observed are, e.g., nuclear fusion or second harmonic generation from a single electron in vacuum. Effects that are expected for the future are, e.g., nonlinear optics involving the Dirac sea with effects like e^-e^+ pair production in vacuum with two counter propagating light beams (see, e.g., [98M1, 00C1, 02M1, 03W1, 04W1] and references therein).

Obviously, exciting phenomena are emerging in this field.

19.4 Problems

1. Try to make a rough guess of the electric field strength due to the positively charged nucleus of Si at the distance of the outer electron shell. At what light intensities would you expect to observe coherent optical nonlinearities assuming that the electric field in the light beam is about 0.1 of the electric field from the nucleus. How do the field strength and the light intensity change if you consider excitons instead of atoms?
2. Calculate the light intensity necessary to create in a semiconductor with $E_g \approx 1.5\,\mathrm{eV}$ a stationary density n_{p} of $10^{18}\,\mathrm{cm}^{-3}$ electron–hole pairs, if their lifetime is 0.1 ns (direct gap semiconductor) or 1 µs (indirect gap semiconductor).
3. Give a schematic drawing as in Fig. 19.2 for a three-photon absorption process and for third-harmonic (3ω) generation. Give the formula for the three-photon absorption coefficient β_3 and for the third-harmonic generation in perturbation theory. Can third-harmonic generation occur in centrosymmetric crystals?
4. Try to find some more processes which contribute to $\mathrm{Re}\{\tilde{n}_2(\omega)\}$.
5. Calculate the critical densities $n_{\mathrm{p}}^{\mathrm{c}}$ for some semiconductors using the formula (19.20).
6. Let a parallel laser beam with a lateral Gaussian envelope fall on a sample with $\mathrm{Re}\{\tilde{n}_2\} \neq 0$. Explain qualitatively how self-focussing or defocussing works.
7. If the pulse of problem 6 also has a short temporally Gaussian envelope, one expects so-called self-phase modulation, which broadens the spectrum beyond the Fourier limit (the so-called chirp mentioned in the text). Try to figure out qualitatively how this effect works.

References to Chap. 19

[31G1] M. Göppert-Mayer, Annalen der Physik **9**, 273 (1931)
[37R1] I.I. Rabi, Phys. Rev. **51**, 652 (1937)

[51S1] J. Schwinger, Phys. Rev., **82**, 664 (1951)
[61K1] W. Kaiser and C.G.B. Garrett, Phys. Ref. Lett., **7**, 229 (1961)
[62B1] R. Braunstein, Phys. Rev. **125**, 475 (1962)
[62M1] P.D. Maker et al., Phys. Rev. Lett. **8**, 21 (1962)
[63F1] P.A. Franken and J.F. Ward, Rev. Mod. Phys. **35**, 23 (1963)
[64B1] L.S. Brown and T.W.B. Kibble, Phys. Rev. A **133**, 705 (1964)
[64B2] R. Braunstein and N. Ockman, Phys. Rev. **134**, 499 (1964)
[65B1] M. Bass, P.A. Franken, and J.F. Ward, Phys. Rev. A **138**, 534 (1965)
[65B2] N. Bloembergen, Nonlinear Optics, Benjamin, New York (1965)
[66K1] T.W.B. Kibble, Phys. Rev. A **150**, 1060 (1966)
[67B1] I. Broser in Phys. and Chem. Of II–VI Compounds, M. Aven and J.S. Prener (eds.), North Holland Pupl., Amsterdam p. 487 (1967)
[71K1] C. Klingshirn, Z. Phys. **248**, 433 (1971)
[75A1] L. Allen and J.H. Eberly, Optical Resonance and Two Level Atoms, Wiley Interscience, New York (1975)
[81K1] C. Klingshirn and H. Haug, Phys. Rep. **70**, 315 (1981)
[84S1] Y.R. Shen, The Principles of Nonlinear Optics, Wiley, New York (1984)
[86E1] H.J. Eichler, P. Günter and D.W. Pohl, Laser Induced Dynamic Gratings, Springer Ser. Opt. Sci. Vol. **50**, Springer, Berlin, Heidelberg (1986)
[89N1] Nonlinear Optics of Organics and Semiconductors, Kobayashi, T., ed., Springer Proc. in Physics **36** (1989)
[89S1] S. Schmitt-Rink, D.A.B. Miller and D.S. Chemla, Adv. Phys. **38**, 9 (1989)
[90B1] P.N. Butcher and D. Cotter, The Elements of Nonlinear Optics, Cambridge Studies in Modern Optics Vol. 9, Cambridge University Press, Cambridge (1990)
[90G1] E.O. Göbel and K. Ploog, Progr. Quantum Elektron. **14**, 289 (1990)
[90H1] H. Haug and S.W. Koch, Quantum Theory of the Optical and Electronic Properties of Semiconductors, World Scientific, Singapore (1990)
[91C1] R. Cingolani and K. Ploog, Adv. Phys. **40**, 535 (1991)
[91M1] D.L. Mills, Nonlinear Optics, Springer, Berlin, Heidelberg (1991)
[92B1] U. Becker et al., J. Crystal Growth **125**, 384 (1992)
[93H1] S.E. Harris, J.J. Macklin, and T.W. Hänsch, Opt. Commun. **100**, 487 (1993)
[94P1] K.-H. Pantke and J.M. Hvam, Intern. Journ. of Modern Physics **B8**, 73 (1994)
[95A1] V.M. Axt et al., phys. stat. sol. (b) **188**, 447 (1995)
[95M1] S. Mukamel, Principles of Nonlinear Optical Spectroscopy, Oxford University Press, New York (1995)
[96J1] A. Jauho, Phys. Rev. Lett. **76**, 4576 (1996)
[97C1] J. Cerne et al., Appl. Phys. Lett. **70**, 3543 (1997)
[97G1] E.N. Glezer and E. Mazur, Appl. Phys. Lett. **71**, 882 (1997)
[98M1] G.A. Mouron et al., Physics Today, January Issue p. 22 (1998)
[99N1] Nonlinear Optics in Semiconductors I and II, E. Garmire and A. Kost (eds.), Semicond. and Semimetals **58** and **59** (1999)
[00A1] V.M. Axt et al., phys. stat. sol. (b) **221**, 205 (2000)
[00C1] A.H. Chin, Phys. Rev. Lett. **85**, 3293 (2000)
[00H1] T.-H. Her et al., Appl. Phys. **70**, 383 (2000)
[00N1] Nonlinear Optical Effects and Materials, Günter, P., ed., Springer, Berlin (2000)
[01A1] V.M. Axt et al., Phys. Rev. B **63**, 115303 (2001)

[01C1] A. Cavalleri et al., Phys. Rev. B **63**, 193306 (2001)
[01P1] P.M. Paul et al., Science **292**, 1689 (2001)
[01M1] O.D. Mücke et al., Phys. Rev. Lett. **87**, 057401 (2001)
[01Z1] M. Zudov et al., Phys. Rev. B **64**, 121204 (R) (2001)
[02M1] G.A. Mouron and D. Umstadter, Scientific American, May Issue p 63 (2002)
[02M2] O.D. Mücke et al., Opt. Lett. **27**, 2127 (2002)
[02S1] W. Schäfer and M. Wegener, Semiconductor Optics and Transport Phenomena, Springer, Berlin (2002)
[02S2] K. Sokolowski-Tinten et al., Springer Series in Chemical Physics **71**, p. 36, 102, 499 (2002)
[02S3] M.Y. Su et al., Appl. Phys. Lett. **81**, 1564 (2002)
[03C1] S.T. Cundiff and J. Ye, Rev. Mod. Phys. **75**, 325 (2003)
[03H1] F. He, Phys. Rev. Lett. **90**, 055002 and **91**, 195001 (2003)
[03W1] M. Wegener, Intern. School on Frontiers of Optical Spectroscopy, Erice, May (2003), Kluwer, Dordrecht, in press
[04O1] Optics of Semiconductors and their Nanostructures, H. Kalt and M. Hetterich (eds.), Springer Series in Solid State Science **146**, (2004)
[04W1] M. Wegener, in Ref. [04O1], p. 171

20

The Intermediate Density Regime

In the following sections we present selected examples from the intermediate density regime where excitons biexcitons and trions are still good quasiparticles.

20.1 Two-Photon Absorption by Excitons

We have mentioned already the two-photon absorption (TPA) to the exciton level in connection with $\boldsymbol{k}$-space spectroscopy of polaritons (Sect. 13.1.4) and with the magnetic properties of excitons (Sect. 16.1). The process is described in the weak coupling limit by (19.9), i.e., a first photon virtually excites an intermediate state which is converted by a second photon to the exciton in the final state. In the strong coupling limit, one would say that two photon-like polaritons merge to form one exciton-like polariton. The momentum conservation

$$\hbar\boldsymbol{k}_1 + \hbar\boldsymbol{k}_2 = \hbar\boldsymbol{k}_{\mathrm{exc}} \tag{20.1}$$

usually allows only the longitudinal exciton and the upper polariton branch (UPB) to be reached. The two-photon selection rules are different from those for one photon. Consequently it is sometimes possible by TPA to reach states that are forbidden in one-photon absorption and vice versa. Examples are given in Chap. 26 on group theory. Experimental result for the two-photon spectroscopy of excitons have already been given in Figs. 13.20, 13.23, 13.30 and in Fig. 16.5. More details, including three-photon absorption, are found in [81F1].

More recent data on two- and three-photon absorption spectroscopy e.g. of ZnTe or ZnO and on the influence of stress are found in [93F1, 98W1]. Some examples for two-photon spectroscopy of excitons in quantum-wells can be found in [01L1] of Chap. 1.

The two-photon absorption coefficient β of (19.8), (19.9) has been investigated for various bulk semiconductors in the band-to-band transition

region in, e.g., [80K1]. More data for this quantity as well as for second harmonic-, sum- and difference frequency generation in bulk or at surfaces or for the dc-effect can be found e.g. in [71K1, 84S1, 97C1, 01Z1, 02S3] and the books [65B2, 75A1, 86E1, 90B1, 95M1, 00N1] both of Chap. 19. Some more recent data on second harmonic generation including quantum structures are given e.g. in [98W1, 01L1, 01W1]. The creation of squeezed light via an n_2 process (see (19.3)) has been reported in [95F1]. We will come across further examples of two-photon absorption and $\chi^{(3)}$ and higher processes in Sects. 20.3 and 23.2.

20.2 Elastic and Inelastic Scattering Processes

If we increase the density of really excited excitons, it can happen that two excitons meet during their lifetimes on their diffusive motion through the sample and scatter via their mutual dipole-dipole interaction while a dipole-monopole interaction would describe the scattering of excitons with free carriers. Scattering of excitons with LO-phonons has already been introduced in Sect. 13.1. For an early description of such scattering processes see [69B1] and for a review [81K1]. All of these scattering processes will disturb the phase of the exciton (exciton-like polaritons). We will meet an example of this phenomenon in Chap. 23.

The scattering processes themselves can be categorized into elastic and inelastic. For elastic scattering the sums of kinetic energies before and after the collision are equal, in addition to momentum conservation. For an elastic scattering event between two excitons we obtain, e.g.,

$$\boldsymbol{k}_{\mathrm{i},1} + \boldsymbol{k}_{\mathrm{i},2} = \boldsymbol{k}_{\mathrm{f},1} + \boldsymbol{k}_{\mathrm{f},2} \tag{20.2a}$$

$$\frac{\hbar^2}{2M}\left(\boldsymbol{k}_{\mathrm{i},1}^2 + \boldsymbol{k}_{\mathrm{i},2}^2\right) = \frac{\hbar^2}{2M}\left(\boldsymbol{k}_{\mathrm{f},1}^2 + \boldsymbol{k}_{\mathrm{f},2}^2\right) \tag{20.2b}$$

in analogy to the definition in classical mechanics. These processes show up mainly in the reduction of the phase relaxation time T_2 i.e. in an increase of the homogeneous broadening. In the inelastic processes on which we shall concentrate now, an exciton is scattered into a higher excited state with principal quantum number $n_{\mathrm{B,f}} \geq 2$, while another is scattered on the photon-like part of the polariton dispersion and leaves the sample with high probability as a luminescence photon, when this photon-like particle hits the surface of the sample. This process is shown schematically in Fig. 20.1. The momentum conservation law is given by (20.2a) with one of the $\boldsymbol{k}_{\mathrm{f}}$ being ≈ 0. If we assume that both excitons are initially in states $n_{\mathrm{B,i}} = 1$ and that the momentum of the photon-like polariton is zero, then energy conservation reads

$$\hbar\omega_{P_{n_{\mathrm{B,f}}}} = E_{\mathrm{ex}}(n_{\mathrm{B}} = 1, \boldsymbol{k} = 0) - E_{\mathrm{ex}}^{\mathrm{b}}\left(1 - \frac{1}{n_{\mathrm{B,f}}^2}\right) - \frac{\hbar^2}{M}\boldsymbol{k}_{\mathrm{i},1}\boldsymbol{k}_{\mathrm{i},2}\,. \tag{20.3}$$

The resulting emission bands are usually called P-bands with an index given by $n_{\mathrm{B,f}}$. The bands are broadened by averaging over the last term on the right of (19.3) and by the fine structure of the exciton states, e.g., the splitting of the $n_{\mathrm{B}} = 2$ exciton into states with S, P_0 and P_1 envelope functions.

The transition probability into the continuum (P_∞) decay rather fast with increasing excess energy. A summary of the calculations for the transition probabilities and further references are given in [69B1, 81K1, 94K1].

In the simplest approximation, one expects that the luminescence intensity of these scattering processes increases quadratically with the density of excitons. Indeed one finds a superlinear increase of these bands with increasing pump power with exponents ranging from 1.5 to 2.

In Fig. 20.2 we give an example for ZnO showing the P_2 and P_∞ bands. References to further experimental results are compiled in [75K1, 81K1, 85H1, 94K1]. Fo r more recent data in semiconductors like ZnO or CuI see [97Z2,02T1]

At higher lattice (and exciton) temperatures a fraction of the excitons will be thermally ionized. In this situation a similar inelastic scattering process becomes possible, in which an $n_{\mathrm{B}} = 1$ exciton-like polariton is again scattered onto the photon-like branch of the dispersion curve, while a free carrier (electron or hole) is scattered under energy and momentum conservation into a higher state in the respective band. A characteristic feature of the resulting rather broad and unstructured emission band is that its maximum shifts with increasing temperature considerably faster to lower energies than the band gap does. Examples are given in Fig. 20.3, for CdS and ZnO. While the basis for exciton-exciton scattering is the dipole-dipole interaction, we have to consider dipole-monopole interaction for the scattering between excitons and free carriers. For exaples see [76H1, 78K1, 78M1, 81K1].

In the literature one can find many other inelastic scattering processes not considered here in detail involving both free and localized states, for example, biexciton-biexciton scattering, scattering between a bound-exciton complex and a free carrier, and scattering processes involving phonons. A review is given in [81K1].

Most of these inelastic scattering processes give rise to the appearance of new emission bands (which usually grow more than linearly with the generation rate), to induced absorption and, eventually, to optical gain. The latter point will be addressed specifically in Chap. 22.

The inelastic scattering processes have been studied in great detail in the hexagonal and cubic II–VI semiconductors, but they have also been observed in III–V and I–VII compounds [69B1, 73H1, 74H1, 75G1, 75K1, 76H1, 78K1, 78M1, 81K1, 86U1, 90P1, 91C1, 94K1, 97Z2, 02T1].

A density-dependent scattering between exciton polartons and LA phonons leads to the concept of phonoritons, e.g., in Cu_2O [99H1].

Inelastic scattering processes between excitons or between excitons and free carriers have also been observed in quantum wells [01L1] and references therein.

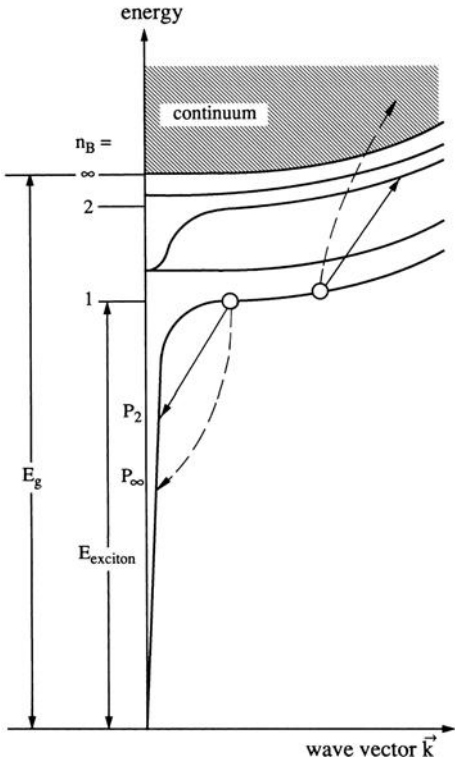


Fig. 20.1. Schematic representation of the inelastic exciton-exciton scattering processes P_2 and P_∞

Presently we see a lot of activities to extend or to reinvent the concept of inelastic scattering processes to cavity polaritons, introduced in (17.1). Examples for exciton-exciton or polariton-polariton scattering are found e.g. in [00B1, 00T1, 02B4, 02B5, 02H1, 02S5] and for electron-polariton scattering in [03L1]. Energy conservation has to be valid again similarly to (20.2) while $\boldsymbol{k}$ conservation is restricted to the two-dimensional $\boldsymbol{k}_\parallel$-space. As it is frequently the case in such a situation, partly also the nomenclature changes. The photon-like lower polariton branch, which shows more a parabolic dispersion in microcavities (Fig. 17.2a) than a light cone, is named $\boldsymbol{k}$-space trap, the fulfillment of $\boldsymbol{k}$ conservation involves a "magic" angle, and the polaritons themselves are introduced as "mysterious" particles, which are half light, half matter [02B4]. See also Fig. 13.2

20.3 Biexcitons and Trions

We start with the discussion of biexcitons and trions for bulk material in Sect. 20.3.1 and proceed to structures of reduced dimensionality in Sect. 20.3.2

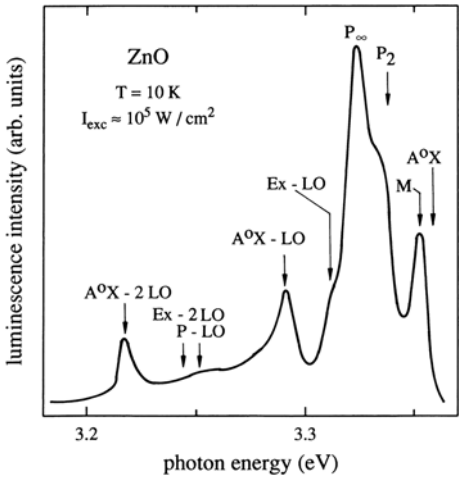


Fig. 20.2. The P_2 and P_∞ bands in the luminescence spectra of ZnO. According to [73H1, 74H1]

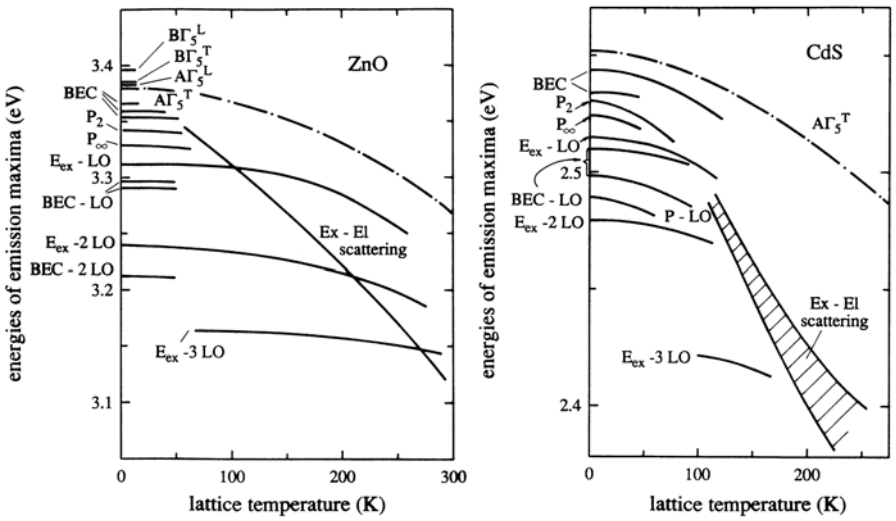


Fig. 20.3. The temperature dependence of various emission band maxima in CdS (**b**) and ZnO (**a**). According to [75K1, 81K1, 94K1]

including in both cases also external fields, generally magnetic fields. The concept of excitons and trions has already been introduced in Sect. 9.4.

20.3.1 Bulk Semiconductors

The idea that biexcitons might exist as bound states of two excitons has been put forward in [58L1, 58M1]. Relatively recent reviews of this topic are,

e.g., [81K1,85H1,86U1,98I1]. Data for bulk semiconductors and for structures of reduced dimensionality are compiled also in [99L1, 02L1] and [01L1, 04L1], respectively.

The biexciton or excitonic molecule is, as already mentioned, a quasiparticle consisting of two electrons and two holes, in analogy to the H_2 or the positronium molecule.

Its dispersion relation is given, in the simplest approximation, by

$$E_{\mathrm{biex}}(\boldsymbol{k}) = 2E_{\mathrm{ex}}(n_{\mathrm{B}} = 1, \boldsymbol{k} = 0) - E^{\mathrm{b}}_{\mathrm{biex}} + \frac{\hbar^2 \boldsymbol{k}^2}{4M}\,, \tag{20.4}$$

where one assumes that the envelope function in the ground state is symmetric under the exchange of equal particles, while the combination of the electrons and of the holes is antisymmetric including spin.

General calculations have been made of the biexciton binding energy normalized by the exciton binding energy as a function of the ratio of electron and hole mass $\sigma = m_{\mathrm{e}}/m_{\mathrm{h}}$ (Fig. 20.4). The experimental data points for E_{biex} are normalized to the experimentally determined values of $E^{\mathrm{b}}_{\mathrm{ex}}$ or to the calculated excitonic Rydberg energy. We have seen already in Sect. 9.2 that these two quantities partly disagree. What is important from the theoretical point of view is that the biexciton exists as a bound state for all values of σ and that the curve $E^{\mathrm{b}}_{\mathrm{biex}}/\mathrm{Ry}^*$ decreases monotonically in the range $0 \leq \sigma \leq 1$. The data points scatter around the theoretical predictions with an accuracy comparable to that of the various calculations.

The semiconductor Cu_2O (see Sect. 13.2.1.2) is possibly an exception. Calculations indicate that the biexciton binding energy vanishes, though the exciton binding energy is about 150 meV [76B1]. This fact is cused by the large ortho-para exciton splitting.

In the following we concentrate on optical nonlinearities connected with the creation and the decay of biexcitons. For further reading see [76N1, 81K1, 85H1, 86U1, 91C1, 94K1, 01K1].

The probability of creating two electron–hole pairs with one photon is very low, but the probability for two-photon excitation is rather high and is sometimes said to have a "giant" oscillator strength, since the virtually created intermediate state is almost resonant with the exciton if photons with $\hbar\omega_{\mathrm{exc}} \approx E_{\mathrm{biex}}/2$ are used.

In connection with Fig. 20.5 we discuss this process as a two-polariton transition. We give the real part of $\boldsymbol{k}$ and schematically the imaginary part. The latter is close to zero for low damping away from the exciton resonance (see Figs. 4.5 and 5.1). If we shine (laser-) light of an energy $\hbar_{\mathrm{exc}}$ onto the sample, we populate this state on the lower polariton branch (LPB). A second polariton with an energy

$$\hbar\omega_{\mathrm{abs}} = E_{\mathrm{biex}} - \hbar\omega_{\mathrm{exc}} \tag{20.5}$$

can accomplish the transition from there to the biexciton. This means that we get an absorption dip or a peak in Im k at $\hbar\omega_{\mathrm{abs}}$ with an oscillator-

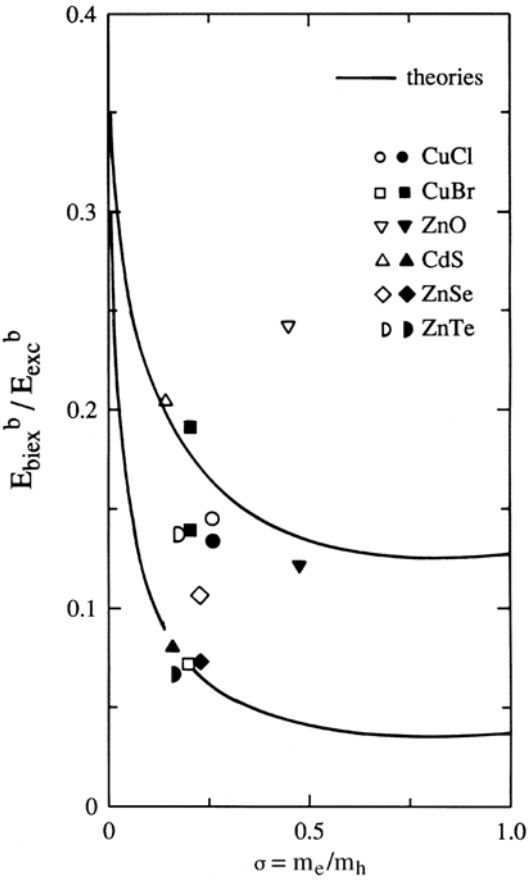


Fig. 20.4. Calculations of the binding energy of the biexciton in units of the excitonic Rydberg energy Ry^*. Experimental data for various semiconductors, normalized to the experimentally observed exciton binding energy (*open symbols*) and to the calculated Ry^* energy (*full symbols*). According to [81K1]. Theories according to [72A1, 72A2, 72B1, 73H2]

strength which increases with the population at $\hbar\omega_{\rm exc}$ on the LPB, evidently an optically nonlinear effect. The peak in Im k is necessarily connected with a resonance-like structure in the real part of the dispersion relation. In the following we give examples of both phenomena, but first we should mention that there is another way to create biexcitons. This starts from exciton-like polaritons, which may even have an incoherent, e.g., thermal population. In the latter case an exciton is converted into a biexciton by absorption of a photon of energy $\hbar\omega'_{\rm abs}$ fulfilling energy and momentum conservation. Considering the different curvatures, i.e., effective masses, of exciton and biexciton dispersion, this process yields an induced absorption at

$$\hbar\omega'_{\rm abs} = E_{\rm ex}(n_{\rm B} = 1, \boldsymbol{k} = 0) - E^{\rm b}_{\rm biex} - \frac{\hbar^2 \boldsymbol{k}_{\rm i}^2}{4M}\,, \tag{20.6}$$

where M is the translational exciton mass, $2M$ the biexciton mass and $\boldsymbol{k}_{\rm i}$ the momentum of the exciton-like polariton in the initial state, which equals the momentum of the biexciton in the final state if we neglect the wave vector

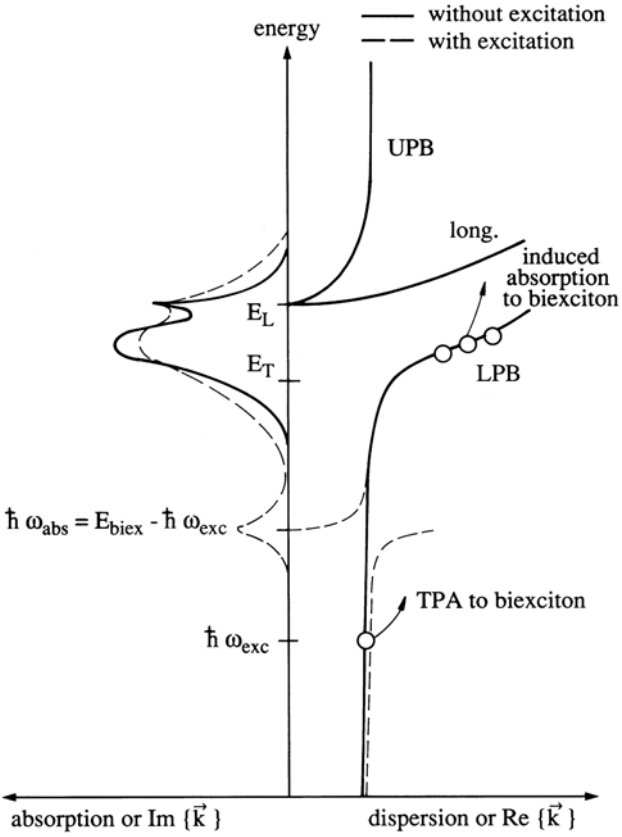


Fig. 20.5. The two-photon (or two-polariton) transition from the crystal ground state to the biexciton state and the two-step process starting from the exciton state

of the photon-like polariton. This process is also known as a two-step process to distinguish it from the two-polariton (or two-photon) transition discussed first. See also Sect. 19.1.

Actually one difference lies in the fact that in the latter case one of the polaritons is either on the photon- or on the exciton–like branch, respectively. A second difference is that the laser at $\hbar\omega_{\rm exc}$ generally produces a more-or-less coherent population on the LPB while the population on the exciton-like branch produced, e.g., by some resonant or non-resonant pump laser, tends to lose its coherence within a few ps (see Sect. 23.2).

In Fig. 20.6 we give an example of the various processes in a pump and probe beam experiment for CuCl. One can see the scattered, spectrally narrow pump laser light $\hbar\omega_{\rm exc}$ and the absorption dips in the probe continuum, which correspond to a bound-exciton complex labelled BEC, the TPA process of (20.5) which shifts oppositely to the energy of the laser, and the two-step or induced absorption process of (20.6) which is fixed in energy and has a larger

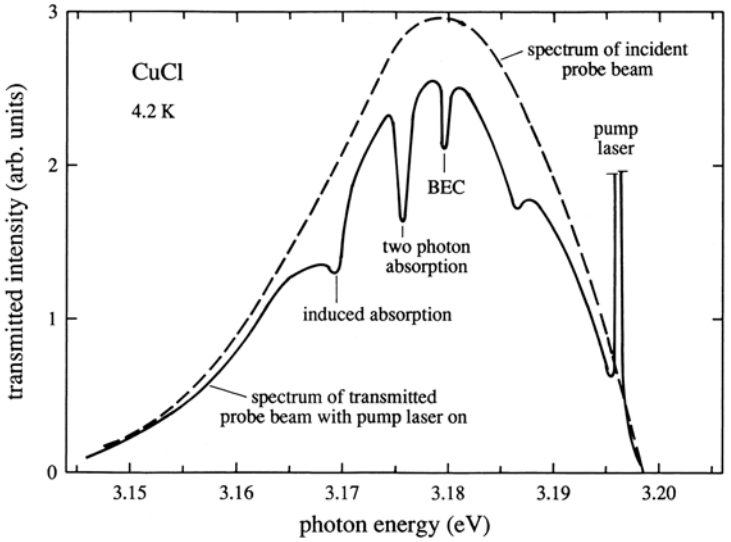


Fig. 20.6. Two-photon and two-step transitions to the biexciton observed in CuCl. According to [85H1]

width due to the incoherent distribution of the excitons on their dispersion curve.

By varying the wave vectors of pump and probe beams relative to each other, it is even possible to measure the dispersion relation of the biexciton in a TPA experiment. An example is shown in Fig. 20.7 for CuBr. Due to the four-fold degeneracy of the $\boldsymbol{\Gamma}_8$ valence band, the biexciton ground state splits into three states with symmetries $\boldsymbol{\Gamma}_1$, $\boldsymbol{\Gamma}_5$ and $\boldsymbol{\Gamma}_3$. The $\boldsymbol{\Gamma}_1$ state shows the simple dispersion of (20.4), while the higher states show some $\boldsymbol{k}$-dependent splitting.

The dispersion of the biexciton has been also measured in CuCl [84K1, 85H1] or CdS [82L1]. In [98I1] it has been shown, that the polariton-polariton interaction in a biexciton may result in a modification of the dispersion relation of the biexciton. For experimental data see, e.g., [01M1].

The decay of biexcitons into an exciton-like and a photon-like polariton is just the reverse process of the two-step transition of (20.6).

Energy and momentum conservation show, that these luminescence bands, which are frequently called M-bands, have an inverted Boltzmann line shape. Consequently the binding energy of the molecule has to be deduced not from the emission maximum but from its high energy edge (see, e.g., [81K1]).

Furthermore it should be noted that the biexciton luminescence generally occurs in the same spectral range as the emission from bound exciton complexes like A°X and D°X, which are also made up of two positively and two negatively charged particles. Side bands of these complexes originating, e.g., from emission of acoustic phonons or inelastic scattering with free carriers

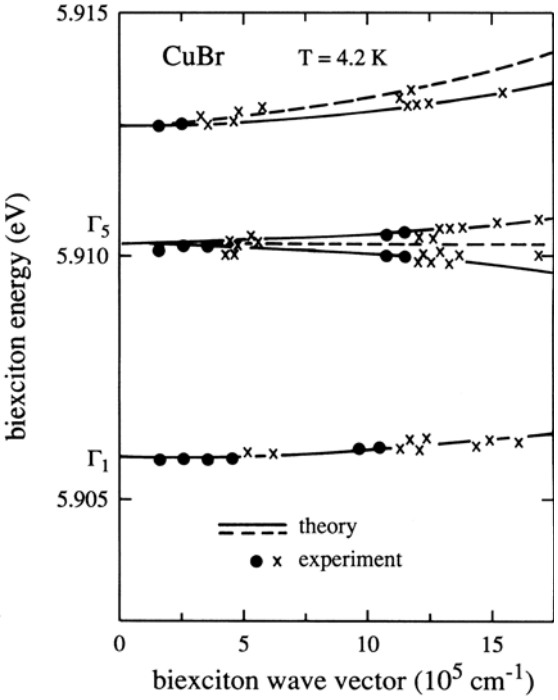


Fig. 20.7. The dispersion of the lowest biexciton states in CuBr. According to [85H1]

often look very similar to biexcitonic M bands and can be easily confused. For more details on this topic see [81K1].

Figure 20.8 shows biexciton luminescence spectra of CuCl under non-resonant and resonant high excitation. In the first case excitons are created, which form biexcitons with a thermal distribution. The temperature of the exciton and the biexciton gas can be higher than the lattice temperature T_{L} which describes the phonon system and, in the situation of Fig. 20.8, lies around 25 K.

In the decay process a longitudinal exciton or a transverse exciton-like polariton can appear in the final state together with the photon-like luminescence. Consequently one observes in cubic materials the so-called M_{L} and M_{T} bands. Under resonant two-photon excitation of the biexciton, one creates a biexciton gas with a narrow, non-thermal distribution, the decay of which gives rise to narrow emission structures N_{T} and N_{L} at the high energy sides of the M_{L} and M_{T} bands, respectively. This spectral position corresponds to the recombination of biexcitons with small wave vectors [76L1, 79H1, 85H1]. Similar effects are also known for CuBr [85H1, 86U1] or for CdS [79S1] or for β-ZnP_2 [01D1].

For (inconclusive) attempts to observe a Bose–Einstein condensation of biexcitons see [76N1, 82P1] and Sect. 20.5.

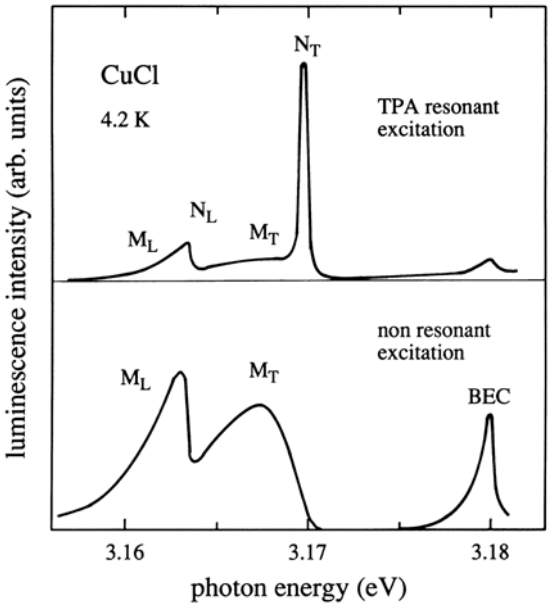


Fig. 20.8. Biexciton luminescence of CuCl under resonant and non-resonant excitation. According to [85H1]

As stated in Chap. 26, the angular momentum can be used (with considerable care) for arguments concerning selection rules, especially for cubic materials. This fact has been used for beautiful experiments with circularly polarized light. Starting from the Γ_1 ground state a Γ_1 biexciton state can only be reached by two quanta with opposite circular polarization as verified in [78I1, 82I1, 82K3, 84K1, 86I1, 86K1, 96H1]. This approach was also used later in quantum wells wires and dots to prove the assignment of an observed transition to biexcitons. See e.g. [96P1, 99W1, 03P1] and references therein.

The diamagnetic shift of the biexciton has been measured, e.g., in CdS [82K1]. It is larger than that of the exciton due to the larger radius of the biexcitons.

In the II–VI compounds, the emission bands of biexcitons are less structured due to the orientation-dependent mixed mode final states (see Sect. 13.1 and [81K1,85H1]). Furthermore, in many II–VI compounds the biexciton emission is spectrally almost degenerate with other recombination processes, especially ones involving bound-exciton complexes like inelastic BEC-free-carrier scattering or the acoustic phonon side-band of EEC. Therefore a clear-cut proof of the existence and the properties of biexcitons is difficult to obtain in these materials from luminescence alone. Details of this topic are given in [81K1]. Some recent work on biexcitons in epitaxial layers of wide gap semiconductors such as ZnO, ZnS or GaN is found in [97Z1,00K1,00Y1,00Y2,03A1, 03H2, 04A1]. As earlier in bulk ZnO [83H1] biexcitons have been observed re-

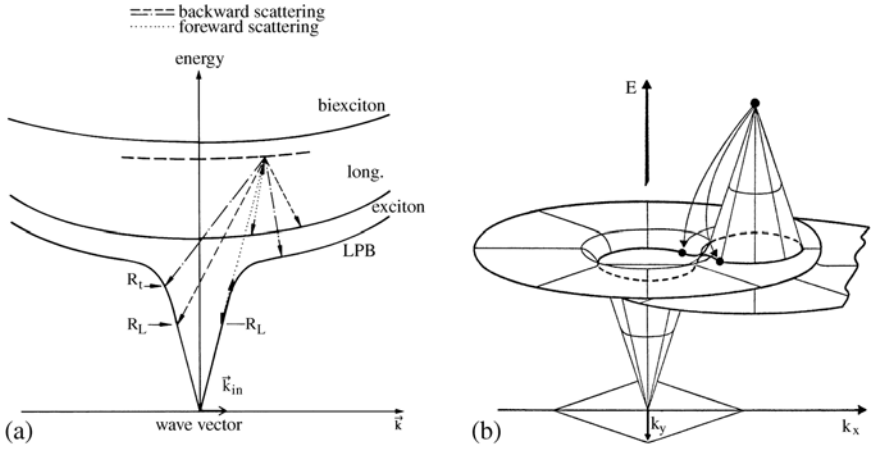


Fig. 20.9. The two-photon Raman scattering via intermediate exciton and biexciton states in a backward scattering geometry (**a**) and for the lower polariton branch over a two-dimensional $\boldsymbol{k}$ space (**b**) [82K2]

cently with countain two holes from the A valence band, one hole from the A and one from the B valence band for ZnO and GaN and even two holes from the B valence band in ZnO [03A1,03H2,04A1]. Slight discrepancies exist between the energies of the three levels in ZnO, especially concerning the BB biexciton in [83H1,03H2,04A1]. The data in [83H1] have been consistently observed by different groups and in different samples, while it is known [83H1], that resonances in two-photon Ramon scattering, a technique from the group of four-wave mixing experiments (see Chap. 25), may lead to erroneous results.

Excitonic molecules have also been observed in the indirect materials Ge and Si in the intermediate density regime before the onset of plasma formation [79G1, 79M1, 82T1, 84T1, 01R1]. Furthermore, biexcitons have not only been found in I–VII, II–VI and group-IV materials, but also in other less common semiconductors like HgI_2 or AgBr [81K1, 85H1, 99L1, 02L1], so that the formation of biexcitons can be considered a rather general feature in the intermediate density regime at low temperatures. At higher temperatures biexcitons are rapidly ionized and their resonances become so broad that they can generally no longer be detected.

Trions have been p redicted for bulk material, e.g., in [76S1, 97S1], but there is not much of experimental evidence for their existence in contrast to quantum wells (see Sect. 20.3.2).

To conclude this subsection on biexcitons in bulk semiconductors we describe in some detail a coherent process which is doubly resonant and which belongs into the group of $\chi^{(3)}$ phenomena. We refer to the two-photon or hyper-Raman scattering and the associated (non-) degenerate four-wave mixing. The idea is described in the following in both the weak and the strong coupling limits (see Fig. 20.9a,b and compare with Fig. 19.2).

An incident photon, which is almost resonant with the exciton state, is converted by a second photon into a virtually excited biexciton. This biexciton has to decay again after a time Δt determined by the energy-time uncertainty relation. If it decays into two photons, which are identical to the incident ones, this process would describe an intensity induced change of the phase velocity, i.e., a contribution to $\mathrm{Re}\{\tilde{n}_2\}$ in (19.3). However, there are many other decay processes which fulfill energy and momentum conservation.

In a backward scattering process, the virtually excited biexciton at wave vector $\left(\boldsymbol{k}_\mathrm{i}^{(1)} + \boldsymbol{k}_\mathrm{i}^{(2)}\right)$ can decay into a transverse or a longitudinal exciton at $\boldsymbol{k}_\mathrm{f}$ and a Raman-like photon with $\hbar\omega_{\mathrm{R_{t,l}}}$ according to

$$\boldsymbol{k}_\mathrm{R} = \boldsymbol{k}_\mathrm{i}^{(1)} + \boldsymbol{k}_\mathrm{i}^{(2)} - \boldsymbol{k}_\mathrm{f}\,, \tag{20.7a}$$

$$\hbar\omega_{\mathrm{R_{t,l}}} = \hbar\omega_\mathrm{i}^{(1)} + \hbar\omega_\mathrm{i}^{(2)} - E_\mathrm{l,t}\,, \tag{20.7b}$$

Since the transverse and longitudinal eigenenergies do not vary much with k in this range, we get for $\hbar\omega_\mathrm{i}^{(1)} = \hbar\omega_\mathrm{i}^{(2)}$ simply

$$\hbar\omega_{\mathrm{R_{t,l}}} = +2\hbar\omega_\mathrm{i} - E_\mathrm{l,t}\,, \tag{20.7c}$$

where the factor 2 on the right explains the name two-photon Raman scattering.

In forward scattering and for all intermediate scattering geometries ((20.7)) would be the same in the weak coupling limit. In the polariton picture this statement is only true if one particle in the final state is a longitudinal exciton. Indeed, the two-polariton state created by the two incident beams, which is almost resonant with the biexciton, can decay either into a longitudinal exciton and a transverse polariton or into two polaritons. In the latter situation both final state particles are on the transverse branches. One particle is exciton-like in the backward configuration, but moves down into the bottle neck if the scattering geometry is changed towards forward scattering. In this case both final state particles have a certain chance of leaving the sample as a photon. For $\omega_\mathrm{i}^{(1)} = \omega_\mathrm{i}^{(2)}$ they will be situated in scattering geometries close to forward scattering slightly above and below ω_i and they will be therefore called $\hbar\omega_\mathrm{R}^{+}$ and $\hbar\omega_\mathrm{R}^{-}$, respectively. The resulting relation between $\hbar\omega_\mathrm{R}$ and $\hbar\omega_\mathrm{i}$ deviates significantly from the slope-two relation of (20.7c).

In Fig. 20.9b we show a graphical solution. The k_x-k_y plane is the scattering plane which contains $\boldsymbol{k}_\mathrm{i}^{(1)} = \boldsymbol{k}_\mathrm{i}^{(2)} = \boldsymbol{k}_\mathrm{i}$ and $\boldsymbol{k}_\mathrm{f}$. The lower polariton branch is plotted centered around $\boldsymbol{k} = 0$ and again in negative energy direction, starting from $2\boldsymbol{k}_\mathrm{i}$, $2\hbar\omega$; according to (20.7a), (20.7b). The intersection line between both energy surfaces or dispersion relations gives just the solutions of (20.7a-c). An avoided crossing has not been considered in Fig. 20.9b. It leads to the additional modification shown in [98I1] and mentioned above in the context of bipolaritons. We come to various (self-) renormalization effects below, but first discuss the solution obtained, if they are neglected.

Examples for forward and backward scattering are given in Fig. 20.10 for CuCl and various geometries. A self-consistent fit of the data with calculated dispersion curves for the exciton polaritons allows one to determine their dispersion curve with high accuracy, as was mentioned in Sect. 13.4 on $\boldsymbol{k}$-space spectroscopy of excitonic polaritons. It should be stressed that the deviations from the slope-two relation in Fig. 20.10c can be only understood in the polariton or strong coupling picture.

The excitation-induced resonance of Fig. 20.5 shows up as an anomaly in the otherwise smooth relation between $\hbar\omega_{\rm i}$ and $\hbar\omega_{\rm R}$ if one of the incoming or outgoing energies coincides with it. An example of this density dependent renormalization of the polariton dispersion is shown in Fig. 20.11.

The fact that the splitting between $\hbar\omega_{\rm exc}$ and $\hbar\omega_{\rm R_{T\pm}}$ does not converge to zero for higher intensities in exact forward scattering indicates that there are also self-renormalization effects of the dispersion at $\hbar\omega_{\rm exc}$ in the sense of an n_2 effect in (19.3) [85H1]. For further examples of renormalizations of the polariton dispersion with increasing excitation see, e.g., [78I1, 79M2, 82I1, 82K3, 83H1, 84K1, 86K1, 96H2] and the reviews [81K1, 85H2, 98I1]. These references also give examples and/or references for hyper- or two-photon Raman

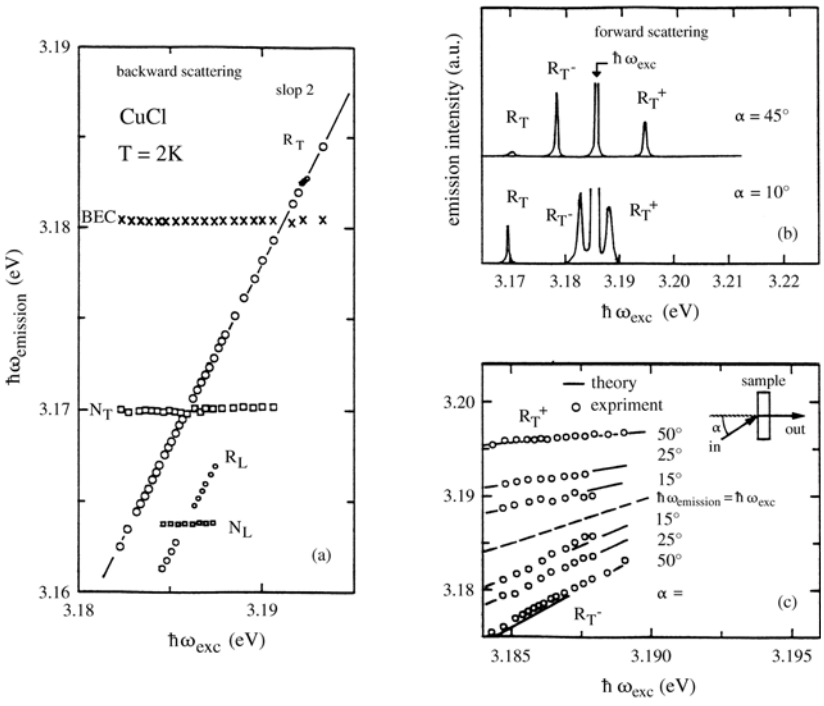


Fig. 20.10. The dependence on $\hbar\omega_{\rm exc}$ of the two-photon Raman emission via an intermediate biexciton state in CuCl for various scattering geometries (**a**,**c**) and a set of spectra for forward scattering (**b**). According to [85H1]

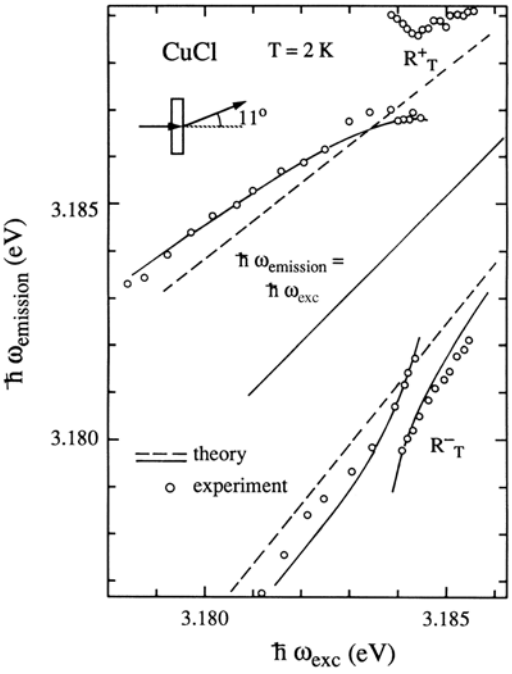


Fig. 20.11. The anomaly in the relation between $\hbar\omega_{\rm exc}$ and $\hbar\omega_{\rm emission}$ caused by the TPA resonance explained with Fig. 20.5. The dashed and solid lines give the theory without and with this resonance. According to [85H1]

scattering in semiconductors other than CuCl, like CuBr or various II–VI compounds.

Until now we have been considering the "spontaneous" decay of the virtually excited biexciton in the intermediate state. It is also possible to stimulate this decay by directing a third beam onto the sample which coincides in direction and energy with one of the particles in the final state. The emission of the other particle is then stimulated and we get a typical $\chi^{(3)}$ process in the sense of Fig. 19.2d. Alternatively we can describe this process as (non-) degenerate four-wave mixing (N)DFWM or as a diffraction of the third beam from a coherent laser-induced grating set up by the interference of the other two incident beams. Actually the close relation between two-photon Raman scattering and (N)DFWM has been verified by various groups, e.g., for CuCl. Details are given, e.g., in [85H1, 86U1]. For a measurement of second order nonlinear susceptibility in various II-VI compounds see [98W2].

20.3.2 Structures of Reduced Dimensionality

Trions, biexcitons and even higher complexes have been observed in many different quantum wells, wires and dots. We give a small selection of results here

and refer the reader for further data to [98I1, 01L1, 04L1] and the references given therein.

In III–V quantum wells and superlattices like GaAs/$Al_{1-y}Ga_yAs$ biexcitons have been predicted and observed experimentally. See, e.g., [83K1, 91B1, 92B1, 92O1, 92P1, 92P2, 93P1, 94P1, 96B1, 97M1, 01E1]. Though there is still some controversy concerning the calculation of their binding energy [83K1,96S1,98I1,99D1,02R1], there is general agreement that the binding energy of the biexciton is around 1.8meV for a GaAs well with a width around 10 nm. Actually it has been found that the ratio of the width dependent binding energies of biexcitons and excitons is roughly constant [96B1, 96S1]. For $In_{1-y}Ga_yAs$ wells see [99B1].

$$E^{\mathrm{b}}_{\mathrm{biex}}/E^{\mathrm{b}}_{\mathrm{ex}} \approx 0.23 \tag{20.8}$$

In Fig. 20.12 we show normalized emission spectra of an $l_z = 10\,\mathrm{nm}$ GaAs/$Al_{1-y}Ga_yAs$ MQW sample at low temperature and for various pump powers.

The peak labeled M on the low energy side of the exciton emission grows more strongly than that of the exciton as shown in the insert. It is attributed to the radiative decay of a thermal population of biexcitons into photons and excitons. The line shape fit gives $E^{\mathrm{b}}_{\mathrm{biex}} = (1.75 \pm 0.05)$ meV in good agreement with other experimental findings, e.g., quantum beats (see Sect. 23.2 or [92O1, 92P1]).

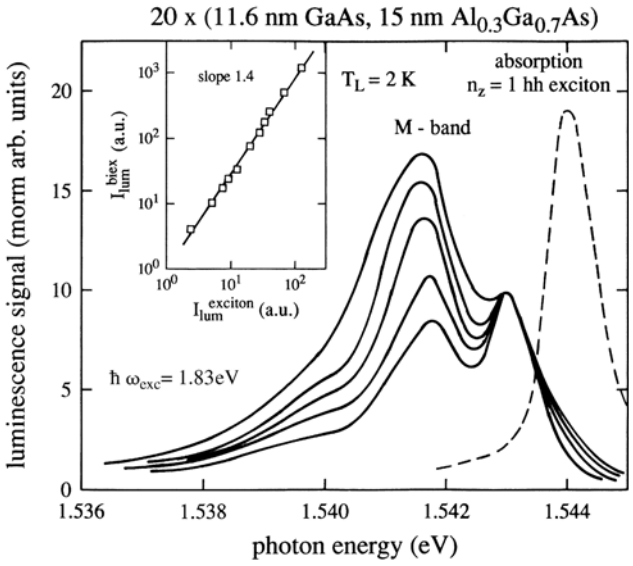


Fig. 20.12. The absorption and normalized emission spectra of an AlGaAs MQW sample, showing the biexciton luminescence band M, which grows superlinearly as shown in the insert. According to [92O1, 93P1, 94P1, 96B1]

Similarly as for excitons, one realized that also biexcitons tend to be localized by, e.g., well width fluctuations. This effect may even enhance the biexcitons binding energy [99L2, 03P1] opening a bridge to the investigation of biexcitons in quantum islands (see below).

Biexcitons have been investigated in external magnetic fields in GaAs and $In_{1-y}Ga_yAs$ wells [94B2, 01B1] partly revealing an unexpected field independence of the biexciton binding energy.

Biexcitons have also been observed in quantum wells embedded in micro cavities (see Sect. 17.1.2) resulting in an additional resonance due to the exciton to biexciton transition [00D1, 03B3].

Trions have been observed in quantum wells by many groups. This fact contrasts the situation in bulk samples (see above) and has the following origin. Quantum wells (or wires) can contain a moderate density of electrons or of holes by slightly doping the barrier. Excitons created by optical pumping then have a rather high probability to bind another carrier resulting in $X^- = (e, e, h)$ or $X_2^+ = (e, h, h)$ trion complexes.

Trions are usually observed in luminescence. Their binding energy is slightly smaller than that of the biexciton. Examples can be found, e.g., in [95F2,96F1,96O1,99L4,00E1,03P1] and in the references given in [01L1,04L1]. For charge transport by X^- see, e.g., [01S1].

Until now we have mainly presented references for III–V, GaAs-based quantum wells, a group of semiconductors in which biexcitons are hardly detectable in bulk materials due to their small binding energy.

In contrast, biexcitons are already well known in bulk II–VI semiconductors as shown in Sect. 20.3.1. Consequently they must also exist in II–VI quantum wells with an even confinement enhanced binding energy. In the following we give a small selection of references. More will be compiled in [03N2, 04L1].

The binding energy of biexcitons reaches values up to 6 meV in ZnSe-based quantum wells as shown, e.g., in [02W1]. The ratio $E^{\mathrm{b}}_{\mathrm{biex}}/E^{\mathrm{b}}_{\mathrm{ex}}$ varies here systematically from 0.10 to 0.23 with decreasing l_z in contrast to (20.8), but is at least in the same range.

For localized excitons and biexcitons values up to $E^{\mathrm{b}}_{\mathrm{biex}} \approx 8,5\,\mathrm{meV}$ are reported in [96P1, 97L1, 99G1, 99L2, 00W1] in agreement with findings for GaAs-based structures mentioned above.

Biexcitons containing one hh and one lh have been identified using the selection rules for circularly polarized light in [96P1, 99W1]. This finding is in close analogy as for biexcitons in bulk ZnO, which may contain two holes from the A valence band, one from the A and another from the B valence band, or two B holes as discussed above [83H2].

Biexcitons in semi-magnetic semiconductors are treated in [01M1] and stimulated emission in [95K1]. Biexcitons in type II ZnSe/BeTe superlattices are observed in [01M1, 04J2]. Eletrons and holes reside in the two different materials.

The existence and the properties of trions in II-VI quantum wells are also well established. They show up in luminescence, reflection and transmission spectra [02A1, 02G1].

Their binding energies tend to be larger for X^- compared to X_2^+. The influence of a magnetic field on this quantity has been investigated in [02A1]. The differences between electron singlet and triplet states of X^- are elaborated in [02G1]. The trion reflection signal has been used to determine the carrier density [02A2] and optical gain connected with the recombination of trions is reported in [01P1].

For quantum wires, the situation is similar as for quantum wells. The existence of biexcitons and of trions is well established in nanostructured, T-shaped and V-groove wires. For biexcitons see, e.g., [98B1, 98V1, 99L3, 01L1, 02W1, 04L1]. It has been found in [98B2] that the binding energy of the biexciton is 40% higher than in a corresponding quantum well.

For trions, data can be found in [04K1]. For recent reviews see [93B1, 97W1, 01L1, 04L1].

Biexcitons in quantum dots in glasses have been observed both in luminescence and in pump-and-probe beam spectroscopy. The biexciton ground state is lower in energy than twice the energy of a single exciton in a dot [90H1]. Due to the three-dimensional confinement there is also a large number of excited states of the biexciton or two electron–hole pair system in a dot [90U1, 93B1, 94K1, 97W1].

In the following we show two examples. Figure 20.13a gives luminescence spectra of CuBr dots in a glass matrix for increasing excitation intensity.

The peak on the high energy side is due to the recombination of single excitons in the dots. The peak on the low energy side, which grows more rapidly with increasing I_{exc} than the other one, is attributed to the recombination of a biexciton, giving a photon and a single exciton in the dot. This interpretation is not only confirmed by the superlinear increase of the biexciton luminescence band but also by the dependence of the emission maxima on the (inverse squared) average dot radius R. The extrapolation to large dot radii converges to the energies of the free exciton and the lowest biexciton luminescence band in bulk samples. The binding energy of the biexciton as a function of R agrees well with theory (Fig. 20.13c).

In Fig. 20.14 we show the inhomogeneously broadened absorption spectra of two (photo-darkened) glass samples containing CdSe quantum dots and the change of the optical density for various pump photon energies $\hbar\omega_{\mathrm{exc}}$.

In Fig. 20.14a,b we observe a bleaching or spectral hole burning at $\hbar\omega_{\mathrm{exc}}$ in the inhomogeneously broadened absorption band, and at a second energy for the dot size of Fig. 20.14a. These two features correspond to the saturation of the transitions form the two uppermost quantized hole states to the lowest conduction band state.

The two structures of induced absorption marked by arrows, correspond to transitions in dots containing one exciton to the ground and an excited state of the biexciton, respectively.

Quantum dots of various II–VI compounds have been also produced by precipitation in organic solvents. For biexcitons in these materials see, e.g., [97W1, 02M1, 02M2].

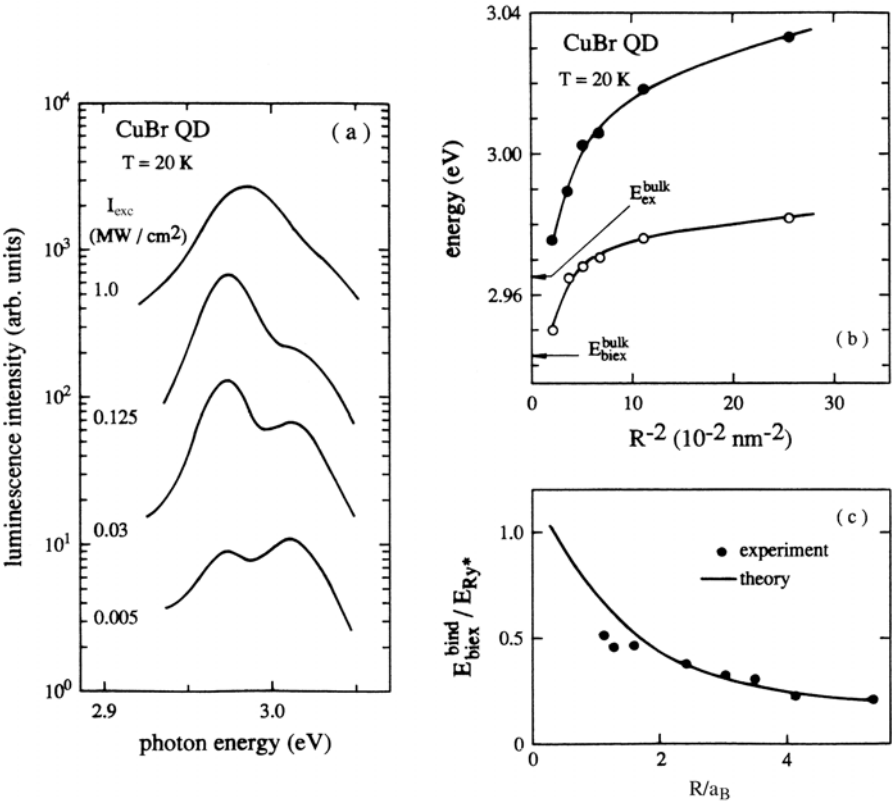


Fig. 20.13. The normalized emission of CuBr quantum dots for various excitation intensities (**a**); the spectral positions of the emission maxima as a function of the average dot radius $\bar{R}$ (**b**); and the binding energy of the biexciton compared to theory (**c**). According to [90H1, 94W1]

Presently, self-assembled quantum islands are preferentially investigated and dots grown on the top of nanopyramids. Model substances for the first type are InAs/GaAs and CdSe/ZnSe and for the second GaAs/$Al_{1-y}Ga_yAs$ but other combination are also used. For some recent data see [01W2, 02A3, 03K2, 03P1, 04K2, 04L2] and Sects. 8.13 and 15.4. After the observation of biexcitons in many different groups of semiconductors and in all (quasi-) dimensions from three to zero, their verification in self assembled islands was not really a surprise. Currently the work concentrates on the spectroscopy of single or a few dots or islands, generally under the influence of a magnetic field and under exploitation of the selection rules for (circularly) polarized light. Examples for such type of investigations are, e.g., [98K1] where binding energies of biexcitons in $In_{0.6}Ga_{0.4}As$ islands of 3.1 meV have been reported. For trions see, e.g., [01B2, 02B1].

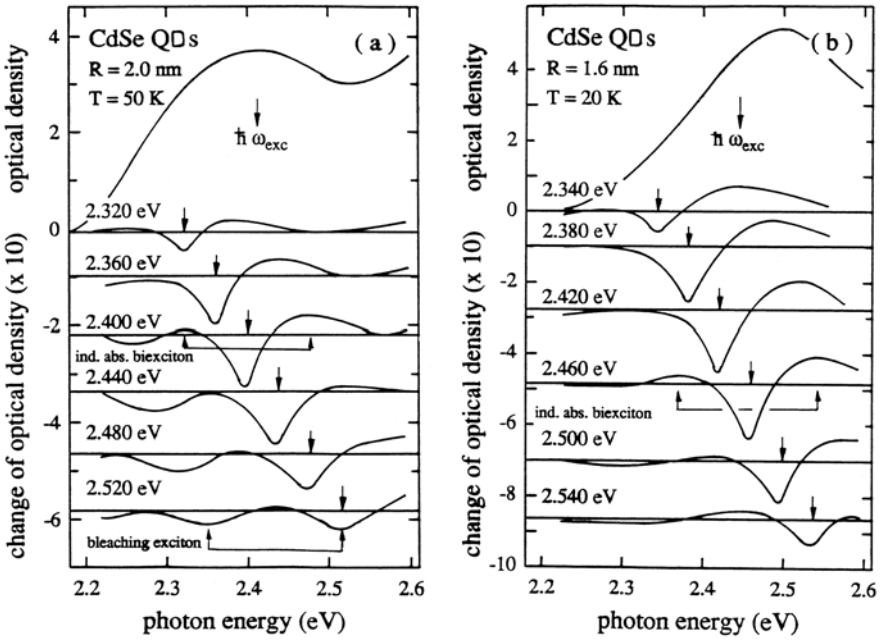


Fig. 20.14. The absorption spectra of two photo-darkened glass samples containing dots of CdSe and the differential transmission spectra for various pump-photon energies. According to [93G1]

Phonon-assisted biexciton generation or recombination is a topic treated both in III–VI [00F1, 00F2, 01Z1] and II–VI islands [99G1].

Multiexciton complexes in quantum islands also form a recent topic of research [98B1, 00D2, 00D3, 00F2, 01W2, 01W3]. These complexes have a certain similarity to the multiexciton complexes bound to point defects like P in Si (see Sect. 14.1).

For singly or even higher charged excitons see, e.g., [00H1, 01B1, 01H1, 01W2, 01Z1, 02S1, 02S2] and for internal transitions in biexcitons [01P2].

In II–VI islands the binding energy of excitons is partly larger than that of excitons in the parent bulk material. Values above 20 meV have been reported, e.g., in [99W2, 00W1, 01H2].

Optical gain resulting from the recombination of excitons, biexcitons or of even higher complexes has been reported, e.g., in [92D1, 95W1, 95W2, 96G1, 96W1, 02M1, 02M2, 03H1]

To conclude, we state that biexcitons, trions and mutliexciton complexes are well established phenomena in semiconductors.

Correlated photons have been observed in the single InAs dot emission from the optical transitions biexciton → exciton and subsequently exciton → ground state [01M2, 02S1, 02S2, 03U1]. We shall briefly return to this topic in Chap. 23.

20.4 Optical or ac Stark Effect

A further example of optical nonlinearities in the intermediate density regime is the optical, ac or dynamical Stark effect. The name implies that we are dealing with a shift of the exciton resonance caused by the electric field of a light beam in the sample. For recent reviews see, e.g., [88O1, 89C1, 90Z1, 93Z1, 01B4] and references therein.

This phenomenon is well known in atomic spectroscopy [92C1] and is frequently described in the "dressed atom" picture, i.e., as an atom in the presence of photons. A simple explanation, outlined in the following, is in terms of level repulsion, a phenomenon which we encountered already several times and which generally occurs in quantum mechanics for two energetically close lying levels which interact with each other.

We assume that the energy of the incident photons is chosen slightly below the exciton resonance and use the weak coupling picture. Then a state $|m, n = 0\rangle$ containing m photons and $n = 0$ excitons is almost degenerate with a state $|m - 1, n = 1\rangle$. Since the two states are coupled via dipole interaction, an energetic repulsion of the two levels occurs.

The frequency of the photons ω is held constant from the outside by the incident laser. Consequently the exciton resonance shifts in a model which includes only the 1 s exciton resonance to slightly higher energies for $\hbar\omega - E_{\mathrm{exc}} < 0$ and possibly to lower energies for $\hbar\omega - E_{\mathrm{exc}} > 0$ by an amount δE. This quantity has been calculated by various authors [88O1, 89C1, 89S1, 90Z1, 93Z1, 01B4]. We cite here the result from [89S1]:

$$\delta E = \frac{2\left|er_{\mathrm{cv}}E_{\mathrm{p}}\right|^2 \left|\phi_{1\mathrm{s}}\left(\boldsymbol{r}_{\mathrm{e}} - \boldsymbol{r}_{\mathrm{h}} = 0\right)\right|^2}{\left(E_{1\mathrm{s}} - \hbar\omega_{\mathrm{p}}\right) N_{\mathrm{S}}^{\mathrm{PSF}}}\,, \tag{20.9}$$

where er_{cv} is the band-to-band transition matrix element or polarizability, E_{p} the field amplitude of the incident laser at frequency ω_{p}, $\left|\phi_{1\mathrm{s}}\left(\boldsymbol{r}_{\mathrm{e}} - \boldsymbol{r}_{\mathrm{h}} = 0\right)\right|^2$ describes the (Sommerfeld) enhancement of the transition probability of the 1s exciton, $E_{1\mathrm{s}} - \hbar\omega_{\mathrm{p}}$ the detuning, and $N_{\mathrm{S}}^{\mathrm{PSF}}$ the density of electron–hole pairs necessary to block the exciton resonance by phase-space filling (Chap. 21).

The term $e\boldsymbol{r}_{\mathrm{cv}}\boldsymbol{E}_p\phi_{1\mathrm{s}}(\boldsymbol{r}_{\mathrm{e}} - \boldsymbol{r}_{\mathrm{h}})$ gives the so-called Rabi frequency (see also Chap. 27) for the 1s exciton. This is basically a field-dependent beat frequency with which the excitation oscillates coherently back and forth between the electronic excitation and the photon state.

The Rabi frequency also gives the splitting if $\hbar\omega$ and the exciton energy coincide.

Actually there are more terms in the shift that include, apart from the renormalization of the 1s exciton resonance, the continuum transitions, phase space filling, etc. with the consequence that the exciton resonance might shift to blue even if $\hbar\omega$ is chosen above the exciton [96N2,00M2]. In the experiments one is generally restricted to the situation $\hbar\omega < E_{1\mathrm{s}}$ to avoid real excitation, which would mask the optical Stark effect. Exceptions are the experiments

on ZnSe [96N2], which allowed also to tune the laser in the transparent window above the exciton energy and gave clear evidence for dominating level repulsion according to (20.9) in this situation.

We show in Fig. 20.15 two different experiments and two calculations, namely in Fig. 20.15a subpicosecond pump-and-probe absorption spectra for zero delay for a MQW sample with 10 nm GaAs and $Al_{1-y}Ga_yAs$ barriers from [86M1]. The pump beam was tuned significantly below the *hh* exciton resonance. A clear blue shift of the *hh* exciton and a less pronounced one of the *lh* exciton are observed. The rather strong reduction of the oscillator strength might be due to some real excitation, e.g., via two photon absorption. Figure 20.15c and d show some calculations from [96H1] using Boltzmann and quantum kinetics for a single resonance. The fact that the latter shows better agreement with experiments underlines the importance of this approach, to which we come back in Sect. 23.2.

Fig. 20.15b finally shows, that very intense (generally several $10\,\mathrm{MW/cm^2}$) subpicosecond lasers are not necessary to observe the optical Stark effect. Nanosecond pulses properly chosen concerning intensity and detuning work also. The fact that the differential transmission signal is almost symmetric and coincides very well with the derivative of a Lorentzian fit to the absorption spectrum indicate, that one observes a blue-shift without a change in oscillator strength under quasi-stationary conditions.

The quantity

$$\sigma = \frac{\delta E(E_{1\mathrm{hh}} - \hbar\omega_\mathrm{p})}{I_\mathrm{P}}\,, \tag{20.10}$$

deduced from Fig. 20.15b is $\sigma_\mathrm{exp} = 8.5 \times 10^{-8}\,(\mathrm{meV})^2\,\mathrm{cm}^2\,\mathrm{W}^{-1}$ and compares favorably with the theoretical value and ps experiments giving $\sigma_\mathrm{theory} \approx 5.8 \times 10^{-8}\,(\mathrm{meV})^2\,\mathrm{cm}^2\,\mathrm{W}^{-1}$ [90K1].

A better description would again be in terms of polariton-polariton interaction [92K1]. This picture describes not only the shift of the exciton resonance by an amount δE but also the appearance of optical gain positioned symmetrically to the low energy side of the pump laser, i.e., at

$$\hbar\omega_\mathrm{g} = \hbar\omega_\mathrm{P} - (E_\mathrm{ex} + \delta - \hbar\omega_\mathrm{p}) \tag{20.11}$$

Equation (20.11) follows from the simple picture that a state with two polaritons at $\hbar\omega_\mathrm{p}$ decays into a Stark-shifted exciton-like polariton and a photon-like polariton at $\hbar\omega_\mathrm{g}$ [88S1].

The optical Stark effect has also been observed in bulk materials like Cu_2O [85F1] and GaN [02C1] and in quantum wires [03F1].

To conclude, we may state that the optical or ac Stark-effect is also nowadays well understood for excitons in semiconductors. Apart from the specific reviews [88O1, 89S1, 90Z1] mentioned above more details may also be found in [96H1, 96S1, 01L1].

If we pump the samples even harder we leave the intermediate density regime and enter the high density regime where the optical properties are

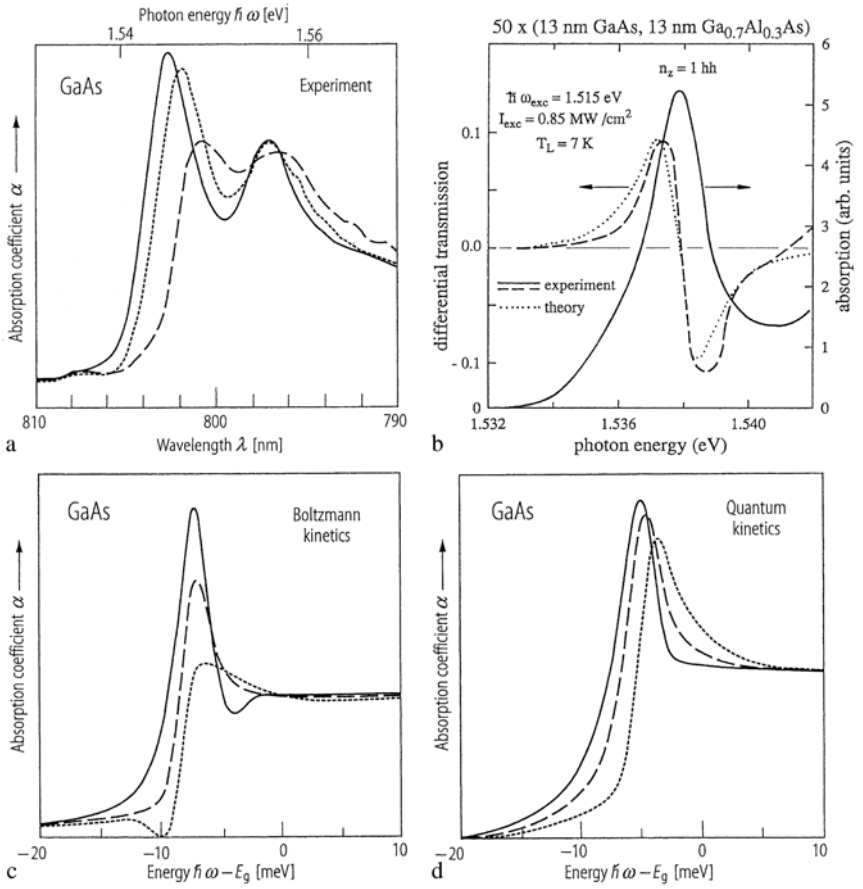


Fig. 20.15. Pump-and-probe beam spectra of a 10 nm GaAs/10 nm $Al_{1-y}Ga_yAs$ MQW sample showing the optical Stark effect [86M1] (**a**) an absorption spectrum without excitation and the differential transmission using 10ns pulses showing the same effect according to [90K1] (**b**). Calculations using Boltzmann (**c**) and quantum kinetics (**d**) [96H1]

determined by the electron–hole plasma. This concept was introduced in Sect. 19.2. However, before treating this aspect we discuss briefly the Bose–Einstein condensation of excitons and biexcitons and introduce another group of optical nonlinearities, the so-called photo-thermal nonlinearities which also generally belong to the intermediate density regime.

20.5 Excitonic Bose–Einstein Condensation

Excitons are Bosons to a good approximation, but are made up of two fermions i.e. they are composite Bosons. Actually, the creation and annihilation oper-

ators of excitons obey the commutation rules for Bosons, however, with an additional correction term of the order $0\left(n_{\mathrm{p}}a_{\mathrm{B}}^3\right)$ where a_{B} is the excitonic Bohr radius and n_{p} is the density of electron hole pairs.

This fact already triggered predictions of excitonic Bose Einstein condensation (BEC) and excitonic superfluidity (ESF) forty years ago [62B1,62M1,68K1]. For older and more recent reviews of this topic see [77H1, 95B2, 00M1, 03P2].

On the other hand, the fermionic constituents result in a transition to an electron–hole plasma at the highest intensities (see Fig. 19.3 and Chap. 21). While the existence of the EHP is well established, excitonic BEC was, at the time of writing this manuscript, still a controversially discussed topic [04I1], in contrast to the situation of (alkali-)atoms in a trap (see, e.g., [95B2, 02K2] and the references given therein). Alkali atoms are also composite Bosons, but the ratio between the ionisation energy and the thermal energy $k_{\mathrm{B}}T_{\mathrm{C}}$ (see (20.12)) is by many orders of magnitude larger than for excitons.

In the following we concentrate on attempts and possibilities to observe excitonic BEC and ESF.

We start with some general considerations in Sect. 20.5.1, and then treat the attempts to observe it in bulk (Sect. 20.5.2) and quasi-two-dimensional semiconductors (Sect. 20.5.3). The section will be completed with short comments about excitonic insulators and so-called "driven BEC".

20.5.1 Basic Properties

BEC is a macroscopic population of one quantum mechanical state by Bose particles (ideally, non- or weakly interacting) in thermal (quasi-)equilibrium [24B1, 24E1]. It occurs if either the temperature T is lowered below a critical temperature T_c at constant particle density n, or if n is raised above n_c at constant T.

For non-interacting ideal Bosons, one finds the following relation between n_c and T_c [00M1]:

$$n_c = 2.612g\left(\frac{mk_BT_c}{2\pi\hbar^2}\right)^{3/2} \tag{20.12}$$

where m and g are the mass of the particles and the degeneracy of the state, respectively. The condensate can show superfluidity.

The fact that the mass of excitons is comparable to the free electron mass, while those of, e.g., alkali atoms are ten to 100 times the proton mass and that excitons can easily be created by pulsed lasers in a density range up to $10^{17}\,\mathrm{cm}^{-3}$, allows one to expect values of T_C up to around 10 K, while successful experiments to observe atomic BEC in traps required T as low as a few 10 µK.

The weak interaction should be slightly repulsive, to avoid condensation of the particles in real space, since a BEC is a condensation in $\boldsymbol{k}$-space.

The only example of a BEC of weakly interacting Bosons is the condensation of alkali and H atoms in a trap treated, e.g., in the reviews [95B2, 02K2].

There are further examples of BEC for Bosons with rather strong interaction like the superfluidity of ^{4}He and systems that involve the pairing of Fermions and the formation of a gap, like superconductors or superfluid ^{3}He.

Excitons would evidently be another example for weakly interacting Bosons. Before we start to treat experimental attempts to observe excitonic BEC and the objections against these interpretations, first we illustrate how (20.12) can be understood.

We showed in Fig. 2.5 three occupation probabilities for particles with an effective mass m, namely the Boltzmann statistics for distinguishable particles and the Fermi–Dirac and Bose–Einstein statistics for indistinguishable particles with half or integer spin, respectively. They are given by

$$f_B\,(T, E, \mu) = \frac{1}{\mathrm{e}^{(E-\mu)/k_B T}}, \tag{20.13a}$$

$$f_{FD}\,(T, E, \mu) = \frac{1}{\mathrm{e}^{(E-\mu)/k_B T} + 1} \tag{20.13b}$$

$$f_{BE}\,(T, E, \mu) = \frac{1}{\mathrm{e}^{(E-\mu)/k_B T} - 1} \tag{20.13c}$$

where μ is the chemical potential or for Fermions the Fermi energy. Evidently, they coincide for $(E - \mu)/k_B T > 2$ and show characteristic differences below.

For particles with a well defined density n, μ is defined by (20.14)

$$n = \int_{E_0}^{\infty} D\,(E) \cdot f_{B,FD,BE}\,(T, E, \mu)\,\mathrm{d}E. \tag{20.14}$$

For particles like photons in black body radiation or phonons in thermal equilibrium, the density of which is not conserved, one has $\mu \equiv 0$. In Fig. 8.20 we showed the square root density of states of particles with a finite and constant (effective) mass m in three dimensions starting at energy E_0

$$D\,(E) \approx \sqrt{E - E_0}\,. \tag{20.15}$$

If one now performs the integration (20.14) for a low density, then μ, i.e., the origin of Fig. 2.5, is situated below E_0. If the density increases, μ shifts towards E_0. Eventually μ coincides with E_0. There is no singularity for Boltzmann particles and Fermions. The density one reaches under this condition for Fermions is given by

$$n_{\mathrm{eff}} = g\left(\frac{m k_B T}{2\pi\hbar^2}\right) \tag{20.16}$$

and is known as effective density of states (see Sect. 8.9). For further increasing n, μ shifts into the band. For Fermions, the distribution is then said to be degenerate since the Pauli principle reduces the occupation probability from 1 to 1/2 at μ and limits it to 1.

For Bosons the situation is different. The singularity at $E - \mu = 0$ does not allow μ to shift beyond E_0. For $\mu = E_0$ the integrated population is finite and corresponds to n_c in (20.12).

For further increasing $n > n_c$ a macroscopic and coherent population of the lowest state develops and accommodates all particles beyond n_c. This is the Bose–Einstein condensation. The condensed phase may show superfluidity, i.e., loss- or frictionless motion.

Apart from a prefactor of the order of unity, the densities for the onset of a degenerate population for Fermions and for the onset of BEC for Bosons are equal.

20.5.2 Attempts to find BEC in Bulk Semiconductors

We now present various attempts to observe excitonic BEC in bulk semiconductors and the objections against these interpretations brought forward. Since the topic of excitonic BEC and/or superfluidity appears almost regularly every few years in literature since its prediction in 1962 we cannot treat all examples. We briefly mention some older ones and then concentrate on recent experiments for Cu_2O.

The macroscopic population of one state should or could show up in a narrow luminescence peak. Therefore, several attempts have concentrated on spectrally narrow emission lines. Unfortunately, not only a condensed exciton phase gives rise to narrow emission lines but other processes, too. The narrow emission bands from bound exciton complexes (which are, by the way, also abbreviated BEC) have been misinterpreted as excitonic BEC, e.g., in AgBr, CdS or CuCl [74C1, 74J1, 75A1, 75N1, 76W1, 77H1, 77W1].

The recombination of biexcitons into a photon and an exciton gives rise to well-known and well-understood emission bands in different semiconductors like the Cu-halides, II–VI compounds or Si as described above in Sect. 20.3. Under resonant two-photon excitation of the biexciton narrow emission bands appeared, which also have been interpreted as a BEC of biexcitons [76N1]. It was shown, however, by two independent groups, that these narrow lines result from a cold but uncondensed gas of biexcitons and/or from resonant two-photon or Hyper-Raman scattering [76L1, 78D1, 78H1, 78O1, 79K1, 85H1]. In the meantime cold biexcitons have been also identified in β-ZnP_2 [01D1]. However, it was possible to verify the Bosonic character of biexcitons in CuCl in the sense that they are preferentially scattered into a state, that is strongly populated by an external laser pump source [82P1, 83P1], but again no spontaneous BEC could be reached.

For a short while the disappearance of excitonic features from the reflection and transmission spectra of CdSe under high excitation was considered as an indication for an excitonic BEC [70A1], but as we will learn in Chap. 21, this is actually an indication for a transition to an EHP.

In Ge the formation of an EHL at low temperatures could be suppressed by the application of stress and magnetic field. Nevertheless, it was also not

possible in this case to reach an excitonic BEC with increasing pump power, although a Bosonic line narrowing was observed at intermediate densities [83T1].

After this short overview of older work we now concentrate on Cu_2O. This is a direct gap semiconductor with a large exciton binding energy of 150 meV, but with parity forbidden band-to-band transition. See Sect. 13.2.1.2 and references therein. The lowest exciton state ($n_B = 1$) is split by exchange interaction by 12 meV in a lower para–exciton, which is optically forbidden to all orders. Only its $\boldsymbol{\Gamma}_5^-$-LO phonon satellite is seen very weakly in the luminescence spectra. The ortho-exciton is only quadrupole allowed, nevertheless resulting in a beautiful polariton dispersion leading continuously from the ortho-exciton branch over the bottleneck and the photon-like branch to $E \to 0$ for $\boldsymbol{k} \to 0$ [81U1, 91F1]. It is weakly seen both in absorption and emission and strongly in two-photon absorption (see Fig. 13.30). Furthermore its various LO phonon replica show up strongly in luminescence and absorption. The LO-phonon assisted absorption shows nicely the square root dependence of the density of states. This absorption band is superimposed by the $n_B P$ exciton states, which are weakly dipole allowed due to the odd parity of the P-envelope function. In good samples the $n_B P$ series can be observed up to $n_B = 9$ and beyond.

The LO-assisted luminescence bands, which reflects the distribution of the excitons in their bands, prove that the excitons reach thermal equilibrium with the lattice down to temperatures well below 10 K, at least under low excitation. Examples for all of these statements can be found in Sect. 13.2.1.2 and the references given therein.

Since the relaxation of ortho-excitons to para-excitons is rather slow at low temperatures [01J1] there was a hope that ortho-excitons might be pumped to beyond the critical density n_c of (20.12) at low crystal temperatures. A certain change of the emission line shape from a simple Boltzmann-type to a narrower Bose-type has been found [83H1, 83M1, 83M2, 90S1, 00M1]. However, various authors agree, that the density approaches the $n_c(T_c)$ curve from low densities, but never reaches it [87F1, 90S2, 96K1, 96N1, 00M1, 01K1]. For calculations of the relaxation dynamics see [98E1, 00T2].

The next approach came from transport measurements of presumably para-excitons [95B1, 96M1]. The experimental setup and the main findings are the following. A brick-shaped Cu_2O sample is exited on one side by intense ns pulses in the LO phonon continuum, producing a cloud of excitons. Their arrival at the opposite side is monitored by a current pulse produced by their field ionization in a Cu_2O/Cu Schottky barrier contact. The resulting current pulse is monitored on an oscilloscope. The signal starts after a delay time given by the LA velocity of sound v_{LA}. The following signal is temporally rather broad, indicating a diffusive transport of the excitons over the sample length of about 3.5 mm. However, the signal becomes steeper and gets shorter if the density is increased at low temperature or if the temperature is lowered at high excitation.

The interpretation given in [95B1, 96M1] is that the excitons undergo a BEC into a superfluid state, when the critical values of n_c and T_c are met, and the condensed cloud propagates with v_{LA} through the sample to the Schottky barrier.

A first small problem in this interpretation is that the signal area saturates with increasing intensity while one would rather expect the opposite behavior for a BEC.

As is usual in the field of excitonic BEC, various alternative explanations have been put forward and BEC and superfluidity have seriously been questioned in various ways.

Due to the weak absorption in the 1s ortho-exciton state, excitons have to be excited in the LO-phonon or ionization continuum, if one wants to reach high densities. Consequently, a large non thermal population of optical and then of acoustic phonons is created by the relaxation of the excitons to the (para-) exciton ground state, and later on by their dominantly non-radiative recombination [01J1]. The phonon cloud created near the surface propagates into the sample and it has been calculated by [96K2, 97T1, 02J1] that an exciton flux caused by the phonon wind and very similar to the experimentally observed signal can be expected at the detector, i.e., a (quasi-) ballistic motion of the excitons driven by the expanding nonthermal cloud of LA phonons. This interpretation does not involve any superfluidity or BEC of the excitons and explains easily why the signal onset delay at the detector coincides with the time of flight of LA phonons.

Spatially and temporally resolved pump and probe beam spectroscopy of a thin Cu_2O rod has been performed in [02J2] using the broadening and bleaching of the higher n_BP excitons states caused by a high density of excitons [98J1, 00J1] as a measure of the density of excitons along the crystal as a function of distance from the excited surface and of time.

The results show exciton transport over distances of around 1 mm, but the whole propagation process could be simulated within experimental error by classical diffusive propagation for a temperature range from 2 to 30 K and the average propagation velocity of the exciton cloud was well below v_{LA}.

Another criticism came from a detailed analysis of the intensity and dynamics of the ortho-exciton luminescence in [99O1, 99O2] and led the authors to the conclusion, that the Bose–Einstein luminescence line shape is due to inhomogeneities of the exciton population and more importantly, that the Auger-recombination of the excitons and the heating of the exciton gas are so dramatic, that one cannot even come close to the conditions of an excitonic BEC.

This huge Auger cross-section has in turn been questioned since the Bohr radius is small and the excitons do not even carry an electric dipole moment. Instead, an efficient ortho $\rightarrow$ para conversion mechanism has been put forward in theoretical investigations [00K2] for high densities. Experimentally, it has been found independently in [02D1, 02J2] that the Auger cross-section is indeed small, however this topic has been introduced again in [04W1].

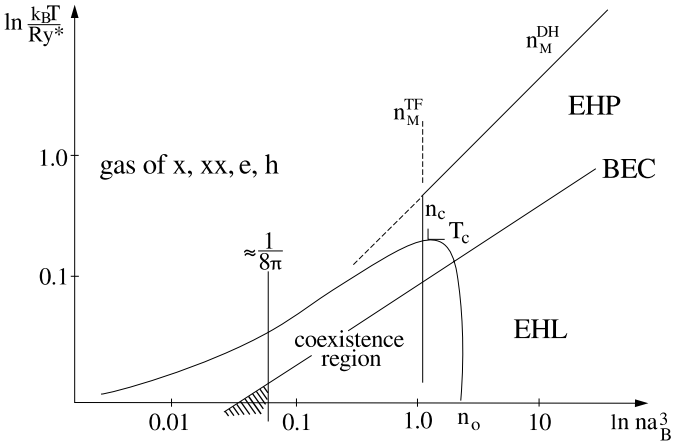


Fig. 20.16. pedestrian approach to a phase diagram [04K2]

Even more recently it has been shown [02K1] that the Schottky barrier exciton detector used in [95B1, 96M1] to monitor the arrival of the excitons might be driven under the highest excitation conditions by several orders of magnitude into saturation, explaining both qualitatively the variation of the time dependence of the electrical signal with increasing excitation and the saturation behaviour of its time integral. The discussion on this aspect was still going on during the writing this manuscript [03K3, 03M1, 04K2]. See also [80A1] in this context.

First attempts failed to confine excitons at sufficiently high densities in Cu_2O in a potential trap caused by externally applied inhomogeneous stress [86T1], but new experiments in this direction using two-photon pumping partly connected with stress-induced wells started recently [01S2, 02N1, 02N2, 03N1]. They present possibly the most promissing approch towards the experimental verification of excitonic BEC in analogy to [02K2], but have not yet resulted in conclusive evidence for excitonic BEC.

The same statement is true for excitonic inter-subband spectroscopy (i.e., the investigation of the excitonic Lyman series) for which theory predicts features characteristic of excitonic BEC [01J2, 03J1, 04J1]. See also the discussion with Fig. 13.36b. The influence of the recently discovered anisotropy of the exciton masses in Cu_2O [03D1] on the issue of excitonic BEC has only to be explored.

Though it is not possible to cite all experimental work published during the last forty years, the selection shown here makes clear, that there is currently no clear-cut and generally accepted proof for excitonic BEC. On the other hand, there is no generally accepted theoretical result explaining why it should not occur [74J1]. So the search for proof will and does continue, including structures of reduced dimensionality. Before we go to such structures in Sect. 20.5.3 we consider a simple phase diagram in Fig. 20.16.

A transition to an EHP (see for more details Chap. 21) occurs beyond a certain density n_M or if the screening length l_c becomes comparable to the excitonic Bohr radius a_B (see Chap. 22).

In the Debye–Hückel approximation, l_c is given by

$$l_\mathrm{DH} = \left(\frac{\varepsilon\varepsilon_0 k_\mathrm{B} T}{e^2 n_\mathrm{P}} \right)^{1/2} \tag{20.17a}$$

resulting in

$$n_\mathrm{M}^\mathrm{DH} = (1.19)^2 \, \frac{\varepsilon\varepsilon_0 k_\mathrm{B} T}{e^2 a_\mathrm{B}^2} = (1.19)^2 \, \frac{k_\mathrm{B} T}{2 a_\mathrm{B}^3 \mathrm{Ry}^*} \,. \tag{20.17b}$$

This approach is only valid as long as the exciton and carrier gases can be described by classical statistics. For very low temperatures it would give the unphysical result that a vanishing density already results in a screening of the exciton. We therefore give n_M^DH in Fig. 20.16 for higher temperatures only. At low temperatures, where the carrier gas is degenerate, the Thomas–Fermi screening length is more appropriate. It gives a temperature-independent value of n_M^TF given roughly by

$$n_\mathrm{M}^\mathrm{TF} a_\mathrm{B}^3 \approx 1 \,, \tag{20.17c}$$

which is also given in Fig. 20.16. This means than an EHP exists to the right and below these two curves.

Now we give the criterion for a BEC for ideal Bosons according to (20.12). We can anticipate, that excitons will no longer behave as ideal Bosons close to the Mott transition to a Fermi gas.

In [71F1] it is claimed that excitons behave as weakly interacting Bosons for

$$n_\mathrm{P} a_\mathrm{B}^3 \ll \frac{1}{8\pi} \approx 0.04 \,. \tag{20.17d}$$

The limit 0.04 is shown as a vertical line in Fig. 20.16, too.

The shaded region where an excitonic BEC can be expected already becomes rather small.

Now we add the phase diagram for EHL formation from Chap. 21 for a finite binding $\Phi > 0$ of the EHL and see that the possibility to observe an excitonic BEC disappears completely. This is also stated in ([00M1], p 12) but is then set aside rather quickly. The author feels, that some more consideration should be given to this argument.

In Si, Ge and diamond the EHL has been clearly observed (see Sect. 21.3.2). Even if Φ was brought close to zero by external fields, no excitonic BEC occurred.

In all III–V and II–VI compounds the plasma forms a bound state. The fact that no phase separation was observed is not due to the fact that the plasma is not bound, but that the carrier lifetime is too short for the phase separation to form. Possibly it is also too short for the BEC condensation to take place, which would explain the failure to observe it in this rather large group of materials.

In CuCl the creation of an EHP has been verified experimentally [85H2] and recently even the formation of an EHL [02N3], however our knowledge of its properties is too limited to give a definite statement about Φ. In Cu_2O no report of an EHP is known to the author. Under conditions where it is possible to bleach all n_BP exciton states [00J1], the 1s excitons are possibly far away from being idealized, weakly interacting Bosons.

To conclude this section, it should be mentioned, that more complex phase diagrams are discussed in theory, which allow for an excitonic BEC pocket and are reviewed, e.g., in [00M1]. However there does not yet seem to be much of an experimental verification of these diagrams.

20.5.3 Structures of Reduced Dimensionality

Before discussing recent results on coupled quantum wells, we want to have a short look into two-dimensional systems in general.

Shifting the step-like density of states function of the two-dimensional system of massive particles in Fig. 8.20 relative to the Bose–Einstein distribution function in Fig. 2.5 shows immediately that, strictly speaking, a BEC is not possible in two dimensions. Either μ is below E_0, and there is no BEC or μ coincides with E_0 and a divergence of the particle density arises from the product of a finite density of states and the divergence in their occupation probability.

Actually, the DOS is not step-like in the sense of a mathematical heavy side function $\Theta(E-E_0)$. One has a (often exponential) tail of localized states below E_0. These tail states behave, however, unlike Bosons. They can be empty and they can be occupied by one exciton. If two excitons (i.e., a biexciton) are placed in this state the energy of both shifts to the red. Higher occupancies lead, if possible, to a blue shift (see also Sect. 20.3.2).

On the other hand it would also be difficult to imagine what a BEC of localized excitons means when every particle sits in another place and at a different energy.

Though there is strictly no BEC in two dimensions for massive particles, a transition to a superfluid state is possible according to Kosterlitz and Touless [00M1] (KTS) for densities n_c above or temperatures below T_c given by

$$T_c \approx \frac{\left(\hbar^2/2m\right) 4\pi n_c}{\ln\ln\left(1/n_c a^2\right)} \tag{20.18a}$$

resulting, to a good approximation, in

$$n_c \approx \frac{0.32 g m k_\mathrm{B} T_c}{\hbar^2}\,. \tag{20.18b}$$

In the following we describe to which extent these ideas could be realized in a system not completely different from the one proposed in [75L1].

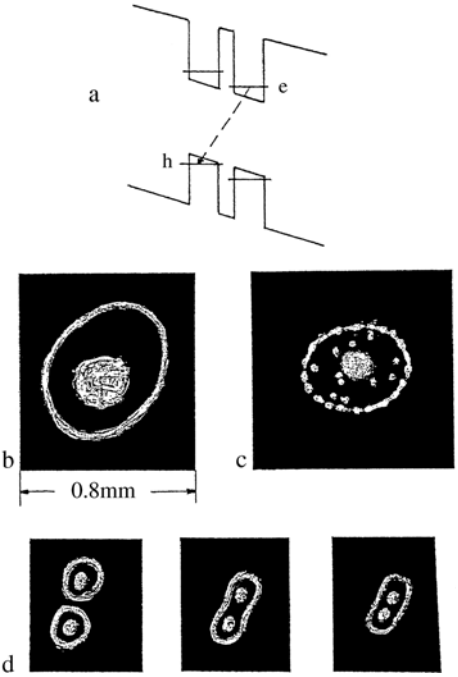


Fig. 20.17. Schematic sketch of the band alignment of the samples used by [98B3, 01B3, 02B2, 02B3, 02S3, 02S4] (**a**) schematic drawings of the luminescence ring structure observed by [02S3, 02S4] (**b**) and by [98B3, 01B3, 02B2, 02B3] (**c**) and the behavior observed by both groups when two excitation spots are used with decreasing distance (**d**)

There have been recent attempts by two groups (Butov et al. and Snoke et al.) to observe excitonic superfluidity in the frame of KTS using two coupled and quantum wells, tilted by an external electric field [94B1, 98B3, 01B3, 02B2, 02B3, 02S3, 02S4]. The appearance of narrow emission bands has also been described in [02L2]. The basic idea of the samples is the same in both cases and is shown in Fig. 20.17a. In a n^+in$^+$ structure, two coupled wells with a narrow barrier are incorporated in the intrinsic layer. A voltage applied to the n^+ cladding layers tilts the band structure and spatially separates the lowest electron state from the highest hole state as shown in Fig. 20.17a. Butov et al. use n^+GaAs and $Al_{1-y}Ga_yAs$ barriers around and between the two GaAs wells in the intrinsic region while Snoke et al. use $In_{1-y}Ga_yAs$ wells.

In both cases excitons are formed, which carry a permanent dipole moment due to the spatial separation of electrons and holes under the action of an external field. This separation increases the exciton lifetime to values in the 100 ns regime, but also reduces the exciton binding energy [03S1]. The parallel-oriented excitonic dipole moments create a repulsive interaction that prevents biexciton formation.

At low temperatures, and under cw excitation, both authors observe rather similar phenomena. With increasing excitation in the barrier, luminescence is not only observed from the excitation spot but also from a bright ring with no detectable emission between Fig. 20.17b,c except for some localization centers in the case of [98B3,01B3,02B2,02B3]. The diameter of the ring increases with increasing pump power reaching diameters in the mm regime. If the excitation spot is moved on the sample, the ring structure follows the excitation spot.

Abrupt temporal variations of the excitation intensity result in variations of the luminescence in the ring structure indicating radial velocities on the order of 1×10^6 to 5×10^6 cm/s which is considerably larger than the LA velocity of sound in GaAs [02S3, 02S4, 03S2].

Snoke at al. investigate the ring at temperatures of a few K but can follow it up to $T \approx 90$ K. Butov et al. works preferentially at lower temperatures and observes a fragmentation of the ring into bright spots, which are rather equally spaced along the ring with some bright spots between the directly excited area and the ring resulting from localization sites in the MQW structure [98B3, 01B3, 02B2, 02B3, 02S3, 02S4, 03S2] (see Fig. 21.7c).

If two widely separated spots are excited, they are surrounded by individual rings. If the excitation spots are brought closer together, both groups find that the rings are deformed and merge into a single structure as shown in Fig. 21.17d.

Though the experimental findings are rather similar, the interpretations of both groups are very different.

Butov et al. claim that the mutual repulsion of the excitons (which is even enhanced by the depolarization field) results in a rapid, presumably ballistic expansion of the excitons with $\boldsymbol{k}_{\parallel} > \boldsymbol{k}_{\text{light}}$ so that they cannot radiate because $\boldsymbol{k}_{\parallel}$ is conserved at a plane interface. The velocities given by Snoke are consistent with such arguments. After that expansion, the excitons undergo a BEC transition resulting in the ring and the bright dots along the ring are considered superfluid vortices.

In contrast Snoke claims that the excitons are in the KTS superfluid state in the dark range between the excitation spot and the ring. The bright ring results from the transition of the excitons into their normal state due to dilution and subsequent recombination. While being in the superfluid state, emission is forbidden for the excitons for some reason, e.g., because of their large value of $\boldsymbol{k}_{\parallel}$. The argument of Snoke against the interpretation of Butov is that a purely ballistic propagation of excitons in these structures over distances up to one mm is extremely unlikely. If, on the other hand, scattering is allowed, then excitons will also be created at smaller values of $\boldsymbol{k}_{\parallel}$ and a measurable luminescence should be observed between the exciton spot and the ring.

The merging of the rings is more difficult to explain in the model of Butov, because the density should be higher in the overlapping region, while the observation of the equally spaced bright spots is difficult to explain in Snoke's model.

At a recent workshop as part of the seventh "Nonlinear Optics and Excitation Kinetics in Semiconductors" meeting (NOEKS 7) [03N2], Snoke came up with a completely different interpretation of the data [03S2]. He observed, that the ring only appears if the sample is excited in the barrier. If one excites the sample directly in the wells, no ring structure appears, even if the incident intensity is increased to compensate for the reduced absorption.

Snoke's present interpretation is, that the formation of the ring has nothing to do with a superfluid state at all. If excited in the well, carriers essentially recombine at the excitation spot. When excited in the barrier, the holes are captured in their well, but electrons partly escape from the wells to the n^+ layer or are not captured at all due to their smaller effective mass. This has the consequence that a two-dimensional "puddle" of holes forms in the well around the excitation spot. Apart from the directly excited area, hardly any electrons exist in this "puddle" and consequently no luminescence appears. The luminescence only occurs at its boarder when electrons reach it, coming, e.g., from the n^+ layers. This model easily explains the merging of the rings in Fig. 20.17d by a merging of the two hole "puddles". However, this model cannot explain the equally spaced bright dots on the ring. Butov still explains these bright dots superfluid vortices [03B1, 03B2]. Future investigation must clarify the origin of the dots. See also the proceedings [04I1], where the ring has been reported also by a third group [04R2] and where Snoke introduced the possibility of a strain induced trap also for quantum wells [99N1, 04S1].

As a short addendum it can be stated that similar rings have been observed by V. Lyssenko in GaAs single quantum wells and superlattices [03L1]. The appearance and disappearance of theses rings can be influenced by varying the focus of the excitation beam.

Luminescence ring structures have also been observed in thin CdS platelets under pulsed excitation and at low temperatures [78L1] but have been interpreted as light scattering under conditions of stimulated emission.

Evidently we are left in quasi two-dimensional systems with the same situation as in bulk material. There is presently no clear, generally accepted evidence for excitonic BEC nor for superfluidity.

There are possibly some unknown reasons why it does not occur, which is a challenge to the theoreticians, or, as L.V. Keldysh stated [03K1] "Possibly we had it all the time without noticing it, because its influences on the optical properties are completely unspectacular," which is a challenge to experimentalists.

Furthermore, Snoke stated as a joke during NOEKS 7 [03N2] that not every circular emission is connected with KTS or BEC, as the reader can confirm by entering "Hoag's Object" in a search engine on the Web.

20.5.4 Driven Excitonic Bose–Einstein Condensations

Since it is obviously difficult to find and to uniquely prove excitonic BEC or (KTS) superfluidity in semiconductors, some authors invented the term "driven BEC".

A general property of these suggestions, which are briefly reviewed, e.g., in [02L3, 02S3, 02S4] is, that no process is occurring in thermal (quasi-) equilibrium distribution and thus no BEC.

Instead, one is either creating a coherent population by an external laser into which or out of which particles are scattered, as already found for the biexciton case in CuCl mentioned above, or it involves lasing on the (photon-like) part of the exciton-polariton dispersion.

It is also a common feature that very similar phenomena have been found one or a few decades ago under different, less spectacular names and are now being reinvented under the name "driven BEC".

For example, all of the processes shown in Figs. 19.3 and 22.1 in the intermediate density regime can result in stimulated emission on the lower polariton branch as found theoretically and verified experimentally in bulk materials and in structures of reduced dimensionality including cavity polaritons [81K1, 85H1, 94H1, 01E1] but at that time nobody had the idea to call such a phenomenon a BEC.

As already mentioned, the scattering of biexcitons into a state which is populated coherently by an external laser source was considered as proof of the Bosonic character of biexcitons, but not as a BEC [82P1, 83P1, 84W1] (also see above).

Another phenomenon connected with biexcitons or more generally two-polariton states, which has been observed in several I–VII and II–VI compounds [81K1, 85H1], has been introduced as two-photon or hyper Raman scattering. See Sects. 13.1.4 and 20.3.1. This process can occur spontaneously, but it can also show gain and stimulated emission in the sense of an optical parametric amplifier. Actually, it is a process belonging to the group of $\chi^{(3)}$ or of higher order processes like four-wave mixing (FWM).

The decay can also be stimulated by sending an external laser beam into one of the outgoing channels resulting in an enhanced scattering into this and also into the other channel. This would then be a typical example for non-degenerate four-wave mixing (NDFWM) or electronic coherent (anti-) Stokes Raman scattering (CARS).

Apart form the fact that one has transfered the experiments from bulk samples to quasi two-dimensional systems and cavities they are extremely similar to the processes described in connection with Figs. 20.9 to 20.11.

20.5.5 Excitonic Insulators and Other Systems

To conclude this section on BEC we shortly mention another concept of excitonic BEC, namely the so-called excitonic insulators. There are two scenarios for the occurrence of this phenomenon, that will be briefly outlined below. More details and many references can be found in [00M1].

One scenario occurs in narrow gap semiconductors with E_g^n, or semimetals ($E_g = 0$). If the binding energy of excitons is still finite, the system may lower its energy at sufficiently low temperatures by the spon-

taneous formation of excitons provided that the following inequalities are fulfilled

$$\mathrm{Ry}^* > E_g^n \quad \text{and} \quad \mathrm{Ry}^* \gg k_B T\,. \tag{20.19}$$

Since a gas of excitons is insulating compared to the gas of free carriers, usually present in semimetals and narrow gap semiconductors, the transition to the excitonic insulator should show up in a characteristic variation of the resistivity.

One candidate is (or was) the semimetal grey tin but no experimental indication of spontaneous exciton formation nor of their condensation is known to the author.

In [91B2,03W1] experiments have been reported for $TmSe_{0.45}Te_{0.55}$, where the resistivity varies by roughly three orders of magnitude when tuning the gap by external pressure.

Other systems, where a transition to an excitonic insulator can occur, are wide-gap semiconductors in which a degenerate EHP has been created. If the effective masses of electrons and holes are equal (a situation generally not met in real semiconductors) a gap in the sense of a BCS-theory could open simultaneously in the degenerate conduction and valence band populations with decreasing temperature resulting in a vanishing of the total conductivity of the condensates.

The BEC of magnons in $TlCuCl_3$ has been reported in [00N1]. For quantum-Hall exciton condensates in bilayers see [02B6, 02B7, 03S3, 04K3]

20.5.6 Conclusion and Outlook

We shall see in the next chapter, that the existence and the properties of the electron–hole plasma in semiconductors are well established both experimentally and theoretically. Concerning excitonic BEC and superfluidity, the situation is still unclear and controversially discussed in the literature even 40 years after the first theoretical prediction and it is to some extent a question of taste or of the personal experiences of the scientist if he stresses the "pros" or "cons." Obviously, the author is presently more on the "contra" side. In any case, there is much more theoretical work on the topic than hard experimental facts or, as Littlewood stated [03L2], the "smoking gun" argument is still missing. Maybe it will come in the future, to finish with an optimistic statement.

20.6 Photo-thermal Optical Nonlinearities

The first interaction of visible light with matter is usually via the electronic system, i.e., an incoming photon is absorbed creating an electron–hole pair. This pair recombines after a while and, since most recombination processes in semiconductors are nonradiative, energy is transferred to the lattice, i.e.,

to the phonon system. In simpler words, the lattice is heated. An increase of the lattice temperature in turn results in changes of the absorption spectrum, e.g., via the Urbach rule of (13.12) and of the real part of the refractive index (Chap. 6).

We show an example in Fig. 20.18. The transmission spectrum (a) of a CdS sample is measured at RT with a weak cw probe beam from an incandescent lamp. We see Fabry–Perot modes of the platelet type sample in the transparent region and the onset of the excitonic absorption. When we illuminate the sample with the green (514.5 nm) line of a cw Ar^+ laser with a power of about $2\,kW/cm^2$ we find the spectrum (b). The sample temperature increases in the laser spot by roughly 50 K and the consequences are seen in the spectrum. The absorption edge becomes less steep and shifts to the red. The Fabry–Perot modes also show a slight red shift revealing an increase of the refractive index.

At the position of the Ar $^+$-laser line, which is indicated by an arrow, we find a strong excitation-induced increase of absorption.

The photo-thermal optical nonlinearities are thus much less complex than some of the electronic effects of preceding sections and they usually have rather long relaxation time constants, often in the ms range. If one wants to investigate the electronic properties of semiconductors, thermal effects are often unwanted and disturbing. But, as we shall see in Chap. 24, it is precisely because of these properties that they can be used as model systems for the applicability of (electro-) optical nonlinearities in optical bistability or in (electro-) optic data handling. We come back to these aspects in Chap. 23

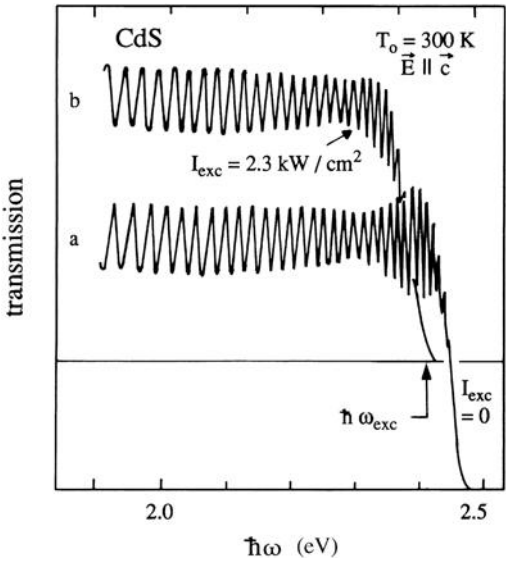


Fig. 20.18. Transmission spectra of a CdS platelet at a surrounding temperature of 300 K without (**a**) and with (**b**) excitation by the green line of an Ar^+ laser. According to [92K2]

20.7 Problems

1. If a crystal has inversion symmetry, then the parity of the eigenstates is well defined. Which states can you reach from the crystal ground state (even parity) in one-, two- and three-photon absorption processes? Compare your results later with the information given in Chap. 26.
2. Assume that the scattering cross section of excitons is determined simply by their Bohr radius. Try to estimate the density at which an exciton has a probability of 0.1 of hitting another exciton during its lifetime. Which other parameters of the exciton do you need to make this estimate? Compare with values of the Mott density in Chap. 21.
3. Calculate the spectra of the luminescence of the biexciton and of the induced absorption. From which feature can you deduce the binding energy of the biexciton?
4. What are the best conditions for observing the optical Stark effect? Check if they have been met in some of the experiments described in the references.
5. Calculate the increase of the lattice temperature of a semiconductor under pulsed excitation: $I_{\mathrm{exc}} = 10^6\,\mathrm{W/cm^2}$; $\hbar\omega_{\mathrm{exc}} = 2\,\mathrm{eV}$; pulse duration $= 10\,\mathrm{ns}$; repetion rate $= 10\,\mathrm{s^{-1}}$; $\alpha(\hbar\omega_{\mathrm{exc}}) = 10^5\,\mathrm{cm^{-1}}$; diffusion length of excitons $T_d = 1\,\mu\mathrm{m}$. Consider bath temperatures of 5 K and of 300 K and use Debye's approximation for the specific heat with $l_D = 300\,\mathrm{K}$. Can you make a modified guess if you know that the velocity of sound in solids is of the order of $3 \times 10^3\,\mathrm{m/s}$?
6. Calculate the critical density for the onset of BEC in a typical semiconductor (electron mass $0.1\,\mathrm{m_0}$, hole mass $0.6\,\mathrm{m_0}$ dielectric constant $\varepsilon = 10$) for $T = 4.2$ and 100 K. Campare with the densities for the transition to an electron–hole plasma given in Chap. 21.

References to Chap. 20

[24B1] S.N. Bose, Z. Physik **26**,178 (1924)
[24E1] A. Einstein, Sitzungsber. der königl. Preuss. Akademie der Wissenschaft (1924)
[58L1] M.A. Lampert, Phys. Rev. Lett. **1**, 450 (1958)
[58M1] S.A. Moskalenko, Optics and Spectroscopy **5**, 147 (1958)
[62B1] J.M. Blatt, K.W. Böer and W. Brandt, Phys. Rev. **126**, 1691 (1962)
[62M1] S.A. Moskalenko, Fiz. Tverd. Tela (Sov. Phys. Solid State) **4**, 276 (1962)
[63H1] J.J. Hopfield and D.G. Thomas, Phys. Rev. B **132**, 563 (1963)
[68K1] L.V. Keldysh and A.N. Kozlov, Sov. Phys. JETP **27**, 521 (1968)
[69B1] C. Benoit à la Guillaume, J.M. Debever and F. Salvan, Phys. Rev. **177**, 567 (1969)
[70A1] I.Kh. Akopyan and B.S. Razbirin, JETP Lett. **12**, 251 (1970)
[71F1] A.L. Fetter and J.D. Wlecka, Quantum Theory of Many Particle Systems, Mc Graw Hill, New York (1971)

[72A1] O. Akimoto and E. Hanamura, J. Phys. Soc. Jpn. **33**, 1537 (1972)
[72A2] O. Akimoto and E. Hanamura, Solid State Commun. **10**, 253 (1972)
[72B1] W.F. Brinkmann, T.M. Rice and B. Bell, Phys. Rev. B **8**, 1570 (1972)
[73H1] J.M. Hvam, Solid State Commun. **12**, 95 (1973)
[73H2] W.T. Huang, phys. stat. sol. (b) **60**, 309 (1973)
[74C1] W. Czaja and C.F. Schwerdtfeger, Solid State Communication **15**, 87 (1974)
[74H1] J.M. Hvam, phys. stat. sol. (b) **63**, 511 (1974)
[74J1] W.D. Johnyston, Jr. and K.L. Shaklee, Solid State Commun. **15**, 73 (1974)
[75A1] T. Anzai, T. Goto and M. Ueta, J. Phys. Soc. Japan **38**, 774 (1975)
[75G1] E.O. Göbel, K.L. Shaklee and R. Epworth, Solid State Commun. **17**, 1185 (1975)
[75K1] C. Klingshirn, phys. stat. sol. (b) **71**, 547 (1975)
[75L1] Yu.E. Lozovik and V.I. Yudson, JETP Lett. **22**, 274 (1975)
[75N1] N. Nagasawa et al., J. Phys. Soc. Japan **38**, 593 (1975)
[76B1] F. Bassani, and M. Rovere, Solid State Commun. **19**, 887 (1976)
[76H1] B. Hönerlage, C. Klingshirn and J.B. Grun, phys. stat. sol. (b) **78**, 599 (1976)
[76L1] R. Levy et al., phys. stat. sol. (b) **77**, 381 (1976)
[76N1] N. Nagasawa et al., Journ. of Luminesc. **12/13**, 587 (1976)
[76S1] B. Stébé and C. Comte, Solid State Commun. **19**, 1237 (1976)
[76W1] J. Weber, phys. stat. sol. b **78**, 699 (1976)
[77H1] E. Hanamura and H. Haug, Phys. Rep. **33 c**, 209 (1977)
[77W1] J. Weber and H. Stolz, Solid State Commun. **24**, 707 (1977)
[78D1] Vu Duy Phach et al., phys. stat. sol. b **86**, 159 (1978)
[78H1] B. Hönerlage, Vu. Duy Phach and J.B. Grun, phys. stat. sol. b **88**, 545 (1978)
[78I1] T. Itoh, T. Suzuki and M. Ueta, J. Phys. Soc. Jpn., **44**, 345 (1978)
[78K1] S.W. Koch et al., phys. stat. sol. **89**, 431 (1978)
[78L1] V.G. Lyssenko and C. Klingshirn, unpublished (1978)
[78M1] W. maier and C. Klingshirn, Sol. State Commun. **28**, 13 (1978)
[78O1] M. Ojima et al., J. Phys. Soc. Japan. **45**, 884 (1978)
[79G1] P.L. Gourley and J.P. Wolfe, Phys. Rev. B **20**, 3319 (1979)
[79H1] B. Hönerlage et al., J. Lumin. **18/19**, 683 (1979)
[79K1] T. Kushida, Solid State Commun. **32**, 209 (1979)
[79M1] A.E. Mayer and E.C. Lightowlers, J. Phys. **C12**, L507 (1979)
[79M2] V. May, K. Henneberger and F. Henneberger, phys. stat. sol. (b) **94**, 611 (1979)
[79S1] H. Schrey, V.G. Lyssenko and C. Klingshirn, Sol. State Commun. **31**, 299 (1979)
[80A1] V.F. Agekyan and A.A. Berezhnaya, Sov. Phys. Semicond. **14**, 1002 (1980) and references therein
[80K1] G. Kobbe and C. Klingshirn, Z. Physik **B37**, 9 (1980)
[81F1] D. Fröhlich, Festkörperprobleme / Adv. Solid State Phys. **XXI**, 363 (1981)
[81K1] C. Klingshirn and H. Haug, Phys. Rep. **70**, 315 (1981)
[81U1] Ch. Uihlein, D. Fröhlich and R. Kenklies, Phys. Rev. B **23**, 2731 (1981)
[82I1] T. Itoh and T. Katohno, J. Phys. Soc. Jpn., **51**, 707 (1982)
[82K1] G. Kurtze, V.G. Lyssenko and C. Klingshirn, phys. stat. sol. (b) **110**, K103 (1982)

[82K2] G. Kurtze, Ph.D. Thesis, Karlsruhe (1982)
[82K3] M. Kuwata and N. Nagasawa, J. Phys. Soc. Jpn., **51**, 2591 (1982)
[82L1] V.G. Lyssenko et al., Sol. State Commun. **42**, 401 (1982)
[82P1] N. Peyghambarian, L.L. Chase and A. Mysyrowicz, Optics Commun. **41**, 178 (1982)
[82T1] V.B. Timofeev, in Excitons, ed. By E.I. Rashba, M.D. Sturge, North-Holland, Amsterdam, Cahp. 9, p. 349 (1982)
[83H1] H. Haug and H.H. Kranz, Z. Phys. B **53**, 151 (1983)
[83H2] J.M. Hvam et al., phys. stat. sol. (b) **118**, 179 (1983)
[83K1] D.A. Kleinmann, Phys. Rev. B **28**, 871 (1983)
[83M1] A. Mysyrowicz et al., Phys. Rev. B **27**, 2562 (1983)
[83M2] A. Mysyrowicz, Phys. Rev. B **27**, 2325 (1983) and in ref. [81A1]a of chapt.1 p659
[83P1] N. Peyghambarian, L.L. Chase and A. Mysyrowicz, Phys. Rev. B **27**, 2325 (1983)
[83T1] V.B. Timofeev, V.D. Kulakovskii and I.V. Kukushkin, Physica B + C **117/118**, 327 (1983)
[84K1] M. Kuwata, J. Phys. Soc. Japan. **53**, 4456 (1984)
[84T1] M.L.W. Thewalt and W.G. Mc Mullan, Phys. Rev. B **30**, 6232 (1984)
[84W1] J.P. Wolfe and A. Mysyrowicz, Sci. Am. **250** (3), 70 (1984)
[85F1] D. Fröhlich, A. Nöthe and K. Reimann, Phys. Rev. Lett. **55**, 1355 (1985)
[85H1] B. Hönerlage et al., Phys. Rep. **124**, 161 (1985)
[85H2] D. Hulin et al., J. Luminesc., **30**, 290 (1985)
[86I1] M. Inoue, Phys. Rev. B **33**, 1317 (1986)
[86K1] M. Kuwata, T. Itagaki and N. Nagasawa, Jpn. J. Appl. Phys., **25**, 1382 (1986)
[86M1] A. Mysyrowicz et al., Phys. Rev. Lett. **56**, 2748 (1986)
[86U1] M. Ueta et al., Excitonic Processes in Solids, Springer Ser. Solid-Stte Sci. **60**, Springer, Berlin, Heidelberg (1986)
[86T1] D.P. Trauernicht, J.P. Wolfe and A. Mysyrowicz, Phys. Rev. B **34**, 2561 (1986)
[87F1] D. Fröhlich, K. Reimann and R. Wille, Europhys. Lett. **3**, 853 (1987)
[88O1] Optical Switching in Low Dimensional Solids, H. Haug, L. Banyai, eds., NATO ASI Ser. B **194**, (1988)
[88S1] W. Schäfer, Festkörperprobleme / Advances in Solid State Physics **28**, 63 (1988)
[89C1] D.S. Chemla et al., J. Lumin. **44**, 233 (1989)
[89S1] S. Schmitt-Rink, D.S. Chemla and D.A.B. Miller, Adv. Phys. **38**, 89 (1989)
[90H1] Y.Z. Hu, M. Lindberg and S.W. Koch, Phys. Rev. B **42**, 1713 (1990)
[90K1] C. Klingshirn, Semicond. Sci. Technol. **5**, 457and 1006 (1990)
[90P1] N. Presser, phys. stat. sol. (b) **159**, 443 (1990)
[90U1] A. Uhrig et al.: Z. Phys. B **81**, 385 (1990)
[90S1] D.W. Snoke, J.P. Wolfe and A. Mysyrowicz, Phys. Rev. B **41**, 11171 (1990)
[90S2] D.W. Snoke and J.P. Wolfe, Phys. Rev. B **42**, 7876 (1990)
[90Z1] R. Zimmermann, Festkörperprobleme / Adv. Solid State Phys. **30**, 295 (1990)
[91B1] S. Bar-Ad and I. Bar Joseph, Phys. Rev. Lett. **66**, 2591 (1991)
[91B2] B. Bucher, P. Steiner and P. Wachter, Phys. Rev. Lett. **67**, 2717 (1991)
[91C1] R. Cingolani and K. Ploog, Adv. Phys. **40**, 535 (1991)

[91F1] D. Fröhlich et al., Phys. Rev. Lett. **67**, 2343 (1991)
[92B1] S. Bar-Ad and I. Joseph, I., Phys. Rev. Lett. **68**, 349 (1992)
[92C1] C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, Atom-Photon Interactions: Basic Processes and Applications, Interscience, New York (1992)
[92D1] V.S. Dneprovskii et al., Solid State Commun. **81**, 227 (1992)
[92K1] L.V. Keldysh, phys. stat. sol. (b) **173**, 119 (1992)
[92K2] C. Klingshirn, J. Grohs and M. Wegener, In Nonlinear Dynamics in Solids, ed. by H. Thomas (Springer, Berlin, Heidelberg 1992), p. 88 and references therein
[92O1] D. Oberhauser et al., phys. stat. sol. (b) **173**, 53 (1992)
[92P1] R.T. Phillips et al., Phys. Rev. B **45**, 4308 (1992)
[92P2] R.T. Phillips et al., Phys. Rev. Lett. **68**, 1880 (1992)
[93B1] L. Banyai and S.W. Koch, Semiconductor Quantum Dots World Scientific Series on Atomic, Molecular and Optical Physics, Vol. 2, World Scientific, Singapore, (1993)
[93F1] D. Fröhlich et al., phys. stat. sol. (b) **177**, 379 (1993)
[93G1] S.V. Gaponenko et al., J. Opt. Soc. Am. B **10**, 1947 (1993)
[93P1] K.-H. Pantke et al., Phys. Rev. B **47**, 2413 (1993)
[93Z1] R. Zimmermann and D. Fröhlich, p. 51 in Optics of Semiconductor Nanosstructures, F. Henneberg, S. Schmitt-Rink and E.O. Göbel eds., Akademieverlag, Berlin (1993)
[94B1] L.V. Butov et al., Phys. Rev. Lett. **73**, 304 (1994)
[94B2] S. Bar-Ad et al., Phys. Rev. B **50**, 18375 (1994)
[94H1] R. Houdre et al., Physical Review Letters **73**, 2043 (1994)
[94K1] C. Klingshirn, Adv. Materials for Optics and Electronics **3**, 103 (1994) and references given therein
[94K2] V. Klimov, S. Hunsche and H. Kunze, Phys Rev. B **50**, 8110 (1994)
[94P1] K.-H. Pantke and J.M. Hvam, Intern. Journ. Of Modern Physics B **8**, 73 (1994)
[94W1] U. Woggon et al., J. Lumin. **59**, 135 (1994)
[95B1] E. Benson, E. Fortin and A. Mysyrowicz, phys. stat. sol. (b) **191**, 345 (1995)
[95B2] Bose–Einstein Condensation, D.W. Snoke and S. Stringari, eds., Cambridge University Press, Cambridge, UK (1995)
[95F1] A.M. Fox et al., Phys. Rev. Lett. **74**, 1728 (1995)
[95F2] G. Finkelstein et al., Phys. Rev. Lett. **74**, 976 (1995)
[95K1] F. Kreller et al., Phys. Rev. Lett. **75**, 2420 (1995)
[95W1] U. Woggon, phys. stat. sol. b **189**, 285 (1995)
[95W2] U. Woggon et al., Jpn. J. Appl. Phys. **34**, 232 (1995) and phys. stat. sol. b **188,** 221 (1995)
[96B1] D. Birkedahl et al., J.M.: Phys. Rev. Lett. **76**, 672 (1996)
[96F1] G. Finkelstein et al., Phys. Rev. B **53**, R1709 and 12593 (1996)
[96G1] H. Giessen et al., J. Opt. Soc. Am. **13**, 1039 (1996), Chemical Physics **210**, 71 (1996) and Optics Letters **21**, 1043 (1996)
[96H1] H. Haug and A.-P. Jauho, Quantum Kinetics in Transport and Optics of Semiconductors, Springer Series in Solid State Sciences **123**, Springer, Berlin (1996))
[96H2] M. Hasuo, H. Kawano and N. Nagasawa, J. Liminesc. **66/67**, 396 (1996)

[96K1] G.M. Kavoulakis, G. Baym and J.P. Wolfe, Phys. Rev. B **53**, 7227 and 16625 (1996)
[96K2] G.A. Kopelevich et al., Sov. Phys. JETP **82**, 1180 (1996)
[96M1] A. Mysyrowicz, E. Benson and E. Fortin, Phys. Rev. Lett. **77**, 896 (1996)
[96N1] N. Naka, Crystal Growth and Charact. **33**, 89 (1996)
[96N2] U. Neukirch and K. Wundke, Solid State Commun. **99**, 607 (1996)
[96O1] J.L. Osborne et al., Phys. Rev. B **53**, 13002 (1996)
[96P1] J. Puls, H.-J. Wünsche and F. Henneberger, J. Chem. Phys., **210**, 235 (1996)
[96S1] J. Singh et al., Phys. Rev. B **45**, 15909 (1996)
[96S2] J. Shah, Ultrafast Spectroscopy of Semiconductor Nanostructrures, Springer Series in Solid State Sciences **115**, Springer Berlin (1996)
[96W1] U. Woggon et al., Phys. Rev. B **54**, 17681 (1996), J., Luminesc. **70**, 269 (1996) and Festkörperprobleme / Adv. Solid State Phys. **35,** 175 (1996)
[97L1] W. Langbein et al., Phys. Rev. B **55**, R7383 (1997)
[97M1] V. Mizeikis et al., Phys. Rev. B **55**, 5284 (1997)
[97S1] B. Stébé et al., Phys. Rev. B **56**, 12454 (1997)
[97T1] S.G. Tikhodeev et al., Phys. Rev. Lett. **78**, 3225 (1997)
[97W1] U. Woggon, Springer Tracts in Modern Physics **136** (1997)
[97Z1] R. Zimmermann et al., Phys. Rev. B **56**, R12722 (1997)
[97Z2] P. Zu et al., Solid State Commun. **103**, 459 (1997)
[98B1] T. Baars et al., Phys. Rev. B **58**, R1750 (1998)
[98B2] M. Bayer et al., Phys. Rev. B **58**, 4740 (1998)
[98B3] L.V. Butov and A.I. Filin, Phys. Rev. B **58**, 1980 (1998)
[98E1] C. Ell, A.L. Ivanov and H. Hang, Phys. Rev. B **57**, 9663 (1998)
[98I1] A.L. Ivanov, H. Haug, and L.V. Keldysh, Phys. Reports **296**, 237 (1998)
[98J1] A. Jolk and C. Klingshirn, phys. stat. sol. b **206**, 841 (1998)
[98K1] A. Kuther et al., Phys. Rev. B **58**, R7508 (1998)
[98V1] F. Vouilloz et al., Solid State Commun. **108**, 945 (1998)
[98W1] J. Wrzesinski and D. Fröhlich, J. Cryst. Growth **184/185**, 686 (1998)
[98W2] H.P. Wagner et al., Phys. Rev. B **58**, 10494 (1998)
[99B1] P. Borri et al., Phys. Rev. B **60**, 4505 (1999)
[99D1] R. Denschlag and R. v.Baltz, phys. stat. sol. (b) **215**, 287 (1999)
[99G1] F. Gindele et al., Phys. Rev. B **60**, 8773 and R2157 (1999)
[99H1] L. Hanke et al., Phys. Rev. Lett. **83**, 4365 (1999)
[99L1] Landolt-Börnstein, New Series, Group III, Vol. **41** A2, Rössler, U. ed., Springer, Berlin (1999)
[99L2] W. Langbein and J.M. Hvam, Phys. Rev. B **59**, 15405 (1999)
[99L3] W. Langbein, H. Gislason and J.M. Hvam, Phys. Rev. B **60**, 16667 (1999)
[99L4] W. Langbein, Appl. Phys. Lett. **75**, 2150 (1999)
[99N1] V. Negoita, D.W. Snoke and K. Eberl, Appl. Phys. Lett. **75**, 2059 (1999)
[99O1] K.E. O'Hara, J.R. Gullingsrud and J.P. Wolfe, Phys. Rev. B **60**, 10872 (1999)
[99O2] K.E. O'Hara, L.Ó. Súillebháin and J.P. Wolfe, Phys. Rev. B **60**, 10565 (1999)
[99W1] H.P. Wagner, Phys. Rev. B **59**, 4584 (1999)
[99W2] H.P. Wagner et al., Phys. Rev. B **60**, 10640 (1999)
[00B1] F. Boeuf et al., Phys. Rev. B **62**, R2279 (2000)
[00D1] G. Dasbach et al., phys. stat. sol. (b) **221**, 319 (2000)

[00D2] E. Dekel et al., Phys. Rev. B **61**, 11009 (2000)
[00D3] E. Dekel et al., Phys. Rev. B **62**, 11038 (2000)
[00E1] A. Esser, Phys. Rev. B **62**, 8232 (2000)
[00F1] F. Findeis et al., Phys. Rev. B **61**, R10579 (2000)
[00F2] F Findeis et al., Solid State Commun. **114**, 227 (2000)
[00H1] A. Hartmann, Phys. Rev. Lett. **84**, 5648 (2000)
[00J1] A. Jolk et al., phys. stat. sol. (b) **221**, 295 (2000)
[00K1] H.J. Ko et al., Appl. Phys. Lett. **77**, 537 (2000)
[00K2] G.M. Kavoulakis and A. Mysyrowicz, Phys. Rev. B **61**, 16619 (2000)
[00M1] S.A. Moskalenko and D.W. Snoke, Bose–Einstein Condensation of Excitons and Biexcitons (and Coherent Nonlinear Optics with Excitons), Cambridge University Press, Cambridge, UK (2000)
[00M2] T. Meier et al., Phys. Rev. B **62**, 4218 (2000)
[00N1] T. Nikuni et al., Phys. Rev. Lett. **84**, 5868 (2000)
[00T1] F. Tassone and Y. Yamamoto, phys. stat. sol. A **178**, 119 (2000)
[00T2] D.B. Tran Thoial and H. Hang, Sol State Commun. **115**, 379 (2000)
[00Y1] Y. Yamada et al., Phys. Rev. B **61**, 8363 (2000)
[00Y2] Y. Yamada et al., ICICE Transaction on Elektronics, Japan, **E83-c**, 605 (2000)
[00W1] U. Woggon et al., Phys. Rev. B **61**, 12632 (2000)
[01B1] T. Baars et al., Phys. Rev. B **63**, 153312 (2001)
[01B2] M. Baier et al., Phys. Rev. B **64**, 195326 (2001)
[01B3] L.V. Butov et al., Phys. Rev. Lett. 86, 5608 (2001)
[01B4] P. Brick et al., Phys. Rev. B, 64, 075323 (2001)
[01D1] I. Dmitruk et al., Phys. Rev. B **64**, 045207 (2001)
[01E1] J. Erland et al., Phys. Rev Lett. 86, 5791 (2001)
[01H1] K. Hinzer et al., Phys. Rev. B **63**, 075314 (2001)
[01H2] K. Hild et al., phys. stat. sol. (b) **224**, 379 (2001)
[01J1] M. Jörger et al., Phys. Rev. B **64**, 113204 (2001)
[01J2] K. Johnsen and G.M. Kavoulakis, Phys. Rev. Lett. **86**, 858 (2001)
[01K1] G.M. Kavoulakis, Phys. Rev. B **65**, 035204 (2001)
[01L1] Landolt- Börnstein, New Series, Group III, Vol. 34 C1, C. Klingshirn ed., Springer, Berlin (2001)
[01M1] Ch. Mann et al., Phys. Rev. B **64**, 235206 (2001)
[01M2] E. Moreau et al., Phys. Rev. Lett. **87**, 183601 (2001)
[01M3] H. Mino et al., Physica B **298**, 421 (2001)
[01P1] J. Puls et al., phys. stat. sol. (b) **229**, 637 (2001)
[01P2] R. Perez, J. Physics, Cond. Matter **13**, L539 (2001)
[01R1] A.A. Rogachev, Progress in Quant. Electronics **25**, 141 (2001)
[01S1] D. Sanvitto et al., Science **294**, 837 (2001)
[01S2] Y. Sun, G.K. Wong and J.B. Ketterson, Phys. Rev. B **63**, 125323 (2001)
[01W1] G. Wang, G.K. Wong and J.B. Ketterson, Applied Optics **40**, 5436 (2001)
[01W2] C. Watatani et al., phys. stat. sol. (b) **224**, 353 (2001)
[01W3] A.J. Williamson, A. Franceschitti and A. Zunger, Europhys. Lett. **53**, 59 (2001)
[01Z1] A. Zrenner et al., Physica B **298**, 239 (2001)
[02A1] G.V. Astakhov et al., Phys. Rev. B **65**, 165335 (2002)
[02A2] G.V. Astakhov et al., Phys. Rev. B **65**, 115310 (2002)
[02A3] I.A. Akimov et al, Appl. Phys. Lett. **81**, 4730 (2002)

[02B1] M. Bayer et al., Phys. Rev. B **65**, 195315 (2002)
[02B2] L.V. Butov et al., Nature 417, 47 (2002)
[02B3] L.V. Butov, A.C. Gossard and D.S. Chemla, Nature 418, 751 (2002)
[02B4] J.J. Baumberg, Physics World **15**, 37 (2002) or in [04E1]
[02B5] J.J. Baumberg et al., Physica A **13**, 358 (2002)
[02B6] A.A. Burkov and A.H. MacDonald, Phys. Rev. B **66**, 115320 (2002)
[02B7] E. Bascoes, A.A. Burkov and A.H. MacDonald, Phys. Rev. Lett. **89**, 086401 (2002)
[02C1] C.K. Choi et al., Phys. Rev. B, 65, 155206 (2002)
[02D1] S. Denev and D.W. Snoke, Phys. Rev. B **65**, 085211 (2002)
[02G1] J. Gutowski et al., phys. stat. sol. (b) **229**, 653 (2002)
[02H1] R. Huang et al., Phys. Rev. B **65**, 165314 (2002)
[02J1] A.D. Jackson and G.M. Kavoulakis, Europhys. Lett. **59**, 807 (2002)
[02J2] A. Jolk, M. Jörger and C. Klingshirn, Phys. Rev. B **65**, 245209 (2002)
[02K1] C. Klingshirn, T. Fleck and M. Jörger, phys. stat. sol. (b) **234**, 23 (2002)
[02K2] W. Ketterle, Intern. J. Mod. Phys. B **16**, 4537 (2002) and Rev. Mod. Phys. **74**, 1131 (2002)
[02L1] Landolt-Börnstein, New Series, Group III, Vol. 41 A1B, Rössler, U., ed., Springer, Berlin (2002)
[02L2] A.V. Larionov et al., JETP Lett. **75**, 570 (2002)
[02L3] P.B. Littlewood et al., phys. stat. sol. (b) **234**, 36 (2002)
[02M1] A.A. Mikhalovsky et al., Appl. Phys. Lett., **80**, 2380 (2002)
[02M2] A.V. Malko et al., Appl. Phys. Lett., **81**, 1303 (2002)
[02N1] N. Naka and N. Nagasawa, Phys. Rev. B **65**, 075209 (2002)
[02N2] N. Naka and N. Nagasawa, Phys. Rev. B 65, 245203 (2002)
[02N3] M. Nagai and M. Kuwata-Gonokami, J. Luminesc. **100**, 233 (2002)
[02R1] C. Riva et al., phys. stat. sol. (b) **234**, 50 (2002)
[02S1] Ch. Santori et al., Phys. Rev. B **66**, 045308 (2002)
[02S2] Ch. Santori et al., Phys. Rev. B **65**, 073310 (2002)
[02S3] D. Snoke et al., Nature **418**, 754 (2002)
[02S4] D. Snoke, Science **298**, 1368 (2002)
[02S5] M.S. Skolnick et al., Mat. Science and Engineering C **19**, 407 (2002)
[02T1] I. Tanaka and M. Nakayama, Molecular Crystals & Liquid Crystals Science & Technology B **29**, 507 (2002)
[02W1] H.P. Wagner et al., phys. stat. sol. (b) **231**, 11 (2002)
[03A1] S. Adachi et al., Phys. Rev. B **67**, 205212 (2003) and phys. stat. sol. b **240**, 348 (2003)
[03B1] L.V. Butov, phys. stat. sol. b **238** (2003)
[03B2] L.V. Butov, Solid State Commun., **127**, 89 **(**2003**)**
[03B3] P. Borri et al., Semicond. Science and Technol. **18**, 351 (2003)
[03D1] G. Dasbach et al., Phys. Rev. Lett. **91**, 107401 (2003) and phys. stat. sol. b **238**, 541 (2003) and Phys. Rev. B (2004) in press
[03F1] Qu. Fanyao and P.C. Morais, Phys. Lett. A **310**, 460 (2003)
[03H1] H. Htoon et al., Appl. Phys. Lett. **82**, 4776 (2003)
[03H2] K. Hazu, Phys. Rev. B **68**, 033205 (2003)
[03J1] M. Jörger et al., phys. stat. sol. b **238**, 470 (2003)
[03K1] L.V. Keldysh, private communication (2003)
[03K2] K.V. Kavokin, phys. stat. sol. a **195,** 592 **(2003)**
[03K3] C. Klingshirn, T. Fleck and M. Jörger, phys. stat. sol. b **239**, 261 (2003)

[03K4] E. Kapon in Ref. [81A1]l of Chap. 1
[03L1] V.G. Lyssenko, private communication (2003)
[03L2] P.B. Littlewood, private communication (2003)
[03L3] P.G. Lagoudakis et al., Phys. Rev. Lett. **90**, 206401 (2003)
[03M1] A. Mysyrowicz, P.J. Leblanc and E. Fortin, phys. stat. sol. b **239**, 257 (2003)
[03N1] N. Naka and N. Nagasawa, phys. stat. sol. b **238**, 397 (2003) and in Refs. [04I1, 04E1]
[03N2] Nonlinear Optics and Excitation Kinetics (NOEKS 7), Karlsruhe: Proc. published in phys. stat. sol. C, **0** issue 5 (2003)
[03P1] B. Patton, W. Langbein and U. Woggon, Phys. Rev. B **68**, 125316 (2003)
[03P2] L. Pitaevskii and S. Stringari, Bose–Einstein Condensation, Clarendon Press, Oxford (2003)
[03S1] M.H. Szymanska and P.B. Littlewood, Phys. Rev. B 67, 193305 (2003)
[03S2] D. Snoke, phys. stat. sol. b **238**, 389 (2003)
[03S3] S.H. Simon et al., Phys. Rev. Lett. **91**, 646803 (2003)
[03U1] S.M. Ulrich et al., Appl. Phys. Lett., **83**, 1848 (2003)
[03W1] P. Wachter, B. Bucher and J. Malar, Europhys. Lett. **62**, 343 (2003)
[04A1] S. Adachi, EXCON '04, Cracow, to be published in J. Luminesc.
[04E1] EXCON '04, Cracow, Proc. to be published in J. Luminesc.
[04I1] Intern. Conf. on Spontaneous Coherence in Excitonic Systems (ICSCES), Pitsburgh, 2004 to be published in Solid. State Commun.
[04J1] M. Jörger PhD Thesis, Karlsruhe (2004) and EXCON '04, Cracow, to be published in J. Luminesc.
[04J2] Z. Ji et al., EXCON '04, Cracow, to be published in J. Luminesc.
[04K1] A. Klochikin et al., 27^{th} ICPS, Flagstaff (2004) to be published
[04K2] C. Klingshirn et al., in Ref. [04I1]
[04K3] M. Kellog et al., Phys. Rev. Lett **93**, 036801 (2004)
[04L1] Landolt-Börnstein, New Series, Group III, Vol. C2, C. Klingshirn, ed., Springer, Berlin (2004), in press
[04L2] W. Langbein et al., Phys. Rev. B **69**, 161301(R) (2004)
[04R2] R. Rapaport et al., Phys. Rev. Lett. **92**, 117405 (2004) and Ref. [04I1]
[04S1] D. Snoke et al., in Ref. [04I1]
[04W1] J. Wolf in Ref. [04I1]

21

The Electron–Hole Plasma

Having introduced the basic idea of the electron–hole plasma (EHP) in Sect. 19.2, we now give details of some of its properties, e.g., the density at which the transition from an exciton gas to an EHP occurs in Sect. 21.1, the renormalization of the band gap in the EHP and its thermodynamic properties in Sect. 21.2. Then we present results for indirect and direct-gap semiconductors showing some characteristic differences in these properties. The next section brings the results of the EHP in semiconductor structures of reduced dimensionality followed by aspects of inter-subband transitions in bulk semiconductors and in structures of reduced dimensionality (Sect. 21.5).

For early ideas of the EHP see, e.g., [68K1, 74T1, 77H1, 77R1] and the references therein. The concept of the Mott-transition was introduced for highly doped semiconductors and later adapted for highly excited ones [74M1]. Strong input also came from the physics of metals including the jellium model. Some reviews for further reading beyond the selected topics addressed below are, e.g., [81K1, 84H1, 84K1, 85B1, 85E1, 88Z1, 90C1, 90H1, 92K1, 94C1, 94K1, 94K2, 96K1, 04O1].

21.1 The Mott Density

Equation (19.20) and (20.17) gave a very crude approximation of the density of electron–hole pairs n_c at which the transition from an excitonic system to an EHP can be expected. In this section we give some more refined considerations.

The transition to the EHP can be tackled in the following way. We consider one exciton in a sea of free carriers (electrons and holes) of density n_p. The free carriers screen the Coulomb potential of the exciton transforming it to a Yukawa-type potential. For the derivation of this and the following formula see [84H1, 84K1, 88Z1, 90H1] and references therein.

$$\frac{1}{4\pi\varepsilon_0\varepsilon}\frac{e^2}{|\boldsymbol{r}_\mathrm{e}-\boldsymbol{r}_\mathrm{h}|} \Rightarrow \frac{1}{4\pi\varepsilon_0\varepsilon\,(n_\mathrm{P})}\exp\left\{-\frac{|\boldsymbol{r}_\mathrm{e}-\boldsymbol{r}_\mathrm{h}|}{l}\right\}\cdot\frac{\mathrm{e}^2}{|\boldsymbol{r}_\mathrm{e}-\boldsymbol{r}_\mathrm{h}|} \tag{21.1}$$

where some thought is necessary concerning the value of ε entering in this equation. In the simplest approximation one could use the static dielectric constant (static screening). A better approximation is obtained if both the dependence of ε on frequency and on n_p are taken into account and thus the contributions of plasmons or of plasmon–phonon mixed states to $\varepsilon(\omega, n_\mathrm{P})$, the so-called dynamic screening [81K1, 85B1, 88Z1]. If the screening length l falls below a certain value l_c, the Yukawa potential no longer has a bound state, at least in three-dimensional systems. The excitonic Bohr radius and l_c are connected with each other by

$$a_\mathrm{B} l_\mathrm{c}^{-1} = 1.19\,; \quad l^{-1} =: k_\mathrm{s}\,. \tag{21.2}$$

The inverse screening lenght l^{-1} is also called screening-"vector".

By considering an electron–hole gas described by classical Boltzmann statistics, one can find the so-called Debye–Hückel screening length l_DH [23D1] and the density n_M at which the EHP starts to exist or, in other words, where the excitons cease to exist as individual quasi-particles, is given by

$$l_\mathrm{DH} = \left(\frac{\varepsilon_0 \varepsilon k_\mathrm{B} T}{e^2 n_\mathrm{P}}\right)^{1/2}, \tag{21.3a}$$

$$n_\mathrm{M} = (1.19)^2 \frac{\varepsilon\varepsilon_0 k_\mathrm{B} T}{e^2 a_\mathrm{B}^2} = (1.19)^2 \frac{k_\mathrm{B} T}{2 a_\mathrm{B}^3 \mathrm{Ry}^*}\,, \tag{21.3b}$$

since we have

$$a_\mathrm{B}\mathrm{Ry}^* = \frac{\hbar^2 \varepsilon\varepsilon_0}{\mu e^2} \cdot \frac{e^4 \mu}{2\left(\varepsilon\varepsilon_0\right)^2 \hbar^2} = \frac{e^2}{2\varepsilon\varepsilon_0}\,, \tag{21.3c}$$

where μ is the reduced mass of electron and hole.

Though this is a reasonable approach at higher temperatures at which a large fraction of excitons is thermally ionized, it gives the physically unreasonable result that in the limit $T \Rightarrow 0\, l_\mathrm{c}$ is reached already for vanishing values of n_p, which can be orders of magnitude smaller than what is given by (19.20).

For the case of a degenerate electron–hole plasma, which is more likely at low temperatures, (see, (8.41)), one obtains the Thomas–Fermi screening length l_TF [27T1, 28F1] and correspondingly a different value for n_M

$$l_\mathrm{TF}^{-2} = \frac{e^2}{\varepsilon\varepsilon_0} \sum_{i=\mathrm{e,h}} \frac{\partial n_i}{\partial E_\mathrm{F}^i}\,, \tag{21.4a}$$

where E_F^i stands for the quasi Fermi levels of electrons and holes, respectively, and the terms $\partial n_i / \partial E_\mathrm{F}^i$ give the densities of states at the quasi Fermi energies. In the effective mass approximation (21.4a) simplifies for isotropic, non-degenerate bands to:

$$l_\mathrm{TF}^{-1} = \left[\frac{3e^2}{\varepsilon\varepsilon_0\hbar^2}\left(m_\mathrm{e} + m_\mathrm{h}\right)\left(\frac{1}{3\pi^2}\right)^{2/3} n_\mathrm{P}^{1/3}\right]^{1/2}$$

and

$$n_{\mathrm{M}} = (1.19)^6 \, a_{\mathrm{B}}^{-6} \left(\frac{\varepsilon\varepsilon_0 \hbar^2}{3e^2} \frac{1}{m_{\mathrm{e}} + m_{\mathrm{h}}} \right)^3 \left(3\pi^2\right)^2 \tag{21.4b}$$

The density n_{M} given by (21.3),(21.4) is also known as the Mott density, because it describes the transition from an insulating gas of excitons at lower densities to the metal-like state of an EHP at higher densities [74M1].

A description of the screening on a more sophisticated theoretical level involves the so-called random-phase approximation (RPA) or the simplified version of the single plasmon pole (SPP) approximation. In the latter case the contribution of the free carriers in the EHP to $\varepsilon(\omega)$ is described by a pole at the plasma frequency ω_{PL} (Sect. 10.1) or more precisely by the plasmon–phonon mixed state (Sects. 10.1 and 12.4). These topics are however beyond the scope of this book and we refer the reader to [77H1, 77R1, 81K1, 84H1, 84K1, 88Z1] for earlier calculations, to [98M1, 98M2, 98P1, 98S1] for more recent ones and references therein.

21.2 Band Gap Renormalization and Phase Diagram

In addition to the screening of the Coulomb interaction in the exciton, there are further important renormalization effects of the electronic eigenstates in an EHP. They will be treated in this section and we shall show how they result in the formation of a liquid-like state of the plasma below some critical temperature T_{c}, at least under thermodynamic quasi-equilibrium conditions.

The width of the forbidden gap is a monotonically decreasing function of the electron–hole pair density n_{P} in the plasma, due to exchange and correlation effects as shown in Fig. 21.1, where we give various energies as a function of n_{P}. This statement can be explained qualitatively in the following way. In the plasma we have Coulomb energies which attract carriers of opposite charge and repel those of like charge. If electrons and holes were completely randomly distributed in the sample, then the Coulomb attraction and repulsion energies would cancel exactly and the width of the forbidden gap would be independent of n_{P}.

In reality, the carriers are not randomly distributed. The Pauli principle which is a consequence of the exchange interaction of identical fermions forbids two electrons with parallel spin from sitting in the same unit cell. Since this situation would occur for a random distribution, but does not for fermions, we can conclude that the exchange energy increases the average distance between electrons with parallel spin and consequently reduces their total repulsive Coulomb energy. The reduction of a repulsive energy term means a lowering of the total energy of the electron system. The same arguments also hold for the holes.

The correlation energy is spin independent and describes the fact that the electron–hole pair system can lower its energy further, if the distribution of

electrons and holes relative to each other is not random, but if in the vicinity of a hole an electron is found with higher probability than another hole and vice versa. This correlation really occurs. It is a "reminder" of the Coulomb interaction between electron and hole which is responsible for the formation of excitons at low densities. Usually the sum of these two terms is called exchange and correlation energy.

There are some universal formula that describe the reduction of the band gap ΔE_{g} normalized by the excitonic Rydberg energy as a function of the normalized plasma density n_{P} including the exchange (or Hartree–Fock) and the correlation energy [82V1,88Z1,88Z2]. Usually one chooses a dimensionless quantity called r_{s} in which the volume occupied by a carrier pair in the plasma is compared with the volume of an exciton

$$r_{\mathrm{s}} = \left(\frac{4\pi a_{\mathrm{B}}^3}{3} n_{\mathrm{P}}\right)^{-1/3} . \tag{21.5}$$

One should note that r_{s} decreases with increasing n_{P} due to the negative exponent. In an EHP in semiconductors values of r_{s} generally lie between 1 and 4.

In Fig. 21.2 we give approaches to the universal behavior of

$$\Delta E_{\mathrm{g}}/\mathrm{Ry}^* = f(r_{\mathrm{s}}) \tag{21.6}$$

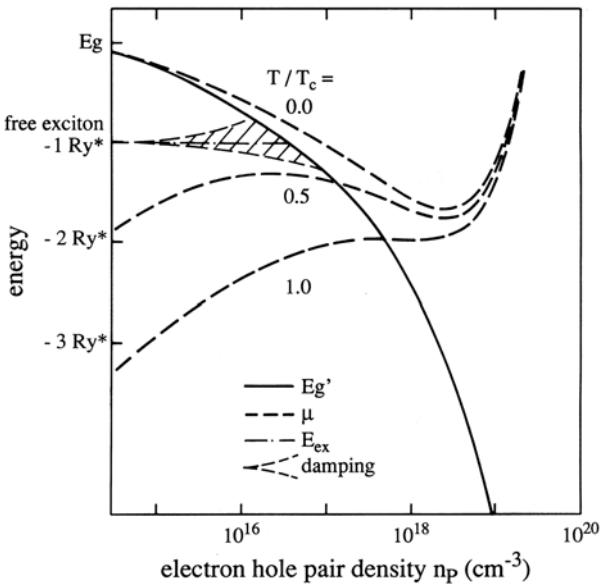


Fig. 21.1. Schematic dependence of the width of the forbidden gap E_{g}', of the exciton energies, and of the chemical potential of the electron–hole pair system μ on the electron–hole pair density n_{P}. Numerical data for CdS [88Z1]

from the three references [82V1,88Z1,88Z2]. It can be stated that the band-gap renormalization is almost temperature independent, and that the universal curves describe the experimental situation very well in the more covalently bound semiconductors like Si, Ge, GaAs, GaP etc. For clarity these are not shown in Fig. 21.2 since the experimental data for these substances coincide closely with the theoretical curves. In contrast, further corrections have to be considered for substances with increasing degree of ionic binding like the II–VI materials CdS or ZnSe. We come back to this aspect later.

In addition to E'_{g} (n_{P}) Fig. 21.1 shows the energy of the 1s exciton as a function of n_{P}. It has been found that the excitons roughly maintain their absolute energy with increasing n_{P} until the binding energy vanishes at the Mott density. Sometimes a small red- or blue-shift is observed with increasing n_{P}, especially in narrow QW see, e.g., [00M1] and references therein. This finding is due to the almost complete compensation of two effects. One is the decrease of E'_{g} with increasing n_{P} which should shift the exciton energy to the red. The other is the decrease of the exciton binding energy with increasing n_{P} due to the screening of the Coulomb energy mentioned in Sect. 21.1 which shifts the exciton closer to the gap, i.e., towards the blue.

This screening of the binding energy leads to an increase of the excitonic radius and thus to a decrease of the overlap of electron and hole wavefunction and a decrease of oscillator-strength with increasing n_{P}. Furthermore, the damping of the exciton resonances increases with n_{P} due to an increasing scattering rate which reduces the phase relaxation time T_2. This effect is indicated schematically by the hatched area in Fig. 21.1. The Mott density introduced in Sect. 21.1 is, in the representation of Fig. 21.1, the density

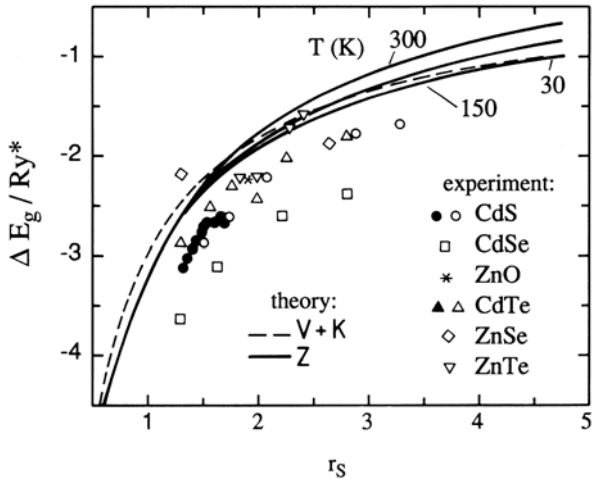


Fig. 21.2. The universal relation between the band-gap renormalization ΔE_{g} measured in units of the excitonic Rydberg energy Ry^* and the normalized carrier density $r_{\mathrm{s}}(n_{\mathrm{P}})$ together with data for various II–VI compounds [82K1, 88Z2, 90K1]

at which the curves $E_{\rm ex}(n_{\rm P})$ and $E'_{\rm g}(n_{\rm P})$ cross. This density decreases with increasing quantum number $n_{\rm B}$. For experimental verifications of these facts see, e.g., [78L1, 00J1]. Usually one concentrates on the lowest exciton, i.e., on $n_{\rm B} = 1$.

Until now we have discussed mainly the screening of the Coulomb interaction between electrons and holes as a reason for the disappearance of the exciton as a bound electron–hole pair state. Actually there is another process which contributes, namely phase-space filling [88C1]. As has been explained in connection with (8.10) and (9.4),(9.6), one needs electron and hole wavefunctions from a certain regime in $\boldsymbol{k}$-space to build up the exciton wavefunction. If these states are occupied by electrons and holes, they are blocked and can no longer be used for the construction of the exciton. This phase-space filling is again a consequence of the fermionic character of the constituents of the exciton. It is already present in three-dimensional semiconductors but plays a more important role in QW and in structures of even lower quasi-dimensionality, where the screening of the Coulomb interaction is reduced, since it is not so easy to screen the electric field lines between electron and hole which propagate through the barriers.

The next quantity given in Fig. 21.1 is the chemical potential μ of the electron–hole pair system. In a pumped system away from thermodynamic equilibrium the distribution of electrons and holes in their bands can no longer be described by a single Fermi energy (or chemical potential) $E_{\rm F}$. Instead individual quasi Fermi energies $E_{\rm F}^{\rm e}$ and $E_{\rm F}^{\rm h}$ have to be introduced for electrons and holes, respectively. The chemical potential of the electron–hole pair system is just the energetic distance between $E_{\rm F}^{\rm e}$ and $E_{\rm F}^{\rm h}$. If we measure them from the extrema of the renormalized bands, we get

$$\mu(n_{\rm P}, T_{\rm P}) = E_{\rm F}^{\rm e}(n_{\rm P}, T_{\rm P}) + E_{\rm F}^{\rm h}(n_{\rm P}, T_{\rm P}) + E'_{\rm g}(n_{\rm P})\,, \tag{21.7}$$

where we assume that the carriers have a thermal distribution in their respective bands and that the temperatures of electrons $T_{\rm e}$ and of holes $T_{\rm h}$ are equal. We call this temperature then the plasma temperature $T_{\rm P}$:

$$T_{\rm e} = T_{\rm h} = T_{\rm P}\,. \tag{21.8}$$

Depending on the material and the excitation conditions, $T_{\rm P}$ can be higher than the lattice temperature $T_{\rm L}$ [81K1, 04C1]. This means that a thermal distribution of electrons and holes is established by mutual scattering (this generally happens in an EHP on a ps time scale) but that this distribution is not in equilibrium with the phonon system described by $T_{\rm L}$.

Without pumping, μ is zero. It increases with increasing density. At some density, μ will exceed $E'_{\rm g}(n_{\rm P})$. In this case the quasi Fermi level of the electrons and possibly also of the holes are situated in the bands. This is the onset of population inversion between conduction and valence band. We speak in this situation of a degenerate EHP and the use of Fermi statistics is mandatory.

When neither $E_{\mathrm{F}}^{\mathrm{e}}$ nor $E_{\mathrm{F}}^{\mathrm{h}}$ are situated in the respective bands, Boltzman statistics are a good approximation.

In the degenerate case, the mean kinetic energy of the particles increases rapidly with density. At zero temperature a simple relation holds in the effective mass approximation

$$\bar{E}_{\mathrm{kin}}^{\mathrm{e,h}} = \frac{3}{5} E_{\mathrm{F}}^{\mathrm{e,h}} \,. \tag{21.9}$$

The steep increase of μ with increasing n_{P} is due to this effect and at high densities (21.9) overcomes the decrease of $E_{\mathrm{g}}'(n_{\mathrm{P}})$. In the literature one often finds the term "Burstein–Moss shift". It means a blue shift of the absorption edge for highly doped or excited semiconductors and it is explained by a generally degenerate population of, e.g., electrons in the conduction band, which blocks optical transitions close to the band extremum. The inspection of Fig. 21.1 tells us that the situation is more complex because we have the interplay of band gap renormalization, which shifts the onset of absorption to the red, and state filling, which shifts it to the blue. The application of a simple Burstein–Moss shift would only be justified for a situation where the band-gap renormalization is much smaller than the band filling.

The free energy $F(n_{\mathrm{P}}, T_{\mathrm{P}})$ of the EHP is connected with $\mu(n_{\mathrm{P}}, T_{\mathrm{P}})$ by [74T1]

$$\mu\,(n_{\mathrm{P}}, T_{\mathrm{P}}) = \left(\frac{\partial n_{\mathrm{P}} F\,(n_{\mathrm{P}}, T_{\mathrm{P}})}{\partial n_{\mathrm{P}}}\right)_{T,V} . \tag{21.10a}$$

On the other hand, it is given by

$$F(n_{\mathrm{P}}, T_{\mathrm{P}}) = U(n_{\mathrm{P}}, T_{\mathrm{P}}) - T_{\mathrm{P}} S(n_{\mathrm{P}} T_{\mathrm{P}}) \tag{21.10b}$$

with S being entropy and U being the internal energy of the EHP. The pressure $p(n_{\mathrm{P}}, T_{\mathrm{P}})$ of the plasma is given by

$$p\,(n_{\mathrm{P}},\, T_{\mathrm{P}}) = n_{\mathrm{P}}^2 \left(\frac{\partial F\,(n_{\mathrm{P}},\, T_{\mathrm{P}})}{\partial n_{\mathrm{P}}}\right)_{N,T} \tag{21.10c}$$

where N is the number of the electron hole pairs in the system.

Below a certain critical temperature T_{c} the dependence of μ on n_{P} is not monotonic but may go through a maximum and a minimum as shown in Fig. 21.1. In this situation quasi-equilibrium thermodynamics predict a first-order phase transition to an electron–hole liquid (EHL) [74T1, 81K1, 84H1, 84K1, 88Z1]. Thus, similar to the case of a real or van der Waals gas, below T_{c} we expect a phase separation into a liquid-like EHL surrounded by a gas phase of excitons and free carriers. Liquid-like means essentially that the density of the plasma is constant for constant T_{P}. An increase of the pump-power, i.e., of the average density then merely increases the ratio of the volumes filled by the EHL and the gas. The phase diagram of this transition follows from a Maxwell-type construction and is shown in Fig. 21.3 for the simplest case, where we have a coexistence region of gas and liquid below T_{c} while

a distinction between gas and liquid is not physically meaningful above T_c. Modifications of this simple phase diagram may be introduced if one considers excitons and biexcitons, possible regions where these quasi-particles might undergo a Bose condensation, the influence of the finite lifetime, etc. These complications have already been partly discussed in Sect. 20.5 in connection with Fig. 20.16 and are therefore not repeated here.

There are some scaling rules for the phase diagram assuming parabolic bands, a temperature independent exchange and correlation energy (see Fig. 21.2) and a free energy varying quadratically with temperature

$$\frac{E_0}{k_B T_c} = \text{const} \tag{21.11a}$$

$$\frac{n_c}{n_0} = \text{const} \tag{21.11b}$$

where E_0 is the ground state energy of the EHP at $T = 0\,\mathrm{K}$ with respect to the gap at $n_P = 0$, i.e.,

$$E_0\,(n_0) = E'_g\,(n_0, T = 0) + \frac{3}{5} E_F\,(n_0, T_P = 0) - E_g\,(n_P = 0, T = 0) \tag{21.11c}$$

and where T_c and n_c are data of the critical point.
The binding energy $\Phi(n_0, T_P = 0)$ of the EHP is defined as the difference

$$\Phi(n_0, T_P = 0) = E_{ex} - \mu(n_0, T_P = 0)\,. \tag{21.11d}$$

The EHL has indeed been observed in some semiconductors in the form of small liquid-like EHP droplets (EHD), especially if the lifetime of the carriers is long enough that the spatial separation into a liquid and a gas phase can develop. This is often the case for indirect gap materials like C, Si, Ge, GaP or $Al_{1-y}Ga_yAs$ ($y > 0.45$) [75R1, 77H1, 77H2, 77R1, 84H1, 84K1, 85B1, 85E1, 92K1, 96K1, 02N3, 02N4, 02S1, 03N1].

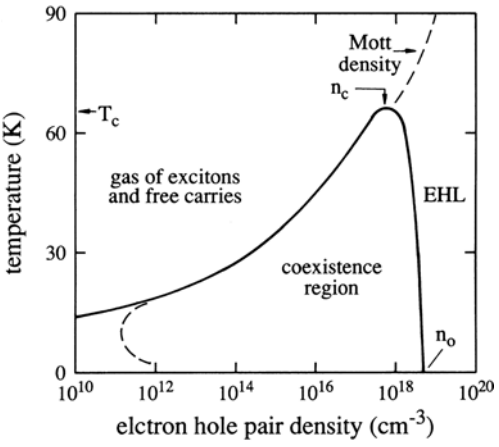


Fig. 21.3. The phase diagram for an EHP under thermodynamic quasi-equilibrium conditions [88Z1]. Compare with Fig. 20.16

In the direct gap materials with dipole allowed band-to-band transition the lifetime of the carriers is generally so short ($\tau \approx 0.1\,\mathrm{ns}$) that no liquid phase can develop, though an EHP is created under sufficiently strong pumping. It has a positive binding energy Φ as for the indirect gap materials [78B1, 79C1]. We shall see examples of both cases (the formation of a liquid-like phase or not) in the next sections. Recently EHL formation has even been claimed in CuCl under resonant excitation [00N1, 02N1].

Before that we comment briefly on the description of the EHP used so far. We considered an exciton in a sea of electrons and holes. This description is adequate at higher temperatures, where a significient fraction of the excitons is thermally ionized, or at densities around n_M, or under band-to-band excitation, where primarily free carriers are created, which form excitons only after relaxation from the continuum to the bound states [04C1]. If we excite excitons resonantly, we should consider the screening and phase-space filling of excitons by excitons and not by free carriers. Furthermore, the curve of $\mu(n_\mathrm{P})$ should start for $T_\mathrm{P} = 0\,\mathrm{K}$ and $n_\mathrm{P} \to 0$ not from E_g but from E_ex. The corresponding calculations and experiments have been made and indicate that the above-outlined scenario of the EHP formation remains qualitatively valid except for a possible shift of n_M to higher values [81S1, 82F1, 82F2, 84M1, 98M1, 98M2, 98P1, 98S1, 00M2].

21.3 Electron–Hole Plasmas in Bulk Semiconductors

We preset now some selected results starting with indirect gap semiconductors, which have usuallly sufficiently long carrier lifetimes to develop a phase separation to an electron–hole liquid below T_c.

21.3.1 Indirect Gap Materials

The first observation of an EHP was in Si. The "history" of its investigation is outlined briefly in reviews [77H1, 77R1, 85E1]. In Fig. 21.4 we show a luminescence spectrum of Ge under high excitation. Due to the indirect bandstructure, the emission bands shown involve a momentum-conserving LA phonon. We see the recombination of free excitons and a broad band which is the emission of the EHL. Due to the long carrier lifetime in pure samples, which is in the μs to ms regime, the phase separation can develop and the carriers can cool down to equilibrium with the lattice so that very low values of T_P close to T_L can be reached. As expected for the liquid phase the shape of this emission does not change with increasing pump intensity below T_c but the ratio of free exciton to EHP emission changes according to the changing volume fractions that the two phases occupy. The density of electron–hole pairs can be estimated from the generation rate and the lifetime. The onset of plasma luminescence for a fixed T_L then gives the low density side of the coexistence region in the phase diagram. The high density side

can be deduced more easily from an analysis of the lineshape of the plasma luminescence.

Since a momentum-conserving phonon is involved, the emission at a certain energy $\hbar\omega$ is given simply by an integral over all transitions separated by $\hbar\omega + \hbar\Omega_{\mathrm{LA}}$, independent of the wave vectors of electron and hole, weighted by the density of states and the occupation probabilities of electrons and of holes in their respective bands.

For spontaneous emission we get in the simplest approximation

$$I_{\mathrm{lum}}(\hbar\omega) = c_1 \int_0^\infty \int_0^\infty \left(E_{\mathrm{kin}}^{\mathrm{e}}\right)^{1/2} \left(E_{\mathrm{kin}}^{\mathrm{h}}\right)^{1/2} f_{\mathrm{e}} f_{\mathrm{h}} \delta\left[\hbar\omega - \left(E_{\mathrm{g}}' + E_{\mathrm{kin}}^{\mathrm{e}} + E_{\mathrm{kin}}^{\mathrm{h}} + \hbar\Omega_{\mathrm{LA}}\right)\right] \mathrm{d}E_{\mathrm{kin}}^{\mathrm{e}}\, \mathrm{d}E_{\mathrm{kin}}^{\mathrm{h}}\,, \tag{21.12}$$

where we have assumed parabolic bands, i.e., $E_{\mathrm{kin}}^{\mathrm{e,h}} = \hbar^2 k_{\mathrm{e,h}}^2 / 2m_{\mathrm{e,h}}$ resulting in the square-root dependence of the density of states and the energy conservation contained in the δ-function. The quasi Fermi functions are given by

$$f_{\mathrm{e,h}}(\hbar\omega) = \left[1 + \exp\left(\pm \frac{E_{\mathrm{kin}}^{\mathrm{e,h}} - E_{\mathrm{F}}^{\mathrm{e,h}}}{k_{\mathrm{B}} T}\right)\right]^{-1}. \tag{21.13}$$

Furthermore the transition probability is assumed in (21.12) to be energy independent and to be contained in c_1 together with other constant terms.

The term for stimulated emission or reabsorption looks similar, except that the product $f_{\mathrm{e}} f_{\mathrm{h}}$ in (21.12) is replaced by $(1 - f_{\mathrm{e}} - f_{\mathrm{h}})$. Stimulated emission however is of no importance for most indirect gap materials since the transition probability is very small. This prevents the standard semiconductor Si from being used as a material for laser diodes.

Recently it has been shown, however, that stimulated emission from the EHP can be obtained in indirect $\mathrm{Al}_{1-x}\mathrm{Ga}_x\mathrm{As}$ and similar compounds for compositions x close to the direct-indirect transition due to a strong coupling

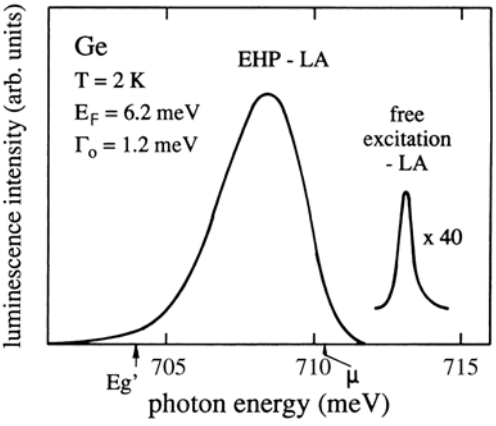


Fig. 21.4. The luminescence of Ge in the presence of an EHP (LA phonon replica) [77M1]

of the direct and the indirect conduction band minima due to alloy scattering [88K1, 90F1, 92K1, 92K2, 92P1, 94K1, 96K1, 98D1].

Some modifications can be introduced in (21.12), e.g., by final state damping Γ_0. We come back to these effects in connection with direct gap materials in Sect. 21.3.2.

A detailed analysis of luminescence spectra like that of. Fig. 21.4 along the lines outlined above gives the values of T_P, $E_\mathrm{F}^{\mathrm{e,h}}$ and consequently the densities n_P and $E'_\mathrm{g}(n_\mathrm{P})$. This allows one to determine the high density side of the phase diagram of the EHL. The low density side can be deduced from the onset of EHL luminescence with increasing excitation, i.e., generation rate. Examples of such results are shown in Fig. 21.5 for Ge and Si. The critical temperatures T_c for the EHL in Si and Ge have been determined to be around 23 and 6.5 K, respectively. In addition, the band-gap renormalization $E'_\mathrm{g}(n_\mathrm{P})$ is found. It agrees nicely with the universal formula of Fig. 21.2.

The multi-valley structure of the conduction band reduces the kinetic energy and helps to stabilize the plasma. By applying homogeneous uniaxial stress to the samples it is possible to lift the degeneracy of the multi-valley conduction band structure. The results elucidated the influence of the band peculiarity on the kinetic energy and on $\mu(n_\mathrm{P}, T_\mathrm{P})$ [81F1,82F1]. More complex phase diagrams are discussed, e.g., in [95S1].

As an example for a less common semiconductor and for an indirect gap compound semiconductors we give a references for the EHP phase diagram in C and GaP [79M1, 02N4, 02S1, 03N1]. For the discussion of EHL in the wide gap, direct semiconductor CuCL see [02N1].

The size and the number of carriers in the EHD has been determined by various techniques including disintegrating them in the internal field of a pn-

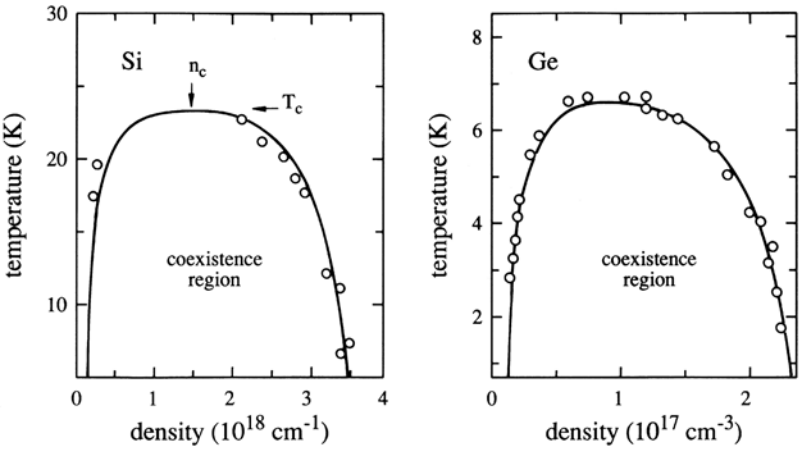


Fig. 21.5. The phase diagrams of the electron–hole liquid in Si and Ge. According to [81F1] and [74T1], respectively. Note the linear and logarithmic scales on the x-axes of Figs. 21.3 and 21.5

junction and measuring the current pulses connected with this event. Usually one finds about 10^3 carrier pairs per droplet in Si or Ge [77H1, 77R1, 83C1, 85E1].

The EHD can be pinned by impurities such as those which produce the multi-exciton complexes (see Sect. 14.1) and which may act thus as nucleation centers for the EHD condensation [79M1]. In pure materials the EHP may move through the sample in a diffusive way or be driven by a heat flow (phonon wind) if there is a spatial gradient in the phonon populations [79W1, 80G1]. In Fig. 21.6a and b we show an example. The phonons are created in the excitation spot by the nonradiative recombination of carriers. Since the phonons propagate (ballistically) into some preferential directions determined by their dispersion, they push the droplets also into these directions. The expansion of the droplet cloud is therefore in-homogeneous as seen in Fig. 21.6 a where the luminescence from the EHD is recorded directly and even better in b where contours of equal luminescence brightness are given.

We want to conclude the discussion of indirect bulk semiconductors with a beautiful experiment concerning the so-called γ-drops in Fig. 21.6c,d. The

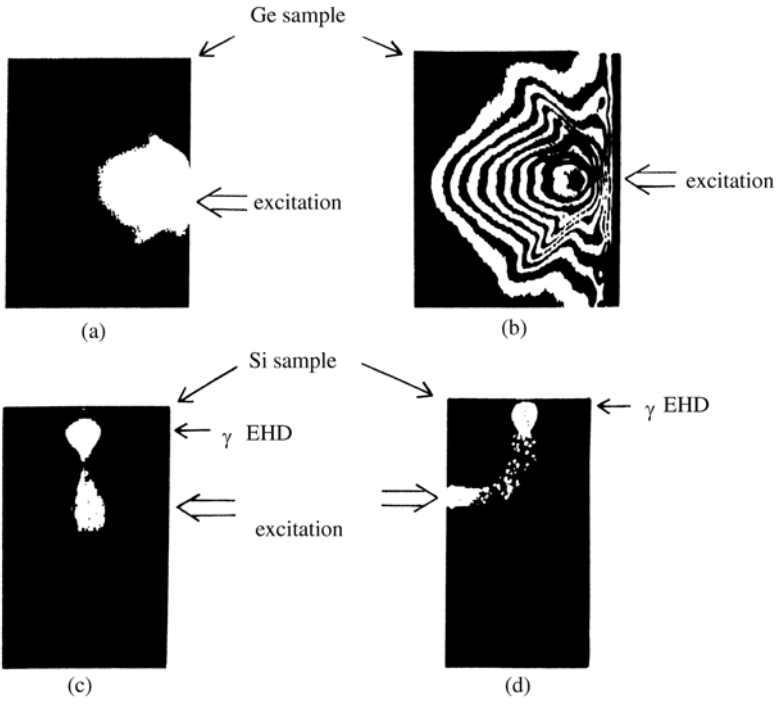


Fig. 21.6. The expansion of a plasma droplet cloud under the influence of the phonon wind (**a**, **b**) and the flow of excitons and plasma droplets into a γ-drop. Front and side views of the sample (**c**, **d**) [77M2, 80G1]

basic idea is the following. The width of the forbidden gap depends on strain as already mentioned above. If inhomogeneous stress is applied to a sample in an appropriate orientation, a situation can be created in which a minimum of E_g appears in the volume of a sample and not at its surface. If the sample is strongly pumped at the surface (e.g., by a blue laser) then the excitons and the droplets move under the influence of the gradient of E_g and accumulate in the potential minimum, forming there a macroscopic drop called a γ-drop. The diameter of such drops can be as large as one millimeter. In Fig. 21.6c,d this situation is seen recorded again by imaging the luminescence of the electron–hole pair recombination.

21.3.2 Electron–Hole Plasmas in Direct-Gap Semiconductors

Many properties of the EHP in direct-gap materials are comparable to those in indirect-gap semiconductors, for example, the renormalization of the gap and the transition from a degenerate to a non-degenerate EHP with increasing plasma temperature at constant density or with decreasing density at constant temperature. Calculations in the limit of quasi-equilibrium thermodynamics predict for direct gap materials, too, a first-order phase separation into an EHL. For CdS these calculations yield a value of T_c of 64 K [81K1, 88Z1]

However, there are also characteristic differences between the EHP in direct and indirect gap materials. One is the appearance of strong stimulated emission in a degenerate EHP in semiconductors with a direct, dipole-allowed band-to-band transition. Gain values of up to $10^4\,\mathrm{cm}^{-1}$, i.e., $1\,\mu\mathrm{m}^{-1}$ can be expected. Another difference is the short carrier lifetime, which is in the (sub)-ns regime for a nondegenerate EHP and may be as short as 100 ps or below in the degenerate case.

We now discuss the consequences of these two differences and other aspects of the EHP in direct gap semiconductors. The EHP has been observed in many direct gap III–V, II–VI and I–VII compounds [81K1, 82F1, 85H2, 90K1, 90K2, 93D1, 93P1, 94I1, 94K2, 95G1, 02N1, 04C1, 04K2, 04P1]. We concentrate here on CdS and GaAs for bulk materials.

The gain in the degenerate EHP leads to stimulated emission which distorts the luminescence spectra so that their evaluation becomes very difficult. Only if extremely small volumes are excited with a few μm length in all three dimensions is it possible to observe sometimes the spontaneous emission of the EHP. In Fig. 21.7 we give an example where an excitation spot of a few $\mu\mathrm{m}^2$ has been used and even there some laser modes start to show up. We stress that the spectral width of the emission increases with increasing pump intensity, in contrast to the situation in Si or Ge below T_c.

In order to overcome the problem of stimulated emission, the gain spectra have been investigated directly. The experimental technique is pump-and-probe beam spectroscopy. One measures the transmission or reflection spectra with a weak, spectrally broad probe beam, once without and once with a spectrally narrow, intense pump beam on. The difference between the two spec-

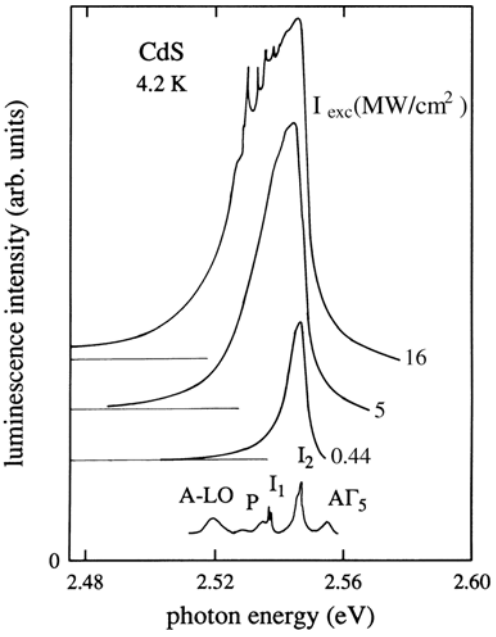


Fig. 21.7. The plasma emission of CdS for increasing excitation using an extremely small excited volume [75L1]

tra gives information about optical nonlinearities and renormalization effects. With pulses of several ns duration one observes a quasi-stationary situation, while measurements with ps excitation allow the investigation of the decay dynamics of the EHP. A tuning of the pump energy $\hbar\omega_{\mathrm{exc}}$ can be used for excitation spectroscopy of the plasma gain, and even spatially resolved pump and probe beam experiments are possible [81B1,84K2,85M1]. For more details see Chap. 25.

We discuss now in connection with Fig. 21.8 the changes of the absorption spectra which we expect in a direct gap semiconductor when we go from a low-density exciton gas to an EHP. Without excitation we observe in absorption (or reflection) the series of exciton resonances and the continuum transitions above E_{g}. In the presence of a plasma the exciton resonances are gone, the forbidden gap $E'_{\mathrm{g}}(n_{\mathrm{P}})$ is renormalized and the states are filled in the degenerate case up to the chemical potential $\mu(n_{\mathrm{P}}, T_{\mathrm{P}})$. As will be explained in more detail below, we expect optical gain between E'_{g} and μ and absorption due to band-to-band transitions above. For a detailed analysis of the gain spectra we consider the recombination rate $r(\hbar\omega)$ at a certain photon energy $\hbar\omega$. Since in direct gap semiconductors no momentum-conserving phonon is involved, we have to consider here recombination under energy and momentum conservation in contrast to the indirect semiconductors. For the recombination rate, which contributes to the emission at a certain photon energy $\hbar\omega$ we obtain

$$r\left(\hbar\omega\right) \propto f_{\mathrm{e}} f_{\mathrm{h}} \left(1+N\right) - \left(f_{\mathrm{h}}-1\right)\left(f_{\mathrm{e}}-1\right) N\,. \tag{21.14}$$

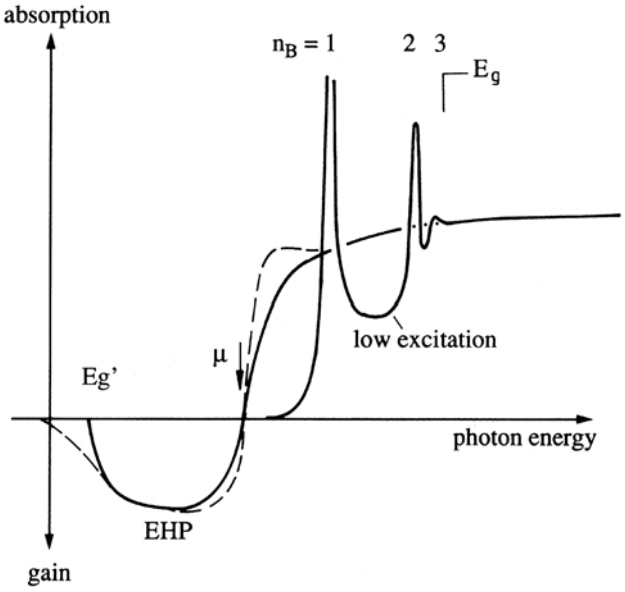


Fig. 21.8. Schematic drawing of the low temperature absorption spectra of a direct-gap semiconductor at low excitation and in the presence of an electron–hole plasma

The first term describes spontaneous and stimulated emission, where N is the density of photons in the (laser) mode with frequency ω under consideration. The second term describes the reabsorption.

Equation (21.14) can be rewritten as a term independent of N giving the spontaneous emission and a term linear in N giving the net rate of stimulated emission and of reabsorption

$$r \propto f_\mathrm{e} f_\mathrm{h} - N\left(1 - f_\mathrm{e} - f_\mathrm{h}\right) \tag{21.15a}$$

$$\propto r_\mathrm{spont} + r_\mathrm{stim}\,. \tag{21.15b}$$

Net gain evidently results for $f_\mathrm{e} + f_\mathrm{h} > 1$ and this simply means

$$\mu\left(n_\mathrm{P}, T_\mathrm{P}\right) > E'_\mathrm{g}\left(n_\mathrm{P}\right) \tag{21.16}$$

as mentioned above.

Using now the square-root dependence of the combined density of states (Sects. 8.9 and 13.1) and also momentum conservation, we get for the spectra of spontaneous emission $I_\mathrm{lum}(\hbar\omega)$ and of gain $g(\hbar\omega)$ for $\hbar\omega \geq E'_\mathrm{g}(n_\mathrm{P})$

$$I_\mathrm{lum}\left(\hbar\omega\right) = c_2\left[\hbar\omega - E'_\mathrm{g}\left(n_\mathrm{P}\right)\right]^{1/2} f_\mathrm{e} f_\mathrm{h}\,, \tag{21.17a}$$

$$g\left(\hbar\omega\right) = c_2\left[\hbar\omega - E'_\mathrm{g}\left(n_\mathrm{P}\right)\right]^{1/2}\left(f_\mathrm{e} + f_\mathrm{h} - 1\right)\,, \tag{21.17b}$$

where $\hbar\omega$ is given by

$$\hbar\omega = E'_{\mathrm{g}}\,(n_{\mathrm{P}}) + \frac{\hbar^2 \boldsymbol{k}_{\mathrm{e}}^2}{2m_{\mathrm{e}}} + \frac{\hbar^2 \boldsymbol{k}_{\mathrm{h}}^2}{2m_{\mathrm{h}}} \tag{21.18}$$

with $\boldsymbol{k}_{\mathrm{e}} = -\boldsymbol{k}_{\mathrm{h}}$, and c_2 is a proportionality constant including once more the transition matrix element, which is assumed to be $\boldsymbol{k}$-independent. The expression (21.17) leads to a square-root increase of the gain spectra at the low energy side and a transition to absorption at $\hbar\omega = \mu$ see Fig. 21.8.

Some modifications of the simple shape of the spectra given in (21.17) are necessary. One is the so-called Landsberg or final-state damping explained on the right of Fig. 21.9. If an electron recombines with a hole there is an empty state in the Fermi sea of electrons in the conduction band and an occupied state in the sea of holes in the valence band. Both states relax very rapidly as indicated by the dashed arrows. This short lifetime enters as an energy-dependent broadening which is also known as Landsberg damping. Consequently the gain spectrum has to be convoluted with a Lorentzian. The parameter Γ in this Lorentzian depends on energy [81K1]. It has a minimum at the quasi Fermi energies and increases below and above roughly linearly

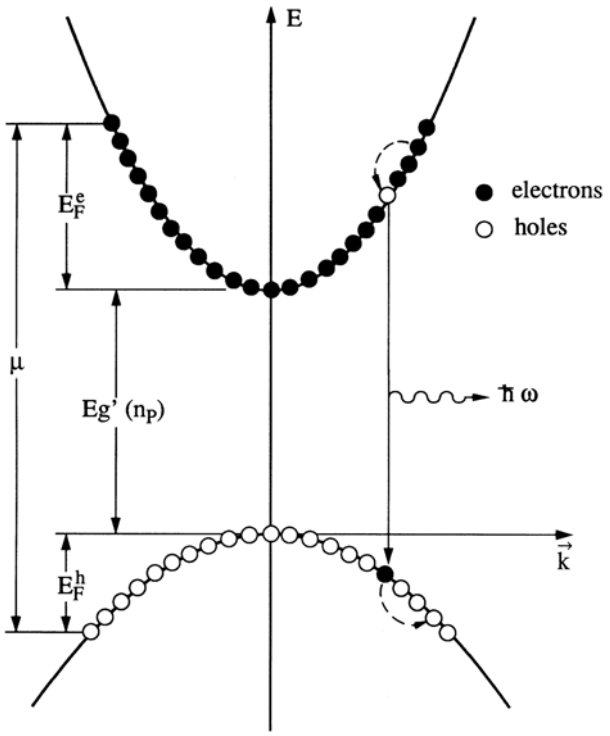


Fig. 21.9. Schematic explanation of the recombination and of the final state damping in a plasma

with energy; see Fig. 21.10a. This final-state damping is the reason why the gain and luminescence spectra do not actually start at E'_g with a square-root dependence but with a smooth tail extending below E'_g as indicated in Fig. 21.8. Further contributions to the broadening include the scattering with other quasiparticles such as carriers or plamsons. This phenomenon is also known as "Fermi-sea shake up". It has been stressed recently in [93S1] and references therein and is a "remainder" of or analogy to the inelastic scattering processes discussed in the previous chapter.

The other correction is the so-called excitonic enhancement [81K1, 88Z1] which is in metals also known as Fermi-edge singularity or Mahan exciton. See [88Z1, 89K1, 93C1] and references therein. It describes an enhancement of the oscillator strength around μ as a multiplicative term $\rho(\hbar\omega)$ to the gain, absorption or luminescence spectra Fig. 21.10b. The physical origin is the following. Around the quasi Fermi levels there are close-lying occupied and empty states. Electrons and holes in the EHP can therefore perform around μ under their mutual Coulomb attraction at least pieces of a Bohr-orbit-like motion. This correlation effect is reminiscent of the exciton and enchances the oscillator strength. In Fig. 21.10 we give as examples the $\boldsymbol{k}$-dependence of the damping for the conduction and valence bands in GaAs and the excitonic enhancement $\rho(\hbar\omega)$. More recent discussions of the optical properties of an EHP can be found, e.g., in [98M1, 98M2, 98P1, 98S1, 04C1, 04K2]. Especially it has been shown in [04C1, 04K2] that an exciton-like peak can appear in a (dilute) EHP at the spectral position of the exciton due to the e-h Coulomb correlation. See also the above information about the role of the excitonic enhancement in the transmission spectrum.

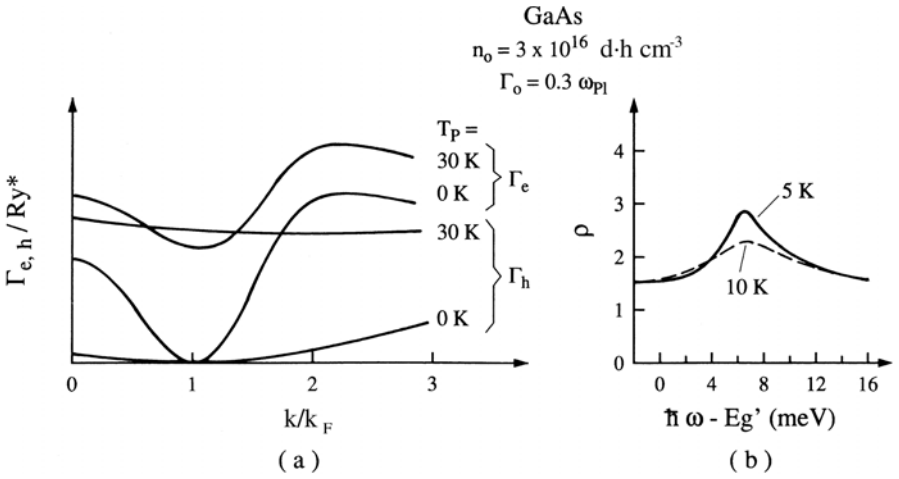


Fig. 21.10. The damping in the conduction and valence bands of GaAs in the presence of a plasma (**a**) and the excitonic enhancement (**b**). From [81K1]

After these general considerations we want to see some examples. In Fig. 21.11 we show transmission spectra of a CdS single crystal platelet with a thickness of a few μm without and with additional excitation in the band-to-band transition region.

The plane parallel surfaces of the sample give rise to Fabry–Perot modes in the transparent spectral regions. At low temperatures we see without excitation the absorption dip of the $A\Gamma_5$ exciton (cf. Fig. 13.9), the transparent window above, and the onset of the $B\Gamma_5$ exciton resonance. For higher lattice temperatures the absorption becomes less steep and shifts to the red according to the Urbach rule of (13.12) or Fig. 13.10. With pump, we see that the exciton resonance disappears at low temperatures. This fact is also confirmed in the reflection spectra not shown here [81B1, 85M1]. In addition we clearly see optical amplification. In Fig. 21.11a the chemical potential μ coinciding with the crossover from gain to absorption is situated around 2.54 eV and the reduced gap is at 2.49 eV compared to 2.58 eV in the unexcited sample. For higher temperatures, population inversion is no longer reached (see the curve for $T_L = 300\,K$) because of the temperature dependence of the effective density of states (8.41) which gives the onset of degeneracy. Nevertheless it is still possible to bleach the tail of the absorption.

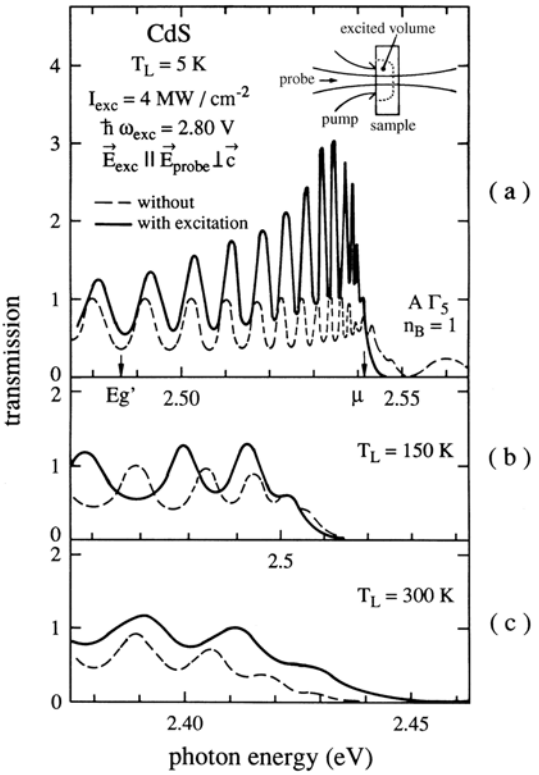


Fig. 21.11. Transmission spectra of a CdS platelet-type sample without (- - -) and with (——) additional excitation and for various lattice temperatures [88S1]

In addition we see a blue-shift of the Fabry–Perot modes at all temperatures. This means that the refractive index decreases. This finding is easily intelligible with the help of the Kramers–Kronig relations of Chap. 6, if we consider that the exciton resonance disappears from the absorption spectrum. For more recent details of this phenomenon see, e.g., [82K1, 83S1].

From raw data as in Fig. 21.11 it is possible to deduce the gain spectra. Examples are given in Fig. 21.12 for CdS for different values of I_{exc} and for GaAs. The fit including the above mentioned phenomena gives n_{P}, T_{P}, $E'_{\mathrm{g}}(n_{\mathrm{P}})$ and $\mu(n_{\mathrm{P}}, T_{\mathrm{P}})$. The agreement between calculated data for these quantities and experiment is shown in Fig. 21.12b. Figure 21.12c gives finally the gain spectrum of a 1 μm thick epitaxial GaAs layer. The energies of the exciton and of the gap without excitation are given and the data for E'_{g} and μ under excitation. The density reached here is considerable higher than in bulk GaAs because of the confirment of the EHP to a 1 μm thick layer. We now list several typical features of the EHP observed in many direct gap semiconductors, but without giving figures or detailed references in every case. They may be found in [78B1, 80B1, 81K1, 82D1, 84H1, 85B1, 88Z1, 89K4, 89K5, 90H1, 90K2, 94K2,

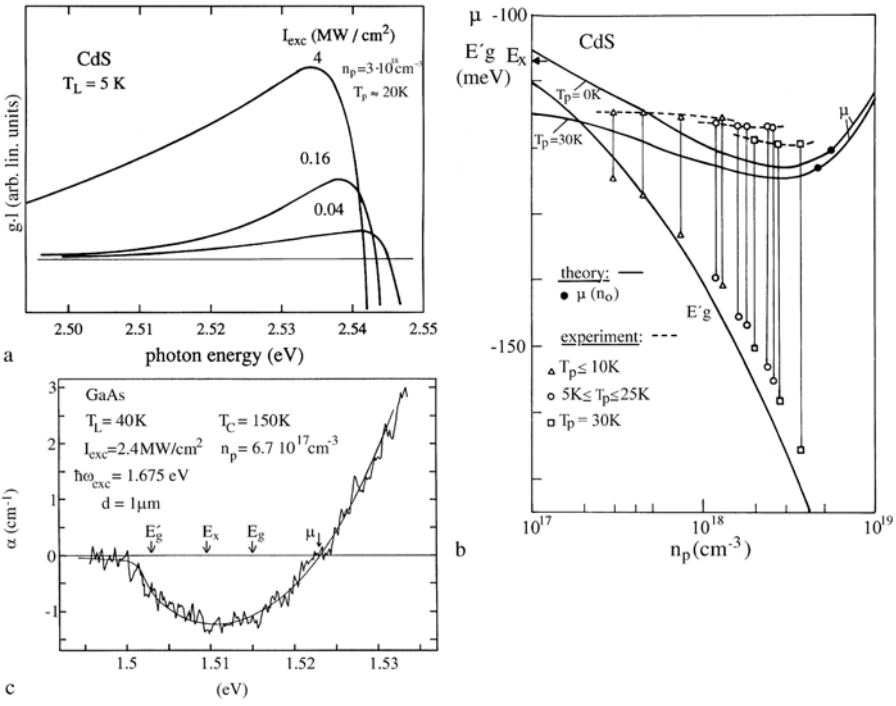


Fig. 21.12. Gain spectra of CdS at low temperatures for various pump intensities (**a**) the reduced gap $E'_{\mathrm{g}}(n_{\mathrm{P}})$ and chemical potential $\mu(n_{\mathrm{P}}, T_{\mathrm{P}})$ deduced from spectra like in (**a**), (**b**) and a gain spectrum for a thin GaAs epitaxial layer (**c**) [81B1,80B1, 89K4]. Note the different directions of positive gain in (**a**) and (**c**)

02N2, 02N3]. The width of the gain spectra increases under quasi-stationary nanosecond excitation with the generation rate as long as there is no pinning of the quasi Fermi levels by stimulated emission, and decreases in time after picosecond excitation. This means that the density is not constant, although T_P is clearly below T_c e.g. for CdS. In Fig. 21.12 the values for the highest excitation are $n_P = 3.10^{18}\,cm^2$; $T_P = 20\,K$ ($T_L = 7\,K, T_c = 64\,K$). Evidently no liquid phase is formed. The reason is the short lifetime of the carriers of about 100 to 200 ps in a degenerate EHP [81B1]. This time is not long enough for the formation of a spatial phase separation into liquid like droplets and a gas phase. There will be strong spatio–temporal density fluctuations and in some regions (e.g., in the vicinity of fast recombination centers in the volume or at the surface) the density will stay lower than in others, possibly even below the Mott density. No clear evidence of a phase separation has been observed in a binary direct gap semiconductor up to now. A pinning of n_P with increasing excitation and spatial density inhomogeneities are often results of stimulated emission and not of a phase separation. However, recently EHL formation has been claimed in the direct gap semiconductor CuCl, based among others on the dynamics of luminescence and IR-reflection spectra [02N1].

The plasma temperature T_P usually lies above the lattice temperature T_L. The difference is only a few ten K in the more ionic II–VI compounds with good coupling of the carriers to the lattice. However, after pulsed excitation the difference can reach several tens of K up to around a hundred K in a first transient cooling period. In the more covalent III–V compounds the differences between T_P and T_L are generally higher. This temperature difference also increases with the excess energy of the excitation, i.e., with $\hbar\omega_{exc} - E_{ex}$. A rather cool electron–hole plasma can consequently be created under resonant excitation in the exciton region or even slightly below.

Since the absorption coefficient is rather high in direct gap semiconductors ($10^4 - 10^5\,cm^{-1}$) the penetration depth of the pump light is $\lesssim 1\,\mu m$. Consequently a strongly inhomogeneous spatial density distribution is created in thicker bulk samples. The gradient of the chemical potential (or of the density) causes an expansion of the plasma. Typical values of the drift or diffusion length l_D of a degenerate EHP, e.g., in CdS or CdSe, are $l_D \approx (10 \pm 5)\mu m$, while the values for the nondegenerate case are closer to $1\,\mu m$ [81B1, 84K2, 88S1, 89R1]. Depending on the experimental conditions one therefore often has to consider an inhomogeneous density distribution in the evaluation of luminescence or gain spectra. In quasi-stationary pump and probe beam experiments it is therefore advisable to probe only in the spatial and temporal center of the pump pulse and to use thin ($\lesssim 10\,\mu m$) samples. In narrow gap materials such as InSb values of l_D up to $60\,\mu m$ have been reported due to the longer carrier lifetimes and the smaller effective masses in this group of materials [85H1].

In Fig. 21.2 we compare finally the calculated, normalized band-gap renormalization as a function of r_s [see (21.5)] with experimental data. As already mentioned, good agreement is found for the mainly covalently bound direct

and indirect gap semiconductors like Si, Ge, GaAs, GaP etc. For clarity we do not give data points for them here. In the case of several populated (in-) equivalent minima, the distribution of the carriers over the various minima has to be considered explicitly [82F2,88K1,92K2,96K1]. Beautiful experiments along these lines exist for $Al_{1-y}Ga_yAs$ in the transition region from a direct to an indirect band structure around $y = 0.43$ [88K1, 92G1, 92K2, 98D1].

The data for various II–VI semiconductors are shown explictly in Fig. 21.2. They follow the trend of the theory, but with an average shift of about 0.5 Ry^* to lower energies. This discrepancy has two main causes. The ionic part of the binding in the II–VI stabilizes the plasma, e.g., via plasma-phonon mixed states. This contribution is not contained in the general formula. Additionally the values of Ry^* and of the experimentally determined exciton binding energies do not exactly coincide, as was mentioned in Sects. 9.2 and 20.3.1 in connection with the biexciton. Similar arguments also hold for a_B which appears in r_s.

In mixed crystals such as $CdS_{1-x}Se_x$ and $Al_{1-y}Ga_yAs$ it has been found that the many-particle effects are strongly reduced as long as the electron–hole pairs occupy only the localized states. One finds essentially a bleaching of the absorption tail of the localized states with increasing density, but not much of a band-gap renormalization [87M1]. If the extended states are populated, however, the properties of the band renormalization and band filling are very similar to those of the binary compounds CdS, CdSe and GaAs [85C1,87M1,89K1,96K1]. The population of the extended states occurs e.g. under sufficiently strong pumping so that the total density of electron–hole pairs exceeds the density of localized states. N_{loc}. This situation can be achieved quite easily in $Al_{1-y}Ga_yAs$ but not in $CdS_{1-x}Se_x$ where N_{loc} is between 10^{18} and $10^{19}\,cm^{-3}$. In the latter case, however, the extended states

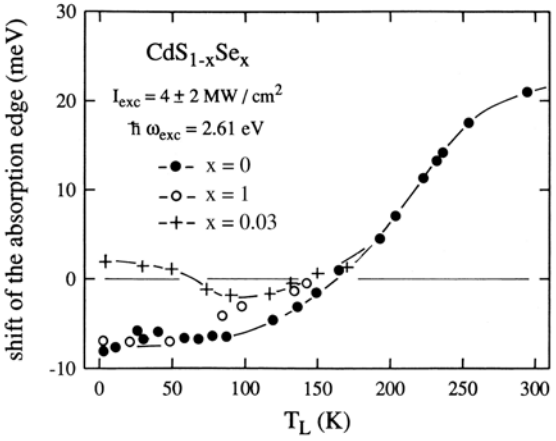


Fig. 21.13. The shift of the absorption edge in CdS, CdSe, and $CdS_{1-x}Se_x$ under high excitation as a function of temperature [89K1]

are also populated to a considerable degree for $T_P \gtrsim T_L > \varepsilon_{loc}/k_B$ where ε_{loc} is the tailing parameter introduced in (14.7). In Fig. 21.13 we show the shift of the absorption edge under high excitation for CdS (see also Fig. 21.11), CdSe, and $CdS_{1-x}Se_x$,. In $CdS_{1-x}Se_x$ one finds at low temperatures a blue shift, due to state filling without renormalization, but a higher temperatures a behavior similar to that for the EHP in the binary compounds.

For the observation of plasmons or plasmon-phonon mixed states in unipolar plasmas created by high doping levels see e.g. Fig. 12.1 and 12.4 or [97G2], for bipolar plasmas Fig. 12.3 or for more recent work in Si, C, GaAs, GaN or CuCl [00N1, 02N2, 02N3, 02N4, 02S1]. As a short preview to Chap. 23, we mention that the formation time of Coulomb screening and band-gap renormalisation has been found to be a few ten fs [01H1].

21.4 Electron–Hole Plasma in Structures of Reduced Dimensionality

In this section, we look at the electron–hole plasma in systems of reduced dimensionality. For recent reviews of this topic see [90C1, 90H1, 94C1, 94K1, 94K2, 94K3, 96K1, 01L1, 04C1, 04K1, 04O1] and references therein.

In Fig. 21.14a we show pump and probe spectra of an $Al_{1-y}Ga_yAs/GaAs$ MQW sample, the GaAs substrate of which has been removed by a selective etching technique. Without excitation we see the now familiar features of the $n_z = 1$ and $2hh$ and lh exciton resonances. With increasing pump power, again using quasi-stationary ns excitation, a scenario develops similar to that for the three-dimensional system: The exciton resonances disappear, and there is population inversion between the reduced gap and the chemical potential. The weak modulation in the resulting gain spectra comes from residual Fabry–Perot modes of the structure. The slope of the transition from gain to absorption gives the plasma temperature if the excitonic enhancement is also properly taken into account. The approximately constant gain value reflects the constant density of states for effective mass particles in a two-dimensional system (see Fig. 8.20). In agreement with the statements above, T_P is higher than in the more ionic II–VI compounds. In Fig. 21.14b we show the density dependence of E'_g and of μ and a calculation of E'_g for a strictly two-dimensional system using some effective parameters of the excitonic Rydberg energy to account for the finite well thickness l_z. The band gap renormalization shown in Fig. 21.14b has been confirmed for lower densities (0.5 to $2 \times 10^{11}\,cm^{-2}$) in [01I1]

Looking closer, we find some differences that are characteristic of the two-dimensional system. The $n_z = 2hh$ resonance is still seen quite clearly at densities where the $n_z = 1hh$ feature has already vanished. The screening of the Coulomb interaction by the free carriers would act on occupied and empty states in the same way. Consequently we may conclude that phase-space filling, which depends on the occupation of the states, is more important in quasi-2d

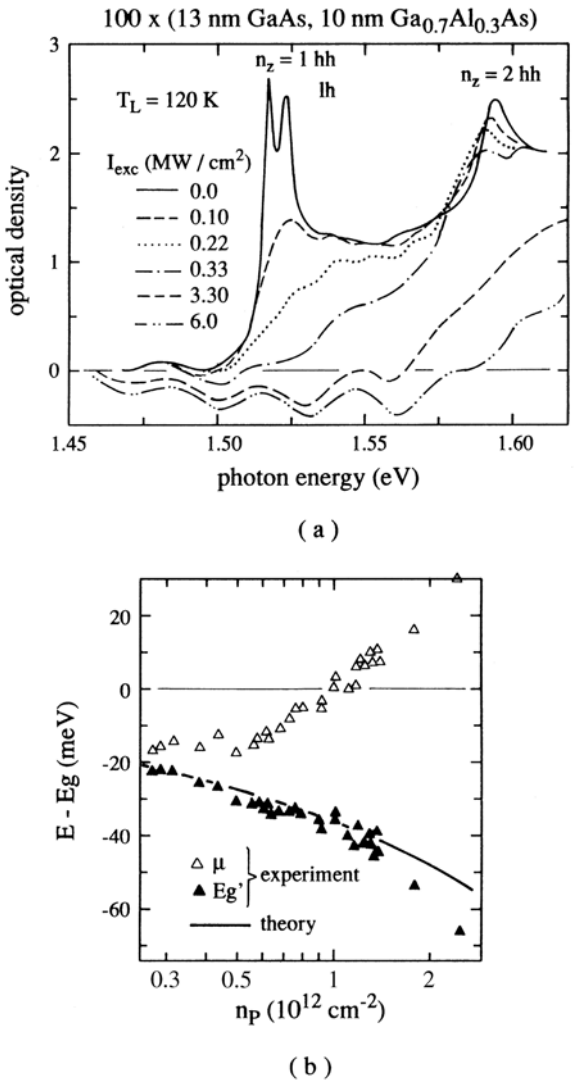


Fig. 21.14. Transmission spectra of an $Al_{1-y}Ga_yAs/GaAs$ MQW sample for various pump intensities (**a**) and the resulting dependence of the reduced gap E'_g (**b**) and of the chemical potential μ on the electron–hole pair density per unit area n_P [89K4, 91S1]

systems than in 3d ones. This interpretation is confirmed by the fact that the $n_z = 2hh$ resonance disappears only when the chemical potential comes close to it, i.e., when the occupation of the higher subbands starts.

Qualitatively this finding can be understood as follows: The Coulomb interaction between electron and hole can be screened in a 3d system in all

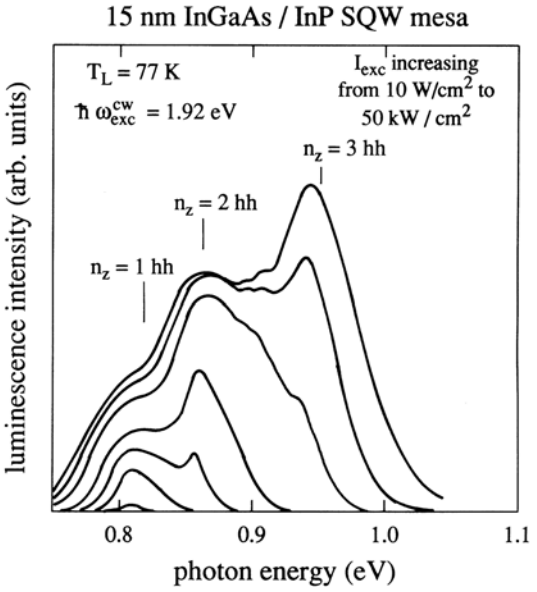


Fig. 21.15. Luminescence spectra of small mesas of InGaAs MQW samples [89K2]

directions of space, whereas the screening in a QW affects preferenentially the field lines in the well. It is more difficult to screen the field lines which go through the barrier since here the probability of finding carriers is reduced to the exponential tails of wave function shown in Figs. 8.22 and 23.

The other point concerns the fact that the renormalization of the higher subbands ($n_z \geqslant 2$) in a QW system is smaller than that of the fundamental gap, as long as the higher subbands are not populated. One reason for this finding [88C1, 89K2, 89K3, 89K4, 91S1] is that electrons (and holes) in different subbands are no longer subject to the exchange interaction, which forms an important contribution to the renormalization of the gap, as outlined in Sect. 21.2.

In Fig. 21.15 we give as a further example for the above phenomena luminescence spectra of an $In_{1-y}Ga_yAs$ MQW sample. In this case the plasma has been confined laterally by the formation of mesa-structures of about $10\,\mu m \times 10\,\mu m$, resulting in a very homogeneous excitation. The band renormalization and the successive filling of states and of higher subbands with increasing pump power are clearly visible even under cw excitation. The band gap renormalization in short period type II GaAs/AlAs superlattices (see Sect. 15.2) has been investigated in [95L1, 96K1]. Due to the redistribution of the carriers between the various band extrema, even a negative, differential band gap renormalization occurs.

The influence of a unipolar plasma on the optical spectra of an $In_{1-y}Ga_yAs$ single quantum well has been presented in [87B1, 88C1]. The two-dimensional

carrier density has been varied by the gate voltage in a MODFET. See also [04C1, 04K2] for a discussion of the luminescence spectra of this material.

Obviously electron–hole plasmas also exist in $Si/Si_{1-x}Ge_x$ quantum wells. In [92X1] a band gap renormalization of about 15 meV has been observed for wider wells, which tend to vanish for narrower wells.

Recent results on the EHP in quasi-one-dimensional quantum wires are compiled in [90C1, 94C1, 96K1, 01L1, 04K1].

In quantum wires, the screening is even less important than in quasi two-dimensional systems. To a good approximation it can be neglected. Calculations of the band gap renormalization (BGR) by exchange and correlation effects can be found, e.g., in [96B1, 97G1, 99P1]. Experiments for, e.g., GaAs- or InAs-based quantum wires are found, e.g., in [95G1, 95G2, 99B1, 99P1, 01L1] and the references given therein.

In quantum dots, it is to some extent a question of semantics to distinguish between (possibly charged) multiexciton complexes (see Sect. 20.3.2) and a plasma. We give here some references for type II Ge/Si dots [02Y1]. References for GaAs- and InAs-based dots are found, e.g., in [95G2, 96B2, 96S1, 98H1, 01L1].

In this section we have mentioned several times the appearance of optical gain or of stimulated emission in an EHP. Actually this phenomenon does not only appear in an EHP but also in connection with other high-excitation processes. We give more information about these cases in the next chapter.

21.5 Inter-subband Transitions in Unipolar and Bipolar Plasmas

In Sect. 13.3 we gave examples for intra-excitonic transitions.

When the excitons disappear in a unipolar or bipolar plasma, there is still the possibility to observe transitions between various subbands both in bulk semiconductors and in structures of reduced dimensionality. These transitions are known as inter-valence band or inter sub-band transitions.

21.5.1 Bulk Semiconductors

In Fig. 21.16 we show the valence band structure of a cubic group IV or binary semiconductor including heavy and light holes and the spin orbit split-off (so) valence band.

The dashed line corresponds to a situation without interaction between the *lh* and the so bands and the solid lines include such an interaction resulting in an avoided crossing (also see Sect. 8.8). This interaction increases with decreasing spin orbit splitting Δ_{so}. The calculation in Fig. 21.6 has been made for InP [99D1] where $\Delta_{so} \approx 100\,\text{meV}$.

If holes are present, e.g., in the *hh* band created either by p-doping or by optical interband excitation of electrons from the valence to the conduction

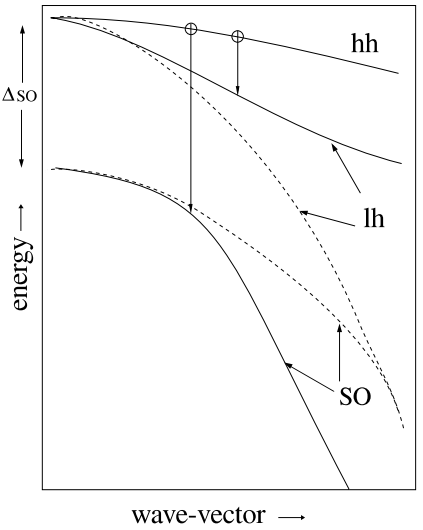


Fig. 21.16. Valence band of a cubic semiconductor with heavy hole (hh), light hole (lh), spin orbit split-off band (so) and transitions of holes [99D1]

band, these holes can be excited to deeper bands by absorption of a photon as shown by the two arrows.

At elevated temperatures, these transitions usually lead to rather broad and unstructured absorption bands. Examples for InP are found, e.g., in [99K1]. Data for other materials including Si or Ge can be found in [02L1].

The electron and hole states can already be quantized for bulk materials by applying a sufficiently strong magnetic field. If electrons and/or holes are created by doping, optical excitation or carrier injection the inter-Landau level transitions or cyclotron transitions can be observed in luminescence or absorption. A few out of the many references in this field are, e.g., [76B1, 83L1, 92U1, 02L1].

21.5.2 Structures of Reduced Dimensionality

In structures of reduced dimensionality the valence and conduction bands are already split without external magnetic field into various subbands. For an early example see [85W1].

The selection rules for inter-subband transitions in a simple conduction band are simple

$$\Delta n_z = \pm 1; \quad \text{for} \quad \boldsymbol{E} \perp \quad \text{well}\,. \tag{21.19}$$

The selection rule between different subbands connects states of opposite parity and the orientation of the electric field results from the orientation of the dipoles normal to the plane of the well as becomes obvious, e.g., by inspecting the wave functions of the first two quantized electron states in Fig. 8.21.

We show various examples in Fig. 21.17. Figure 21.17a gives the absorption of an n-type modulation doped, 6 nm wide GaAs multiple quantum well

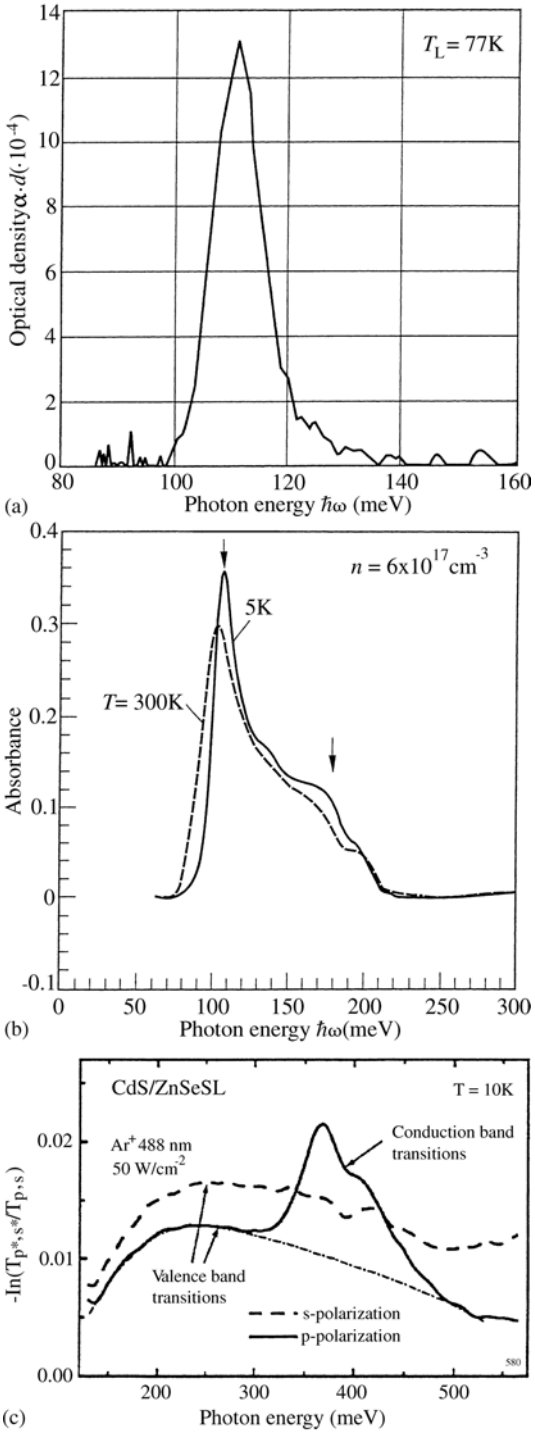


Fig. 21.17. (**a**) The inter-subband absorption of an n-modulation doped $GaAs/Al_{1-y}Ga_yAs$ MQW sample. (**b**) The inter-miniband absorption of a homogenously n doped 200 × (7.5 nm GaAs/2.5 nm $Al_{1-y}Ga_yAs$) superlattice [95T1, 93H1] and (**c**) of an optically pumped 140 × (2 nm CdS/1.9 nm ZnSe) superlattice [00G1, 02G1]

sample. The absorption band for the $n_z = 1 \rightarrow n_z = 2$ transitions is rather narrow. If the two conduction subbands are parallel, the width comes from homogeneous broadening and an additional inhomogeneous contribution caused by interface disorder.

The situation is quite different for superlattices. The miniband dispersion results in a characteristic broadening of the inter-miniband absorption as shown for the conductions bands of a III–V and a II–VI superlattice in Fig. 21.17b, c.

Figure 21.17c is also an example for the verification of the selection rules for conduction subbands. The sample is undoped and optical excitation creates both electrons and holes [00G1, 02G1]. In this reference data are also given for multiple quantum wells and many particle effects and their influence on inter-subband transitions are discussed.

Inter- and intra-subband transitions in GaAs/$Al_{1-y}Ga_yAs$ and in GaN/$Al_{1-y}Ga_yN$ multiple quantum wells are discussed in [98Z1, 03C1] and data on inter-subband transitions in Si/$Si_{1-x}Ge_x$ multiple quantum wells are found in [94F1]. Inter-miniband transitions are reviewed in [95H1].

For quantum wires see, e.g., [00C1]. Further examples and references are given in [01L1].

21.6 Problems

1. Compare the values of the Mott density of CuCl, CdS, and InSb at 0 and at 300 K using the appropriate formulas and material parameters. What are the excitation intensities needed to reach the Mott density under quasi-stationary conditions if you excite with $\hbar\omega_{\mathrm{exc}} \gtrsim E_{\mathrm{g}}$ and assume a diffusion length of the carrier of 1, 10 or 50 µm and a lifetime of 0.1 ps?
2. Develop formulas to deduce the gain spectra from raw data as in Fig. 21.11 assuming that the diffusion length of the plasma is greater than (less than) the thickness of the sample. If you get stuck, try to make reasonable assumptions.
3. How would the gain spectra (21.17) change qualitatively, if you assume a drift of the carriers with a velocity $v = \frac{1}{2}v_{\mathrm{F}}$? Here, v_{F} is the Fermi velocity.
4. Deduce the effective plasma temperature from some of the spectra of Figs. 21.12, 21.14 and 21.15, neglecting excitonic enhancement and final-state damping.

References to Chap. 21

[23D1] P. Debye and E. Hückel, Phys. Z. **24**, 185, 305 (1923)
[27T1] L.H. Thomas, Proc. Camb. Phil. Soc. **23**, 542 (1927)
[28F1] E. Fermi, Z. Phys. **48**, 73 (1928)

[68K1] L.V. Keldyh, Proc. 9th Intern. Conf. Phys. Semiconductors, p. 1303, Moscow (1968)
[74M1] N.F. Mott, Metal-Insulator Transitions, Taylor, Francis, London (1974)
[74T1] G.A. Thomas, T.M. Rice and J.C. Hensel, Phys. Rev. Lett. **33**, 219 (1974)
[75L1] V.G. Lyssenko et al., Sov. Phys. JETP **41**, 163 (1975)
[75R1] T.L. Reinecke and S.C. Ying, Phys. Rev. Lett. **35**, 311 (1975)
[76B1] G. Bauer et al., phys. stat. sol. (b) **75**, 543 (1976)
[77H1] T.C. Hensel, T.G. Phillips and G.A. Thomas, Solid State Phys. **32**, 88 (1977)
[77H2] E. Hanamura and H. Haug, Phys. Rep. **33C**, 209 (1977)
[77M1] R.W. Martin, Solid State Commun. **19**, 373 (1976); ibid. **22**, 523 (1977)
[77M2] R.S. Markiewicz, J.P. Wolfe and C.D. Jeffries, Phys. Rev. B **15**, 1988 (1977)
[77R1] T.M. Rice, Solid State Phys. **32**, 1 (1977)
[78B1] G. Beni and T.M. Rice, Phys. Rev. B **18**, 768 (1978)
[78L1] V.G. Lysenko and V.I. Revenko, Sov. Phys. Sol. State **20**, 1238 (1978)
[79C1] M. Combescot, Solid State Commun. **30**, 81 (1979)
[79M1] H. Maaref et al., J. Lumin. **18/19**, 547 (1979)
[79W1] R.M. Westervelt, J.C. Culbertson and B.S. Black, Phys. Rev. Lett. **42**, 269 (1979)
[80B1] K. Bohnert, G. Schmieder and C. Klingshirn, phys. stat. sol. (b) **98**, 175 (1980)
[80G1] M. Greenstein and J.P. Wolfe, Solid State Commun. **33**, 309 (1980)
[81B1] K. Bohnert et al., Z. Phys. B **42**, 1 (1981)
[81F1] A. Forchel et al., Phys. Rev. Lett. **46**, 678 (1981)
[81K1] C. Klingshirn and H. Haug, Phys. Rep. **70**, 315 (1981)
[81S1] H. Stolz, R. Zimmermann and G. Röpke, phys. stat. sol. (b) **105**, 585 (1981)
[82D1] R. Dai et al., Z. Physik B **46**, 189 (1982)
[82F1] A. Forchel, Ph.D. Thesis, Stuttgart (1982)
[82F2] G.W. Fehrenbach et al., Phys. Rev. Lett. **49**, 1281 (1982)
[82K1] A. Kreissl et al., phys. stat. sol. (b) **114**, 537 (1982)
[82V1] P. Vashista and R.K. Kalia, Phys. Rev. B **25**, 6492 (1982)
[83C1] M. Combescot and C. Benoit à la Guillaume, Solid State Commun. **46**, 579 (1983)
[83L1] G. Lindemann et al., Phys. Rev. B **28**, 4693 (1983)
[83S1] S. Schmitt-Rink et al., Physica B **117/118**, 339 (1983)
[84H1] H. Haug and S. Schmitt-Rink, Prog. Quantum Electron **9**, 3 (1984)
[84K1] S.W. Koch, Dynamics of First-Order Phase Transitions in Equilibrium and Nonequilibrium Systems, Lecture Notes Phys. **207**, Springer, Berlin, Heidelberg (1984)
[84K2] K. Kempf and C. Klingshirn, Solid State Commun. **49**, 23 (1984)
[84M1] G. Manzke, V. May and K. Henneberger, phys. stat. sol. (b) **125**, 693 (1984)
[85B1] D. Bimberg, in Landolt-Börnstein New Series, Group III, Vol. **17i**, 297 (1985)
[85C1] M. Capizzi et al., Helv. Phy. Acta **58**, 272 (1985)
[85E1] Electron–Hole Droplets in Semiconductors, C.D. Jeffries and L.V. Keldysh, (eds.), Mod. Probl. Cond. Mat. Sci., Vol. 6, North-Holland, Amsterdam (1985)

[85H1] D.J. Hagan et al., Opt. Lett. **10**, 187 (1985)
[85H2] D. Hulin et al., J. Luminesc. **30**, 290 (1985)
[85M1] F.A. Majumder et al., Phys. Rev. B **32**, 2407 (1985)
[85W1] L.C. West and S.J. Eglash, Appl. Phys. Lett. **46**, 1156 (1985)
[87B1] I. Bar-Joseph et al., Phys. Rev. Lett. **59**, 1357 (1987)
[87M1] F.A. Majumder et al., Z. Phys. B **66**, 409 (1987)
[88C1] D.S. Chemla et al., IEEE J. QE **24**, 1664 (1988)
[88K1] H. Kalt, A.L. Smirl and T.F. Bogges, phys. stat. sol. (b) **150**, 895 (1988)
[88S1] H.E. Swoboda et al., Z. Phys. B **70**, 341 (1988) and Phys. Rev. B **39**, 11019 (1989)
[88Z1] R. Zimmermann, Many Particle Theory of Highly Excited Semiconductors, Teubner Texte Phys., Teubner, Leipzig (1988)
[88Z2] R. Zimmermann, phys. stat. sol. (b) **146**, 371 (1988)
[89K1] C. Klingshirn et al., NATO ASI Seri. B **194**, 353 (1989)
[89K2] V.D. Kulakovskii et al., Phys. Rev. B **40**, 8087 (1989)
[89K3] H. Kalt et al., Phys. Rev. B **40**, 12017 (1989)
[89K4] C. Klingshirn et al., NATO ASI Ser. B **200**, 167 (1989)
[89K5] C. Klingshirn et al., NATO ASI Ser. B **194**, 353 (1989)
[89R1] M. Rinker et al., Solid State Commun. **69**, 887 (1989)
[90C1] R. Cingolani and K. Ploog, Adv. Phys. **40**, 535 (1990)
[90F1] H. Fieseler, R. Schwabe and J. Staehli, phys. stat. sol. (b) **159**, 411 (1990)
[90H1] H. Haug and S.W. Koch, Quantum Theory of Optical and Electronic Properties of Semiconductors, World Scientific, Singapore (1990)
[90K1] C. Klingshirn, Festkörperprobleme / Adv. Solid State Phys. **30**, 335 (1990)
[90K2] M. Kunz et al., J. Crystal Growth **101**, 734 (1990)
[91S1] K.-H. Schlaad et al., phys. stat. sol. (b) **159**, 173 (1990) and Phys. Rev. B **43**, 4268 (1991)
[92G1] S. Gürtler et al., phys. stat. sol. (b) **173**, 441 (1992)
[92K1] H. Kalt, Festkörperprobleme / Adv. Solid State Phys. **32**, 145 (1992)
[92K2] H. Kalt and M. Rinker, Phys. Rev. B **45**, 1139 (1992)
[92P1] P.P. Paskov, Europhys. Lett. **20**, 143 (1992)
[92U1] K. Unterrainer et al., Semicond. Sci and Technol. **7**, B 604 (1992)
[92X1] X. Xiao et al., Appl. Phys. Lett. **60**, 1720 (1992)
[93C1] M. Combescot and Ch. Tanguy, in Proc. 21st Int. Conf. on Physics of Semiconductors, Beijing 1992, Ping Jiang, Hou-Zhi Zheng eds., World Scientific, Singapore, p. 141 (1993)
[93D1] F. Daminger et al., In Proc. 21st Int. Conf. on Physics of Semiconductors, Beijing 1992, Ping Jiang, Hou-Zhi Zheng eds., World Scientific, Singapore, p. 1293 (1993)
[93H1] M. Helm et al., Phys. Rev. B **48**, 1601 (1993)
[93P1] K. Ploog, R. Nötzel and O. Brandt, In Proc. 21st Int. Conf. on Physics of Semiconductors, Beijing 1992, Ping Jiang, Hou-Zhi Zheng eds., World Scientific, Singapore, p. 1297 (1993)
[93S1] M.S. Skolnick et al., In Proc. 21st Int. Conf. on Physics of Semiconductors, Beijing 1992, Ping Jiang, Hou-Zhi Zheng eds., World Scientific, Singapore, p. 41 (1993)
[94C1] R. Cingolani, Physica Scripta T **49** B, 470 (1994)
[94F1] T. Fromherz et al., Phys. Rev. B **50**, 15073 (1994)
[94I1] M. Illing et al., J. Crystal Growth **138**, 638 (1994)

[94K1] E. Kapon, Semiconductors and Semimetals **40**, 259 (1994)
[94K2] C. Klingshirn, NATO ASI Series B **339**, 327 (1994)
[94K3] H. Kalt, J. Lumin. **60/61**, 262 (1994)
[95G1] M. Grundmann et al., J. Nonlin. Opt. Phys. Mater. **4**, 99 (1995)
[95G2] H.F. Ghaemi et al., Superlattices and Microstructures **17**, 15 (1995)
[95H1] M. Helm, Semicond. Sci. and Technol. **10**, 557 (1995)
[95L1] W. Langbein et al., Phys. Rev. B **51**, 1946 (1995)
[95S1] L.M. Smith and J.P. Wolfe, Phys. Rev. B **51**, 7521 (1995)
[95T1] E. Towe et al., Superlattices and Microstructures **17**, 129 (1995)
[96B1] M. Bayer et al., Phys. Rev. B **55**, 13180 (1996)
[96B2] K. Bollweg et al., Phys. Rev. Lett. **76**, 2774 (1996)
[96K1] H. Kalt, Optical Properties of II–V Semiconductors: The Influence of Multi-Valley Band Structures Springer Series in Solid State Sciences **120**, Springer, Berlin (1996)
[96S1] R. Steffen et al., Surf. Sci **361**, 805 (1996)
[97G1] A. Grindt et al., phys. stat. sol. (b) **202**, 725 (1997)
[97G2] H. Göppert et al., J. Luminesc. **72–74**, 430 (1997)
[98D1] U. Dörr et al., J. Appl. Phys. **83**, 2241 (1998)
[98H1] R. Heitz et al., Phys. Rev. B **57**, 9050 (1998)
[98M1] G. Manzke et al., phys. stat. sol. (b) **206**, 37 (1998)
[98M2] G. Manzke et al., Phys. Rev. Lett. **80**, 4943 (1998)
[98P1] M.F. Pereira, Jr. and K. Henneberger, phys. stat. sol. (b) **206**, 477 (1998)
[98S1] R. Schepe et al., phys. stat. sol. (b) **206**, 273 (1998)
[98Z1] Q.-S. Zhu et al., Phys. Rev. B **57**, 12388 (1998)
[99B1] C.R. Bennet et al., Physica B **263-264**, 546 (1999)
[99D1] A. Dargys, Optics Commun. **170**, 47 (1999)
[99K1] S. Kakimoto and H. Watanabe, J. Appl. Phys. **85**, 1822 (1999)
[99P1] C. Piermarocchi et al., Sol. State Commun. **112**, 433 (1999)
[00C1] S. Calderon et al., Phys. Rev. B **62**, 9935 (2000)
[00G1] M. Göppert et al., J. Crystal Growth **214/215**, 625 (2000) and Physica E **7**, 89 (2000)
[00J1] A. Jolk et al., phys. stat. sol. (b) **221**, 295 (2000)
[00M1] T. Meier et al., Phys. Rev. **62**, 4218 (2000)
[00M2] T. Meier and S.W. Koch, phys. stat. sol. (b) **221**, 211 (2000)
[00N1] M. Nagai and M. Kuwata-Gonokami, phys. stat. sol. b **221**, 261 (2000)
[01I1] H. Ichida et al., Proceedings of 2000 Internat. Conf on Excitonic Processes in Condensed Matter (= EXCON 2000), Osaka, Japan, August 22–25, 2000, K. Cho and A. Matsui eds., World Scientific, Singapore, New Jersey, London, Hong Kong, p. 225 (2001)
[01L1] Landolt-Börnstein New Series, Group III, Vol 34C1, C. Klingshirn, ed., Springer, Berlin (2001)
[01H1] R. Huber et al., Nature **414**, 286 (2001)
[01N1] M. Nagai, R. Shimano and M. Kuwata-Gonokami, Phys. Rev. Lett. **86**, 5795 (2001)
[02G1] M. Göppert et al., Phys. Rev. B **65**, 115334 (2002)
[02I1] T.J. Inagaki and M. Aihara, Phys. Rev. B **65**, 205204 (2002)
[02L1] Landolt-Börnstein New Series, Group III, Vol. 41 A1 B, U. Rössler, ed., Springer, Berlin (2002)
[02N1] M. Nagai and M. Kuwata-Gonokami, J. Luminesc. **100**, 233 (2002)

[02N2] M. Nagai, K. Ohkawa and M. Kuwata-Gonokami, Appl. Phys. Lett. **81**, 484 (2002)
[02N3] M. Nagai and M. Kuwata-Gonokami, J. Phys. Soc. Japan **71**, 2276 (2002)
[02N4] M. Nagai et al., phys. stat. sol. b **238**, 509 (2002)
[02S1] R. Shimano et al., Phys. Rev. Lett. **88**, 057404 (2002)
[02Y1] A.I. Yakimov et al., Physica E **13**, 1026 (2002)
[03C1] G. Chen et al., phys. stat. sol. b **240**, 384 (2003)
[03N1] M. Nagai, phys. stat. solidi B, **238**, 509 (2003)
[04C1] S. Chatterjee et al., Phys. Rev. Lett. **92**, 067402 (2004)
[04K1] E. Kapon in ref. [81A1]l in press
[04K2] S.W. Koch and M. Kira in Ref. [04O1]p 1
[04O1] Optics of Semiconductors and Their Nanostructures, H. Kalt and M. Hetterich eds., Springer Series in Solid State Science **146** (2004)
[04P1] H. Priller et al., phys. stat. sol. b **241**, 587 (2004)

22

Stimulated Emission and Laser Processes

Stimulated emission from semiconductors is usually identified in experiments by the occurrence of one or several of the following criteria: a strongly superlinear increase of the optical luminescence output I_{lum} as a function of the pump power I_{exc} above a certain threshold $I_{\mathrm{exc}}^{\mathrm{th}}$ with slopes α in $I_{\mathrm{lum}} \propto I_{\mathrm{exc}}^{\alpha}$ of three and more; a simultaneous spectral narrowing of the emission, often accompanied by the appearance of laser modes imposed by some cavity length and a spatially directed emission above $I_{\mathrm{exc}}^{\mathrm{th}}$. The increase of the coherence length of the emitted light is also characteristic for laser emission, but is less used to identify laser emission from semiconductors, though semiconductors laser with external, wavelength-selective cavities are developed with extremly narrow spectral emission.

Many of the high excitation effects outlined in Chaps. 19 to 21 can lead to stimulated emission under suitable conditions; see, e.g., [73B1, 75K1, 81K1, 81S1, 92E1, 93K1, 94C1, 95K1, 97N1, 99M1, 03I1, 03R1, 03R2] and references therein. In this chapter we treat those aspects of the high excitation phenomena of Chaps. 19 to 21 which are specific for stimulated emission.

Pumping can be accomplished by excitation with intense (laser-) light [81K1, 94C1], ($I_{\mathrm{exc}} \gtrsim 10^4\,\mathrm{W/cm^2}$ depending on the material and the pump conditions), with electron beams [64B1, 73B1, 92B1, 92N1] ($j \gtrsim 10\,\mathrm{A/cm^2}$, $U \gtrsim 30\,\mathrm{keV}$), by a flash-lamp [83B1] or by a forward biased *pn* junction [81S1, 92E1, 92N1, 04H1], provided that it is possible to grow the material in the desired highly *p*- and *n*-doped versions. Optical pumping is usually the proper choice for scientific investigations of the gain processes, since it allows resonant excitation of some species, e.g., certain exciton or biexciton levels. Pumping by electron beams, on the other hand, is unselective and an energy of about $3\,E_{\mathrm{g}}$ is deposited for the creation of one electron–hole pair. Electron–beam pumping was widely used before high power lasers were available [64B1, 73B1] and more recently for the realization of color projection TV tubes with high brightness [92B1, 92N1].

Pumping by a *pn* (hetero-) junction biased in the forward direction is obviously the best choice for most technical applications in laser diodes, which

are found, for example, in every CD player, laser scanner, laser printer or bar code reader [81S1, 92E1].

Stimulated emission has, until now been almost exclusively limited to direct gap semiconductors. An exception, namely $Al_{1-y}Ga_yAs$ close to the direct-indirect crossover, was mentioned in Sect. 21.3 and in [96K1] and we shall come back to this later. Our considerations are restricted therefore to direct gap materials where not stated otherwise.

22.1 Excitonic Processes

The processes which lead to stimulated emission can either be of intrinsic nature, i.e., they involve only excitations of the perfect lattice, or of extrinsic nature, involving some impurity or defect states. We shall see in the following examples of both groups.

We use the language of the weak coupling limit for reasons of simplicity, but state, that we are actually dealing with polaritons. We start with intrinsic processes in the intermediate density regime, namely the recombination of an exciton with emission of one or more LO phonons, or with inelastic scattering by another exciton or a free carrier, or finally the decay of a biexciton into a photon and an exciton. All these processes were already introduced in Chap. 20. In Fig. 22.1 we summarize these effects again actually drawing them in the strong coupling or polariton picture.

In all four cases some initial state decays under emission of a photon and leaves behind in the crystal another excitation, for example an exciton in its ground state in the case of biexciton decay, an LO phonon in the case of the phonon replica, or an exciton in an excited state $n_B \gtrsim 2$ or a free carrier at higher kinetic energy in the cases of inelastic exciton-exciton or-free carrier scattering, respectively. All these processes have the low threshold for population inversion typical for four-level laser systems as long as this excited final state in the sample is not (thermally) populated. If there are no excitons thermally excited into $n_B = 2$ state, or no thermally populated LO phonons, then population inversion is reached in the presence of two or even one exciton, respectively. See also (22.2c). We consider now the rate equations for these processes, assuming that for a given photon energy $\hbar\omega$ there is just one (laser-) mode containing photons with a density N_{ph}. The general rate equation which holds is [81K1]:

$$\frac{dN_{Ph}}{dt} = -2\kappa N_{Ph} + \sum \frac{2\pi}{\hbar}\delta(\Delta E)\,|W|^2\,Q\,, \tag{22.1}$$

where $-2\kappa N_{Ph}$ contains all losses of the resonator, e.g., due to finite reflection or diffraction. The second term on the right of (22.1) contains a sum over all (scattering) processes contributing to the emission at $\hbar\omega$,$\delta(\Delta E)$ stands for the energy conservation, Q is a population factor into which we can integrate the $\boldsymbol{k}$-conservation, and $|W|^2$ finally is the transition matrix element. As a first

approach to the strong coupling limit one can also incorporate into $|W|^2$ the probabilities that the exciton polariton is exciton-like in the initial state and photon-like in the final state. The next step would involve the use of the polariton dispersion curves of Fig. 22.1 instead that of excitons and photons. See [81K1].

We consider here as an example the P_2 line, i.e., the inelastic scattering between two excitons in the $n_\mathrm{B} = 1$ state, from which one is scattered under energy and momentum conservation into a state with $n_\mathrm{B} = 2$, while the other one appears as a photon. In the strong coupling limit one would state, that there occours an inelastic polariton-polariton scattering process between two exciton polaritons, from which one is scattered onto the photonlike (lower) polariton branch and the other into an excited exciton-like polariton state.

In this case we have

$$\begin{aligned}\Delta E = {} & E_\mathrm{ex}\,(n_\mathrm{B} = 1, \boldsymbol{k}_1) + E_\mathrm{ex}\,(n_\mathrm{B} = 1, \boldsymbol{k}_2) \\ & -\hbar\omega - E_\mathrm{ex}\,(n_\mathrm{B} = 2, \boldsymbol{k}_1 + \boldsymbol{k}_2)\ ,\end{aligned} \tag{22.2a}$$

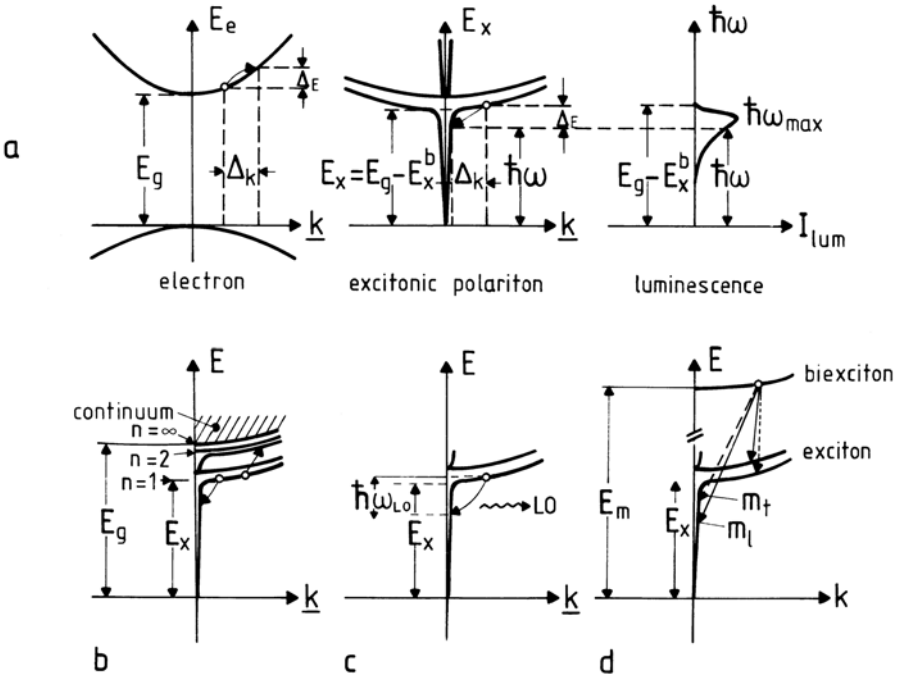


Fig. 22.1. Inelastic scattering processes in the intermediate density regime: exciton-LO phonon emission (**c**), inelastic exciton–exciton scattering (**b**), inelastic exciton–electron scattering (**a**) and biexciton decay (**d**) [81K1]

and the population factor Q reads

$$\begin{aligned} Q = & N_{\mathrm{ex}}\left(\boldsymbol{k}_1, n_{\mathrm{B}}=1\right) \times N_{\mathrm{ex}}\left(\boldsymbol{k}_2, n_{\mathrm{B}}=1\right) \\ & \times\left[1+N_{\mathrm{ex}}\left(\boldsymbol{k}_1+\boldsymbol{k}_2, n_{\mathrm{B}}=2\right)\right]\left(1+N_{\mathrm{Ph}}\right) \\ & -N_{\mathrm{Ph}} N_{\mathrm{ex}}\left(\boldsymbol{k}_1+\boldsymbol{k}_2, n_{\mathrm{B}}=2\right)\left[1+N_{\mathrm{ex}}\left(\boldsymbol{k}_1, n_{\mathrm{B}}=1\right)\right] \\ & \times\left[1+N_{\mathrm{ex}}\left(\boldsymbol{k}_2, n_{\mathrm{B}}=1\right)\right] , \end{aligned} \tag{22.2b}$$

where we consider photons and excitons as bosons.

We can now decompose Q into two terms. One which is independent of N_{Ph} and describes the spontaneous emission and another one which is linear in N_{Ph} and gives the net gain or absorption. For the latter, Q_{stim}, we find

$$\begin{aligned} Q_{\mathrm{stim}} = & N_{\mathrm{Ph}}\left\{N_{\mathrm{ex}}\left(\boldsymbol{k}_1, n_{\mathrm{B}}=1\right) N_{\mathrm{ex}}\left(\boldsymbol{k}_2, n_{\mathrm{B}}=1\right)\right. \\ & -N_{\mathrm{ex}}\left(\boldsymbol{k}_1 \boldsymbol{k}_2, n_{\mathrm{B}}=2\right) \\ & \left.\times\left[1+N_{\mathrm{ex}}\left(\boldsymbol{k}_1, n_{\mathrm{B}}=1\right)+N_{\mathrm{ex}}\left(\boldsymbol{k}_2, n_{\mathrm{B}}=1\right)\right]\right\} . \end{aligned} \tag{22.2c}$$

If we assume thermal equilibrium between the excitons in the various states, we see that inversion occurs, i.e., $Q_{\mathrm{stim}} > 0$ at low temperatures, even if we have only two excitons which collide as mentioned already above.

The luminescence and gain increase roughly quadratically with the exciton density and thus superlinearly with the pump intensity at low temperature until the gain overcomes the losses and stimulated emission sets in. At higher temperatures the $n_{\mathrm{B}} = 2$ states will be also populated, reabsorption occurs and may even overcome the gain in some spectral regions, resulting in excitation-induced absorption.

The biexciton decay and the exciton-free carrier scattering also have gain values which increase superlinearly with density, while the gain of the ex-n LO process grows essentially linearly with the generation rate until laser emission sets in.

If Fig. 22.2 we show the calculated temperature dependences of the laser threshold densities for three of the above-mentioned processes for a given constant value of κ. A variation of κ will shift the various quadratic processes with respect to the linear one. The increase of the thresholds with increasing temperature comes from the thermal population of the final states as indicated above. The high threshold at low temperatures of the process involving free carriers originates from the assumption of thermal equilibrium. In this case no free carriers (i.e., excitons in the continuum state) are present at low temperatures. The exciton-free carrier-scattering process can be influenced by doping. The fact that the calculated carrier densities are relatively high and may exceed the Mott density (Chap. 21) is due to the use of rather high losses in the calculations. Lower loss values reduce the calculated threshold densities.

On the other hand this finding shows that stimulated emission occurs frequently in the transition region between intermediate and high densities. The

exciton binding energy decreases with increasing density, i.e., the excitons become "soft" and inelastic scattering processes between carriers may take place in an EHP or the emission of phonons or plasmons in the recombination process of an electron–hole pair. While scientists tried to separate these various processes in the 1970s and 1980s it is now, at least from the theoretical side, the aim to present calculations that are valid continuously from low densities to the EHP as already mentioned in Chap. 21 where some references are also given. A technical term to describe these effects is gain and loss in a "strongly Coulomb correlated electron–hole plasma". For some more experimental and theoretical references see, apart from the references given in Chap. 21, e.g. [96G1, 96H1, 96K1, 97P1, 98M1, 98P1].

For the purpose of didactics and clarity, we still separate these processes here, but one should be aware of the continuous transition from the intermediate to the high density regime, especially in direct-gap semiconductors.

In Fig. 22.2 we give experimental data for the excitation intensity at threshold for CdS which show the trends predicted by theory and dicussed above. The low value of the threshold around 80 K comes from a cooperative effect, since various processes are spectrally degenerate in CdS at this temperature; see Fig. 20.3. Similar data are known for other semiconductors like ZnO. See [81K1] and reference therein. For more recent data on GaN and ZnO see Sect. 22.3.

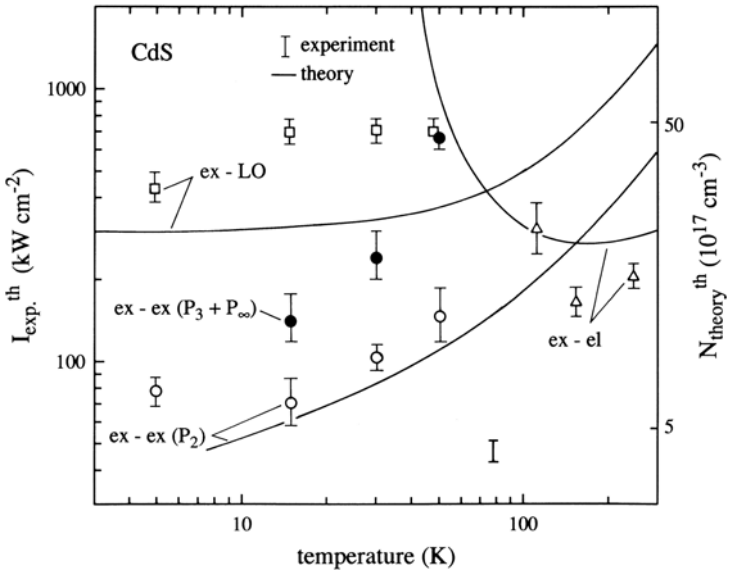


Fig. 22.2. The calculated threshold density $N^{\rm th}_{\rm theory}$ for various gain processes in the intermediate density regime of CdS as a function of temperature, and the experimentally observed excitation intensity at threshold $I^{\rm th}_{\rm exc}$ as a function of the lattice temperature [81K1]

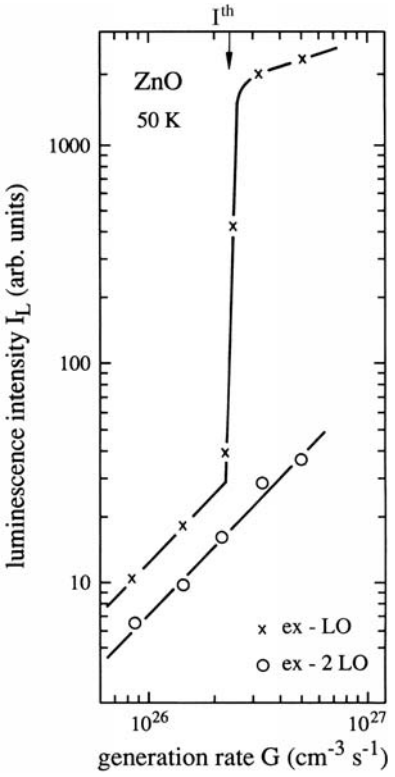


Fig. 22.3. The luminescence intensity of the first two LO phonon replicas of ZnO as a function of the generation rate G, showing the laser threshold [75K1, 80W1]

In Fig. 22.3 we show an example of the optical input-output curves, showing clearly the laser threshold. As an example, the exciton-LO process in ZnO has been chosen [80W1].

Similar input–output curves are obtained for the light output as a function of the forward current (density) in laser diodes. Often one defines a differential or slope efficiency which is the (maximum) of the derivative of curves like Fig. 22.3. It must be mentioned that theses slope efficencies may raech unit but they say nothing about the absolute internal or external luminescence yield of the device.

Recent publications also stressing (bi-) excitonic gain processes in quantum structures are, e.g., [95K1, 95K2, 97C1, 01L1, 01R1], which are even partly claimed to coexist with EHP phenomena.

Inelastic scattering processes between two excitons or free carriers are treated in quantum structures, e.g., in [87F1, 89F1, 94C2, 96R1, 97F1].

We come back to the laser emission in structures of reduced dimensionality below and in Sect. 22.3.

Another group of intrinsic laser processes, still in the intermediate density regime, involves only interaction processes of virtually and coherently created particles. As an example we take the two-photon or hyper-Raman scattering

already introduced in Sects. 20.3 and 13.1.4. There we presented a process where a biexciton is created virtually by two quanta and decays under energy and momentum conservation.

If we now send an additional third quantum $\hbar\omega_\mathrm{s}$ into the sample with momentum and energy coinciding with those of a possible decay channel, this quantum can stimulate the decay of the virtually excited biexciton into another photon $\hbar\omega_\mathrm{s}$, $\boldsymbol{k}_\mathrm{s}$. A second photon $\hbar\omega_\mathrm{f}$, $\boldsymbol{k}_\mathrm{f}$ must then necessarily be simultaneously emitted according to (20.7). See [85H1].

If the third quantum lies energetically below the two pump quanta, $\hbar\omega_\mathrm{f}$ is necessarily located above, and the whole process is an example of electronic CARS (coherent anti-Stokes Raman scattering). This latter process also occurs with optical phonons. The sample in this case is again illuminated with some pump photon and with quanta corresponding to the Stokes emission, resulting in stimulated emission of the anti-Stokes line.

Alternatively these processes can be named non-degenerate four-wave mixing (NDFWM) and can also be considered as diffraction from a moving laser-induced grating written by the interference of one pump and one stimulating quantum, $\hbar\omega_\mathrm{p}$ and $\hbar\omega_\mathrm{s}$, respectively, and read by one of the two incident beams resulting in a Doppler-shifted diffracted signal. After the investigation of the above mentioned inelatic scattering or coherent NDFWM processes with respect to their laser properties in bulk material and in quantum structures (see above or [01L1]), we see now that analogous phenomena are investigated involving cavity polaritons, as outlined already in Sect. 20.2. The experiments are beautiful and the introduction of a new language in connection with the rediscovery continues. Stimulated emission due to polariton-polariton scattering in a microcavity in analogy to Fig. 20.1 or 22.1b is named polariton laser or PLASER, and the low theshold well known for four level laser systems mentioned above comes now partly under names like thesholdless lasing or lasing without inversion. See e.g. the Ref. [00B1, 00T1, 02B4, 02B5, 02S5, 03L1] of Chap. 20 or for more recent considerations of microcavity lasers [02K3, 02Z1]. The close connection between stimulated emission and Bose–Einstein condensation of cavity polaritons has already been mentioned in Sect. 20.5.4. Further references are e.g. [01O1, 02L2].

We now give two examples of stimulated emission in disordered systems. The tail of localized states in $CdS_{1-x}Se_x$. typically has a maximum tailing parameter $E_0(x)$ of about 5 meV and contains roughly 10^{18} states per cm^3. So the lower portion of this tail can be easily filled by optical pumping at low temperatures where thermal excitation into the extended states with a much higher density of states is prevented. If the gain value, i.e., the inversion, of the filled states is sufficiently large, laser emission sets in.

We show schematically in Fig. 22.4 (left-hand side) the density of states with the mobility edge ME and the chemical potential, μ, which indicates the energy up to which the states are filled at the highest excitation intensity. The right-hand side gives observed emission spectra showing the spikes of laser modes slightly above threshold.

Another aspect of disordered systems has been exploited recently to observe stimulated emission in indirect gap $Al_{1-y}Ga_yAs$ and similar materials [90C1, 94C1, 94K1, 94W1, 96K1] as mentioned briefly in Chap. 21 and references therein. GaAs is a direct-gap material with $E_g \approx 1.4\,eV$ at 4 K and AlAs an indirect one with $E_g \approx 2.2\,eV$. There exist alloys of all compositions y. For $y < y_c = 0.57$, the minimum in the conduction band at the X point in the first Brillouin zone becomes lower than that at Γ. Under high (optical) pumping most of the electrons therefore sit in the X minimum. However, the alloy disorder couples the states and Γ and at X so that the electron–hole pairs can recombine without participation of phonons. This fact enhances the transition rate so strongly that stimulated emission has been reached for $y < 0.57$ at a wavelength of 620 nm corresponding to 2.2 eV, i.e., already in the orange part of the spectrum with gain values up to $200\,cm^{-1}$ [94W1].

The coupling between Γ and X states can also be presented in another way. The alloying destroys to some extent the translational invariance of the lattice. As a consequence the $\boldsymbol{k}$-conservation, which is based on this translational invariance as shown in Sect. 3.1.3, is partly relaxed, allowing recombination processes which violate a strict $\boldsymbol{k}$-conservation rule.

We should like to mention that some recombination processes which involve impurity centers also lead to optical gain. In fact, one of the first theoretical considerations of stimulated emission from semiconductors started with the recombination of bound-exciton complexes (BEC) like excitons bound to

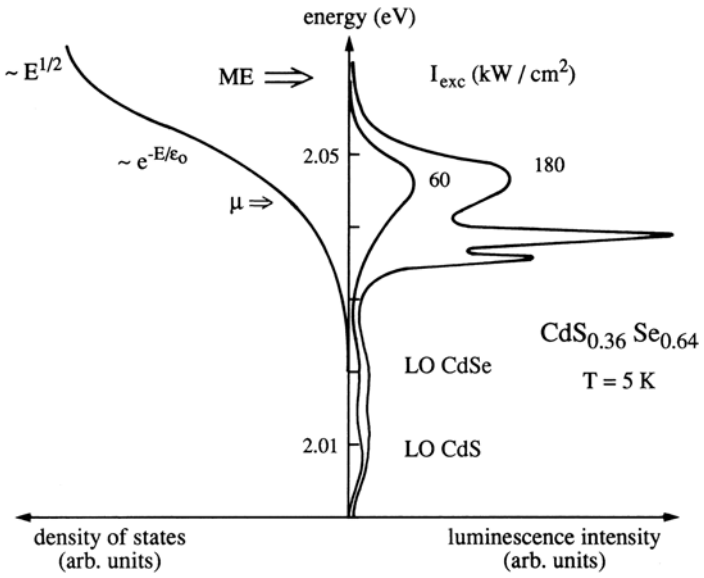


Fig. 22.4. *Left:* The density of localized and extended states in $CdS_{1-x}Se_x$ (schematic). *Right:* Observed emission spectra below and above threshold [87M1]

neutral acceptors or donors in CdS [62T1], because of their small spectral width. However, direct stimulated emission from BEC is observed only very rarely, but a slightly more complex mechanism gives optical gain. This is the recombination of a BEC under emission of a photon and an acoustic phonon, the so-called acoustic sideband. See Sect. 14.1 and [81K1]. This process also becomes possible because of the disturbed translational invariance. It is again a four-level process with low threshold as long as the acoustic phonon states are not (thermally) populated.

The two-electron transition explained in connection with Fig. 14.5 also yields gain as shown in [75K1] for ZnO.

One expects some advantages for semiconductor lasers in structures of reduced dimensionality. We come back to this aspect in Sect. 22.3. As a consequence of these expectations, gain processes have been widely investigated in structures of reduced dimensionality. We already mentioned a few examples above. Often the parameters of the gain are provided, such as its spectral position, spectral width, maximum value, temperature dependence and the pump power necessary to observe it, but no detailed information is given about the recombination process. We give in the following a further, rather limited selection of references to gain measurements, allowing the reader to enter deeper into the field. Apart from [92E1, 94C1, 99K1, 01L1] we mention the following.

The gain in GaN-based structures or more generally group III-nitrides is investigated, e.g., in [97N1, 98O1, 99M1, 03R1]. These structures resulted in commercially available light-emitting and laser diodes for the short wavelength part of the visible spectrum including the near-UV.

For the green spectral range, gain processes in $Zn_{1-y}Cd_ySe$-based structures have been investigated [92N1, 95K1, 95U1, 96G1, 97C1, 97P1, 98M1, 02G1] as well as ZnTe-based structures [94M1, 01C2, 01S3]. Laser diodes based on this material combination are still awaiting their commercial application, because the device lifetime is still limited to unacceptably low values by the creation of dark line defects under operation.

For laser emission in the IR based on lead salts or $Ga_{1-y}In_yN_{1-x}As_x$ see [01L1, 03I1] and [02K4, 04K1, 04P2, 04S1], respectively.

Some examples for the investigation of gain and lasing in quantum wires of III–V and II–VI materials can be found, e.g., in [90C1, 94C1, 94C2, 94K1, 99K1].

A further reduction of the dimensionality leads to gain and lasing from quantum dots including dots in insulators or self-assembled islands in semiconductors [92N1, 93F1, 93M1, 95G1, 95W1, 99E1, 99I1, 99K1, 01B1, 01E1, 02E1, 02S1, 03C1], in [97W1] of Chap. 1 and [95W1, 95W2, 96G1, 96W1] of Chap. 20.

Other materials for gain and laser emission in the blue and near-UV, which have become fashionable again, are ZnO [75K1, 81K1, 04P1] and ZnO-based quantum structures including nanorods (see, e.g., [02K3, 02Z1] or [97S1, 01H1, 02K4] of Chap. 15) allowing, similarly to GaN, partial lasing well above room temperature [98B1, 00O1, 04P1].

ZnO powders are also used to investigate lasing in "active random media", a process that involves enhanced backscattering of light at the ZnO grains

[99C1,00S1,01C1,01M1,01S1,04H2,04Q1]. This process has been discussed in Sects. 8.15, 9.6 and 14.4 in connection with localization of free carriers and/or excitons.

22.2 Electron–Hole Plasmas

The laser process which has presently the highest importance with respect to technical applications in laser diodes is the stimulated emission of a degenerate electron–hole plasma as outlined in Sect. 21.4. Population inversion occurs when the chemical potential μ of the electron–hole pair system is located above the reduced gap E'_g. See Fig. 21.1 or (21.16). It should be mentioned that some of the inelastic scattering processes or recombination under emission of a phonon or of a plasmon-phonon mixed state quasiparticle mentioned above may also occur in a plasma, contributing to the long-wavelength part of the gain spectra, or at densities below those fulfilling (21.16).

Population inversion can be achieved much more easily in indirect gap semiconductors like Si or Ge due to the longer carrier lifetimes, but the indirect nature of the transition makes the gain values so small that impracticably large volumes have to be pumped. However, far-infrared lasing has been reported in p-Ge between the heavy and light hole valence bands, or in a magnetic field between hole Landau (or cyclotron) levels [91G1]. Light emission and possibly gain from (partly doped) Si nanostructures are discussed in [01D1, 01K2, 02L1, 03D1] and the references given therein. On the other hand, gain values of up to $10^4\,\mathrm{cm}^{-1}$ can be reached with direct gap semiconductors, so that devices with a length of the active material of about 100 µm can be pumped in a forward-biased *pn* junction to give power densities in the $10^5\,\mathrm{W/cm}^2$ range at the surface of the laser. Differential internal quantum efficiencies deduced from the slope of the light power output versus electrical current input in excess of 50% have been reported.

We come now to one of the reasons why structures of reduced dimensionality are in principle favorable for the use as active materials in laser diodes.

In Fig. 22.5 we show calculated gain spectra resulting from an EHP for idealized $In_{.53}Ga_{.47}As/InP$ quantum structures of various quasi-dimensions d from 3 down to 0 as indicated in the figure.

The density is kept constant and the confining linear lengths are always 10 nm. It is obvious that the absolute gain values increase and that the widths of the gain spectra decrease with decreasing d. This effect would allow smaller active volumes and/or lower threshold currents for decreasing d.

We stress that these calculations are valid for idealized structures. The calculations include, e.g., only homogeneous broadening.

In reality, inhomogeneous broadening due to alloying and fluctuations of the width in the confinement direction(s) tend to increase the width and to

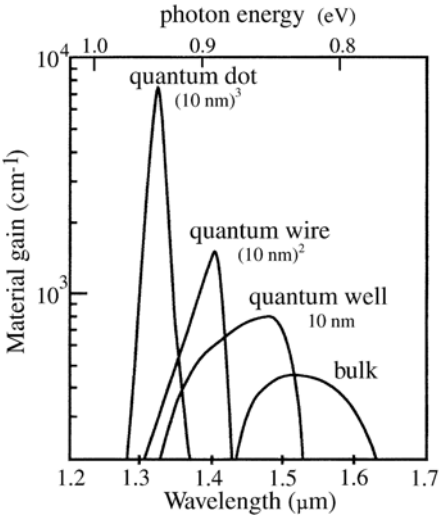


Fig. 22.5. Calculated gain spectra of idealized $In_{1-y}GaAs/InP$ quantum structures for various quasi-dimensions but with constant electron–hole pair density $n_P = 3 \times 10^{18}\,cm^{-3}$ [86A1, 99K1]

decrease the height of the gain spectra with decreasing d. Furthermore the confining potentials are sometimes rather moderate, e.g., for self-assembled islands so that carriers can escape at elevated temperatures. Consequently, commercially available laser diodes are based on double-heterostructures or quantum wells. The future will show if the various types of wires, e.g., etched, T-shaped or V-grooves (see Sects. 8.11 to 13) or even quantum dots or islands will make it into application.

22.3 Basic Concepts of Laser Diodes and Present Research Trends

Though this is not a textbook on semiconductor technology or devices, we outline in the first part of this section some developments and trends in the design of light emitting and laser diodes. A detailed discussion of this topic is, however, beyond the scope of this book and may be found, e.g., in [81S1, 92E1, 94C1, 99K1] and references therein. In the second part we list some of the present research trends in the field of semiconductors lasers.

The simplest way to build a laser diode would be to strongly bias a pn junction in the forward direction. In order to fulfill the condition $\mu > E'_g$ for simple band-to-band recombination at least one (better both) of the n and p doped layers had to be doped so highly, that the Fermi level is in the band, i.e., that the population is degenerate (see Fig. 22.6a).

This design has the disadvantage that neither the carriers nor the photons are confined or guided in any direction. Consequently, the threshold current density of such structures is extremely high so that such devices could only be operated in a pulsed mode and at reduced temperatures. This means that

they were essential to prove that the concept of laser diodes works, but they were only of very limited practical use.

A major breakthrough was the invention of double-heterostructures embedded in the intrinsic region of a *pin* diode (Fig. 22.6b). An undoped GaAs layer is, e.g., grown between *n*- and *p*-doped $Al_{1-y}Ga_yAs$ layers. The reduced band gap of this layer allows one to reduce the doping levels, as can be seen from a comparison between Figs. 22.6a and b. Furthermore the carriers are confined and the light quanta are already guided in one dimension since the larger band gap material has a smaller refractive index. Lateral guiding of the light can be achieved, e.g., by gain guiding (Fig. 22.6c) using a strip-like contact. Consequently, the current is injected along this line and photons experience maximum amplification along the direction of this line, which coincides trivially with the axis of the resonator, which in turn consists generally in the cases shown in Fig. 22.6a-d of cleaved semiconductor surfaces. Photons emitted under an oblique angle with respect to the resonator axis are lost for the laser, but are only amplified over a shorter distance. A lateral wave guiding can be achieved, e.g., by etching a ridge under the upper contact, possibly followed by a coating or overgrowth with high-band-gap (i.e., low-index) material to reduce surface recombination. This design already allows cw room temperature operation.

Since the threshold current in the forward-biased pin laser diode depends, among other things, on the volume in which inversion has to be reached, a more advanced design has been realized in which the electrons and holes are

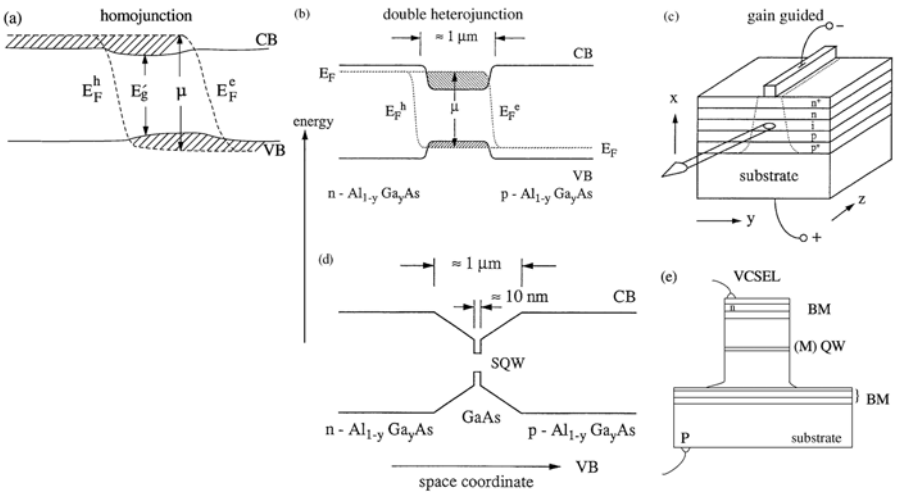


Fig. 22.6. Schematic drawings of a simple *pn* junction laser biased in the forward direction (**a**) of a double-heterostructure (**b**) of the effect of gain guiding (**c**) of a GRINSCH structure (**d**) and of a VCSEL (**e**)

confined in one (or a few) quantum wells with a typical width of about 10 nm, while the light quanta are guided in a structure with a width comparable to their wavelength, i.e., about 1 μm [81S1, 92E1, 96K1]. These graded index separate confinement heterostructure (GRINSCH) structures are shown in Fig. 22.6d. The funnel-shaped bandstructure of the optical waveguide helps to collect the injected carriers in the quantum well. Minimum threshold currents reached with these and similar structures are in the 1 mA range. A certain drawback of these structures is the limited spatial overlap of light-field and electron wavefunction.

Another currently very lively field of research concerns the development of surface-emitting laser diodes which can be arranged in one- and two-dimensional arrays and addressed individually. These arrays are very important ingredients for high brightness displays and in parallel electro–optic data handling (Sect. 24.2).

Among the most promising concepts are so-called vertical cavity surface emitting lasers (VCSELs) shown in Fig. 22.6e. These are monolithic devices, where two stacks of Bragg mirrors (BM) are grown epitaxially together with the cavity, which contains one or a few quantum wells at the positions of the antinodes of the cavity (also see Sect. 17.1). The current is injected, e.g., through n^+ and p^+ doped Bragg reflectors forming the cavity. Ideally two-dimensional arrays of VCSELs are fabricated, in which every single laser structure can be addressed individually to allow, for high-brightness two-dimensional displays. The concept of VCESL's is closely related to lasing involving cavity polaritons and their interaction processes. See Sect. 22.1 [04K1] or Refs. [00B1, 00T1, 02B4, 02B5, 02S5, 03L1] of Chap. 20.

Laser diodes have developed into various directions. Linear arrays are optimized towards maximum output to pump, e.g., Nd-solid-state lasers. Others are optimized towards minimum threshold currents or towards maximum modulation band width for data transfer. A basically simple way to modulate the output power is via current modulation. Another possibility has been realized by sending short optical pulses from the laser [98H1] with sufficient excess energy to heat the carrier gas to temperatures that reduce μ below E'_g. As a consequence, the laser switches off, although the electron-hole pair density is increased. After cooling the carriers the laser starts operating again.

Concerning the emission wavelength there are the two windows in glass fiber communication at 1.55 μm and 1.3 μm that can be covered by $Ga_{1-y}In_y As_{1-x}P_x$-based structures and more recently by $Ga_{1-y}In_yN_xAs_{1-x}$-based ones. $Ga_{1-y}Al_yAs$-based structures emit in the (nearinfra-) red.

For full color displays, diodes in the yellow and green (e.g., those based on $Cd_{1-y}Zn_ySe$) and in the blue (based on $Al_{y1}Ga_{1-y1-y2}In_{y2}N$ or on $Cd_{y1}Mg_{y2}Zn_{1-y1-y2}O$) are under investigation. References for these materials have been given above, but some can be added, e.g., [96R1, 96R2] for ZnO (though this is definitely not the "first report on lasing in ZnO by optical

pumping" as can be seen from [75K1, 81K1] and the references given therein) or [01C2, 01S3] for ZnTe-based materials and [00F1, 00T1] for $Zn_{1-y}Cd_ySe$-based structures or [02A1] for CdS-based ones.

Shorter emission wavelengths also allow one to increase the density of bits per unit area n_b, e.g., on CDs. There is a rough scaling law $n_b \sim \lambda^{-2}$. This means that a transition from the near-IR to the near-UV allows one to increase n_b by a factor often.

Organic semiconductors or phosphors are also being investigated for the purpose of (full color) displays. Presently these are mainly in the shape of light emitting diodes or electroluminescent devices, i.e., devices that operate below the onset of stimulated emission [02T1, 03M1, 04H1]. In this sense, one also investigates Si nanocrystals as they occur in porous Si (see, e.g., [01K2] and references therein).

Now we give a few references on some further present research trends in semiconductor lasers.

The incorporation of CdSe-based quantum islands as active materials might help to improve the lifetime of $Cd_{1-y}Zn_yS_{1-x}Se_x$-based II–VI lasers [01K1, 02K1].

Inter-subband or inter-miniband transitions like the ones described in Sect. 21.5 can be inverted to give laser emission. Examples for quantum structures can be found, e.g., in [96W1, 02K2]. We already mentioned an example for bulk p-Ge [91G1].

Quantum cascade lasers are unipolar devices that use, e.g., an electron several times for the emission of photons. See, e.g., [96F1, 02K2] and references therein. The basic idea is roughly the following. Electrons are injected in a p^+in^+ structure, which contains a periodic array of superlattices and quantum wells in the intrinsic region. The electron reaches the $n_z = 2$ state in the quantum well through the miniband states of the SL and performs a $n_z = 2 \rightarrow n_z = 1$ transition under emission of a photon, e.g., in the mid-IR, and tunnels from the $n_z = 1$ state into the miniband of the next SL, where the process repeats itself. For the incorporation of quantum cascade lasers in photonic crystals see [03C1].

Another topic are lasers without inversion, treated, e.g., in [89H1, 96F1, 01B2] and references therein.

Concerning the resonators, semiconductor laser resonators are made in the simplest case by cleaved surfaces. They may also contain Bragg mirrors like the cavities in Sect. 17.1 or the VCSEL structures mentioned above or in [01E1, 02P1], or other distributed feedback (DFB) structures [81S1, 92E1].

A recently investigated topic in this field is microdisc lasers emitting into so-called whispering gallery modes (see, e.g., [94J1, 00C1, 00L1, 03R1]).

For quantum kinetics, spatio-temporal dynamics and quantum fluctuations in semiconductor lasers see, e.g., [96H1, 03G1].

The last point we want to address is the temperature dependence of the laser threshold. As seen, e.g., from Fig. 22.2 there may be various types of de-

pendencies. For laser diodes one often finds for the forward current at threshold an empirical relation

$$I_{\mathrm{th}}(T) = I_0 \exp\{T/T_0\} \ . \tag{22.3}$$

The art of device development and materials engineering is currently to make I_0 as small as possible, e.g., by reducing losses including the density dependent Auger-recombination and to make T_0 large so that the properties of the laser do not change much with varying operating temperature.

With this small excursion into more application-oriented topics in semiconductor laser research and development we close this chapter.

22.4 Problems

1. How would you expect the curves of Fig. 22.2 a to shift with respect to each other if the total losses of the cavity increases or decreases by a factor $\sqrt{10}$?
2. Why is lasing via a degenerate EHP at room temperature more likely in standard III–V than in II–VI compounds? To answer this question calculate the effective density of states for electrons and holes in various 2d and 3d materials. Why should you do so?
3. Calculate the gain spectra for a 3d degenerate EHP of a direct and an indirect-gap semiconductor with otherwise identical parameters. Do the same for a quasi-2d direct gap material.
4. Why does lasing generally occur not at the maxima of the gain spectra but rather on their low-energy sides?

References to Chap. 22

[62T1] D.G. Thomas and J.J. Hopfield, J. Appl. Phys. **33**, 3243 (1962)
[64B1] N.G. Basov and O.V. Bogdankevich, In Proc. 7th Int. Conf. Phys. Semicond., Paris, 1964, Dunod, Paris, p. 225 (1964)
[73B1] J. Bille, Festkörperprobleme/Adv. Solid State Phys. **13**, 111 (1973)
[75K1] C. Klingshirn, phys. stat. sol. (b) **71**, 547 (1975)
[80W1] W. Wünstel and C. Klingshirn, Opt. Commun. **32**, 269 (1980)
[81K1] C. Klingshirn and H. Haug, Phys. Rep. **70**, 315 (1981)
[81S1] S.M. Sze, Physics of Semiconductor Devices, 2nd e., Wiley, New York (1981) and Semiconductor Devices, Physics and Technology ibid. (2002)
[83B1] M.S. Brodin et al., Sov. Phys. Techn. Phys. **9**, 1852 (1983)
[85H1] B. Hönerlage et al., Phys. Rep. **124**, 161 (1985)
[86A1] M. Asada, Y. Miyamoto and Y. Suematsu, IEEE J. QE **22**, 1915 (1986)
[87F1] Y.-P. Feng and H.M. Spector, J. Phys. Chem. Solids **48**, 593 and 1191 (1987)
[87M1] F.A. Majumder et al., Z. Phys. B **66**, 409 (1987)

[89F1] Y.-P. Feng, Z. Huang and H.N. Spector, J. Phys. Chem. Solids **50**, 117 (1989)
[89H1] S.E. Harris, Phys. Rev. Lett. **62**, 1033 (1989)
[90C1] R. Cingolani and K. Ploog, Adv. Phys. **40**, 535 (1990)
[91G1] E. Gornik, K. Unterrainer and C. Kremser, Optical and Quantum Electronics **23**, 267 (1991)
[92B1] R.B. Bhargava, J. Cryst. Growth **117**, 894 (1992)
[92D1] V.S. Dneprovskii et al., phys. stat. sol. (b) **173**, 405 (1992)
[92E1] K.J. Ebeling, Integrierte Optoelektronic, 2nd edn., Springer, Berlin, Heidelberg (1992)
[92N1] A.S. Nasibov et al., J. Cryst. Growth **117**, 1040 (1992)
[92N2] A.V. Nurmikko and R.L. Gunshor, phys. stat. sol. (b) **173**, 291 (1992)
[93F1] P. Faller et al., Opt. Mater. **2**, 39 (1993)
[93K1] C. Klingshirn, Adv. Mater. Opt. Electron. **2** (1993)
[93M1] Y. Masumoto, T. Kawamura and K. Era, Appl. Phys. Lett. **62**, 225 (1993)
[94C1] W.W. Chow, S.W. Koch and M. Sargent III, Semicondcutor Laser Physics, Springer, Heidelberg Berlin (1994)
[94C2] R. Cingolani, Physica Scripta T **49** B, 470 (1994)
[94I1] A. Imamoglu and R.J. Ram, Optics Letters, **19**, 1744 (1994)
[94J1] F. Jahnke and S.W. Koch, SPIE Proc. **2146**, 354 (1994)
[94K1] E. Kapon, Semicond. and Semimetals **40**, 259 (1994)
[94M1] F.A. Majumder et al., phys. stat. sol. b **186**, 591 (1994)
[94W1] A. Wörner, H. Kalt and R. Westphäling, Proc. 22th ICPS, Vancouver, p. 301 (1994)
[95G1] H. Giessen, N. Peyghambarian and U. Woggon, Optics and Photonics News, **6**, (dec. issue) 34 (1995)
[95K1] H. Kalt, Materials Science Forum **182-184**, 329 (1995)
[95K2] F. Kreller et al., Phys. Rev. Lett. **75**, 2420 (1995)
[95W1] U. Woggon et al., Japanes J. of Appl. Physics **34**, Supp. 232 (1995)
[95U1] M. Umlauff et al., Phys. Rev. B **52**, 5063 (1995)
[96F1] J. Faist et al., Phys. Rev. Lett. **76**, 411 (1996) and Appl. Phys. Lett. **68**, 3680 (1996)
[96G1] H. Grempel et al., phys. stat. sol. (b) **194**, 199 (1996)
[96H1] K. Henneberger and S.W. Koch, Phys. Rev. Lett. **76**, 1820 (1996)
[96K1] H. Kalt, Optical Properties of II–V Semiconductors: The Influence of Multi-Valley Band Structures, Springer Series in Solid State Sciences **120**, Springer, Berlin (1996)
[96R1] D. Robart et al., JOSA B **13**, 1000 (1996)
[96R2] D.C. Reynolds, D.C. Look and B. Jogai, Solid State Commun. **99**, 873 (1996)
[96W1] J. Wang et al., Superlattices and Microstructures **20**, 245 (1996) and IEEE Photonics Technol. Lett. **8**, 1001 (1996)
[97C1] L. Calcagnile et al., Mater. Sci. Eng. B **43**, 71 (1997)
[97F1] A.I. Filin et al., JETP Lett. **65**, 656 (1997)
[97N1] S. Nakamura and G. Fasol, The Blue Laser Diode, Springer, Berlin (1997)
[97P1] M.F. Pereira and K. Henneberger, phys. stat. sol. (b) **202**, 751 (1997)
[98B1] D.M. Bagnall et al., Appl. Phys. Lett. **73**, 1038 (1998)
[98H1] S.G. Hense, M. Elsässer and M. Wegener, Festkörperprobleme/Advances in Solid State Physics **37**, 207 (1998)

[98M1] P. Michler et al., J. Crystal Growth **184/185**, 575 (1998) and phys. stat. sol. (b) **206**, 399 (1998)
[98O1] Optical Properties of GaN and Related Materials, S.J. Pearton ed., Gordon and Break, Amsterdam (1998)
[98P1] Q.Y. Peng, phys. stat. sol. (b) **206**, 419 (1998)
[99C1] H. Cao et al., Phys. Rev. Lett. **82**, 2278 (1999)
[99E1] R. Engelhardt et al., J. Appl. Phys. **86**, 5578 (1999)
[99I1] S.V. Ivanov et al., Appl. Phys. Lett. **74**, 498 (1999)
[99K1] E. Kapon, in Semicondcutor Lasers, E. Kapon ed., p. 291, Academic Press, Amsterdam (1999)
[99M1] H. Morkoc, Nitride Semiconductors and Devices, Springer Series in Materials Science **32**, Springer, Berlin (1999)
[00C1] H. Cao et al., Appl. Phys. Lett. **76**, 3519 (2000)
[00E1] P.G. Eliseev et al., Appl. Phys. Lett **77**, 262 (2000)
[00F1] W. Faschinger and J. Nürnberger, Appl. Phys. Lett., **77**, 187 (2000)
[00L1] K.J. Luo et al., Appl. Phys. Lett. **77**, 2304 (2000)
[00O1] A. Ohtoma et al., Appl. Phys. Lett. **77**, 2204 (2000)
[00S1] Y. Sun et al., Appl. Phys. Lett. **77**, 2322 (2000)
[00T1] M.C. Tamargo et al., J. Crystal Growth **214/215**, 1058 (2000)
[01B1] P. Borri et al., Appl. Phys. Lett. **79**, 2633 (2001)
[01B2] A. Belyanin et al., Phys. Rev. A **64**, 013814 (2001)
[01C1] H. Cao et al., Phys. Rev. Lett. **86**, 4524 (2001)
[01C2] J.H. Chang et al., Appl. Phys. Lett. **78**, 566 (2001)
[01D1] L. Dal Negro et al., Opt. Materials **17**, 41 (2001)
[01E1] J. Erland et al., Phys. Rev. Lett. **86**, 5791 (2001)
[01H1] M.H. Huang et al., Science **292**, 1897 (2001)
[01K1] M. Klude et al., Electronics Lett. **37**, 1119 (2001) and SPIE Proc. **4594**, 260 (2001)
[01K2] D. Kovalev et al., Opt. Materials **17**, 35 (2001)
[01L1] Landolt-Börnstein, New Series Group III, Vol. 34 C1 and 2, C. Klingshirn, ed., Springer, Berlin (2001) and (2004)
[01M1] A. Mitra and R.K. Thareja, J. Appl. Phys. **89**, 2025 (2001)
[01O1] A. Olaya-Castro et al., Phys. Rev. Lett. **87**, 246403 (2001)
[01R1] J. Rubio et al., Solid State Commun. **120**, 423 (2001)
[01S1] G. v. Soest et al., Phys. Rev. Lett. **86**, 1522 (2001)
[01S2] R.L. Sellin et al., Appl. Phys. Lett. **78**, 1207 (2001)
[01S3] M. Schmidt et al., Proc. 25^{th} ICPS, N. Miura and T. Ando eds., Springer Proc. in Physics **87**, 1557 (2001)
[02A1] T. Abe et al., phys. stat. sol. (b) **229**, 1015 (2002)
[02E1] H.-J. Eisler et al., Appl. Phys. Lett. **80**, 4614 (2002)
[02G1] J. Gutowski et al., phys. stat. sol. (b) **234**, 70 (2002)
[02K1] M. Klude et al., phys. stat. sol. (b) **229**, 1029 (2002)
[02K2] R. Köhler et al., Nature **417**, 156 (2002)
[02K3] A. Kakovin et al., phys. stat. sol. (a) **192**, 212 (2002)
[02K4] J. Kvietkova et al., Proc. ASDAM '02 p 175, IEEE, Piscataway (2002)
[02L1] K. Luterova, J. Appl. Phys. **91**, 2896 (2002)
[02L2] Yu.E. Lozovik et al., Phys. Rev. B **66**, 075124 (2002) and JETP **98**, 582 (2004)
[02P1] Th. Passow, Adv. in Solid State Phys. **42**, 13 (2002)

[02S1] K. Sebald et al., phys. stat. sol. (a) **190**, 593 (2002)
[02T1] N. Tessler et al., Science **295**, 1506 (2002)
[02Z1] M. Zamfirecu, Phys. Rev. B **65**, 161205 (R) (2002)
[03C1] R. Colombelli, Science **302**, 1374 (2003)
[03C2] S.A. Crooker et al., Appl. Phys. Lett. **82**, 2793 (2003)
[03D1] J. De la Torre et al., Physica E **17**, 604 (2003)
[03I1] A. Ishida and H. Fujiasu, Optoel. Properties of Semicond. and Superlattices **18**, 533 (2003)
[03G1] E. Gehring and O. Hess, Spatio-Temporal Dynamics and Quantum Fluctuations in Semiconductor Lasers, Spriner Tracts in Modern Physics **189**, Springer, Heidelberg, Berlin (2003)
[03H1] H. Htoon et al., Appl. Phys. Lett. **82**, 4776 (2003)
[03M1] C.D. Müller et al., Nature **421**, 829 (2003)
[03R1] Y.P. Rakovich et al., Semicond. Science and Technol. **18**, 914 (2003)
[03R2] W.P. Risk, T.R. Gosnell and A.V. Nurmikko, Compact Blue-Green Lasers, Cambridge University Press, Cambridge (2002)
[04H1] W.E. Howard, Scientific Am. **290** (2), 64 (2004)
[04H2] G. Hackenbroich, Physik Journal **3**, 25 (2004)
[04K1] J. Kováč et al., Laser Physics **14**, 521 (2004)
[04P1] H. Priller et al., phys. stat. sol. (b) **241**, 587 (2004)
[04P2] S.H. Park, Appl. Phys. Lett. **85**, 890 (2004)
[04Q1] F. Quochi et al., Apll. Phys. Lett. **84**, 4454 (2004)
[04S1] S.A. Smith et al., Electronics Lett. **40**, 935 (2004)

23

Time Resolved Spectroscopy

Until now we have considered the linear and nonlinear optical properties of semiconductors mainly under quasi-stationary conditions, i.e., the excitation or measuring pulse lengths were assumed to be longer than the lifetime of the excited species.

The development of ps sub ps and finally of fs lasers pulses, with pulse durations down to only two cycles of the electromagnetic field [03T1] during the last two decades has allowed the dynamics of the optical excitations in semiconductors to be measured directly in the time domain. Since the characteristic time constants can be as short as a few fs, this field of research is also known as "ultrafast Spectroscopy". The field developed very rapidly through the competitive work of many research groups worldwide who were investigating intra- and inter-subband transitions in bulk semiconductors and quantum structures. When the pulse duration reached the limit of only two cycles of light (FWHM) and when all (quasi) dimensions had been explored from $d = 3$ down to $d = 0$ and the spatial resolution has been stretched to the resolution limit, the field reached, though still very active, a certain degree of maturity. The development of the field can be nicely followed in the Proceedings of the Series of International Conferences on the Physics of Semiconductors (ICPS) mentioned in the introduction, in Ultrafast Phenomena [93U1], in Nonlinear Optics and Excitation Kinetics (NOEKS) [88N1] and in some monographs, reviews and data collections such as [96H1, 96S1, 01L1, 02L1, 02S1, 03Q1, 04O1] or some older ones like [77L1, 78S1, 84S1, 91D1]. The diversity of topics in the various time regimes mentioned below is so large that it is completely impossible to cover all phenomena in all types of materials or structures in a general textbook on semiconductor optics. We therefore present here a didactic, but also to some extent, arbitrary selection of topics and examples.

We start with an introduction of the basic time constants and then proceed from shorter to longer times, or in other words, from coherent processes over intra- and inter-subband relaxation to the lifetime of electron–hole pairs.

23.1 The Basic Time Constants

We first introduce the basic time constants on a very elementary level and refine these concepts later to some extend. We concentrate again on the properties of electronic excitations in semiconductors and present the basic mechanisms with the help of Fig. 23.1, again using the weak-coupling limit if not stated otherwise.

With a short laser pulse an electronic excitation is created, here excitons in the continuum states.

The polarization, $\boldsymbol{P}$, produced by the incident electromagnetic field of this pulse is initially in phase with this field, i.e. the two waves are coherent. The first scattering processes which occur destroy this coherence. The characteristic time in which the fraction of the polarization that is still in phase with the exciting pulse decays to e^{-1} is called the phase-relaxation time T_2 ore more precisely the decoherence time (see also Chap. 4 and Sect. 3.1.5). For the part of the polarization wave which is still coherent with the exciting pulse we can write, to a first approximation

$$\boldsymbol{P}_{\mathrm{coh}} = \boldsymbol{P}_0 \exp\left(-t/T_2\right) . \tag{23.1}$$

Here we assume that the decay of the polarization is simply exponential and that this decay starts immediately after the excitation due to coupling to some "bath". For details see the so-called fluctuation-dissipation theorem treated, e.g., in refs. [81A1]k,l of Chap. 1. These are the so-called Markovian processes.

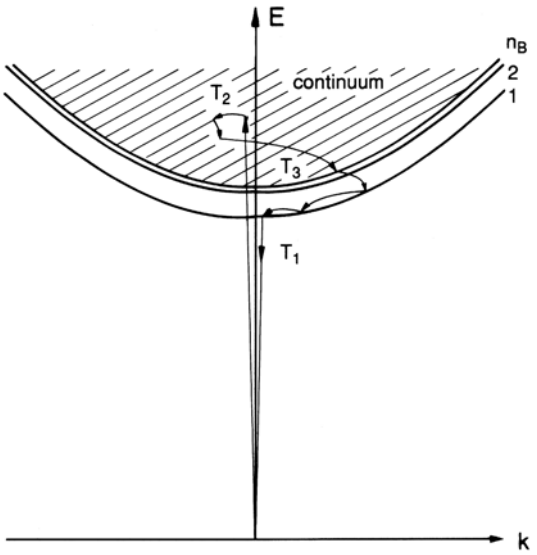


Fig. 23.1. A schematic drawing of the dispersion relation of excitons and of the main time constants

Deviations from this simple behavior then obviously form the non-Markovian processes. We shall see examples for both cases below [01B3, 01W1, 04O1].

The scattering processes which destroy the phase include:

(i) scattering with phonons. The interaction mechanisms are the Fröhlich coupling with the electric field of (preferentially longitudinal) optical phonons, the deformation potential and the piezo-coupling to acoustic and optical phonons. For the interaction processes with phonons see Sect. 8.6.
(ii) scattering with other electronic excitations, e.g., elastic and inelastic exciton-exciton or exciton-electron scattering,
(iii) scattering at impurities or other lattice defects. These processes also include the interface roughness scattering in quantum wells and scattering with alloy fluctuations in mixed crystals,
(iv) recombination. If there are no other phase-destroying processes, the phase will ultimately be lost in the recombination process, which can be described by a lifetime T_1 of the excited species with density $N(t)$. We assume again a simple exponential decay in this case we have as an upper limit

$$N = N_0 \mathrm{e}^{-t/T_1} \propto \boldsymbol{P}^2 \propto \left[\boldsymbol{P}_0 \exp\left(-t/T_2\right)\right]^2 \tag{23.2}$$

since the number-density N is proportional to the amplitude squared. From (23.2) follows the inequality

$$T_2 \leqslant 2T_1 \,. \tag{23.3a}$$

The "$<$" sign is valid if there are phase disturbing processes other than recombination, the equality is valid if recombination is the only phase-destroying process. We shall see in the future examples of both cases.

Sometimes one wants to investigate the "pure" dephasing T_2^*, i.e., all processes that destroy the phase other than recombination. In this case one can write

$$\frac{1}{T_2} = \frac{1}{2T_1} + \frac{1}{T_2^*} \tag{23.3b}$$

where T_2^* can exceed T_1 by more than the factor 2 present in (23.3a).

The first process (i) obviously obtains increasing importance with increasing temperatures, and the second one (ii) with increasing optical excitation or doping. The third process (iii) is especially important for free excitons in ordered structures and for excitons in extended states in disordered structures, while it can be considerably reduced at low temperatures for spatially localized excitons occurring, e.g., as bound-exciton complexes (Sects. 8.14 and 14.1, 2) or as localized excitons in alloys or quantum structures (Chap. 15).

To make this simple picture slightly more "true" in the sense of the statement at the beginning of this book, we state that there is a collection of times describing the phase relaxation.

The first aspect we mention is homogenous versus inhomogeneous broadening.

If all oscillators have exactly the same frequency, as introduced, e.g., in Chap. 4, the decay times of the polarization of the ensemble and of every individual oscillator are the same. The damping or the width Γ of the (Lorentzian) resonance are given by [77L1, 91D1, 96S1]

$$T_2 = \frac{2\hbar}{\Gamma} \,. \tag{23.4a}$$

In numbers this means that

$$\Gamma = 1\,\mathrm{meV} \mathrel{\hat{=}} T_2 \approx 1.3\,\mathrm{ps} \,. \tag{23.4b}$$

On the contrary, for a periodic harmonic oscillation a period T of 1 ps corresponds to

$$\hbar\omega = h\nu = \frac{h}{T} \tag{23.4c}$$

resulting in

$$T = 1\,\mathrm{ps} \mathrel{\hat{=}} \hbar\omega \approx 4\,\mathrm{meV} \,. \tag{23.4d}$$

If we have, on the other hand, an inhomogeneously broadened system, where every oscillator has a slightly different eigenfrequency distributed over an interval $\Delta\omega_{\mathrm{in}}$ around ω_0, e.g., according to a Gaussian distribution, the decay of the polarization of the ensemble is shorter than the decoherence of every individual oscillator for the following reason.

At the beginning, all oscillators are excited in-phase by a short pulse. Due to the slightly different eigenfrequencies a phase shift develops with time between the various oscillators with the consequence that the polarizations of the individual oscillators cancel. The dephasing of the macroscopic polarization or the "free polarisation" decay T_{pd} is given roughly by

$$T_{\mathrm{pd}} \cdot \Delta\omega_{\mathrm{in}} \approx 1 \,, \tag{23.4e}$$

which also depends on details of the frequency distribution. The decoherence time of the individual oscillator T_2 can be considerably longer than T_{pd}.

The amplitude of the radiation emitted according to (2.26), e.g., from an ensemble of oscillating dipoles, decays for inhomogeneous broadening with T_{pd} due to dephasing of the ensemble. This effect is called free polarisation decay in analogy to the free induction (or magnetization) decay of spin systems, e.g., in electron paramagnetic resonance (EPR, ESR) or nuclear magnetic resonance (NMR) experiments (see Chap. 27 for details and references).

Inhomogeneous broadening is important for localized or bound excitons, e.g., in alloys or quantum structures at low temperature, while homogeneous broadening dominates for free exciton resonances in high-quality bulk samples or at elevated temperatures when the homogenous broadening due to inelastic scattering with optic and acoustic phonons dominates over inhomogeneous contributions.

We shall see below that it is possible to distinguish, e.g., in four-wave mixing experiments, between homogenous and inhomogeneous broadening and to deduce the decoherence time T_2 of individual oscillators even in the presence of inhomogeneous broadening.

The coherence in the time interval from the excitation pulse to the dephasing is also known under the keyword quantum coherence and manipulations carried out during this time interval under coherent control.

The discussion has so far concentrated on the dynamics of the so-called interband polarization, which is valid, e.g., for excitons or band-to-band transitions, i.e., for the optical properties in the vicinity of the fundamental gap.

For electrical transport properties, e.g., through nanowires or point contacts another quantity is relevant, the so-called intraband coherence. This coherence is destroyed only by inelastic scattering processes. Elastic scattering of the charge carriers at defects does not destroy this type of coherence as becomes clear if we consider, e.g., the weak localization by enhanced backscattering introduced in Sects. 8.15 or 9.6. This topic is beyond the scope of this book, but we mention it here so that the reader is familiar with the terms.

A further aspect concerns the difference of the coherence of the interband polarisation of an (in)homogeneously broadened transition and the dephasing or decoherence of the spin of either the exciton as a whole or of an individual carrier (i.e., electron or hole) in the exciton. Since spins interact only weakly with phonons and electric polarisation, the coherence of the spin may be maintained even if the coherence of the interband polarization is destroyed, e.g., by some scattering process. This is especially true if the spatial part of the wave function and the spin part are not strongly coupled, e.g., via spin-orbit interaction. As we shall see below it is also possible to measure the dephasing of the spin independently of the dynamics of the spatial part of the wave function.

To conclude this short tour d'horizon on dephasing we mention a more philosophical aspect. If we make the system under consideration larger and larger (in the ultimate limit the whole universe) there is no more "bath" left by the coupling to which dephasing is introduced. Instead, all degrees of freedom oscillate with their proper energy factors $\exp\{i\omega t\}$. At the end, this aspect may lead to similar philosophical problems like the deterministic point of view in classical physics, but this aspect leads even further beyond the scope of this book than intraband coherence above.

The above aspects are treated in great detail in [96S1, 99A1, 02S1] and the references given therein, in addition to some of the text books mentioned at the beginning of this chapter.

Now we return to the discussion of the longer time constant.

A next step during the lifetime of an electron–hole pair is the intraband relaxation. This process is sometimes desbribed by a time constant T_3. It is of special importance when the excess energy with which the pair has been created is considerably larger than $k_{\mathrm{B}}T$.

The usual intraband relaxation for excess energies above $\hbar\omega_{\mathrm{LO}}$ takes place by emission of optical phonons. This process is rather fast with typical time constants in the (sub-) ps regime. The rest of the excess energy has to be dissipated by emission of acoustic phonons. This process is usually much slower and takes progressively more time with decreasing energy- and increasing momentum transfer towards the bottom of a parabolic dispersion.

After a few scattering processes among themselves and with the lattice, free electron–hole pairs (excitons) reach a thermal distribution which can be described by a temperature T_{e} and (usually) by Boltzmann statistics. (This statement is not true at low temperatures for excitons in localized tail states [93K1, 99K1]).

If the lifetime of the exciton is sufficiently long and the coupling to phonons sufficiently strong, the excitons thermalize with the lattice. This means that the temperatures T_{e} and T_{L} describing the distribution of excitons and of phonons in their respective bands become equal.

If the excitons are created with some excess energy $E_{\mathrm{excess}} \gg k_{\mathrm{B}}T_{\mathrm{L}}$, thermalization means a decrease of the average kinetic energy towards a value of $3/2k_{\mathrm{B}}T_{\mathrm{L}}$. If, in contrast, excitons are created resonantly at the bottom of the lowest band, thermalization means an increase of energy and a spreading in $\boldsymbol{k}$-space, since the kinetic energy connected with the photon momentum $\boldsymbol{k}_{\mathrm{ph}}$ is usually much smaller than $k_{\mathrm{B}}T_{\mathrm{L}}$ even for $T_{\mathrm{L}} \approx 4.2$ K, i.e., at liquid He temperature. The quantity k_{B} "times" 4.2 K is roughly 0.4 meV while $\hbar^2\boldsymbol{k}_{\mathrm{ph}}^2/2M$ is in the 10 μeV range depending on the material parameters.

Finally the excitons recombine radiatively or non-radiatively with a time constant T_1. T_1 is generally in the ns regime for direct gap semiconductors and reaches values in the μs or even ms range for indirect materials.

In most semiconductors, the recombination process is predominantly non-radiative. Only for some high quality quantum well or alloy samples has the luminescence yield η (i.e., the ratio of the number of emitted photons to excited electron–hole pairs) been claimed to reach unity [98W5,03F1]. Otherwise typical values of η are $10^{-3} \lesssim \eta \lesssim 10^{-1}$ for good direct gap bulk materials with dipole-allowed band-to-band transition. For materials containing a lot of non-radiative centers (so-called luminescence killers like Fe or Cu ions in some compounds) or in indirect gap materials like Si and Ge or in direct ones with dipole-forbidden band-to-band transition like Cu_2O, η can be even considerably smaller.

This means that the decay time of the luminescence generally does not give, i.e., for all cases with $\eta < 0.5$, the radiative decay time in contrast to frequent claims of this sort found in the literature. The fact that good agreement between theories assuming $\eta = 1$ and experiment is nevertheless often found can possibly be reasoned as follows. The radiative and many of the nonradiative recombination channels depend on the electron–hole overlap integral. Since this quantity is generally not very precisely known in absolute units, it is frequently used as a fitting parameter, with the result that both radiative and nonradiative recombination pro-

cesses are taken into account implicitly but unintentionally in the theoretical description.

Apart from the "mono-molecular" recombination processes which result in an exponential decay described by T_1 in (23.2), there are other recombination processes like inelastic scattering described in Sect. 20.2 or the nonradiative Auger recombination. In the latter process one electron–hole pair (exciton) recombines and transfers all of its energy to a third particle, e.g., a free electron, as additional kinetic energy. This process can limit the quantum efficiency in high density electron–hole pair systems, e.g., in laser diodes [75H1,79H1]. Still other processes involve the subsequent capture of a carrier, e.g., an electron at some deep center and the subsequent recombination with or capture of a hole at the same center. All these processes lead to a nonexponential decay that can in principle not be described by a single time constant T_1 but involves more complex models of recombination kinetics, which generally tend to contain a lot of unknown cross-sections or reaction constants. Sometimes it is at least possible to give an effective lifetime T_1^{eff} averaged over a certain density or spectral interval. In the following, we shall see examples for the various situations.

We give here a short summary of various decay relations for a population $N(t)$

$$N = N_0 e^{-t/T_1} \tag{23.5a}$$

is the simple (and single-) exponential decay obtained for mono-molecular decay processes.

Sometimes one observes a multi-exponential decay

$$N = \sum_i N_{0i} e^{-t/T_1} \tag{23.5b}$$

which goes over to

$$N = \int_0^\infty \frac{N(T_1)}{T_1} e^{-t/T_1} \, \mathrm{d}T_1 \tag{23.5c}$$

if the number of different T_{1i} gets large.

In some cases one observes a stretched exponential decay

$$N = N_0 e^{-(t/T_1)^\beta} \tag{23.5d}$$

especially for localized states.

In some other cases power laws appear

$$N = \sum_i N_{0i} (t/T_{1i})^{-\alpha} \tag{23.5e}$$

with $-\alpha = -1$ for a bimolecular decay process.

In order to be also to make a reliable distiction between the various possibilities (23.5a to e), one should be able to follow the decay of $N(t)$ over at least three orders of magnitude. For a smaller variation a reasoable fit i possible with various models by an appropriate selection of the parameters.

The possibilities to introduce the various time constants, like T_2, and to decribe the various phenomena like four wave mixing in the frame of optical or semiconductor Bloch equations are given in Chap. 27 and the references listed there.

After this general introduction, we present some selected and we trust, representative examples together with some information about the relevant experimental techniques. For the latter topic see also Sect. 25.3.

23.2 Decoherence and Phase Relaxation

We start with some selected examples of the measurements of the T_2-time and then present examples for non-Markovian processes and coherent control.

23.2.1 Determination of the Phase Relaxation Times

The most widely used techniques to determine T_2-times are based on four-wave mixing experiments. Therefore we start with them and then present other, sometimes even older techniques.

23.2.1.1 Four-Wave Mixing Experiments

The standard mathematical description of what follows is in terms of optical or semiconductor Bloch equations and the density matrix formulation. These aspects are presented in a didactic fashion in Chap. 27. Additionally we refer the reader to [46B1, 57F1] and textbooks like [96H1, 96S1, 02S1, 03Q1, 04O1] or to conference proceedings like [88N1, 93U1, 03C1]. Here we give an intuitive description.

In Fig. 23.2a–c we show the principle setup and in (b) the temporal evolution of various quantities for an homogeneously broadened situation. In this case all oscillators in the medium have exactly the same eigenfrequency ω_0 and the finite width of a resonance comes exclusively from its finite T_2 time. In Fig. 23.2c we give the frequently more realistic case of inhomogeneous broadening. In this case the eigenfrequencies are additionally distributed over a certain frequency range $\Delta\omega_{\text{inh}}$ around ω_0.

We start with the homogeneous case. A first incident pulse with wave vector $\boldsymbol{k}_1$ and frequency ω_1 produces in the sample a polarization $\boldsymbol{P}_1$, the coherent part of which decays with T_2. During its decay, this polarization radiates according to the well-known law

$$I_{\text{rad}} \propto \left|\ddot{\boldsymbol{P}}\right|^2 . \tag{23.6}$$

This radiation was known as “free induction decay” in analogy to corresponding techniques in magnetic (spin-) systems. A more appropriate and nowadays

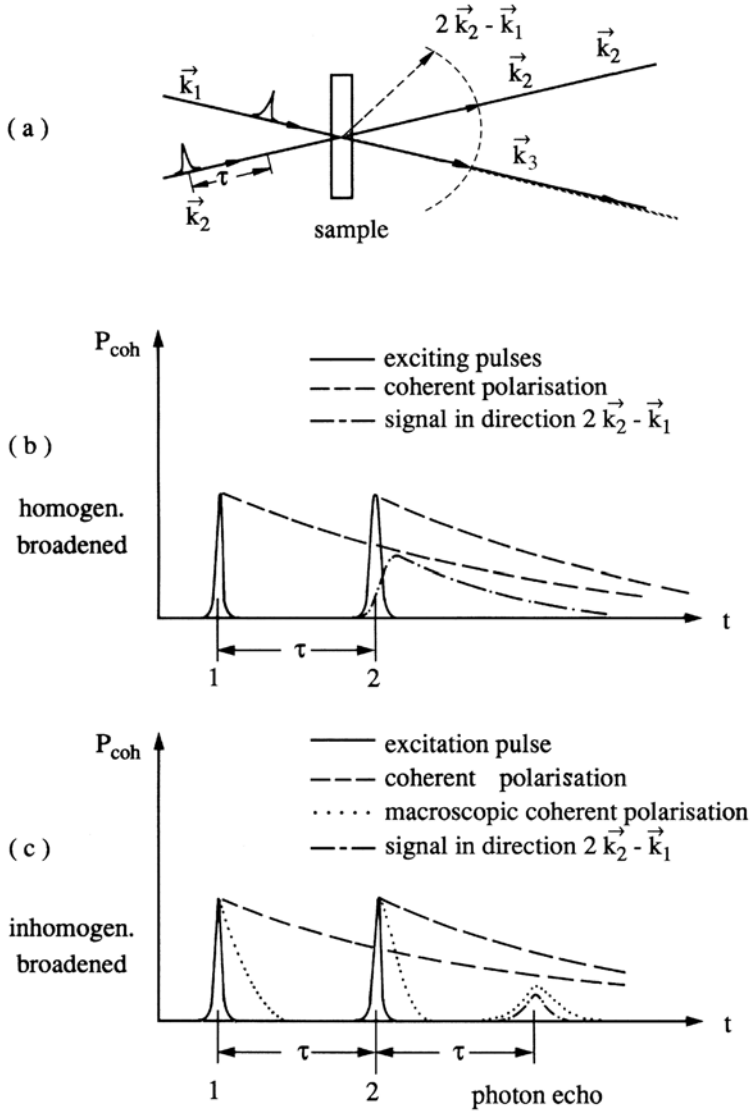


Fig. 23.2. The setup for a photon-echo experiment (**a**), and the time dependence of the polarization for a homogeneously (**b**) and an inhomogeneously (**c**) broadened two-level system

generally used name is "free polarization decay". Usually radiation has only a minor influence on the damping, i.e., T_2 is usually limited by scattering processes rather than by the radiative lifetime. Exception will be mentioned later.

After a delay τ a second pulse $\boldsymbol{k}_2$ arrives which is coherent with the first one and (usually) has the same frequency but a different direction. To achieve

this coherence the two pulses are generally produced from a single pulse using a beam splitter and a variable optical delay for one of them.

The second pulse creates a polarization $\boldsymbol{P}_2$, which interferes with the coherent part left over from the first one. This interference produces a laser-induced grating (LIG) of the polarization which radiates its first orders in the directions given by (23.7), assuming a thin grating, i.e., the Raman–Nath regime:

$$\boldsymbol{k}_{\text{diff}}^{(1)} = 2\boldsymbol{k}_2 - \boldsymbol{k}_1; \boldsymbol{k}_{\text{diff}}^{(-1)} = \boldsymbol{k}_1 \,. \tag{23.7}$$

The Raman–Nath regime is characterized to a first approximation by the following inequality between the thickness of the sample d, the spacing of the grating Λ and the wavelength of the incident beams λ:

$$d^2 < \lambda\Lambda \,. \tag{23.8}$$

The negative first-order signal is difficult to detect since it coincides with the direction of the transmitted part of the pulse 1. So we concentrate on the other, background-free direction.

Since, in the homogeneously broadened case, all oscillators have the same eigenfrequency, the coherent parts of the polarizations all have the same phase and the scattered signal I_{s} starts to develop immediately when the second pulse arrives. It reaches its maximum I_{sp} when the second pulse is over, assuming that the duration τ_{p} of both pulses is much shorter than T_2, i.e.,

$$\tau_{\text{p}} \ll T_2 \,. \tag{23.9}$$

The diffracted signal in the case of homogeneous broadening depends on the delay time τ in the following way

$$I_{+1}^{\text{hom}} \propto |\boldsymbol{P}(\tau)|^2 \propto \left[\exp\left(-\tau/T_2\right)\right]^2 \propto \exp\left(-2\tau/T_2\right) \,, \tag{23.10a}$$

and the same relation holds for the time-integrated signal as shown below in Fig. 23.3a.

For the inhomogeneously broadened case of Fig. 23.2b, the part of the polarization amplitude of every oscillator which is still coherent with pulse 1 again decays with T_2. But since all oscillators have (slightly) different eigenfrequencies, they lose phase with one other in a time inversely proportional to the inhomogeneous broadening $(\Delta\omega_{\text{in}})^{-1}$, resulting essentially in destructive interference of the radiation of all oscillators. The emission of the free polarization decay is then also limited by this quantity as shown in Fig. 23.2b. When the second pulse arrives after τ, a rephasing starts in the sense of phase conjugation or cum grano salis of a time reversal. See Chap. 27. After another time τ all oscillators are in phase again and radiate the so-called photon-echo (Figs. 23.2c and 23.3b). Its temporal width is limited by the free polarization decay and thus by τ_{p} or by $(\Delta\omega_{\text{in}})^{-1}$, and the (time-integrated) signal intensity decays with τ according to

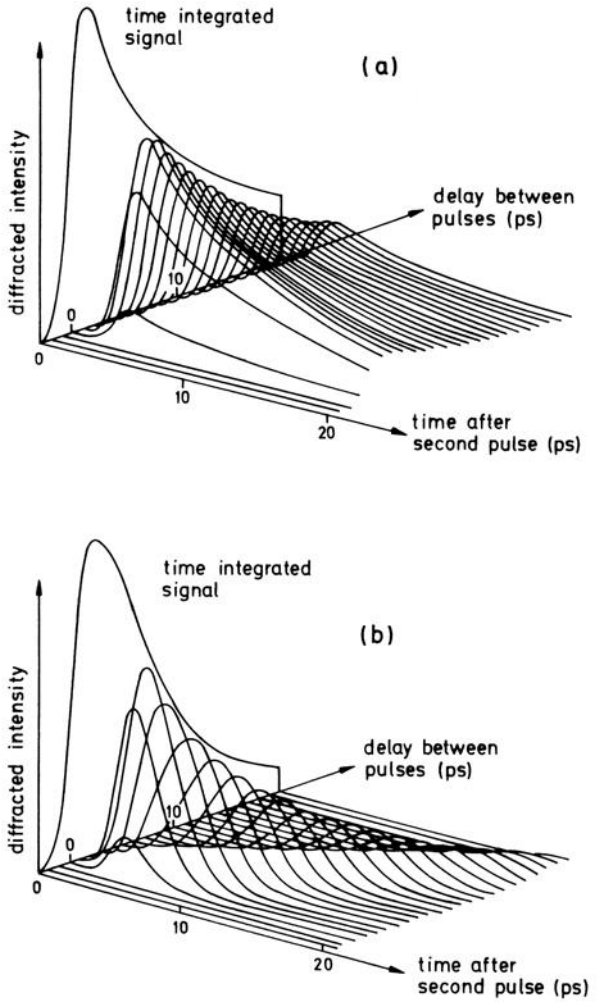


Fig. 23.3. The temporal evolution of the diffracted order as a function of the delay between the two pump pulses and the time after the second pump pulse for a homogeneously (**a**) and an inhomogeneously broadened system (**b**). The time-integrated diffracted signal is also given [89D1]

$$I_{\mathrm{s}(0)}^{\mathrm{inhom}} \propto |\boldsymbol{P}\,(t=2\tau)|^2 \propto [\exp\,(-2\tau/T_2)]^2 \propto \exp\,(-4\tau/T_2)\ . \qquad (23.10\mathrm{b})$$

The appearance of an echo as shown in Figs. 23.2c and 23.3b is thus a clear indication of an inhomogeneous broadening.

The formula to calculate the diffracted orders and their time evolution tend to get rather lengthy and can be treated only numerically when going from the optical to the semiconductor Bloch equations ore when using realistic laser pulses of finite duration and not just δ-pulses (see Chap. 27). Therefore

we refer the reader to the books mentioned at the beginning of this chapter, in addition to Chap. 27, for approximate formulae on the basis of two- and three-level systems.

The appearance of a photon echo can be demonstrated in a didactic way also by Fig. 23.4 (see Chap. 27).

A bunch of runners representing the polarisation of the various oscillators starts in a stadion at $t = 0$, i.e., at the arrival time of the first pulse. Since they all have different speeds (inhomogeneous broadening) their angular positions spread with time. At the time of the second pulse arriving at $t = \tau$ they turn around instantaneously and run back with the same speed. This results in a collective arrival at the starting point $t = 2\tau$, (emission of the echo). The decay of the echo amplitude with τ can be visualized in this picture if we assume that some runners stumble during their way and fall down (recombination). Even if they get up again after a while and continue to run they will not arrive back at the starting point in time. The same holds for runners who change their speed (these two processes would correspond to inelastic scattering) or who arbitrarily change their direction without stopping or changing $|v|$ (inelastic scattering).

A peculiarity of the description in the weak coupling limit in the framework of $\chi^{(3)}$ and of the optical Bloch equations is the fact that no signal appears in the direction $2\boldsymbol{k}_2 - \boldsymbol{k}_i$ for negative delay, i.e., if pulse 2 arrives before pulse 1. This fact is surprising since a negative delay corresponds to an interchange of pulse 1 and 2 and the first diffracted orders of pulse 1 and 2 go in the direction of $\boldsymbol{k}_2$ and $2\boldsymbol{k}_1 - \boldsymbol{k}_2$ and $\boldsymbol{k}_1$ and $2\boldsymbol{k}_2 - \boldsymbol{k}_1$, respectively. If the polariton picture or the semiconductor Bloch equations are used which include among others the Coulomb interaction between the carriers (see Chap. 27), a signal is also expected for negative delays. First calculations show [91W1] that the temporal buildup of the signal for negative delays has twice the slope of the decay for positive τ.

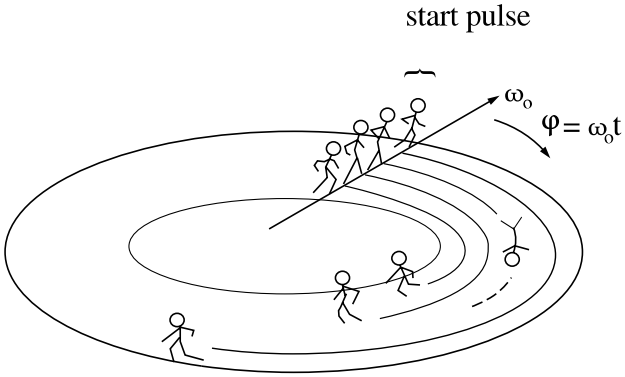


Fig. 23.4. The illustration of the appearance of a photon echo with runners in a stadion

We have mentioned already that most transitions in semiconductors are inhomogeneously broadened, at least at low temperatures and densities. Bound exciton complexes are inhomogeneously broadened since the surroundings of every impurity are slightly different due, e.g., to other impurities in the neighborhood. For the localized tail states in disordered systems and for quantum structurs at low temperatures the inhomogeneous broadening is obvious. For free excitons or exciton polaritons the question of homogeneous or inhomogeneous broadening is less clear since a different eigenenergy belongs to every $\boldsymbol{k}$-vector due to dispersion. The continuum states of the excitons or the band-to-band transitions in an electron–hole plasma again form, to a good approximation, an inhomogeneously broadened system with, however, very short T_2 times. At higher densities or temperatures the T_2 times get shorter and may become so short that they dominate the broadening, i.e., we get a transition to homogeneous broadening.

We now present first experimental results for bulk samples, namely FWM results obtained in the ($n_\mathrm{B} = 1$) $A\Gamma_5$ free exciton resonance of CdSe at low temperature. Since even samples of a few μm thickness are opaque in this regime, the diffracted signal in reflection was used (Fig. 23.5a). With increasing pump power, the signal increases, too, but the decay gets faster. This is due to inelastic exciton-exciton scattering. Using these data and others obtained in the exciton to biexciton transition (Fig. 23.5b), the curve of Fig. 23.5c is obtained which shows a linear increase of the dephasing rate T_2^{-1} with increasing excitation. The extrapolation to zero density (and temperature) gives T_2 values around 50 ps. The slope of the curve in Fig. 23.5c allows one to determine additionally the exciton-exciton scattering cross-section via

$$T_2^{-1} = T_{02}^{-1} + \delta N_\mathrm{ex} \tag{23.11a}$$

and one finds with [89D1]

$$\sigma_\mathrm{ex-ex} = \delta v_\mathrm{th} \tag{23.11b}$$

$\sigma_\mathrm{ex-ex} = 6.87\pi a_\mathrm{B}^2$, i.e., a value slightly larger than the geometrical one. In [91W1, 98W1] a linear dependence of the dephasing rate on the density has been found both for bulk-like and quasi 2d ZnSe structures, with some dependence on the fact that if the exciton population is still coherent or not. It should be mentioned that for much higher densities corresponding to those reached in an electron–hole plasma and short times one finds a decay time T_s of the FWM mixing signal on the electron–hole pair density n_eh given by [88B1, 91W1, 01W2]

$$T_\mathrm{s}^{-1} = T_0^{-1} + c n_\mathrm{eh}^{1/3} \tag{23.11c}$$

for all quasi-dimensions.

The temperature dependence of the dephasing rate is given in all dimensions by the sum over a low temperature value Γ_0, a term linear in T due to

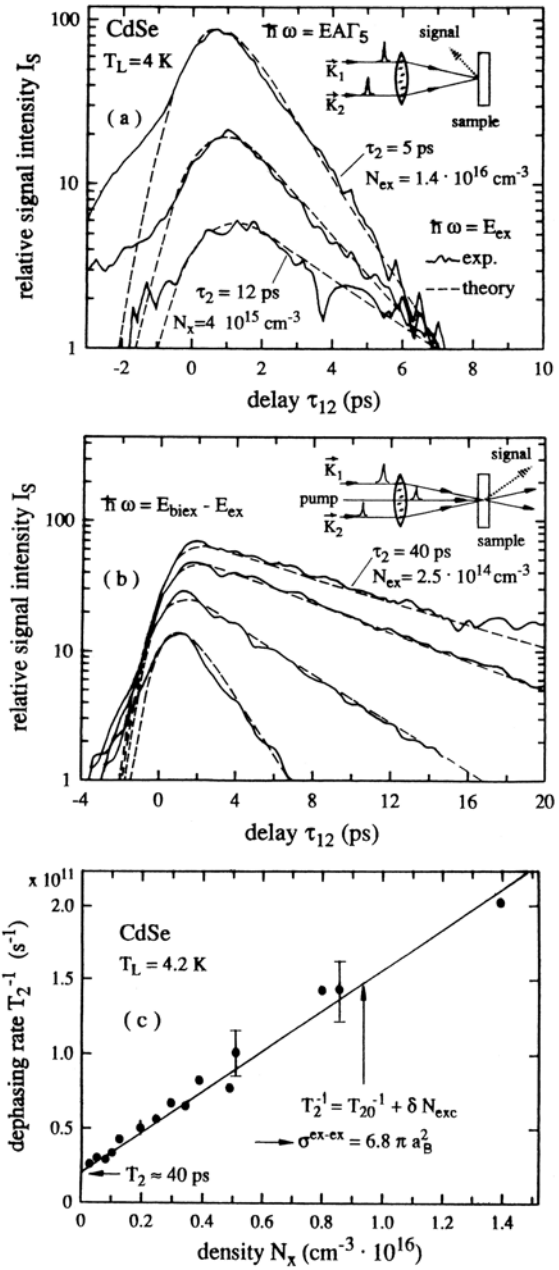


Fig. 23.5. The evolution of the time-integrated photon echo as a function of τ in the $n_{\mathrm{B}} = 1A\Gamma_5$ exciton resonance of CdSe for different pump powers in reflection geometry (**a**) in the exciton $\rightarrow$ biexciton transition, where a well-defined exciton density is created by a prepulse in transmission geometry (b) and the dependence of T_2^{-1} on the exciton density (**c**) [89D1]

scattering with acoustic phonons and a Bose term proportional to the occupation probability of LO phonons [92O1, 98W1, 04H1], i.e.,

$$T_2^{-1} = \Gamma_0 + \beta_{\mathrm{ac}} T + \frac{\beta_{\mathrm{LO}}}{(\exp(\hbar\omega_{\mathrm{LO}}/k_{\mathrm{B}}T) - 1)} \,. \tag{23.11d}$$

The onset of the signal for negative delays is not covered by the simple theory, as mentioned above, but by the improvements given in [91W1, 98W1].

Photon-echo experiments at low temperature in GaAs revealed T_2 values around 10 ps [89K1, 92O1]. Diffusion measurements with laser-induced population gratings in CdS gave a diffusion constant $D(5\,\mathrm{K}) \approx 20\,\mathrm{cm}^2\mathrm{s}^{-1}$ and with the Nernst–Townsend–Einstein relations,

$$D = \mu \frac{k_{\mathrm{B}}T}{e} \,, \quad \mu = e\frac{T_2'}{M} \,, \quad D = \frac{k_{\mathrm{B}}T}{M} T_2' \,, \tag{23.11e}$$

for classical effective mass particles, a time between scattering events $T_2' \approx 30\,\mathrm{ps}$ [88W1]. Here, μ is the mobility of excitons and T_2' the time between two scattering processes. Evidently this quantity is closely related to the phase relaxation time. We shall describe a more detailed experiment concerning this point later in connection with MQW. Here, we can state that T_2 times in the range a ps up to a few tens of ps are obviously typical for the lowest free exciton states in high quality bulk (and MQW, see below) samples at low temperature and density in agreement with data from reflection spectroscopy. The T_1 values under these conditions are in the 0.3–3 ns regime, i.e., we have $T_2 \ll T_1$. T_2 decreases with increasing density and with increasing temperature. In the latter case T_2 values in the range of 100 fs are reached at room temperature, so that the homogeneous dominates the inhomogeneous broadening.

The damping of polaritions has been deduced in CuCl by FWM in [83M1, 91V1]. Polaritons and bipolaritons have been analyzed for CdS in [01M1] reaching the $T_2 = 2T_1$ limit of (23.3a) for the biexciton with times of a few ps.

Now we consider, still in 3d semiconductors, excitons which cannot move freely through the sample. We start with an example of a bound-exciton complex (Sects. 9.5, 14.1), more precisely with an exciton bound to a neutral acceptor (A^0X) in CdSe.

In Fig. 23.6a we show directly the time-resolved photon-echo for the laser tuned into the A^0X resonance of CdSe at low temperatures. One sees clearly the scattered intensities of the two pump pulses $P_{1,2}$ and the photon-echo P_{s}. A first inspection already shows that the T_2 times of this BEC are much longer than for free excitons. Extrapolation of the data to $T = 0\,\mathrm{K}$ (Fig. 23.6b) gives $T_2 \approx 600\,\mathrm{ps}$. Since the lifetime T_1 of this BEC in this same sample has been determined to be around 400 ps, we come here again close to the limit given by (23.3a). The comparison of this value, which corresponds roughly to a homogeneous width of 20 μeV, with the spectral width of the BEC luminescence, which in good samples is between 0.1 and 1 meV, together with the observation of a photon-echo, clearly shows the inhomogeneous broadening of the BEC resonances.

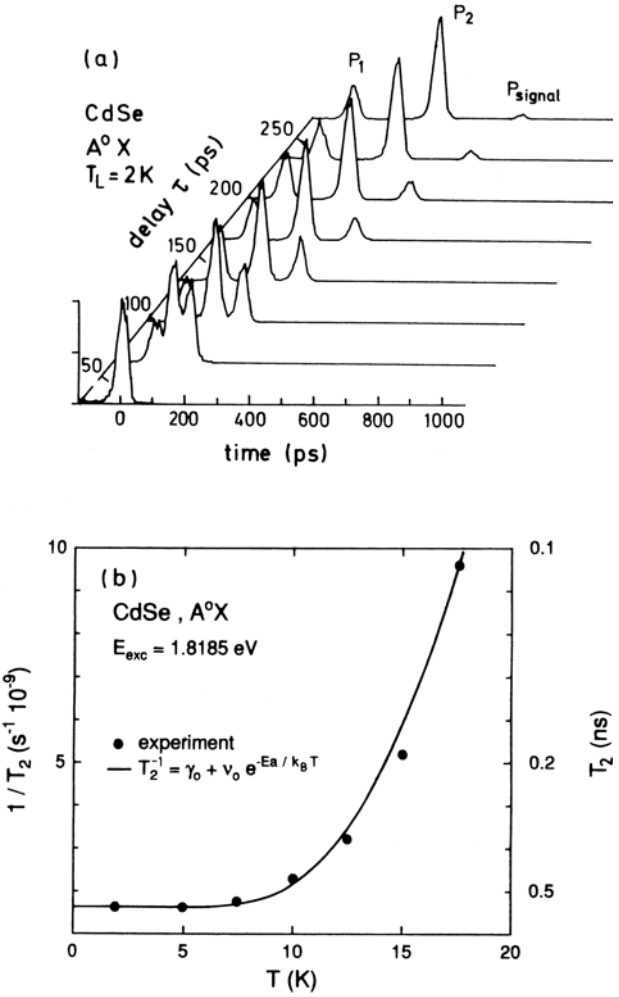


Fig. 23.6. Photon-echo observed in the resonance of a bound exciton complex in CdSe (**a**) and the dephasing rate as a function of temperature (**b**) [92P1,92P2,94P1]

The temperature dependence of T_2 time is given in Fig. 23.6b. It can be fitted by an activation law.

$$T_2^{-1} = T_{02}^{-1} + \nu_0 \exp(-E_\mathrm{a}/k_\mathrm{B}T) \tag{23.12}$$

with $T_{02} \approx 600\,\mathrm{ps}$, $\nu_0 = 3 \times 10^{11} s^{-1}$ and an activation energy of $E_\mathrm{a} \approx 5.6\,\mathrm{meV}$. This value corresponds approximately to the binding energy of the exciton to the complex. Since T_{02} is slightly above the observed luminescence decay time and thus of T_1 [92P2], we can conclude that the main phase destroying processes for this complex at low temperature are the recombination

and at higher temperatures the thermal ionization of the exciton from the neutral acceptor by absorption of a phonon.

In CdS a T_2 value of 400 ps has recently been observed for the D^+X complex with the quantum-beat technique to which we shall return later [91S1].

The difference in the T_2 times of BEC compared to free excitons at low temperatures can be qualitatively understood in the following way. A free exciton moves through the sample with a certain velocity and has a good chance of hitting a defect like an impurity or a dislocation. The resulting scattering process destroys the phase coherence with the driving laser pulse. A resonantly created bound exciton sits on the defect, oscillates in phase with the generating light field and cannot do much more at low temperatures until it recombines, provided that the density of other centers is so low that there is no interaction between them. The same type of arguments seems to hold for localized excitons in some disordered systems (Sects. 9.6 and 14.4) such as $CdS_{1-x}Se_x$ to which we turn now. See, e.g., [92S1,92S2] and references therein.

In Fig. 23.7 we show the luminescence spectrum of a $CdS_{0.65}Se_{0.35}$ sample at $T = 2\,K$ excited at the high energy edge of the zero phonon luminescence band, i.e., in the transition region from extended to localized states. One observes the zero phonon emission from the localized states around 2.21 eV and the CdSe and CdS LO phonon replica around 2.19 and 2.18 eV, respectively. The full dots give T_2 values at the respective energies measured with the photon-echo technique. The values range from $T_2 \lesssim 400\,ps$ close to the "mobility edge" up to $T_2 \approx 3\,ns$ for deeper localized states. Again these values are comparable with the intraband relaxation times T_3 and the interband recombination times T_1. Data for these latter quantities have been determined from time- and spectrally resolved luminescence and range from some hun-

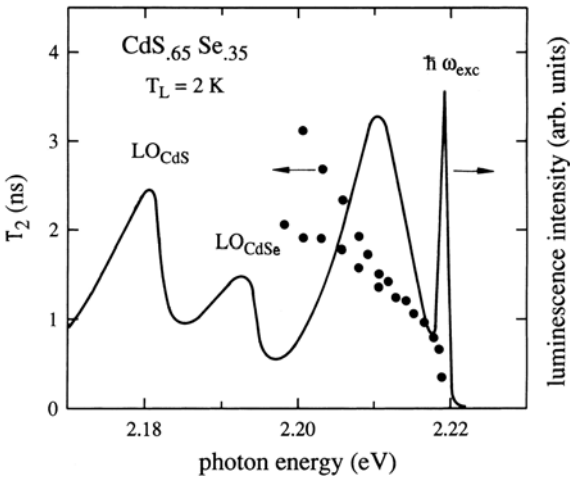


Fig. 23.7. The low temperature luminescence spectrum of a $CdS_{1-x}Se_x$, sample and the energy dependence of the T_2 times deduced from photon-echo experiments [92S1]

dreds of ps to several ns; see [92S1] and references therein. An increase of the lattice temperature leads once more, via an increasing phonon scattering, to a decrease of T_2.

An interesting aspect of ultrafast time-resolved spectroscopy can be treated in connection with Fig. 23.8, where we give the temporal decay curves of the time-integrated photon echo for $CdS_{1-x}Se_x$ using laser pulse durations from 0.12 to 10 ps at low temperatures ($4.2\,\mathrm{K} \leqslant T \leqslant 10\,\mathrm{K}$). The density of excited carrier pairs is in all cases roughly equal and in two cases even the same sample has been used.

Two effects are striking. The exponential tail from which we deduce T_2 is missing or at least below the detection limit for $\tau_\mathrm{p} = 0.12\,\mathrm{ps}$. However, it becomes increasingly stronger and longer for increasing τ_p. Furthermore the ratio of the first initial spike, which corresponds roughly to the autocorrelation function of the laser pulse and which is also known as the "coherent artefact", to the slower decaying part decreases with increasing τ_p. The interpretation follows Fig. 23.8e where we show the density of states for extended and for localized states, and the spectral shape of a short ($\tau_\mathrm{p} = 0.12\,\mathrm{ps}$) and a long ($\tau_\mathrm{p} = 10\,\mathrm{ps}$) laser pulse imposed by the relation

$$\Delta E \cdot \tau_\mathrm{p} \geqslant \hbar\,. \tag{23.13}$$

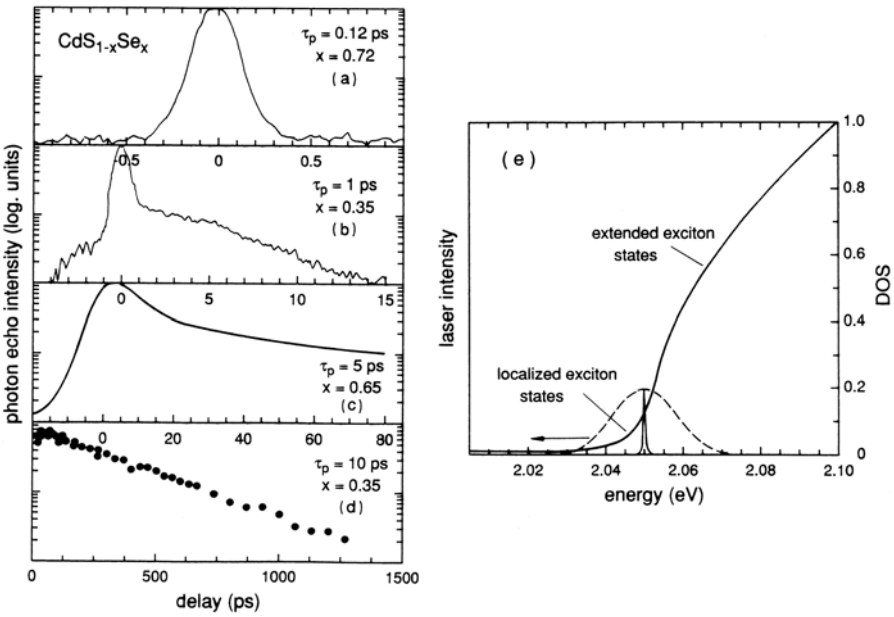


Fig. 23.8. The decay of the time-integrated photon-echo as a function of the delay between the two pump pulses for various durations of the pump laser (**a**–**d**) and the density of localized states and the minimum spectral width of the laser pulses for $\tau_\mathrm{laser} = 10\,\mathrm{ps}$ and $\tau_\mathrm{laser} = 0.12\,\mathrm{ps}$, respectively (**e**) [92S1, 99S3]

A spectrally narrow, rather long pump pulse tuned to the spectral range of localized states allows one to measure the true T_2 of these states. If the pulse length is reduced, the high energy tail of the laser extends into the region of extended states. In alloy semiconductors even at low temperatures these states have a very short T_2 time of the order of 0.1 ps due to alloy disorder scattering [91S1, 92S2]. This value explains the short autocorrelation spike. Furthermore, the excitons in the extended states which are created by a short, spectrally broad pulse also scatter with the excitons in the localized states and thus reduce the T_2 values of the latter. Consequently the T_2 values obtained with rather long, spectrally narrow pulses with ΔE from (23.14) much smaller than the tailing parameter ε_0

$$\Delta E \ll \varepsilon_0 \tag{23.14}$$

will give the most reliable results. See also the discussion in connection with Fig. 23.11.

The spectral proximity of states with different T_2 times occurs not only for disordered systems and necessitates some considerations concerning the selection of a laser for a planned experiment.

It must be mentioned here that short phase relaxation times have been observed in other disordered systems like $Al_{1-y}Ga_yAs$ crystals, amorphous (α-) Si, and InGaAs MQW [91W1, 92S2]. The T_2 values at low temperature are only around or below a few ps. The difference is presently explained as follows. If the disorder causes fluctuations essentially only in one band (in $CdS_{1-x}Se_x$ this is the valence band originating mainly from the 3p and 4p orbitals of S^2 and Se^2, respectively) a stationary state with long T_2 time can be constructed from a localized carrier of one type (here a hole) binding a free carrier of the other type by Coulomb attraction to form a localized exciton. If, on the other hand, fluctuations and localization occur in both bands, as is the case for the more covalently bound materials $Al_{1-y}Ga_yAs$ and α-Si, it is not possible to get a stationary state with two localized carriers plus Coulomb interaction. Here alloy disorder scattering always appears as a phase-destroying process. More details about these topics are found in [92S2] and references therein. Furthermore, the short pulses used in these experiments may play a role in the sense of the discussion in connection with Fig. 23.8 and with data for quantum wells cited below. Before we present results on structures of reduced dimensionality we introduce another aspect of FWM experiments namely (quantum) beats.

In this case a short laser pulse of sufficient spectral width simultaneously excites coherently two spectrally narrow, well-separated resonances. This means the spectral width of the laser ΔE, the spectral separation of the two resonances ΔE_r and the homogenous width of each of them ΔE_h should obey the inequality

$$\Delta E \gtrsim \Delta E_\mathrm{r} \gg \Delta E_\mathrm{h}\,. \tag{23.15}$$

Under these conditions a beating occurs, as known already from two coupled pendula in classical mechanics, with a beat frequency ω_b

$$\hbar\omega_\mathrm{b} = \Delta E_\mathrm{r} \,. \tag{23.16}$$

In our case a modulation of the intensity with ω_b can be observed, e.g., in a FWM experiment or in the free polarization decay signal. In Fig. 23.9 we give examples of the beating between the $n_z = 1$ heavy- and light-hole exciton resonances (a, b) and the beating between the exciton resonance and the induced transition exciton $\rightarrow$ biexciton (c, d). Since the beating is a coherent process, the decay of the envelope of the beat signal contains information on the T_2 of the resonance with the lower T_2 value.

The beat frequency just fulfils (23.17). These quantum beats have recently been observed in many systems, such as excitons in the indirect gap semiconductor AgBr [91S1], between strain-split heavy and light hole excitons in ZnSe layers [96W1], in bound exciton complexes in CdS and CdSe [91S1, 92P1, 93P1, 93P2, 94E1, 94P1, 96B1] and even between the lower and upper polariton branches of the lowest dipole-forbidden $n_\mathrm{B} = 1$ exciton resonance in Cu_2O, which has a longitudinal transverse splitting of

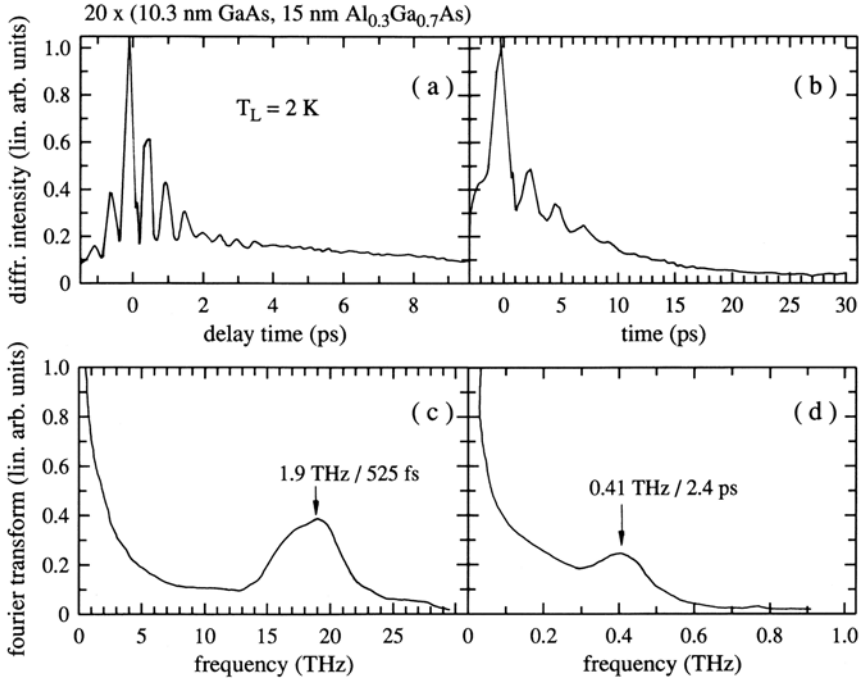


Fig. 23.9. Quantum beats and their Fourier transforms for the beating between the heavy- and light-hole exciton in an $Al_{1-y}Ga_yAs/GaAs$ MQW (**a**, **c**) sample and for the exciton $\rightarrow$ biexciton transition (**b**, **d**) [92O1, 93P1]

a few μeV only [91F1]. Other experiments concern the beating between various polariton branches in CdSe [93P2, 94E1, 96B1] or in InGaAs MQW [91W1] or between the exciton and the biexciton state in AlGaAs MQW; see [92O1, 93P1, 93P2, 94E1, 96B1] and references therein.

Recently it has been possible to distinguish between a beating of two electronic transitions and a mere interference of the electromagnetic waves radiated from two independent, slightly detuned oscillators [91B1, 92G1, 92K1, 92P1, 94P1].

For a "real" quantum beat, i.e., a beating between to spectrally close-lying transitions that have at least one state in common or, in other words, that occur in one quantum mechanical system, one expects maxima in a plot of the FWM intensity over a two-dimensional plane of real time t and delay time τ for

$$T = \tau + nT_{\mathrm{B}} \tag{23.17a}$$

where T_{B} is the beat period. Meanwhile, the interference of the radiation emitted from two independent systems on the detector results in

$$t = 2\tau + n_{\mathrm{B}}\,. \tag{23.17b}$$

In Fig. 23.10 examples are given for both situations. In Fig. 23.10a one observes upconverted and time resolved quantum beats between the light and

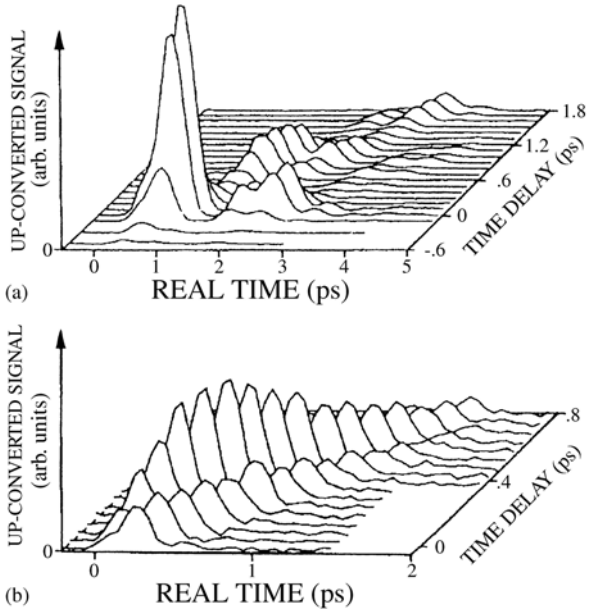


Fig. 23.10. Quantum beats between *hh* and *lh* exciton in a MQW sample (**a**) Polarization interference from the *hh* resonance of a sample containing quantum wells of two different thicknesses [92G1, 92K1]

heavy hole over a two-dimensional t, τ plane. Figure 23.10b gives, in contrast, the polarization interference of the hh excitons of spatially separated QWs with $l_z = 9\,\text{nm}$ and 8 nm. The main maxima follow the relations (23.18a) and (23.18b), respectively.

This picture is very clear, but not completely true, since the whole semiconductor sample has one common ground state and since even spatially separated oscillating dipoles have some, though possibly weak, radiative coupling. This coupling is, in a classical description, responsible for the Clausius–Massotti correction to the dielectric function (see Sect. 4.2).

Consequently, in calculated or measured data that show a pronounced behavior according to (23.17b), one usually observes weaker intensity modulations obeying (23.17a) and vice versa.

In the picture of strong exciton–photon coupling, a generalized description of the quantum beats can be expected that includes both of the abovementioned aspects as limiting cases.

In Fig. 23.11 we present a FWM experiment on a GaAs/$Al_{1-y}Ga_yAs$ sample containing wells of different thicknesses, as shown in the absorption spectrum.

The laser pulse is so short that it covers spectrally several transitions. For the 8 nm QW only the *hh* is excited. One observes the rise and the slow signal decay of predominantly localized excitons. As compiled, e.g., in [01L1] the T_2 times range at low temperatures up to several tens of ps. For the 10 nm QW the *hh* and the *lh* exciton are excited simultaneously, resulting in quantum beat features and again a slow signal decay. For the 13 and 16 nm-wide wells,

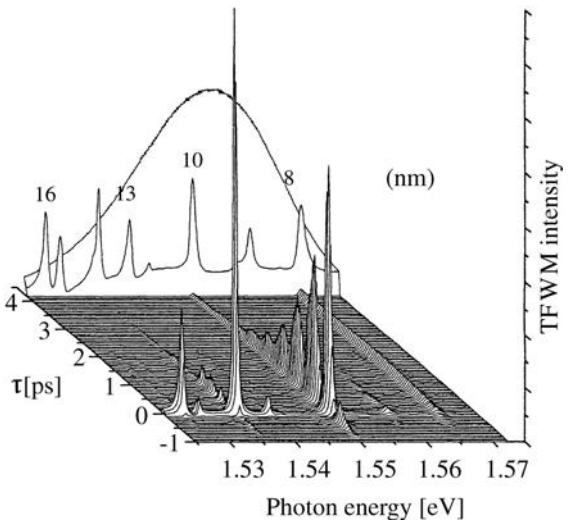


Fig. 23.11. Time integrated FWM spectra as a function of time delay for a sample containing quantum wells of various thicknesses as indicated in the absorption spectrum [97E1, 98H1]

both the *hh* and the *lh* excitons are excited as well as the continuum states with even higher intensity. As shown in [97E1] this has the consequence that a FWM signal appears spectrally at the position of the exciton, but first with a rapid decay given by the inverse of the spectral width of the laser pulse rather than by the T_2 time of the exciton. Note that the spectral widths of the exciton absorption features are not compatible with the short decay times via (23.4). After this first decay, a much weaker, but slowly decaying signal is observed that gives the actual T_2 time. It is possible that this slow and weak part of the signal was partly below the detection limit of the experiments described in connection with Fig. 23.8.

In [01W1] the FWM signal has been analyzed as a function of delay time between the pulses and the real time after the second pulse, revealing details of the FWM process, which go beyond both the simple explanations given here and the scope of this book.

Recently, beautiful experiments have been reported on FWM on a graded superlattice, i.e., one where the well width varied systematically from well to well. By spectral filtering of the laser, excitons in adjacent wells or next-nearest neighbors could be excited deliberately and the beating between the resonances in one well as well as the interactions between wells could be observed in the spectrally resolved, time-integrated signal as a function of the delay time between the pulses [04L1].

Now we give a small selection of further results of FWM in various quantum structures: The dephasing in the quasi-two-dimensional exciton-biexciton system of $GaAs/Al_{0.3}Ga_{0.7}As$ single quantum wells has been treated in [00L1] and for $In_{0.18}GA_{0.82}As$ SQW in [99B2]. The contributions of acoustic and optic phonons to the dephasing have been analyzed and separated in [99B1] for $In_{1-y}Ga_yAs/GaAs$ SQW.

The oscillations of wave packets in space and time in $In_{1-y}Ga_yAs/GaAs_{1-x}P_x$ MQW samples are investigated in [99O1] and the influence of strong magnetic fields on excitons and biexcitons has been detailed in [94B1] and references therein.

For data on exciton and biexciton dephasing in various types of III–V and II–VI quantum wires see, e.g., [98W2, 98W3, 99L1] and references therein.

In self-assembled III–V quantum islands, very long dephasing times have been measured for excitons, biexcitons and multiexcitons for the reasons given above for localized states [01B1, 01B2, 02B1, 02B3, 04L2]. For localized exciton and biexciton states in II–VI structures, data are found in [97L1, 99W1], where it has been shown that an inhomogeneous broadening of the biexciton binding energy results in a fast decay of the FWM signal that is not related to the much longer dephasing of the biexciton.

Excitons in single dots are probed in [98B1, 01F1, 01F2]. Data for the dephasing in electrically pumped $InAs/In_{1-y}Ga_yAs$ amplifiers are given in [00B2].

More data for III–V compounds can be found, e.g., in [01L1].

It should be noted that the observation of higher diffracted orders than the first one in FWM experiments are indications of $\chi^{(2n+1)}$ processes with $n > 1$.

In $CdS_{1-x}Se_x$ it was, e.g., easy to observe three diffraction orders [89D1,89D2, 91S2], but these higher orders also contribute to the properties of the first diffracted orders. Recently $\chi^{(5)}$ and $\chi^{(7)}$ processes have been investigated in GaAs-based and ZnSe-based quantum wells [01A1, 01L2].

We move on now to a third experimental technique to determine T_2 which is again connected with FWM and laser-induced gratings. This is nondegenerate four-wave mixing (NDFWM), and the system to which we apply it are quantum dots in glass matrix. It is a method to determine T_2 from spectroscopy, i.e., in the frequency rather than in the time domain.

In this case one writes a grating with two laser beams of different frequencies ω_1 and ω_2. The coherence time of each laser must be long enough to provide a sufficiently coherent overlap between them. One generally uses ns dye lasers with an additional intracavity etalon to reduce their linewidths. In the polariton picture one would describe the diffraction process in the following way: Two quanta of beam 1 (or 2) create an intermediate two-polariton state, the decay of which is stimulated by a polariton from beam 2 (or 1) resulting in the emission of a quantum in the direction

$$2\boldsymbol{k}_1 - \boldsymbol{k}_2 \quad (\text{or} \quad 2\boldsymbol{k}_2 - \boldsymbol{k}_1) \tag{23.18a}$$

with energies

$$\hbar(2\omega_1 - \omega_2) \quad (\text{or} \quad \hbar(2\omega_2 - \omega_1))\,, \tag{23.18b}$$

as shown schematically in Fig. 23.12. Higher diffracted orders correspond to higher $\chi^{(n)}$ processes and show frequency shifts which are integer multiples of $(\omega_1 - \omega_2)$.

In the weak coupling limit, the first order corresponds to a process of the type (see Sect. 19.3)

$$\chi^{(3)}\,(\omega : \omega_1, \omega_1, -\omega_2)\ . \tag{23.18c}$$

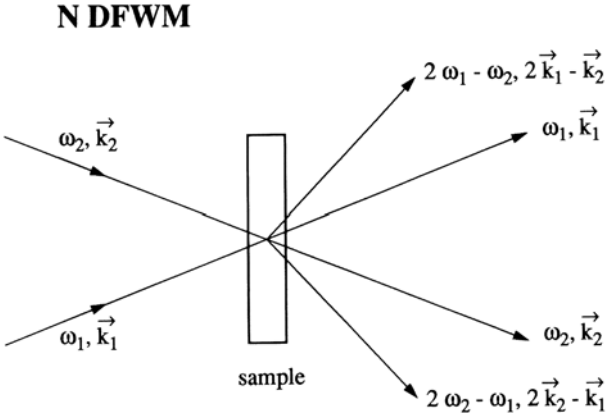


Fig. 23.12. The schematics of a non-degenerate four-wave mixing (NDFWM) experiment

In a classical wave picture we have the following description. The two incident waves interfere to form a moving interference pattern. It moves laterally with a speed v_G given by

$$v_G = \frac{\omega_1 - \omega_2}{|\boldsymbol{k}_1 - \boldsymbol{k}_2|} \,. \tag{23.19}$$

The orders diffracted off this moving grating are Doppler shifted and a quantitative analysis of these ideas leads to exactly the same results as given in (23.18).

If there is an optical nonlinearity due to the real or virtual excitation of some species with a lifetime T_1 or T_2, respectively, then the grating will be washed out if it moves over one period Λ in a time shorter than $T_{1,2}$. Consequently the efficiency drops, when

$$\Lambda v_G^{-1} < T_1 \quad \text{or} \quad T_2 \,. \tag{23.20}$$

A quantitative analysis leads to the following relation for the efficiency $\eta_{\pm 1}$ of the $\pm$ first diffracted order [78Y1].

$$\eta_{\pm 1} \propto \left|\chi^{(3)}\right|^2 \propto \left[\left(1 + \Omega^2 T_1^2\right)\left(1 + \Omega^2 T_2^2\right)\right]^{-1} \tag{23.21}$$

with $\Omega = \omega_1 - \omega_2$. This means that there is a narrow central spike, whose decay with increasing detuning Ω is determined by T_1 and wider wings from which one can deduce T_2 in the case $T_1 \gg T_2$.

In Fig. 23.13 we give as an example results for a semiconductor-doped glass with CdS crystallites with an average radius of 7.5 nm. The narrow central

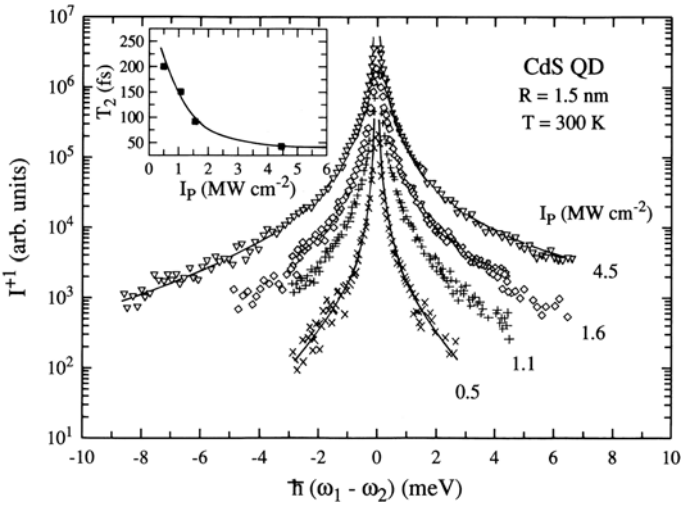


Fig. 23.13. The NDFWM signal as a function of the detuning in a silicate glass doped with CdS nanocrystals for various pump powers I_p [94W1]

spike is limited rather by the spectral width of the lasers than by T_1. The T_1 time of a few hundred ps typical for such a system was measured independently by time-resolved luminescence and would give a width below 0.1 meV. The wider wing can be fitted with (23.21) to give a T_2 time of the resonantly excited electron–hole pairs in these quantum dots of about 100 fs, decreasing with increasing pump power as shown in the inset. These values increase only slightly when going to lower temperatures. The reason why this value is so short in comparison with bound exciton complexes is not completely clear at present. The stronger coupling to the optical phonons in QD compared to the bulk material may be of importance or the scattering at, or the relaxation of the carriers into, some interface states between the semiconductor and the surrounding matrix.

Similar values of T_2 have been obtained from spectral hole burning (see below) and also from time-resolved FWM experiments. It should be mentioned, however, that much longer dephasing times of several tens of ps have been observed by various techniques for quantum dots capped by or embedded in organic matrices [97W1].

23.2.1.2 Other Techniques and Coherent Processes

In the following sections we briefly present some further techniques to determine dephasing times and some other coherent phenomena including, e.g., the spin system, to give the reader an impression of the width and wealth of the field and present research activities.

23.2.1.2.1 Propagation Effects

An inherent problem in time-resolved four-wave mixing experiments is the fact that the pulses have to have a certain time delay. In bulk material the polarization propagates together with the electromagnetic part of the light pulse as introduced in the polariton concept in Chap. 5. If the pulses are short, their shape in the samples is pancake-like. A 100 fs pulse may have, depending on the parameters for group and phase velocities, a thickness of only a few μm with some dispersion during propagation through the sample. With a delay between the pulses of several ps the spatio-temporal overlap of the two pulses is an issue, which must be carefully checked and depends on the angle between the beams $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ in the sample and on the lateral diameter of the pulses. This is a reason why little FWM data exists in the exciton resonances of direct-gap bulk semiconductors with dipole-allowed band-to-band transitions. Another reason is the strong absorption in the exciton resonance of many semiconductors. A way out also used in Fig. 23.5a is to measure the reflection geometry since the reflection signals originate from the polarization close to the surface.

In contrast, for quantum structures the polarization stays essentially confined in the quantum well, wire or dot for the usual close to normal incidence geometry, except for some polarization or radiation coupling between

the wells. The spatial coincidence of the polarization created by subsequent pulses is therefore unproblematic.

An adequate method for bulk semiconductors would therefore be to investigate the polariton propagation. This can and has been done. We already presented corresponding data for various bulk materials for the case of propagation quantum beats and for time of flight experiments in Sects. 5.3 and 13.1.4, respectively. We add here some references for a thin ZnSe layer [96N1] for the layered semiconductor InSe [97N1] and for photonic crystal wave guides [02N1]. We come back on propagation effects under the aspect of transport properties in Sect. 23.2.1.2.5 and 2.3.3.

23.2.1.2.2 Rayleigh Scattering and Speckle Analysis

An effect that has been considered for a long time as more disturbing than useful in semiconductor spectroscopy is Rayleigh scattering. This means coherent and elastic scattering of light without a frequency shift in contrast to Brillouin or Raman scattering. Rayleigh scattering requires some deviations from the strict periodicity of a perfect crystal lattice, which are generally always present even in high quality samples due to, e.g., defects, strain fields or interface roughness for quantum structures.

In pioneering work summarized in [94S1] it has been shown that the dynamics of resonant Rayleigh scattering involved information on the dephasing of, e.g., bound exciton complexes, confirming the long T_2-times mentioned above. For examples of more recent work see [98B2, 99S1, 04M1, 04M2, 04Z1].

The light scattered resonantly and coherently from various regions of the sample can interfere constructively or destructively depending on the scattering direction. This effect leads for rather small illuminated areas to the appearance of bright and dark spots in the scattered light field, know as speckle.

The dependence of the speckles on the direction and their temporal dynamics under pulsed excitation allow one to distinguish pure dephasing from lifetime (see (23.3b)) [99L2,02K1,03M1] or to investigate polaronic effects [00P1] and the references therein.

23.2.1.2.3 Linear Spectroscopy

If one has no further information it is difficult to learn something about the dephasing time T_2 of a system from linear absorption, reflection or luminescence spectra.

If one does not know, e.g., if an absorption band is only homogeneously or additionally inhomogeneously broadened, its spectral width gives only a lower limit for the decoherence time T_2 or for the time of the free polarization decay T_{PD} via (23.4) and (23.5).

However, the spectrally extremely narrow emission lines from individual localization sites of excitons in quantum wells observed at low temperatures or the narrow emission features of bound exciton complexes, which we have

seen in some of the preceding chapters, indicate that the T_2 times must be in the range of several tens to hundreds of ps as is indeed found from photon echoes as shown above. The temperature dependence of the spectral half width of the emission line of excitons in self-assembled $In_{0.6}Ga_{0.4}As$ islands grows from values of 2 μeV ($T_2 > 0.7$ ns) to almost 10 meV at RT ($T_2 \approx 0.13$ ps) [02B2, 02B3, 04L2].

The reflection spectra of free exciton resonances of high-quality bulk samples (see, e.g., Chap. 13) can be fitted at low temperatures usually by damping parameters around or below 1 meV. Therefore T_2 times measured by FWM in the range of some ps (see above) did not come as a big surprise [04H1].

23.2.1.2.4 Spectral Hole Burning and Related Techniques

A further technique to determine the homogeneous width of a transition is spectral hole burning. This method is frequently used in the spectroscopy of atoms in gases, solutions or solids but has only more recently been used in semiconductor optics [77L1, 91D1, 93W1, 94M2, 94W1]. We explain shortly the basic idea before going to examples.

A homogeneously broadened absorption band has a Lorentzian line shape for the simplest case of a two-level system (see Fig. 23.14a). Its width is given by (23.4). Pumping at any energy $\hbar\omega_P$ results in a decrease of the whole absorption feature if a certain fraction of the carriers is transferred from the ground to the excited state.

For an inhomogeneously broadened system (see Fig. 23.14b) the absorption band is composed of different transitions as shown schematically on the r.h.s. of this figure. If every single transition has a Lorentzian profile, and if the inhomogeneous distribution is Gaussian, the total line shape is a Voigt profile as long as both widths are comparable.

Pumping at $\hbar\omega_P$ saturates the transition at this energy, but does not influence the others, resulting in a spectral hole. If the pump laser is spectrally narrower than the homogeneous width of every individual absorption peak in the inhomogeneous broadened distribution, the width of the spectral hole gives the homogeneous width according to (23.4).

Care has to be taken that the laser is spectrally sufficiently narrow and sufficiently weak to avoid the spectral wings saturating adjacent transitions. This effect is known as power broadening and results in an overestimation at the homogeneous width (for more details see, e.g., [94M2]).

Examples of spectral hole burning are experiments on $CdS_{1-x}Se_x$ quantum dots in glass matrices, revealing homogeneous widths on the order of 10 meV [93W1, 93W2, 97W1], corresponding to T_2 times of about 100 fs, in agreement with the NDFWM data of Fig. 23.13 or time resolved FWM [93S1, 94M1]. Note that the dots in organic matrices may show much longer T_2 times [96E1, 97W1] as mentioned already above.

Other examples are the spectral hole burning experiments in the gain region of an inverted bulk semiconductor and of quantum dots [93M1, 96G1, 98M1].

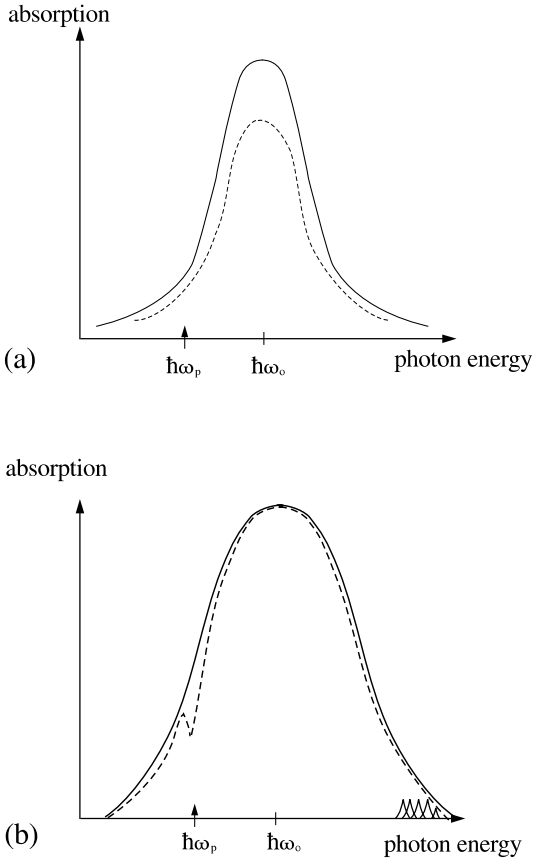


Fig. 23.14. Schematic drawing of the absorption of a homogeneously (**a**) and an inhomogeneously (**b**) broadened absorption band centered around $\hbar\omega_0$ and the modification induced by optical pumping at $\hbar\omega_P$ (*solid* and *dashed lines*, respectively)

A similar technique has been used in [02W1] for ZnSe quantum wells, where it has been explicitly shown that the homogeneous width of the exciton increases with increasing excitation intensity and that the data deduced from the fit of a density-dependent Voigt profile and from FWM coincide.

For spectral hole burning and bleaching dynamics in the semiconductor Sr_2CuO_3, which is related to high T_c superconductors see [00O1].

23.2.1.2.5 Transport Measurements

In transport measurements, another time constant appears in the frame of a relaxation approach, namely the momentum relaxation time T_2', i.e., the time of free flight between random collisions. We give an example of how nonlinear optics can contribute to the determination of transport properties.

In Fig. 23.15 we compare the results for phase-destroying processes measured with two different techniques. The crosses are data for the scattering rate, i.e., for T_2^{-1}, obtained from photon-echo experiments. They start with values around 30 ps at 2 K and T_2 decreases with increasing temperature essentially due to scattering with (acoustic) phonons in the temperature interval of Fig. 23.15. The open circles are obtained from diffusion measurements. The idea of this experiment is the following: two coherent pulses falling simultaneously on the sample interfere and produce a spatially modulated population grating (laser-induced grating, LIG). If this population produces any optical nonlinearity, light will be diffracted from this grating. The grating efficiency decays for times longer than T_2 due to recombination and due to diffusion. The first process does not depend on the grating period Λ of the LIG, but the second one does. By variation of Λ the two contributions can be separated. If two short pulses ($\tau_\mathrm{p} \ll T_1$) are used, the diffracted signal intensity, which is measured by a third delayed pulse, decays with a characteristic time τ_s given by

$$\frac{1}{\tau_\mathrm{s}} = \frac{2}{T_1} + \frac{8\pi^2 D}{\Lambda^2}\,, \tag{23.22}$$

which allows the determination of D and via (23.11d) of T_2'. The values given by open circles in Fig. 23.15 have been obtained in this way.

As a short aside, we mention here the following alternate method: For quasistationary conditions (i.e., $\tau_\mathrm{p} > T_1$), one can state as a rule of thumb that the grating efficiency (i.e., the intensity of the diffracted order) decays with varying Λ to one half for $\Lambda_{1/2}$ (see e.g. [88W1])

$$\Lambda_{1/2} = l_\mathrm{D} 2\pi\,, \tag{23.23}$$

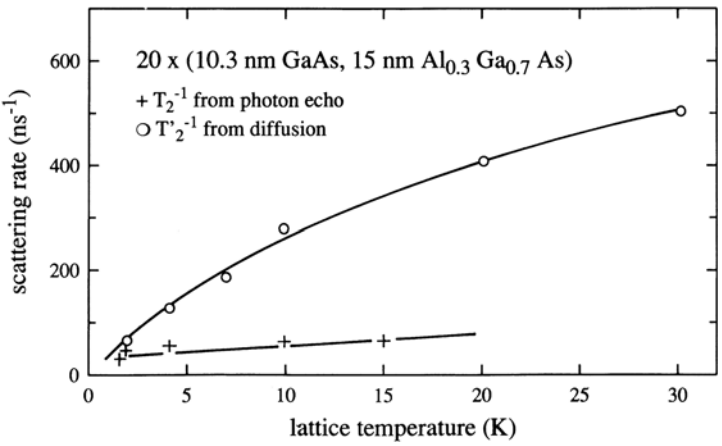


Fig. 23.15. The inverse T_2 times deduced by photon echo experiments and the inverse momentum relaxation times T_2' or scattering rates deduced from diffusion experiments for $Al_{1-y}Ga_y$, As MQW samples. From [93O1]

where l_D is the diffusion length of the relevant particles, i.e., excitons. This quantity is connected with the diffusion constant D via

$$l_D \approx (DT_1)^{1/2} \tag{23.24}$$

and (23.11d).

We return now to the ps experiments.

As can be seen in Fig. 23.15, the values of T_2 and T_2' coincide approximately for T around 5 K. For higher temperatures T_2' becomes significantly shorter than T_2.

The reason is that at least one additional scattering process contributes to T_2, which is relevant for transport properties and which is absent or much smaller in photon-echo experiments. One of these processes is the interface roughness scattering in the quantum wells. The scattering rate increases with the (thermal) velocity v_{th} of the excitons, i.e., with $T^{1/2}$. In photon-echo experiments the excitons are created optically and have only extremely small in-plane velocities. The excitons acquire thermal equilibrium with the lattice and considerably higher velocities after one (or a few) scattering processes with phonons, but then they no longer contribute to the coherent photon-echo signal. Thermalization with the lattice means in this case evidently an increase of energy and momentum of the optically created excitons, as discussed in Sect. 22.1. The solid curves in Fig. 23.15 have been calculated without and with inclusion of interface roughness scattering, respectively. For more details see [92O1, 93O1].

For more recent quasiballistic transport measurements of excitons in ZnSe quantum wells and the dynamic properties deduced from them see, e.g., [02Z1, 04Z2].

23.2.1.2.6 Bloch Oscillations and Terahertz Spectroscopy

In Sect. 16.2.2 we mentioned the Wannier Stark Ladder appearing in superlattices under the influence of an electric field. We treat this topic here, together with the Bloch oscillations and add some further aspects of THz spectroscopy because it involves many aspects of ultrafast spectroscopy.

The basic idea is rather old and goes back to the 1930s [28B1, 34Z1, 65S1]. We shall first introduce the concept of Bloch oscillations, explain why they have been observed only recently, and then present various approaches to understand the phenomenon and give experimental data. The presentation here is essentially based on various recent review articles [92F1, 92L1, 94L1, 94R1, 94V1, 02L1]. Then we outline some analogies and differences between the concept of Bloch oscillations in the presence of an electric field and the cyclotron resonance in the presence of a magnetic field. For the latter topic see also Sect. 16.1. We finish this subsection with some references of THz spectroscopy.

The basic concept of Bloch oscillations is based on the motion of a wave packet centered around $\boldsymbol{k}$ in a periodic lattice under the influence of an exter-

nal force $\boldsymbol{F}$, e.g., an electric field $\boldsymbol{E}$, neglecting for the moment all scattering processes; see [28B1] or Sect. 8.5

$$\hbar\dot{\boldsymbol{k}} = \boldsymbol{F} = e\boldsymbol{E}\,. \tag{23.25}$$

Under the above conditions the electron will perform a periodic orbit in $\boldsymbol{k}$-space in the direction of $\boldsymbol{E}$ as shown schematically in Fig. 23.16.

The carrier starts, e.g., at point A at rest, and is accelerated by the external field moving thus towards B. At B the sign of the effective mass changes and the carrier is decelerated by the field but still increasing its momentum and energy according to (23.25) and reaching the right-hand edge of the first Brillouin zone at point C with zero group velocity. Instead of considering wave vectors beyond the first Brillouin zone, we can assume that the electron is transferred by a reciprocal lattice vector $\boldsymbol{G}$ from π/a to $-\pi/a$, i.e., to point D; see Sects. 7.2 and 8.1. From D, the wave packet is accelerated until it reaches point E and decelerated from E to A reaching point A again with zero group velocity. The duration of one orbit in $\boldsymbol{k}$-space T_B and the frequency ω_B are given by [92F1]

$$\omega_\mathrm{B} = 2\pi/T_\mathrm{B} = ae\,|\boldsymbol{E}|\,\hbar^{-1}\,. \tag{23.26}$$

This periodic motion in $\boldsymbol{k}$-space is necessarily connected with a periodic motion in real space in the direction of $\boldsymbol{E}$ with an amplitude l_B [92F1, 02L1]:

$$l_\mathrm{B} = B/2e\,|\boldsymbol{E}|\ , \tag{23.27}$$

where B is the width of the band. The periodic motions in real and in momentum space are known as Bloch oscillations or as "Zener Pendeln" [28B1, 34Z1, 65S1].

This phenomenon has been used to explain (partly erroneously) the I–V characterisitcs of Zener diodes in the blocking direction. It was assumed that the electrons in the valence band under the influence of the blocking voltage

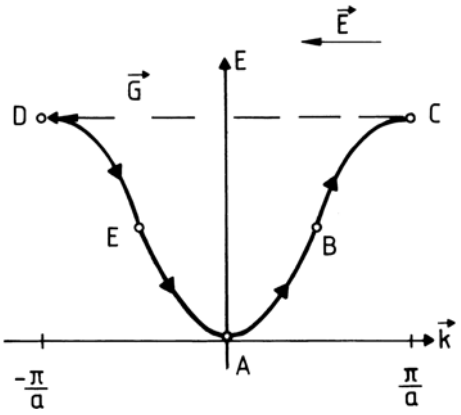


Fig. 23.16. Schematic presentation of the Bloch oscillation or "Zener Pendeln" of a charge carrier in an electric field in the absence of scattering

over the pn junction execute such a periodic orbit, and tunnel with a certain probability to the conduction band whenever they hit the top of the valence band.

Actually, the carriers in bulk semiconductors have no chance to complete such an orbit in a partly occupied band (in a completely filled band, there is no net motion at all due to Pauli's principle), since the T_2 times are too short. Even at low temperatures the T_2 values at the bottom of the band are only a few tens of ps and they tend towards ten fs deeper in the band; see Sect. 23.2.1. In order to fulfill the condition

$$\omega_{\mathrm{B}} \cdot T_2 \geqslant 1 \tag{23.28}$$

which would be necessary to be able to observe Bloch oscillations, one would need with (23.26) dc electric fields in excess of $10^7\,\mathrm{V/m}$ which would result in rapid destruction of the sample.

The way out of this dilemma follows from an inspection of (23.26). For a given T_2 an increase of the spatial period of the system results in a decrease of the electric field strength $\boldsymbol{E}$ necessary to fulfill (23.28). Since the lattice constant a of all simple semiconductors is comparable to the atomic diameters and thus of the order of 0.5 nm only the use of man-made artificial superlattices can help. With periods d around 10 nm we can expect a reduction of the field strength by more than an order of magnitude. In connection with Fig. 23.17, we recall and extend what we have learned about superlattices in Sect. 8.10.

The solid line in Fig. 23.17 gives the lower edge of the conduction band of the two materials forming the superlattice. The width of the barriers is so narrow that the electron wavefunctions of adjacent wells overlap. This overlap integral results in a finite band width (Sects. 5.4 and 8.1). The width of these minibands is directly proportional to the overlap integral and can thus be varied in this Kronig–Penney-like system by the thickness of the barriers. Typical widths of the minibands are of the order of a few meV, in contrast to about 10 eV for the width of the bands in bulk material. Note that the dispersion $E(\boldsymbol{k})$ and the motion of the carriers in the plane normal to the growth direction is still almost bulk-like. The formation of minibands in the

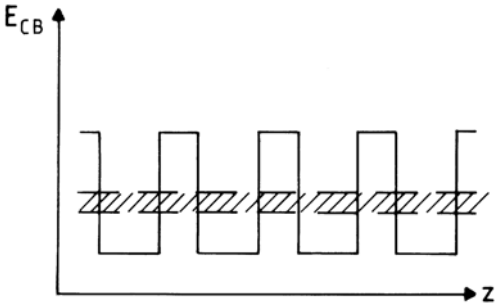


Fig. 23.17. The conduction-band edge of a superlattice and the resulting miniband for $k_x = k_y = 0$ indicated schematically by the *hatched area*

valence band is generally much less pronounced, due to the greater effective masses of the holes, which reduces the tunneling into the barriers and consequently the overlap integral. The appearance of the minibands in the conduction bands also modifies the density of states in the effective-mass approximation from the step-like function of a two-dimensional system to the one given in Fig. 8.26e.

Since superlattices usually have some ten to one hundred periods, the number of different $\boldsymbol{k}$-states in the growth direction in the interval from $-\pi/d$ to $+\pi/d$ is also some ten to one hundred, in contrast to values above 10^7 for a bulk crystal with a length of about one centimeter.

If such a superlattice is incorporated into a pn junction biased in the blocking direction, electric fields can be applied in the growth direction which are sufficient to cause optically excited carriers to perform Bloch oscillations with frequencies ω_B which fulfil (23.28).

The experimental observation of these oscillations can be accomplished by time-resolved four-wave mixing in the way discussed with Figs. 23.2 or 23.9. There is an overall temporal decay of the diffracted signal intensity as a function of the delay between the two pump pulses, which is governed essentially by the dephasing time T_2, but this decaying signal is modulated with the frequency ω_B. A rather early example for such an experiment is given in Fig. 23.18.

The frequency of the modulation varies in this type of experiments linearly with the applied field according to (23.26) giving support to the underlying concepts.

In a more elaborate model, detailed in [92F1, 94L1], the occurrence of the oscillations can be understood along the following lines. A pump pulse of duration of about 100 fs has a (Fourier-limited) spectral width comparable to, or larger than, the width of the minibands. It thus excites coherently all the several ten to one hundred band-to-band transitions in the miniband.

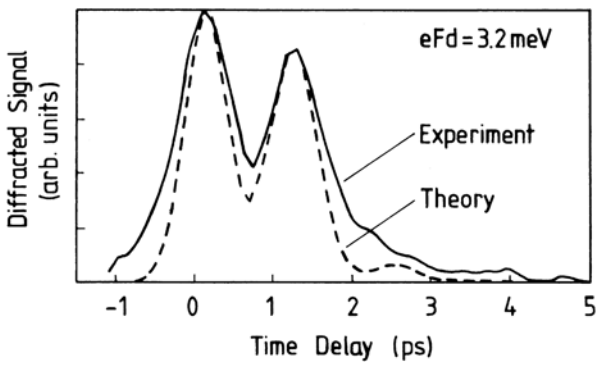


Fig. 23.18. The temporal evolution of the FWM signal in a $\mathrm{GaAs/Al_{1-y}Ga_yAs}$ superlattice with a period of 11 nm under the influence of an electric field of 0.3×10^7 V/m [92F1]

Since each of these transitions has a different frequency, the polarizations connected with the transitions run out of phase with respect to each other. It can be shown, however, that the influence of an external field is such that the polarization of all these transitions rephase again after integer multiples of T_B. Consequently the signal peaks for delay times of the two pump beams which are integer multiples of T_B.

This picture has another consequence: If there is a macroscopic polarization in the medium, which oscillates with frequency ω_B, this frequency must be radiated. Indeed it was recently possible to observe the THz emission at frequency ω_B direclty [94R1,94V1] by the use of a micro-dipole antenna which is gated in a sampling technique by a part of the 100 fs pump pulse. The experimental setup is shown in Fig. 23.19. In Fig. 23.20 the signal of the antenna is shown as a function of the delay time of the gate pulse.

The temperature of the GaAs/$Al_{1-y}Ga_yAs$ superlattice with a period of 11.4 nm was 10 K resulting in a reasonably long T_2 time. Beautiful oscillations with frequency ω_B can be seen. With increasing temperature, the decay of the envelope gets faster due to a decrease of T_2. More details of this experiment are given in [94R1,94V1].

Before we proceed to have a look at Bloch oscillations from another point of view we make the following statement. Until now we presented the Bloch oscillations in the picture of free carriers. This is the original approach and it is easily intelligible from the didactic point of view. Actu-

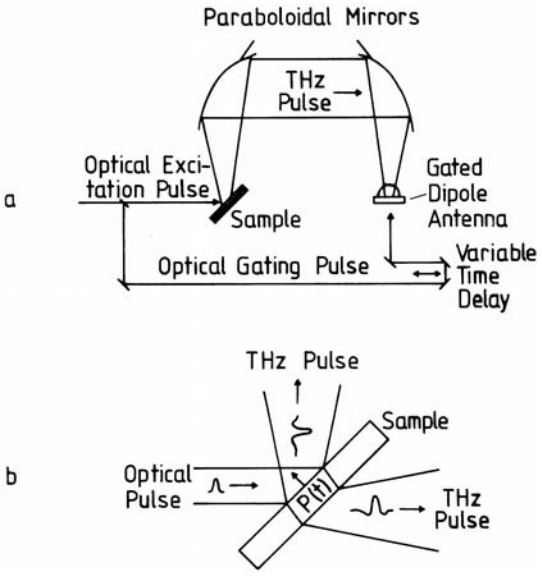


Fig. 23.19. An experimental setup to observe the Bloch oscillation directly. Schematic overview of the experiment (**a**) and of the emission of the THz pulse at ω_B (**b**) [94R1]

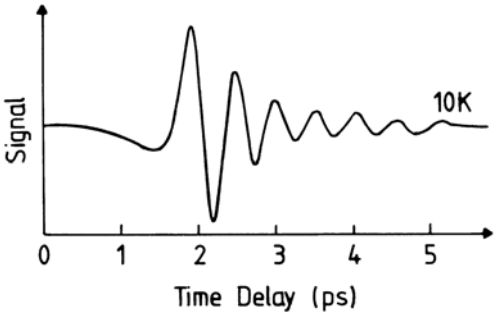


Fig. 23.20. Electric field at the antenna as a function of the delay of the gate pulse [94R1]

ally we should remember, however, that by optical pumping, electron–hole pairs or excitons are created. At a first glance it seems difficult to understand the concept of Bloch oscillations for the neutral quasi-particle exciton. However, it can be shown theoretically and experimentally [97L2, 00L3, 03D2, 03H1] that the frequency ω_{B} survives when the exciton wave function is constructed from a superposition of electron and hole wave packets. This statement is a good example for the motto given at the beginning of this book.

More recent aspects of Bloch oscillations include, e.g., the direct observation of the spatial motion of the carriers according to (23.27) [97L2, 98S1, 02L1], the coupling to phonons or plasmons [91D1, 01L2], the appearance of a dc current [00L2, 00L3] or Zener breakdown [01R1].

Now we come back to the formation of the minibands. They appear, as already stated, from the overlap of the wavefunctions in the various quantum wells which would have all the same energy without this overlap in the some way as the bands appear from atomic orbitals in the bulk (Fig. 8.1).

If we apply a constant electric field along the structure of Fig. 23.17 the whole structure is tilted (Fig. 23.21), and the eigenstates of the carriers in the various wells no longer have the same eigenenergy.

As a consequence, a miniband is no longer formed and the wavefunctions, which are extended and of the Bloch-type without electric field, become localized in the growth direction extending only over a few periods of the superlattice. They have discrete energy levels.

$$E_n = E_0 + ne\,|\boldsymbol{E}|\,d, \quad \text{with} \quad n = 0, \pm 1, \pm 2, \ldots \tag{23.29}$$

This picture is especially appropriate for electric fields, which produce a potential drop over one superlattice period $e|\boldsymbol{E}|d$ which is comparable to or larger than the width of the miniband without field.

The formation of the localized states is also known as Stark localization and the equally spaced ladder of eigenenergies of (23.29) is known as the Wannier–Stark ladder. In this picture, the temporal modulation of the FWM signal in Fig. 23.18 can be understood as quantum beats between adjacent band-to-band transitions with $\Delta n = \pm 1$. The (coherent) emission of quanta

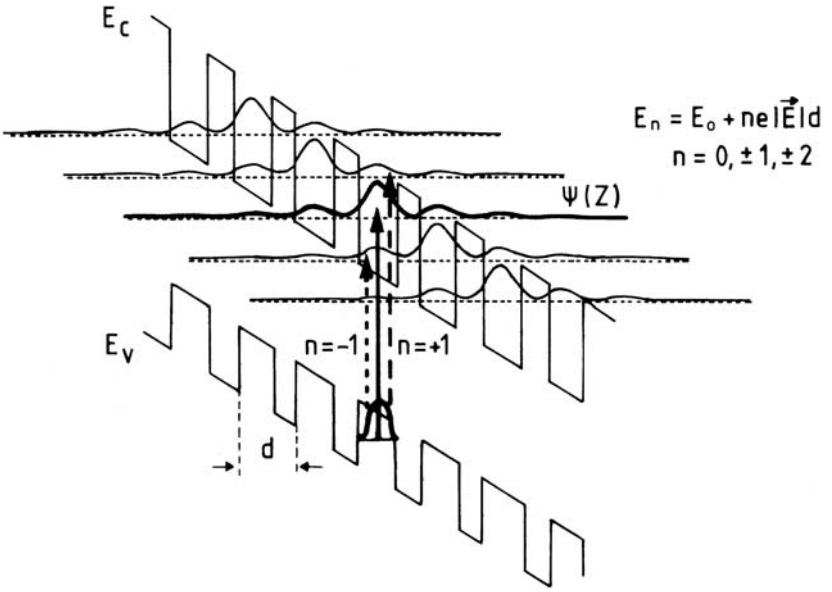


Fig. 23.21. The conduction bandstructure (similar to Fig. 23.16) and the valence bandstructure of a type-I superlattice, however now under the influence of a constant electric field in the growth direction. A wave function in the valence band is shown and several in the conduction band and three different transitions with energies E_0 and $E_0 \pm e|\boldsymbol{E}|d$ [94L1, 94R1, 02L1]

$\hbar\omega_{\mathrm{B}}$ in Fig. 23.20 is then just a result of transitions between adjacent rungs of the ladder.

In addition, the transitions indicated in Fig. 23.21 suggest that the structures should also be visible in simple linear absorption spectroscopy and indeed they are. An example is given in Fig. 23.22. Photocurrent excitation spectra are shown instead of absorption because the substrate of the $\mathrm{GaAs/Al_{1-y}Ga_yAs}$ superlattice which is incorporated in a pin junction is opaque; see Chap. 25.

The parameter in Fig. 23.22 is the bias voltage over the diode which is closely related to the electric field strength over the depletion region. A rather complex fan of optical transitions evolves with increasing blocking voltage. The following transitions can be identified: the quantized heavy hole to electron transition for $n = 0, \pm 1, -2$ denoted by $1hh, 0hh, -1hh$ and $-2hh$, two transitions involving the second quantized hh level denoted by $0hh2$ and $-1hh2$ and one transition starting from the first quantized light hole level. The fan of the various transitions again follows at least for higher fields, the linear relations of (23.26) (with a being replaced by the superlattice period d) and (23.29).

The Bloch oscillation and the Wannier–Stark ladder which we have discussed now from various points of view evidently allow one to perform beau-

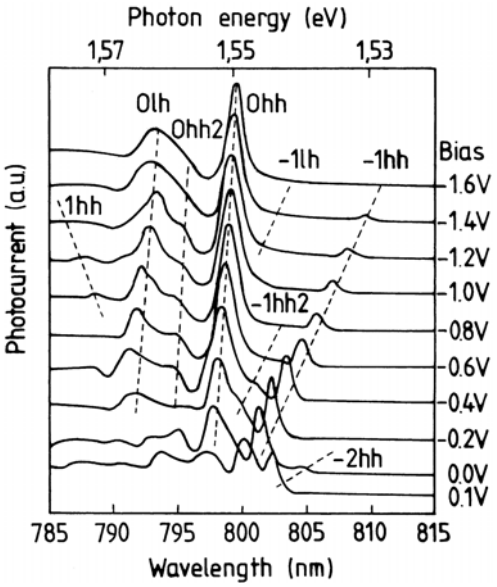


Fig. 23.22. The photocurrent excitation spectrum of a superlattice incorporated in a pin junction for various bias voltages [94L1]

tiful experiments and they may also offer prospects for application, especially concerning the creation of THz radiation. More information on these topics is found in [92F1, 92L1, 94L1, 94R1, 94V1, 01L1, 02L1] and references therein.

The THz pulses created by this and other techniques (see Chap. 25) start to allow ultrafast time resolved spectroscopy in the energy range of a few meV, which earlier was only accessible by Fourier spectroscopy, which is an inherently slow technique. Just to give one example we mention the THz emission of an oscillating wave packet in asymmetric coupled double quantum wells [92R1, 96B2, 00H1].

For the observation of Bloch oscillations in ultra cold atoms, its analogy to the Josephson Effect and other aspects of "nonlinear optics" with atoms, see [99J1, 99M2, 00E2, 00R2].

Now we address a rather well-known magnetic-field-induced effect, which has a lot of similarities but also some differences to the above-discussed electric-field-induced effects, namely the cyclotron resonance. See also Sect. 16.1. Under the influence of a magnetic field $\boldsymbol{B}$ a particle with a velocity v performs a periodic orbit in the plane perpendicular to $\boldsymbol{B}$. For particles with an isotropic effective mass this orbit is simply a circle (again assuming negligible scattering) and the frequency around this circle is independent of v (we assume $|v| \ll c$) and given by the cylotron resonance ω_c

$$\omega_\mathrm{c} = 2\pi/T_\mathrm{c} = e\,|\boldsymbol{B}|\,/m_\mathrm{eff}\,. \tag{23.30}$$

The circular motion in real space is necessarily connected with a periodic orbit in $\boldsymbol{k}$-space. In contrast to the Bloch oscillations, however, these orbits occur at constant energy since the force $e\boldsymbol{v} \times \boldsymbol{B}$ is always normal to the orbit.

If scattering processes are included, the condition for the observability of the cyclotron orbit is, in analogy with (23.28),

$$\omega_c T_2 \geqslant 1\,. \tag{23.31}$$

In analogy to the THz emission of the Bloch oscillation, the cyclotron resonance can be observed in absorption or emission at frequency ω_c depending on the relative phase of the electromagnetic radiation and the particle in the cyclotron orbit.

In the regime of the excitonic continuum or of the band-to-band transition, the combination of electron and hole cyclotron energies can be seen, as shown in Fig. 16.6, which corresponds in this comparison to Fig. 23.22.

The analog of the Wannier–Stark ladder are here the Landau levels. The motion of the carrier is quantized in the plane normal to $\boldsymbol{B}$ resulting in a harmonic-oscillator-like ladder

$$E = E_0 + (n + 1/2)\,\hbar\omega_c + \hbar^2 \boldsymbol{k}^2_{\parallel \boldsymbol{B}}/2m_{\text{eff}} \quad n = 0, 1, 2, \ldots\,. \tag{23.32}$$

The quasi one-dimensional free motion parallel to $\boldsymbol{B}$ results, in the effective mass approximation, in a DOS for every level of the ladder which varies with the inverse square-root of the kinetic energy, as found in Sect. 2.6 and shown in Fig. 23.23.

Quantum beats induced by a magnetic field are described in [91B2, 92B1, 94B1], though the explanation is in most cases not quite as simple as outlined here in connection with the Landau levels.

Further aspects of intersubband transitions in a magnetic field and related topics are treated in [82H1, 86H1, 87M1, 89H1, 90M1] and references therein, which leads us back to Chap. 16.

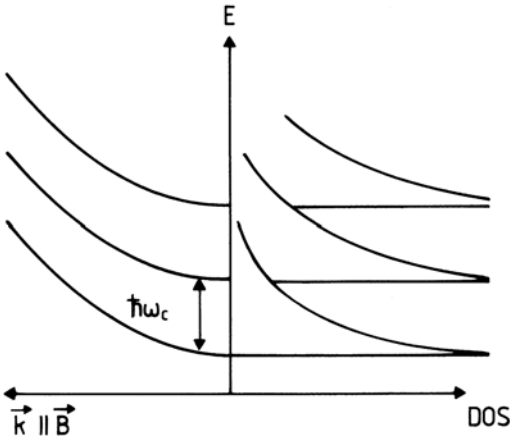


Fig. 23.23. The dispersion of the subbands resulting from the Landau levels and the corresponding density of states

23.2.1.2.7 Spin Dynamics

We mentioned in Sect. 23.1 that the decay time of the spin polarization may be different from the interband decoherence time. Therefore we present here examples of the measurement of the decay of the electron spin and then some references to work on spin dynamics.

The basic idea of the experiment is outlined in Fig. 23.24a. A 25 nm-wide GaAs quantum well is kept at low temperatures in a magnetic field of $2T$ normal to the directions of luminescence excitation and detection.

The excitation is either in the continuum states, in the $n_z = 1lh$ or in the hh exciton resonances. The circularly polarized excitation preferentially produces spin-polarized electron–hole pairs. While the hole spin is assumed to randomize rapidly due to spin–orbit coupling, the electron spin has a long dephasing time of several hundred ps, comparable to the luminescence decay time and produces, by its Larmor precession (r.h.s. of Fig. 23.24a), a modula-

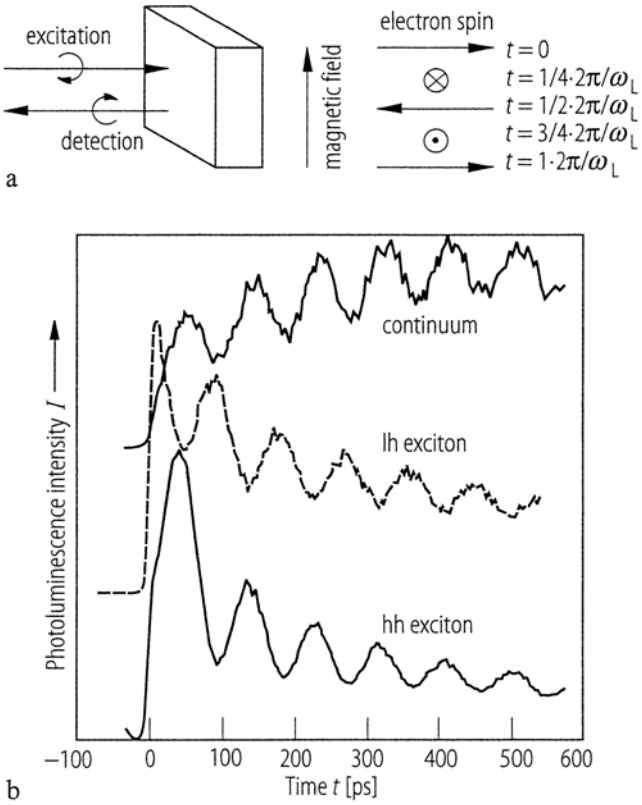


Fig. 23.24. The basic idea to measure the spin dephasing time by observing the Larmor precession for circularly polarized excitation and detection (**a**) and the temporal evolution of the hh-luminescence intensity for various excitation energies (**b**) [96O1]

tion of the luminescence intensity in the circularly polarized detection channel (Fig. 23.24b). The long value of the spin decoherence time, exceeding evidently (Fig. 23.24b) several hundred ps, contrasts with the interband decoherence time, which hardly exceeds the range of a few tens of ps at low temperature in the lowest exciton resonance and decreases rapidly with increasing excess energy and temperature.

Evidently, the electron spin does not even lose its coherence in an intersubband transition. Therefore it can be concluded that the electron spin (but not the hole spin) is strongly decoupled from the spatial part of the electron wave function.

This effect, which is known as the Hanle effect, has been observed even more pronouncedly in the (stimulated) emission from a semiconductor microcavity laser [97O1]. An extension of electron and hole spin dynamics in GaAs quantum wells towards quantum computing is suggested in [01H2, 01H3].

The long spin dephasing times have also been observed in self-assembled InAs quantum islands [01P1, 02C1].

The density-dependent spin dynamics in ZnSe quantum wells and dots have been investigated, e.g., in [00K1, 01O1, 02T1].

In semi-magnetic semiconductors the interaction of the spin of the carriers with the Mn^{2+} spins seems to possibly shorten the interband decoherence time [96C1], while the spin relaxation times of the free induction decay times seem to be long [97C1].

23.2.1.2.8 Phonon Dynamics

To conclude this survey of coherence times, we give an example of the phonon dynamics in Fig. 23.25. Similarly to exciton resonances in reflection spectroscopy (see Sect. 23.2.1.2) it is possible to determine the damping parameter of optically active phonons from a fit of their reststrahlbande (see Chap. 11). Recently it became possible to determine this quantity in the time domain by, e.g., impulsive Raman scattering [85Y1, 92Z1].

The data shown in Fig. 23.25 have been obtained from phonon generation via ultra fast screening of surface space charge fields [90C1, 92K2].

Many semiconductors have a strong electric field close to the surface resulting in band bending. If this field is screened in a time that is short compared to the oscillation period of optical phonons by creating a high density of electron–hole pairs with a sufficiently short pulse (< 50 fs) see (23.4d), (23.4e), then two things happen. The reflectivity of the surface changes over a time until the surface field recovers (which can be rather long) and a coherent pulse of LO phonons is excited.

As long as this pulse stays coherent, it modulates the reflectivity with a frequency $\hbar\omega_{\mathrm{LO}}$, e.g., via an electro-optic effect [92K2]. Both effects are seen in Fig. 23.25 for a GaAs:Cr surface.

The oscillation period of 8.84 THz corresponds to the LO phonon energy of GaAs of about 36 meV and the decay of the periodic part of the signal

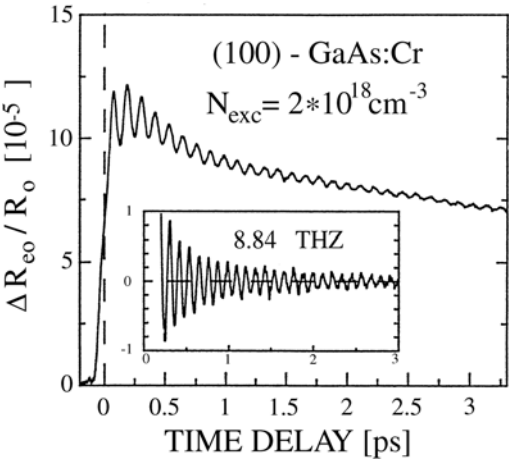


Fig. 23.25. Time resolved reflection changes of a GaAs:Cr sample after excitation of $2 \times 10^{18}\,\mathrm{cm}^{-3}$ of carrier pairs with a 50 fs pulse [92K2]

corresponds to a dephasing time of the order of 1 ps in agreement with data from reflection spectroscopy. For very high excitation densities it was also possible to excite the TO phonon mode resulting in beating [92K2]. For more recent data see, e.g., [01D1, 01Y1].

With this example we finish the short overview of possibilities to experimentally determine various dephasing times.

Before we come to the generally longer relaxation and recombination times we give some examples of experiments in the regime of quantum coherence.

23.2.2 Quantum Coherence, Coherent Control and Non-Markovian Decay

In the following we treat some aspects of quantum kinetics. In the first part we give an intuitive and didactic example of the distinction between Markovian and non-Markovian damping. Then we proceed with an example of non-Markovian damping and coherent control. In the third subsection we explain the concept of Rabi oscillations. Contributions to these topics can be found e.g. in [01W1, 02S1, 03Q1, 04O1].

23.2.2.1 Markovian versus Non-Markovian Damping

We closely follow an idea presented by M. Wegener in [01W1].

We introduced in Chap. 4 the model of Lorentz-oscillators to understand, on a basic level, the spectra of complex dielectric functions and the complex index of refraction and reflection.

For convenience we repeat (4.6) here for the case of vanishing driving force

$$\ddot{x}(t) + \gamma \dot{x}(t) + \omega_0^2(t) = 0\,. \tag{23.33}$$

A mechanical realization would be a pendulum that is for small amplitudes, to a very good approximation, a harmonic oscillator and the damping could be realized by attaching a sphere to the pendulum, which is submersed in a water tank as shown in Fig. 23.26. The friction proportional to the velocity force F_f would then be given by Stokes's law

$$F_\mathrm{f} = 6\pi\eta r\dot{x}\,(t) \tag{23.34}$$

where r and η are the radius of the sphere and the viscosity of the water, respectively.

The damping is evidently instantaneous in time and does not depend on the past. This approach is used, e.g., in the Boltzmann equation or in Fermi's golden rule.

The second order differential equation says that $x \equiv 0$ for all times $t \geq t_0$ if we realized at $t = t_0$ for an infinitely short time $x = 0$ and $\dot{x} = 0$, i.e.,

$$x\,(t_0) = 0; \quad \dot{x}\,(t_0) = 0 \quad \Rightarrow \quad x\,(t) \equiv 0 \quad \text{for} \quad t > t_0\,. \tag{23.35}$$

If we stop the oscillating pendulum for a very short time at its equilibrium position, then the result will be different. The bath itself flows with the moving pendulum and keeps flowing for a while when we stop it, since the energy transferred to the water is not immediately or instantaneously dissipated into a huge (ideally infinite) number of degrees of freedom of the bath.

If we release the pendulum after the above-mentioned very short time, it will be pushed by the still moving water and continue to perform a damped harmonic oscillation, although with a much smaller amplitude than before. This phenomenon is not described by (23.33), (23.35) but requires a memory kernel of the form

$$\ddot{x}\,(t) + \int\limits_{-\infty}^{t} \gamma\,(t - t')x\,(t')\,\mathrm{d}t' + \omega_0 x\,(t) = 0\,. \tag{23.36}$$

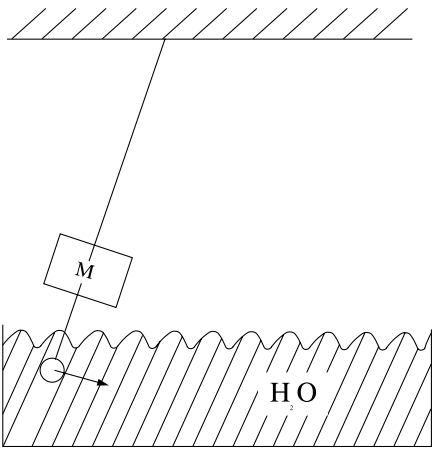


Fig. 23.26. A mechanical model to realize the difference between Markovian and non-Markovian damping. According to [01W1]

The "history" is taken into account by the integral over the function $\gamma(t-t')$. The equation is non-local in time (see Chap. 5, where we introduced non-locality in space).

In contrast to the solutions of (23.33) the amplitude of $x(t)$ does not decay immediately exponentially with time.

A simple approach to the memory function $\gamma(t-t')$ would be an exponential

$$\gamma\,(t-t) \sim \exp\left\{\frac{-\,(t\;-\;t')}{T_{\mathrm{M}}}\right\} \tag{23.37}$$

where T_{M} is obviously the memory time.

Setting T_{M} to zero or replacing it by a δ function $\delta\,(t-t')$ would bring (23.36) back to (23.33).

The limit $T_{\mathrm{M}} = 0$ corresponds to a Markovian and a memory function extending over a finite time to non-Markovian relaxation.

In the following we give experimental data from [01W1,02S1] as an example of the damping of the interband polarization of GaAs bulk layers by emission of LO phonons.

23.2.2.2 Damping by LO Phonon Emission and Other Processes

Various experimental and theoretical examples of quantum kinetics and of coherent control can be found, e.g., in [96H1, 96S1, 01L1, 02S1, 03Q1] and the references given therein. We concentrate here first on an example of LO emission from the inter-subband transition in thin, bulk-like GaAs epilayers [95B1, 98W3, 98W4].

In Fig. 23.27a,b we show schematically the experimental setups for the experiments described below and in Fig. 23.27c the band-to-band excitation and the LO phonon emission process.

Figure 23.27a gives the standard setup for a two-beam four-wave mixing experiment (also see Chap. 26). In Fig. 23.27b there are two coherent pulses on one beam line. The delay time $\tau_{11'}$ between them can be tuned with extremely high precision with a technique based on the Pancharatnam phase [97W2].

Finally, Fig. 23.27c shows the band-to-band excitation at $\hbar\omega_{\mathrm{P}}$ under $\boldsymbol{k}$-conservation. Under the assumption that the electron emits the LO phonon alone, the transition energy of the final E_{f} state is given by

$$\hbar\omega_{\mathrm{P}} - E_{\mathrm{f}} = \hbar\omega_{\mathrm{mod}} = \hbar\Omega_{\mathrm{LO}}\left(1 + \frac{m_{\mathrm{e}}}{m_{\mathrm{h}}}\right) \tag{23.38}$$

independent of $\boldsymbol{k}$.

There are several reasons why the inclusion of excitonic effects can be neglected in a first approximation for GaAs. The experiments described below have been performed at 77 K. So for bulk GaAs the inequality $E_{\mathrm{ex}}^{\mathrm{bind}} \leq k_{\mathrm{B}}T$ holds and the exciton resonance is already significantly damped. In addition the short Fourier limited pulses with a duration of about 15 fs have a spectral

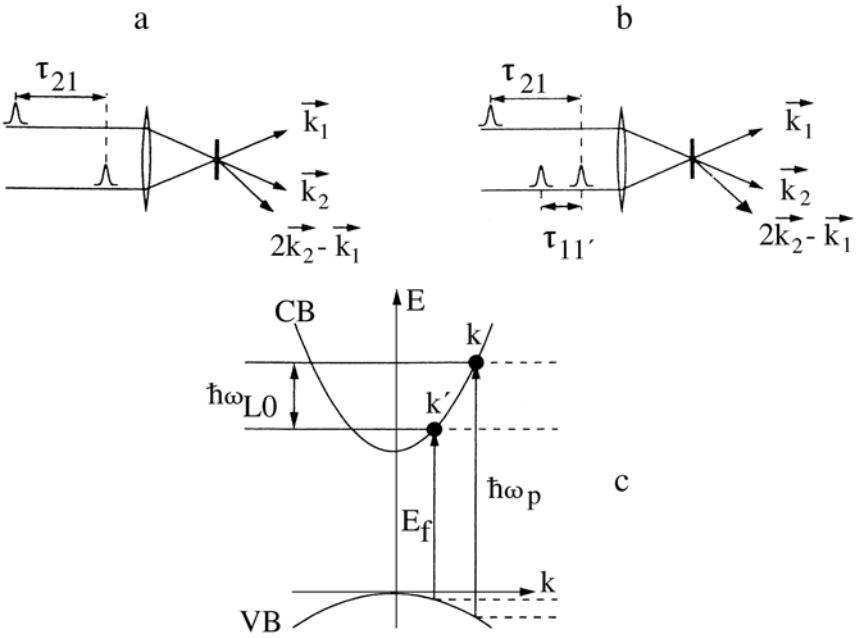


Fig. 23.27. Schematic experimental setups for a two-beam four-wave mixing experiment (**a**), a coherent control experiment (**b**), and the band-to-band excitation (**c**) [98W3, 98W4, 01W1]

width of almost 100 meV (see (23.4)), thus covering the energy difference given by (23.38) and exceeding the exciton binding energy by a factor of 20. This means that excitons and continuum states are excited simultaneously or, in other words, on such a short timescale that the electron–hole pair does not yet "know" if it is bound to an exciton due to the time-energy uncertainty relation.

In Fig. 23.28 we show experimental and theoretical results [95B1, 98W3, 98W4, 01W1].

Figure 23.28a gives results of two-beam four-wave mixing for various electron–hole pair densities from $1.2 \times 10^{16}\,\mathrm{cm}^{-3}$ to $6.3 \times 10^{16}\,\mathrm{cm}^{-3}$. AC is the autocorrelation trace of the 15fs laser. The insert shows the laser spectrum and the spectrum of the four-wave mixing signal for $\tau_{1,2} = 0$, which is centered on the exciton energy as already discussed in connection with Fig. 23.11. The experimental solid traces show the overall decay of the coherent interband polarization, but additionally a modulation with a frequency $\hbar\omega_\mathrm{m}$ given by (23.38) in agreement with theory (dots). In the simplest approximation, the beating can be understood as a coherent superposition of a state at $\hbar\omega_\mathrm{P}$, which has not yet emitted a LO phonon and a state at E_f which has. The decay of the oscillation results from a damping of the LO-phonon, i.e., a decay to lower energy lattice vibra-

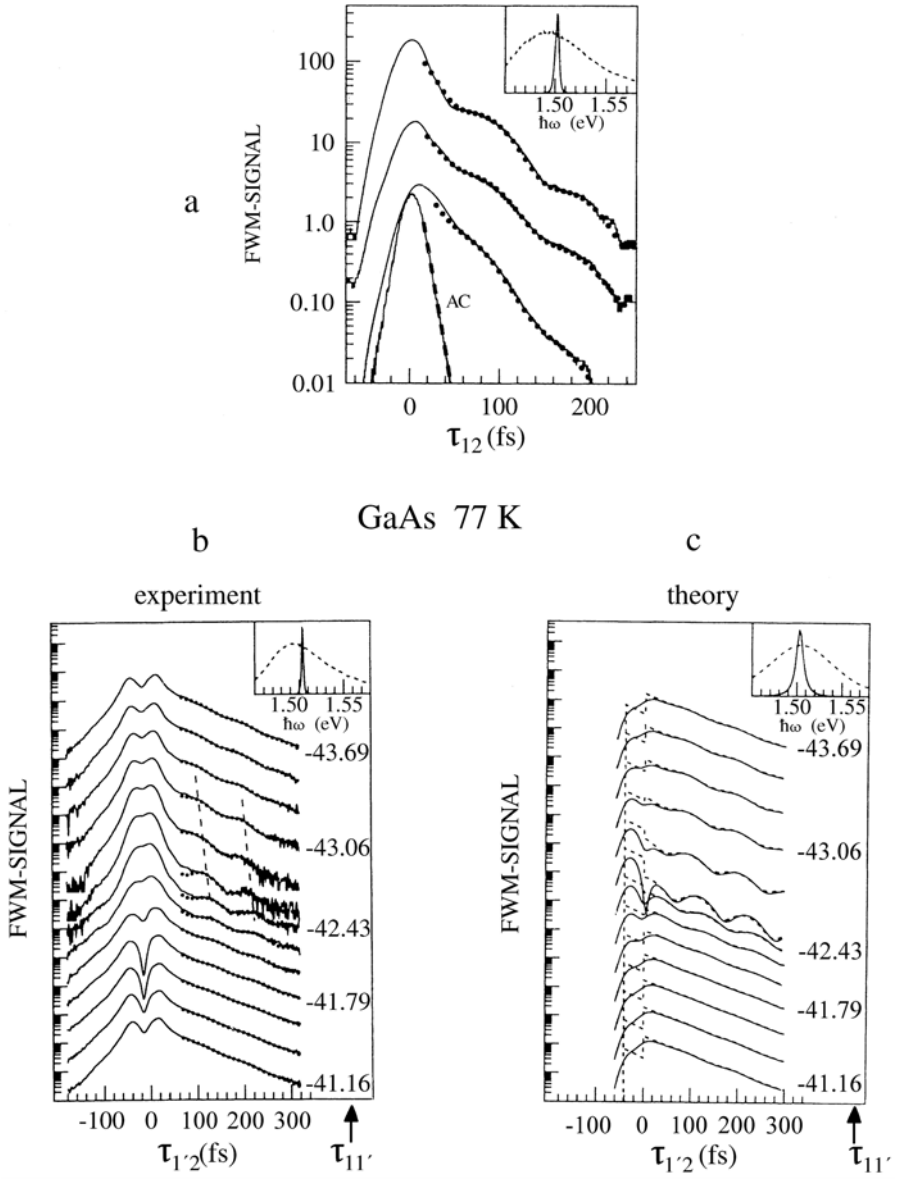


Fig. 23.28. Spectrally and temporally integrated two-beam four-wave mixing experiments on GaAs (**a**), experiments for coherent control (**b**), and theoretical modeling (**c**) [95B1, 98W3, 98W4, 01W1]

tions corresponding to an increasing number of degrees of freedom of the heat bath. Only after these processes have occurred, the absorption process is irreversibly completed. This is the reason why we stressed the im-

portance of phase relaxation or a scattering process in the absorption process in Sect. 3.2.

Once the LO phonon has decayed in (many) other degrees of freedom, decoherence of the inter-band polarization is completed.

The modulation of the oscillatory signal in Fig. 23.28 a decays faster than in Fig. 23.25, among others, because of the lower sample temperature in the latter case, reducing inelastic phonon-phonon scattering.

Similar data have been also obtained for GaAs quantum wells [98W4] and for ZnSe epitaxial layers [99S2], in the latter case with the important difference that the modulation occurs directly with the LO phonon frequency of ZnSe and not with the mass corrected term of (23.28). This finding proves that in ZnSe the electron–hole pair or exciton emits the LO-phonon as a whole and not a single carrier as can also be expected from the higher exciton binding energy and the stronger polar coupling to LO phonons.

The emitted, but still coherent phonon obviously corresponds to the flowing water in the tank of Fig. 23.26.

As long as the absorption process or the dephasing are not completed, it should be possible to influence its decay by a third coherent pulse. This is the idea of coherent control in the regime of quantum coherence.

In Fig. 23.28b and c we show measured and calculated spectrally and temporally integrated four-wave mixing signals as a function of the time delay τ_{12} but now with the delay $\tau_{11'}$ as a parameter (see Fig. 23.27b). Depending on the type of interference of the polarizations created by pulses 1 and $1'$ (constructive or destructive) the height, the decay time and the degree of modulation of the signal vary, with a periodicity given by the light frequency [01W1]. The experimental data can be reproduced to a good approximation by theory.

Some further examples of quantum coherence and coherent control can be found for the interaction with phonons in [00W1] with plasmons in [98M2, 00H2, 00H3, 00V1] or for Coulomb quantum kinetics in the regime of band-to-band transitions [99H1] and references therein, for exciton polaritons and biexcitons in quantum wells [95H1, 97B1, 03V1] or single quantum dots [98B3, 98B4] or for inter-subband transition [00S1].

A rather early example was the manipulation of the THz emission of two coupled and asymmetric quantum wells by the delay between two coherent incident pulses [93P3].

After this excursion to coherent control we turn to the last, but not least, effect in the quantum coherent regime, namely to Rabi oscillations.

23.2.2.3 Rabi Oscillations

The idea of coherent oscillations of the population of a two-level system was introduced by Rabi for atomic systems [37R1,39R1]. For a detailed description of two-level systems see, e.g., [75A1]. The basic idea is outlined in Fig. 23.29.

We consider an ensemble of identical (i.e., only homogeneously broadened) two-level systems with a transition energy

$$\hbar\omega_0 = E_2 - E_1 \,. \tag{23.39}$$

This ensemble is illuminated by a monochromatic light beam of frequency ω resulting in a normalized detuning d defined by

$$d = \frac{\hbar\,(\omega_0 - \omega)}{\hbar\omega_\mathrm{R}} \tag{23.40}$$

where ω_R is the Rabi frequency. It is given here, in the simplest case, by the product of the transition dipole matrix element μ multiplied by the amplitude $\boldsymbol{E}_\mathrm{l}$ of the incident light field

$$\hbar\omega_\mathrm{R} = \mu \boldsymbol{E}_\mathrm{l} \,. \tag{23.41a}$$

It should be noted that in the literature there are additions to this simple term due to, e.g., exciton resonances (see (20.9)), damping of the transition or many particle effects [96H1, 96S1, 02S1, 03Q1].

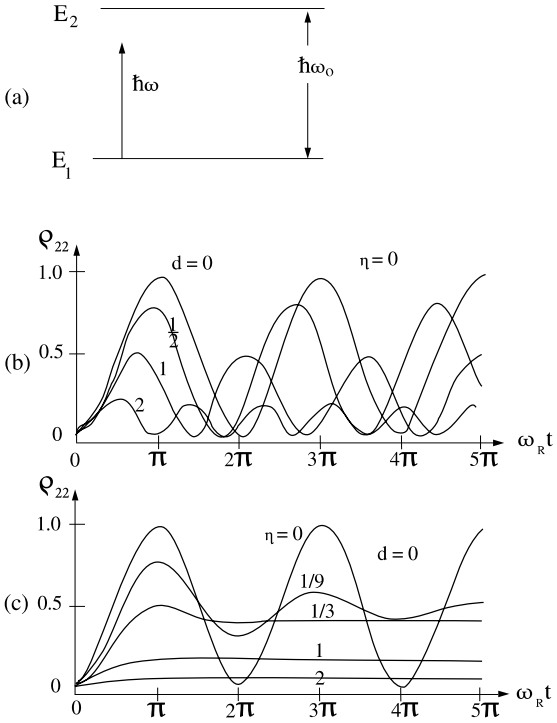


Fig. 23.29. A two level system illuminated with coherent light of frequency ω (**a**) the normalized population of the upper states ρ_{22} as a function of normalized time for various normalized detunings d and vanishing damping (**b**) and for zero detuning and various normalized dampings (**c**) [03D1] and references given therein

If the dephasing time is infinitely long, i.e., $\eta = \gamma = 0$ in Fig. 23.29, the driving field pumps all systems coherently from the ground state to the excited state and back again and so on. The characteristic frequency for this process, the so-called Rabi oscillation, is just the Rabi frequency of (23.40). If the detuning d is zero the population of the two-level system can be completely coherently driven between the ground and excited state. For finite detuning, the maximum value of ρ_{22}, i.e., the population of the upper state, decreases as shown in Fig. 23.29b.

If the population is flipped once or a few times from the ground state to an excited state by a short pulse, one speaks of Rabi flopping rather than of Rabi oscillations.

Zero damping is, of course, an idealization. For finite (normalized) damping the oscillations disappear with time and one reaches a steady state population under cw excitation, which can be described by simple recombination kinetics using, e.g., Einstein's coefficients. In this case the maximum value of ρ_{22} is 0.5 after the Rabi oscillations die out (see Fig. 23.29c).

The sample is transparent under these conditions, which is why population inversion and lasing are usually not possible in a pure two-level system.

While the dephasing times can be rather long in atomic systems and Rabi oscillations and the optical Stark effect (see Sect. 20.4) have been known for a long time [75A1], Rabi flopping could only be observed in semiconductors after the invention of intense fs lasers.

In early experimental examples in $In_{0.08}Ga_{0.92}As$ multiple quantum wells and CdSe epilayers indications of Rabi flopping have been deduced from the shape of the transmitted pulse [94C1,95C1,97F1,98G1,01N1] with a duration of about 100 fs and a peak intensity in the GW cm^{-2} range. We show here an example of a 20-period $In_{0.1}Ga_{0.9}Ga/GaAs$ multiple quantum well sample from [99S3] in Fig. 23.30. In Fig. 23.30a the transmission spectrum exhibits the $n_z = 1hh$ and light hole transitions. Figure 23.30b gives the spectra and polarization of the 770 fs pump pulse tuned to the hh transition and the 150 fs probe pulse, centered on the lh resonance.

The almost periodic oscillations of the transmission in Fig. 23.30c are attributed to Rabi oscillations. This interpretation is based on agreement with calculations including nonlinear Coulomb renormalization of the transitions (Fig. 23.30d) and the fact that the oscillation frequency increases with the square root of the energy flux of the pulse, i.e., linearly with the field amplitude.

Rabi oscillations in quantum dots have been reported, e.g., in [01S1,02B2, 02B3,02K2] and for μ-cavities in [01D2].

Usually one has the situation

$$\hbar\omega_R \ll \hbar\omega_c \approx E_g \tag{23.42a}$$

where ω_R is the Rabi frequency and ω_c the carrier frequency of the laser pulse.

According to (23.40), if the laser intensity is sufficiently high, however, one may reach a situation

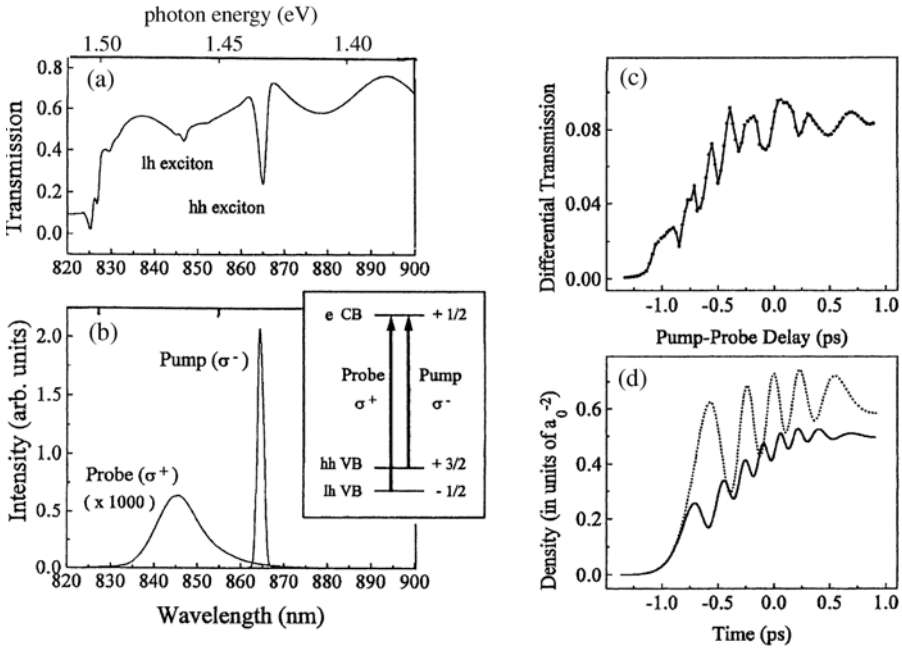


Fig. 23.30. (**a**) The transmission spectrum of the $In_{0.1}Ga_{0.9}As/GaAs$ MQW sample. (**b**) The spectral shapes and polarizations of pump and probe pulses with the relevant transitions in a band-to-band model in the *inset*. (**c**) The differential transmission signal as a function of pump-probe delay. (**d**) Calculations neglecting (*dotted line*) and including (*solid line*) Coulomb renormalization [99S3]

$$\hbar\omega_{\mathrm{R}} \approx \hbar\omega_{\mathrm{c}} \approx E_{\mathrm{g}} \tag{23.42b}$$

known as carrier wave Rabi flopping.

This new phenomena, which appears under this unusual situation, belongs to the regime of extreme nonlinear optics (see Sect. 19.3) and can only be observed for lasers with a duration of a few light cycles to avoid sample damage or (excitation-induced) dephasing. The Rabi doublets of the fundamental and the third harmonic may overlap around the second harmonic, producing signals that depend on the absolute optical phase between the carrier wave and the envelope, i.e., whether the maximum of the envelope coincides with a maximum of the carrier wave or with a zero transition. Examples for this regime are found in [01M2, 03S1].

23.3 Intra-Subband and Inter-Subband Relaxation

After this overview of coherent processes we review incoherent intraband and inter-subband relaxation and finish this subsection with some transport properties. For the latter topic see also Sect. 23.2.1.2.5

23.3.1 Formation Times of Various Collective Excitations

The first processes we consider concern the formation times of some quasiparticles or collective states. These times are partly as short as the decoherence times and involve concepts of quantum kinetics. A recent review connecting the regimes of 23.2 and the one here in is [01C1].

The first example is the formation time of excitons. They can be either directly resonantly excited by a spectrally sufficiently narrow laser (i.e., smaller than Δ_{LT}) as stationary eigenstates, or can from after band-to-band excitation i.e. excitation in their ionization continuum with some delay.

In GaAs the exciton binding energy is small $\left(E_{\mathrm{x}}^{\mathrm{b}} \approx 4.5\,\mathrm{meV}\right)$ compared to the LO phonon energy ($\hbar\omega_{\mathrm{LO}} \approx 36\,\mathrm{meV}$), and the polar coupling to phonons is weak.

Consequently the formation time of excitons is rather slow in the range of 100 ps for electron–hole pair excitation in the continuum [98K2, 02H1, 03K1, 03K2, 04C1, 04K2, 04O1] in contrast to resonant excitation [03K2]. The presence of excitons has been deduced for GaAs bulk samples or $\mathrm{In}_{1-y}\mathrm{Ga}_y\mathrm{As}$ quantum wells from the 1s $\rightarrow$ 2p transition at about 6.3 meV in time-resolved pump and THz probe spectroscopy [03K1]. In these experiments it was also possible to show the thermal ionization of excitons at a lattice temperature of 60 K to occur on a time scale around 10 ps.

The rather slow formation time of excitons in bulk GaAs and in GaAs- or InAs-based wells has even provoked the extreme question of whether excitons exist at all. For a recent review see, e.g., [03K2] or [04C1] and references therein. This question has a philosophical aspect, namely "what does the existence of a quasiparticle in a solid mean", which we do not attempt to answer here. The more down-to-earth aspect is if excitons exist in a similar way as other quasi-particles like phonons, plasmons and magnons. Here, we think the answer is rather clear and yes.

All the experiments showing the free motion of excitons (see below in Sect. 23.4), the direct measurements of their (polariton) dispersion curves or of their thermal distribution on this dispersion as detailed in Chaps. 9, 13 and 15 or the formation of biexcitons (20.3) make it difficult to discard the concept of excitons as quasiparticles.

Furthermore (time-resolved) photo-luminescence excitation spectra of excitons show in more ionic bound semiconductors with higher binding energy comparable to the LO phonon energy like in CdTe or ZnSe, that excitons are formed rapidly, even in the ionization continuum. We already addressed this point in Sect. 23.2.2 and will see a further example below.

The dynamics of LO phonon emission from heavy holes occurring even for a kinetic energy below $\hbar\omega_{\mathrm{LO}}$ has been investigated as a function of the Fröhlich coupling parameter α in GaAs and CdTe both experimentally and theoretically using quantum kinetics (see [01B4] and references therein). These results essentially confirm the more didactically formulated statements above and also give, for the example of the *hh* in CdTe, some information on the

formation times of the polaron of several 10 fs. For the polaron concept see Sect 8.6.

The formation time of the magnetic polaron introduced in Sect. 10.2 has been investigated by various groups (see e.g. [87Z1,95D1,96H2] and references therein). The formation time depends on the Mn concentration, on temperature, on the presence of an external magnetic field and, for quantum wells, also on the well width. Typical formation times range from several tens of ps to a few ns. Thus the formation time of magnetic polarons is substantially longer than that of (phonon) polarons mentioned above.

The last example for formation times is the build-up of Coulomb screening in an electron–hole plasma. As explained in Chap. 21 a random spatial distribution of electrons and holes does not result in band gap renormalization.

Recently it was possible to observe in beautiful experiments with time resolved band-to-band pump and THz probe beam spectroscopy the build-up of Coulomb screening and plasmon scattering on a timescale of a few 10 fs. For details on the experiment and theory see, e.g., [98B3,98B4,01C1,01H1,01H2, 01H3]. These buildup times are roughly equal to the inverse of the plasma frequency ω_{PL}, a fact that is not totally surprising.

23.3.2 Intraband and Inter-subband Relaxation

We now leave the regime of ultrashort times and of quantum kinetics and turn to "usual" intraband and inter-subband relaxation, starting with bulk materials, proceeding to alloys and then to structures of reduced dimensionality.

We start with rather old measurements on the exciton polariton in CdS [75W1]. Figure 23.31 shows the polariton dispersion of the $n_{\mathrm{B}} = 1A\Gamma_5$ exciton, the luminescence spectra for two polarizations, and the lifetime measured from the luminescence decay at various energies after band-to-band excitation. The excitons relax in about 1 ps by LO-phonon emission onto the lower polariton brand (LPB). There they can relax further, e.g., by emission of acoustic phonons. This process becomes slower and slower with decreasing slope of the dispersion, since less and less energy transfer is allowed per unit of momentum transfer. For this reason the excitons accumulate in the indicated area, which is known as the "bottle neck". See also Sect. 13.1.3. As a loss mechanism, polaritons can be transmitted from the UPB and the LPB through the surface, but only if the parallel component of their wave vector $\boldsymbol{k}$ satisfies

$$k_{\parallel} \lesssim \frac{2\pi}{\lambda_{\mathrm{vac}}} \tag{23.43}$$

with a probability depending on the reflection coefficient and on the squared amplitude of the photon-like part of the wavefunction. Additionally the excitons can be trapped at defects and recombine there (non-) radiatively or be thermally reexcited into extended states. They can be scattered onto the photon-like branch of the dispersion curve under emission of one or more

LO phonons (Sect. 13.1.3). A further relaxation to the regime below the bottle-neck by emission of acoustic phonons becomes rather improbable due to the rapidly decreasing density of final states; see Sect. 2.6 and Fermi's golden rule.

A similar scenario exists also for the cavity polaritons introduced in Sets. 17.1, 20.2 and 22.1.

The accumulation of the exciton-like polaritons in the bottle-neck during the relaxation in the band directly explains the increase of their radiative decay times in Fig. 23.31c at these energies.

The relaxation of localized excitons through the localized tail states shows up in the time dependence of the emission line shape and its maximum position after pulsed excitation. Some examples have already been given in Sect. 14.4 and another will be given in Sect. 23.4. The thermally activated relaxation from metastable states to deeper ones at intermediate temperatures (e.g., 10 K to 50 K) and the thermal excitation to higher, generally extended states at even higher temperatures caused by the higher density of these states result in the well-known non-monotonous shift of the emission maximum with temperature. Examples for various II–VI or III–V materials and theoretical modeling are given for bulk materials and quantum films in [87S1,87S2,92S1,92S2,92S4,92S5,99K1,00K2,01T1,02K4,02R1,03R1,04K1] and the references given therein.

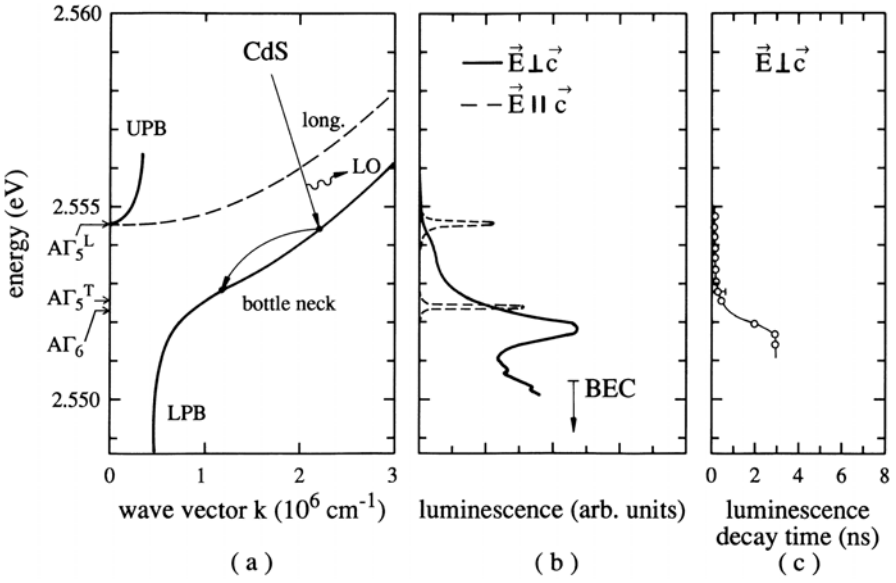


Fig. 23.31. The dispersion of the $n_B = 1A\Gamma_5$ exciton polariton in CdS in the bottle-neck region (**a**); the luminescence spectra for the dipole allowed orientation $E \perp c$ and the forbidden one $E \parallel c$ (**b**); and the decay time of the luminescence as a function of photon energy (**c**) [75W1] compare also with Figs. 13.4, 11, 14, 15 and 29

In Fig. 23.32 we give another example of intraband relaxation, this time of a AlGaAs MQW sample. The differential transmission between unexcited and excited sample is measured with 50 fs increments when the sample is excited around $t = 0$ with a pump pulse of 80 fs duration. There is some bleaching due to partial blocking of the band-to-band transitions by the bunch of electron–hole pairs excited in the continuum (dotted areas). The temporal evolution of the intraband relaxation can be seen nicely. There is a decrease of excitonic absorption in the region of the $n_z = 2hh$ transition around 1.57 eV. This change is rather small and constant on the time scales shown here ($\ll T_1$) and is due to direct screening of the Coulomb interaction in the exciton by the carriers. This screening is largely independent of the energy and distribution of the free carriers in the bands. A similar behavior is found for the $n_z = 1hh$ and lh resonances around 1.46 eV in the three lowest traces. When the excited electrons relax down into the exciton states, the transmission increases drastically. This finding is attributed to phase-space filling and exchange interactions which become effective for these exciton resonances only after the excited species start to populate the respective states. Figure 23.32 thus demonstrates qualitatively

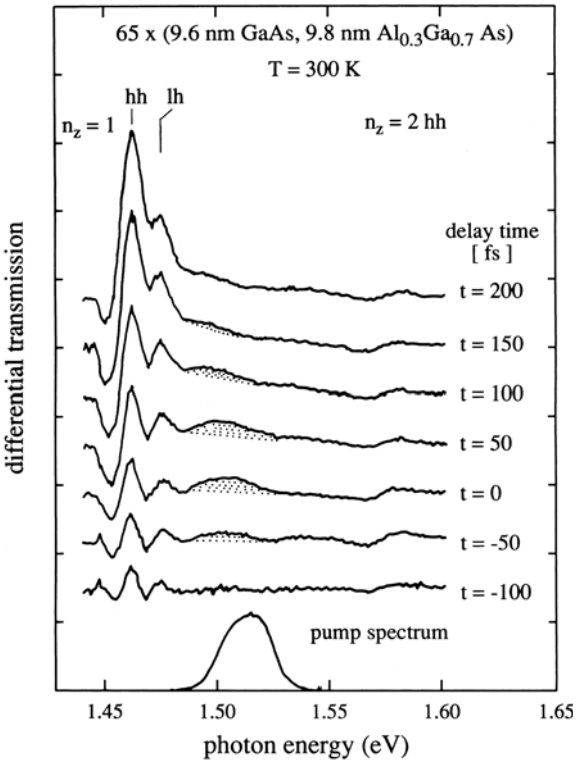


Fig. 23.32. Intraband relaxation of electron–hole pairs in the $Al_{1-y}Ga_yAs/GaAs$ MQW measured by time-resolved pump and probe experiments [88C1]

that direct screening of the Coulomb interaction is less efficient in quasi two- (and one-)dimensional systems than exchange interaction and phase-space filling. This difference to three-dimensional materials has already been pointed out Sect. 21.2 and 4 in connection with electron–hole plasmas.

As a further example of quantum wells we provide data in Fig. 23.33 for a $ZnSe/ZnS_{1-x}Se_x$ superlattice from [98U1].

The l.h.s. of Fig. 23.33a shows schematically the luminescence spectrum. There is a zero-phonon line (ZPL) resulting from the emission of excitons localized in tail states and of a free exciton with small $\boldsymbol{k}$ vector (23.43). The LO phonon sideband monitors in a similar fashion as for bulk samples in Figs. 13.12 and 13.13 the whole distribution of excitons since radiative recombination through emission of a LO phonon is possible from every $\boldsymbol{k}$ state as indicated on the r.h.s. of Fig. 23.33a, while Fig. 23.33b shows a time-integrated photoluminescence spectrum showing free and localized *lh*, *hh* and bound-exciton (BE) luminescence and their LO phonon replicas. The photoluminescence excitation spectrum taken at the position of the *hh* exciton shows the lh exciton resonance and peaks at integer multiples of LO phonon energies above the *hh*. Up to four such LO phonon resonances are visible. The absence of a mass renormalization factor (23.38) shows, that the electron–hole pairs

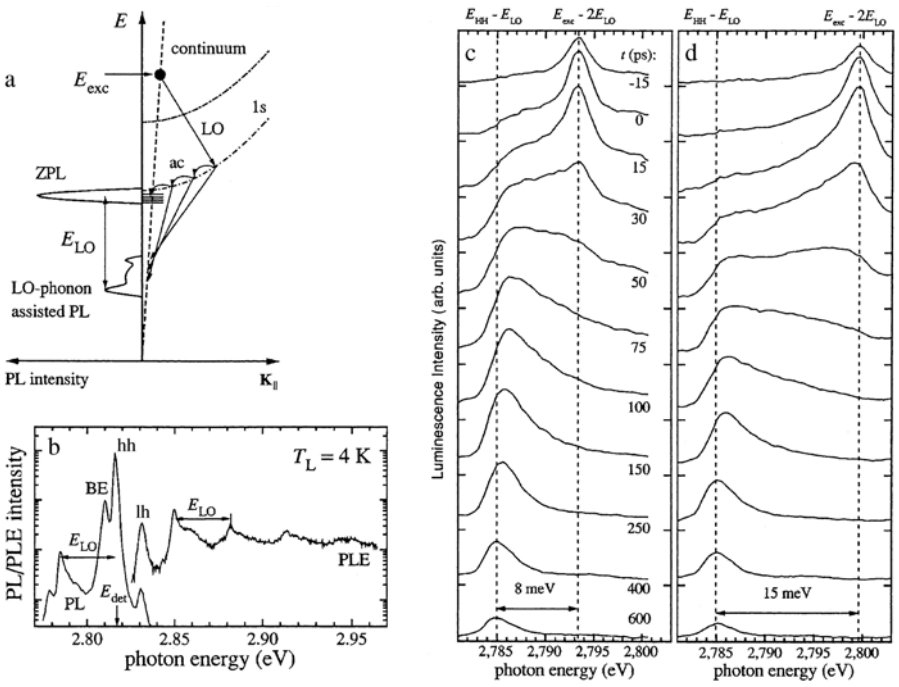


Fig. 23.33. (**a**) Schematics of the luminescence spectrum and of the exciton relaxation process in $ZnSe/ZnS_{1-x}Se_x$ superlattices (**b**) A time-integrated luminescence spectrum **c**, d Time-resolved spectra for two different excess energies [98U1]

emit the LO phonons together, i.e., that they form a rapidly bound state that relaxes on the exciton dispersion curve as shown schematically on the r.h.s. of Fig. 23.32a. This interpretation is confirmed by time-resolved measurements of the LO phonon replica after ps-pulsed excitation in Fig. 23.32c,d where the excitation energy is 1.5 and 1.2 LO phonon energies above the *hh* band. Within the temporal resolution luminescence appears 2 LO phonons below the excitation. Then a slower relaxation of the excitons by emission of acoustic phonons (see r.h.s. of Fig. 23.32a) leads to a population of hh excitons on a time scale of 100 ps followed by a lifetime-limited decay of this line of about 0.5 ns. Calculations in [98U1] not shown here agree with the experimental data and further corroborate this model.

Further relaxation processes include, e.g., the capture of carriers into quantum wells, wires or dots. For some references see, e.g., [96S2,98R1,98U2,99F1, 00L4,00R1,01R1,01L1] and references therein. For the inter-subband or inter-level relaxation in the quantum structures see, e.g., [99F1, 99H1, 99H2, 99M1, 00L4, 01L1, 01M3, 02B2, 02B3, 02B4, 02D1, 02K2, 02K3, 04O1] and references therein.

We give here an example for the inter subband relaxation of *hh* excitons in a GaAs/$Al_{1-y}Ga_yAs$ MQW sample, which is excited into the third subband as shown schematically in the inset of Fig. 23.34.

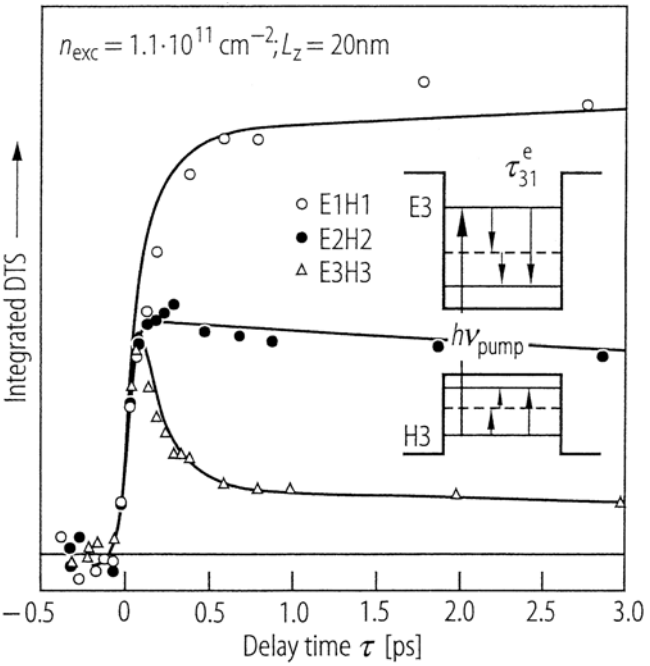


Fig. 23.34. Time-resolved differential transmission (bleaching) signal of the first three *hh* exciton resonances [94H1]

The bleaching of the resonantly excited $n_z = 3hh$ exciton decays on a sub ps timescale by relaxation to deeper levels. Simultaneously, the bleaching of the $n_z = 2hh$ and $n_z = 1hh$ builds up. The subsequent decay of the $n_z = 2hh$ resonance bleaching is already much slower, while the $n_z = 1hh$ ground state still shows a slight increase due to feeding from higher states over the time interval shown in Fig. 23.34 of 3 ps. The lifetime of this state is in the range of 100 ps as shown below.

Further topics in the field of relaxation concern, e.g., the relaxation by tunneling processes see [96S1, 96T1, 01L1] and references therein, or the relaxation in an electron–hole plasma [93M1, 96S1, 97M1, 97M2, 98M1] and references therein.

Recent reviews on relaxation processes is also found in [97M2, 04O1] and in other contributions to this book.

23.3.3 Transport Properties

Transport properties are intimately related to relaxation and collision processes, e.g., in the relaxation time approximation of the electrical conductivity resulting in

$$\sigma = \frac{ne^2\tau}{m} \quad \text{and} \quad \mu = \frac{v_\mathrm{D}}{|\boldsymbol{E}|} = \frac{\tau e}{m} \tag{23.44a}$$

where n is the carrier density, m is the effective mass of the carriers, τ is the time between collisions, v_D is the drift velocity, and μ is the mobility, resulting in the expression (23.44b) for the diffusion constant according to the Nernst–Townsend–Einstein Relation

$$D = k_\mathrm{B}T\tau/m = k_\mathrm{B}T\mu/e\,. \tag{23.44b}$$

A more elaborate theoretical concept is given in terms of the Boltzmann equations, treated in many textbooks of solid state or semiconductor physics.

Electrical transport and conductivity as well as heat conductivity are obviously not topics of this book. However we give below some examples of exciton or electron–hole pair transport in semiconductors. See also 23.2.1.2.5 above.

Actually we have already given several examples of such transport measurements, e.g., for excitons in Cu_2O in Sect. 20.5.2 of the electron–hole plasma in Sect. 21.3.1 or above with references [92O2, 02Z1, 03Z1] including transition from coherent or balistic transport to diffusive one [04D1, 04Z2].

Furthermore the coherent polariton transport has been discussed, e.g., in connection with propagation quantum beats in Sect. 5.2. In the following we give some more examples.

An early example is exciton diffusion in CdS [59B1]. The transport of excitons via the decay of laser-induced population grating (also see above) has been investigated for CdS in [88W1, 88W2] revealing a temperature- and density-dependent diffusion length on the order of 1 μm.

The influence of the polariton dispersion on the exciton transport has been investigated by the same technique as for $ZnSe/ZnS_{1-x}Se_x$ quantum wells in [98N1] showing the influence of the polariton group velocity on the transport length.

In contrast, the diffusion length of excitons tends to zero for zero temperature due to localization effects in alloys like $CdS_{1-x}Se_x$ as found in [92S1, 92S2, 92S4]. For the influence of disorder in quantum wells on transport see [00Z1] and references therein.

The expansion of an electron–hole plasma in direct gap semiconductors like CdS or CdSe has been investigated by spatially resolved pump-and-probe beam spectroscopy resulting in diffusion lengths in the 4 to 10 μm regime [84K1, 85M1].

23.4 Interband Recombination

Now we discuss some measurements of the interband recombination time T_1. For the definition of this quantity see the discussion with (23.5) or Chap. 27.

To measure this quantity it is best in principle to observe effects to which all excited species contribute by roughly the same amount. One could follow, e.g., the further temporal evolution of the bleaching of the exciton resonances in Figs. 23.32 and 23.34 and one would find that the differential transmission signal disappears with a time constant of about 0.5 ns.

Very often one relies, however, simply on the temporal evolution of the luminescence, e.g., of excitons or biexcitons, as a monitor of the population of the respective species, although it is well known that the luminescence yield η for most semiconductors is considerably smaller than 1. For high quality direct gap samples one finds, as already mentioned several times,

$$10^{-1} \gtrsim \eta \gtrsim 10^{-3}\,; \tag{23.45a}$$

only for localized excitons in some selected samples of bulk alloys or quantum structures or for some laser diodes have values of η close to 1 been reported. For indirect gap materials or direct ones with dipole forbidden band-to-band transitions or for direct ones containing "luminescence killers" like Fe or Cu ions η can even be much smaller than in (23.45) (see, e.g., [97W3,98W5,01J1,03F1,03T1] and the references given therein for examples). This means that with luminescence decay measurements only, one monitors the fate of a minority of the exited species and assumes quietly that they represent the majority.

Furthermore, radiative decay involves the conservation of the component of $\boldsymbol{k}$ parallel to the surface (see below) so that large $\boldsymbol{k}$ excitons hardly contribute to the zero phonon line. Finally the luminescence decay is often not simply exponential after pulsed excitations (see (23.5)), so a definition of a lifetime is, strictly speaking, not possible. Only something like an effective decay time over a certain time – or intensity interval – can be given.

Looking closely, the problem of radiative decay turns out to be rather complex. An exciton polariton in a perfect three-dimensional crystal cannot decay radiatively, since it is already the quantum of the mixed state of exciton and photon.

It can move (diffusively or ballistically) through the sample and be transmitted through or reflected at the surface of the crystal when it hits it with a certain probability. In the first case it appears as luminescence on the other side. Transmission through the surface is only possible if the parallel component of $\boldsymbol{k}$, which is conserved at the surface (see Sect. 3.1.3), is smaller than

$$k_{\parallel} \leq k_{\mathrm{vac}} = 2\pi\omega/c\,. \tag{23.45b}$$

(see, e.g., Fig. 13.11). For $|k| \gg k_{\parallel}$ given by (23.45b) the transmission probability decreases rapidly. One can either argue that the polariton is essentially exciton-like or that the refractive index of these states is very large resulting with (3.19) in a high reflectivity or substantial ranges of total internal reflection.

What can happen in a perfect crystal is that the exciton polariton is scattered by emission of n LO phonons on the photon-like branch of the dispersion curve, where it travels with a high group velocity essentially ballistically through the sample and has a good chance of being transmitted through the surface when it hits it, appearing as luminescence as shown in Fig. 13.12.

If there are defects in the sample (or at its surface), the exciton polariton can be trapped at these defects and recombine there either nonradiatively resulting in (23.45a) or emit a photon-like polariton, which appears with high probability as a luminescence photon. The same holds for capture in localized tail states in bulk alloys or quantum structures. Similar arguments are valid for cavity polaritons.

In quantum wells and wires $k_{\parallel}$ is the only component of $\boldsymbol{k}$, this means that free exciton (-polaritons) with large $\boldsymbol{k}$ cannot radiate at all. Those with small $\boldsymbol{k}$ produce photons outside the quantum structure very quickly (in about 10 ps) and particles with higher $\boldsymbol{k}$ have then to relax down to small $\boldsymbol{k}$ values.

This process and the restriction of the emission of free excitons to the light cone defined by (23.45b) results in a law for the luminescence decay time T_{L} as a function of the lattice temperature T_{lattice} ([91A1,91F2,92O1,92O2,94A1, 98L1, 00Z2, 04P2] and references given therein)

$$T_{\mathrm{L}} \; \alpha \; T_{\mathrm{lattice}}^{d/2} \tag{23.45c}$$

where d is the quasidimensionality of the system. We give a detailed derivation for $d = 2$ below. The linear increase of the luminescence decay time with temperature for $d = 2$, i.e., for quantum wells, is frequently used as an argument for dominant radiative decay. This is a clear over-interpretation of the data since similar samples show different slopes in the relation (23.45c) as shown below in Fig. 23.35 and samples with $\eta_{\mathrm{lum}} \ll 1$ can also show such behavior [91O1, 92O1, 92O2, 92S5, 03F1]. If one assumes that the capture of

excitons or exciton-like polaritons into nonradiative and radiative centers, including localization sites, decrease rapidly with increasing $\boldsymbol{k}$ it becomes clear that a similar law to (42.343 b) also holds for these cases, no longer restricting (23.45c) to dominant radiative recombination. However, the decrease of the luminescence yield observed for many high-temperature samples (e.g., for GaAs QW is usually above 100 K) implies that excitons or electron–hole pairs with high thermal kinetic energy find an increasing number of nonradiative decay channels.

The situation is slightly different for quasi-zero-dimensional quantum dots, islands and other localization sites since there is no more $\boldsymbol{k}$-conservation.

In spite of all these implications it is experimentally quite easy to measure the temporal decay of a luminescence signal, and therefore many data on T_1 of electronic excitations of semiconductors are obtained by this technique. If one remains aware of the above implications it may be a useful technique.

We start with a few typical exciton lifetimes in bulk semiconductors.

The exciton lifetimes are at low temperature and density in bulk semiconductors with direct, dipole-allowed band-to-band transitions typically in the range of a few hundred ps to a few ns (see Fig. 23.31) and for dipole-forbidden materials from several tens of ns to 5 ms (see Sect. 20.5). For Cu_2O the values for the lifetime reported for the forbitten para exciton range from around 10 ns to beyond 10 ms. See the ref. given in Sects. 13.2.1.2 and 20.5.2.

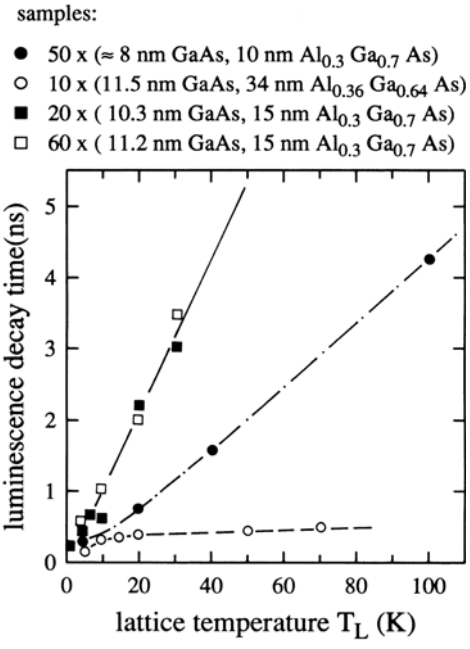


Fig. 23.35. The decay time of the luminescence in various $Al_{1-y}Ga_y$ As/GaAs MQW samples as a function of the lattice temperature [92O1]

In indirect gap materials values from a few tens of ns to ms are adequate. Increasing temperatures and increasing densities tend to decrease the lifetimes, e.g., due to increasing average $\boldsymbol{k}$-vectors and inelastic scattering processes like Auger recombination [75H1, 79H1], respectively. Stimulated emission also reduces the effective lifetime down to the 100 fs range [85M1].

The localized excitons in alloy semiconductors show that at low temperature there is a strong dependence of the luminescence decay time on the photon energy (see e.g. [89G1, 91U1, 98K1, 99K1, 04K1] and references therein). However this decay is governed by a complex interplay of temperature dependent relaxation between various localized states (see Sects. 23.4, 9.6 and 14.4) and radiative as well as nonradiative recombination. Typical lifetimes are again in the range of a few ns but may reach 100 ns for deeply localized states and especially for distant pairs [92S4, 98K1, 99K1, 04K1].

The same holds true for localized excitons in quantum structures, especially in II–VI quantum wells and superlattices. Examples are found in [93K2, 01S2, 04O1]. In type II structures the electron–hole overlap depends strongly on the well widths and therefore also on the lifetime. An example for CdS/ZnSe structures is found in [01S2, 03S2, 04P1].

In piezo-superlattices like CdSe/CdS the band bending and consequently the electron–hole overlap integral depend not only on the layer widths but also depend strongly on the carrier density via a screening of the piezo fields resulting in an extremely nonlinear recombination dynamics and in a temporal shift of the emission photon energy over several hundreds of meV after pulsed excitation [94L2, 95L1].

Now we look in more detail into $GaAs/Al_{1-y}Ga_yAs$ multiple quantum well samples with less pronounced localization effects.

In the following example (Figs. 23.35 and 23.36) we consider a MQW sample of high quality with narrow absorption and emission lines and virtually no Stokes shift between emission and absorption. In emission under cw excitation and at a low temperature this sample shows the $n_z = 1hh$ exciton and weakly, on its high energy side, the $n_z = 1lh$ exciton. A shoulder on the low energy side is presumably due to a bound exciton complex or to an excitonic molecule or trion (Sect. 20.3) and a band around 1.53 eV due to a free-to-bound transition (Sect. 14.2). The latter structure saturates rapidly with increasing pump power. In Fig. 23.36a,b we see the decay of the $n_z = 1hh$ luminescence after resonant and non-resonant ps excitation for various pump powers. The most striking feature is that a simple exponential decay is almost never observed.

For low, resonant excitation (curve 1) one observes a rapid decay of the luminescence, which can be attributed to a rapid capture of excitons into some deep, presumably non radiative centers. For increasing excitation this process saturates and for $t > 250$ ps one gets in curve 2 a simple exponential decay with $T_1 \approx 200$ ps. This order of magnitude seems to be representative for the radiative and nonradiative decay of excitons in good MQW samples at low temperatures ($T \lesssim 10$ K). For even higher pump power (curve 3) this value is

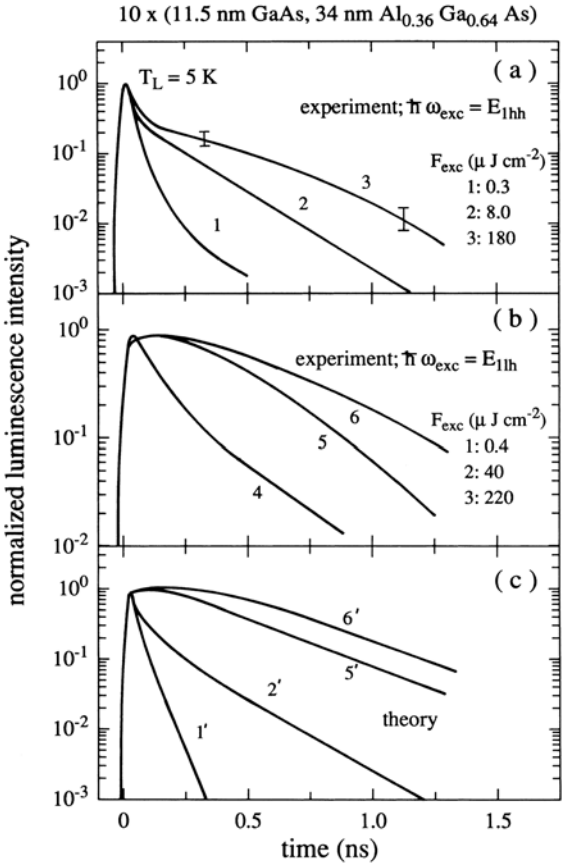


Fig. 23.36. The decay curves of the $1hh$ exciton luminescence for various excitation fluences for resonant (**a**) and non-resonant (**b**) excitation of the $n_z = 1hh$ exciton in an AlGaAs MQW sample, and some calculated decay curves (**c**) [92O1]

reached only at later times and a hump develops in the decay curves, which is even more pronounced in the case of non-resonant excitation with some excess energy (curves 4 to 6).

This feature can be interpreted as a thermalization (or T_3) process. Increasing I_{exc} or $\hbar\omega_{\mathrm{exc}}$ leads to the creation of a hot gas of excitons with a certain average kinetic energy and with finite $\boldsymbol{k}_{\parallel}$ vectors. In contrast to the three-dimensional case, $\boldsymbol{k}_{\parallel}$ is the total momentum $\boldsymbol{k}$ of the excitons in quasi two-dimensional systems, as mentioned above. Due to the conservation of $\boldsymbol{k}_{\parallel}$ in the radiative recombination process only excitons with $\boldsymbol{k}_{\parallel}$ around zero (or excitons in the small tail of localized states caused by well-width fluctuations) can participate in radiative recombination without emission of LO phonons, a process which is usually much weaker in standard III–V materials compared, e.g., to II–VI semiconductors. The humps or plateaux in curves 7 and 8 reflect

the cooling or the formation process of the exciton gas from states with larger towards states with smaller $\boldsymbol{k} = \boldsymbol{k}_\parallel$.

A modeling of the decay along the lines above gives the curves 1′, 2′, 5′, and 6′, in Fig. 23.14c, which correspond to the unprimed experimental ones. The good agreement gives some support to these ideas.

A related experimental finding concerns the fact that in quantum wells one often observes a linear increase of the luminescence decay time T_1 time with lattice temperature for $T_\mathrm{L} \gtrsim 10\,\mathrm{K}$ up to around 100 K. Examples are given in Fig. 23.35 or in [91A1, 91F2, 91G1, 92O1].

$$T_1 = \frac{T_1' k_\mathrm{B} T_\mathrm{L}}{E_\mathrm{m}} \quad (10\,\mathrm{K} \leqslant T_\mathrm{L} \lesssim 100\,\mathrm{K}) \,. \tag{23.46}$$

where T_1'/E_m is for the moment only a characteristic parameter of the material. We will present a simple model from [92O1] which accounts for this effect. For details and partly similar, partly complementary approaches see [91A1, 91F2, 91G1, 92O1, 96P1, 98L1, 00Z2].

The relation to the effect discussed above comes from the consideration that the exciton temperature T_ex has the lattice temperature T_L as a lower limit and is close to this limit for low excitation intensities and small excess energies. In the following we thus assume $T_\mathrm{ex} \approx T_\mathrm{L}$. It is also reasonable to assume that only excitons with small (kinetic) energy $0 \leqslant E \leqslant E_\mathrm{m}$ participate in the recombination process. This condition is fulfilled for radiative decay for free excitons with [see also (23.45b)]

$$\left|\boldsymbol{k}_\parallel\right| \lesssim \frac{2\pi}{\lambda_\mathrm{v}} \approx 8 \times 10^4\,\mathrm{cm}^{-1} \tag{23.47}$$

where $\boldsymbol{k}_\parallel$ is the in-plane wave vector and λ_v the luminescence wavelength in vacuum, due to the conservation of the parallel component of $\boldsymbol{k}$.

The numerical values refer to $\mathrm{Al}_{1-y}\mathrm{Ga}_y\mathrm{As}/\mathrm{GaAs(M)QW}$ and $\boldsymbol{k}_\parallel = 0$ and $\boldsymbol{k}_\parallel = 2\pi/\lambda_\mathrm{v}$ correspond to normal and grazing emission, respectively.

The contribution to E_m resulting from (23.47) is

$$E_\mathrm{m}^0 = \frac{\hbar^2 \boldsymbol{k}_\parallel^2}{2M} = \frac{\hbar^2 2\pi}{M\lambda_\mathrm{v}^2} \approx 10\,\mu\mathrm{eV} \,. \tag{23.48}$$

The value of E_m^0 can be increased by damping and by inhomogeneous broadening. This means that we can include a tail of localized states, provided that its tailing parameter E_0 fulfills the condition $E_0 \lesssim k_\mathrm{B} T_\mathrm{L}$, i.e., the localized excitons are in equilibrium with the free ones. From the absorption and luminescence linewidth of good samples (Fig. 15.3a) we can conclude that

$$E_0 < E_\mathrm{m} \approx 0.5\text{–}1\,\mathrm{meV} \tag{23.49}$$

so that the above conditions are fulfilled for $T_\mathrm{L} \gtrsim 10\,\mathrm{K}$.

For $k_B T_L < E_m$ the excitons decay with an average lifetime around 200 ps (Fig. 23.36b). For $T > 10\,K$ we give the following treatment.

The two-dimensional density of states in effective mass approximation *D(E)* is given according to Sect. 2.6 by

$$D(E) = \begin{cases} D_0 & \text{for} \quad E \geqslant 0 \\ 0 & \text{for} \quad E < 0 \end{cases} \tag{23.50a}$$

with

$$D_0 = g \cdot \frac{1}{2\pi} \frac{2M}{\hbar^2} \approx 10^{12}\,(\mathrm{cm^2 meV})^{-1} \tag{23.50b}$$

possibly including a small exponential tail as mentioned above.

The number of excitons per unit interval of energy $N(E)$ is then given by

$$N(E) = D_0 \exp[-(E-\mu)/k_B T] \quad E \geqslant 0\,, \tag{23.51}$$

where μ is the chemical potential of excitons and where Boltzmann statistics and the weak coupling approach are used.

The total density N_0 at a time t follows from integration over (23.51):

$$N_0(t) = k_B T D_0 \exp[\mu(t)/k_B T] \tag{23.52a}$$

This means that the chemical potential μ can be expressed in terms of N_0 and T

$$\mu = k_B T \ln \frac{N_0}{D_0 k_B T}\,, \tag{23.52b}$$

or

$$N(E) = D_0 \exp\left(-\frac{E}{k_B T} + \ln \frac{N_0}{D_0 k_B T}\right)\,. \tag{23.53}$$

For the Boltzmann approximation to be valid we must have $\mu < 0$ or

$$N_0 < D_0 k_B T \approx 10^{12}\,\mathrm{cm^2} \quad \text{for} \quad T = 10\,K\,. \tag{23.54}$$

This condition is very well fulfilled for low excitation.

For a "monoatomic" recombination process one generally uses the ansatz

$$-\frac{\mathrm{d}N_0}{\mathrm{d}t} = \frac{1}{T_1'} N_0\,. \tag{23.55}$$

This ansatz has to be modified because of momentum conservation, since only a fraction γ of the excitons can participate in the recombination, namely those which are sitting in the interval E_m around the bottom of the band:

$$\gamma = \frac{N(E=0) E_m}{N_0}\,. \tag{23.56}$$

Equation (23.55) has then to be rewritten as

$$-\frac{\mathrm{d}N_0}{\mathrm{d}t} = \frac{1}{T_1'}\gamma N_0 = \frac{1}{T_1'}\frac{N(E=0)E_\mathrm{m}}{N_0}N_0\,, \tag{23.57}$$

which means we have an effective lifetime T_1

$$T_1 = T_1' N(E=0)E_\mathrm{m}/N_0\,. \tag{23.58}$$

With (23.49) and (23.52), (23.53) we finally obtain

$$T_1 = T_1'\frac{k_\mathrm{B}T}{E_\mathrm{m}} \approx 200\,\mathrm{ps}\frac{k_\mathrm{B}T}{1\,\mathrm{meV}} \quad \text{for} \quad T > 10\,\mathrm{K}\,, \tag{23.59}$$

in agreement with the experimental finding.

The linear increase of T_1 with temperature is sometimes used to argue that the recombination of excitons in quantum wells is essentially radiative [91F1], as mentioned above. This assumption may hold in some selected samples in which the spectral width of the free exciton emission is governed by E_m^0 in (23.48). See [03T1]. For other samples, the above approach seems to be more realistic, especially if one recalls that MQW samples with comparable parameters can show different slopes in the relation $T_1 = f(T_\mathrm{L})$ (Fig. 23.35) and that some II–VI superlattices with rather low luminescence yields also show a linear relation as mentioned already above.

We refer the reader for more details and examples to e.g. [88N1,91S1,92P1, 92S1, 92S2, 93U1, 96S1, 01L1, 03P1, 04O1] and references therein.

To conclude this section we give some more references to the recombination dynamics of excitons, trions or biexcitons as a function of sample parameters and excitation [96L1, 98L1, 98Z1, 00E1, 00Y1, 00Z1, 00Z2, 02G1] for various II–VI and III–V quantum wells.

For $Si_{1-x}Ge_y$/Si multiple quantum wells lifetimes on the order of 0.1 µs have been found [95Z1].

For V grooves see, e.g., [98B5, 99O2] and references therein,for ZnO nanorods [04P2] and for quantum dots or islands see [97W1, 99B3, 00B1, 00D1, 00F1, 02S2, 03C2, 03H2]. For the dynamics of eletron–hole plasma see e.g. [00N1, 02N2, 02N3].

Exciting and beautiful new results emerged recently as shown here or can be expected in the field of ultrafast spectroscopy of semiconductors in the near future partly in connection with (extreme) nonlinear optics. They will comprise e.g. the first elementary interaction processes of the electromagnetic radiation field with the electronic system, where memory effects or non-Markovian behaviour occur, or the coherently driven polarization field in the sense of Rabi flopping or the consistent description of the phenomena in the strong coupling limit, i.e. in the polariton picture. An important step into this direction is the transition from the optical Bloch equations to the semiconductor Bloch equations, though this approach involves an

extremely high computational effort see Chap. 27. First results into these directions will be found e.g. in the recent and future issues of [88N1] or in [04O1].

23.5 Problems

1. If you prefer the mathematical description of physics to the intuitive picture given here, then go through the concept of the optical Bloch equations, e.g., with the help of [75A1, 78S1, 84S1] or Chap. 27, and find out what the Rabi frequency, a $\pi/2$- and a π-pulse are, and what the relation to a spin system is (spin echo). By the way, the photon echo also works if pulses other than $\pi/2$ and π are used, but the description is more difficult.
2. Find some data for the electrical conductivity of standard semiconductors. Use the equations given in this chapter to deduce the momentum relaxation of free carriers in electrical transport at various temperatures and compare with the data for excitons. Are the above-mentioned formulas also adequate to analyse the electrical conductivity of a metal?
3. Many of the above effects of ultrashort time-resolved spectroscopy like Rabi oscillations, photon echoes, etc., have been previously found in atomic systems. Find some examples. Why have they been detected so late for semiconductors?
4. There are several ways to create THz pulses. Find some of them in the literature.

References to Chap. 23

[28B1] F. Bloch, Z. Phys. **52**, 555 (1928)
[34Z1] C. Zener, Proc. Royal. Soc. London A **145**, 523 (1934)
[37R1] I.I. Rabi, Phys. Rev. **51**, 652 (1937)
[39R1] I.I. Rabi et al., Phys. Rev. **55**, 526 (1939)
[46B1] F. Bloch, Phys. Rev. **70**, 460 (1946)
[57F1] R.P. Feynmann, F.L. Vornon and R.W. Hellwarth, J. Appl. Phys. **28**, 49 (1957)
[59B1] I. Broser and R. Broser-Warminsky, J. Phys. Chem. Solids **8**, 177 (1959)
[65S1] E. Spenke, Elektronische Halbleiter 2nd edn., Springer, Berlin, Heidelberg (1965)
[75A1] L. Allen and J.H. Eberly, Optical Resonance and Two-Level Atoms, Wiley, New York (1975)
[75H1] A. Haug, Festkörperprobleme **XII**, 411 (1975)
[75W1] P. Wiesner and U. Heim, Phys. Rev. B **11**, 3071 (1975)
[77L1] V.S. Letokhov and V.P. Chebotayev, Nonlinear Laser Spectroscopy, Springer Ser. Opt. Sci. **4**, Springer, Berlin, Heidelberg (1977)
[78S1] R.L. Shoemaker, in Coherence and Laser Spectroscopy, Plenum, New York, p. 197 (1978)

[78Y1] T. Yajima and H. Souma, Phys. Rev. A **17**, 309, 324 (1978)
[79H1] A. Haug, J. Lumin. **20**, 173 (1979) and references therein
[82H1] R. Höpfel et al., Surf. Sci. **113**, 118 (1982)
[83M1] Y. Masumoto, S. Shionoya and T. Takagahara, Phys. Rev. Lett. **51**, 923 (1983)
[84K1] K. Kempf and C. Klingshirn, Solid State Commun. **49**, 23 (1984)
[84S1] Y.R. Shen, The Principles of Nonlinear Optics, Wiley, New York (1984)
[85M1] F.A. Majumder et al., Phys. Rev. B **32**, 2407 (1985)
[85Y1] Y.X. Yan, E.B. Gamble and K.A. Nelson, J. Chem. Phys. **83**, 5391 (1985)
[86H1] D. Heitmann, Surf. Sci. **170**, 332 (1986)
[87M1] U. Merkt, Festkörperprobleme / Adv. Solid State Phys. **27**, 109 (1987)
[87S1] S. Shevel et al., J. Luminesc. **37**, 45 (1987)
[87S2] H.E. Swoboda et al., J. Luminesc. **38**, 79 (1987)
[87Z1] J.J. Zayhowski et al., Phys. Rev. B **35**, 6950 (1987)
[88B1] P.C. Becker et al., Phys. Rev. Lett. **61**, 1647 (1988)
[88C1] D.S. Chemla, S. Schmitt-Rink and D.A.B. Miller, in Optical Nonlinearities and Instabilities in Semiconductors, ed. By H. Haug, Academic, New York, p. 83 (1988)
[88N1] Series of the International Conference "Nonlinear Optics and Excitation Kinetics in Semiconductors", published in:
NOEKS I: phys. stat. sol. (b) **146**, 311–391; **147**, 699–756 (1988)
NOEKS II: phys. stat. sol. (b) **159**, 1–484 (1990)
NOEKS III: phys. stat. sol. (b) **173**, 1–478 (1992)
NOEKS IV: phys. stat. sol. (b) **188**, 9–587 (1995)
NOEKS V: phys. stat. sol. (b) **206** (1997)
NOEKS VI: phys. stat. sol. (b) **221** (2000)
NOEKS VII: phys. stat. sol. (c) **0** (5) (2003)
NOEKS VIII: in preparation
[88W1] Ch. Weber et al., Z. Phys. B **72**, 379 (1988)
[88W2] Ch. Weber et al., Appl. Phys. B **45**, 113 (1988)
[89D1] C. Dörnfeld and J.M. Hvam, IEEE J. QE-**25**, 904 (1989)
[89D2] C. Dörnfeld, PhD Thesis, Kaiserslautern (1989)
[89G1] C. Gourdon and P. Lavallard, phys. stat. sol. (b) **153**, 641 (1989)
[89H1] M. Helm et al., Phys. Rev. Lett. **63**, 74 (1989)
[89K1] J. Kuhl et al., Festkörperprobleme / Adv. Solid State Phys. **29**, 157 (1989)
[90C1] G.C. Cho, W. Kütt and H. Kurz, Phys. Rev. Lett. **65**, 764 (1990)
[90M1] U. Merkt, Festkörperprobleme / Adv. Solid State Phys. **30**, 77 (1990)
[91A1] L.C. Andreani, Solid State Commun. **77**, 641 (1991)
[91B1] J.-Y. Bigot et al., Phys. Rev. Lett. **66**, 1138 (1991)
[91B2] S. Bar Ad and I. Bar-Joseph, Phys. Rev. Lett. **66**, 2491 (1991)
[91D1] W. Demtröder, Laserspektroskopie, 2nd edn., Springer, Berlin, Heidelberg (1991)
[91F1] D. Fröhlich et al., Phys. Rev. Lett. **67**, 2343 (1991)
[91F2] J. Feldmann et al., Phys. Rev. Lett. **67**, 2355 (1991)
[91G1] M. Gurioli et al., Phys. Rev. B **44**,3115 (1991)
[91O1] D. Oberhauser et al., Superlat. Microstruct. **9**, 107 (1991)
[91S1] H. Stolz, Festkörperprobleme / Adv. Solid State Phys. **32**, 219 (1991)
[91S2] H. Schwab, PhD Thesis, Kaiserslautern (1991)
[91U1] M. Urban, H. Schwab and C. Klingshirn, phys. stat. sol. (b) **166**, 423 (1991)

[91V1] F. Vallée, F. Bogani and C. Flytzanis, Phys. Rev. Lett. **66**, 1509 (1991)
[91W1] M. Wegener et al., Phys. Rev. A **44**, 2124 (1991)
[92B1] S. Bar Ad et al., Phys. Rev. Lett. **68**, 349 (1992)
[92F1] J. Feldmann, Festkörperprobleme / Adv. Solid State Phys. **32**, 81 (1992)
[92G1] E.O. Göbel et al., phys. stat. sol. (b) **173**, 21 (1992)
[92K1] M. Koch et al., Phys. Rev. Lett. **69**, 3631 (1992)
[92K2] W. Kütt, Festkörperprobleme / Adv. Solid State Phys. **32**, 113 (1992)
[92L1] K. Leo, Festkörperprobleme / Adv. Solid State Phys. **32**, 97 (1992)
[92O1] D. Oberhauser et al., phys. stat. sol. (b) **173**, 53 (1992)
[92O2] D. Oberhauser et al., J. Lumin. **53**, 409 (1992)
[92P1] K.-H. Pantke et al., phys. stat. sol. (b) **173**, 69 (1992)
[92P2] K.-H. Pantke, Proc. 21th ICPS, Ping Jiang and Hou-Zhi Zheng eds. World Scientific, Singapore, p. 129 (1992)
[92R1] H.G. Roskos et al., Phys. Rev. Lett. **68**, 2216 (1992)
[92S1] H. Schwab et al., phys. stat. sol. (b) **172**, 479 (1992)
[92S2] U. Siegner et al., Phys. Rev. B **46**, 4564 (1992)
[92S3] H. Schwab and C. Klingshirn, Phys. Rev. B **45**, 6938 (1992)
[92S4] H. Schwab et al., Phys. Rev. B **46**, 7528 (1992)
[92S5] W. Sack et al., J. Luminesc. **53**, 409 (1992)
[92Z1] H.J. Zeiger et al., Phys. Rev. B **45**, 768 (1992)
[93K1] H. Kalt et al., Physica B **191**, 90 (1993)
[93K2] M. Koch et al., Phys. Rev. B **47**, 1532 (1993)
[93M1] K. Messner et al., Phys. Rev. B **48**, R15472 (1993)
[93O1] D. Oberhauser et al., Phys. Rev. B **47**, 6827 (1993)
[93P1] K.-H. Pantke et al., Phys. Rev. B **47**, 2413 (1993)
[93P2] K.-H. Pantke et al., Phys. Rev. Lett. **70**, 327 (1993)
[93P3] P.C.M. Planken et al., Phys. Rev. B **48**, 4903 (1993)
[93S1] R.W. Schoenlein et al., Phys. Rev. Lett. **70**, 1014 (1993)
[93U1] Ultrafast Phenomena VIII, eds.: J.-L. Martin, A. Migus, G.A. Mourou and A.H. Zewail, Springer Ser. Chem. Phys. **55**, Springer, Berlin, Heidelberg (1993) and previous Vols. **4**, **14**, **23**, **38**, **46**, **48**, **53** of this series and Ultrashort Processes in Condensed Matter W.E. Bron ed., NATO ASI Series B **314**, (1993)
[93W1] U. Woggon et al., Phys. Rev. B **47**, 3684 (1993)
[93W2] S. Weiss et al., Appl. Phys. Lett. **63**, 2567 (1993)
[94A1] H. Akyiama et al., Phys. Rev. Lett. **72**, 924 (1994)
[94B1] S. Bar Ad et al., Phys. Rev. Lett. **72**, 776 (1994)
[94C1] S.T. Cundiff et al., Phys. Rev. Lett. **73**, 1178 (1994)
[94E1] J. Erland et al., J. Crysal Growth **138**, 800 (1994)
[94H1] S. Hunsche et al., Phys. Rev. B **50**, 5791 (1994)
[94L1] P. Leisching et al., Phys. Rev. B **50**, 14389 (1994)
[94L2] W. Langbein et al., Appl. Phys. Lett. **65**, 2465 (1994)
[94M1] D.M. Mittlemen et al., Phys. Rev. B **49**, 14435 (1994)
[94M2] R.M. Macfarlane in refs. [81A1] of Chap. 1 g p 151 and h p 471
[94P1] K.-H. Pantke and J.M. Hvam, J. Mod. Phy. B **8**, 73 (1994)
[94R1] H.G. Roskos, Festkörperprobleme / Adv. Solid State Phys. **34**, 297 (1994)
[94S1] H. Stolz, Time Resolved Light Scattering from Excitons, Springer, Berlin (1994)
[94V1] K. Victor, H.G. Roskos and Ch. Waschke, J. Opt. Soc. Am. B (1994) **11**, 2470 (1994)

[94W1] U. Woggon et al., J. Crystal Growth **138**, 976 and 988 (1994)
[95B1] L. Banyai et al., Phys. Rev. Lett. **75**, 2188 (1995)
[95C1] S.T. Cundiff et al., phys. stat. sol. (b) **188**, 307 (1995)
[95D1] T. Dietl et al., Phys. Rev. Lett. **74**, 474 (1995)
[95H1] A.P. Heberle, J.J. Baumberg and K. Köhler, Phys. Rev. Lett. **75**, 2598 (1995)
[95L1] W. Langbein, M. Hetterich and C. Klingshirn, Phys. Rev. B **51**, 9922 (1995)
[95Z1] A. Zrenner et al., Phys. Rev. B **52**, 16608 (1995).
[96B1] D. Birkedal et al., Phys. Rev. Lett. **76**, 672 (1996)
[96B2] A. Bonvalet et al., Phys. Rev. Lett. **84**, 4392 (1996)
[96C1] S.T. Cundiff et al., JOSA B **13**, S. 1263 (1996)
[96E1] S.A. Empedokles, D.J. Norris and M.G. Bawendi, Phys. Rev. Lett. **77**, 3873 (1996)
[96G1] H. Giessen et al., JOSA B **13**, 1039 (1996)
[96H1] H. Haug and A.-P. Jauho, Quantum Kinetics in Transport and Optics of Semiconductors, Springer Series in Solid State Sciences **123**, Springer Berlin (1996)
[96H2] R. Hellmann et al., J. Crystal Growth **159**, 976 (1996)
[96L1] W. Langbein, H. Kalt and J.M. Hvam, Phys. Rev. B **54**, 14589 (1996)
[96N1] U. Neukirch et al., phys. stat. sol. (b) **196**, 473 (1996)
[96O1] M. Oestreich, S. Hallstein and M. Rühle, IEEE J. Selected Topics Quantum Electron. **2**, 747 (1996) and Phys. Rev. B **53**, 7911 (1996)
[96P1] C. Piermarochi et al., Phys. Rev. B **53**, 15834 (1996)
[96S1] J. Shah, Ultrafast Spectroscopy of Semiconductors and Semiconductor Nanostructures, Springer Series in Solid State Sciences **115**, Springer, Berlin (1996)
[96S2] H. Schneider et al., Superlattices and Microstruct. **19**, 347 (1996)
[96T1] S. Ten et al., Phys. Rev. B **53**, 12637 (1996)
[96W1] K. Wundke et al., Phys. Rev. B **53**, 10973 (1996)
[97B1] J.J. Baumberg et al., phys. stat. sol. (b) **204**, 9 (1997)
[97C1] S.A. Croober et al., Phys. Rev. B **56**, 7574 (1997)
[97E1] K. El-Sayed et al., Phys. Rev. B **55**, 2456 (1997)
[97F1] C. Fürst et al., phys. stat. sol. (b) **204**, 20 (1997)
[97L1] W. Langbein et al., Phys. Rev. B **55**, R7383 (1997)
[97L2] V.G. Lyssenko et al., Phys. Rev. Lett. **79**, 301 (1997)
[97M1] G. Manzke, U. Moldzio and K. Henneberger, phys. stat. sol. (b) **202**, 961 (1997)
[97M2] E. Mazur, in Spectroscopy and Dynamics of Collective Excitations in Solids, NATO ASI series B **356**, 417 (1997)
[97N1] S. Nüsse et al., Phys. Rev. B **55**, 4620 (1997)
[97O1] M. Oestreich et al., Festkörperprobleme / Adv. Solid State Phys. **37**, 245 (1997)
[97W1] U. Woggon, Optical Properties of Semiconductor Quantum Dots, Springer Tracts in Mod. Physics **136** (1997)
[97W2] M.U. Wehner, M.H. Ulm and M. Wegener, Opt. Lett. **22**, 1455 (1997)
[97W3] R. Westphäling et al., J. Luminesc. **72-74**, 980 (1997)
[98B1] N.H. Bonadeo et al., Phys. Rev. Lett. **81**, 2759 (1998)
[98B2] D. Birkedal, J. Shah, Phys. Rev. Lett. **81**, 2372 (1998)
[98B3] N.H. Bonadeo et al., Science **282**, 147 (1998)

[98B4] L. Banyai et al., Phys. Rev. Lett. **81**, 882 (1998)
[98B5] J. Bellessa et al., Phys. Rev. B **58**, 9933 (1998)
[98G1] H. Giessen et al., Phys. Rev. Lett. **81**, 4620 (1998)
[98H1] J.M. Hvam, NATO ASI Series B **372**, 357 (1998)
[98K1] A.A. Klochikhin et al., Physics of Solid State **40**, 821 (1998)
[98K2] M. Kira, F. Jahnke and S.W. Koch, Phys. Rev. Lett. **81**, 3263 (1998)
[98L1] P. Lefebvre et al., Phys. Rev. B **57**, R9447 (1998)
[98M1] P. Michler et al., phys. stat. sol. (b) **206**, 399 (1988)
[98M2] U. Moldzio, G. Manzke and K. Henneberger, phys. stat. sol. (b) **206**, 265 (1998)
[98N1] U. Neukirch et al., Phys. Rev. B **57**, 9208 (1998)
[98R1] A. Richter et al., Appl. Phys. Lett. **73**, 2176 (1998)
[98S1] M. Sudzius et al., Phys. Rev. B **57**, R12693 (1998)
[98U1] M. Umlauff et al., Phys. Rev. B **57**, 1390 (1998)
[98U2] A.V. Uskov et al., Appl. Phys. Lett. **72**, 58 (1998)
[98W1] H.P. Wagner et al., Phys. Rev. B **57**, 1791 (1998)
[98W2] H.P. Wagner et al., Phys. Rev. B **57**, 1797 (1998)
[98W3] M.W. Wehner et al., Phys. Rev. Lett. **80**, 1992 (1998)
[98W4] M.U. Wehner, D.S. Chemla and M. Wegener, Phys. Rev. B **58**, 3590 (1998)
[98W5] R. Westphäling et al., J. Appl. Phys. **84**, 6871 (1998) and J. Crystal Growth **184/185**, 1072 (1998)
[98Z1] S. Zimmermann et al., Appl. Phys. Lett. **73**, 154 (1998)
[99A1] D.D. Awschalom and J.M. Kikkawa, Physics Today **52**, 33 (1999)
[99B1] P. Borri et al., Phys. Rev. B **59**, 2215 (1999)
[99B2] P. Borri et al., Phys. Rev. B **60**, 4505 (1999)
[99B3] G. Bacher et al., Phys. Rev. Lett. **83**, 4417 (1999)
[99F1] R. Ferreira and G. Bastard, Appl. Phys. Lett. **74**, 2818 (1999)
[99H1] W.A. Hügel et al., Phys. Rev. Lett. **83**, 3313 (1999)
[99H2] P. Hawker, J. Kent and M. Henini, Appl. Phys. Lett. **75**, 3832 (1999)
[99J1] S. Jorda, Phys. Blätter **55**, issue 2, p 12 (1999)
[99K1] A.A. Klochikhin et al., Phys. Rev. B **59**, 12947 (1999)
[99L1] W. Langbein, H. Gislason and J.M. Hvam, Phys. Rev. B **60**, 16667 (1999)
[99L2] W. Langbein, J.M. Hvam and R. Zimmermann, Phys. Rev. Lett. **82**, 1040 (1999)
[99M1] D. Morris, N. Perret and S. Fafard, Appl. Phys. Lett. **75**, 3593 (1999)
[99M2] J.H. Müller, Phys. Blätter **55**, issue 5, p 14 (1999)
[99O1] R. Otremba et al., Sol. State Commun. **109**, 317 (1999)
[99O2] D.Y. Oberli et al., Phys. Rev. B **59**, 2910 (1999)
[99S1] A.V. Shchegrov, D. Birkedal and J. Shah, Phys. Rev. Lett. **83**, 1391 (1999)
[99S2] D. Steinbach et al., Phys. Rev. B **60**, 12079 (1999)
[99S3] A. Schülzgen et al., Phys. Rev. Lett. **82**, 2346 (1999)
[99W1] H.P. Wagner et al., Phys. Rev. B **60**, 10640 (1999)
[00B1] D. Birkedahl et al., Appl. Phys. Lett. **77**, 2201 (2000)
[00B2] P. Borri et al., Appl. Phys. Lett. **76**, 1380 (2000)
[00D1] E. Dekel et al., Phys. Rev. B **62**, 11038 (2000)
[00E1] A. Esser et al., Phys. Rev. B **62**, 8232 (2000)
[00E2] T. Esslinger, I. Bloch and T. Hänsch, Phys. Blätter **56**, issue 2, p 47 (2000)
[00F1] A. Fiore et al., Appl. Phys. Lett. **76**, 3430 (2000)
[00H1] P.G. Huggard et al., Phys. Rev. Lett. **84**, 1023 (2000)

[00H2] W.A. Hügel, M.F. Heinrich and M. Wegener, Phys. Rev. B **62**, 2686 (2000)
[00H3] W.A. Hügel, M.F. Heinrich and M. Wegener, phys. stat. sol. (b) **221**, 473 (2000)
[00K1] H. Kalt et al., phys. stat. sol. (b) **221**, 477 (2000)
[00K2] A. Klochikhin et al., JETP Lett. **72**, 320 (2000)
[00L1] W. Langbein and J.M. Hvam, Phys. Rev. B **61**, 1692 (2000)
[00L2] F. Löser et al., Phys. Rev. B **61**, R13373 (2000)
[00L3] F. Löser et al., Phys. Rev. Lett. **85**, 4763 (2000)
[00L4] S. Lan et al., Phys. Rev. B **61**, 16847 (2000)
[00N1] N. Nagai and M. Kuwata-Gonokami, phys. stat. sol. b **221**, 261 (2000)
[00O1] T. Ogasawara et al., Phys. Rev. Lett. **85**, 2204 (2000)
[00P1] J.P. Prineas et al., Phys. Rev. Lett. **85**, 3041 (2000)
[00R1] S. Raymond et al., Phys. Rev. B **61**, R16331 (2000)
[00R2] M. Rauner, Phys. Blätter **56**, issue 2, p 10 (2000)
[00S1] G.B. Serapiglia et al., Phys. Rev. Lett. **84**, 1019 (2000)
[00V1] Q.T. Vu et al., Phys. Rev. Lett. **85**, 3508 (2000)
[00W1] U. Woggon et al., Phys. Rev. B **61**, 1935 (2000)
[00Y1] S. Yamaguchi et al., Phys. Rev. B **61**, 10303 (2000)
[00Z1] K.C. Zeng et al., Appl. Phys. Lett. **76**, 864, 1728 and 3040 (2000)
[00Z2] B.P. Zhang et al., J. Appl. Phys. **88**, 4916 (2000)
[01A1] V.M. Axt et al., Phys. Rev. B **63**, 115303 (2001)
[01B1] D. Birkedal, K. Leosson and J.M. Hvam, Phys. Rev. Lett. **87**, 227401 (2001)
[01B2] P. Borri et al., Phys. Rev. Lett. **87**, 157401 (2001)
[01B3] R. v. Baltz, in Advances in Energy Transfer Processes, B. Di Bartolo and X. Chen eds. World Scientific, p. 73, Singapore (2001)
[01B4] M. Betz et al., Phys. Rev. Lett. **86**, 4684 (2001)
[01C1] D.S. Chemla and J. Shah, Nature **411**, 549 (2001)
[01D1] Th. Dekorsky, Phys. Blätter **57**, issue 7, 8, p. 67 (2001)
[01D2] B. Devaud et al., Comptes Rendus de l'Academie des Sciences, Series IV, France, Vol. **2**, 1439 (2001)
[01F1] T. Flissikowski et al., Phys. Rev. Lett. **86**, 3172 (2001)
[01F2] T. Flissikowski et al., Phys. Rev. B **68**, 161309 (R) (2001)
[01H1] D. Hägele et al., Sol. State Commun. **120**, 73 (2001)
[01H2] R. Huber et al., Nature **414**, 286 (2001)
[01H3] H. Haug, Nature **414**, 261 (2001)
[01J1] M. Jörger et al., Phys. Rev. B **64**, 113204 (2001)
[01L1] Landolt-Börnstein, New Series, Group III, Vol. 34C1 and the forth coming subvolumes C2 and C3, C. Klingshirn, ed., Springer, Berlin (2001)
[01L2] W. Langbein et al., JOSA B **18**, 1318 (2001)
[01M1] Ch. Mann et al., Phys. Rev. B **64**, 235206 (2001)
[01M2] O.D. Mücke et al., Phys. Rev. Lett. **87**, 057401 (2001)
[01M3] T. Müller et al., Appl. Phys. Lett. **79**, 2755 (2001)
[01N1] N.C. Nielsen et al., Phys. Rev. B **64**, 245202 (2001)
[01O1] M. Oestreich et al., Festkörperprobleme / Adv. Solid State Phys. **41**, 173 (2001)
[01P1] M. Paillard et al., Phys. Rev. Lett. **86**, 1634 (2001)
[01R1] B. Rosam et al., Phys. Rev. Lett. **86**, 1307 (2001)
[01S1] T.H. Stievater, Phys. Rev. Lett. **87**, 133603 (2001)
[01S2] M. Schmidt et al., phys. stat. sol. (b) **229**, 643 (2001)

[01T1] S.A. Tarasenko et al., Semicon. Science and Technol. **16**, 486 (2001)
[01W1] M. Wegener, in Advances in Energy Transfer Processes, B. Di Bartolo and X. Chen, eds. World Scientific, p. 215, Singapore (2001)
[01W2] M. Wegener, Adv. In Solid State Physics **41**, 89 (2001)
[01Y1] K.J. Yee et al., Phys. Rev. Lett. **86**, 1630 (2001)
[02B1] P. Borri et al., Phys. Rev. Lett. **87**, 187401 (2002)
[02B2] M. Bayer and A. Forchel, Phys. Rev. B **65**, 041308 (2002)
[02B3] P. Borri et al., Phys. Rev. B **66**, 081306(R) (2002)
[02B4] M. Betz et al., phys. stat. sol. (b) **233**, 401 and ibid. **231**, 181 (2002)
[02C1] S. Cortez et al., Phys. Rev. Lett. **89**, 207401 (2002)
[02D1] B. Dal Don et al., phys. stat. sol. (b) **229**, 463 (2002)
[02G1] J. Gutowski et al., phys. stat. sol. (b) **229**, 653 (2002)
[02H1] G.R. Hayes and B. Ohkawa, phys. stat. sol. a **190**, 637 (2002)
[02L1] V.G. Lyssenko, in ref. [81A1]k of Chap. 1 p 323
[02K1] G. Kocherscheidt et al., Phys. Rev. B **66**, 161314 (2002)
[02K2] H. Kamada et al., phys. stat. sol. (a) **190**, 485 (2002)
[02K3] T. Kuroda et al., Phys. Rev. B **66**, 121302 (2002)
[02K4] E. Kurtz et al., Thin Solid Films 412 **89** (2002)
[02N1] M.C. Netti et al., Appl. Phys. Lett. **81**, 3927 (2002)
[02N2] M. Nagai and M. Kuwata-Gonokami, Appl. Phys. Lett. **81**, 484 (2002)
[02N3] M. Nagai and M. Kuwata-Gonokami, J. Phys. Soc. Japan **71**, 2276 (2002)
[02R1] A. Reznitsky et al., phys. stat. sol. (c) **0**, issue 1, 280 (2002)
[02S1] W. Schäfer and M. Wegener, Semiconductor Optics and Transport Phenomena, Springer, Berlin (2002)
[02S2] Ch. Santori et al., Phys. Rev. B **65**, 073310 (2002)
[02T1] E. Tsitsishvili, R. v. Baltz and H. Kalt, Phys. Rev. B **66**, 161405 (2002)
[02W1] S. Wachter et al., Phys. Rev. B **65**, 205314 (2002) and Physica B **314**, 309 (2002)
[02Z1] H. Zhao, S. Moehl and H. Kalt, Appl. Phys. Lett. **80**, 1391; **81**, 2794 (2002) and Phys. Rev. Lett. **89**, 097401 (2002)
[03C1] Coherence and Quantum Optics VIII, N.P. Bigelow, J.H. Eberly, C.R. Stroud Jr., and I.A. Wamsley eds. Kluwer, Dordrecht (2003) and the preceeding volumes of this series
[03C2] S.A. Crooker et al., Appl. Phys. Lett. **82**, 2793 (2003)
[03D1] D. Di Bartolo, in ref. [81A1]l of Chap. 1
[03D2] M.M. Dignam and M. Hawton, Phys. Rev. B **67**, 35291 (2003)
[03F1] T. Fleck, M. Schmidt and C. Klingshirn, phys. stat. sol. (a) **198**, 248 (2003)
[03H1] M. Hawton and M.M. Dignam, Phys. Rev. Lett. **91**, 267402 (2003)
[03H2] H. Htoon et al., Appl. Phys. Lett. **82**, 4776 (2003)
[03K1] R.A. Kaindl et al., phys. stat. sol. (b) **238**, 451 (2003)
[03K2] S.W. Koch et al., phy. stat. sol. (b) **238**, 404 and 443 (2003)
[03M1] G. Mannarini et al., phys. stat. sol. (b) **238**, 494 (2003)
[03P1] B. Patton, W. Langbein and U. Woggon, Phys. Rev. B **68**, 125316 (2003)
[03Q1] Quantum Coherence, Correlation and Decoherence in Semiconductor Nanostructures, T. Takagahara, ed., Academic Press, Amsterdam (2003)
[03R1] A. Reznitsky et al., phys. stat. sol. (c) **0**, issue 5, 1544 (2003)
[03S1] T. Serenyi, C. Benedek and M.G. Benedict, Fortschr. Phys. **51**, 226 (2003)
[03S2] M. Schmidt et al., Proc. 26th ICPS in Institute of Physics Conf. Series **171**, 82 (2003)

[03T1] T. Tritschler et al., SPIE Proc. **4992**, 33 (2003) and phys. stat. sol. (b) **238**, 561 (2003)
[03V1] T. Voss et al., Optics Commun. **218**, 415 (2003)
[03Z1] H. Zhao et al., Phys. Rev. B **67**, 353300 (2003)
[04C1] S. Chatterjee et al., Phys. Rev. Lett. **92**, 067402 (2004)
[04D1] B. DalDon et al., phys. stat. sol. c **1** (3), 462 (2004)
[04H1] K. Hazu et al., J. Appl. Phys. **96**, 1270 (2004)
[04K1] A. Klochikhin et al., Phys. Rev. B **69**, 085304 (2004)
[04K2] M. Kira, W. Hoyer and S.W. Koch, Solid State Commun. **129**, 733 (2004)
[04L1] V. Lysenko et al., in Ref. [81A1]l of Chap. 1, in press
[04L2] W. Langbein, P. Borri and U. Woggon, Phys. Rev. B **70**, 033301 (2004)
[04M1] G. Mannarini et al., Phys. Rev. B **69**, 085326 (2004)
[04M2] G. Mannarini, W. Langbein and R. Zimmermann, phys. stat. sol. c **1** (3), 489 (2004)
[04O1] Optics of Semiconductors and Their Nanostructures, H. Kalt and M. Hetterich (eds.), Springer Series in Solid State Science **146**, (2004)
[04P1] H. Priller et al., phys. stat. sol. (c) **1**, issue 4, 747 (2004)
[04P2] H. Priller et al., Proc. EXCON '04, Cracow, J. Luminesc., in press (2004)
[04P3] P. Płochocka et al., Phys. Rev. Lett. **92**, 177402 (2004)
[04T1] E. Tsitsishvili, R.v. Balz and H. Kalt, phys. stat. sol. c **1** (3), 564 (2004)
[04Z1] R. Zimmermann and E. Runge, Phys. Rev. B **69**, 155307 (2004)
[04Z2] H. Zhao and H. Kalt, Phys. Rev. B **69**, 233305 (2004)

24

Optical Bistability, Optical Computing, Spintronics and Quantum Computing

In this chapter we present some of the properties of optical bistability, an effect that is not limited to semiconductors, and some of the concepts of digital optical computing. This concept is based on optical nonlinearities and bi- or multistability. We also explain why digital optical computing did not make its way into broad commercial use. Then we proceed to spintronics and quantum computing. The latter concept especially relies, to a limited extent only, on semiconductor optics. We introduce these ideas here because the author feels that beautiful physics has been accomplished in all three fields, though he also sees strong analogies to the development and the fate of the concept of digital optical computing for the combinations of semiconductor optics with spintronics and quantum computing.

24.1 Optical Bistability

Optical bistability in semiconductors, but also in other materials like Na vapor or liquid crystals, was a hot topic of research in the 1980s and the beginning of the 1990s. Although there are presently still about 500 papers per year on optical bistability and 400 on optical computing, the interest in optical bistability in semiconductors has faded, especially when it became clear that the hope for a digital optical computer based on optically bistable elements is unrealistic.

We nevertheless address this topic here in some breadth because beautiful physics came out of this research and because something can also be learned from it for the topics of spintronics and quantum computing, which we present towards the end of this chapter.

The development of optical bistability and optical computing is nicely documented in the proceedings of a series of conferences and summer schools that we mention here [80O1, 84O1, 84O2, 85G1, 86O1, 87F1, 88O1, 88O2, 88O3, 88P1, 89O1, 92N1, 93O1, 98D1], where also the history of this field is described.

24.1.1 Basic Concepts and Mechanisms

An optically bistable element shows a static hysteresis loop in the relation between incident and transmitted (or reflected) light intensities. In Fig. 24.1 we show two possible input-output characteristics. In the case Fig. 24.1a the device switches with increasing incident power at $I\uparrow$ from a state of low, into a state of high transmission and stays there for further increasing incident intensity I_{in}, but also for decreasing intensity. Only at $I_{\text{in}} = I\downarrow < I\uparrow$ does the system switch back to the state of low transmission.

In Fig. 24.1b the device switches with increasing incident intensity at $I\downarrow$, from a high transmission a lower one and back at $I\uparrow < I\downarrow$. This means that the hysteresis loops in Fig. 24.1a and b are rotated anticlockwise and clockwise, respectively.

In both cases one has two stable output states between $I\uparrow$ and $I\downarrow$ and which one of them is realized depends on the history, i.e., on whether one comes from higher or lower intensities. This is the essence of optical bistability.

The dashed lines in Fig. 24.1 show a third, unstable branch. If the system is prepared by some means on this branch it can in principle stay there forever. However, the slightest deviation from this unstable branch (e.g., due to fluctuations) causes the system to move into a state on the upper or lower stable branch, as indicated by the dotted arrows. The dotted line thus separates the "basins of attraction" of upper and lower branches and is therefore also called "separatrix".

Only two ingredients are necessary to create optical bistability, namely a sufficiently strong (electro-)optical nonlinearity, e.g., of the types we have seen in Chaps. 19–22 or Sect. 16.2, and a suitable feedback.

It is not necessary that the optically nonlinear material is a semiconductor, it can equally be a gas like sodium vapor, but in line with the title of this book we concentrate here exclusively on semiconductors.

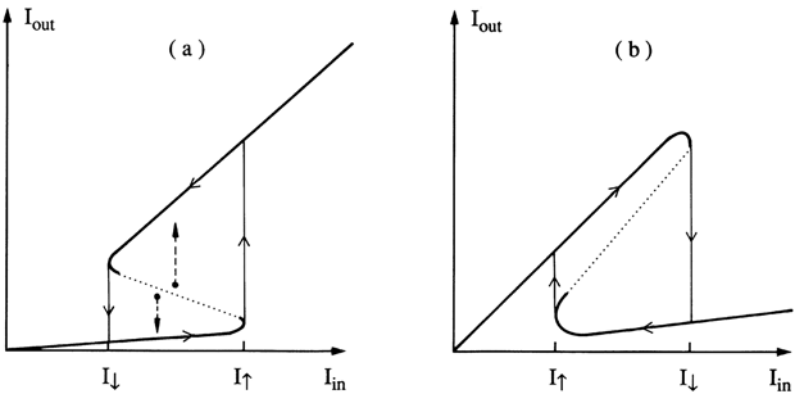


Fig. 24.1. Input–output characteristics of optically bistable elements (**a**) anticlockwise; (**b**) clockwise

A hysteresis loop which is rotated anticlockwise as in Fig. 24.1a can be obtained in transmission with a dispersive nonlinearity and the feedback from a Fabry–Perot resonator. In reflection one then observes a hysteresis which is rotated clockwise as in Fig. 24.1b. The hysteresis of Fig. 24.1a in transmission can be also obtained with bleaching of absorption and again a Fabry–Perot, while a hysteresis like Fig. 24.1b is obtained in transmission with excitation-induced absorption. In this case the feedback is "built-in", as we shall see later.

To further clarify the terminology we can say that dispersive nonlinearity means a change of the real part of the refractive index with increasing light intensity I i.e. $\Delta n(I) \gtrless 0$, bleaching means an excitation-induced decrease of the absorption coefficient or of the imaginary part of $\tilde{n}$ i.e., $\Delta\alpha(I) < 0$ or $\Delta\kappa < 0$, and induced absorption means finally an increase of α or κ with increasing excitation, i.e., $\Delta\alpha(I) > 0$ or $\Delta\kappa(I) > 0$.

There are also some other mechanisms which can lead to optical bistability, but those mentioned above are the most important ones and we shall concentrate on these in the following examples.

24.1.2 Dispersive Optical Bistability

For dispersive optical bistability one starts with an incident wavelength corresponding to a transmission minimum of the Fabry–Perot (FP) resonator (Sect. 3.1.6). In this situation one has destructive interference of the partial waves reflected from the partly transmitting mirrors at the front and back of the FP. Consequently most of the incident intensity I_{in} is reflected and the intracavity intensity I_{intra} as well as the transmitted intensity I_{out} are much smaller then I_{in}. Now we increase I_{in}. Then I_{intra} increases, too. If I_{inlra} becomes so large that it changes the refractive index n of the medium in the cavity, the positive feedback sets in: a change on n causes a change of the phase shift per round trip δ. As a consequence, the interference of the partial waves becomes more constructive whatever the sign of Δn, since we started with maximum destructive interference. Thus, I_{intra} increases, which changes δ even more, and so on. If this positive feedback is sufficiently strong, at a certain incident intensity $I\uparrow$, it can cause an abrupt transition to a state of high transmission close to a transmission maximum of the FP. Then we have a high I_{intra} which can even exceed I_{in}. We can then lower I_{in} below $I\uparrow$ and still keep the system in the highly transmitting state. Only at some value $I\downarrow$ of I_{in} with $I\downarrow < I\uparrow$ does the system return to the low transmission state.

To complement these intuitive arguments we now describe the process mathematically following the sets of equations given in several of the above references. We have two equations for the transmission of the Fabry–Perot resonator which have to be solved simultaneously. One is the (Airy-) function from Sect. 3.1.6, repeated here for convenience:

$$T_{\text{FP}} = (1-R)^2\{[\exp(\alpha L/2) - R\exp(-\alpha L/2)]^2 + 4R\sin^2\delta\}^{-1} = \frac{I_{\text{out}}}{I_{\text{in}}} \quad (24.1a)$$

$$T_{\mathrm{FP}} \approx \frac{(1-R)^2}{(1-R)^2 + 4R\sin^2\delta} \quad \text{for} \quad \alpha L \ll 1\,, \tag{24.1b}$$

where R is the reflectivity of one mirror.

Equation (24.1b) holds for weak absorption, and we assume for the moment that α is small and does not change much with I_{intra} so that (24.1b) can be used.

The refractive index and thus the phase shift δ is assumed to depend on I_{intra} in the simplest case in a linear way (see Sects. 19.1 and 19.2):

$$\delta = \delta_0 + \Delta\delta(I_{\mathrm{intra}}) = n\omega L c^{-1}(1 + n_2 I_{\mathrm{intra}}) \tag{24.2a}$$

$$\text{or} \quad \Delta\delta \propto I_{\mathrm{intra}}\,. \tag{24.2b}$$

The second equation for the transmission reads

$$T_{\mathrm{FP}} = I_{\mathrm{intra}} \left(\frac{1-R}{1+R}\right) \frac{1}{I_{\mathrm{in}}} \tag{24.3}$$

Equations (24.1) and (24.2) give the transmission as a function of I_{intra}, and (24.3) gives a straight line with a slope inversely proportional to I_{in}.

In Fig. 24.2a we show both curves. The solutions are the intersections and are shown in Fig. 24.2b. If we start with small I_{in} we have an almost vertical line from (24.3) and only one solution. With increasing I_{in} two new solutions appear, but the system still remains in the low transmission state. At $I_{\mathrm{in}} = I\uparrow$ this lower solution disappears (point A) and the FP jumps to the new solution (point B). The spike in Fig. 24.2b is a transient feature which can be observed if the transition from A to B is so slow that the system can be followed through the transmission maximum. For further increasing I_{in} we remain on the upper branch. For decreasing I_{in} this remains true down to $I\downarrow$ where this solution disappears (point C) and the system has to return to the lower state at point D.

The intermediate solution appearing in Fig. 24.2a, e.g., at point E and indicated in Figs. 24.1a and 24.2b by dotted lines, can be shown to be unstable in the sense mentioned above.

Since for low absorption the intensity reflected from a FP is just the complement of the transmitted intensity, i.e.,

$$\frac{I_{\mathrm{r}}}{I_0} = 1 - \frac{I_{\mathrm{t}}}{I_{\mathrm{in}}}\,, \tag{24.4}$$

one can observe a hysteresis which is rotated clockwise in reflection.

In Fig. 24.3a,b we show this situation for an idealized, lossless FP and Fig. 24.3c, gives an example. The nonlinear medium is a semiconductor-doped glass, coated on both plane parallel surfaces by dielectric mirrors to form the FP. The optical nonlinearity is a photo-thermal one (Sect. 20.6), namely the increase of the refractive index with increasing temperature caused by the small absorbed fraction of the intracavity intensity.

Figure 24.4 shows another example of dispersive optical bistability. In this case the dispersive nonlinearity of CdS observed under high excitation is exploited (Fig. 21.11). The surfaces of the platelet type sample are coated with dielectric mirrors with $R \approx 0.6$. In this case pulses of a duration of only 15 ns are used. The switching processes, which occur on a (sub)ns time scale, are clearly seen in Fig. 24.4a,b. The resulting hysteresis loop is given in Fig. 24.4c.

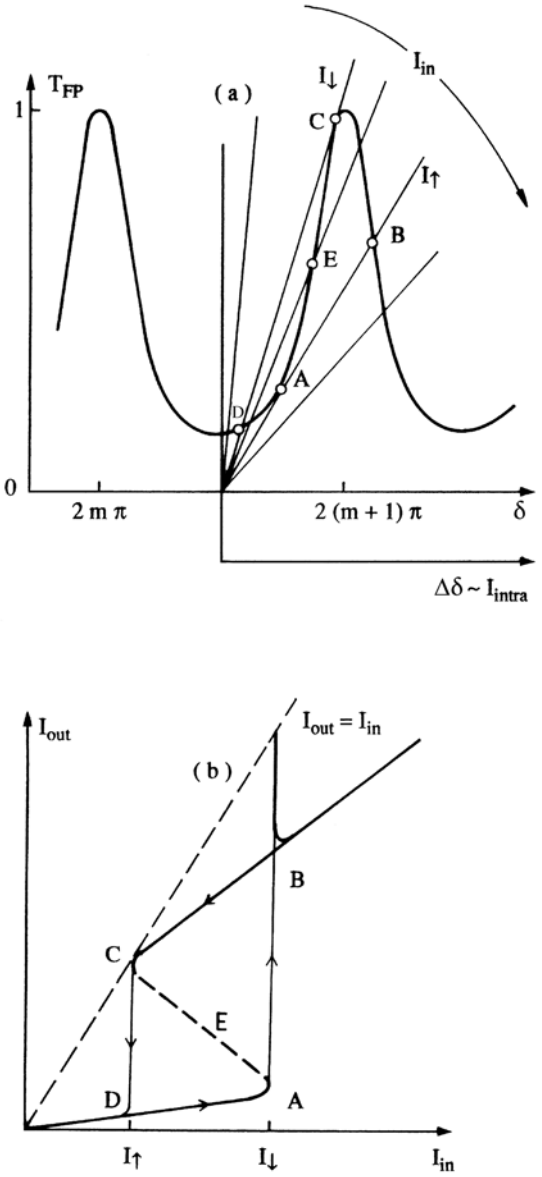


Fig. 24.2. The graphical solution of ((24.1b), (24.2), (24.3)) (**a**) and the resulting input–output characteristic (**b**)

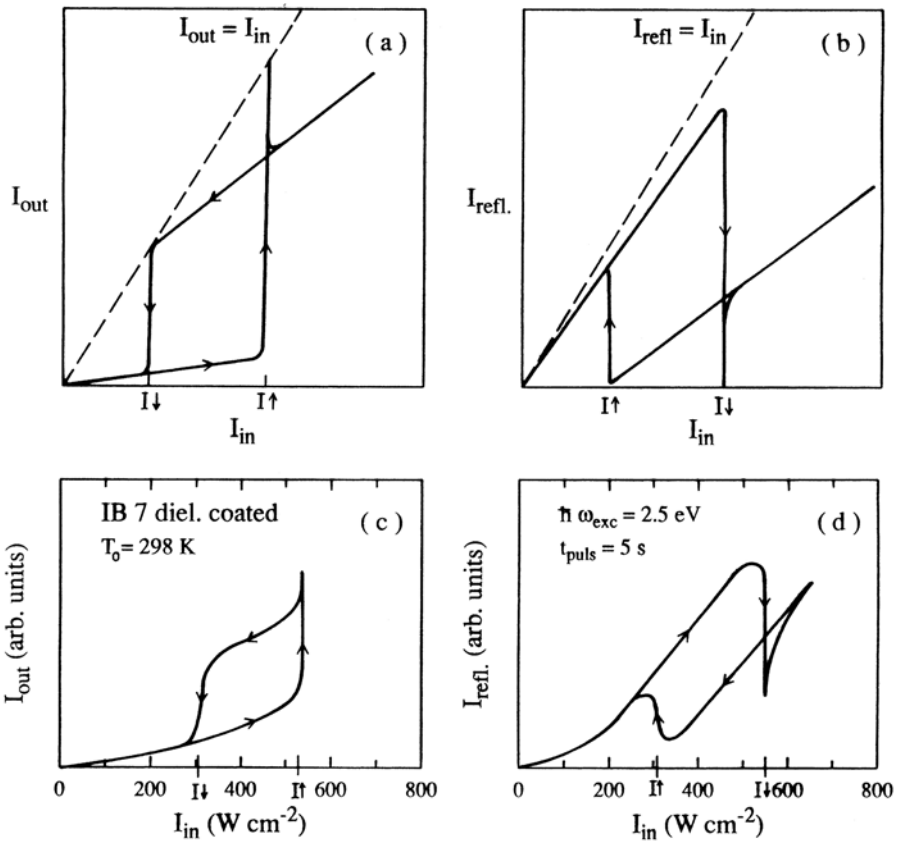


Fig. 24.3. The bistable hysteresis loops resulting from a dispersive optical nonlinearity in a Fabry–Perot resonator in transmission and reflection; idealized (**a**, **b**) and experimentally observed (**c**, **d**) [92G1]

Many other examples are found e.g. in [80O1, 84O1, 86O1, 88B2, 88O1, 04B1] and the references given therein.

24.1.3 Optical Bistability Due to Bleaching

A hysteresis loop similar to those of Figs. 24.1a, 24.2b, and 24.4c can be created by bleaching of absorption.

We start with a FP resonator with an absorbing medium and use a wavelength for which a transmission maximum would occur if the medium in the resonator were transparent.

This means

$$\sin^2 \delta = 0 \quad \text{but} \quad T_{\mathrm{FP}} \approx (1-R)^2 \exp(-\alpha L) \ll 1\,. \tag{24.5}$$

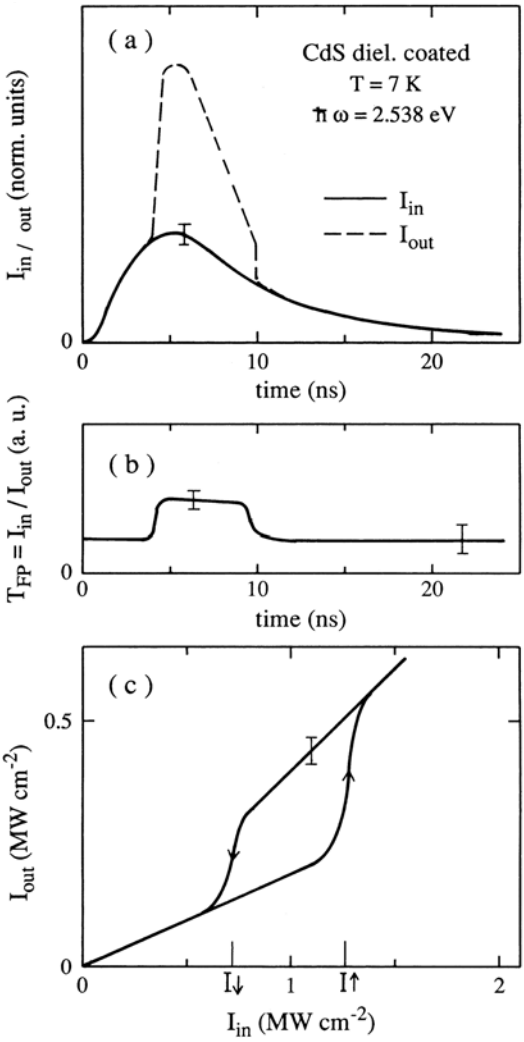


Fig. 24.4. The dispersive optical bistability of a CdS platelet with dielectric coatings observed in the temporal evolution of the incident and transmitted pulses (**a**) and in the trans-mision (**b**). Part (**c**) shows the resulting hysteresis loop [86W1]

The absorption of the medium with $\alpha L \gg 1$ destroys the action of the FP. If the medium starts to bleach with increasing I_{in}, however, the constructive interference in the FP comes into play and increases I_{intra} even further, causing an even stronger bleaching, and so on.

I_{inlra} can again become larger than I_0 and T_{FP} reaches unity.

The construction of the bistable loop is similar to that described in Sect. 24.1.2 above with the main difference that the nonlinearity now oc-

curs in the $\alpha(I_{\text{intra}})$ instead of the $\delta(I_{\text{intra}})$. In Fig. 24.5 we give an experimental example. The nonlinearity is the bleaching of the absorption edge in CdS at room temperature due to the formation of an electron–hole plasma (Fig. 21.11c). Dielectric mirrors have been evaporated directly onto the plane parallel surfaces of the platelet type crystal.

Since bleaching is often connected with dispersive changes due to Kramers–Kronig relations, hysteresis loops like that in Fig. 24.5 are often the result of a cooperative effect of bleaching and dispersive nonlinearities.

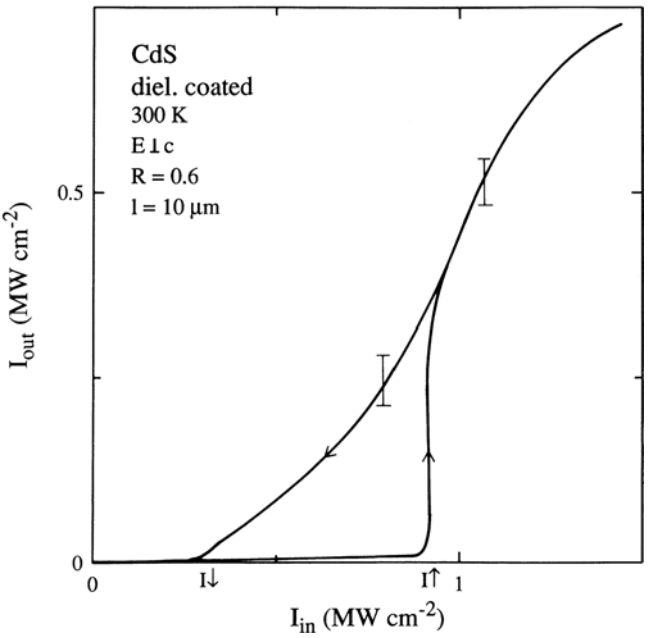


Fig. 24.5. Bistable hysteresis loop caused by bleaching of absorption [86W1]

24.1.4 Induced Absorptive Bistability

We first explain the process of induced absorptive optical bistability in words. The idea is the following: A sample is used whose absorption increases as a function of the density N of some excited species. These may be electron–hole pairs as in Fig. 21.11a around 2.541 eV or simply an increase of the phonon population, i.e., of the lattice temperature T_{L} as shown in Fig. 20.18. One starts at low excitation at a wavelength where the sample is almost transparent. To give a number, let us assume that 10% of the incident light is absorbed. If we increase I_{in}, then the absorbed power increases, too, and is eventually sufficient to raise the density of the excited species N, which causes

the sample to absorb more. This gives a built-in positive feedback, because an increase of the absorption in turn increases N. If the effect is sufficiently strong the sample can switch at $I\downarrow$. in Fig. 24.1b to a state of low transmission and high absorption. We assume now that 80% of the incident light is absorbed. Because of this high value, we can lower I_{in} to values below $I\downarrow$ and still keep the sample on the low transmission branch down to $I\uparrow$ ($< I\downarrow$) where the sample switches back again. In our example the two intensities would obviously be connected by

$$I\downarrow \cdot 0.1 = I\uparrow \cdot 0.8\,. \tag{24.6}$$

Note the interchange of $I\downarrow$ and $I\uparrow$ between Figs. 24.1a and b and between Figs. 24.2b and 24.6b.

The most important difference between the induced absorptive bistability and those relying on bleaching or on dispersive nonlinearities is the fact that no external FP resonator is necessary to produce the positive feedback since it is "built–in".

To calculate the hysteresis we again have to solve two simultaneous equations, as shown in Fig. 24.6a. One is the transmission as a function of N:

$$T = \exp\{-\alpha(\hbar\omega, N)L\} = \frac{I_{\text{t}}}{I_0} \tag{24.7}$$

represented in Fig. 24.6 by the heavy line.

The other equation states that the increase of N, i.e.,

$$\Delta N = N - N_0 \tag{24.8}$$

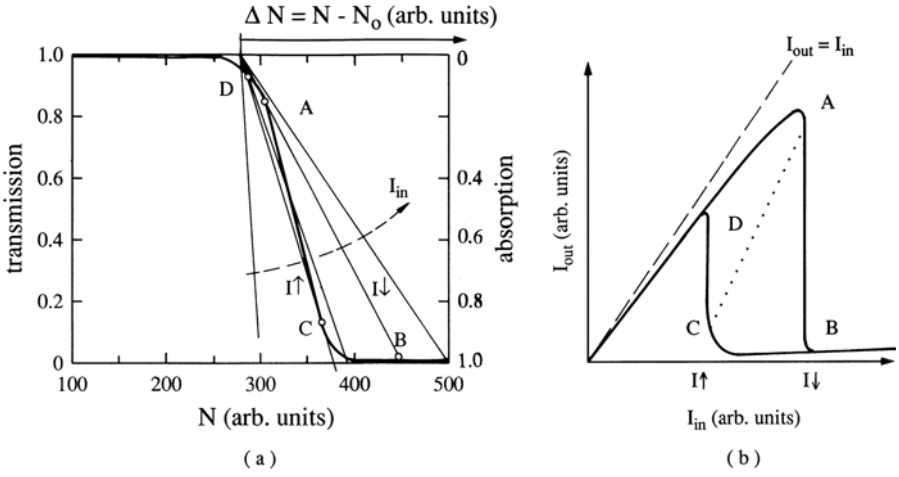


Fig. 24.6. The graphical solution of (24.7)–(24.9) (**a**) and the resulting input–output characteristic (**b**). In this case N is the lattice temperature and the data refer approximately to a CdS platelet with $d < 10\,\mu\text{m}$, $\boldsymbol{E} \perp c$ and $\hbar\omega \approx 2.4\,\text{eV}$ [92K1]

is a function of the absorbed fraction of the incident light. In the simplest case a linear relation can again be assumed

$$\Delta N = A I_{\text{in}} \,, \tag{24.9a}$$

with

$$A = 1 - T = 1 - \exp[-\alpha(\hbar\omega, N)L] \tag{24.9b}$$

or

$$T = 1 - \frac{\Delta N}{I_{\text{in}}} \,. \tag{24.9c}$$

So we once more have a straight line with a negative slope inversely proportional to I_{in}.

The procedure is now similar to that shown in Fig. 24.2a. For low I_{in} we start in Fig. 24.6a with one solution. With increasing I_{in} two new ones appear. At $I\downarrow$ (point A) the initial solution disappears and the system has to jump onto the low transmission branch to point B. There it remains for further increasing I_{in} and also for I_{in} below $I\downarrow$, down to $I\uparrow$ (point C) where the solution on the low transmission branch disappears and the system jumps to point D. Again, the intermediate dotted branch is unstable.

It should be noted that the width of the hysteresis, or indeed the fact if there is optical bistability at all and not just a strongly nonlinear, but monostable input–output characteristic, depends critically on the initial value of the absorption at N_0 and on the steepness of the $T(N)$ curve [84M1]. We shall see later in Sect. 24.1.6 an example where this phenomenon will be used.

In Fig. 24.7 we show two experimental results, again for CdS, based in (a) on the increase of absorption on the low energy flank of the lowest free exciton with increasing exciton density, which is shown schematically in Figs 20.18 and 21.11. In Fig. 24.7b the increase of the absorption due to photothermal nonlinearities is exploited, (see Fig. 20.18). For early examples of induced optical bistability see e.g. [83B1, 84H1, 84M1, 86H1] and references therein.

In Fig. 20.18 we have also seen that photothermal nonlinearities can lead to an increase of both n and of κ. In Fig. 24.8 this fact is used to demonstrate a "butterfly" hysteresis of a CdS sample with dielectric reflecting coatings. The sample is first in a transmission minimum of the FP. With increasing I_0 it switches to the highly transmitting state at $I_1\uparrow$ due to dispersive optical bistability. A further increase of I_0 results at $I_1\downarrow$ in an induced-absorptive switch down, which simultaneously switches off the FP. If we now decrease I_0 we have at $I_2\uparrow$ the switch back of the induced-absorptive bistability and the system is again on the upper branch of the dispersive bistability from which it switches down at $I_2\downarrow$. For more details of this phenomenon see [87W1, 88K1, 90G1, 92G1, 92G2].

There are several further topics connected with (photothermal induced-absorptive) optical bistability which we mention here only briefly, referring the reader for more details to the literature.

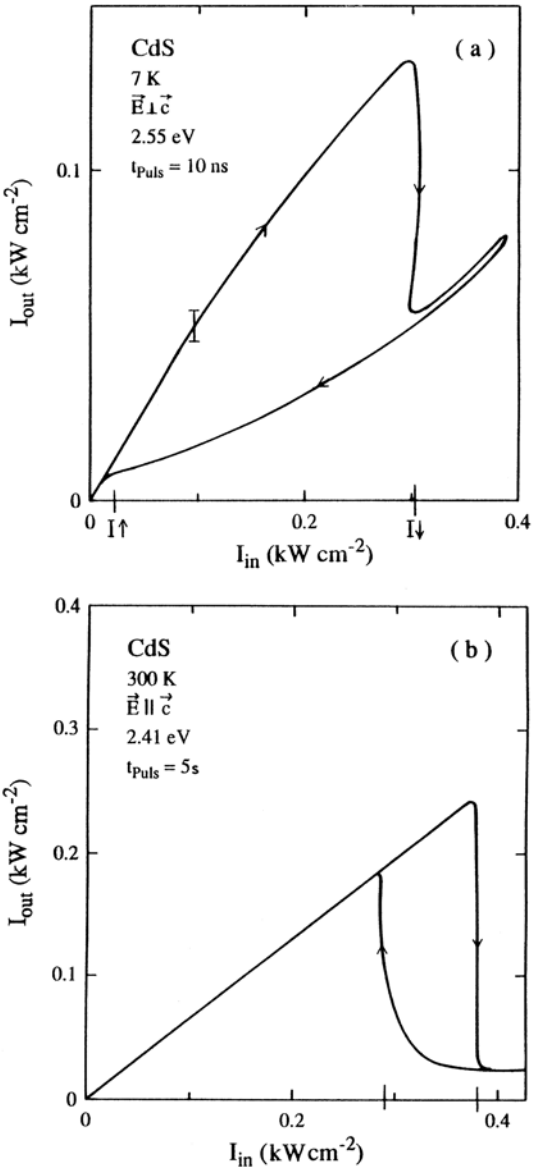


Fig. 24.7. The induced absorptive optical bistability demonstrated with a photoelectronic (**a**) and a photothermal nonlinearity (**b**) [87K1] and [87W1], respectively

Lateral structure formation occurs, e.g., when a spatially wide holding beam with an intensity in the bistable regime is switched in a smaller spot. The resulting "switching waves" have been treated in [90G1, 92G1]. The use of photothermal optical bistability as a temperature sensor is investigated

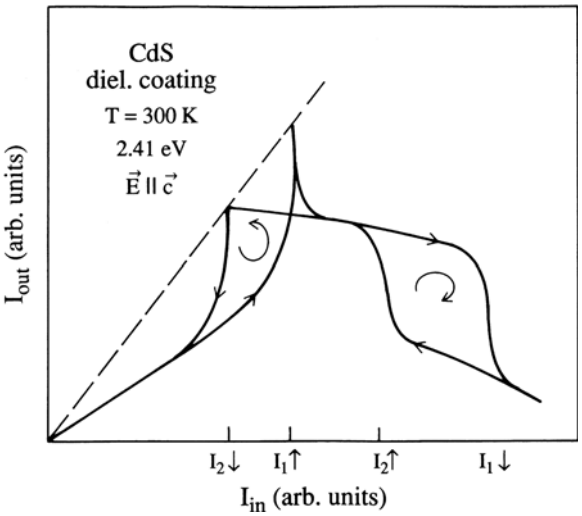


Fig. 24.8. "Butterfly" type hysteresis loops in CdS [87W1]

in [91G1, 92G1] and the relation between the dynamic width of the hysteresis loop and the duration of the incident pulse is analyzed in [91G1] and references therein.

24.1.5 Electro-Optic Bistability

As a further example of optical bistability, we present an electro-optic device that relies on the quantum-confined Stark effect (QCSE) introduced in the Sect. 16.2. The mechanism is the excitation-induced increase of absorption [88C1, 89M1].

The basic concept is outlined in Fig. 24.9a. A constant voltage U_0 is applied in the blocking direction to a MQW pin structure through a resistor R. If the semiconductor is not or only weakly illuminated, it has an internal resistivity $R_\mathrm{i} \gg R$ so that most of the voltage drops over the diode. Consequently, the exciton resonances are red-shifted due to the QCSE as shown in Fig. 16.15 or in Fig. 24.9b. The photon energy $\hbar\omega$ of the incident light is chosen so that it corresponds to the spectral position of the free hh $n_z = 1$ exciton without electric field. This means that the absorption is weaker with an electric field over the MQW than without. This effect is now exploited. One starts at low light input I_in with an applied voltage and relatively high transmission (Fig. 24.9b,c). With increasing I_in free carriers are created in the MQW, i.e., in the intrinsic region of the pin diode, through thermal ionization of the excitons produced by the absorbed fraction of I_in. These carriers give rise to a photocurrent through the system. With increasing I_in, and thus increasing photocurrent, a greater voltage drops over R and the electric field in the MQW decreases. This reduces the QCSE and the exciton resonance

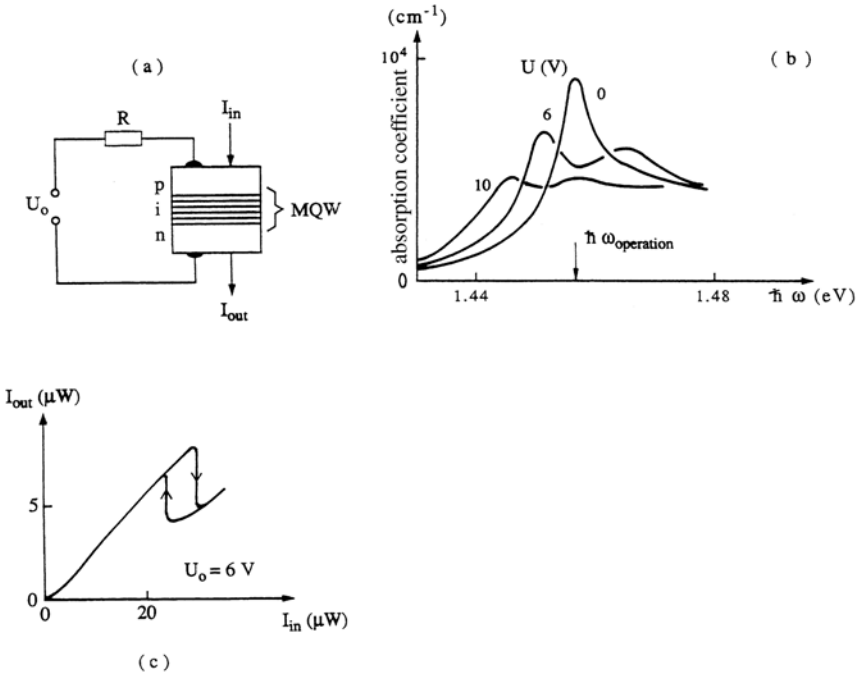


Fig. 24.9. The basic concept of a self-electro-optic effect device (SEED) (**a**), the QCSE in an AlGaAs MQW pin diode (**b**), its input–output characteristics (**c**) [89M1]

shifts towards $\hbar\omega$, i.e., the absorption at this photon energy increases and consequently the photocurrent, too. This is the positive feedback which results in optical bistability if it is sufficiently strong. An example is given in Fig. 24.9c.

These devices have been called self-electro-optic effect devices (SEED) and were considered to be the most promising candidates for application in optical data handling (see below). The devices used in this field usually have a more complex structure integrating the diode with field effect transistors to enhance the width and height of the hysteresis loop and to lower the necessary optical input power. See e.g. [89M1, 89O1].

Finally it should be mentioned that photo–thermal SEEDs have been developed on the basis of Si, CdS, and other semiconductors [87F2, 88B1, 88E1, 88W1, 90K1, 93L1, 93Z1, 93Z2, 95G1, 95K1]. One again uses the induced absorptive optical bistability shown in Fig. 24.6b and evaporates some transparent (ohmic) contacts onto the sample. An applied voltage helps, via the photocurrent, to increase the sample temperature. This means that the battery delivers a part of the energy necessary to switch the sample. The incident light intensity can be correspondingly reduced. Furthermore, the device can be driven through a bistable loop by varying the incident light intensity at constant applied voltage and vice versa.

Other systems that show strong optical nonlinearities and therefore have been used or proposed for (electro-) optically bistable elements are nipi structures, sometimes with an incorporated multiple-quantum-well structure (see, e.g., Sects. 8.10 and 15.2 and [86D1]) and piezo superlattices [93C1, 94L1].

24.1.6 Nonlinear Dynamics

A field that in recent years has developed partly in parallel to, and partly independent of optics is that of nonlinear dynamics and synergetics. In this field one considers, e.g., phase transitions in driven systems, the formation of temporal and/or spatial patterns in dissipative systems under a constant inflow of energy, and other cooperative effects.

Such investigations are of rather general nature and, in addition to physical systems, may also involve biological and sociological systems. References to this fascinating field are [83H1, 84K1, 88S1, 92N1, 98D1].

Examples in which the concepts of nonlinear dynamics are applied to solids have been collected in a recent school [92N1]. We concentrate here on some selected examples in which the nonlinear optical properties of semiconductors and especially optical bistability play an essential role.

First, we repeat schematically in Fig. 24.10a the well known phase diagram of a real or van der Waals gas in the p–V plane with the coexistence region below T_c and Maxwell's construction showing the familiar aspects of a first order phase transition below T_c.

A first example of a pumped system was already discussed in Chap. 21. The transition from an exciton gas to an electron–hole plasma below T_c is also a first-order phase transition in a driven system with a coexistence region (Fig. 24.10b) similar to that of a van der Waals gas (Fig. 24.10a).

Another example is the optical bistability. The hysteresis is again a typical indication of a first-order phase transition in a driven system; it is shown in Figs. 24.10c and d. The bistable or "coexistence region" depends on some control parameter such as the initial temperature of the sample in the case of Figs. 24.2b or 24.6b. We compare it in Fig. 24.10 with the usual p–V phase diagram of a "real" or van der Waals gas. The usual isothermic way through the coexistence region follows the Maxwell construction and is shown by a thin horizontal line. In principle, however, one could go, on the curve given by the van der Waals equation along the solid lines to the points A and B, resulting in a hysteresis loop similar to Figs. 24.10c and d. These states correspond to a supersaturated vapor or superheated liquid. They can occur if care has been taken to remove all nucleation centers. The dotted line, however, is really unstable; if the system could be prepared on this branch by some trick, the slightest fluctuation would cause it to develop into a stable state.

In the case of optical bistability, which is an open, dissipative, driven system one often has transitions from the extrema and not the "mean-field" solution of the van der Waals gas given by Maxwell's construction and imposed by the law of energy conservation. However, the mean-field solution pertains

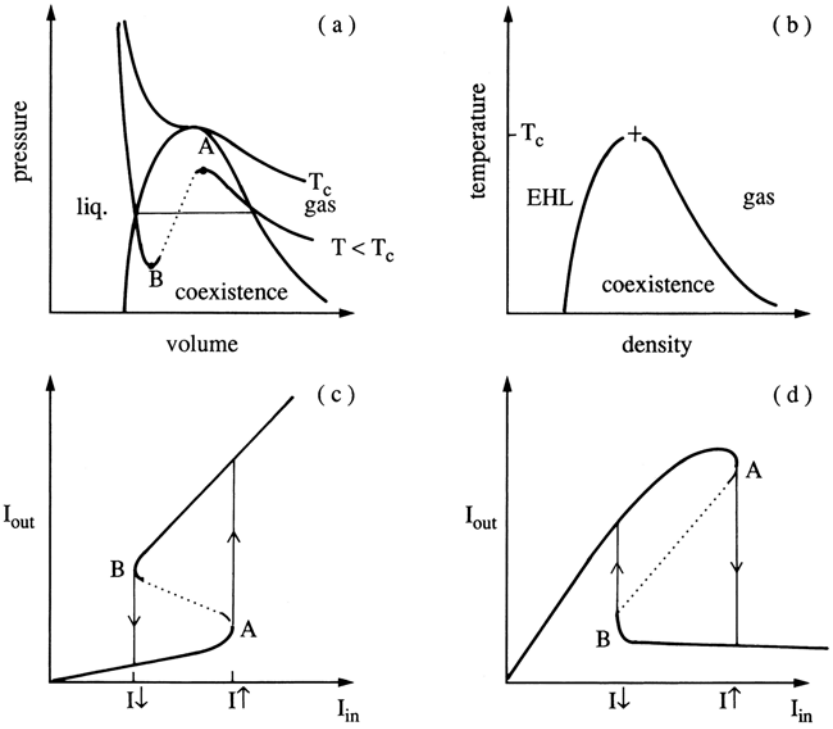


Fig. 24.10. The phase diagram of a van der Waals gas (**a**) the co-existence region of an EHP (**b**) and the input–output characteristics of optically bistable devices (**c**, **d**)

if a sufficient amount of "noise" is added to the optical input beam. For more details of these aspects, see [84K1].

A characteristic of transitions from the extrema (and also of second-order phase transitions) is the so-called critical slowing down. This term denotes the following: If we prepare the system in a state, e.g., on the upper branch in Fig. 24.10d, and increase the incident intensity at $t = 0$ to a value above $I\downarrow$ by an amount δ, it takes some time t_{switch} until the system switches to the lower branch. For small values of $\delta(\ll I\downarrow)$ one obtains a logarithmic singularity of t_{switch} with δ, i.e.,

$$t_{\text{switch}} \propto -\ln \delta / I\downarrow \quad \text{for} \quad \delta / I\downarrow \ll 1\,. \tag{24.10}$$

In Fig. 24.11 we show as an example for the critical slowing the case of induced absorptive photothermal optical bistability of Fig. 24.6b. The experimental data points coincide nicely with the logarithmic singularity. The singularity can be understood in this case at least qualitatively in the following way. As long as the system is on the upper branch, all power which is deposited (i.e., absorbed) from the input beam can be dissipated while the system stays on the upper branch. If we now proceed stepwise to an input power $I\downarrow + \delta$, only

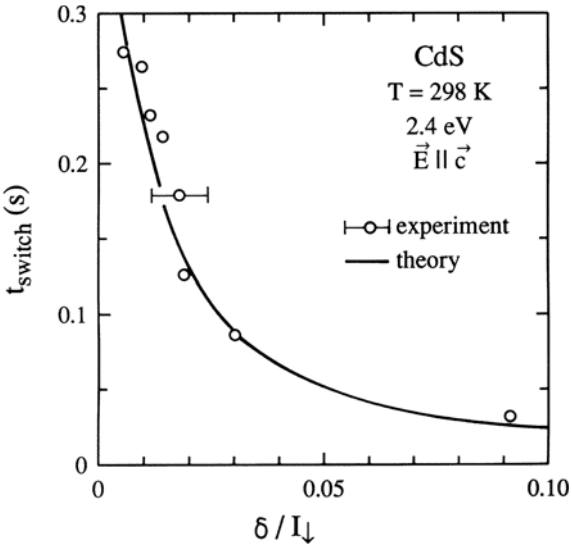


Fig. 24.11. The critical slowing down for the induced absorptive optical bistability of Fig. 24.6b [87W1]

the small fraction δ will be used to drive the system to the new stable state on the lower branch. Since a finite amount of energy is needed to go from one state to the other, i.e., to go to a higher temperature, the transition time t_{switch} increases with decreasing δ. The quantitative analysis then gives the logarithmic singularity of (24.10).

Other features of this driven or non-equilibrium phase transition include the dynamical blowing-up of the hysteresis loop, the measurement of the unstable branch, and the temporal expansion of the switching front in the case of inhomogeneous excitation. We do not want to go into details here but refer the reader to [86H1, 87F1, 87W1, 88C1, 88K1, 88W1, 89M1, 90G1, 90K1, 91G1, 92G1, 92G2, 93Z1, 95G1, 95K1] and references therein.

Instead we now introduce a setup used to study another group of effects that are characteristic of nonlinear dynamics, namely self-oscillations. The setup is a ring resonator containing an induced absorptive element, which can be bistable or monostable. This system is thus complementary to the Ikeda resonator [80I1, 82I1, 86M1], which contains a dispersive nonlinear element in a ring resonator.

The basic idea [86L1, 87G1, 87W1, 87W2] is shown in Fig. 24.12a. We have a ring resonator whose roundtrip time τ_{R} is long compared to the switching time of the nonlinear device. The incident intensity I_0 is chosen such that the part transmitted through the first beam splitter TI_0 is below the switch back intensity $I\uparrow$ of the bistable device and the sample is assumed to be in the highly transmitting state at the beginning. This means that the intensity falling on the sample $I(t)$ is almost completely transmitted. (We neglect

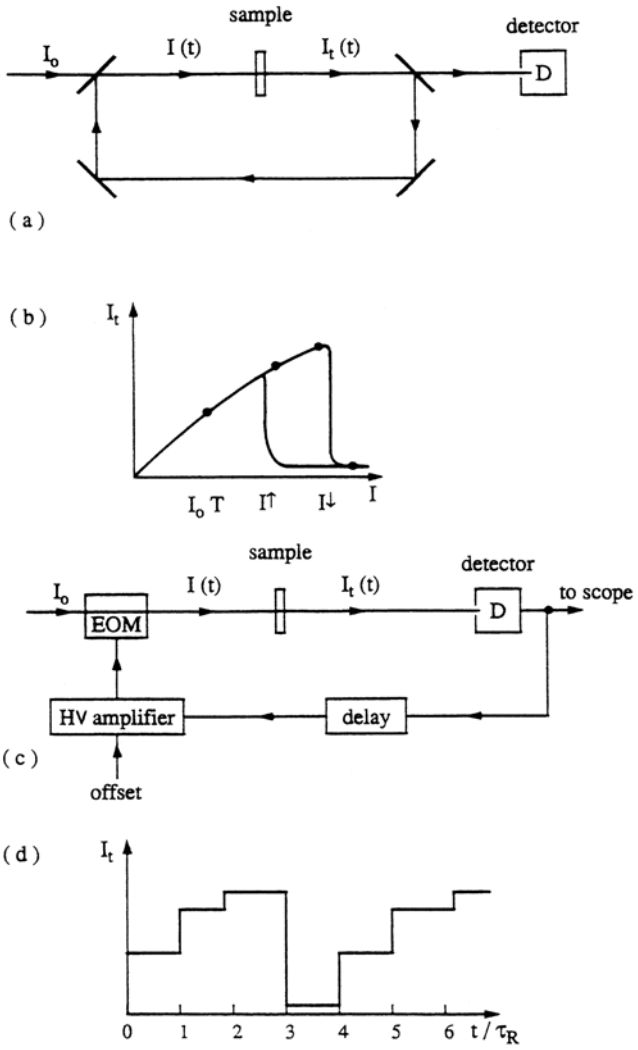


Fig. 24.12. The ring resonator with a bistable element (**a**); the input–output characteristics of this element (**b**); the realisation of (**a**) by a electro-optic hybrid resonator (**c**); and the resulting self-oscillations (**d**) [87W1]

the reflectivity of the sample itself for the moment for reasons of simplicity.) A small fraction of the transmitted intensity I_t, is coupled through the next mirror and detected to see what is going on in the resonator. After one round trip, the incident and the transmitted intensity fall on the sample, and we assume, again for simplicity, that we have to add intensities, i.e., that τ_R is much larger than the coherence length of the laser. The above sequence repeats itself several times, each with a step-like increase of the intensities falling on the

sample and being transmitted through it, I and I_t, respectively. After several round trips I eventually exceeds $I\downarrow$ and the crystal switches to the absorbing state. This means, that almost no more light is transmitted. Consequently only TI_0 falls on the sample after another round trip time and the crystal switches again into the transmitting state and the above described sequence begins a new, i.e., for constant incident intensity we observe self-oscillation (a temporal structure) with a period which is an integer multiple of τ_R.

For photothermal optical nonlinearities the condition $\tau_\mathrm{R} \gg \tau$ would necessitate ring resonators with a delay time of at least 100 ms, equivalent to approximately one trip around the globe. Such devices tend to get unwieldy and therefore the electro-optic hybrid system of Fig. 24.12c has been set up. The incident light passes through an electro-optic modulator (EOM), operated in the linear regime. A constant offset U_0 on the high-voltage amplifier (HVA), which drives the EOM models the constant incident intensity I_0T of Fig. 24.12c. The optical nonlinearity is really the photothermal nonlinearity of a CdS platelet type sample using the green line (514.5 nm) of an Ar^+ laser. The transmitted intensity I_t, is detected by a photodiode, monitored by an oscilloscope and sent through an electronic delay line τ_R. Then the signal returns via the high voltage amplifier on the EOM, which opens according to the applied signal. The choice of the amplification in the delay time allows one to monitor the total reflectivity of the set of four mirrors/beamsplitters in Fig. 24.12a. In this ring resonator, as assumed above, we lose the phase information of the light, and we neglect lateral structures which may appear due to partial switching of the sample.

The system is described by the following set of equations [87W1]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta T_\mathrm{L} = A\left(\Delta T_\mathrm{L}\right) I(t) \frac{1}{c'L}\,, \tag{24.11a}$$

$$A\left(\Delta T_\mathrm{L}\right) = 1 - \exp[-\alpha(\Delta T_\mathrm{L})]\,, \tag{24.11b}$$

$$\begin{aligned} I(t) &= I_0T + R^2 I_\mathrm{t}(t-\tau_\mathrm{R}) \\ &= I_0T + R^2 I(t-\tau_\mathrm{R}) \exp\{-\alpha\left[\Delta T_\mathrm{L}(t-\tau_\mathrm{R})\right]\}\,, \end{aligned} \tag{24.11c}$$

where ΔT_L is the increase of the lattice temperature of the sample in the illuminated spot above the surrounding temperature $T_{0\mathrm{L}}$ at time t

$$\Delta T_\mathrm{L}(t) = T_\mathrm{L}(t) - T_{0\mathrm{L}} \tag{24.11d}$$

A,c' and L are the absorption, the heat capacity, and the sample thickness, respectively; α is the temperature-dependent absorption coefficient; R^2 the total reflectivity of the loop, and I_0T the constant intensity which drives the whole system.

Equation (24.11a) is the rate equation containing the thermal relaxation time τ_0 and, in the generation term, the absorbed fraction A of the intensity

$I(t)$ falling on the sample. A is given in (24.11b) while (24.11c) is the delay (or iteration) equation of the system. The experimental results shown below for a system such as that of Fig. 24.12c can be obtained qualitatively and in most cases quantitatively by solving the set of equations (24.11) using realistic material parameters for $\alpha(T_\mathrm{L})$, τ_0, c' and L.

In Fig. 24.13 we show three examples of the many different scenarios which can be realized with the above resonator. In Fig. 24.13a the CdS platelet is used in the polarization $\boldsymbol{E} \parallel \boldsymbol{c}$. It shows for sample thicknesses below $10\,\mu\mathrm{m}$, $T_{0\mathrm{L}} < 300\,\mathrm{K}$ and the green line of the Ar^+ laser, photothermally induced absorptive bistability. In a certain range of incident intensities self-oscillations occur as described in Fig. 24.12. Depending on R^2 or on I_0T various oscillation modes can be observed, which lock into multiples of τ_R for $\tau_\mathrm{R} \gg \tau_0$. A mode with a total period $n\tau_\mathrm{R}$ and m different maxima per period is designated an n/m mode. With this definition we see in the row (a) of Fig. 24.13 for different incident intensities a 3/1 and a 2/1 mode. For intermediate incident intensity we get a 5/2 mode. This new mode can be predicted from those of the preceding generation by independently adding the numerators and denominators of the mode. This is exactly the way to construct a Farey tree and indeed it can be shown using (24.11) that the modes follow this Farey tree pattern and that the stability ranges of I_0T for the various modes form a devil's staircase [87W1, 87W2, 92G1]. Indeed, this ring resonator with a bistable element seems to be another rather universal system besides the circle map for the investigation of Farey trees [88S1]. Farey trees also occur under periodic modulation of the input intensity I_0T. With other forms of hysteresis loop both ascending and descending steps can be realized and a short perturbation of the incident intensity reveals mode coexistence [87W1, 90G1, 92G1, 92G2, 93Z1].

If the polarization of the light with respect to the crystallographic $\boldsymbol{c}$ axis is rotated by 90° [row (b) of Fig. 24.13], the input–output characteristic is no longer bistable but monostable, due to the dichroism of CdS and the higher initial absorption for $\boldsymbol{E} \perp \boldsymbol{c}$ compared to $\boldsymbol{E} \parallel \boldsymbol{c}$. Under favorable conditions it is then possible to observe another scenario. For a certain value of I_0T we again observe a periodic oscillation locking into a multiple of τ_R (here $2\tau_\mathrm{R}$). For decreasing TI_0 we find period doubling to $4\tau_\mathrm{R}$ showing two different maxima.

Finally, for further decreasing TI_0, we reach a non periodic behavior, i.e., a deterministic chaotic oscillation. This sequence of period doublings is the so-called Feigenbaum scenario, which represents one of the routes leading from periodic to deterministic chaotic behavior. The "strange attractor" on which the system moves in the case of the nonperiodic oscillations was found to have a fractal dimensionality of about 2.6 [87W1]. The important point here is the parabolic maximum of the iteration procedure, i.e., of the input–output characteristic, which is common to the generic system for the Feigenbaum scenario, namely the logistic equation given by

$$x_{n+1} = ax_n(1 - x_n)\,. \tag{24.12}$$

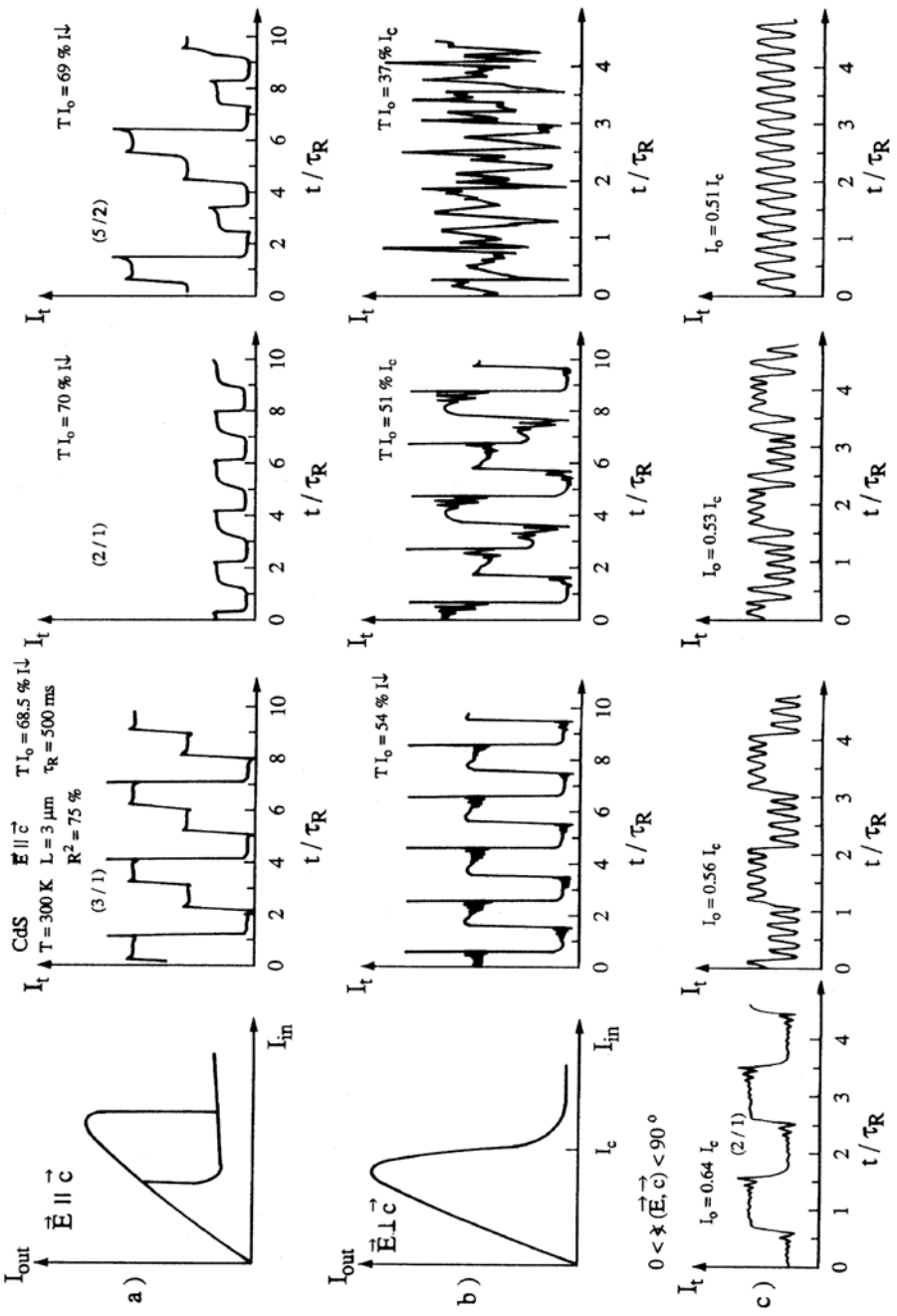


Fig. 24.13. A selection of the rich variety of oscillation modes of the hybrid ring resonator of Fig. 24.12c using bistable and monostable input–output characteristics of the optically nonlinear element [87W1, 90G1, 92G1, 93Z1]

The fast relaxation oscillations which can be seen on the plateaux can prevent the transition to chaotic behavior. Such a situation is shown in part (c) of Fig. 24.13. Here, the angle between $\boldsymbol{E}$ and $\boldsymbol{c}$ was chosen to yield a monostable input–output characteristic with a very steep descending branch. We start again with a regular oscillation, which is locked into a multiple of τ_{R}. With changing input parameters, the amplitude of the fast relaxation oscillation increases until more and more of the spikes result in a modulation over the whole intensity regime. Finally we are left with a regular oscillation, the period of which is determined approximately by the period of the relaxation oscillation τ_0 and not by τ_{R}, however in such a way that an integer multiple of τ_0 equals τ_{R} and with every peak having a shape similar to the initial oscillation governed by τ_{R}.

The first two scenarios shown in Figs. 24.13a,b can be deduced from (24.11) even in the adiabatic limit, i.e., $\tau_0/\tau_{\mathrm{R}} \Rightarrow 0$ whereas the full dynamics are necessary to describe and understand the observations of part (c) of Fig. 24.13.

Some examples of spatio-temporal structure formation in optical bistability in semiconductors are found in [88O2, 92N1] and references therein. They represent a further aspect of nonlinear dynamics which can be investigated with optically bistable systems, but we do not want to go into details here. We, nonetheless, hope that the reader has acquired the feeling that nonlinear semiconductor optics can contribute significantly to the fascinating field of nonlinear dynamics and synergetics.

24.2 Device Ideas, Digital Optical Computing and Why It Failed

As mentioned in the introduction, semiconductors have a very wide field of applications especially in electronics and opto-electronics. It is not the aim of this book to cover this field, but we refer the reader to [81S1, 92E1] and references therein.

What we want to do here is to outline an aspect of application of (electro-) optical bistability in (electro-) optic data processing, also known as (digital) "optical computing" [84O2, 85G1, 87F1, 89O1, 93O1, 98D1].

An optically bistable device with an input-output characteristic such as those shown in Fig. 24.1 can be considered as a binary memory if, in the bistable region, we identify the states of low and high transmission with the logic states "zero" and "one", respectively.

By a suitable choice of the nonlinearity and the excitation conditions, the width of the bistable loop can be made rather narrow. In this case the devices can be used as "AND", "OR" and "NOT" gates (or inverters) as shown schematically in Fig. 24.14. These three functions, together with a memory, are the basic ingredients needed to construct a computer. When this fact became clear in the early nineteen-seventies, there was a very enthusiastic

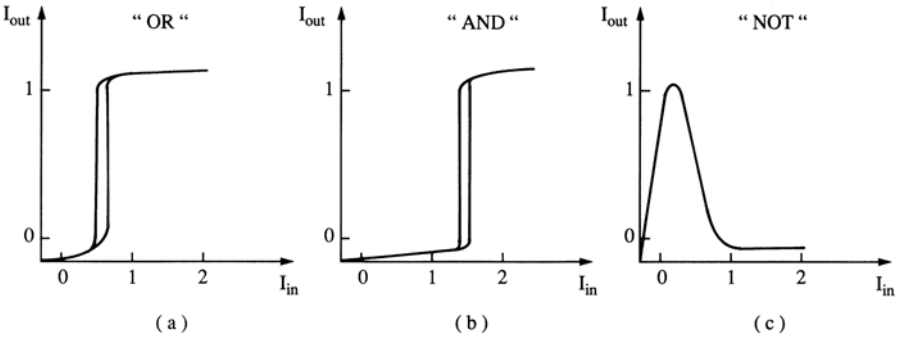


Fig. 24.14. The use of nonlinear optical devices as logic gates

period, where scientists hoped that it would be relatively easy to build an all-optical computer based on the above devices [85G1, 87F1, 89O1, 93O1]. Indeed, such an optical computer would in principle have one big advantage over electronic computers and this is massive parallel data handling. What one wants to do basically in a computer is the following (Fig. 24.15): One has a certain input signal which is processed in a logic unit. The processing depends on the information stored in a large number of memories and results in an output signal.

In electronic computers one faces the following problem: If there are N memory cells, the processing unit has to be connected with $2N$ wires if one

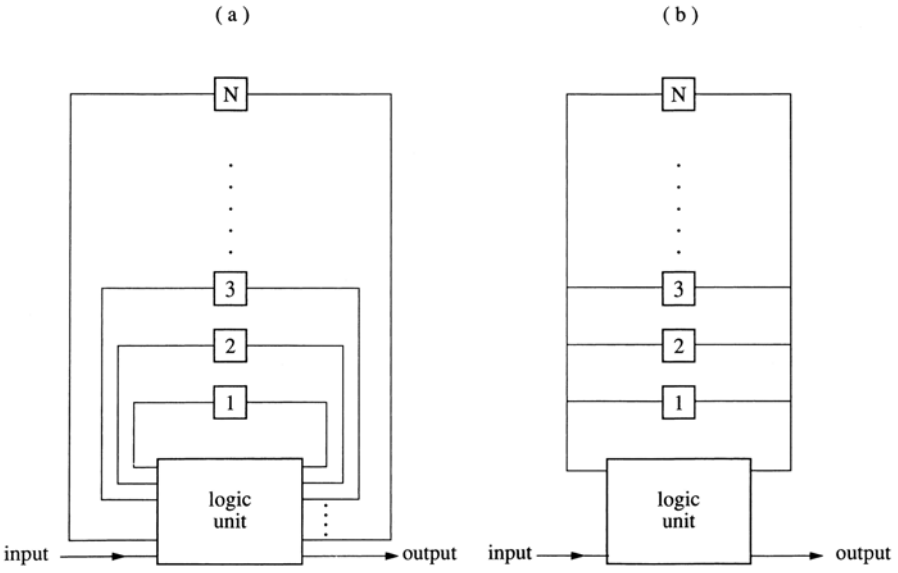


Fig. 24.15. Parallel (**a**) and serial (**b**) data handling

wants to simultaneously know the contents in all memories for the logic decision, since in one wire only one signal can propagate at one (cycle-)time (Fig. 24.15a). If N becomes large ($N \geqslant 10^6$) this concept breaks down because it is impossible to have 2×10^6 pins on one chip. The ingenious way out of this problem was serial data processing. Here one has only two connections between the logic unit and the memory and one uses an address and requests the information in the memories serially, i.e., one after the other (Fig. 24.15b). Unfortunately this takes time and, together with the propagation time in the connections, it is the limiting factor for presently available "calculation speed" in electronic computers. This problem is also known as the "von Neumann" bottleneck.

In optics, on the other hand, one simple lens can be used to handle many data-streams in parallel, since photons in vacuum usually do not interact. So one could imagine a two-dimensional array with let us say $N = 10^3 \times 10^3$ data points (Fig. 24.16). All these data points can be focussed simultaneously by one lens or by a holographic grating (not shown here) onto a logic plane which simultaneously reads data from a memory and performs some logic.

One (or several) further planes can be used to interchange the pixels, e.g., with the help of permanent or reconfigurable holographic gratings, to obtain a "perfect shuffle" function. The output obtained in this way can be processed further or fed back as input to go several times through a loop.

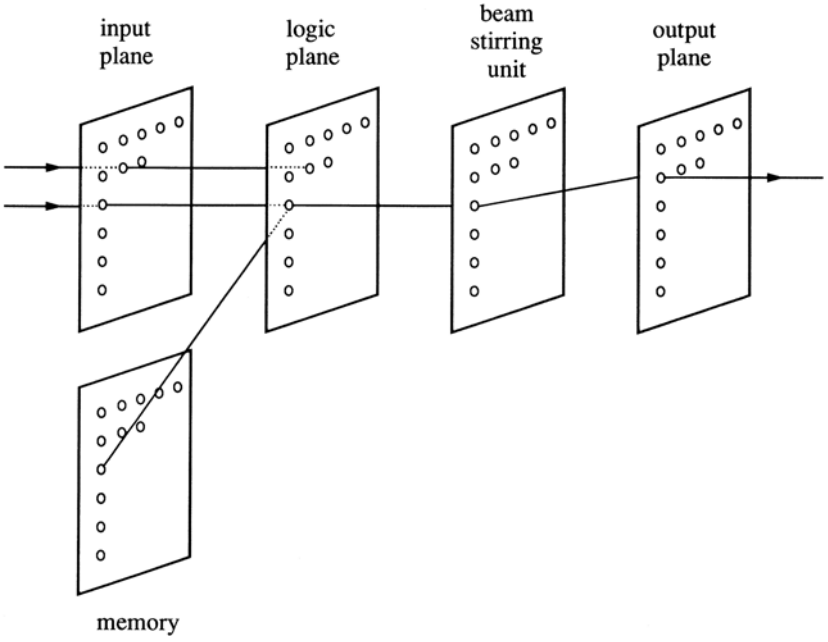


Fig. 24.16. Schematic sequence of two-dimensional optical logic arrays

Another approach to optical data handling involves data transport through (glass) fiber optics. In this case, data have to be brought into an optical form anyhow and one can try to do progressively more operations in the optical regime. Here the demand for parallelism is reduced, but the switching speed has to be high.

At the end of the nineteen-seventies, expectations, especially in digital all-optical computing, were strongly damped for two reasons. Amplification of optical signals with high frequency is very difficult. But this is a crucial requirement for fan-out, i.e., the output of one device must be able to drive several others in the next logic plane. The first way found to solve this problem uses "threshold devices" which, with a constant optical holding beam, are kept just below the switching point and can then be switched with a small input signal. However, it was found that the stability of such a "threshold logic" is not sufficient for practical applications. The other disappointment was that optical switches were found to be inferior to electronic switches concerning speed and power consumption. Optical switches which need low input powers are usually slow (μs to ms) like the photothermal ones. The fast ones, exploiting for example the formation of an electron–hole plasma can have switching times in the ps to ns regime but need powers that are too high for dc operation and sometimes also low temperatures.

Apart from the unsolved fan out problem there was another shortcoming. There was no consistent concept of hard- and software or of devices and computer architecture for (digital) optical computing. Obviously it did not make sense to copy the architecture of serial electronic computers and merely use optical pulses instead of electric ones. A genuine new architecture was missing that makes use of the massive parallelism offered by optical based realistic devices.

At the end there was a "demonstrator", which could optically add $2 + 2$ but used one or several Ar^+ lasers with an electrical power consumption of several kW and some PCs to control the whole system, which filled a complete lab.

Although there is presently still some ongoing research into using semiconductors for optical memory and switches (e.g., [99Z1]) and though beautiful and excellent physics has been made with optically nonlinear or bistable samples or "devices", the author does not expect (digital) optical computing to become a widely used technology in the next few decades.

What emerges, however, and what might spread into common use are electro-optic devices. With the increasing use of optical fiber communication systems, electronic data must be transformed into optical signals and back and we may eventually see an increasing amount of optical data handling at the interconnects.

Another field of optical data handling in a wider sense is making good progress in the form of luminescent semiconductor displays, laser scanners and related devices.

24.3 Spintronics

As already demonstrated in Sect. 16.1 magnetic fields and the spin of carriers, including diluted magnetic semiconductors (DMS), are exciting topics of research. Some further examples, which are chosen rather arbitrarily, are the control of ferromagnetism in materials like $In_{1-y}Mn_yAs$ or $Ga_{1-y}Mn_yAs$ by the carrier concentration in field effect transistor like structures [00A1, 00O1], injection and transport properties of electron spin through a sample or across hetero interfaces in semiconductors [96C1, 98H1, 99F1, 99K1, 00J1, 00O1, 01G1, 01O1], the generation of spin currents without electric charge transport [00B1, 03H1, 03S1], and spin coherence and relaxation phenomena [97A1, 97K1, 97T1, 98K1, 99A1, 99O1, 99O2, 00K1, 01O1, 01S1, 03T1].

Based on such results, the idea of spintronics was born, i.e., analogous to electronics but using the spin of the carriers (generally electrons) in addition to or instead of their electric charge. See, e.g., [02S1, 02S2, 04S1] for reviews or [95N1, 98P1, 99P1, 00B1, 01H1, 01S1] for a few individual results.

In a wider sense there is a real success of spintronics in the application of read heads for magnetic disks, based on the giant magnet resistance (GMR) that occurs if a nonmagnetic metal layer, e.g., Cu is sandwiched between ferromagnetic layers (see, e.g., [03A1, 04S2] and references therein).

On the other hand the author feels that it will be very unlikely that spin precision transistors [03R1] or other electronic devices based on DMS will be able to compete with the highly developed Si technology. The devices are much more complicated, the relevant effects are often observed in semiconductors only at low temperatures (e.g., ferromagnetism in $Ga_{1-y}Mn_yAs$ or ZnO:Co), the spin coherence times are usually short, rarely exceeding a few ns, except for a few exceptions [97K1, 02C1] and the spin injection from metallic ferromagnets into semiconductors is very difficult and requires tunnel barriers since the density of states (or the impedance) in the metal and the semiconductor are too different. Furthermore, and this is an analogy to the failure of digital optical computing mentioned above, there is no consistent concept for the use of spins instead of the electric charge, which would show advantages over the established Si-technology. The author feels that he is not the only scientist sharing this skepticism [03D1].

24.4 Quantum Computing

In a conventional electronic computer there are two states presenting the logic "0" and "1". Their realization is by electric current or voltage "off" and "on", respectively. Recently a new idea emerged, namely to use a coherent superposition of two quantum mechanical states, a so-called quantum- or q-bit

$$a|0\rangle + b|1\rangle \tag{24.13}$$

and to carry out manipulations or logic operations in a time shorter than the decoherence time (see Sect. 23.2) of states $|0\rangle$ and $|1\rangle$.

For an overview see, e.g., [02N1, 02S1]. The realization of the two states $|0\rangle$ and $|1\rangle$ is considered via Josephson junctions [99M1], via single atoms in a cavity or atoms in a coherent state in a trap [95B1, 95D1, 95T1, 01C1] or via electron or nuclear spins or excitons in semiconductors and semiconductor structures of reduced dimensionality like quantum dots [98K2, 98L1, 01B2, 01H1,01S2,02S3,03P1,03Q1] and this is why we address this topic here briefly.

The situation seems to be very similar to the one of optical computing from two decades ago or of spintronics.

There is a huge amount of novel, beautiful and exciting basic research being performed in this field. However, the realization of the promised or at least envisaged applications, here via quantum computing, seems rather unlikely to the author.

There is again no consistent approach for hard- and software or computer architecture. Many of the possible systems for the realization of q-bits require low temperatures and they are intrinsically very sensitive to all types of decoherence processes.

To the best of the author's knowledge, the only case where an advantage of quantum computing over conventional computing has been shown theoretically is the factorization of large numbers. It is unlikely that industry is willing to invest too much money in this niche.

The other field of applications of quantum computing including the presently very fashionable field of entangled states is quantum cryptography. Though this might be a huge market, practical realizations are a long way away, if they come at all.

Another approach to the continuing miniaturization of electronic devices is to use molecular structures including carbon nanotubes [00J2], but this field leads beyond the scope of this book.

24.5 Problems

1. Find a simple relation for the "sufficiently strong" excitation induced increase of the absorption (Sect. 24.1.4) needed to produce optical bistability.
2. Is it possible for longitudinal structures to appear in the switching process of induced absorptive optical bistability in thick samples, under the assumption that the excited species responsible for the increase of absorption have negligible diffusion length?
3. Can induced absorptive optical bistability occur if the increase of the absorption depends not on the density of some excited species, but directly and instantaneously on the amplitude or intensity of the light field? Is your answer also true for dispersive and bleaching optical bistability in a Fabry–Perot resonator?

4. Give the conditions on the finesse of the Fabry–Perot and the bleaching required to get optical bistability from the combination of both.
5. Iterate the logistic equation (24.12) to get the Feigenbaum scenario.
6. Find some additional information about the "circle map" and the Farey tree, e.g. in [88S1].
7. Find out what an "entangled state" is. Compare it to the description of two or more identical Fermions by Slater's determinate, which was developed in the 1930s soon after the introduction of quantum mechanics.

References to Chap. 24

[80I1] K. Ikeda, H. Daido and O. Akimoto, Phys. Rev. Lett. **45**, 709 (1980)
[80O1] Optical Bistability, Ch.M. Bowden, M. Ciftan, H.R. Robl, eds., Plenum, New York (1980)
[81S1] S.M. Sze, Physics of Semiconductor Devices, 2nd, eds., Wiley, New York (1981) and Semiconductor Devices, Physics and Technology 2nd ed. (2002)
[82I1] K. Ikeda, H. Daido and O. Akimoto, Phys. Rev. Lett. **48**, 617 (1982)
[83H1] H. Haken, Synergetics, an Introduction, Springer Ser. Syn., Vol. 1 and subsequent volumes, Springer, Berlin, Heidelberg (1983)
[83B1] K. Bohnert, H. Kalt and C. Klingshirn, Appl. Phys. Lett. **43**, 1088 (1983)
[84H1] F. Henneberger and H. Rossmann, phys. stat. sol. (b) **121**, 685 (1984)
[84K1] S.W. Koch, Dynamics of First-Order Phase Transitions in Equilibrium and Nonequilibrium Systems, Lecture Notes Phys. **207**, Springer, Berlin, Heidelberg (1984)
[84M1] D.A.B. Miller, J. Opt. Soc. Am. **B1**, 857 (1984)
[84O1] Optical Bistability II, P.Ch.M. Bowden, H.M. Gibbs, S.L. McCall, eds., Plenum, New York (1984)
[84O2] Optical Bistability, Dynamical Nonlinearity and Photonic Logic, B.S., Wherrett and S.D. Smith, eds., The Royal Society, Cambridge University Press, Cambridge (1984)
[85G1] H.M. Gibbs, Optical Bistability, Controlling Light with Light, Academic, New York (1985)
[86D1] G.H. Doehler, IEEE J. QE-**22**, 1682 (1986)
[86H1] F. Henneberger, phys. stat. sol. **137**, 371 (1986)
[86L1] M. Lindberg, S.W. Koch and H. Haug, J. Opt. Soc. Am. B **3**, 751 (1986)
[86M1] J.V. Moloney, Phys. Rev. A **33**, 4061 (1986)
[86O1] Optical Bistability III, P. Mandel, N. Peyghambarian, S.D. Smith, eds., Springer Proc. Phys. Vol. 8, Springer, Berlin, Heidelberg (1986)
[86W1] M. Wegener et al., Semicond. Sci. Technol. **1**, 366 (1986)
[87F1] From Optical Bistability Towards Optical Computing, P. Mandel, S.D. Smith, B. Wherett, eds., North-Holland, Amsterdam (1987)
[87F2] F. Forsmann, D. Jäger and W. Niessen, Opt. Commun. **62**, 193 (1987)
[87G1] I. Galbraith and H. Haug, J. Opt. Soc. Am. B **4**, 1116 (1987)
[87K1] C. Klingshirn, in [87F1]
[87W1] M. Wegener and C. Klingshirn, Phys. Rev. A **35**, 1740, 4247 (1987)
[87W2] M. Wegener, C. Klingshirn and G. Müller-Vogt, Z. Phys. B **68**, 519 (1987)
[88B1] S.V. Bogdanov and V.G. Lyssenko, Sov. Tech. Phys. Lett. **14**, 270 (1988)

[88B2] T. Brand, H.J. Eichler and B. Smandek, SPIE Proc. **1017**, 200 (1988)
[88C1] D.S. Chemla, D.A.B. Miller and S. Schmitt-Rink, in [88O3] p 83
[88E1] H.J. Eichler et al., SPIE Proc. **1017**, 90 (1988)
[88K1] S.W. Koch, in [88O3] p 273
[88O1] Optical Bistability IV, W. Firth, N. Peyghambarian, A. Tallet, eds., J. Phys., Paris **49**, C2 suupl. au no. 6 (1988)
[88O2] Optical Switching in Low-Dimensional Solids, H. Haug, L. Banyai, eds., NATO ASI Ser. B **194** (1988)
[88O3] Optical Nonlinearities and Instabilities in Semiconductors, H. Haug, ed., Academic, New York (1988)
[88P1] Photonic Switching, T.K. Gustafson and P.W. Smith, eds., Springer Series in Electronics and Photonics **25**, Springer, Berlin (1988)
[88S1] H.G. Schuster, Deterministic Chaos, 2nd eds., Physik Verlag, Weinheim (1988)
[88W1] A. Witt et al., IEEE J. QE-**24**, 2500 (1988)
[89M1] D.A.B. Miller, in [89O1] pp 55 and 71
[89O1] Optical Computing, B.S. Wherett and F.A.P. Tooley, eds., The Scottish Universities Summer School in Physics **34**, Edinburgh University Press, Edinburgh (1989)
[90G1] J. Grohs et al., Opt. Commun. **78**, 77 (1990)
[90K1] V. Kazukauskas et al., Z. Phys. B **79**, 149 (1990)
[91G1] J. Grohs, H. Issler and C. Klingshirn, Opt. Commun. **86**, 183 (1991)
[92E1] K.J. Ebeling, Integrierte Optoelektronik, 2nd eds., Springer, Berlin, Heidelberg (1992)
[92G1] J. Grohs et al., Int'l J. Bifurcation Chaos **2**, 861 (1992)
[92G2] J. Grohs et al., SPIE Proc. **1807**, 192 (1992)
[92K1] C. Klingshirn, J. Grohs and M. Wegener, in [92N1]
[92N1] Nonlinear Dynamics in Solids, H. Thomas, ed., Springer, Berlin (1992)
[93C1] J. Cibert et al., Phys. Scr. T **49**, B487 (1993)
[93L1] N. Linder et al., Appl. Phys. Lett. **62**, 1916 (1993)
[93O1] Optical Information Technology, S.D. Smith and R.F. Neale, eds., Esprit Basic Research Series, Springer, Berlin (1993)
[93Z1] U. Zimmermann et al., in [93O1]
[93Z2] U. Zimmermann et al., Semicond. Science and Tech. **8**, 1399 (1993)
[94L1] W. Langbein et al., J. Crystal Growth **138**, 191 (1994)
[95B1] A. Barenco et al., Phys. Rev. Lett. **74**, 4083 (1995)
[95D1] D.P. Di Vincenco, Science **270**, 255 (1995)
[95G1] J. Grohs et al., Z. Phys. B **98**, 133 (1995)
[95K1] V. Kazukauskas et al., Optics Comm. **122**, 83 (1995) and phys. stat. sol. (b) **187**, 241 (1995)
[95N1] Y. Nishikawa et al., Appl. Phys. Lett. **66**, 839 (1995)
[95T1] Q.A. Turchette et al., Phys. Rev. Lett. **75**, 4710 (1995)
[96C1] S.A. Croober et al., Phys. Rev. Lett. **77**, 2814 (1996)
[97A1] T. Amand, Phys. Rev. Lett. **78**, 1355 (1997)
[97K1] J.M. Kikkawa et al., Science **277**, 1284 (1997)
[97T1] A. Tackeuchi, O. Wada and Y. Nishikawa, Appl. Phys. Lett. **70**, 1131 (1997)
[98D1] C. Denz, Optical Neuronal Networks, Vieweg, Braunschweig (1998)
[98H1] D. Hägele et al., Appl. Phys. Lett. **73**, 1580 (1998)

[98K1] J.M. Kikkawa and D.D. Awschalom, Phys. Rev. Lett. **80**, 4313 (1998)
[98K2] B.E. Kane, Nature **393**, p. 133, May (1998)
[98L1] D. Loss and D.P. Di Vincenco, Phys. Rev. A **57**, 120 (1998)
[98P1] G.A. Prinz, Science **282**, 1660 (1998)
[99A1] D.D. Awschalon and J.M. Kikkawa, Physics Today, p 33, June (1999)
[99F1] R. Fiederling et al., Nature **402**, 787 (1999)
[99K1] J.M. Kikawa and D.D. Awschalom, Nature **397**, 139 (1999)
[99M1] Y. Makhlin, G. Schön and A. Shnirman, Nature **398**, 305 (1999)
[99O1] M. Oestreich et al., Appl. Phys. Lett. **74**, 1251 (1999)
[99O2] Y. Ohno et al., Phys. Rev. Lett. **83**, 4196 (1999) and Nature, **402**, 791 (1999)
[99P1] S.S.P. Parkin et al., J. Appl. Phys. **85**, 5828 (1999)
[99Z1] S. Zimmermann et al., Science **283**, 1292 (1999)
[00A1] H. Akinaga et al., Appl. Phys. Lett. **76**, 97 (2000)
[00B1] R.D.R. Bhat and J.E. Sipe, Phys. Rev. Lett. **85**, 5432 (2000)
[00J1] B.T. Jonker et al., Phys. Rev. B **62**, 8180 (2000)
[00J2] C. Joachim, J.K. Gimzewski and A. Aviram, Nature **408**, 541 (2000)
[00K1] H. Kalt et al., J. Crystal Growth **214/215**, 630 (2003) and phys. stat. sol. b **221**, 477 (2000)
[00O1] H. Ohno et al., Nature **408**, 944 (2000)
[01B1] B. Beschoten et al., Phys. Rev. B **63**, 121202 (2001)
[01B2] M. Bayer et al., Science **219**, 451 (2001)
[01C1] B.E. Cole et al., Nature **410**, 60, (2001)
[01G1] M. Ghali et al., Solid State Commun. **119**, 371 (2001)
[01H1] D. Hägele et al., Solid State Commun. **120**, 73 (2001)
[01O1] M. Oestreich et al., Festkörperpbrobleme / Adv. Solid State Physics **41**, 173 (2001)
[01S1] J.S. Sandhu et al., Phys. Rev. Lett. **86**, 2150 (2001)
[01S2] G. Salis et al., Nature **414**, 619, (2001)
[02C1] S. Cortez et al., Phys. Rev. Lett. **89**, 207401 (2002)
[02N1] M.A. Nielsen, Scientific American, p 49, Nov. (2002)
[02S1] Semiconductor Spintronics and Quantum Computing, D.D. Awschalom, D. Loss and N. Samarath, eds., Springer, Berlin (2002)
[02S2] Semiconductor Spintronics, H. Ohno, ed., special issue, Semicod. Science and Technol. **17**, 275 pp (2002)
[02S3] R.M. Stevensen et al., Phys Rev. B **66**, 081302 (R) (2002)
[03A1] M. Albrecht, J.-U. Thiele and A. Moser, Phys. Journal, p 25, Okt. (2003)
[03D1] M.I. Dyakonov, Optika i Spektroskopiya **95**, 279 (2003) and Proc. Future Trends in Microelectronics, Workshop, Corsica, (2003), in press
[03H1] J. Hübner et al., Phys. Rev. Lett. **90**, 216601 (2003) and phys. stat. solidi (a) **195**, 3 (2003)
[03P1] E. Pazy et al., phys. stat. sol. (b) **238**, 411 (2003)
[03Q1] Quantum Computing and Quantum Bits in Mesoscopic Systems, A. Leggett, B. Ruggiero and P. Silvestrini eds., Kluwer, Dordrecht (2003)
[03R1] E.I. Rashba and Al.L. Efros, Appl. Phys. Lett. **83**, 5295 (2003)
[03S1] M.J. Stevens et al., Phys. Rev. Lett. **90**, 136603 (2003)
[03T1] E. Tsitsishivili, R.V. Baltz and H. Kalt, Phys Rev. B **67**, 205330 (2003)
[04B1] A. Baas et al., Phys. Rev. A**69**, 023809 (2004)
[04S1] Spin Electronics, D.D. Awschalom et al., eds., Kluwer, Dordrecht (2004)
[04S2] S.A. Solin, Sci. American, p 45, July (2004)

25

Experimental Methods

Warning: Many of the experimental techniques and set-ups generally used in semiconductor optics, some of which being described below, can involve some risks. For example lasers beams may damage the eyes, high voltages used in many lasers are hazardous, gases and liquids used as laser materials may be poisonous etc. Before you build any set-up or start any experiment read the relevant safety instructions and regulations and stick to them!

This chapter is based both on the contribution of the author to [89L1, 01L1] and on the preceding edition of this book. Further references may be found, e.g., in [82L1,86L1,89S2,92E1,93O1,93H2] or in [86L1,89L1,90K1,91D1,92E1, 92S1, 96O1, 96Y1, 97E1, 98D1, 01H1, 02S1, 03T1] of Chap. 1.

The purpose of this chapter is to review experimental techniques, which have been or can be used for the optical spectroscopy of semiconductors.

In agreement with the concept of this book outlined in the introduction, it is not the purpose of this chapter to present all possibilities of optical spectroscopy, which have ever been used. This would fill a whole volume by its own. Instead we concentrate on typical, widely used techniques.

It is also not the purpose to cite the first, the most recent or the "best" application of a certain technique of spectroscopy, but rather one or a few typical examples, which are by no means exhaustive. In this sense the selection of examples is and must be to some extend arbitrary.

We divide the topic in five main sections namely

25.1 Linear optical spectroscopy
25.2 Nonlinear optical spectroscopy
25.3 Time-resolved spectroscopy
25.4 Spatially resolved spectroscopy
25.5 Spectroscopy under the influence of external fields

and give not only information on the equipment, which has to be used and the experimental techniques but also some information, which properties or material parameters can be obtained with the respective technique.

An important aspect is that many of the techniques can be applied for various fields or in various combinations. We want to illustrate this statement with a simple example.

The observation of the photoluminescence (PL) spectrum of a sample under low non-resonant, cw excitation is a method belonging generally to the field of linear optical spectroscopy. If the photon energy of the exciting light beam is tuned, one has the method of photoluminescence excitation spectroscopy (PLE), which may give some information on the absorption spectrum or on excited states of the luminescing complex (see below). If the excitation power is increased, one may reach the intermediate or even the high-density regimes where e.g. biexcitons and trions or even an electron–hole plasma are created, see Chaps. 19 to 21. Since these phenomena are usually connected with strong optical nonlinearities, the PLE has been shifted or applied to nonlinear spectroscopy. If, on the other hand, the luminescence is excited with a short pulse, it is possible to monitor the luminescence rise and decay times leading us to Sect. 25.3 and to Chap. 23. If the short exciting pulses are strong enough, and tunable, the technique of PLE, which originates from linear spectroscopy, ends up in a combination of nonlinear and time resolved spectroscopy. Even spatial resolution techniques or external fields can be combined with PLE.

It is obvious that we cannot list all these combinations and/or modifications, but restrict here to the basic concepts and leave it to the fantasy and the experimental skill of the reader to use or even to invent new cases of such cross-linked techniques.

At the end of this introduction to Chap. 25 we want to give again some references to further articles or books in which techniques of optical spectroscopy have been presented already. Optics in general is outlined e.g. in [98L1, 99H1] and various aspects on experimental techniques used for semiconductors can be found e.g. in [81K2, 82L1, 85H1, 86L1, 89S2, 89L1, 91D1, 92E1, 93O1, 96Y1, 97K1, 98K1, 98K3, 01A1].

25.1 Linear Optical Spectroscopy

The most widely performed experiments in linear spectroscopy are transmission-, reflection- and luminescence spectroscopy. Other techniques include, e.g., photoluminescence excitation spectroscopy, ellipsometry, modulation spectroscopy, measurements of the luminescence yield or Raman spectroscopy.

Linear spectroscopy means in general that the intensities of all incident light-beams are so low that the optical properties of the sample under investigation are not changed by the illumination or in other words that the electric and magnetic susceptibilities do not depend on the field strength of the incident light beams (see Chaps. 2 and 19).

The range or upper limit of the incident energy flux densities of the light (often called light intensity) for the regime of linear spectroscopy cannot be

given in general, since they depend strongly on the sample under investigation, on the lifetime of the excited species and on the experimental conditions like the spectral position of the incident light beams relative to the fundamental absorption edge of the sample, on the sample temperature, on the duration of the excitation pulse etc. There is, however, a rather simple test if one is in the regime of linear optics: One measures the spectra of reflection or transmission for a given incident intensity and then again for an intensity which is one or at least half an order of magnitude different. If the spectra do not change, one is in the linear regime. The same holds, if the emission spectra grow with the excitation intensity but do not change their spectral shape, or in other words, if normalized spectra coincide [81K2].

25.1.1 Equipment for Linear Spectroscopy

The standard set-up for linear optical spectroscopy consists of a light source, the sample under investigation, frequently placed in a cryostat, a monochromator or spectrometer to disperse the light, a detection unit and a PC to handle and evaluate the data. The light sources for reflection, transmission and luminescence spectroscopy can be incandescent lamps for the visible including the near IR and UV, glow-bars for the IR, low or high pressure gas discharge lamps, and cw or pulsed lasers. Without trying to be complete we mention gas lasers like Ar^+, He-Ne for the visible, CO and CO_2 lasers for the IR, Ar^+, He-Cd, N_2 and excimer lasers for the UV, solid state lasers like the Nd-based ones for the near IR including their harmonics in the green and near UV, or tunable titanium sapphire lasers for the red and near IR, dye lasers, which can cover the whole spectral range from the near IR over the visible to the near UV depending on the dye and the pump source. Free electron lasers work in the IR while synchrotron radiation covers the spectral region from the IR over the visible to the soft X ray regime.

Standard diode lasers are based on $Ga_{1-y}Al_yAs$, $(Al_{1-y}Ga_y)_{1-x}In_xP$ or $In_{1-y}Ga_yAs$ quantum structures and cover the range from the red to the near IR (see Chap. 4) while lead salt diodes emit deeper in the IR (see Chap. 7). GaN-based laser diodes for the blue or near UV are now also commercially available [97N1].

Some more information on these various light sources and references to detailed literature can be found e.g. in [98K1]. Luminescence is generally excited optically, especially in fundamental research, but excitation with electron beams [73B1], by injection of carriers in pn junctions (see e.g. the Sect. 22.3 on laser diodes) or by impact ionization in electroluminescent devices [89S1, 96M1] are also common.

Very recently the use of short THz pulses has been introduced for spectroscopy in this range i.e. in the range up to a few meV photon energy. See e.g. [94S1, 97N2, 98C1, 98K7, 00H1, 00L1, 04C1] and references therein.

The samples under investigation are frequently kept in a cryostat. The heads of closed cycle refrigerators reach, dependent on the heat inflow, e.g.

through windows, temperatures down to 5 or 10 K; cold fingers of He-cryostats, on which the samples are attached in a vacuum chamber, have also lower temperatures around 5 K. If the sample is immersed in liquid He under normal pressure one finds 4.2 K but has the problem that the He is boiling. Pumping of the He below the λ-point avoids boiling in the volume and results in temperatures ≤ 2 K.

Though semiconductor devices are generally operated at room temperature or above, it is very useful in fundamental research to start spectroscopic experiments at low temperature, since the spectral width of absorption-, reflection- and luminescence peaks is frequently much smaller at low temperatures and facilitates interpretation. Following the spectral features with increasing temperature allows then to observe thermal broadening or the (dis-)appearance of spectral features with temperature.

For the spectral dispersion of the reflected, transmitted or emitted light grating monochromators or spectrometers are almost exclusively used in the visible, near UV and near IR. The theoretical resolution limit $\Delta\lambda$ of a grating is given at a wavelength λ by

$$\frac{\lambda}{\Delta\lambda} = Nm \tag{25.1}$$

where N is the number of coherently illuminated grooves of the grating and m is the diffraction order. In practice the resolution is rather limited by the width of entrance and exit slits, by aberrations in the optical system of the spectrometer/monochromator or by the spatial resolution of the detector array. Due to these facts the resolution depends on the focal length of the spectrometer and on the number of grooves of the grating per mm.

In the visible a resolution of $\Delta\lambda \leq 0.1$ nm is usually reached with a focal spectrometer length around 0.5 m and a grating with 1200 to 2400 lines per mm.

A grating blazed at a wavelength λ_B can be generally used for the spectral range

$$\frac{2}{3}\lambda_B \leq \lambda \leq 1.5\lambda_B \tag{25.2}$$

in first order and for a corresponding interval around $\lambda_B/2$ in second order. Care has to be taken that the various orders of spectrally broad features may overlap, i.e. the direction, into which the first order is diffracted at wavelength λ, coincides with the second order at $\lambda/2$. This may cause spurious signals in the spectra. A prevention is to use filters which cut short wavelengths and which are available from the near IR over the visible to the near UV.

Though the efficiency of gratings varies rather smoothly with λ over the range given in (25.2) for unpolarized light, it must be kept in mind that the efficiency curves are different for light with the electric vector polarized parallel and perpendicular to the grooves of the grating, respectively [82H1, 98L1, 99H1]. In polarization sensitive experiments it is therefore either necessary to correct the spectra with the spectral efficiency curves for both

polarizations or to have always the some state of polarization at the entrance slit of the spectrometer e.g. by inserting suitable delay plates or other devices like polarization scramblers.

Prism spectrometers/monochromators have the advantages to produce only one spectrum and to have often high optical through put. This makes them a good tool for the preselection of a wavelength interval or a individual line of a lamp. Due to their nonlinear dispersion, they are no longer much in use in spectroscopy itself.

Interference filters may also serve for a preselection of light.

Fabry–Perot resonators with high finesse can be used in contrast as extremely high resolution devices which can easily exceed $\lambda/\Delta\lambda = 10^6$ but which have usually only a very small free spectral range [88H1, 98L1, 99H1].

Fourier spectrometers are ideal for the range from the far IR to the red part of the spectrum [72B1, 99H1].

We conclude this subsection with a few words on the detector side of the set up. Photographic plates and films are generally out dated due to their nonlinear characteristics and the long processing time between exposure and the availability of the data.

Vacuum- and semiconductor photodiodes have typical sensitivities of 0.1 A per Watt of incident light power. Especially vacuum photodiodes can be very fast with rise- and decay-times in the 0.1 ns regime. Avalanche diodes, photomultipliers or channeltrons are much more sensitive due to the built-in multiplication of optically created carriers. Photomultipliers can reach single photon sensitivity. However they are generally used behind the exit slit of monochromators. This means that a spectrum can be recorded only sequentially e.g. by turning the grating. This procedure, often connected with lock-in techniques [83M1] to improve the sensitivity or the signal to noise ratio, makes the recording of a wider spectrum generally a rather time consuming procedure.

If a one- or two-dimensional diode array or a vidicon tube (eventually combined with an image intensifier e.g. a channel plate) are used at the exit of a spectrometer, one has a significant multiplex advantage, i.e. one records simultaneously the spectrum over a certain interval, which is determined by the dispersion of the spectrometer and the geometrical dimension of the sensitive area of the detector array. The sensitivity can reach a few photons especially for cooled detector arrays.

A similar multiplex advantage holds also for a Fourier spectrometer. The difference is that the spectral resolution is preset with a diode array and the signal to noise ratio increases with integration time, while the signal level is essentially given in a Fourier spectrometer but the spectral resolution increases with increasing shift of the mirror in the Michelson interferometer.

The useful spectral range of the detectors is generally limited on the low photon energy side by the work function for photo-cathodes and the band gap (or the ionisation energy of a defect) for semiconductor-based detectors.

On the high photon energy side, the limitation comes often from the window material (glass, silica, sapphire, etc) (see e.g. also [98K1]).

25.1.2 Techniques and Results

In the following we give some examples of techniques of linear spectroscopy and the information, which can be deduced from them. Many further examples have been given already in Chaps. 11 to 17.

Luminescence spectra give usually information on the deepest radiative states of a system including their optical phonon replica or acoustic phonon wings. At low temperatures these are often defect states or localized states resulting from some disorder like spatial fluctuations of the well width or of the composition of alloys. At higher temperatures extended states become also accessible in luminescence. It should be noted that the quantum efficiency i.e. the average number of luminescence photons created per excited electron–hole pair is in most samples considerably below unity [98W1, 03F1]. Luminescence spectra give therefore generally only information on the fate of a minority of excited species as discussed already in Sect. 23.4. Furthermore it should be noted that the thermal distribution function (generally a Boltzmann term) can, if reached during the lifetime of the excited species, enhance at low temperatures the luminescence of low lying but forbidden transitions (e.g. triplet or so-called dark excitons) compared to energetically slightly higher allowed transitions (see e.g. Sect. 13.2 or Chap. 14).

The spectra of the transmitted and reflected light allow to deduce in various degrees of sophistication [82L1] the spectra of absorption andreflection.

The absorption- and reflection structures give information of optically allowed transitions from the occupied ground state in the unexcited sample to excited states [92D1]. The area under the absorption peak is a measure for the oscillator strength of the transition and the width for the inhomogeneous or homogeneous broadening whatever is larger. These transitions can be the various exciton series and their continuum states close to the band gap (for examples see Chaps. 13 to 15), intersubband transitions e.g. in doped quantum structures (see e.g. Sects. 13.3 and 21.5), phonon transitions (e.g. [98G1] or Chap. 11) or plasmons (see e.g. [97G1] or Chap. 12). The quantitative evaluation of reflection spectra needs generally more complex model theories. Examples are found for excitons in [88U1] or in Sects. 13.1 and 15.1, or for plasmons in [97G1] and in Chap. 12.

Usually only transitions with rather high oscillator strength show up in reflection. An experimental trick to enhance the signal is to observe not under normal incidence, which is the geometry most easily to evaluate, but close to Brewster's angle of the barrier or substrate material with light polarized in the plane of incidence. For an example see Sect. 15.1.

Fabry–Perot modes, which appear often in quantum structures due to parallel surfaces or interfaces of substrate, buffer layers etc., can be partly

avoided by anti-reflection coatings [81C1, 99H1], otherwise they may produce spurious structures [95W1].

A way to deduce extremely weak absorption features using a generalized Planck's law together with Kirchhoff's law has been elaborated in [95D1, 95W3]. Also see in this context [57K1, 59M1, 04C1].

Ellipsometry allows in principle to deduce the spectra of the real and imaginary parts of the dielectric function or of the complex index of refraction. While the method is basically simple and described essentially by Fresnel's formulae for a single surface against vacuum, it becomes increasingly complex if (partly unknown) absorbate layers and multi-layers structures are investigated [77A1, 98K1, 04K1, 04S1].

A classical method to enhance weak signals is modulation spectroscopy [69C1, 73S1, 94P1, 97H2, 98D1, 98D2]. The basic idea is, to modulate the spectra slightly by a weak, temporally periodic external perturbation like an electric field, (electro-modulation), a heating of the sample (thermo-modulation), a variation of the detection wavelength (wavelength-modulation) etc. and to detect the reflected or transmitted light with a lock-in amplifier at the frequency of the modulation. In the simplest case, this results in the observation of the derivative of the spectrum and allows to observe weak structures on a broad background or dark current. Observation at twice the modulation frequency gives the second derivative. The increasing availability of digital data handling reduces the use of analog modulation techniques.

A widely used technique is the excitation spectroscopy of photoluminescence (PLE) or of the photo current in pin-diode structures (PCE). The idea is to monitor e.g. the intensity of a certain luminescence feature (e.g. the exciton luminescence of a multiple quantum well) or the photocurrent and to tune the spectrally narrow excitation source ideally keeping the incident photonflux density constant. These PLE or PCE spectra show often peaks at positions of stronger absorption simply because more electron–hole pairs are created at these energies. For examples see Sects. 13.1, 14.1 and 15.1 or [81B1, 81B2, 87B1]. It is remarkable that exciton features show up in PCE spectra at temperatures much too low for thermal ionisation of these excitons. Field ionisation in some Schottky (surface-) barriers possibly plays a role (see, e.g., [80A1, 82S1, 83S1, 02K1, 03K1] and references therein).

In this sense PLE or PCE spectra are related to the absorption spectrum. The technique is very useful in cases where the absorption spectrum is not directly accessible, e.g. because the substrate on which the quantum structure has been grown is opaque at the spectral region of interest (e.g. in AlGaAs structures grown on GaAs) and cannot be removed by selective chemical etching, or because the optical density is too low to be detected (e.g. for a single quantum well or for forbidden transitions).

However it should be noted that the PLE and PCE spectra are related to but not identical with the absorption spectrum [87B1]. Processes which enter are the reflection spectrum in the range in which the exciting photon energy

is tuned, the absorption length of the exciting light, the diffusion length of the excited species, the relaxation into the observed luminescent channel including the branching ratio for relaxation into other states, and finally the efficiency and the escape depth of the luminescence. Similar arguments hold for PCE. In [97K1, 98U1] it has been shown e.g. that the PLE spectra depend strongly on the selected spectral position of the exciton luminescence of ZnSe-based quantum wells. This dependence has been even used to study the intra(sub)band relaxation processes.

Measurements of the absolute luminescence efficiency can be performed in various ways:

A calculation from the signal intensity of the recorded luminescence spectrum for a given excitation power involves a quantitative knowledge of the spectral through-put or the efficiency of all optical components like detector, monochromator / spectrometer, lenses, mirrors, beamsplitters, etc. and of the geometry like the solid angle of acceptance of the spectrometer, the size of the image of the excitation spot on the entrance slit of the spectrometer and the spatial radiation characteristics of the sample. It is obvious that there are many possibilities to obtain erroneous results.

The comparison of a measured luminescence spectrum of the sample under investigation with a standard, which has luminescence efficiency close to unity (usually a laser dye or some luminescence centers in a solid matrix) necessitates identical geometries for both experiments including the excitation depth and the spatial radiation characteristics. These conditions are usually hard to fulfill.

Placing the sample under investigation in an integrating or Ulbricht sphere, fitted into a cryostat, eliminates the radiation characteristics of the sample and allows to obtain even absolute luminescence efficiencies per spectral interval with rather simple calibration techniques. Examples are found in [90L2, 93V1, 97W1, 98W1, 03F1].

In an alternative method, the heat is measured, which is deposited in the sample under investigation. If the absorbed energy is known, it is easy to calculate the luminescence efficiency. Examples for this calorimetric absorption spectroscopy (CAS) are found in [88J1, 91B1, 94F1]. However there are two draw-backs: one is usually limited to very low temperatures, to have sufficiently sensitive bolometers and the other is that the luminescence may be emitted in a spectral range which is not detected e.g. via deep centers resulting in an over estimation of the luminescence yield of the monitored luminescence channels.

For other thermal methods see e.g. [81J1, 82I1].

Another general caveat is that often differential efficiencies are given in literature (partly without explicitely mentioning it), e.g. the slope in the relation between emitted number of photons versus excited electron–hole pairs. If this relation is strictly linear through the origin, the absolute efficiency and the differential efficiency are the same. If the relation is nonlinear, e.g. above the threshold of stimulated emission of a laser diode, the differential

efficiency may be close to unity while the absolute efficiency is still much smaller.

While transition energies from the near IR through the visible to the UV are usually determined directly by luminescence, absorption and reflection spectroscopy, one uses often inelastic light scattering techniques for the spectroscopy of transition energies in the range below a few hundred meV. These techniques are known as Raman scattering for creation or annihilation of optical phonons (see [76R1, 96Y1, 98K1, 01L1] or Chap. 11) including confined ones or backfolded acoustic phonons [99D1], electronic excitations like intersubband transitions (e.g. [92J1]), spin flips (see e.g. Sect. 16.1), or plasmons (Chap. 12). The energies of such excitations and consequently the energy shifts resulting from such excitations in Raman scattering are usually situated in the range from one or a few meV up to a few hundred meV. Brillouin (-Mandelstamm) scattering means the scattering under emission or absorption of acoustic phonons close to the origin of their dispersion relation. It is more restricted to bulk materials Sect. 13.1.4 and the frequency shift is usually in the (sub-) meV range.

Raman scattering requires a stable and spectrally very narrow ($\Delta\hbar\omega <$ 0.1meV) laser irradiation. The background of this laser beam caused e.g. by amplified spontaneous emission (ASE) should be as low as possible. Possibly it makes sense to send the laser over a grating and through a distant ($\approx 1\,\mathrm{m}$) pinhole or diaphragm. Frequently used lasers are the various lines of Ar^+ lasers or tunable dye lasers pumped by Ar^+ lasers. Especially in the second case, the ASE problem has to be considered.

One observes the scattered light frequently in a back scattering geometry or under a right angle. The scattering geometry is often described by the so-called Porto-notation

$$\boldsymbol{k}_i\,[(\boldsymbol{e}_\mathrm{i}, \boldsymbol{e}_\mathrm{s})]\,\boldsymbol{k}_\mathrm{s} \tag{25.3}$$

where the $\boldsymbol{k}$ and $\boldsymbol{e}$ give the directions of wavevectors and polarizations ($\boldsymbol{E}$ fields) and the indices stand for **i**ngoing and **s**cattered light. In the spectral dispersion of the inelastically scattered light, stray light from the incident beam has to be suppressed as far as possible to observe the weak Raman- (or Brillouin-) signals. Therefore double – or even triple spectrometers are used, followed by a sensitive (single photon) detection. For examples see [76R1, 94D1, 95V1, 96Y1, 98K1, 99D1].

The characteristic energies which can be deduced are e.g. those of Raman active phonons, electronic intersubband transitions, plasmons or plasmon-phonon mixed states.

Alternatively, these low laying excitations can be observed in the IR by Fourier spectroscopy if they are IR active. The selection rules are partly complimentary for both techniques. If the sample has a center of inversion, phonons are either Raman active or IR active. Without center of inversion they may be both [76R1, 96Y1].

Since the geometrical and optical thicknesses of many quantum structures are often very small, and since some transitions are allowed only for an electric field polarized perpendicular to the well, like intersubband transitions between neighbouring conduction subbands, one often uses samples with a "wave-guide" geometry as shown in Fig. 25.1, which allows multiple pass and an appropriate orientation of the electric field vector. For examples see [90O1, 00G1].

Coupling to surface or quasi two-dimensional excitations with larger wavevectors than offered by the light can be obtained by producing a grating on top of the sample with a period Λ. The wavevectors of excited states accessible in the sample $\boldsymbol{k}_{\text{sample}}$ are then given by

$$\boldsymbol{k}_{\text{sample}} = \boldsymbol{k}_{\parallel} + n2\pi\Lambda^{-1}, n = 0, \pm1, \pm2, ... \tag{25.4}$$

where $\boldsymbol{k}_{\parallel}$ is the parallel component of the wavevector of the incident light beam. For an example in combination with Raman scattering see [88Z1].

For excitation energies around and below a few meV, which tend to be hardly accessible by Fourier spectroscopy one can use THz spectroscopy with sources described e.g. in [94S1, 97N2, 98C1, 98K1, 00H1, 00L1, 04C1, 04K1], or in Fig. 23.12.

The scattering of light without frequency shift is known as Rayleigh scattering. The observation of Rayleigh scattered light gives e.g. information on (static) disorder in the sample but also on phase coherence times, etc. Examples are found in [94S2]. The spatial structure of the Rayleigh scattered light is known as "speckle". One started recently to extract the information contained therein for quantum wells [99L2]. For recent discussions of Rayleigh scattering see contributions to Sect. 23.2.1.2.2.

The influence of the barrier thickness between quantum wells in MQW samples compared to the wavelength λ has a significant influence on the optical properties. The limiting cases of $\lambda/4$ or $\lambda/2$, so-called Bragg-structures, are presented e.g. in [96M2].

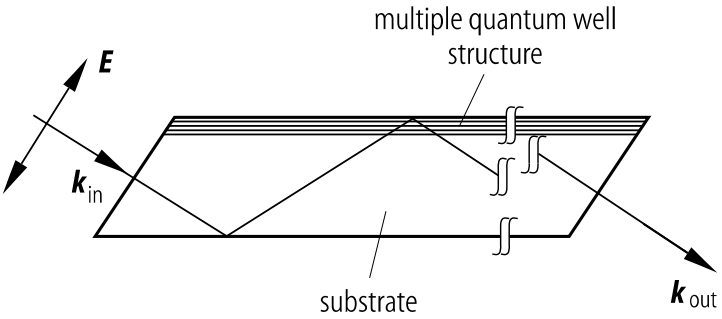


Fig. 25.1. The "waveguide" geometry to allow for multiple pass of the probe light through the quantum structure with a significant component of the electric field vector normal to this quantum well structure

25.2 Nonlinear Optical Spectroscopy

The term "nonlinear optics" includes all phenomena in which the optical properties of a sample are modified by (usually optical) excitation in a reversible way. Reversible means in this context that the sample returns to its previous optical properties with a certain relaxation- or decay time, once the excitation source has been switched off.

These modifications include an excitation-induced increase or decrease of absorption in some spectral regions, the latter effect being also known as bleaching, the appearance of optical amplification or gain, excitation-induced changes of the real part of the refractive index (so-called dispersive optical nonlinearities) or changes of the spectral shape of luminescence spectra with increasing excitation, like an excitation-induced broadening of emission bands, the disappearance or saturation of low excitation features and the appearance of new ones.

We have intoduced the main ideas already in Chap. 19. A good overview of the recent development of the field is found e.g. in Ref. [88N1, 04O1] of Chap. 23

The description of the phenomena in the frame of semiconductors or optical Bloch equations is given in Chap. 27 and the references there.

25.2.1 Equipment for Nonlinear Optics

The excitation source for the observation of optical nonlinearities is usually a sufficiently strong laser (cw or pulsed). There is again no rule from which intensity on optical nonlinearities can be expected for the same reasons that did not allow to give a general value when the regime of linear optics of Sect. 25.1 is left.

To some extend, the more or less powerful pump lasers necessary for the observation of optical nonlinearities can be replaced by (cw or pulsed) electron beams or by a forward biased p(i)n junction. In all cases a high density of electron–hole pairs can be created, which modify the optical properties of the semiconductor.

Optical excitation or pumping has the big advantage that lasers exist for a wide range from the IR, visible and near UV (see (25.1)). They can be continuously working (cw) or be pulsed down to a few fs pulse duration [93K1,98F1,98I1,01A1,03C1,03S1,03S2,03T1,03T2,05F1]. Many of them are tunable over a certain spectral range (e.g. dye-, colour center-, diode-, titanium sapphire- or CO_2 lasers) which allows to excite resonantly some species.

In p(i)n junctions the electrons and holes are usually injected close to the band extrema of the doped regions.

The excitation by electron-beams (typically 10 to 100 keV electrons with current densities up to some $10\,\mathrm{A/cm^2}$) is completely unselective. Usually an energy of about $3E_\mathrm{g}$ is deposited to create one electron–hole pair. This means that a lot of heat is introduced into the samples.

The repetition rate ν_R of pulsed excitation sources should be smaller than the inverse of the lifetime of the excited species T_1 i.e.

$$\nu_R \cdot T_1 \ll 1 \tag{25.5}$$

to avoid accumulation effects.

If quasi stationary conditions shall be reached, the pulse duration T_P should fulfill

$$T_P \geq T_1 \tag{25.6}$$

and the opposite, if dynamical processes shall be studied (see Chap. 23). Under quasi stationary conditions, the density of excited species n_P can be approximated by

$$n_P = G \cdot T_1 \tag{25.7}$$

where G is the generation rate i.e. the number of created species per unit volume (or area) and time. T_1 is the "effective" lifetime. The term "effective" has been introduced, since a lifetime is defined strictly only for an exponential decay i.e.

$$\frac{dn_P}{dt} = -n_P \frac{1}{T_1} \quad \text{or} \quad n_P(t) = n_{P0} \exp(-t/T_1) \tag{25.8}$$

but in certain density intervals it is often useful or convenient to operate with an effective lifetime even for the case of a non-exponential decay. See (23.5).

Before we present several techniques of nonlinear spectroscopy we want to give some information on possible general problems.

We defined the regime of nonlinear optics as the one of reversible changes. The investigation of irreversible changes and damages is a research field of its own. For examples see e.g. [97M1, 99C1]. Macroscopic damages can be often identified by observing the sample before and after (but never during!) illumination through a microscope. Surface and volume damages can be often caused by melting and/or evaporation of the sample by the energy deposited by the excitation source.

If the sample has to be viewed in situ during the experiment it is more convenient and much more safer to view it through a microscope with a small TV camera and to observe the image on a CRT screen. If the camera is damaged by an intense laser pulse it can be replaced for a few hundred US $ or less, your eye cannot.

Another simple test is to measure the linear optical properties e.g. the luminescence spectrum before the experiment of nonlinear optics and after it. If the spectra coincide, the sample was most likely not damaged. If they are different, permanent changes or damages have been introduced.

A frequent problem is heating of the sample by the excitation source, while one is looking for effects caused by high carrier densities. A simple change of

the lattice temperature may cause significant changes of the optical properties (see e.g. [82L1]). A test for such phenomena is to reduce the repetition rate of the pump source, or the duration of the excitation pulse (if possible) and to check if the phenomena remain the same. One can also try to measure the decay time of the phenomena and to check if it coincides with the expected or known lifetime of the excited species or rather with the thermal relaxation time constant τ_{th}, which is given for a homogeneous, isotropic medium in simplest approximation by

$$\tau_{\mathrm{th}} = d^2/\kappa \tag{25.9}$$

[95G1] where d is the diameter of the excitation spot and κ the conductivity for temperature.

One can also try to calculate the upper limit of the rise in lattice temperature using the energy deposited per unit of volume and the temperature dependent specific heat. Diffusion of the excited species, heat conductivity or radiative recombination reduce the rise of the lattice temperature and may be introduced in refined models.

Sometimes the experimental findings allow directly to exclude thermal effects as the main origin of optical nonlinearities. An example is an observed blue-shift of an emission band with increasing excitation in a semiconductor the gap of which shrinks with temperature.

A further point which we want to address here are inhomogeneities of the density of excited species, frequently electron–hole pairs or excitons. Such inhomogeneities can occur, if the thickness of the sample under investigation is larger than the absorption length of the exciting radiation or the diffusion length of the excited species.

This is usually not a problem for samples containing one or a few quantum wells, which are excited in the well but can become a topic to be considered for excitation of thicker multiple quantum well stacks, or for excitation in thicker cap or bufferlayers and subsequent capture of the excited carriers into the wells, or for thick samples containing e.g. quantum dots.

Fast surface recombination can even complicate the issue.

Lateral inhomogeneities over the excited area can result from hot spots or other inhomogeneities of the beam profile of the exciting laser, from microscopic or macroscopic defects of the sample, which act as fast recombination centers like dislocations or precipitates, or from local fluctuations of the potential relief, which funnel the carriers or excitons to the local minima.

A point which deserves some consideration is the variation of the intensity of the incident laser beams.

The simplest and cheapest way is to insert neutral density filters, which are either homogeneously doped glasses or thin metal films (Pt or Ir) deposited on glass or silica. The wavelength range over which they have (almost) constant extinction can be seen in the data sheets of the manufacturers like Schott or Corning. If the filters are placed normal to the beam, they can possibly provide an unwanted feedback into the lasers. Therefore they should be inserted

slightly off normal. This produces however a slight lateral displacement of the beam, which can be of importance in spatially resolved experiments. Even worse is the kink in the optical beam resulting from an (un-)intentional slight wedge shape of the filters.

A variation of the laser beam intensity by variation of its pump power is usually discouraged, since it results often in a variation of the beam profile or diameter.

A further possibility to attenuate a laser beam is to use rotatable polarizers. Polaroid foils are easily damaged by high power (pulsed) lasers. The same holds for polarizers which are glued together with Canda balm. Polarizers with an air-gap are the choice for high power laser beams, with lateral openings for the reflected beams. These reflected beams must be carefully blocked to avoid eye-damage, which can arise from a (collimated) laser beam but even from scattered laser light.

Since a rotation of the polarizer prism may also displace the beam, a combination of fixed polarizers and suitably oriented Pockels- or Kerr-cells is recommended but more expensive. The intensity can be then tuned by the applied voltage without mechanically moving parts.

A special technique to vary the intensity of an incident laser beam is the so-called z-scan technique [93S2, 95K1]. In this case the laser beam is focussed by a lens and the sample is moved through the focal point by varying the relative distance between lens and sample. The method is simple and cheap. However, it has the disadvantage that the light intensity in the vicinity of the focal point is often calculated in the diffraction limit assuming aberration free focussing lenses and an ideal laser beam with plane wave-fronts and a laterally Gaussian profile. Deviations from these conditions occur frequently and modify the intensity profile in the vicinity of the focal point. Furthermore, the beam diameter changes necessarily in these experiments, resulting in a variable influence of carrier diffusion.

25.2.2 Experimental Techniques and Results

The experimental methods can be roughly subdivided in one-beam methods, pump-and-probe beam spectroscopy and four-wave mixing or laser-induced gratings. The second and the third one use two (or more) beams. The difference is essentially that the two beams need not (or even must not) be coherent for pump-probe spectroscopy, while they must be coherent for four-wave mixing and laser-induced gratings. See Chaps. 19 to 24 for results in addition to the references given below.

25.2.2.1 One Beam Methods

Experimentally most simple are one beam methods. One sends one, usually spectrally narrow cw or pulsed laser beam on the sample and is able to observe e.g. the following phenomena:

- second, third and higher harmonic generation. (See Chap. 19.) In thin samples with thickness $d \ll k^{-1}$ where k is the normal component of the wavevector of the incident light, the second harmonic signal growths often quadratically with increasing sample thickness. In thicker samples one obtains high signal amplitudes only under so-called phase- or index matching conditions. This means that the wavevector of the second harmonic is twice that of the fundamental wave ($\boldsymbol{k}$ conservation) or that the refractive indices n at ω and 2ω are equal. Since n tends to increase with ω this condition can be usually fulfilled only in some material with suitable birefringence for a selected orientation [65B1, 84S1, 91M1, 95K1].
- the rectification of light. This effect, which is also known as dc-effect, closely related to second harmonic generation (it is just the difference frequency instead of the sum frequency) and to the linear electro-optic effect [71K1].
- two-photon absorption [31G1]. In this effect two photons of an incident beam are absorbed simultaneously to create an excited state. If one- and two-photon absorption, described by the absorption (or extinction) coefficient α and by β, respectively, are simultaneously present, then the attenuation of a beam propagating over an infinitesimal distance $\mathrm{d}x$ in the sample is given by [71K1, 75K1]:

$$-\mathrm{d}I(x) = \left[\alpha I(x) + \beta I^2(x)\right] \mathrm{d}x \tag{25.10}$$

  integration over x form zero to the sample thickness d leads to

$$I(d) = \left[\frac{1}{I_0}\mathrm{e}^{\alpha d} - (1 - \mathrm{e}^{\alpha d})\frac{\beta}{\alpha}\right]^{-1} \tag{25.11}$$

  where reflection at the surfaces is not yet included.
- two step absorption, which proceeds not via a virtually excited intermediate state, but via a really excited or populated state created by the same or another beam. It may show a similar dependence as (25.10), (25.11) [73B2]. For luminescence assisted two-photon absorption spectroscopy (LATS) see [79S1]
- bleaching of absorption. This effect results in a superlinear increase of the transmitted intensity. Note that optical amplification cannot be reached in a one-beam experiment.
- dispersive nonlinearities. This term includes all changes of the real part of the refractive index by the light field i.e. an increase or a decrease of $n(I)$ caused by the light intensity I e.g. according to

$$n(\omega, I) = n_0(\omega) + n_2(\omega)I \tag{25.12}$$

  Depending on the sign of n_2 this may result in self-focusing ($n_2 > 0$) or self defocussing ($n_2 < 0$) of a beam and in "blooming" of the farfield on a screen.

- two-photon- or hyper-Raman scattering [85H1].
- optical bi- or multistability. For details see e.g. Chap. 24 or [86W1] and references therein.
- Excitation-induced variations of the luminescence spectra. The electronic eigenstates close to the fundamental absorption edge are strongly modified with increasing excitation density in the electronic system going e.g. from excitons at low densities to an electron–hole plasma at the highest densities (see e.g. Chap. 19 to 21 for these effects).

If a spectrally broad laser beam is used, which covers absorbing and transparent parts of the transmission spectrum of the sample one can excite it strongly with one spectral part of the incident beam and simultaneously monitor changes in the transmission- or reflection-spectra [88S1]. This technique leads directly to pump-and-probe measurements.

25.2.2.2 Pump-and-Probe Beam Spectroscopy

As the name says, one uses in pump-and-probe beam spectroscopy two beams. The pump beam is usually a spectrally narrow, intense, pulsed or cw-laser beam (possibly replaced by an electron beam or by carrier injection in a forward biased p-n junction) to excite the (electronic system of the) sample. This pump beam can be tuned to a specific transition or resonance like an exciton. The second beam is usually spectrally broad, tuned to the spectral interval of interest and is used to monitor the changes in transmission or reflection induced by the pump beam. The probe beam has to be so weak that it does not itself introduce any optical nonlinearities.

If the change in transmission is small one speaks also about differential transmission spectroscopy (DTS), especially if the signal is the difference of the transmission spectrum without and with excitation. A related technique is differential photoluminescence excitation spectroscopy (DPLE) [92B1].

There are many geometries for pump-and-probe beam spectroscopy. An example is given in Fig. 25.2a.

Both beams can impinge collinearly on the sample. This makes the achievement of spatial coincidence rather simple. If the sample is not completely opaque for the pump, it is better to have the pump beam incident under a finite angle (Fig. 25.2a) or even from the reverse side of the sample to separate the transmitted pump light from the probe beam.

The spatial coincidence of pump-and-probe beam on the sample can be checked by placing a pinhole at the place of the sample or by observing the excitation spot with a small TV camera, possibly through a microscope (but never with the eye!). The temporal coincidence of pulsed pump and probe beams can be verified by placing a fast photodiode at the position of the crystal for pulses longer than the response-time of the detector and display unit, usually an oscilloscope. For very short pulses, correlation techniques have to be used and fine delay stages for the two beams.

If no spatial or temporal resolution is required, it is recommended to make the diameter of the probe beam and its duration (in case of pulsed lasers) narrower and shorter than those of the pump beam, to monitor the changes only during the pump pulse maximum in a region of homogeneous excitation.

The intensity (power per unit area) or the fluence (energy per unit area) of the pump pulse can be varied as mentioned above up to the damage threshold. The intensity of the probe beam has - as already mentioned - to be kept so low that it does not itself induce any optical nonlinearities. This can be checked by varying the probe-beam intensity. A factor 10 difference to the pump beam is minimum. Possibly one has to subtract from the transmitted or reflected probe beam the luminescence caused by the pump beam and of course the dark current of the detector. If optical amplification of the probe beam shall be investigated, care has to be taken, not to saturate the gain by too intense probe beams [93M1, 98K2].

There are some variants of pump-and-probe beam spectroscopy, which use the light emitted from the excited area of the sample as probe beam. One is luminescence assisted, two-photon spectroscopy (LATS) [79S1], the other the so-called stripe-length method [81K2, 98K2]. See Fig.2.2b. The excitation spot is focussed to a narrow stripe of length l and width w with $w \ll l$. The emission spectrum $I_{\mathrm{lum}}(\hbar\omega, l)$ is detected which propagates along this stripe. In case of optical amplification or gain $g(\hbar\omega)$ in the excited stripe, the emitted intensity does not vary linearly with l but according to

$$\frac{I_{\mathrm{lum}}(l, \hbar\omega)}{I_{\mathrm{lum}}(l + \Delta l, \omega)} = \frac{\exp\{g(\hbar\omega)l\} - 1}{\exp\{g(\hbar\omega)(l + \Delta l)\} - 1} \tag{25.13a}$$

if the length l of the excited stripe is varied from l to $l + \Delta l$ [98K2].

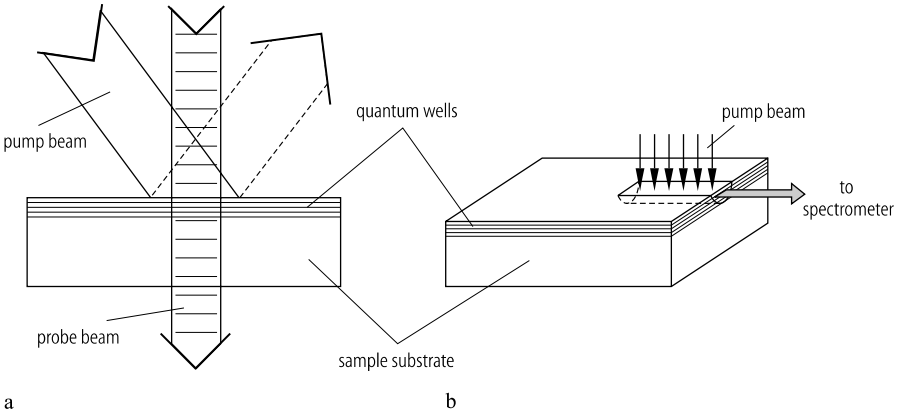


Fig. 25.2. An example for the arrangement of pump-and-probe beam spectroscopy **(a)**, and for the stripe-length method **(b)**

An especially simple situation arises if l is varied between l and $2l$ [78H1, 98K2]

$$g(\hbar\omega) = \frac{1}{l} \ln \left\{ \frac{I_{\mathrm{lum}}(2l, \hbar\omega) - 1}{I_{\mathrm{lum}}(l, \hbar\omega) - 1} \right\} \tag{25.13b}$$

Similar formula hold for absorption. In the case of gain, care has again to be taken that the gain is not saturated by the amplified spontaneous emission. A rough limit is reached for

$$g(\hbar\omega) \cdot l \geq 1 \tag{25.14}$$

Since gain values can exceed $10^3\,\mathrm{cm}^{-1}$ this implies $l < 100\,\mu\mathrm{m}$ and $w \leq 10\,\mu\mathrm{m}$. For processes with even higher gain values pump-and-probe spectroscopy in transmission is recommended. A recent detailed investigation of the stripe-length method can be found in [95B2, 95U1, 98K2].

Phenomena, which can be observed with pump-and-probe spectroscopy are e.g. two-photon or two-step transitions e.g. to biexciton levels [85H1, 97W2], a broadening of the exciton resonances due to collisions with optically excited species (like free carriers or other excitons, the so-called excitation-induced dephasing), a bleaching of the exciton resonances in the transition to an electron–hole plasma, the band gap renormalization ocurring in this process, the gain connected with a degenerate electron–hole plasma, and intersubband transitions in the IR induced by optical pumping across the gap (see e.g. [98S1, 00G1] or examples in Chaps. 20 to 23).

25.2.2.3 Four-Wave Mixing and Laser-Induced Gratings

The basic idea of all four-wave mixing (FWM) and laser-induced grating (LIG) experiments is to have two coherent laserbeams interfere on the sample under investigation. See e.g. [84S1, 85H1, 86E1, 91M1] and references therein. In Fig. 25.3 the wavefronts are shown and the angle θ between the two beams with wavevectors $\boldsymbol{k}_{1,2}$ and (angular) frequencies $\omega_{1,2}$ is defined.

The interference leads to a laterally periodic modulation of field amplitude and intensity. The period Λ of this grating and its vector in reciprocal space $\boldsymbol{G}$ are given by

$$\boldsymbol{G} = \boldsymbol{k}_1 - \boldsymbol{k}_2 \tag{25.15a}$$

$$G = \frac{2\pi}{\Lambda} \tag{25.15b}$$

$$\Lambda = \lambda(2 \sin\theta/2)^{-1} \tag{25.15c}$$

One can consider in the calculation of Λ and $\boldsymbol{G}$ wavevectors, wavelength λ and angle θ of the incident beams in vacuum or of the beams in the sample. The result will be the same.

If the incident light fields produce any optical nonlinearity either via the interference of the coherent polarizations, which they create in the medium, or

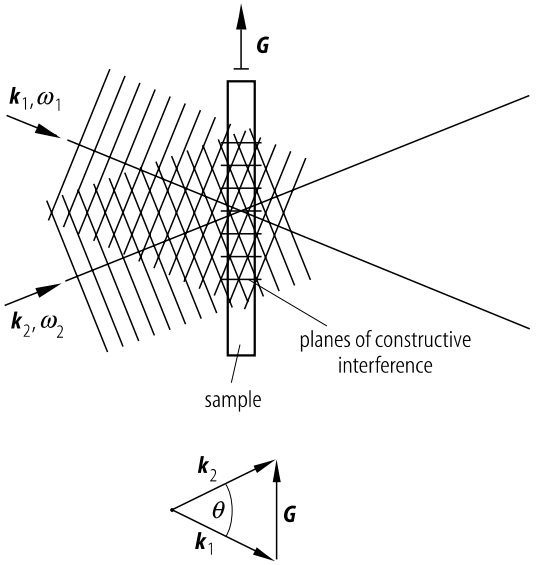


Fig. 25.3. The interference of two laser beams on a sample

via incoherent population, which they excite, one obtains a phase-, amplitude- or mixed grating, which diffracts light.

The names four-wave mixing (FWM) and laser-induced grating (LIG) are often used synonymously. If one wants to differentiate them, one can use FWM for the coherent polarization gratings and LIG for incoherent population gratings.

One can now apply various criteria to distinguish different regimes.

A grating is considered as thick or thin if its thickness d is above or below a critical value d_c, respectively. The quantity d_c is given in simplest approximation by

$$d_c = \Lambda^2/\lambda \tag{25.16}$$

where λ is the wavelength of the light diffracted from a grating with period Λ. Refined formulae can be found in [80M1, 86E1].

The diffraction properties of thick gratings (the so-called Bragg-regime) differ from those of thin gratings (Raman-Nath-regime) as will be shown below for some examples.

The wavevectors of the diffracted orders $\boldsymbol{k}_{\mathrm{D,n}}$ are in principle given by

$$\boldsymbol{k}_{\mathrm{D,n}} = \boldsymbol{k}_i \pm \mathrm{n}\boldsymbol{G}, \quad \mathrm{n} = 0, 1, 2, \ldots \tag{25.17}$$

where $\boldsymbol{k}_i$ is one of the incident beams. The multitude of the values of n is much more restricted for thick than for thin gratings. See below and Fig. 25.4. For quantum structures one will be generally in the regime of thin gratings except e.g. for thick samples filled with semiconductor nanocrystals (e.g. semiconductor doped glasses).

Another criterion is to distinguish between gratings which are created by two beams of equal frequency. This is the so-called degenerate four-wave mixing (DFWM), defined by

$$\text{DFWM} : |\boldsymbol{k}_1| = |\boldsymbol{k}_2| \,; \omega_1 = \omega_2 \tag{25.18}$$

In this case the interference pattern is spatially stationary and the elastically diffracted light shows no frequency shift. The grating is washed out, if the diffusion length of the excited species l_{D} is comparable with Λ i.e. with (25.15b)

$$l_D G \geq 1 \tag{25.19}$$

and the diffracted intensity decays for smaller Λ. For an example see [88W1].

Non-degenerate four-wave mixing (NDFWM) occurs for generally small differences $\Delta\omega$

$$\omega_1 - \omega_2 = \Delta\omega \ll \omega_i; \quad i = 1, 2 \tag{25.20}$$

of the frequencies of the two incident laser beams, resulting in small differences of the $|\boldsymbol{k}_i|$; $i = 1, 2$

$$|\boldsymbol{k}_1| - |\boldsymbol{k}_2| \ll \Delta k = |\boldsymbol{k}_1 - \boldsymbol{k}_2| \ll |\boldsymbol{k}_i| \,; i = 1, 2 \tag{25.21}$$

The grating or interference pattern moves in this case laterally with a speed v_{gr} given by

$$v_{\mathrm{gr}} = \frac{\Delta\omega}{\Delta k} \tag{25.22}$$

The diffracted orders are Doppler–shifted by integer multiples of $\Delta\omega$. In addition sum and difference frequency generation can occur.

Note that the grating structure is washed out, if the grating moves over one period in a time short compared to the characteristic decay time T_c of the species responsible for the grating [78Y1, 78Y2] i.e. if

$$T_c \Delta\omega \geq \underline{1} \tag{25.23}$$

For examples see e.g. [88R1, 92S2, 97W2], and the references therein.

A further criterion is how the gratings are read out. See Fig. 25.4

The simplest possibility is to have the two beams, which form the grating diffracted off it themselves. This so-called self-diffraction method gives diffracted orders in transmission or reflection according to Fig. 25.4a. A thin grating produces several orders if the spatial modulation of the optical properties is non-sinosoidal. The transmission direction $+0_{\mathrm{t}}$ contains the transmitted part of the incident beam "$\boldsymbol{k}_1$" and the $+1$ diffracted order of beam $\boldsymbol{k}_2$. The transmission direction $+1_{\mathrm{t}}$ gives the first diffracted order of beam $\boldsymbol{k}_1$ and the second of $\boldsymbol{k}_2$ and so on. Similar statements hold for the reflected orders. Sometimes orders up to six or higher can be observed [86E1]. In most cases the diffracted amplitude decreases rapidly with increasing order. In this case e.g. "-1_{t}" gives essentially the minus first diffracted order of beam $\boldsymbol{k}_2$.

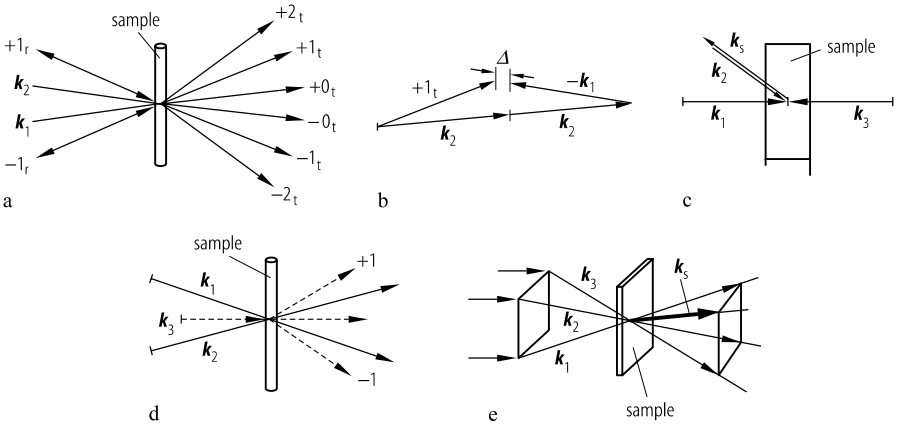


Fig. 25.4. A selection of geometries for four-wave mixing and laser-induced grating arrangements

An exact fulfillment of (25.24), (25.25) resulting from (25.17), (25.18) is not possible in self-diffraction e.g. in DFWM

$$\boldsymbol{k}_{\mathrm{diff}} = \boldsymbol{k}_1 + \boldsymbol{G} = \boldsymbol{k}_1 - (\boldsymbol{k}_1 - \boldsymbol{k}_2) = 2\boldsymbol{k}_2 - \boldsymbol{k}_1 \tag{25.24}$$

and

$$|\boldsymbol{k}_{\mathrm{diff}}| = |\boldsymbol{k}_1| = |\boldsymbol{k}_2| \tag{25.25}$$

As shown in Fig. 25.4b there remains a small misfit of $\boldsymbol{k}$ called $\boldsymbol{\Delta}$, which is possible for a thin grating since the place at which diffraction occurs is limited by the grating or sample thickness d and consequently the normal component of $\boldsymbol{k}$ is defined only according to

$$d\Delta \approx 1 \tag{25.26}$$

Multiplied with $\hbar$, (25.26) is just the uncertainty relation. On the other hand (25.26) is equivalent to (25.16). Note that self-diffraction is also possible for NDFWM.

For thick gratings the acceptable value of Δ tends to zero and consequently there is only diffraction of beam $\boldsymbol{k}_1$ into to direction "-0" i.e. in the direction of $\boldsymbol{k}_2$ and vice versa. Since transmitted and diffracted beams are difficult to separate or to distinguish, one chooses often a different geometry for thick gratings shown in Fig. 25.4c. The beams $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ interfere to form a grating from which beam $\boldsymbol{k}_3$ is diffracted as signal beam $\boldsymbol{k}_{\mathrm{S}}$ counterpropagating to $\boldsymbol{k}_2$. The beams $\boldsymbol{k}_{\mathrm{s}}$ and $\boldsymbol{k}_2$ can be now easily separated by a beam splitter. Note that the role of the three beams $\boldsymbol{k}_1$, $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$ can be interchanged. In all cases the signal beam propagates in direction $\boldsymbol{k}_{\mathrm{s}} = -\boldsymbol{k}_2$. The grating constants are however different e.g. for the interference of $\boldsymbol{k}_1$ and $\boldsymbol{k}_3$ or of $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$, respectively. Consequently the gratings may be differently

influenced by diffusion. See (25.19). Sometimes one calls two of the three incident beams pump beams and the third one test beam. If the thick sample is absorbing, care has to taken to have sufficient spatial overlap between the beams.

With the arrangement of Fig. 25.4c the term "four-wave mixing" becomes clear. Three beams interfere to produce a forth one. In self-diffraction one of the incident beams is used twice as pump and as test beam.

A slightly more elaborate technique also for thin gratings is to write the grating by beams 1 and 2 and to probe it with a third independent beam 3 as shown in Fig. 25.4d. There can be a lot of diffracted orders: not only of $\boldsymbol{k}_3$ diffracted off the grating produced by $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$, but as in Fig. 25.4c all other combinations of self- and non-self-diffraction may occur. A possibility, to separate the various orders is to have the three beams $\boldsymbol{k}_1$, $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$ not coplanar. An especially simple arrangement is to send the beams from three corners of a square onto the interference spot of the sample. The beam which results from an interaction of all three incident beams is then radiated into the direction of the remaining fourth corner in transmission. See Fig. 25.4e.

A further criterion is to distinguish the state of polarization of the incident (and diffracted) beams. If the beams $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ are equally polarized, e.g. both linear and parallel (i.e. collinear), then their interference leads to a real spatial modulation of the field amplitude and intensity in the interference pattern. If they are e.g. linearly but orthogonally polarized, then the light intensity will be spatially constant over the spot, but the state of polarization will vary periodically from linear over elliptic to circular and back.

A different approach to the whole group of phenomena is to stress not the interference of waves and the diffraction off the interference pattern, but the interaction between light quanta, generally polaritons as discussed in Chap. 19.

It is obvious that FWM and LIG offer a huge variety of experiments, which multiplies if time resolution (Sect. 25.3) is added. Therefore we can give only a very small selection of the information, which can be deduced from these experiments. The simple observation of any diffracted order proofs that there must be some optical nonlinearity. These nonlinearities can be due to real and virtual (or incoherent and coherent) excitations in the electronic system of the semiconductor, excitation-induced (or collision-) broadening of exciton resonances, two-photon transitions, or the transition to an electron–hole plasma. A tuning of $\omega_{1,2}$ allows to detect resonances, i.e. frequencies for which certain nonlinearities are especially strong. Non-self-diffraction (Fig. 25.4d) allows to probe e.g. dispersive nonlinearities in the transparent part of the sample caused by a population grating created by the beams $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ situated in the absorbing region.

Similarly to the pump-and-probe beam technique, thermal gratings must be considered, which may result from the deposited energy [91T1].

The intensity of the diffracted signal varies in lowest approximation linearly with the intensity of all three incident beams i.e.

$$I_s \sim I_1 I_2 I_3 \tag{25.27}$$

The proportionality constants contains terms squared which describe the variation of refractive index or absorption coefficient with light intensity, population density, etc. in the wave picture, the (dipole-) transition matrix elements and resonance denominators in the quantum picture. Generally one observes a trend towards saturation for high intensities.

A variation of Λ via θ (25.15) allows to determine the diffusion length of the excited species. See [88W1, 92O1, 92S2] as examples.

Non-degenerate four-wave mixing allows to determine even with long pulses extremely short dephasing times of the order of a few tens of fs [97W2].

25.3 Time-Resolved Spectroscopy

Time-resolved spectroscopy is one of the most powerful tools to investigate the dynamical processes of excited species in semiconductors. Almost steplike progress of the field was often connected with the development of new lasers with shorter pulses and easier handling, the most recent example being the titanium sapphire laser (also Al_2O_3:Ti or Ti-Sa) [93K1, 97G2, 98F1, 98I1, 01A1, 03C1, 03S1, 03S2, 03T1, 03T2, 05F1]. The main dynamical processes which can be studied by time-resolved spectroscopy are presented in detail in Chap. 23.

A good overview of the development of this field is presented again in Ref. [88N1, 04O1] of Chap. 23.

25.3.1 Equipment for Time-Resolved Spectroscopy

To obtain time resolution one needs a pulsed or temporally modulated excitation source and possibly a time-resolving detection system.

We start with the excitation sources. One possibility is to chop or modulate a cw beam. Pulse durations, which can be easily reached with Pockels-cells or acousto-optic modulators go down to the sub μs regime and reach even to a few ns.

The other possibility is to use light sources, which already emit pulses. Flashlamps are available for the ms and µs regimes.

Q-switched neodymium and chromium lasers emit pulses of typically a few ns duration. The same holds for N_2 and excimer lasers and dye lasers pumped by them.

Shorter pulses from 100 ps down to about 10 fs are produced by mode-locking ([98I1] and references therein). Mode-locking means a suitable, phase-stable superposition of various modes of a laser, which results in short bursts

of light. The repetition rate ν_R of mode-locked lasers is given by the inverse of the round-trip time of the pulses in the cavity T_R, and the duration of the pulses T_P is limited by the spectral width $\Delta\omega$ of the optical amplification spectrum of the active- or gain medium of the laser.

$$T_R \nu_R = 1 \tag{25.28a}$$

$$T_P \Delta\omega \approx 1 \tag{25.28b}$$

Mode-locked Ar^+ lasers reach e.g. pulse durations in the 100ps regime. Mode-locking can occur spontaneously like for TiSa lasers or it can be induced actively e.g. by electro-optic or acousto-optic modulators with are tuned to the round-trip time T_R, or passively by saturable absorbers.

If shorter pulses are required than produced by a mode-locked Ar^+ or Nd laser, it is possible to pump with the mode-locked Ar^+ (or with the fundamental or second harmonic of a mode-locked Nd(YAG-) laser) synchronously a dye or colour-center laser. Synchronously means in this context that the round-trip times of both lasers are equal. The dye or colour-center lasers can emit shorter pulses in the sub-ps to the 100 fs regime due to their wider gain spectra [87T1, 88K1, 92S1].

The next step of sophistication were colliding pulse mode-locked (CPM) lasers pumped usually by a cw Ar^+ laser [83F1, 87H1]. Here two counter-propagating pulses are circulating in the same cavity which contains the gain medium (usually a dye jet) and the saturable absorber at places chosen for optimal performance.

Many of these ps to fs lasers became more and more out-dated by the invention of the titanium sapphire laser. The Al_2O_3:Ti crystal is either pumped by a cw Ar^+ laser or the second harmonic of high power semiconductor laser diode arrays. The TiSa laser starts mode-locking by itself via the Kerr-lens mode-locking effect. Typical pulse durations are 100 fs and the corresponding Fourier-limited width is about 20 meV. The laser is tunable from the near IR to the red.

The pulses can be amplified e.g. in regenerative amplifiers and shaped e.g. with pulse stretchers and compressors, the later ones reaching pulse durations below 10 fs.

A detailed description of these techniques is beyond the scope of this book and we refer the reader e.g. to [98F1, 98I1, 01A1, 03S1, 03T1, 03T2, 05F1] and references therein.

Tunable fs emission can be also obtained from optical parametric oscillators (OPO) or -amplifiers (OPA) [97H1].

Presently pulses with a minimum duration of a few fs can be reached in the IR, the red and the blue [87T1,88K1,92S1,98S2,02M1,02S1]. These pulses are necessarily spectrally very broad and react very sensitively on a passage through lenses, windows or other optical elements.

An actual line of development aims at small (shoe-box size) fs-lasers based on semiconductor laser diodes and glass fibers [98I1].

Almost deliberate pulse-shaping becomes possible by the development of diffraction structures combined with suitable apertures [98C1].

The ultimate limit of pulse durations has been discussed in [90K1].

Short THz pulses can be created from Bloch oscillations (see Sect. 23.2) but also from switching processes of strip lines, difference frequency generation etc. [94S1, 97N2, 98C1, 00H1, 00L1, 04K1].

The next aspect is to measure the duration of the laser pulses. These techniques are partly identical to those which are used to investigate e.g. the decay time of the luminescence. So we treat them here partly together.

Time constants from cw down to a few ns can be easily monitored by photomultipliers, semiconductor- or -vacuum photodiodes and a (storage) oscilloscope. Faster, but repetitive processes down to the sub ns regime can still be recorded by sampling oscilloscopes and related techniques.

The range from μs down to a few ps is most conveniently covered by a streak camera. The way in which such a streak camera works, is shown schematically in Fig. 25.5.

The light beam, which shall be temporally dispersed falls on a photocathode. The photo-electrons are accelerated and imaged on a luminescence screen. On their way they are deflected in a capacitor by a voltage varying linearly with time (see upper part). Consequently one obtains a time resolved signal on the screen. The time resolution depends for fixed acceleration voltage and geometry of the set-up on the speed, with which the deflection voltages varies and on the height of the entrance slit. Both parameters can be changed. Usually light amplifier stages are incorporated in the tube of the streak camera.

There are single shot cameras which sweep with a certain time delay after the arrival of a trigger pulse and synchroscan cameras which are synchronized

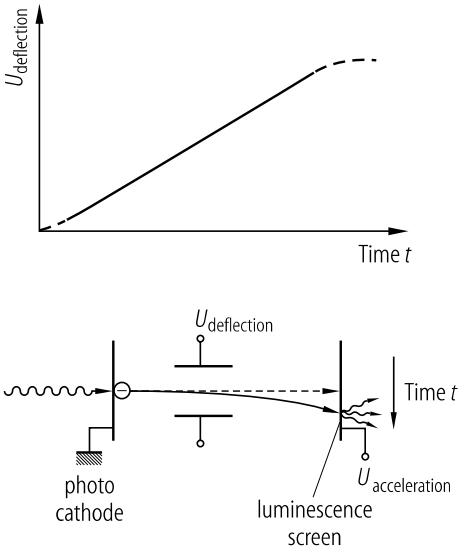


Fig. 25.5. Schematic diagram of the operation of a streak camera

to the repetition rate of a mode-locked laser. These repetition rates are often in the 80MHz range. Normal to the paper of Fig. 25.5 one can have a photon energy axis, e.g. if the streak camera is placed behind a spectrometer. On the screen one obtains then a three-dimensional plot namely signal intensity as a function of time and photon energy. This allows e.g. to monitor the temporal and spectral evolution of the luminescence or of a reflected or transmitted probe pulse. For an example see [98U1].

The regime from a few (tens) of ps down to the shortest optical pulses of a few fs has to be covered by correlation techniques. For the characterization of the laser pulse itself one uses autocorrelation techniques. There are many different ways of these autocorrelation techniques (see e.g. [93K1,97G2,01A1, 03S1,05F1] and references therein). We show here schematically one example in Fig. 25.6a.

The pulse, which has to be investigated, is split into two parts, which can be delayed with respect to each other by a variable delay stage. Then they are focussed together on a nonlinear crystal which produces the second harmonic. Every beam is frequency doubled by itself, but in the so-called background-free direction one obtains a signal which is proportional to:

$$S(\tau) = \left| \int \boldsymbol{E}_1(t)\boldsymbol{E}_2(t+\tau)\mathrm{d}t \right|^2 \tag{25.29}$$

where τ is the relative delay time between the pulses.

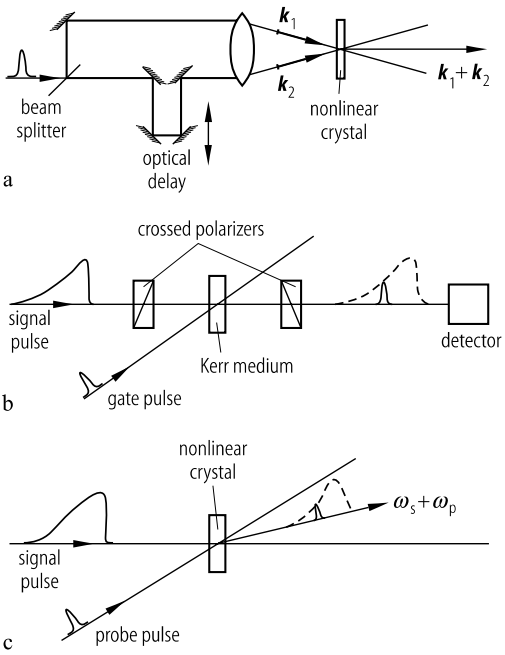


Fig. 25.6. Schematic set-up for an autocorrelation measurement **(a)** and for time resolved luminescence spectroscopy using a Kerr-cell shutter **(b)** or up-conversion **(c)**

Evidently $S(\tau)$ goes to zero, if the pulses do no longer overlap. A problem with many of these techniques is that they do not give the complete information of the pulse and require some additional information e.g. whether the pulse is Fourier limited or not.

Complete information is available with techniques known as FROG (frequency resolved optical gating) See e.g. [93T1,97G2,98J2,98L3] and references therein. The basic idea is to measure via an optical gate the spectrum of the laser pulse at various times. The gate is operated by a part of the laserpulse itself. These so-called FROG traces contain the complete information on the electric field $\boldsymbol{E}(t)$ of the laser pulse and allow to detect also a so-called chirp i.e. a shift of ω of the pulse with time.

Correlation techniques can also be used to determine the temporal development of e.g. the luminescence light or a transmitted probe pulse. One can send the pulse through a sequence of two crossed polarizers with a Kerr-cell in between. This arrangement transmits light only during the interval over which a short, suitably polarized gate pulse induces birefringence in the Kerr medium (Fig. 25.6b). Alternatively the luminescence at frequency ω can be mixed with a short probe pulse of frequency ω_{P} and the sum frequency signal is detected as a function of the relative delay between luminescence and probe pulse (Fig. 25.6c). Examples for these techniques are found e.g. in [96S1, 99L3] and the references given therein. It should be noted that every spectral resolution of the pulses e.g. in a grating monochromator results in a lengthening of the pulse. Therefore it is necessary to perform the time-resolution before the pulse is spectrally dispersed. For details of this aspect see e.g. and [97G2].

The time delay can be controlled with extremely high precision by using the Pancharatnam phase [97W3].

25.3.2 Experimental Techniques and Results

We proceed in this subsection from longer to shorter times i.e. from the lifetime (T_1) over intraband relaxation (T_3) to dephasing times (T_2) and coherent effects.

Before we give some examples for time-resolved spectroscopy, we should stress again significant differences between bulk or 3d materials and systems of reduced dimensionality like single and multiple quantum wells. See also Chap. 23. In bulk material, the wavevector $\boldsymbol{k}$ is a three-dimensional quantity. If an exciton, or more precisely a polariton hits the surface of the sample, only the component of $\boldsymbol{k}$ parallel to the surface is conserved, due to the reduced translational invariance of a surface. Consequently an exciton or polarition that hits the surface can contribute with a certain transmission probability to the emission e.g. to luminescence as long as its parallel component $k_{\|}$ of $\boldsymbol{k}$ is

$$k_{\|} \leq \frac{\omega}{c}; \quad |\boldsymbol{k}| \geq k_{\|} \tag{25.30a}$$

whatever the amount of $\boldsymbol{k}$ is. For ideal quantum wells or wires, the parallel component of $\boldsymbol{k}$ is in contrast to bulk already the total $\boldsymbol{k}$, i.e.

$$k_{\parallel} = |\boldsymbol{k}| \tag{25.30b}$$

For localized states, which tend to appear at the bottom of the bands due to fluctuations of the composition or of the width of wells or wires or for quantum dots and ∼islands, $\boldsymbol{k}$ conservation is, however, relaxed.

This has consequences on the luminescence and its dynamics as discussed e.g. in [92O1] or in Sects. 13.1, 15.1 and 23.4

25.3.2.1 Lifetime Measurements

Luminescence decay measurements are often used to determine the lifetime of some excited species like excitons. For examples see [95M1, 96S1] and the remarks to this topic in Sect. 23.5. The statement given already several times that the luminescence monitors only the fate of a minority of excitons or electron–hole pairs is also valid here. However, luminescence decay measurements are easy to perform and therefore often used as means to learn something about T_1.

The luminescence decay includes radiative and nonradiative processes, characterized in the simpliest case by their respecitve time constants T_{r} and T_{nr}. The luminescence decay time T_{lum} is then given by

$$\frac{1}{T_1} = \frac{1}{T_{\mathrm{lum}}} = \frac{1}{T_{\mathrm{r}}} + \frac{1}{T_{\mathrm{nr}}} \tag{25.31}$$

as long as no complications occur as the relaxation from high energy states to the luminescing ones.

In most cases it will be therefore not correct to identify the luminescence decay time with the radiative lifetime and this quantity (25.31) with the dephasing time.

In ideal, quasi two- or one-dimensional structures without $\boldsymbol{k}$-relaxing localization effects, the guided and antiguided or radiative polariton modes have to be considered. See Sects. 9.3, 13.1 or 15.1. However, at low temperature and excitation density the luminescence comes generally from more or less deeply localized states. This is no longer true at higher excitation levels e.g. in an electron–hole plasma, where all localized states are filled and where the luminescence and gain spectra reflect the density of states of free particles. For an example see e.g. Sect. 21.4.

More reliable methods, which involve all excited species, would be time-resolved pump- and probe beam experiments, which exploit an optical nonlinearity to which all excited species contribute like an excitation-induced collision broadening of a (exciton) resonance or the band-gap renormalization. However, even then the time dependent distribution can influence the nonlinearity.

An additional complication may arise through the stimulated emission, which reduces the effective lifetime of the carriers. In [97M2, 99H2] it has been nicely shown, how this stimulated emission can be manipulated by an additional pulse with some excess energy, which creates on one hand side additional carriers, but transiently heats the carrier gas to temperatures, which turn the carrier gas from an degenerate, inverted population into a non-degenerate one thus switching off the stimulated emission.

The decay of the diffracted intensity of a laser-induced diffraction grating (LIG) (see Sect. 25.2) depends both on the lifetime T_1 and on the diffusion-length l_D or ∼constant D of the excited species. The decay time constant of the signal intensity after pulsed excitation of the LIG τ_s is given in simplest approximation by [78S1]

$$\frac{1}{\tau_\mathrm{s}} = \frac{2}{T_1} + \frac{8\pi D}{\Lambda^2} \tag{25.32}$$

where Λ is the lateral period of the LIG. A plot $\tau_\mathrm{s}^{-1} = f(\Lambda^{-2})$ gives by its slope D and its abcissa T_1. For an example see e.g. [92O1].

A way to determine an effective lifetime T_1 without time resolution is to use the equation

$$N = G \cdot T_1 \tag{25.33}$$

where N is the number (or density) of excited species, and G the generation rate (per unit volume or unit area). This method can be applied if N can be determined in absolute numbers e.g. by a certain optical nonlinearity like a certain amount of excitation-induced broadening or bleaching of an exciton resonance, which can be directly related to the density of present carriers. For an example see e.g. [98J1, 02J1]

25.3.2.2 Intraband and Intersubband Relaxation

The intra- and intersubband-relaxation can be followed most conveniently by time-resolved luminescence spectroscopy. Examples can be found e.g. in [98G1,98K2,98K4,98S1]. Another possibility is time-resolved pump-and-probe spectroscopy.

The relaxation of a population e.g. through a tail of localized and radiative states can be seen in the leading edge of the luminescence dynamics. Often one finds a delayed onset of luminescence at lower photon energies or a red shift of the emission maximum with time [93H2, 93K2, 98U1, 99K2]. A detailed investigation of the intraband relaxation via LO and acoustic phonon scattering has been deduced from time-resolved luminescence spectroscopy e.g. in ZnSe or CdTe-based quantum wells [98U1].

The evolution of the emission from the resonantly excited states over a non-thermal, then thermal but hot distribution (i.e. $T > T_\mathrm{lattice}$) and eventually to a thermal one in equilibrium with the lattice ($T = T_\mathrm{lattice}$), if the lifetime is sufficiently long, is known also as hot luminescence. Since free states with large ($|\boldsymbol{k}| > \omega/c$) cannot contribute to the luminescence in quasi two- or

one-dimensional structures (see above) it is sometimes good to monitor the radiative recombination to a strongly localized state e.g. a hole bound to an acceptor, which relaxes $\boldsymbol{k}$ conservation [94H1, 94H2, 94H3, 95L1, 99S1] or to look for LO phonon replica (see Fig. 23.33). The warning that luminescence spectroscopy monitors often only the fate of a minority of the excited species (electron–hole pairs, excitons etc.) holds also here.

Experimentally more difficult is to monitor the temporal evolution of a cloud of carriers created with some excess energy in time-resolved pump-and-probe spectroscopy or differential transmission spectroscopy. An example is found in [92F1]. Here all excited species contribute to phase-space filling and screening of the Coulomb interaction, however with a weight depending on their distribution function.

25.3.2.3 Coherent Processes

We come now to some examples of coherent processes. The dephasing time T_2 of the interband polarization and the lifetime T_1 are connected by the inequality

$$T_2 \leq 2T_1 \tag{25.34}$$

In most cases one has $T_2 \ll T_1$, but the upper limit for T_2 is given by $2T_1$, if there are no other phase-destroying processes than recombination. The factor 2 comes from the fact that T_2 describes the decay of an amplitude and T_1 of a population.

One should note that there are other, partly independent dephasing times like the one of the intraband polarization or of the spin. See e.g. [92B2, 96B2, 96B3, 96B4] or Sect. 23.2.

Especially the dephasing time of spin can be much longer than the dephasing of the real space part of the wavefunction. This is the case e.g. for electrons with negligible spin orbit coupling (e.g. for $l = 0$; $j = s = \frac{1}{2}$) but not for almost degenerate hole states ($j = \frac{3}{2}, \frac{1}{2}, l = 1, s = \frac{1}{2}$). See e.g. [96O1,96O2].

The most widely used technique to determine T_2 and related phenomena are FWM experiments. See Sect. 23.2 or [89D1]

Before we give examples, we stress here also some differences between bulk samples and quantum structures.

A short pulse propagates in bulk material as a polariton wave packet [83M2, 85H1, 85T1, 93P1, 94B1, 95L2, 95V1, 00D1, 00D2, 01A1] i.e. the polarization and the electric field amplitude propagate together. Consequently the overlap of two short pulses "1" and "2" in a four-wave mixing experiment, which have a certain relative time delay t_{12} and which are supposed to form the grating is an issue, which has thoroughly to be considered for bulk samples [83M2, 85T1, 00D1, 00D2]. For a single or a few quantum wells the same holds for pulse propagation in the plane of the well. The situation becomes however different for the generally used geometry that both pulses hit the sample under almost normal incidence. The (macroscopic) polarization created in a quantum well can decay by radiating an electromagnetic

field, by destructive interference between different oscillators in an inhomogeneously broadened ensemble, by dephasing processes like scattering with phonons, but it stays confined in the well. Therefore there is no problem for the polarization created by a second, delayed pulse in the well to overlap with (the left over of) the polarization of the first pulse. This fact or advantage compared to bulk material is one of the reasons that a dominant part of coherent spectroscopy is performed in quantum structures or thin films.

It should be noted that the situation must develop again to bulk-like, anisotropic behaviour for stacks of quantum wells or superlattices with a total thickness exceeding the wavelength of the light and an internal spacing small compared to this quantity. See e.g. [98K3] but also [96H2, 96H3] for larger spacing.

Another important aspect in time-resolved FWM experiments is to distinguish between the time delay between the incident pulses t_{12}, which does not involve a time resolved detection and the "real time" observation of the diffracted signal appearing at or after the arrival of the second pulse. For examples see [92G1, 96W1] and Sect. 23.2.1.

Now we present some selected experimental techniques. The dephasing time T_2 of the interband polarization can be measured in the time or in the frequency domain.

Linear measurement in the frequency domain include investigations of the spectral width of an absorption peak, which is inversely proportional to T_2, or a fit of the reflection feature. However, in both cases one has to know by other arguments that the spectral feature under investigation is homogeneously but not inhomogeneously broadened. This is often the case for exciton resonances at elevated temperatures (e.g. room temperature) but usually not for quantum structures at very low temperatures, e.g. liquid He temperature.

Spectral hole burning and nondegenerate four-wave mixing (NDFWM) are nonlinear techniques in the frequency domain which allow to measure T_2 values down to a few tens of fs, which became accessible in the time domain only very recently. As a first example we cite the dephasing measurements by spectral hole burning in the gain region of an inverted semiconductor [93M1]. Another example are CdS and CdSe quantum dots in glass matrices. Spectral hole burning [93W1, 97W2], non-degenerate four-wave mixing [97W2] as well as time resolved four-wave mixing [93S1, 94M1] revealed independently and consistently temperature and excitation density dependent dephasing times of a few tens of fs. It should be noted that similar dots in organic matrices can have considerably longer dephasing times, as deduced from spectral widths $\leq 0.1\,\mathrm{meV}$ observed in single dot spectroscopy [96E1]. Spectral hole burning allows with sufficient spectral resolution on the other hand also to deduce extremely long dephasing times [92M1].

Some further examples for NDFWM with ns and fs pulses can be found in [92S2, 98K5].

Four-wave mixing spectroscopy in the two-beam, self-diffraction configuration and using two short incident pulses (see Fig. 25.8a) is the usual way to determine the dephasing time T_2 of a resonance in a quantum well or ∼wire or ∼dot sample in the time domain. See Sect. 23.2.

As an extension of these measurements it is possible to send a prepulse on the sample, which arrives at a time t_{13} before pulse one. This allows to investigate the excitation-induced dephasing, i.e. the decrease of T_2 induced by a well defined incoherent population density if

$$T_1 > t_{13} > T_2 \tag{25.35}$$

and by a polarization for

$$T_2 > t_{13} \tag{25.36}$$

In the latter case also diffracted orders can be observed, which result from the interference of beams 3 and 1 or 2. For examples see e.g. [98K4].

Recent topics include the investigation of exciton (or carrier) phonon and exciton-exciton scattering, and its reversibility in the regime of quantum coherence also known as coherent control (see Sect. 23.2 or [95B3, 99W2]), the investigation of the dynamics of Rayleigh scattering and the speckle associated with it [94D1, 94S2, 96S1, 99L2], or the investigation of the decay of the excitonic polarization [94B1].

Another group of experiments in the coherent regime is the investigation of beat phenomena. See again Sect. 23.2 or [96H2, 96H3, 96M2].

The quantum beat spectroscopy allows to detect a (roughly constant) energy splitting e.g. between exciton and biexciton transition, even if the inhomogeneous broadening of the exciton resonance is comparable to or even larger than this splitting. For examples see e.g. [96A1, 99E1].

In superlattices it is possible to observe in the presence of an electric field applied perpendicular to the layers so-called Bloch oscillations see Sect. 23.2.1.2.6.

The dephasing of spin states can be investigated by the Hanle effect in a magnetic field [96O1, 96O2].

Examples of time resolved spectroscopy in the THz regime are published e.g. in [97N2, 00H1, 00L1, 04C1, 04K2].

25.4 Spatially Resolved Spectroscopy

Spatially resolved spectroscopy is generally used to investigate the diffusion length of excited species (carriers, excitons) both in bulk material and in structures of reduced dimensionality as well as ballistic and tunnelling transport. For the latter group of samples spatially resolved spectroscopy is also frequently employed to separate one or a few localized states in an inhomogeneously broadened resonance.

25.4.1 Equipment for Spatially Resolved Spectroscopy

The main equipment are various types of microscopes, pinholes or other apertures. Conventional microscopes are limited in their resolution roughly by the wavelength of the light. Confocal microscopes reach a resolution $d_{\min}$ which is a factor $1/\sqrt{2}$ smaller [90W1, 98L1, 99H1]

$$\text{Conventional}: \; d_{\min} = 0.61\lambda/\text{NA} \tag{25.37a}$$

$$\text{Confocal}: \; d_{\min} = 0.61\lambda/(\text{NA} \cdot \sqrt{2}) \tag{25.37b}$$

where λ is the used wavelength and NA the numerical aperture. For good microscope objectives NA can reach values around one.

In both cases, it is easier to use microscope objectives with long working distance at least for low temperature measurements. This allows to keep the sample inside a cryostat and the optics outside, which facilitates handling and alignment considerably. Note that the microscope objectives have to be corrected for the cryostat window, if the resolution limit shall be reached. However, better NA can be reached with objectives in the cryostat. For the use of solid state immersion lenses see [99V1, 99V2, 03D1, 03M1].

The set-up for a confocal microscope is shown schematically in Fig. 25.7.

The improvement of the resolution according to (25.37b) results from the fact that the excitation beam is focused on a small volume, and that the emission (luminescence, scattered light, etc.) is re-collected from this excitation spot by the same optics. Excited species, which diffuse out of the focal "point" contribute hardly to the detected signal. Example can be found in [98D1, 98D2, 99L1, 02Z1, 03D1, 03M1, 03Z1].

Cathodoluminescence in a scanning electron microscope (SEM) can be also used for spatially resolved luminescence measurements. If no spatial resolution

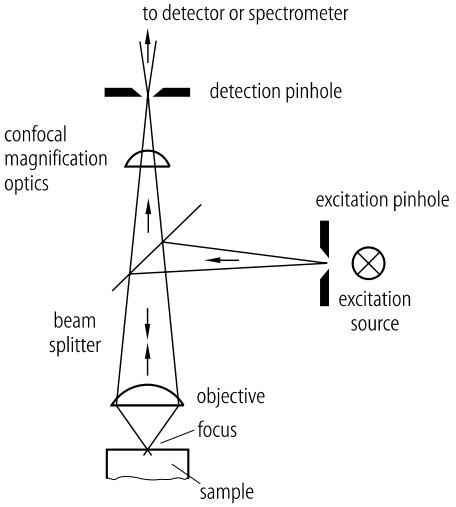


Fig. 25.7. Schematic set-up for a confocal microscope

is introduced in the detection system of the luminescence, the resolution is limited by the diameter of the exciting electron beam or the diffusion length of the excited species, whatever is larger. For an example see [91H1].

While far-field optics does not allow to reach a resolution below the diffraction limit given by (25.37), it is possible in near-field optics.

One can either produce, e.g. by lithographic techniques, pinholes in opaque (metal) layers on the sample, with diameters below the limit given in (25.37). Due to the close contact, an incident excitation beam leaks through the pinhole as long as the thickness of the layer is smaller than its diameter. In turn, luminescence light can leak out of such a pinhole. Problems may arise from polarization-dependent field enhancement effects at the metal edges of the pinhole and from strain induced by the metal layer. For examples see [00H2].

More versatile but also more expensive are so-called near field scanning optical microscopes (NSOM) also called frequently (especially in Europe) scanning near field optical microscope (SNOM) [01K1].

The basic idea is the following (see Fig. 25.8)

A fiber tip is produced by pulling or etching [99A1] a monomode glass fiber. See [95P1] and references therein. The aperture of this fiber can be considerably below λ in the range down to or even below 100nm. If a light beam is sent down the fiber it produces an evanescent light field at the tip, which can couple to the sample, if it is brought very closely (distance $\ll \lambda$) to the sample surface. Alternatively it can collect luminescence from the sample from an area which couples to the evanescent light field. For the approach of the tip to the sample surface and for the scanning over the surface, techniques are used, which are borrowed from scanning force microscopy. For an uncoated tip, the evanescent field originates not only from the end of the tip but partly

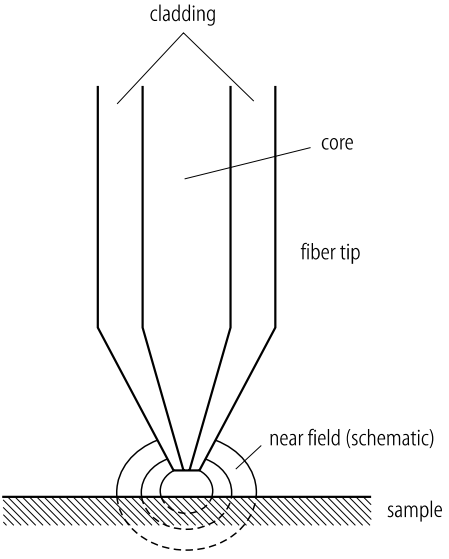


Fig. 25.8. Schematic drawing of the fiber tip of a SNOM and of the evanescent light field

also from its flanks [98F2,98K6]. The aperture can be more tightly defined by a reflecting metal coating (Al) of the flanks [95P1].

The prize one has to pay for circumventing the diffraction limit of resolution is a low optical through-put of the fiber tip (or of the pinhole). It depends on the width of the aperture and on the presence of a metal coating and can be as low as 10^{-4} for a single passage.

Therefore one often excites the sample globally by conventional optics and collects the light locally with the SNOM or vice versa. Local excitation and collection through the SNOM would give the best spatial resolution but involves the limited through-put twice i.e. squared.

Note that the reconstruction of spatial structures from SNOM scanning patterns is a nontrivial task [98F2, 99A1].

Present trends in the use of SNOMS aim towards single molecule [96B1, 96H1, 01K3], single quantum dot resolution [96E1], single localization site spectroscopy [98V1, 00F1] including homogeneity tests [97L1], (ultra-)high time resolution [93W1,95B1,95W2,99L1], low temperature applications [97B1, 00F1], polarization resolved imaging [96L1,99A1,99B1] or transport measurements [97R1, 99R1]. Nearfield photo-current spectroscopy has been reported in [97L2]. A light-induced thermal expansion of the SNOM tip has been found in [96L2].

For a review, also see e.g. [01K2]. In summary, it can be stated that a SNOM with an uncoated tip gives a spatial resolution around 0.2 to 0.3 μm, i.e. not much better than a confocal microscope with a solid-state immersion lens. The first has a larger scanning field on the sample surface, but is much more expensive than the latter.

A SNOM with a metal-coated tip reaches a spatial resolution below 0.1 μm but the throughout through the metal coated tip is very small, as already mentioned.

There are further methods to investigate diffusion and transport e.g. via laser-induced gratings (see Sect. 25.2 and 25.3) which do not need a special equipment.

For other approaches to overcome the for field diffraction limit see [03H1].

25.4.2 Experimental Techniques and Results

In the following, we give some examples of experiments, in which spatial resolution is crucial.

The diffusion length l_D of some excited species can be determined by various techniques.

A rather simple method is to excite the sample in a narrow spot with diameter $< l_D$ and to observe the luminescence resulting from the recombination spatially resolved. This method has the big disadvantage that the luminescence light is scattered in the sample and at its surface or interfaces by any inhomogeneity. Therefore one can in most cases argue, whether a luminescence photon, which is detected at a place outside the excitation spot,

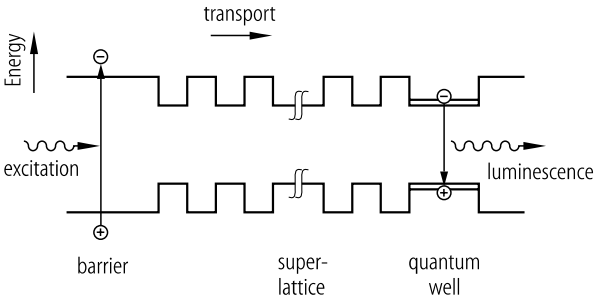


Fig. 25.9. An arrangement, which allows to detect transport of electron–hole pairs through a superlattice

originates from the recombination of an electron–hole pair, which diffused to this place and recombined there radiatively, or it recombined radiatively in the excited area and the luminescence photon has been scattered out of the sample at the place of its detection. For a recent example see [01N1]. Note that in both cases the luminescence signal will decrease with decreasing distance from the excitation spot.

If the density of excited species is so high that they result in some optical nonlinearities, other and more reliable techniques can be used like spatially resolved pump-and-probe experiments or diffraction from laser-induced population gratings (LIG). In the case of LIG measurements the grating constant has to be varied (see Sect. 25.2.2.3). The diffusion length can be deduced under quasi-stationary and pulsed excitation. For examples see e.g. [84K1, 85M1, 92O1, 98J1, 02J1].

The diffusion or tunneling of excited species into the depth of the sample can be monitored e.g. for a superlattice or for coupled quantum wells by exciting at the surface of the sample, e.g. in a cap layer and detecting the luminescence from a wider quantum well with smaller quantization energy in the depth of the sample, which can be reached by diffusion, ballistic transport or tunneling as shown schematically in Fig. 25.9.

Examples of the investigation of ballistic and diffusive transport with the help of a pinhole in a metal coating of the sample are found e.g. in [00H2].

Some further techniques, which have been introduced in the last years for the investigation of quantum structures are described in the following.

Brightness maps can be obtained by scanning e.g. a quantum well sample with an electron microscope or a focussed laser beam on the excitation side, or by a classical, or a confocal microscope or a near-field microscope (SNOM) on the detection side. Usually one obtains an irregular pattern of brighter and less bright spots or areas. For examples see e.g. [89C1, 90L1, 93B1, 98V1, 00F1]. Sometimes, the luminescence spectra are red shifted at the brighter spots compared to the less bright ones. Current interpretations include either the diffusion of excited carriers to and the relaxation into deeper localized states, resulting e.g. from local variations of the well width, or the diffusion to fast, nonradiative recombination centers. These centers can be point or line defects like dislocations [00D3].

Alternatively, information on the homogeneity of a transition can be obtained from reflection mapping [97L1].

The resolution of an inhomogeneously broadened luminescence band into a limited number (≤ 100) of individual localization sites can be achieved also by various types of microscopy or by preparing small and well separated mesa structures. Examples for both techniques are found in [99K1, 99W1, 00D3] or in [98P1], respectively.

See in this context the discussion about the continuous transition from disorder-induced localization sites via so-called self-organized or ∼assembled quantum islands or dots to nanocrystals in Sect. 8.13 or 15.4

Single dots or defects can also be observed by conventional microscopy, if they are sufficiently diluted [96E1, 99L4, 99L5].

The diffusion of excitons from GaAs wells into GaAs wires has been monitored by near-field spectroscopy in [97R2] and the transport in the well in [98L2].

To conclude this point on spatially resolved spectroscopy an aspect should be mentioned which is presently hardly exploited, though it is well known that a lateral confinement of the resolution Δx results in an uncertainly Δk_x according to

$$\Delta x \cdot \Delta k_x \geq 1 \tag{25.38a}$$

In SNOM experiments Δx can be brought considerably below the wavelength of the light λ i.e.

$$\Delta x < \lambda = 2\pi c/\omega \tag{25.38b}$$

Consequently Δk_x reaches values

$$\Delta k_x > 2\pi/\lambda = \omega/c \tag{25.38c}$$

This allows to couple to states which are not accessible in conventional spectroscopy like surface polaritons [97K1] or guided polariton modes in ideal quantum wells or wires

Furthermore, relative transition intensities can be manipulated via (25.38c) e.g. for quadrupole transitions, the transition matrix element of which varies linearly with k.

25.5 Spectroscopy Under the Influence of External Fields

The external fields which we discuss are mechanical stress/strain, electric and magnetic fields.

For experimental results see Chap. 16 and the references given therein and below.

25.5.1 Equipment for Spectroscopy Under the Influence of External Fields

For the spectroscopy itself one needs a set-up for linear or nonlinear spectroscopy and / or for temporally or spatially resolved spectroscopy as outlined in Sects. 25.1 to 25.4. In addition and this is what we discuss below, one needs some means to create and/or to apply the external fields.

Hydrostatic pressure up to some GPa can be applied in diamond anvil cells. See e.g. [94L1, 94L2] where also references for the construction of such cells are given. The samples are kept together with a small piece of ruby for calibration purposes in a liquid or, at low temperatures in liquid/solid He, which has enough plasticity to reach almost hydrostatic pressure conditions. The access to the sample is usually through the diamonds. The pressurized volume is usually tiny for two reasons: big diamonds tend to be expensive and high pressures are difficult to be realized over large areas with experimentally reasonably accessible forces. Typical sample dimensions are in the $100 \times 100\,(\mu\mathrm{m})^2$ range.

Two-dimensional compressive or tensile strain results generally also in some lattice distortion in the third dimension in the sense of uniaxial strain and can be created in different ways, e.g. through

- pseudomorphic growth of thin layers on a substrate with different lattice constant. This growth-mode is possible only over a thickness below a critical thickness l_c. Above l_c lattice relaxation sets in by the formation of dislocations or by the formation of hillocks or dots (see Sect. 8.13). The latter tend to alloying and/or segregation during overgrowth.
- mismatch of the thermal expansion coefficients. Even if a material grows on a lattice matched substrate or is grown in a thickness exceeding l_c, so that it adopts its own lattice constant, strain may occur during the cooling from the growth temperature (usually a few hundreds of °C) to the measurement temperature e.g. room- or liquid He temperature.

In some cases the quantum well sample has been glued on a new substrate and the original one has been removed. This procedure allows within some limits to choose the difference of the thermal expansion coefficients deliberately [89S3].

Inhomogeneous stress with a maximum in the volume of the sample can be created by applying the pressure via a round plunger (see e.g. [77M1, 03N1]).

Magnetic fields can be applied either parallel to the propagation direction of $\boldsymbol{k}$ i.e. of the incident light beam (Faraday configuration) or perpendicular to it (Voigt configuration). Since one has often $\boldsymbol{k}$ also normal to the plain of the quantum wells, the Faraday configuration leads to a complete quantization of the electron states in all three dimensions for sufficiently strong fields in quantum wells.

Design questions, which have to be answered for the design of a magnet and which enter in the costs, concern e.g. the diameter of the bore, the homogeneity

of the field, the volume over which it is homogeneous, and the arrangement of the windows.

For magnetic fields up to approximately 14 T one uses generally superconduction coils. Fields up to 25 T can be reached with resistive coils, e.g. of the Bitter type and 35 T with hybrid magnets which consist of an outer superconducting coil and an inner resistive one. Such set-ups are available e.g. at the high magnetic field laboratory in Grenoble.

Even higher fields up to 100 T can be produced only in a pulsed way by discharging a capacitor bench through the coil. Such facilities exist e.g. at the university of Tokyo or at the Humboldt Universität zu Berlin [81K1]. For recent references see e.g. the Conferenc Proceedings of the International Conferences on High Magnetic Fields in the Physics of Semiconductors [97H3].

Explosive flux compression is also possible, but less frequently used in semiconductor optics, because every shot is connected with the loss of the sample.

Electric fields are most conveniently applied by growing the quantum structure in the intrinsic region of a pin diode, which is then operated in the blocking direction.

Since some semiconductor materials cannot be made n- and p-type, or only with considerable effort, like many of the II–VI compounds, one can also apply electric fields via metal contacts on high resistivity semiconductur structures. See e.g. [89W1].

Finally it should be mentioned that high electric fields can result in strained superlattices or mutliple quantum well structures from the piezoeffect. The piezoeffect appears if a sample is deformed which has at least partly ionic binding and no center of inversion. The electric fields in quantum well structures tend to be higher for materials with the hexagonal wurtzite type structure than for the cubic zincblende type structure. It tends also to be higher for the more ionic bound II–VI compounds compared to the III–V materials. In hexagonal CdSe/CdS and $Ga_{1-y}In_yN/GaN$ superlattices piezofields up to $10^{7...8}$ V/m have been detected [95L3].

25.5.2 Experimental Techniques and Results

Measurements under hydrostatic, bi- or uniaxial stress allow to investigate shifts and splittings of energy levels like the conduction or valence band, or – in other words – to determine the corresponding deformation potentials. For the definition of this quantity see e.g. [82L1] and references therein. A splitting of otherwise (not accidentally) degenerate levels can occur only if the symmetry of the system is lowered i.e. not through hydrostatic pressure. A frequently observed example is the splitting of the Γ_8 valence band of zincblende-type compounds under biaxial stress. See Sect. 16.3 or [01L1]

Another frequently investigated effect is the shift (and splitting) of phonon states with strain, resulting in deformation potentials for phonons.

The dominant part of work is devoted to spectroscopy under the influence of external magnetic fields.

The phenomena and material parameters which can be observed or determined are manifold, and we list a selection of the most prominent ones below. Since there are ample examples and references in Sect. 16.1 we give only very few others here:

- the diamagnetic shift, i.e. a shift quadratic in the magnetic field to higher energies, allows to determine the extension of the wavefunction of the species under investigation,e.g. an exciton, in the plane normal to the magnetic field. In the usual Faraday configuration this is the exciton Bohr radius in the quantum well. For very high magnetic fields this quantity starts to decrease with increasing field strength with consequences e.g. on scattering processes or localization.
- The Zeeman-splitting allows to determine the g-factors of various excited species like excitons. Since it is an off-diagonal element in a matrix representation it mixes also the states and introduces some additional oscillator strength to otherwise forbidden states. Note that a finite zero-field splitting can quench the Zeeman splitting for small fields.

A further aspect is the avoided crossing of a $\boldsymbol{B}$-field-induced fan of levels. This fan is especially obvious for Landau levels into which the higher exciton states and the continuum states develop, if the product of cyclotron frequency ω_c and the dephasing time T_2 fulfill

$$\omega_\mathrm{c} T_2 > 1 \tag{25.39}$$

The observation of the Landau levels e.g. in the interband spectroscopy, or of the cyclotron resonance i.e. the transition between adjacent Landau levels in one band, allows to determine the effective masses, in simplest case via

$$\omega_\mathrm{c} = \frac{eB}{m_\mathrm{eff}} \tag{25.40}$$

The selection rules are for simple parabolic bands complementary i.e. interband transitions are allowed between Landau levels with equal quantization number for electrons and holes

$$n_\mathrm{LLVB} = n_\mathrm{LLCB} \tag{25.41a}$$

while within one series the relations

$$\Delta n_\mathrm{LLVB} = \pm 1 \quad \text{or} \quad \Delta n_\mathrm{LLCB} = \pm 1 \tag{25.41b}$$

hold.

The cyclotron absorption can be observed depending on the material parameters and the magnetic field in EPR (electron paramagnetic resonance)

like experiments, in Fourier IR-spectroscopy, in electronic Raman scattering or in THz spectroscopy. For examples see [94S1].

The g-factor of carriers or of excitons can be also determined from spin-flip Raman scattering or from spin quantum beats or the Hanle effect. The g-factors of defects and doping centres are generally deduced in EPR.

Semimagnetic semiconductors (also known as diluted magnetic semiconductors), i.e. structures containing e.g. manganese or iron ions, can have extremely high g-factors [94G1] and it is even possible to tune the band structure by a magnetic field [96Y2]. Furthermore such structures allow to investigate magnetic polarons, i.e. carriers with a spin polarization cloud. See e.g. [96Y2] Sects. 10.2 or 16.1.2.

Further important experiments, which concern the integer and fractional quantum Hall-effect and its interpretation, are not subject of this subvolume but can be found in the Landolt–Börnstein subvolume on transport in quantum structures [01L2].

The optical spectroscopy of semiconductor quantum structures under electric fields is mainly centered around the following phenomena. See Sect. 16.2

- the quantum-confined Stark effect (QCSE). This effect includes shifts and transfer of oscillator strength of various exciton levels caused by an electric field perpendicular to the wells (or wires). While the Stark effect of excitons is frequently obscured in bulk material by field- or impact ionisation, it can be nicely observed in quantum structures as long as the field is applied normal to the confining potential.
- the observation of Bloch oscillations in the minibands of superlattices and of the Wannier–Stark ladder connected with it.
- tunneling, ballistic or diffusive transport through quantum structures. Though this is predominantly a topic treated in [01L2], there are also several examples that transport can be detected optically.
- electroluminescence. This term is not very precisely determined. In a closer sense it means excitation of some luminescent center by impact excitation or ionization through free carriers accelerated in an applied electric field. This effect is less frequently studied in quantum structures. For an example see [97A1]. In a wider sense, electroluminescence includes all types of light emission caused by the application of a voltage and includes then also luminescence- and laser diodes, which form the main part of applications of semiconductor quantum structures in optics. For nano-optoelectronics see e.g. [92E1, 02N1]
- piezo-superlattices or multiple quantum well have a build-in electric field, resulting from strained layers. This field tilts the bands and shifts the transition energies to the red similarly as in nipi-structures or the exciton ground state in the quantum confined Stark effect. The main tasks are here to determine the field strength, and the band-alignment without field. Under high excitation, the piezo-field can be screened by the spatial separation of carriers. For examples see [95L3].

In some selected cases one applies simultaneously two or three of the external fields discussed here. Examples are given in [01L1].

25.6 Problems

1. What is the blaze wavelength of a grating? How does it work? Why are the efficiencies for + and − first order different even for normal incidence?
2. Which type of lens (plane-convex, bi-convex) would you choose to focus a parallel beam? And which to yield a 1:1 imaging?
3. Is it best to place a crystal polarizer (like a Glan-Thompson) in a parallel, a convergent, or a divergent beam or in the focus?
4. Discuss the advantages and disadvantages of a spectrometer employing a grating or a prism as dispersive element.
5. Inform yourself about the advantages of modulation spectroscopy.
6. What happens if you place an optical component with some coatings or some adhesive layers at the focal point of a high-power laser?
7. What do you see when you look into a phase-conjugating mirror? Solve this problem by thinking, *do not* do it!
8. Have a good idea, which you like yourself, for an experiment on semiconductor optics. Convince some funding agency that your idea has great prospects, both in basic and applied research even if you are yourself rather sceptic about the second aspect, and that it belongs to the best 10% in this universe and this century, and so acquire some 300 kilo units of Euro, US$ or another comparable currency to buy your equipment, pay overheads, hire coworkers etc. Set up your own experiments. If and when it works, please send me some comments or addenda for this or any other chapter based on your own experiences with Murphy's Laws [86M1].

References to Chap. 25

[31G1] M. Göppert-Mayer, Ann. Phys. **9**, 273 (1931)
[57K1] R. Kubo, J. Phys. Soc. Japan, **12**, 570 (1957)
[59M1] P.C. Martin and J. Schwinger, Phys. Rev., **115**, 1342 (1959)
[65B1] N. Bloembergen, Nonlinear Optics, Benjamin, New York (1965)
[69C1] M. Cardona, Modulation Spectroscopy, Academic Press, New York (1969)
[71K1] C. Klingshirn, Z. Phys. **248**, 433 (1971)
[72B1] R.J. Bell, Introductory Fourier-Transform Spectroscopy, Academic Press, New York (1972)
[73B1] J. Bille, Festkörperprobleme / Advances in Solid State Physics **13**, 111 (1973)
[73B2] E. Beckmann, I. Broser, R. Broser, in Luminescence of Crystals, Molecules and Solutions, F. Williams (ed.), Plenum Press, New York 155 (1973)
[73S1] B.O. Seraphin, Modulation Spectroscopy, North-Holland, Amsterdam (1973)

[75K1] C. Klingshirn, phys. stat. sol. (b) **71**, 547 (1975)
[76R1] W. Richter, Resonant Raman Scattering in Semiconductors, Springer Tracts Mod. Phys., Springer, Heidelberg (1976)
[77A1] R.M.A. Azzam, N.M. Bashara, Ellipsometry and Polarized Light, North-Holland, Amsterdam (1977)
[77M1] R.S. Markiewicz, J.P. Wolfe and C.D. Jeffries, Phys. Rev. B **15**, 1988 (1977)
[78H1] J.M. Hvam, J. Appl. Phys. **49**, 3124 (1978)
[78S1] J.R. Salcedom et al.: Phys. Rev. Lett. **41**, 131 (1978)
[78Y1] T. Yajima and H. Souma, Phys. Rev. A **17**, 309 (1978)
[78Y2] T. Yajima, H. Souma and Y. Ishida, Phys. Rev. A **17**, 324 (1978)
[79S1] H. Schrey, V. Lyssenko, and C. Klingshirn, Sol. State Commun. **32**, 897 (1979)
[80A1] V.F. Agekyan and A.A. Berezhnaya, Sov. Phys. Semicond., **14**, 1002 (1980)
[80M1] M.G. Moharam, T.K. Gaylord and R. Magnusson, Opt. Commun. **32**, 19 (1980)
[81B1] K. Bohnert et al.: Z. Phys. B **42**, 1 (1981)
[81B2] G. Blattner et al., phys. stat. sol. (b) **107**, 105 (1981)
[81C1] W.F. Croydon, and E.H.C. Parker, Dielectric Films on Gallium Arsenide, Gordon and Breach, New York (1981)
[81J1] W.B. Jackson et al.: Appl. Opt. **20**, 1333 (1981)
[81K1] G. Kido et al.: J. Phys. E: Sci. Instrum. **14**, 349 (1981)
[81K2] C. Klingshirn and H. Haug, Phys. Rep. **70**, 315 (1981)
[82H1] M.C. Hutley, Diffraction Gratings, Academic Press, London (1982)
[82I1] M. Itoh and I. Ogura, J. Appl. Phys. **53**, 5140 (1982)
[82L1] Landolt-Börnstein, New Series, Group III, Vol. **17a**, O. Madelung, ed. Springer, Berlin (1982), or **41a** (2001)
[82S1] D.G. Seiler et al., Phys. Rev. B **25**, 7666 (1982)
[83F1] R. Fork et al.: IEEE J. Quantum Electron. **19**, 500 (1983)
[83M1] M.L. Meade, Lock-in Amplifiers: Principles and Applications, IEE electrical measurement series, Peregrinus, London (1983)
[83M2] Y. Masumoto, S. Shionoya and T. Takagahra, Phys. Rev. Lett. **51**, 923 (1983)
[83S1] D.G. Seiler et al.: Phys. Rev. B **27**, 2355 (1983)
[84K1] K. Kempf and C. Klingshirn, Sol. State Commun. **49**, 23 (1984)
[84S1] Y.R. Shen, The Principles of Nonlinear Optics, Wiley, New York (1984)
[85H1] B. Hönerlage et al.: Phys. Rep. **124**, 161 (1985)
[85M1] F.A. Majumder et al.: Phys. Rev. B **32**, 2407 (1985)
[85T1] T. Takagahra Phys. Rev. B **31**, 8171 (1985)
[86E1] H.J. Eichler, P. Günter and D.W. Pohl, Laser-Induced Dynamic Gratings, Springer Ser. Opt. Sci. **50**, Springer, Berlin (1986)
[86L1] Laser Spectroscopy of Solids, W.M. Yen and P.M. Selzer, eds., Top. Appl. Phys., Vol. **49**, Springer, Berlin (1986)
[86M1] Murphy‘s Laws, Books I to III, A. Bloch ed., Price/Stern/Sloan, Los Angeles (1986)
[86W1] M. Wegener et al., Semicond. Science and Technol. **1**, 366 (1986)
[87B1] I. Bar-Joseph et al, Appl. Phys. Lett. **50**, 1010 (1987)
[87H1] C. Hirlimann, Rev. Phys. Appl. **22**, 1673 (1987)
[87T1] W.J. Tomlinson and W.H. Knox, JOSA B **4**, 1404 (1987)
[88H1] G. Hernandez, Fabry-Perot Interferometers, Cambridge Studies in Mod. Optics 3, Cambridge (1988)

[88J1] A. Juhl and D. Bimberg, J. Appl. Phys. **64**, 303 (1988)
[88K1] W.H. Knox, IEEE J. Quantum Electron. **24**, 388 (1988)
[88R1] R. Renner et al., J. Cryst. Growth **86**, 581 (1988)
[88S1] H.E. Swoboda et al., Z. Phys. **70**, 341 (1988)
[88U1] I.N. Uraltsev et al., phys. stat. sol. (b) **150**, 673 (1988)
[88W1] Ch. Weber et al., Z. Phys. B **72**, 379 (1988)
[88Z1] J.E. Zucker,A. Pinczuk and D.S. Chemla, Phys. Rev. B **38**, 4287 (1988)
[89C1] J. Christen and D. Bimberg, Rev. Phys. Appl. Colloq. C6, Supplément au no **6**, C6–85 (1989)
[89D1] C. Dörnfeld and J.M. Hvam, IEEE J. Quantum Electron. **25**, 904 (1989)
[89L1] Laser Spectroscopy of Solids II, W.M. Yen ed., Top. Appl. Phys., Vol. **65**, Springer, Berlin (1989)
[89S1] S. Shionoya et al., eds., Springer Proc. Phys. **38** (1989)
[89S2] Spectroscopy of Semiconductors, G. Fasol, A. Fasolino and P. Lugli, eds., NATO ASI Ser. B **206**, Plenum Press, New York (1989)
[89S3] M. Schiekamp et al., Phys. Rev. D **40**, 3077 (1989)
[89W1] Ch. Weber et al., Appl. Phys. Lett. **54**, 2432 (1989)
[90K1] W.H. Knox et al., Optics and Photonics News, April Issue p. 44 (1990)
[90L1] G. Livescu et al.: J. Electron. Mater. **19**, 937.
[90L2] J.D. Lambkin et al., Appl. Phys. Lett. **57**, 1986 (1990)
[90O1] M. Olszakier et al., J. Lumin. **45**, 186 (1990)
[90W1] T. Wilson, ed., Confocal Microscopy, Academic Press, New York (1990)
[91B1] D. Bimberg, T. Wolf, and J. Böhrer, NATO ASI Ser. B **249**, 529, Plenum Press, New York (1991)
[91D1] W. Demtröder, Laserspektroskopie, Springer, Berlin (1991)
[91H1] M.A. Hermann, D. Bimberg and J. Christen, J. Appl. Phys. **70**, R1 (1991)
[91M1] D.L. Mills, Nonlinear Optics, Springer, Berlin (1991)
[91T1] M.L. Thoma, Ch. Weber and C., Klingshirn, Appl. Phys. A **52**, 255 (1991)
[92B1] G. Bongionvanni, A. Mura and J.L. Staehli, phys. stat. sol. (b) **173**, 355 (1992)
[92B2] O. Brandt et al., Phys. Rev. B **45**, 3803 (1992)
[92D1] B. Di Bartolo, ed., Optical Properties of Excited States in Solids, NATO ASI Ser. B **301**, Plenum Press, New York (1992)
[92E1] K.J. Ebeling, Integrierte Optoelektronik, 2nd ed., an English version is available, Springer, Berlin (1992)
[92F1] B.D. Fluegel et al., Sol. State Commun. **83**, 17 (1992)
[92G1] E.O. Göbel et al., phys. stat. sol. (b) **173**, 21 (1992)
[92J1] D.S. Jiang et al., Superlattices Microstruct. **12**, 273 (1992)
[92M1] R.M. Macfarlane, NATO ASI Ser. B **301**, 399, Plenum Press, New York (1992)
[92O1] D. Oberhauser et al., phys. stat. sol. (b) **173**, 53 (1992)
[92S1] G. Sucha, S.R. Bolton and D.S. Chemla, IEEE J. Quantum Electron. **28**, 2163 (1992)
[92S2] H. Schwab et al., phys. stat. sol. (b) **172**, 479 (1992)
[93B1] D. Bimberg and J. Christen, Inst. Phys. Conf. Ser. No **134**, 629 (1993)
[93H1] H. Haug and S.W. Koch, Quantum Theory of the Optical and Electronic Properties of Semiconductors, 2nd ed., World Scientific, Singapore (1993)
[93H2] C.I. Harris et al., J. Phys. IV **3** C5, 171 (1993)
[93K1] W. Kaiser ed., Ultrashort Laserpulses, 2nd.ed., Top. Appl. Phys. **60**, Springer, Berlin (1993)

[93K2] H. Kalt et al.. Physica B **191**, 90 (1993)
[93M1] K. Meissner et al., Phys. Rev. B **48**, 15472 (1993)
[93O1] Optics of Semiconductor Nanostructures, F. Henneberger, S. Schmitt-Rink and E.O. Göbel, eds., Akademie Verlag, Berlin (1993)
[93P1] S. Permogorov et al., Solid State Commun. **88**, 705 (1993)
[93S1] W. Schoenlein et al., Phys. Rev. Lett. **70**, 1014 (1993)
[93S2] R. De Salvo et al., Opt. Lett. **18**, 193 (1993)
[93T1] R. Trebino and D.J. Kane, J. Opt. Soc. Am. A **10**, 1101 (1993)
[93V1] M. Vening, D.J. Dunstan and K.P. Homewood, Phys. Rev. B **48**, 2412 (1993)
[93W1] S. Weiss et al., Appl. Phys. Lett. **63**, 2567 (1993)
[94B1] J.-Y. Bigot et al., in Coherent Optical Interactions in Semiconductors, R.T. Phillips,ed., Plenum Press, New York, p. 245 (1994)
[94D1] E. Dünschede et al., Philos. Mag. B **70**, 443 (1994)
[94F1] F. Fuchs et al., Supperlattices Microstruct. **16**, 35 (1994)
[94G1] J.A. Gaj et al., Phys. Rev. B **50**, 5512 (1994)
[94H1] W. Hackenberg et al., Semicond. Sci. Technol. **9**, 1042 (1994)
[94H2] W. Hackenberg and H.P. Hughes, Semicond. Sci. Technol. **9**, 686 (1994)
[94H3] W. Hackenberg et al., Phys. Rev. B **50**, 10598 (1994)
[94L1] Z.-X. Liu et al., Solid State Electron. **37**, 885 (1994)
[94L2] G.-H. Li et al., Phys. Rev. B **50**, 1575 (1994)
[94M1] D.M. Mittlemen et al., Phys. Rev. B **49**, 14435 (1994)
[94P1] F.H. Pollok, Modulation Spectroscopy of Semiconductors and Semiconductor Microstructures; Handbook of Semiconductors, Elsevier, Amsterdm (1994)
[94S1] D. Some and A.V. Nurmikko, Appl. Phys. Lett. **65**, 3377 (1994)
[94S2] H. Stolz, Time-resolved Light-Scattering from Excitons, Springer Tracts Mod. Phys., Vol. **130**, Springer, Berlin (1994)
[95B1] D. Botkin et al., Rev. Sci. Instrum. **66**, 4130 (1995)
[95B2] G. Bongiovanni, J. Butty and J-L. Staehli, Opt. Eng. **34**, 1941 (1995)
[95B3] L. Bányai et al., Phys. Rev. Lett. **75**, 2188 (1995)
[95D1] E. Daub and P. Würfel, Phys. Rev. Lett. **74**, 1020 (1995)
[95G1] Gerthsen, Physik, 18th ed., Springer, Berlin (1995), presently available 22nd ed. (2003)
[95K1] T.D. Krauss et al., Optics Lett. **20**, 1110 (1995)
[95L1] D.J. Lovering and H.P. Hughes, 22nd Int. Conf. Phys. Semicond., Vancouver, Lockwood, D.J. (ed.), World Scientific, Singapore, Vol. **1**, 209 (1995)
[95L2] V. Langer, H. Stolz and W. von der Osten, Phys. Rev. B **51**, 2103 (1995)
[95L3] W. Langbein, H. Hetterich and C. Klingshirn, Phys. Rev. B **51**, 9922 (1995)
[95M1] G. Mohs et al., Appl. Phys. Lett. **67**, 1515 (1995)
[95P1] D.W. Pohl, Springer Ser. Surf. Sci., 2nd ed., **28**, 232 (1995)
[95U1] M. Umlauff et al., Phys. Rev. B **52**, 5063 (1995)
[95V1] H. Vogelsang et al., 22nd Int. Conf. Phys. Semicond. Vancouver, Lockwood, D.J. (ed.), World Scientific, Singapore, Vol. **2**, 1260 (1995)
[95W1] Th. Weber et al., Semicond. Sci. Technol. **10**, 1113 (1995)
[95W2] S. Weiss et al., phys. stat. sol. (b) **188**, 343 (1995)
[95W3] P. Würfel,S. Finkbeiner and E, Daub, Appl. Phys. A **60**, 67 (1995)
[96A1] T.F. Albrecht et al., Phys. Rev. B **54**, 4436 (1996)
[96B1] D. Botkin et al., Appl. Phys. Lett. **69**, 1321 (1996)

[96B2] K. Bott et al., J. Opt. Soc. Am. B **13**, 1026 (1996)
[96B3] J.J. Baumberg et al., J. Opt. Soc. Am. B **13**, 1246 (1996)
[96B4] D. Birkedal et al., Phys. Rev. B **54**, R14250 (1996)
[96E1] S.A. Empedocles, D.J. Norris and M.G. Bawandi, Phys. Rev. Lett. **77**, 3873 (1996)
[96H1] T. Ha, et al., Proc. Natl. Acad. Sci. USA, Biophysics **93**, 6264 (1996)
[96H2] M. Hübner et al., Phys. Rev. Lett. **76**, 4199 (1996)
[96H3] M. Hübner M. et al., 23rd Int. Conf. Phys. Semicond., Berlin; Scheffler and R. Zimmermann eds., World Scientific, Singapore, **1**, 769 (1996)
[96L1] Ch. Lienau et al., Appl. Phys. Lett. **69**, 2471 (1996)
[96L2] Ch. Lienau et al., Appl. Phys. Lett. **69**, 325 (1996)
[96M1] R.H. Mauch and H.-E. Gumlich eds., Inorganic and Organic Electroluminescence, Wiss.- und Technik-Verlag, Berlin (1996)
[96M2] Merle d'Aubigné et al., Phys. Rev. B **54**, 14003 (1996)
[96O1] M. Oestreich, S. Hallstein and W.W. Rühle, IEEE J. Selected Topics Quantum Electron. **2**, 747 (1996)
[96O2] M. Oestreich et al., Phys. Rev. B **53**, 7911 (1996)
[96S1] J. Shah, Ultrafast Spectroscopy of Semiconductors and Semiconductor Nanostructures, Springer Ser. Solid-State Sci. **115**, Springer, Berlin (1996)
[96W1] M.U. Wehner, D. Steinbach and M. Wegener, Phys. Rev. B **54**, R5211 (1996)
[96Y1] P.Y. Yu and M. Cardona, Fundamentals of Semiconductors, Springer, Berlin (1996)
[96Y2] D.R. Yakovlev and K.V. Kavokin, Comments Condens. Matter Phys. **18**, 51 (1996)
[97A1] M.V. Artemyev, V. Sperling and U. Woggon, J. Appl. Phys. **81**, 6975 (1997)
[97B1] G. Behme et al., Rev. Sci. Instrum. **68**, 3458 (1997)
[97G1] M. Göppert et al., J. Lumin. **72–74**, 430 (1997)
[97G2] E.N. Glezer, NATO ASI Ser. B **356**, 375, Plenum Press, New York (1997)
[97H1] J. Hebeling et al., Optics Commun. **141**, 229 (1997)
[97H2] J.S. Hwang et al., J. Appl. Phys. **82**, 3888 (1997)
[97H3] High Magnetic Fields in the Physics of Semiconductors I and II, G. Landwehr and W. Ossau eds., World Scientific, Singapore (1997)
[97K1] C. Klingshirn, phys. stat. sol. (b) **202**, 857 (1997)
[97L1] W. Langbein et al., phys. stat. sol. (a) **164**, 541 (1997)
[97L2] Ch. Lienau, A. Richter and J.W. Tomm, Appl. Phys. A **64**, 341 (1997)
[97M1] E. Mazur, NATO ASI Ser. B **356**, 4174, Plenum Press, New York (1997)
[97M2] O.D. Mücke et al., phys. stat. sol. (b) **204**, 556 (1997)
[97N1] S. Nakamura and G. Fasol, G., The Blue Laser Diode, Springer, Berlin (1997)
[97N2] M. Nuss and J. Bowers eds., OSA Trends in Optics and Photonics **13**, X, 302 (1997)
[97R1] A. Richter et al., phys. stat. sol. (b) **204**, 247 (1997)
[97R2] A. Richter et al., Surf. Interf. Anal. **25**, 583 (1997)
[97W1] R. Westphäling et al., J. Lumin. **72-74**, 980 (1997)
[97W2] U. Woggon, Optical Properties of Semiconductor Quantum Dots, Springer Tracts Mod. Phys., Vol. **136**, Springer, Berlin (1997)
[97W3] M.U. Wehner, M.H. Ulm and M. Wegener, M., Opt. Lett. **22**, 1455 (1997)

[97Z1] R. Zimmermann, NATO ASI Ser. B **356**, 123, Plenum Press, New York (1997)
[98B1] R. v. Baltz, NATO ASI Ser. B **372**, 323, Plenum Press, New York (1998)
[98C1] T.F. Crimmins, R.M. Koehl and K.A. Nelson, NATO ASI Ser. B **372**, 407, Plenum Press, New York (1998)
[98D1] E. Dekel et al., Physica E **2**, 777 (1998)
[98D2] E. Dekel et al., Phys. Rev. Lett. **80**, 4991 (1998)
[98F1] A.I. Ferguson, NATO ASI Ser. B **372**, 233, Plenum Press, New York (1998)
[98F2] G. v. Freymann, Th. Schimmel and M. Wegener, Appl. Phys. Lett. **73**, 1170 (1998)
[98G1] L.E. Golub et al., phys. stat. sol. (b) **205**, 203 (1998)
[98I1] E.P. Ippen, NATO ASI Ser. B **372**, 213, Plenum Press, New York (1998)
[98J1] A. Jolk and C. Klingshirn, phys. stat. sol. (b) **206**, 841 (1998)
[98J2] M. Jutte, W. von der Osten and H. Stolz, Opt. Commun. **157**, 173 (1998)
[98K1] H. Kuzmany, Solid State Spectroscopy, Springer, Berlin (1998) and more recent editions
[98K2] H. Kalt et al., J. Cryst. Growth **184/185**, 627 (1998)
[98K3] C. Klingshirn, NATO ASI Ser. B **372**, 143, Plenum Press, New York (1998)
[98K4] H. Kalt et al., Acta Phys. Polon. A **94**, 139 (1998)
[98K5] D.S. Kim et al., Phys. Rev. Lett. **80**, 4803 (1998)
[98K6] A. Knorr et al., phys. stat. sol. (b) **206**, 139 (1998)
[98K7] R.A. Kaindl et al., Optics Lett. **23**, 861 (1998)
[98L1] S.G. Lipson, H. Lipson and D.S. Tannhauser, Optical Physics, Cambridge University Press, Cambridge (1998)
[98L2] Ch. Lienau et al., Phys. Rev. B **58**, 2045 (1998)
[98L3] S. Linden, H. Giessen and J. Kuhl, phys. stat. sol. (b) **206**, 119 (1998)
[98P1] W. Petri et al., J. Crystl. Growth **184/185**, 320 (1998)
[98S1] I. Shtrickman et al., Physica E **2**, 65 (1998)
[98S2] D. Steinbach, W. Hügel and M. Wegener, JOSA B **15**, 1231 (1998)
[98U1] M. Umlauff et al., Phys. Rev. B **57**, 1390 (1998)
[98V1] A. Vertikov et al., Appl. Phys. Lett. **72**, 2645 (1998)
[98W1] R. Westphäling et al., J. Appl. Phys. **84**, 6871 (1998)
[99A1] Ch. Adelmann et al., Appl. Phys. Lett. **74**, 179 (1999)
[99B1] S.I. Bozhevolnyi, X. Mufei and J.M. Hvam, J. Opt. Soc. Am. A **16**, 2649 (1999)
[99C1] A. Cavalleri et al., J. Appl. Phys. **85**, 3301 (1999)
[99D1] A. Dinger et al., phys. stat. sol. (b) **215**, 413 (1999)
[99E1] A. Euteneuer et al., Phys. Rev. Lett. **83**, 2073 (1999)
[99H1] E. Hecht, Optik, Oldenbourg, München (1999)
[99H2] S.G. Hense and M. Wegener, Appl. Phys. Lett. **74**, 920 (1999)
[99K1] J.C. Kim et al., Appl. Phys. Lett. **75**, 214 (1999)
[99K2] A. Klochkhin et al., Phys. Rev. B **59**, 12947 (1999)
[99L1] Ch. Lienau et al., Physica B **272**, 96 (1999)
[99L2] W. Langbein, J.M. Hvam and R. Zimmermann, Phys. Rev. Lett. **82**, 1040 (1999)
[99L3] S. Linden, J. Kuhl and H. Giessen, Opt. Lett. **24**, 569 (1999)
[99L4] D. Lüerssen et al., Appl. Phys. Lett. **75**, 3944 (1999)
[99L5] D. Lüerssen et al., Phys. Rev. B **59**, 15862 (1999)
[99R1] A. Richter et al., J. Microsc. **194**, 393 (1999)

[99S1] V.F. Sapega et al., phys. stat. sol. (b) **215**, 379 (1999)
[99V1] M. Vollmer et al., Appl. Phys. Lett. **74**, 1791 (1999)
[99V2] M. Vollmer et al., J. Microsc. **194**, 523 (1999)
[99W1] Q. Wu et al., Phys. Rev. Lett. **83**, 2562 (1999)
[99W2] M. Wegener et al., Festkörperprobleme / Advances in Solid State Physcis **38**, 297 (1999)
[00D1] A. Dinger et al., J. Cryst. Growth **214/215**, 847 (2000)
[00D2] A. Dinger et al., phys. stat. sol. (b) **221**, 485 (2000)
[00D3] A. Dinger et al., J. Cryst. Growth **214/215**, 676 (2000)
[00F1] G. v. Freymann et al., Appl. Phys. Lett. **76**, 203 (2000)
[00G1] M. Göppert et al., J. Cryst. Growth **214/215**, 625 (2000)
[00H1] R. Huber et al., Appl. Phys. Lett. **76**, 3191 (2000)
[00H2] J. Hetzler et al., phys. stat. sol. (b) **221**, 425 (2000)
[00K1] C. Klingshirn, Physik in unserer Zeit **31**, 144 (2000)
[00L1] I.H. Libon et al., Appl. Phys. Lett. **76**, 2821 (2000)
[01A1] Advances in Energy Transfer Processes, B. Di Bartolo ed., World Scientific, Hongkong (2001)
[01K1] S. Kawata, M. Ohtsu and M. Irie, Nano Optics, Springer Series in Optical Sciences **84**, Springer, Berlin (2001)
[01K2] T.A. Klar, E. Engel and S.W. Hell, Phys. Rev. E **64**, 066613 (2001)
[01K3] C. Klingshirn, in ref.[01A1] p165
[01L1] Landolt-Börnstein, New Series, Group III, Vol. **34** C1 and 2, C. Klingshirn (ed.), Springer, Berlin (2001) and (2004)
[01L2] Landolt-Börnstein, New Series, Group III, Vol. **34** B1 and 2, B. Kramer (ed.), Springer, Berlin (2001) and in preparation
[01N1] Y. Nakamura et al., Phys. Rev. B **64**, 075203 (2001)
[02J1] A. Jolk, M. Jörger and C. Klingshirn, Phys. Rev. B **65**, 245209 (2002)
[02K1] C. Klingshirn, T. Fleck and M. Jörger, phys. stat. sol. (b) **234**, 23 (2002)
[02M1] O. Mücke et al., Optics Lett. **27**, 2127 (2002) and Phys. Rev. Lett. **89**, 127401 (2002)
[02N1] Nano-Optoelectronics, M. Grundmann ed., Springer, Berlin, (2002)
[02S1] W. Seitz et al., Opt. Soc. America, TOPS **72**, 66 and 326 (2002)
[02Z1] H. Zhao et al., Appl. Phys. Lett. **80**, 1391 (2002)
[03C1] A. Christ et al., Phys. Rev. Lett. **91**, 183901 (2003) and ibib. **81**, 2794 (2002)
[03D1] B. Dal Don et al., phys. stat. sol. c **0** (4), 1237 (2003)
[03F1] T. Fleck, M. Schmidt and C. Klingshirn, phys. stat. sol. (a) **198**, 248 (2003)
[03H1] S.W. Hell, Nature Biotechnology **21**, 1347 (2003)
[03K1] C. Klingshirn, T. Fleck and M. Jörger, phys. stat. sol. (b) **239**, 261 (2003)
[03M1] S. Moehl et al., J. Appl. Phys. **93**, 6265 (2003)
[03N1] N. Naka and N. Nagasawa, phys. stat. sol. (b) **238**, 397 (2003)
[03S1] Spectroscopy of Systems with Spatially Confined Structures, B. Di Bartolo (ed.), NATO Science Series II, **90**, Kluwer, Dordrecht, (2003)
[03S2] E. Sorokin et al., Appl. Phys. B, **77**, 245 (2003)
[03T1] J. Teipel et al., Appl. Phys. B, **77**, 245 (2003) and CLEO, Baltimore, **CMO3** (2003)
[03T2] T. Tritschler et al., SPIE Proc. **4992**, 33 (2003) and phys. stat. sol. (b) **238**, 561 (2003)
[03Z1] H. Zhao et al., Phys. Rev. B **67**, 035306 (2003) and phys. stat. sol. b **238**, 529 (2003)

[04C1] S. Chatterjee et al., Phys. Rev. Lett. **92**, 067402 (2004)
[04K1] J. Kvietkova et al., Thin Solid Films **455/456**, 228 (2004) and Phys. Rev. B**70**, 045316 (2004)
[04K2] M. Kuwata-Gonokami et al., J. Phys. Soc. Japan **73**, 1065 (2004)
[04S1] M. Schubert, Infrared Ellipsometry on Semiconductor Layer Structures: Phonons, Plasmons and Polaritons, Springer Tracts in Modern Physics, in press (2004)
[05F1] Frontiers of Optical Spectroscopy, B. Di Bartolo ed., Kluwer, Dordrecht, in press (2005)

26

Group Theory in Semiconductor Optics

By K. Hümmer and C. Klingshirn

In this chapter, we shall give an outline of group theory, its connection to quantum mechanics, and its applications in semiconductor optics. We shall present only the most important aspects and rules and generally give no proofs of the various relations.

The information compiled here has been taken from different books on group theory and on crystallography, which are listed at the end this Chapter, namely [60H1,64K1,64T1,67S1,74B1,88L1,90I1,92I1,94H1,96Y1,01L1]. Very good collections of tables found in the books by Koster et al. and by W. Ludwig et al. [63K1, 88L1], from which most of the Tables 26.6 – 26.15 are taken.

26.1 Introductory Remarks

Noether's theorem of Sect. 3.1.3, which we have already used several times in this book, states that a conservation law follows from every invariance of the Hamiltonian. The classic examples are the following.

- If H is invariant under an infinitesmally shift in time, then the total energy of the system described by this Hamiltonian is conserved.
- If H does not depend on a particular spatial coordinate x_i, i.e., if it is invariant under an infinitesimal translation along x_i, then the component of the momentum p_i in the direction x_i is conserved.
- If H has spherical symmetry (as for a single atom, for example) it is invariant with respect to an infinitesimal rotation $\mathbf{d}\boldsymbol{\phi}$ around any axis, and a conservation law for the angular momentum follows for L^2 and for one component of $\boldsymbol{L}$, e.g., L_z.

In a crystal the first condition remains valid in the absence of explicitly time-dependent (perturbative) terms, i.e., energy is still conserved.

The second condition is no longer valid, but is replaced by invariance with respect to translations by integer multiples of the $\boldsymbol{a}_i$ (Sect. 7.2). We have already presented this subject and its consequences in some detail, e.g.,

in Sect. 3.1.3 and Sects. 7.1–5, 8.1, but we shall return to it briefly in this chapter, too.

The third condition concerning infinitesimal rotations likewise no longer applies to crystals. Instead we have, depending on the crystal structure, only invariance with respect to rotations around 2-, 3-, 4-, and 6-fold symmetry axes. This means that strictly speaking $\boldsymbol{L}$ is not a good quantum number in crystals, but later, in connection with compatibility relations, we shall see that arguments based on $\boldsymbol{L}^2$ and/or L_z can still sometimes be used, if with considerable care. In the following sections we give a short introduction to group theory (Sect. 26.2) and to representations and characters (Sect. 26.3). Then we present the crucial part, namely the connection between the Hamiltonian and group theory, in Sect. 26.4. Finally we give applications both of general nature and others which are more or less specific to semiconductors Sects. 26.5, 26.6. The aim of this chapter is not to give the reader a deep insight into group theory and its implications but rather a feeling and understanding of what the γ_i mean that are met in bandstructures as in Sect. 8.6, and to enable him or her to calculate simple selection rules.

26.2 Some Aspects of Abstract Group Theory for Crystals

A group G is a set of elements $\{X_i\}$ and a connection or composition between the elements which is often called "multiplication" with the following properties:

- Closure: if X_i and X_j are elements of G ($X_i, X_j \in G$) then the "product" is element of G.
$$X_i X_j = X_k \in G\,. \tag{26.1}$$
- Associative law:
$$X_i\,(X_j X_k) = (X_i X_j)\,X_k\,. \tag{26.2}$$
- There is a neutral element or identity operation or ∼mapping E in G with
$$EX_i = X_i E = X_i\,. \quad \text{for} \quad X_i \in G\,. \tag{26.3}$$
- There is an inverse element X_i^{-1} for each $X_i \in G$ with
$$X_i^{-1} X_i = X_i X_i^{-1} = E\,. \tag{26.4}$$

The number g of different elements in a group gives the order of the group. There are finite and infinite groups. If the commutative law holds, i.e., if

$$X_i X_j = X_j X_i \quad \text{for all} \quad X_i X_j \in G \tag{26.5}$$

the group is said to be Abelian.

26.2.1 Some Abstract Definitions

A group is called cyclic if it can be generated by one of its elements $G = \{X, XX = X^2, XXX = X^3, \ldots\}$. If one of these elements is the unit element, $X^k = E$, then G is finite and cyclic of order k.

A subset H of a group G is called a subgroup of G if H obeys the group axioms. The whole group G and E are called trivial or improper subgroups; the others are called proper subgroups and one writes $H < G$.

A set of elements from which the complete group can be obtained by composition, is called a set of generators of G.

A cyclic group obviously has exactly one such generator.

26.2.2 Classification of the Group Elements

A distribution of the elements into subsets such that each element belongs to exactly one subset, i.e., two different subsets have no element in common, is called a classification. There are two important methods to classify the elements of a group: (1) coset decomposition and (2) partition into conjugacy classes.

1. A group can be decomposed into left or right cosets relative to a subgroup, say $H = \{Y\}$, by multiplying each $X_i \in G$ with the elements of H from left $X_i\{Y\}$ or right $\{Y\}X_i$, respectively. Exactly one of these cosets is a group, namely the subgroup H.
2. An element X_j is called conjugate to X_i, if an element $X_k \in G$ exists for which the following holds:

$$X_k^{-1} X_i X_k = X_j \,.$$

   If X_j is conjugated to X_i then X_i is also conjugated to X_j with the element $X_k^{-1} \in G$, which must exist. The subset of the elements of G, which are conjugate to X_i, X_k running through G, is the conjugacy class of X_i. The number of elements in a conjugacy class is called its length l, which may be different for different conjugacy classes, but is always a factor of g. Elements of the same conjugacy class have the same order with respect to their cyclic group.

An element for which

$$X_k^{-1} X_i X_k = X_i \tag{26.6}$$

holds for all $X_k \in G$ is called self-conjugate in G. Therefore, in Abelian groups each $X_i \in G$ forms a conjugacy class by itself of length $l = 1$.

Conjugacy is also defined for subgroups. Of special interest are self-conjugate subgroups, also called normal or invariant subgroups for which holds: $X_k^{-1}\{Y\}X_k = \{Y\}$ for all $X_k \in G$ where $\{Y\} = H$ is a proper subgroup of $G : H < G$. Then, $\{Y\}X_k = X_k\{Y\}$, i.e. for a normal subgroup the right and left coset decomposition coincide.

The cosets of G relative to a normal subgroup H are elements of a group, called the factor group F, sometimes denoted as $F = G/H$. The unit element of F is the normal subgroup H.

26.2.3 Isomorphism and Homomorphism of Groups

Isomorphism is a special case of the more general concept of homomorphism. An isomorphism is a one-to-one correspondence between the elements of groups, which is preserved during multiplication. Two groups G and G' with elements X_i and X_i' are isomorphic if (after suitable ordering of their elements) there is a reversible mapping of G onto G': $X_i \leftrightarrow X_i'$, and for each pair X_i, $X_j \in G$ the product of the images $X_i'\ X_j' \in G'$ is equal to the image of the product: $X_i' X_j' = (X_i X_j)'$.

The isomorphism yields a classification of all possible groups into isomorphism classes, which is sometimes called an abstract group. The abstract group displays the common group theoretical features of isomorphic groups independent of the realization of their elements.

In contrast, a homomorphism maps G onto a possibly smaller group G' and is then an irreversible, many-to-one correspondence between G and G', which is preserved during multiplication; more formally: $X_i' X_j' = (X_i X_j)'$.

26.2.4 Some Examples of Groups

1. the positive and negative integers including zero $\{0, \pm 1, \pm 2\}$ with normal addition as connection. They form an infinite Abelian group.
2. the positive rational numbers p/q with $p = 0, 1, 2, 3, \ldots$ and $q = 1, 2, 3, \ldots$ with normal multiplication as connection. They again form an infinite Abelian group.
3. Symmetry operations of any physical object, e.g., atoms, molecules or crystalline solids, by definition map object onto itself. Symmetry operations are geometric mappings that leave all distances invariant, so-called isometries or rigid motions. They form a group, called the symmetry group of the object. In any symmetry group the unit element is the identity mapping that leaves every point fixed.
4. In the following example we shall use a model in order to explain the definitions of abstract group theory given above. The symmetry operations that map an equilateral triangle onto itself form a group (Fig. 26.1). The symmetry operations are: E, the unit element, which simply leaves the triangle unchanged; J and K are rotations of $+120°$ and $240° = -120°$, respectively, around the axis through the center of the triangle and normal to it; L, M, and N are rotations of $180°$ around the axis a, b and c. In the latter case the upper and lower faces of the triangle must be indistinguishable. Alternatively, there can be mirror planes normal to the triangle that contain the a-, b- or c-axis. The symmetry elements at which the rotations are carried out are called either a 3-fold or 2-fold axis as the

cyclic group of J is of order 3: $JJJ = J^3 = E$, $J^2 = J^{-1} = K$. That of L is of order 2: $LL = L^2 = E$, $L = L^{-1}$. In crystallographic groups they are denoted by 3 and 2, a mirror plane by m. In 3-dimensional space these are the symmetry groups of a trigonal trapezohedron, denoted as $D_3 = 32$ (Schönflies and Hermann-Mauguin notation) and of a trigonal pyramid, denoted as $C_{3\mathrm{v}} = 3m$.

The composition (multiplication) of the symmetry operations is to perform one after the other in a well-defined order. A convenient way to display a finite group (of small order) is the multiplication or group table (Table 26.1). It is a square array, where the product X_iX_j is listed at the intersection of the row of operation X_i and the column of operation X_j. X_iX_j means first apply X_j and then X_i on a point. In every row and in every column of the group table, every operation of the group appears once and only once.

The group table of D_3 and $C_{3\mathrm{v}}$, as well as the group of all permutations of three different symbols, say (abc), can be made identical apart from the names or symbols of the operations: $E = (abc)$, $J = (cab)$, $K = (bca)$, $L = (acb)$, $M = (cba)$, $N = (bac)$. There exists a one-to-one correspondence between the operations of the three groups: they are isomorphic. They are of order $g = 6$.

The fact that the group table is not symmetric with respect to the main diagonal shows that the group is non-Abelian. In fact, it is known as the non-Abelian group of lowest order.

The following proper subgroups can be deduced as can be seen from an inspection of the group table:

$$\{E, J, K\},\ g = 3; \quad \{E, L\}, \{E, M\}, \{E, N\},\ g = 2\,. \tag{26.7}$$

Table 26.1. Multiplication table of group D_3. We give the names of the elements and the operations corresponding to Fig. 26.1

	E	J	K	L	M	N
		+120	−120	180_a	180_b	180_c
E	E	J	K	L	M	N
		+120	−120	180_a	180_b	180_c
J	J	K	E	N	L	M
+ 120	+120	−120		180_c	180_a	180_b
K	K	E	J	M	N	L
−120	−120		+120	180_b	180_c	180_a
L	L	M	N	E	J	K
180_a	180_a	180_b	180_c		+120°	−120
M	M	N	L	K	E	J
180_b	180_b	180_c	180_a	−120°		+120
N	N	L	M	J	K	E
180_c	180_c	180_a	180_b	+120°	−120°	

The coset decomposition relative to the subgroup $\{E, J, K\}$ gives one more coset $\{L, M, N\}$, i.e. left and right cosets are identical, and $\{E, J, K\}$ is a normal or invariant subgroup. The coset decomposition relative to the subgroup $\{E, L\}$ gives two additional left cosets $\{J, N\}$, $\{K, M\}$ and two more right cosets $\{J, M\}$, $\{K, N\}$ as can be seen from an inspection of the group table.

In the abstract group mentioned above there are three conjugacy classes, or in short simply C_i classes, as can be verified using the group table:

$$C_1 = \{E\}, l = 1;\ C_2 = \{J, K\}, l = 2;\ C_3 = \{L, M, N\}, l = 3\,. \quad (26.8a)$$

For example, the operations J and K belong to the same class since

$$N^{-1}JN = NM = K \quad \text{and with} \quad N^{-1} = N\,. \quad (26.8b)$$

Geometrically, conjugacy means the search for symmetrically equivalent symmetry elements. They must be of the same type, or from the point of view of group theory, their cyclic groups must be of the same order. In fact, the two 3-fold axes J and K are symmetrically equivalent through the three 2-fold axes or mirror planes L, M, N and vice versa. The 2-fold axes or mirror planes are symmetrically equivalent through the 3-fold axes.

5. The set of all symmetry operations (isometries) of an idealized infinite crystal structure forms a crystallographic space group. In three-dimensional space we distinguish 230 types of crystallographic space groups of order $g = \infty$. The most characteristic feature of a crystal is that its structure can be embedded into a lattice. The existence of a lattice, i.e. the presence of three-dimensional periodicity, gives rise to diffraction effects, as first shown by Laue, Friedrich and Knipping. In terms of group theory this corresponds to the existence of a discrete translation subgroup T of every space group G. The symmetry operations of T are translations by lattice vectors $\boldsymbol{R}_n$. Choosing a suitable basis of lattice vectors, a so-called primitive basis, each lattice vector is represented by a linear combination of the basis vectors with integer coefficients (see also Sect. 7.2):

$$\boldsymbol{R}_n = \sum_i n_i \boldsymbol{a}_i\,. \quad (26.9)$$

Isometries are special affine mappings described by a system of equations

$$\widetilde{\boldsymbol{x}} = \boldsymbol{D}\boldsymbol{x} + \boldsymbol{R} = (\boldsymbol{D}, \boldsymbol{R})\boldsymbol{x}\,. \quad (26.10)$$

The pair $(\boldsymbol{D}, \boldsymbol{R})$ consists of the (3×3)-matrix $\boldsymbol{D}$, the rotational part, and the (3×1)-column $\boldsymbol{R}$, the translational part. For isometries, i.e., distance, preserving mappings, the condition $det\, \boldsymbol{D} = \pm 1$ must be obeyed.

The lattice of translation vectors of a crystal imposes restrictions on the isometries occurring in space groups. The rotational matrix $\boldsymbol{D}$ has to obey

$\boldsymbol{D}^k = \boldsymbol{E}$, with the order $k = 2, 3, 4, 6$. That means a crystal only allows rotations by multiples of 180, 120, 90, and 60°.
Therefore, in crystallographic space groups the following symmetry operations occur. Their symbols are given in the Hermann-Mauguin notation.

a) Identity $(\boldsymbol{E}, \boldsymbol{0})$. Symbol 1: $\boldsymbol{E}$ is the unit matrix.
b) Translations by lattice vectors $(\boldsymbol{E}, \boldsymbol{R}_n)$.
c) Rotations and screw rotations $det\, \boldsymbol{D} = +1$
   i. Rotations. Symbols $k = 2, 3, 4, 6$ according to their order k; for each one exists a line of (invariant) fixed points, the rotation axis.
   ii. Screw rotations. Symbols k_n with $n < k$: $(\boldsymbol{D}, \boldsymbol{R})^k = (\boldsymbol{E}, \boldsymbol{R}_n)$ must hold where $\boldsymbol{R}_n$ is a lattice translation; $(\boldsymbol{D}, \boldsymbol{R})$ can be decomposed into a rotation around an axis (screw axis) and a translation by the screw vector parallel to this axis $\boldsymbol{R}_\parallel = \frac{n}{k}\boldsymbol{R}_0$, where $\boldsymbol{R}_0$ is the shortest lattice vector along the screw axis.
d) Inversion $\boldsymbol{D} = -\boldsymbol{E}$. Symbol $\bar{1}$: exactly one point is fixed, the center of inversion or center of symmetry.
e) Rotoinversions $det\, \boldsymbol{D} = -1$. Symbols $\bar{k}, k \neq 2$: it may be decomposed into a rotation followed by an inversion.
f) Reflection and glide reflection $det\, \boldsymbol{D} = -1$, $\boldsymbol{D}^2 = \boldsymbol{E}$:
   i. Reflection $(\boldsymbol{D}, \boldsymbol{R})^2 = (\boldsymbol{E}, \boldsymbol{0})$. Symbol m: the fixed points form a plane at which reflection of space is performed, the so-called mirror plane. It is a special rotoinversion $\bar{2} = m$ where the axis stands normal to the mirror plane.
   ii. Glide reflection $(\boldsymbol{D}, \boldsymbol{R})^2 = (\boldsymbol{E}, \boldsymbol{R}_n)$. Symbols a, b, c, n, d: may be decomposed into a reflection through a plane (glide plane) and a translation by the glide vector parallel to this plane.

As already mentioned above, the subset of isometries $\{(\boldsymbol{E}, \boldsymbol{R}_n)\}$, i.e., translations by lattice vectors, forms a subgroup T of the space group G. It is normal or invariant since the left and right coset decomposition coincide. The symmetry operations in each coset relative to T have the rotational part $\boldsymbol{D}$ in common since the product of any space group operation $(\boldsymbol{D}, \boldsymbol{R})$ and any lattice translations $(\boldsymbol{E}, \boldsymbol{R}_n)$ does not change the rotational part $\boldsymbol{D}$. The factor group $F = G/T$ where the cosets are the elements of F, is isomorphic to the point group P of the space group G. This point group P of the crystal can thus be obtained directly from the coset decomposition relative to T by considering only the rotational parts given by (3×3)-matrices.
To know the point group P or the crystal class of the space group G of the crystal structure is important with regard to, e.g., the symmetry of the form of the crystal, the symmetry of the lattice and thus of the reciprocal lattice and the symmetry of a diffraction pattern.
There is one further important aspect of this coset decomposition, namely the denotation of space groups by means of the Hermann-Mauguin symbols used in crystallography. The first character is the letter of the type of

the Bravais lattice of the structure given by the translation subgroup T. P means primitive; A, B and C mean one face is centered; I means body centered; F means all faces are centered; R means special centering type in trigonal crystal classes. Then follows a finite set of symbols (maximum three) of generators that when combined with $\{T\}$ permits the generation of all space group operations.

The following are some examples of space groups for important types of semiconductor structures.

1. Wurtzite structure. Prototype: ZnO; other examples CdS, CdSe, GaN.

$$G\colon\ P6_3mc;\ P\colon 6mm = C_{6\mathrm{v}} \quad (\text{Schönflies notation})\,.$$

The point group symbol in the Hermann-Mauguin notation is simply derived by replacing the screw axes by their rotation axes and the glide planes by their mirror planes.
$6mm$ belongs to the hexagonal crystal system. It is of order $g = 12$ with the following symmetry operations. Rotation axis 6 contains the operations 6^+ = rotation by 60°, 6^- = rotation by −60°, 3^+ = rotation by 120°, 3^- = rotation by −120°, and 2 = rotation by 180°. There are two triples of mirror planes. In each triple the symmetrically equivalent mirror planes enclose an angle of 60°. However, the mirror planes of different triples are rotated by 30° against each other. Thus, they are not symmetrically equivalent (there is no symmetry operation rotation by 30°) and they belong to different conjugacy classes, which is why the Hermann-Mauguin symbol $6mm$ has two sets of non-equivalent mirror planes indicated. Therefore, in the single group there are six (conjugacy) classes denoted in the Schönflies notation as (Table 26.6): $E, 2C_6, 2C_3, C_2, 3\sigma_{\mathrm{d}}, 3\sigma_{\mathrm{v}}$.
2. Zincblende structure. Prototype: ZnS, other examples III–V-compounds like GaAs, InP, or I–VII compounds like the Cu-halides or AgBr.

$$G:\ F\bar{4}3m;\ P:\ \bar{4}3m = T_{\mathrm{d}}\,.$$

$\bar{4}3m$ belongs to the cubic crystal system. It is of the order $g = 24$. It is non-centrosymmetric. The classes of the single group contain (Table 26.10), apart from the identity class, the operations 3^+ and 3^- of four diagonal 3-fold axes ($8C_3$) of a cube, 2-fold operations rotation by 180° of three axes $2 = \bar{4}^2(3C_2)$, the operations $\bar{4}^1 = \bar{4}^+$ and $\bar{4}^3 = \bar{4}^-(6S_4)$ of the three $\bar{4}$ rotoinversion axes along the axes of a cube and 2-fold operations at six diagonal mirror planes ($6\sigma_{\mathrm{d}}$)
3. Diamond structure Prototype: C, other examples Ge, Si, Cu_2O, NaCl

$$G\colon\ Fd\bar{3}m;\ P\colon\ m\bar{3}m = O_h$$

$m\bar{3}m$ is centro-symmetric and belongs to the cubic crystal system. It is of order $g = 48$. It can be considered as the direct product of $\bar{4}3m = T_{\mathrm{d}}$ and

the point group $\bar{1} = C_i$. Therefore, we also have the classes (Table 26.14) C_i, $6C_4$ (since $\bar{4} \times \bar{1} = 4$), $8C_{3i}$ (since $3 \times \bar{1} = \bar{3} = C_{3i}$) and $3\sigma_\mathrm{h}$ (since $2 \times \bar{1} = m$).

The outer or direct product of two groups $G = \{X_i\}$ and $G' = \{X'_i\}$ is again a group G'' which consists of all ordered pairs X_i, X'_j. The definition of the connection of the elements of G'' is

$$(X_i, X'_j)(X_k, X'_l) = (X_i X_k, X'_j X'_l)\,, \tag{26.11}$$

and the order of G'' is $g \cdot g'$.

26.3 Theory of Representations and of Characters

Any set of elements which fulfils the multiplication table of a group is called a representation of this group. The operations in Fig. 26.1 can be considered as a representation of D_3.

In a stricter sense, to which we shall stick from now on, we consider a representation to be a set of $(n \times n)$ matrices $(n = 1, 2, 3, \ldots)$ which fulfils the multiplication table with normal matrix multiplication as the connection.

The name of a certain representation is given by a label α used as an index.

With $\boldsymbol{\Gamma}_\alpha$ we denote a set of g matrices that fulfil the multiplication table of group G. $\boldsymbol{\Gamma}_\alpha(X_i)$ is the matrix that represents the element X_i of G and $\boldsymbol{\Gamma}_\alpha(X_i)_{jk}$ is the element of this matrix in row j and column k.

The normal matrix multiplication then results in

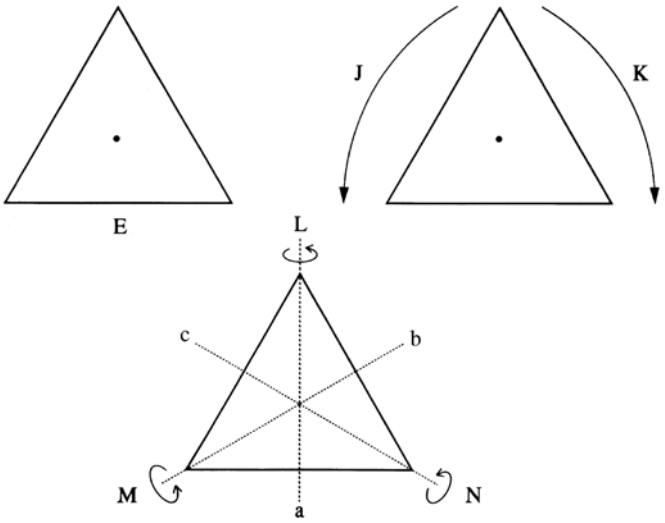


Fig. 26.1. The symmetry operations which map an equilateral triangle into itself

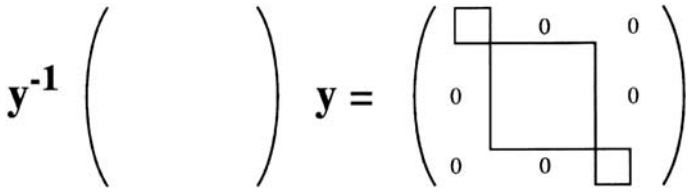


Fig. 26.2. The reduction of a reducible representation into irreducible ones

$$\boldsymbol{\Gamma}(X_i, X_j)_{kl} = \sum_h \boldsymbol{\Gamma}_\alpha(X_i)_{k,h} \boldsymbol{\Gamma}_\alpha(X_j)_{h,l} \,. \tag{26.12}$$

We introduce n_α to denote the dimension of the $(n_\alpha \times n_\alpha)$ matrices of a certain representation α; n_α is the same for all matrices of one representation $\boldsymbol{\Gamma}_\alpha$.

There is an infinite number of representations for one group. If $\boldsymbol{\Gamma}_\alpha$ is a representation and $\boldsymbol{X}$ is a non-singular matrix of the same dimension n_α then $\{\boldsymbol{X}^{-1}\boldsymbol{\Gamma}_\alpha\boldsymbol{X}\}$ is again a representation.

In the following we shall destill from this infinite number of representations some with special properties (the so-called irreducible representations) and present some of their characteristic properties.

Table 26.2 gives four examples of representations of D_3, two one-dimensional ones, a two- and a three-dimensional one. A representation $\boldsymbol{\Gamma}_\alpha$ is called reducible if a non-singular matrix $\boldsymbol{Y}$ can be found such that all matrices $\boldsymbol{Y}^{-1}\boldsymbol{\Gamma}_\alpha(\boldsymbol{X}_i)\boldsymbol{Y}$ acquire the same block form. This means there are square sub-matrices and otherwise zeros; see Fig. 26.2. From the properties of matrix multiplication it follows immediately that only submatrices at identical positions are connected with each other in multiplication. This means that the submatrices of all X_i which are located in the same position, alone form a representation.

A representation $\boldsymbol{\Gamma}_\alpha$ which can be brought into such a block form of submatrices is called reducible. A representation $\boldsymbol{\Gamma}_\alpha$, for which no matrix $\boldsymbol{Y}$ can be found to bring all $\boldsymbol{\Gamma}_\alpha(X_i)$ into the same block form is called irreducible. The representations $\boldsymbol{\Gamma}_1 - \boldsymbol{\Gamma}_3$ in Table 26.2 are irreducible; $\boldsymbol{\Gamma}_4$ is evidently reducible. It is already expressed in the desired block form.

We concentrate in the following on irreducible representations.

Two irreducible representations $\boldsymbol{\Gamma}_\alpha$, and $\boldsymbol{\Gamma}_\beta$ are called equivalent, if there exists a non singular matrix $\boldsymbol{S}$ with

$$\boldsymbol{S}^{-1}\boldsymbol{\Gamma}_\alpha(\boldsymbol{X}_i)\boldsymbol{S} = \boldsymbol{\Gamma}_\beta(\boldsymbol{X}_i) \quad \text{for all} \quad \boldsymbol{X}_i \in \boldsymbol{G} \tag{26.13}$$

With this definition of equivalence we arrive at a very important statement: There exist exactly as many non equivalent (= different) irreducible representations of a group as the group has classes C_i.

For D_3 we found three classes (see above). Consequently D_3 has three nonequivalent irreducible representations. These are the $\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2$ and $\boldsymbol{\Gamma}_3$ in

Table 26.2. Four matrix representations of the group D_3

	Elements of the group					
	E	J	K	L	M	N
$\boldsymbol{\Gamma}_1$	1	1	1	1	1	1
$\boldsymbol{\Gamma}_2$	1	1	1	-1	-1	-1
$\boldsymbol{\Gamma}_3$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
"$\boldsymbol{\Gamma}_4$"	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$

Table 26.2. The irreducible representation in which all elements are represented by 1 is called the trivial representation and is always called $\boldsymbol{\Gamma}_1$. The representations of the identity element are always unit matrices with 1 on the main diagonal and zeros otherwise. For irreducible representations $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$ there exists an orthogonality relation

$$\sum_{X_i} \boldsymbol{\Gamma}_\alpha(X_i)_{kp} \boldsymbol{\Gamma}_\beta(X_i^{-1})_{ql} = \frac{g}{n_\alpha} \delta_{\alpha\beta} \delta_{kl} \delta_{pq} \tag{26.14}$$

with δ_{kl} and δ_{pq} being the Kroneckers symbol and

$$\delta_{\alpha\beta} = \begin{cases} 0 & \text{if } \boldsymbol{\Gamma}_\alpha \text{ and } \boldsymbol{\Gamma}_\beta \text{ are not equivalent} \\ 1 & \text{if } \boldsymbol{\Gamma}_\alpha \text{ and } \boldsymbol{\Gamma}_\beta \text{ are identical} \\ \neq 0 & \text{but undefined if } \boldsymbol{\Gamma}_\alpha \text{ and } \boldsymbol{\Gamma}_\beta \text{ are equivalent} \end{cases}$$

A characteristic quantity which is identical for all equivalent irreducible representations is the trace of the matrices

$$\chi_\alpha(X_i) = Tr\, \boldsymbol{\Gamma}_\alpha(X_i) = \sum_j \boldsymbol{\Gamma}_\alpha(X_i)_{jj} \,. \tag{26.15}$$

These traces are called characters. Two equivalent representations have the same characters since

$$\mathrm{Tr} \boldsymbol{\Gamma}_\alpha(X_i) = Tr\, \boldsymbol{S}^{-1} \boldsymbol{\Gamma}_\alpha(X_i) \boldsymbol{S} \,. \tag{26.16}$$

We can thus write down the characters of all nonequivalent irreducible representations in a table. Since the characters of conjugate elements are the same, as seen from a comparison of (26.6) and (26.16), it is sufficient to give the characters for the different classes. This results in a square scheme of numbers. Table 26.3 gives as a first example the character table for the group D_3. It is identical, by the way, to that of the group $C_{3\mathrm{v}}$ mentioned above.

As already mentioned, $\boldsymbol{\Gamma}_1$, in the first line is always the trivial representation, while the first column gives the dimensionality n_α of the representation $\boldsymbol{\Gamma}_\alpha$.

The nomenclature $\chi_\alpha(C_i)$ means the character of the irreducible representation a for the class C_i.

Introducing the number h_i of elements of the class C_i, with a total number r of classes and of irreducible representations in the group, we find the following relations for the irreducible representations

$$\sum_{i=1}^{r} h_i \chi_\alpha(C_i) \chi_\beta^*(C_i) = g \delta_{\alpha\beta} \,, \tag{26.17a}$$

$$\sum_{\alpha=1}^{r} h_i \chi_\alpha(C_i) \chi_\alpha^*(C_j) = g \delta_{ij} \,. \tag{26.17b}$$

Two representations $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$ are equivalent if $\chi_\alpha(X_i) = \chi_\beta(X_i)$ for all X_i. If a representation $\boldsymbol{\Gamma}$ is given, it contains the irreducible representation $\boldsymbol{\Gamma}_\alpha$ p_α times in the sense of a decomposition into block form:

$$\boldsymbol{\Gamma} = \sum_\alpha p_\alpha \boldsymbol{\Gamma}_\alpha \quad \text{with} \quad p_\alpha = \sum_{X_i} \chi(X_i)\chi^*_\alpha(X_i)\,. \tag{26.18}$$

This means that it is not necessary in practice to try out all possible non singular matrices since one can immediately deduce from (26.18) the decomposition of a given representation into irreducible ones. The reducible representation "$\boldsymbol{\Gamma}_4$" in Table 26.2 can consequently be decomposed into

$$\boldsymbol{\Gamma}_4 = \boldsymbol{\Gamma}_2 \oplus \boldsymbol{\Gamma}_3\,. \tag{26.19}$$

This process also explains the use of the symbol $\oplus$.

The direct product $\boldsymbol{\Gamma}_\alpha \otimes \boldsymbol{\Gamma}_\beta$ of two representations $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$ with dimensions n_α and n_β, respectively, is a new representation of dimension $n_\alpha \cdot n_\beta$. The new set of matrices is obtained in the following way

$$\boldsymbol{\Gamma}_\alpha(X_i) \otimes \boldsymbol{\Gamma}_\beta(X_i) = \begin{pmatrix} \boldsymbol{\Gamma}_\alpha(X_i)_{11}\boldsymbol{\Gamma}_\beta(X_i) & \dots & \boldsymbol{\Gamma}_\alpha(X_i)_{1n_\alpha}\boldsymbol{\Gamma}_\beta(X_i) \\ & \vdots & \\ \boldsymbol{\Gamma}_\alpha(X_i)_{n_\alpha 1}\boldsymbol{\Gamma}_\beta(X_i) & \dots & \boldsymbol{\Gamma}_\alpha(X_i)_{n_\alpha n_\alpha}\boldsymbol{\Gamma}_\beta(X_i) \end{pmatrix} \tag{26.20}$$

The direct product of two representations is again a representation of the group. The direct product of two irreducible representations can be reducible or irreducible.

For the characters of the direct product one finds

$$\chi\left(\boldsymbol{\Gamma}_\alpha \otimes \boldsymbol{\Gamma}_\beta\right) = \chi_\alpha \cdot \chi_\beta\,. \tag{26.21}$$

The decomposition of the products of the irreducible representations of a group is usually given in tables:

$$\boldsymbol{\Gamma}_\alpha \otimes \boldsymbol{\Gamma}_\beta = \sum_\gamma g_{\alpha\beta\gamma}\boldsymbol{\Gamma}_\gamma \tag{26.22}$$

with

Table 26.3. The character table of the group D_3

Irreducible representations	Classes of elements			
	D_3	E	$2C_3$	$3C_2$
	$\boldsymbol{\Gamma}_1$	1	1	1
	$\boldsymbol{\Gamma}_2$	1	1	-1
	$\boldsymbol{\Gamma}_3$	2	-1	0

Table 26.4. Multiplication table for the irreducible representations of the group D_3

D_3	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_3$
$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_3$
$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_3$
$\boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_1+\boldsymbol{\Gamma}_2+\boldsymbol{\Gamma}_3$

$$g_{\alpha\beta\gamma} = \frac{1}{g}\sum_{X_i} \chi_\alpha(X_i)\chi_\beta(X_i)\chi^*_\gamma(X_i)\,. \tag{26.23}$$

We give the direct products of the irreducible representations of D_3 in Table 26.4.

The criterion for a given representation to be irreducible is

$$\boldsymbol{\Gamma}_\alpha \quad \text{irred} \;\Leftrightarrow \sum_{X_i} |\chi_\alpha(X_i)|^2 = g\,. \tag{26.24}$$

A representation of a group is also a representation of each of its subgroups. An irreducible representation of a group can be a reducible or irreducible representation of the subgroup. Which possibility applies can either be checked with formula given above or it can be found in the tables. In Table 26.5 we give these so-called compatibility relations of the irreducible representations of D_3 with the irreducible representations of its two proper subgroups C_2 and C_3.

26.4 Hamilton Operator and Group Theory

In this section we derive the connections between the eigenfunctions and eigenstates of the stationary or time-independent Hamilton operator H and the irreducible representations of a group.

We do not yet specify a particular type of group, but notice that an operator $P(X_i)$ which transforms the system into itself commutes with the Hamilton operator

$$P(X_i)H = HP(X_i)\,, \tag{26.25}$$

where X_i is a symmetry operation of the group. It is known that whenever an operator P commutes with H, there exists a set of functions which are simultaneously eigenfunctions of H and of P. These eigenfunctions are called basis functions. They are given in the tables of Sect. 25.6. We also come back to them in Sect. 25.5.

The group which transforms the system into itself is also known as the group of H.

We consider now an eigenvalue E_α of H with eigenfunction(s) $\psi_{\alpha j}$; $j = 1,\ldots,m$. If the eigenvalue E_α is not degenerate one has $m = 1$. For a degenerate eigenvalue $m > 1$.

Table 26.5. Compatibility table of the irreducible representations of the group D_3 with the irreducible representations of its proper subgroups C_2 and C_3

D_3	Γ_1	Γ_2	Γ_3
C_3	Γ_1	Γ_1	$\Gamma_2 + \Gamma_3$
C_2	Γ_1	Γ_2	$\Gamma_1 + \Gamma_2$

With

$$H\psi_{\alpha j} = E_\alpha \psi_{\alpha j} \tag{26.26}$$

we also have

$$P(X_i)H\psi_{\alpha j} = HP(X_i)\psi_{\alpha j} = P(X_i)E_\alpha \psi_{\alpha j} = E_\alpha P(X_i))\psi_{\alpha j}\,, \tag{26.27}$$

since E_α is just a real number.

This means that together with $\psi_{\alpha j}$ also $P(X_i)\psi_{\alpha j}$ is an eigenfunction with eigenvalue E_α. Since the $\psi_{\alpha j}$ already include all possible eigenfunctions with eigenvalue E_α, the $P(X_i)\psi_{\alpha j}$ can only be linear combinations of the $\psi_{\alpha j}$.

This means

$$P(X_i)\psi_{\alpha k} = \sum_{j=1}^{m} \boldsymbol{\Gamma}_\alpha(X_i)_{jk}\psi_{\alpha j}\,. \tag{26.28}$$

If we perform this procedure for one symmetry operation X_i and for all $\psi_{\alpha k}$ we get a matrix of coefficients $\boldsymbol{\Gamma}_\alpha(X_i)_{jk}$. For a nondegenerate eigenvalue this is just a (complex) number of unit magnitude.

If we then perform the procedure for all symmetry operations X_i of the group of the Hamiltonian we get a set of matrices $\boldsymbol{\Gamma}_\alpha(X_i)$. Now comes the crucial point: this set of matrices forms a representation of the group of H, and generally an irreducible representation.

If this is the case we can identify the eigenvalue E_α which belongs to the eigenstates $\psi_{\alpha j}$ with the label of the irreducible representation $\boldsymbol{\Gamma}_\alpha$, i.e., we say that the eigenvalue E_α has symmetry $\boldsymbol{\Gamma}_\alpha$ or the eigenstates $\psi_{\alpha j}$ of E_α transform according to $\boldsymbol{\Gamma}_\alpha$.

Since most physical systems have an infinite number of eigenstates but finite groups have only a finite number of non-equivalent irreducible representations, the name of one irreducible representation will occur many times as shown schematically in Fig. 26.3.

If the representation of the group produced by the eigenfunctions of one eigenenergy E_α is reducible, we speak of an accidental degeneracy. An example is shown in Fig. 26.3. This means there are actually two different eigenenergies with their sets of eigenfunctions which just happen to be equal for the specific parameters of H, but which could in principle be different.

The important point now is that group theory allows us to make many important predictions, for example, about selection rules or splitting of states under external perturbations, without even knowing the $\boldsymbol{\Gamma}_\alpha(X_i)_{ij}$ nor the eigenvalues or eigenfunctions. There is one drawback, however, group theory

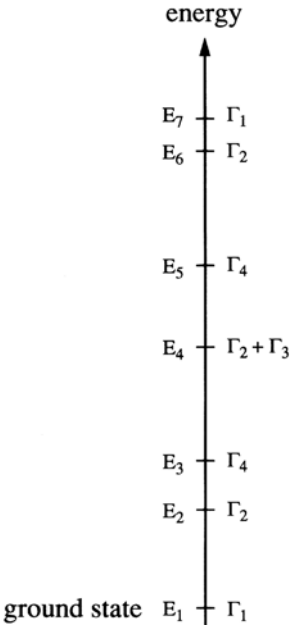


Fig. 26.3. The schematic relation of the eigenenergies E_α, and the corresponding irreducible representation $\boldsymbol{\Gamma}_\alpha$

tells us only whether a certain matrix element is zero or not; it does not say anything about its magnitude if it is non zero. To get quantitative information we must either use physical arguments to deduce how small or large an effect can be, once group theory has told us that it exists, or we just have to calculate the matrix elements. But here group theory helps us again, by telling us which ones are zero.

Until now we have dealt only with real space. For electronic eigenstates we should also consider spin. Spin does not have a representation in real space. Since we do not want to indulge in a lot of mathematics, we simply note that there exists a group $D_{1/2}$ which describes the behavior of particles with spin $1/2$ under point operations (a translation does not affect spin).

To describe electronic wave functions with spin

$$\psi_{\alpha j} = \psi_{\alpha j} \cdot s \,, \tag{26.29}$$

we have to consider the "double group" which is formally just the direct product of $D_{1/2}$ and the corresponding point group $D_{1/2} \otimes$ (point group). This double group is again a group, possessing conjugate elements, classes, irreducible representations, and a character table. We shall finally end up with more classes and more Γ_α compared to the simple point group without spin. In the tables given at the end of this chapter for the three most important point groups of semiconductors one can see from the basis functions which ones are the additional Γ_α and classes. For vibrations (phonons) of molecules and of solids spin is not important. Therefore often only the simple and not the double groups are given in corresponding books.

26.5 Applications to Semiconductors Optics

The two groups that describe crystalline solids are, as already mentioned, the translation group and the space group. The translation group is an infinite Abelian group of all translations $\boldsymbol{R}_n$ as introduced in Sects. 7.2 or 8.1 (also see below).

Additionally, the space group contains all other symmetry operations that map the crystal onto itself like inversion, rotation axes, screw rotation axes, mirror (or reflection) planes or glide reflection planes.

Symmorphic space groups can be decomposed into a semi-direct product of the translation group and a point group that contains symmetry operations that keep at least one point fixed, e.g., inversion, rotation axes or reflection planes.

The generators of symmorphic groups do not contain screw rotation axes or reflection glide planes. Nonsymmorphic groups contain such symmetry operations in their generators and cannot be decomposed in the above way. Details about this topic can be found, e.g., in [94H1].

The translation group is obviously Abelian, i.e.,

$$T(\boldsymbol{R}_1)T(\boldsymbol{R}_2) = T(\boldsymbol{R}_1 + \boldsymbol{R}_2) = T(\boldsymbol{R}_2 + \boldsymbol{R}_1) = T(\boldsymbol{R}_2)T(\boldsymbol{R}_1)\,, \qquad (26.30)$$

where $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are two translations of the lattice and T is the translation operator. Consequently every $\boldsymbol{R}_i$, is self-conjugate and forms a class by itself and all irreducible representations are one-dimensional. A detailed inspection of their properties leads to

$$T(\boldsymbol{R}_i)u(\boldsymbol{r}, k) = \mathrm{e}^{\mathrm{i}k\boldsymbol{R}_i}u(\boldsymbol{r}, \boldsymbol{k}) \qquad (26.31)$$

since the prefactor must be linear in $\boldsymbol{R}_i$ and of the magnitude one. This short argument shows that the Bloch theorem of Sect. 8.1 also follows from group-theoretical considerations.

Since we have already treated this theorem and its consequences in some detail in Chaps. 7 and 8, we need not further pursue it here, but concentrate now on point groups.

We note (again without proof) that according to what we have said above, we can restrict ourselves trivially for symmorphic space groups to the point group.

For nonsymmorphic space groups the following statements are correct for the point group we obtain if we replace screw rotation axes by simple rotation axes and glide reflection planes by simple reflection planes [94H1]. In this sense the point group of the wurtzite crystal structure, which contains a screw rotation axis, is $C_{6\mathrm{v}}$ or $6mm$.

The proper consideration of screw rotation axes or glide reflection planes introduces, e.g., phase factors in X-ray scattering but does not change the statements about matrix elements or selection rules given below.

The point group of a crystal defined in this sense contains all symmetry operations that transform the crystal into itself and keep at least one point fixed. This group also describes the symmetries of the eigenstates for $\boldsymbol{k} = 0$, i.e., at the $\boldsymbol{\Gamma}$ point in reciprocal space that is the center of the first Brillouin zone.

Since this is also the region of $\boldsymbol{k}$-space that is important for the optical properties of most direct gap semiconductors, we stick in the following to $\boldsymbol{k} = 0$ (this is the reason why we use $\boldsymbol{\Gamma}$ to label the irreducible representations) and discuss only later what happens when we go to $\boldsymbol{k} \neq 0$. For the moment we note that a statement which is exact for $\boldsymbol{k} = 0$ will be almost correct in the close vicinity of the $\boldsymbol{\Gamma}$-point.

We now leave the group D_3, which we used above for illustration purposes, and return to the three point groups that are most important for the semiconductors, namely T_d, O_h, and $C_{6\mathrm{v}}$ and which we have introduced already above.

In Sect. 25.6 we give the character tables and the basis functions of these three point groups. The remainder of the 32 point groups can be found in e.g. [63K1, 88L1]. Additionally we give the multiplication tables for the irreducible representations, their compatibility tables with the full rotation group, which is valid for spherical problems like atoms (see below), and with the proper subgroups. There we indicate examples of external perturbations which reduce the symmetry of the point group to that of its subgroups.

The classes with a bar come from the formation of the double group when we include spin and the same is true for the irreducible representations which contain noninteger $\boldsymbol{J}$ and J_z values in their basis functions $\phi(J, J_z)$. Basis functions are examples of functions that possess the symmetry properties of the irreducible representation. Basis functions S_x, S_y and S_z transform like x, y and z, however, without a change of sign under inversion and R means a spherically symmetric function like $x^2 + y^2 + z^2$. A superscript $^+$ or $^-$ appears where inversion C_i is an element of the group and where, consequently, parity is a good quantum number. Obviously $^+$ or $^-$ mean even and odd parity under inversion, respectively.

The information contained in these tables is sufficient to answer many of the questions addressed below.

We present the first important rule. If we have to calculate a transition matrix element M_{if} (see also Sect. 3.2)

$$\boldsymbol{\Gamma}_f \boldsymbol{\Gamma}_s \boldsymbol{\Gamma}_i$$

$$M_{if} \propto \langle f \,|H_s|\, i\rangle = \int \psi_f^* H_\mathrm{s} \psi_i \,\mathrm{d}\tau \,, \tag{26.32}$$

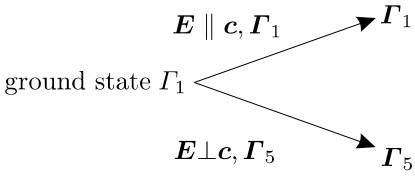
where we start from an initial state $|i\rangle$ of symmetry $\boldsymbol{\Gamma}_i$, and want to reach a final state $\langle f|$ of symmetry $\boldsymbol{\Gamma}_f$ with a perturbation operator H_s that trans-

forms like $\boldsymbol{\Gamma}_s$, then group theory tells us that this matrix element is zero or non-zero according to:

$$\langle f\,|H_\mathrm{s}|\,i\rangle = \begin{cases} \neq 0 & \text{if } \boldsymbol{\Gamma}_f \text{ is contained in the direct product } \boldsymbol{\Gamma}_\mathrm{i} \otimes \boldsymbol{\Gamma}_\mathrm{s} \text{ or} \\ & \text{if } \boldsymbol{\Gamma}_1 \text{ is contained in } \boldsymbol{\Gamma}_f \otimes \boldsymbol{\Gamma}_\mathrm{s} \otimes \boldsymbol{\Gamma}_i \\ & \text{These two statements are equivalent;} \\ 0 & \text{otherwise.} \end{cases} \tag{26.33}$$

This rule enables us, for example, to calculate the optical selection rules in crystals. For full spherical symmetry in atoms these rules read, for electric dipole transitions, $\Delta l = \pm 1$; $\Delta m = 0, \pm 1$. For crystals we have to know the symmetry of H_s. It can be often deduced from inspection of the basis functions. The dipole operator transforms in T_d like $\boldsymbol{\Gamma}_5$, in O_h like $\boldsymbol{\Gamma}_5^-$, and in $C_{6\mathrm{v}}$ like $\boldsymbol{\Gamma}_1$ for $\boldsymbol{E} \parallel \boldsymbol{c}$ and like $\boldsymbol{\Gamma}_5$ for $\boldsymbol{E} \perp \boldsymbol{c}$.

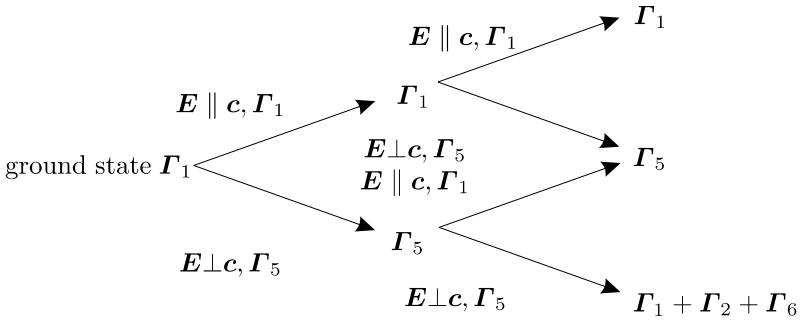
More information about the symmetries of other perturbations like magnetic fields, uniaxial stress etc. is found e.g. in [63K1, 76C1, 77G1, 85H1, 88L1, 97I1] or in the tables of Sect. 26.6. We now use this knowledge to ask which states can be reached by one- and two-photon transitions in $C_{6\mathrm{v}}$ symmetry, starting from the crystal ground state, which has always symmetry $\boldsymbol{\Gamma}_1$. For one-photon transitions, we have to consider matrix elements as in (26.32). Thus for $C_{6\mathrm{v}}$ we have,



For two-photon transitions we have to consider expressions such as (see also Sect. 19.1):

$$M_{if} \propto \sum_{z_j} \frac{\langle f\,|H^{\mathrm{D},2}|\,z_j\rangle\,\langle z_j\,|H^{\mathrm{D},1}|\,i\rangle}{\hbar\omega - (E_{z_j} - E_i)} + c.p. \tag{26.34}$$

and find the following selection rules with the help of the multiplication tables



This means that in two-photon absorption we can reach states with symmetries $\boldsymbol{\Gamma}_1$, $\boldsymbol{\Gamma}_2$, $\boldsymbol{\Gamma}_5$ and $\boldsymbol{\Gamma}_6$ depending on the polarization.

Another field in which the vanishing or otherwise of a matrix element is of importance is perturbation theory. We assume that we have an unperturbed stationary Hamiltonian H with non-degenerate eigenstate E_n^0, ψ_n^0. We now apply a small stationary perturbation H_s to get a total Hamiltonian

$$H = H_0 + H_\mathrm{s}. \tag{26.35a}$$

The perturbed eigenenergy E_n, and wavefunction ψ_n are then given by

$$E_n = E_n^0 + \left\langle \psi_n^0 \left| H_\mathrm{s} \right| \psi_n^0 \right\rangle , \tag{26.35b}$$

and

$$\psi_n = \psi_n^0 + \sum_{k \neq 0} \frac{\left\langle \psi_k^0 \left| H_\mathrm{s} \right| \psi_n^0 \right\rangle \psi_k^0}{E_n^0 - E_k^0} . \tag{26.35c}$$

Equation (26.33) tells us now which perturbations can change the eigenenergy and mix the states and which cannot. As stated earlier, group theory does not tell us how large this effect will be.

In the case of a degenerate level E_n^0 with $\psi_{n,i}^0$, perturbation theory leads to a secular equation given by the determinant

$$\det \left| \left\langle \psi_{n,i}^0 \left| H_\mathrm{s} \right| \psi_{n,j}^0 \right\rangle - E\delta_{ij} \right| = 0 , \tag{26.36}$$

and again (26.33) can predict which matrix elements vanish. The problem of degenerate states has been developed even further with a technique known as "invariant expansion". A discussion of this technique is beyond the scope of this book but the reader may find an elaborate presentation in [76C1, 77G1, 85H1].

Partly equivalent information is obtained if we consider that the perturbation reduces the symmetry of the problem, so that the perturbed problem corresponds to a subgroup of the unperturbed one. If the irreducible representation of E_n^0 is still irreducible in the subgroup, E_n, may shift but does not split. If it decays into more than one irreducible representation in the subgroup, E_n^0 can split into a corresponding number of different levels under the influence of H_s. If some of the new eigenvalues still coincide we have an accidental degeneracy.

To illustrate this fact, inspection of the compatibility relations $C_{6\mathrm{v}}$ shows that an electrostatic field $\boldsymbol{E} \perp \boldsymbol{c}$ ($\boldsymbol{\Gamma}_5$) will result in a splitting of a $\boldsymbol{\Gamma}_5$ level into two states of symmetries $\boldsymbol{\Gamma}_1$ and $\boldsymbol{\Gamma}_2$ of the group C_s.

Another case in which group theory is helpful concerns product wave functions. If, for example, we know that the spatial part of the hole wave function $\phi_h(\boldsymbol{r}_n)$ in T_d transforms like $\boldsymbol{\Gamma}_5$ (this results from an atomic p-state as will be shown below) and we want to add spin σ we get

$$\psi = \phi_n^{\boldsymbol{\Gamma}_5}(\boldsymbol{r}_h) \cdot \boldsymbol{s} . \tag{26.37}$$

In order to discover the symmetry of the total wave function we just form the direct product of the symmetry of the spatial part (here $\boldsymbol{\Gamma}_5$) and of the irreducible representation of spin 1/2. For T_d this reads

$$\boldsymbol{\Gamma}_{\text{total}} = \boldsymbol{\Gamma}_{5,\text{space}} \otimes \boldsymbol{\Gamma}_{1/2} = \boldsymbol{\Gamma}_5 \otimes \boldsymbol{\Gamma}_6 = \boldsymbol{\Gamma}_7 \oplus \boldsymbol{\Gamma}_8 \,. \tag{26.38}$$

This means that the spin results in a splitting into two states of symmetries $\boldsymbol{\Gamma}_7$ and $\boldsymbol{\Gamma}_8$. From physical arguments we know that this is the spin-orbit splitting.

A very important example of product wave functions in semiconductors are exciton wave functions. As shown in Sect. 9.1, they are a product of the electron and hole wavefunctions and of the envelope function.

The possible symmetries of an exciton therefore result from the direct product of the symmetries of the electron, hole and envelope functions, i.e.,

$$\boldsymbol{\Gamma}_{\text{exciton}} = \boldsymbol{\Gamma}_{\text{el}} \otimes \boldsymbol{\Gamma}_{\text{h}} \otimes \boldsymbol{\Gamma}_{\text{env}} \,. \tag{26.39}$$

For the ground state (main quantum number $n_{\text{B}} = 1$) $\boldsymbol{\Gamma}_{env}$ is always $\boldsymbol{\Gamma}_1$. So in T_d, for example, we find excitons formed with the hole in the $\boldsymbol{\Gamma}_7$ or $\boldsymbol{\Gamma}_8$ valance band and the electron in the $\boldsymbol{\Gamma}_6$ conduction band

$$\boldsymbol{\Gamma}_6 \otimes \boldsymbol{\Gamma}_8 \otimes \boldsymbol{\Gamma}_1 = \boldsymbol{\Gamma}_3 \oplus \boldsymbol{\Gamma}_4 \oplus \boldsymbol{\Gamma}_5 \tag{26.40a}$$

$$\boldsymbol{\Gamma}_6 \otimes \boldsymbol{\Gamma}_7 \otimes \boldsymbol{\Gamma}_1 = \boldsymbol{\Gamma}_2 \oplus \boldsymbol{\Gamma}_5 \tag{26.40b}$$

The $\boldsymbol{\Gamma}_5$ is the spin singlet state, which can be reached from the ground state by an electric-dipole transition; the $\boldsymbol{\Gamma}_3$, $\boldsymbol{\Gamma}_4$ and $\boldsymbol{\Gamma}_2$ states are dipole forbidden. Actually they are the triplet states with parallel electron and hole spins.

The number of possible exciton states increases rapidly for $n_B > 1$.

Still another important case is the reduction of symmetry in a quantum well compared to the parent bulk materials.

In Table 26.6 we give the irreducible representations, the compatibility relations and the selection rules when going from a bulk zincblende structure to a QW grown on a (001) surface. For this, but also for other orientations, see, e.g., [96I1, 97I1, 00S1, 02M1]. A single ideal interface between two different materials both having the point group T_d reduces the symmetry to $C_{2\text{v}}$. See [96I1, 97I1] and references therein. A quantum well, i.e., two close-lying interfaces, results in symmetry D_{2d} if the two materials have a common anion or cation, e.g., AB/AC. This situation is treated in Table 26.6. In the case of a well/barrier combination without a common anion or cation AB/CD, the total symmetry may be D_{2d} or C_{2v} depending on the termination of the interfaces. For C_{2v} symmetry, the orthogonal directions [110] and $[1\bar{1}0]$ in the plane of the quantum well are no longer equivalent. This may result in polarization anisotropies of the optical properties or in different band offsets at both interfaces and in internal electric fields if, e.g., the dipole layers on both sides

Table 26.6. Group theoretical relations and selection rules for zincblende-type bulk material and quantum wells grown on (100) oriented substrates [93J1]

Point group (example)	T_d (GaAs, bulk)	D_{2d} GaAs QW on (001) oriented substrate	Comments
Irreducible representation of lowest conduction band	$\boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_6$	
Irreducible representation of highest valence band (does not apply to CuCl or a few other exceptions)	$\boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_6$ $\boldsymbol{\Gamma}_7$	Heavy holes light holes
Excitons with S, i.e., $\boldsymbol{\Gamma}_1$ envelope function	$\boldsymbol{\Gamma}_6^{\mathrm{LB}} \otimes \boldsymbol{\Gamma}_8^{\mathrm{VB}} \otimes \boldsymbol{\Gamma}_1^{\mathrm{env}} = \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_6^{\mathrm{LB}} \otimes \boldsymbol{\Gamma}_6^{\mathrm{VB}} \otimes \boldsymbol{\Gamma}_1^{\mathrm{env}} = \boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_5$	Heavy-hole excitons
		$\boldsymbol{\Gamma}_6^{\mathrm{LB}} \otimes \boldsymbol{\Gamma}_7^{\mathrm{VB}} \otimes \boldsymbol{\Gamma}_1^{\mathrm{env}} = \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5$	Light-hole excitons
Dipole-allowed transitions	$\boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_5$ for $\boldsymbol{E} \perp z$ $\boldsymbol{\Gamma}_4$ for $\boldsymbol{E} \parallel z$	z is the growth direction, i.e., the normal to the QW plane

are different. For more details see, e.g., [97I1, 00S1, 02M1, 03T1] and references therein.

Going one step further, we come to the more complicated entities such as bound-exciton complexes or biexcitons. The procedure is basically as above, i.e.,

$$\boldsymbol{\Gamma}_{\mathrm{biex}} = (\boldsymbol{\Gamma}_{\mathrm{el}} \otimes \boldsymbol{\Gamma}_{\mathrm{el}})^{\pm} \otimes (\boldsymbol{\Gamma}_{\mathrm{h}} \otimes \boldsymbol{\Gamma}_{\mathrm{h}})^{\pm} \boldsymbol{\Gamma}_{\mathrm{env}}^{\mp\mp} . \tag{26.41}$$

However, one now has to consider that these systems contain partly indistinguishable fermions. Therefore the total wave-function must change sign under the exchange of two identical particles. This means that if the combination of the two electrons $(\boldsymbol{\Gamma}_{\mathrm{el}} \otimes \boldsymbol{\Gamma}_{\mathrm{el}})^{-}$ changes sign under exchange, then the envelope function must have even parity under this operation and vice versa. The same holds for the holes. The parity of the combinations can be seen from the coupling coefficients in tables like [63K1]. We do not go into details here but mention that the ground state of the biexciton always has an envelope of $\boldsymbol{\Gamma}_1$ symmetry. The possible combinations for the excitons of (26.40a), (26.40b) are in the ground state $\boldsymbol{\Gamma}_1$ and $\boldsymbol{\Gamma}_1 \oplus \boldsymbol{\Gamma}_3 \oplus \boldsymbol{\Gamma}_5$ [77G1, 85H1], respectively. In $C_{6\mathrm{v}}$ the biexciton containing two holes from the same valence band always has only symmetry $\boldsymbol{\Gamma}_1$ in the ground state, but with a hole from the $\boldsymbol{\Gamma}_9$ and another from a $\boldsymbol{\Gamma}_7$ valence band it is possible to construct in addition $\boldsymbol{\Gamma}_5$ and $\boldsymbol{\Gamma}_6$ biexciton levels [77G1, 85H1].

The bands in semiconductors often still contain some information about the parent atomic orbitals, especially in the lattice periodic part $u_k(\boldsymbol{r})$ of the Bloch function. The atomic orbitals result from a spherically symmetric problem. Therefore it is reasonable to say a few words about this group and its compatibility relations with the point groups of semiconductors. In the full spherical rotation group, the system can be rotated around any axis by any angle and is transformed in to itself. It can be shown that all rotations by the same angle ϕ but around arbitrary axes are in the same class. The full rotation group can therefore be considered as a continuous group (Fig. 26.4) with rotation angle $0 \leq \phi < 2\pi$. Consequently there must be a correspondingly infinite number of irreducible representations $\boldsymbol{\Gamma}_l$. Figure 26.4 gives the character of the representation. For the identity element E we get the dimensionality of $\boldsymbol{\Gamma}_l$ which is obviously $2l+1$. The basis functions are the spherical harmonics $Y_l^m(\phi,\theta)$.

The compatibility relations between the full rotation group for even and odd parity, called $D^{\pm}_{L_z}$ and $D^{\pm}_{J_z}$ and the three point groups T_d, O_h and $C_{6\mathrm{v}}$ are given in Sect. 26.6. These tables may be used as follows: In a tight-binding approximation we would assume that the uppermost valence band of CdS is formed from the filled $3p$ levels of S^{2-}, while the lowest conduction band comes from the empty $5s$ levels of Cd^{2+}. From the compatibility tables we learn that this results in the following band symmetries for $C_{6\mathrm{v}}$, at the $\boldsymbol{\Gamma}$-point:

$$\begin{aligned} &\text{CB}: \quad 5s \text{ levels} \Rightarrow D_{1/2} \Rightarrow \boldsymbol{\Gamma}_7\,; \\ &\text{VB}: \quad 3p \text{ levels} \Rightarrow \begin{cases} D_{3/2} \Rightarrow \boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_9 \\ D_{1/2} \Rightarrow \boldsymbol{\Gamma}_7\,. \end{cases} \end{aligned} \tag{26.42}$$

For T_d symmetry one finds similarly

$$\text{CB}: \ \boldsymbol{\Gamma}_6, \quad \text{VB}: \ \boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_8\,. \tag{26.43}$$

This implies a twofold splitting in the valence band of T_d symmetry, which comes from the spin-orbit interaction as already mentioned, and a threefold splitting in C_{6v} again arising from spin-orbit splitting and, in addition, from the hexagonal crystal field.

Note, that for covalent sp^3 binding the top of the valence band and the bottom of the conduction band also have p- and s-character, thus leading to the same results as for ionic binding.

Group theory cannot tell us the magnitude of the splittings nor the ordering of the bands. It is found, however, that for T_d symmetry the $\boldsymbol{\Gamma}_8$ state usually forms the upper valence band and the $\boldsymbol{\Gamma}_7$ comes below (except in CuCl which has the reverse ordering) and in $C_{6\mathrm{v}}$ one usually has an uppermost valence band of symmetry $\boldsymbol{\Gamma}_9$ and two $\boldsymbol{\Gamma}_7$ levels below (except for ZnO, where the two upper bands are also most probably interchanged). See Sect. 8.8, 13.1, 16.1

It is clear that angular momentum is no longer a good quantum number in solids. The good quantities are the $\boldsymbol{\Gamma}_i$. With great care, however, one can

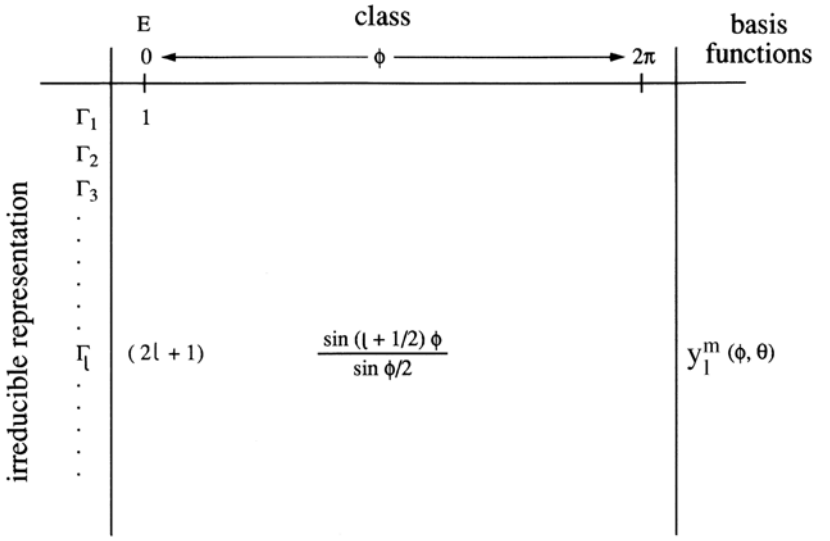


Fig. 26.4. The character table of the spherical rotation group

sometimes still use arguments based on angular momentum. In the wurtzite structure the $\boldsymbol{c}$ axis is the quantization axis and J_z is to some extent usable for discussions. The p-orbitals forming the valence band have $l = 1$ and $m = 0, \pm 1$. The $m = 0$ states have symmetry $\boldsymbol{\Gamma}_1$ (or $\boldsymbol{\Gamma}_2$) the $m = \pm 1$ states have symmetry $\boldsymbol{\Gamma}_5$. With the z-component of the spin of $\pm 1/2$ we can then produce $J_z = \pm\hbar/2$ and $J_z = \pm 3\hbar/2$ states corresponding to $\boldsymbol{\Gamma}_7$ and $\boldsymbol{\Gamma}_9$ respectively. For values of J_z larger than $3\hbar/2$, however, the rather shaky approach above breaks down completely. See also the basis functions in the tables for O_h and T_d.

To conclude this last chapter we consider two more topics namely finite $\boldsymbol{k}$-values and time-reversal symmetry.

What happens if we leave the $\boldsymbol{\Gamma}$ point and go to $\boldsymbol{k} \neq 0$? First we apply all symmetry operations of the point group to the system as shown schematically in Fig. 26.5a for a two-dimensional square lattice. We end up with a number of $\boldsymbol{k}$-vectors, which are known as the "star of $\boldsymbol{k}$". If $\boldsymbol{k}$ is in a "general" position we get as many $\boldsymbol{k}$-vectors in the star as the group has elements. On the other hand, all these symmetry operations transform the system into itself. Consequently the properties must be the same for all elements of the star of $\boldsymbol{k}$. We illustrate this statement in Fig. 26.5b for the dispersion $E(\boldsymbol{k})$. We give contours of constant energy and see clearly that the star of $\boldsymbol{k}$ allows a fourfold band warping as indeed occurs for T_d and O_h symmetries, simply because cubic symmetry is lower than spherical symmetry.

If $\boldsymbol{k}$ is not in a "general" direction but coincides with one of the symmetry lines or planes of the system, then several of the elements of the star of $\boldsymbol{k}$ coincide as shown in Fig. 26.5c. The symmetry operations that transform $\boldsymbol{k}$ into

itself are known as the "group of $\boldsymbol{k}$". This group of $\boldsymbol{k}$ is a subgroup of the full point group. The irreducible representations at the $\boldsymbol{\Gamma}$-point are consequently also representations of the subgroup. If we start at the $\boldsymbol{\Gamma}$-point with a certain level which transforms like $\boldsymbol{\Gamma}_\alpha$ and move along some symmetry direction, the level $\boldsymbol{\Gamma}_\alpha$ may or may not split, depending on whether $\boldsymbol{\Gamma}_\alpha$ becomes a reducible representation or remains an irreducible one in the (sub) group of $\boldsymbol{k}$.

A typical example is the splitting of the fourfold degenerate $\boldsymbol{\Gamma}_8$ state in T_d symmetry for $\boldsymbol{k} \neq 0$ into two twofold degenerate states known as heavy and light-hole bands (Figs. 8.16 and 26.6), where the states $(J, J_z) = (3/2, \pm 3/2)$ correspond to the heavy hole valence band and $(3/2, \pm 1/2)$ corresponds to the light holes (see e.g. [74B1,03T1]). The $\boldsymbol{\Gamma}_6$ and $\boldsymbol{\Gamma}_7$ states do not show such behavior.

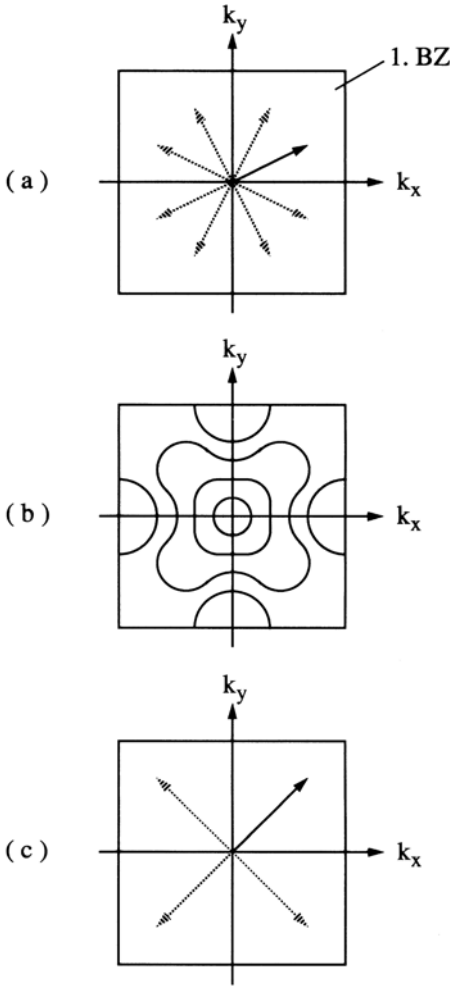


Fig. 26.5. The star of $\boldsymbol{k}$, (**a**) the qualitative shape of surfaces of constant energy $E(\boldsymbol{k})$ (**b**); and a $\boldsymbol{k}$-vector of higher symmetry (**c**)

For a general orientation of $\boldsymbol{k}$ all degeneracies can be lifted. In band structure theory one thus usually calculates $E(\boldsymbol{k})$ in directions of high symmetry (see, e.g., Figs. 8.9 – 8.16) to exploit the advantages of group theory and tries to extrapolate in other directions if necessary.

The final symmetry operation that we address is the invariance of a microscopic physical system under time reversal. If we neglect spin for the moment, invariance under time reversal has the following consequence for dispersion relations:

$$E(\boldsymbol{k}) = E(-\boldsymbol{k})\,, \tag{26.44}$$

even if the point group does not include the inversion. This phenomenon is known as Kramers degeneracy.

In a power series expansion of the dispersion

$$E(\boldsymbol{k}) = \sum_{n=0}^{\infty} a_n \boldsymbol{k}^n \tag{26.45}$$

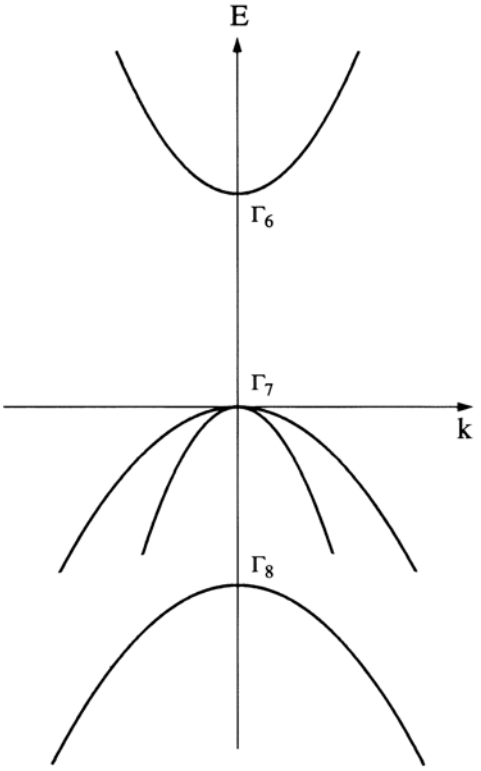


Fig. 26.6. The bandstructure around the $\boldsymbol{\Gamma}$-point for T_d symmetry. Compare with Fig. 8.16

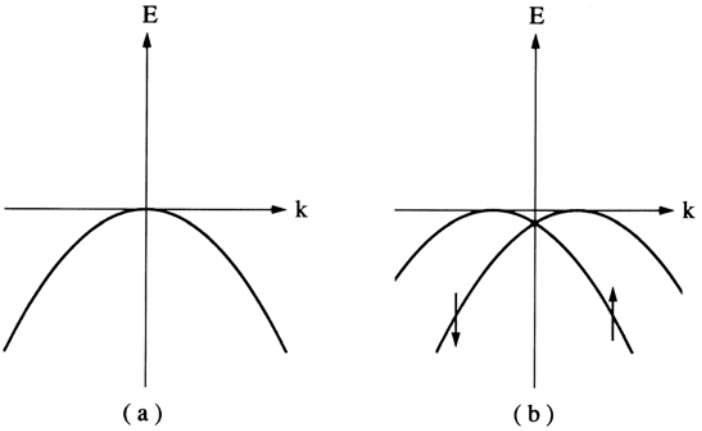


Fig. 26.7. Schematic drawing of the Kramers degeneracy without spin (**a**) and with spin (**b**) as occurs e.g., for Γ_7 bands in C_{6v} symmetry for $\boldsymbol{k} \perp \boldsymbol{c}$

this permits only even values for n (Fig. 26.7a). If we now include the spin, (26.46) changes to

$$E(\boldsymbol{k}, \uparrow) = E(-\boldsymbol{k}, \downarrow)\,. \tag{26.46}$$

This means that time reversal flips the spin. This condition can be also fulfilled with odd powers in the expansion (26.46) and, in particular, allows terms linear in $\boldsymbol{k}$ (Fig. 26.7b).

Detailed group-theoretical investigations show that k-linear terms are possible for Γ_8 states in T_d symmetry or for Γ_7 states $\boldsymbol{k} \perp \boldsymbol{c}$ in $C_{6\mathrm{v}}$ symmetry, but not for $\boldsymbol{k} \parallel \boldsymbol{c}$ and not for $\boldsymbol{\Gamma}_9$ levels.

Indeed such $\boldsymbol{k}$-linear terms are known; for example, for the B-exciton in CdS which contains a hole from the Γ_7 valence band or for the A excitons in ZnO. See Sect. 13.1.

26.6 Some Selected Group Tables

In this section, we give some tables for the three most important groups of semiconductors. We start in Tables 26.7 – 26.10 with the character tables of the point group $C_{6\mathrm{v}}$. In the first column we give the names of the irreducible representations in various notations, in the last one typical basis functions. From them it becomes clear, which representation belongs to the simple and which one to the double group, including spin namely here $\boldsymbol{\Gamma}_1$ to $\boldsymbol{\Gamma}_6$ and $\boldsymbol{\Gamma}_7$ to $\boldsymbol{\Gamma}_9$, respectively. The next tables give the multiplication tables of the irreducible representations and the combatibility relation with the subgroups and with the full rotation group. Tables 26.10 – 26.13 give the same information for T_d. For O_h we give only two tables for characters and the compatibility relation.

Table 26.7. Character table and basis functions of the point group C_{6v} (or $6mm$). Below E_2 or Γ_6 are the additional irreducible representations of the double group which includes spin

C_{6v}		E	$\bar{E}$	C_2	$2C_3$	$2\bar{C}_3$	$2C_6$	$2\bar{C}_6$	$3\bar{\sigma}_d$	$3\bar{\sigma}_v$		Basic functions
A_1	$\boldsymbol{\Gamma}_1$	1	1	1	1	1	1	1	1	1	a	R or z
A_2	$\boldsymbol{\Gamma}_2$	1	1	1	1	1	1	1	-1	-1	a	S_z
B_1	$\boldsymbol{\Gamma}_3$	1	1	-1	1	1	-1	-1	1	-1	a	$x^3 - 3xy^2$
B_2	$\boldsymbol{\Gamma}_4$	1	1	-1	1	1	-1	-1	-1	1	a	$x^3 - 3xy^2$
E_1	$\boldsymbol{\Gamma}_5$	2	2	-2	-1	-1	1	1	0	0	a	$(S_x - \mathrm{i}S_y)$ $-(S_x + \mathrm{i}S_y)$
E_2	$\boldsymbol{\Gamma}_6$	2	2	2	-1	-1	-1	-1	0	0	a	$\boldsymbol{\Gamma}_3 \otimes \boldsymbol{\Gamma}_5$
	$\boldsymbol{\Gamma}_7$	2	-2	0	1	-1	$\sqrt{3}$	$-\sqrt{3}$	0	0	c	$\phi(1/2, -1/2)$, $\phi(1/2, 1/2)$
	$\boldsymbol{\Gamma}_8$	2	-2	0	1	-1	$-\sqrt{3}$	$\sqrt{3}$	0	0	c	$\boldsymbol{\Gamma}_7 \otimes \boldsymbol{\Gamma}_3$
	$\boldsymbol{\Gamma}_9$	2	-2	0	-2	2	0	0	0	0	c	$\phi(3/2, -3/2)$, $\phi(3/2, 3/2)$

Table 26.8. Multiplication table for the irreducible representations of the group $C_{6\mathrm{v}}$

$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_7$	$\boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_9$	
$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_7$	$\boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_9$	$\boldsymbol{\Gamma}_1$
	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_7$	$\boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_9$	$\boldsymbol{\Gamma}_2$
		$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_7$	$\boldsymbol{\Gamma}_9$	$\boldsymbol{\Gamma}_3$
			$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_7$	$\boldsymbol{\Gamma}_9$	$\boldsymbol{\Gamma}_4$
				$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_9$	$\boldsymbol{\Gamma}_8 + \boldsymbol{\Gamma}_9$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_5$
					$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_8 + \boldsymbol{\Gamma}_9$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_9$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_6$
						$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_7$
							$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_5 + \boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_8$
								$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_9$
								$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4$	

Table 26.9. Compatibility table of the irreducible representations of the group $C_{6\mathrm{v}}$ with the irreducible representations of its proper subgroups

$C_{6\mathrm{v}} : E(z)$	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_7$	$\boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_9$
$C_6 : H(z)$	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_5 + \boldsymbol{\Gamma}_6$	$\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_8$	$\boldsymbol{\Gamma}_9 + \boldsymbol{\Gamma}_{10}$	$\boldsymbol{\Gamma}_{11} + \boldsymbol{\Gamma}_{12}$
C_{3v}	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_5 + \boldsymbol{\Gamma}_6$
C_{2v}	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_3$	$\boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_5$	$\boldsymbol{\Gamma}_5$
$C_s : E(x) : H(y)$	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4$
$C_s : H(x) : E(y)$	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_1$	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2$	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4$	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4$

Table 26.10. Compatibility table of the irreducible representations of the full rotation group of angular momentum $L_z \leq 6$ and $J_z \leq \frac{13}{2}$ with the irreducible representations of the group C_{6v}

D_0^+	$\boldsymbol{\Gamma}_1$	D_0^-	$\boldsymbol{\Gamma}_2$
D_1^+	$\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_5$	D_1^-	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_5$
D_2^+	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_5 + \boldsymbol{\Gamma}_6$	D_2^-	$\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_5 + \boldsymbol{\Gamma}_6$
D_3^+	$\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5 + \boldsymbol{\Gamma}_6$	D_3^-	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5 + \boldsymbol{\Gamma}_6$
D_4^+	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5 + 2\boldsymbol{\Gamma}_6$	D_4^-	$\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5 + 2\boldsymbol{\Gamma}_6$
D_5^+	$\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + 2\boldsymbol{\Gamma}_5 + 2\boldsymbol{\Gamma}_6$	D_5^-	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + 2\boldsymbol{\Gamma}_5 + 2\boldsymbol{\Gamma}_6$
D_6^+	$2\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + 2\boldsymbol{\Gamma}_5 + 2\boldsymbol{\Gamma}_6$	D_6^-	$\boldsymbol{\Gamma}_1 + 2\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + 2\boldsymbol{\Gamma}_5 + 2\boldsymbol{\Gamma}_6$
$D_{1/2}^+$	$\boldsymbol{\Gamma}_7$	$D_{1/2}^-$	$\boldsymbol{\Gamma}_7$
$D_{3/2}^+$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_9$	$D_{3/2}^-$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_9$
$D_{5/2}^+$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_8 + \boldsymbol{\Gamma}_9$	$D_{5/2}^-$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_8 + \boldsymbol{\Gamma}_9$
$D_{7/2}^+$	$\boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8 + 2\boldsymbol{\Gamma}_9$	$D_{7/2}^-$	$\boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8 + 2\boldsymbol{\Gamma}_9$
$D_{9/2}^+$	$\boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8 + \boldsymbol{\Gamma}_9$	$D_{9/2}^-$	$\boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8 + 2\boldsymbol{\Gamma}_9$
$D_{11/2}^+$	$2\boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8 + 2\boldsymbol{\Gamma}_9$	$D_{11/2}^-$	$2\boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8 + 2\boldsymbol{\Gamma}_9$
$D_{13/2}^+$	$3\boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8 + 2\boldsymbol{\Gamma}_9$	$D_{13/2}^-$	$3\boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8 + 2\boldsymbol{\Gamma}_9$

Table 26.11. Character table and basis functions for the irreducible representations of the point group T_d (or $\bar{4}3\,m$)

T_d		E	$\bar{E}$	$8C_3$	$8\bar{C}_3$	$3C_2$	$6S_4$	$6\bar{S}_4$	$6\sigma_d$	Time reversal	Basis functions
A_1	$\boldsymbol{\Gamma}_1$	1	1	1	1	1	1	1	1	a	R or xyz
A_2	$\boldsymbol{\Gamma}_2$	1	1	1	1	1	−1	−1	−1	a	$S_x S_y S_z$
E	$\boldsymbol{\Gamma}_3$	2	2	−1	−1	2	0	0	0	a	$(2z^2 - x^2 - y^2)$
											$\sqrt{3}(s^2 - y^2)$
F_1	$\boldsymbol{\Gamma}_4$	3	3	0	0	−1	1	1	−1	a	S_x, S_y, S_z
F_2	$\boldsymbol{\Gamma}_5$	3	3	0	0	−1	−1	−1	1	a	x, y, z
$\bar{E}_1$	$\boldsymbol{\Gamma}_6$	2	−2	1	−1	0	$\sqrt{2}$	$-\sqrt{2}$	0	c	$\phi(1/2, -1/2)$
											$\phi(1/2, 1/2)$
$\bar{E}_2$	$\boldsymbol{\Gamma}_7$	2	−2	1	−1	0	$-\sqrt{2}$	$\sqrt{2}$	0	c	$\boldsymbol{\Gamma}_6 \times \boldsymbol{\Gamma}_2$
$\bar{E}_3$	$\boldsymbol{\Gamma}_8$	4	−4	−1	1	0	0	0	0	c	$\phi(3/2, -3/2)$
											$\phi(3/2, -3/2)$
											$\phi(3/2, 1/2)$
											$\phi(3/2, 3/2)$

Table 26.12. Multiplication table of the irreducible representations of the group T_d

Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8	
Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8	Γ_1
	Γ_1	Γ_3	Γ_5	Γ_4	Γ_7	Γ_6	Γ_8	Γ_2
		$\Gamma_1+\Gamma_2+\Gamma_3$	$\Gamma_4+\Gamma_5$	$\Gamma_4+\Gamma_5$	Γ_8	Γ_8	$\Gamma_6+\Gamma_7+\Gamma_8$	Γ_3
			$\Gamma_1+\Gamma_3+\Gamma_4+\Gamma_5$	$\Gamma_2+\Gamma_3+\Gamma_4+\Gamma_5$	$\Gamma_6+\Gamma_8$	$\Gamma_7+\Gamma_8$	$\Gamma_6+\Gamma_7+2\Gamma_8$	Γ_4
				$\Gamma_1+\Gamma_3+\Gamma_4+\Gamma_5$	$\Gamma_7+\Gamma_8$	$\Gamma_6+\Gamma_8$	$\Gamma_6+\Gamma_7+2\Gamma_8$	Γ_5
					$\Gamma_1+\Gamma_4$	$\Gamma_2+\Gamma_5$	$\Gamma_3+\Gamma_4+\Gamma_5$	Γ_6
						$\Gamma_1+\Gamma_4$	$\Gamma_3+\Gamma_4+\Gamma_5$	Γ_7
							$\Gamma_1+\Gamma_2+\Gamma_3$	Γ_8
							$2\Gamma_4+2\Gamma_5$	

Table 26.13. Compatibility table of the irreducible representations of T_d with the irreducible representations of its proper subgroups

T_d	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8
T	Γ_1	Γ_1	$\Gamma_2+\Gamma_3$	Γ_4	Γ_4	Γ_5	Γ_5	$\Gamma_6+\Gamma_7$
D_{2d}	Γ_1	Γ_3	$\Gamma_1+\Gamma_3$	$\Gamma_2+\Gamma_5$	$\Gamma_4+\Gamma_5$	Γ_6	Γ_7	$\Gamma_6+\Gamma_7$
$C_{3v}: E(w)$	Γ_1	Γ_2	Γ_3	$\Gamma_2+\Gamma_3$	$\Gamma_1+\Gamma_3$	Γ_4	Γ_4	$\Gamma_4+\Gamma_5+\Gamma_6$
$S_1: H(z)$	Γ_1	Γ_2	$\Gamma_1+\Gamma_2$	$\Gamma_1+\Gamma_3+\Gamma_4$	$\Gamma_2+\Gamma_3+\Gamma_4$	$\Gamma_5+\Gamma_6$	$\Gamma_7+\Gamma_8$	$\Gamma_5+\Gamma_6+\Gamma_7+\Gamma_5$
$C_3: H(w)$	Γ_1	Γ_1	$\Gamma_2+\Gamma_3$	$\Gamma_1+\Gamma_2+\Gamma_3$	$\Gamma_1+\Gamma_2+\Gamma_3$	$\Gamma_4+\Gamma_5$	$\Gamma_4+\Gamma_5$	$\Gamma_4+\Gamma_5+2\Gamma_6$
$C_{2v}: E(z)$	Γ_1	Γ_3	$\Gamma_1+\Gamma_3$	$\Gamma_2+\Gamma_3+\Gamma_4$	$\Gamma_1+\Gamma_2+\Gamma_4$	Γ_5	Γ_5	$2\Gamma_5$
$C_s: E(v): H(v)$	Γ_1	Γ_2	$\Gamma_1+\Gamma_2$	$\Gamma_1+2\Gamma_2$	$2\Gamma_1+\Gamma_2$	$\Gamma_3+\Gamma_4$	$\Gamma_3+\Gamma_4$	$2\Gamma_3+2\Gamma_4$

Table 26.14. Compatibility table of the irreducible representations of the full rotation group for $L_z \leq 6$ and $J_z \leq \frac{13}{2}$ with the representations of the point group T_d

D_0^+	$\boldsymbol{\Gamma}_1$	D_0^-	$\boldsymbol{\Gamma}_2$
D_1^+	$\boldsymbol{\Gamma}_4$	D_1^-	$\boldsymbol{\Gamma}_5$
D_2^+	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_5$	D_2^-	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4$
D_3^+	$\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5$	D_3^-	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5$
D_4^+	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5$	D_4^-	$\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5$
D_5^+	$\boldsymbol{\Gamma}_3 + 2\boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5$	D_5^-	$\boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + 2\boldsymbol{\Gamma}_5$
D_6^+	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_3 + \boldsymbol{\Gamma}_4 + 2\boldsymbol{\Gamma}_5$	D_6^-	$\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_3 + 2\boldsymbol{\Gamma}_4 + \boldsymbol{\Gamma}_5$
$D_{1/2}^+$	$\boldsymbol{\Gamma}_6$	$D_{1/2}^-$	$\boldsymbol{\Gamma}_7$
$D_{3/2}^+$	$\boldsymbol{\Gamma}_8$	$D_{3/2}^-$	$\boldsymbol{\Gamma}_8$
$D_{5/2}^+$	$\boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_8$	$D_{5/2}^-$	$\boldsymbol{\Gamma}_6 + \boldsymbol{\Gamma}_8$
$D_{7/2}^+$	$\boldsymbol{\Gamma}_6 + \boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_8$	$D_{7/2}^-$	$\boldsymbol{\Gamma}_6 + \boldsymbol{\Gamma}_7 + \boldsymbol{\Gamma}_8$
$D_{9/2}^+$	$\boldsymbol{\Gamma}_6 + 2\boldsymbol{\Gamma}_8$	$D_{9/2}^-$	$\boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8$
$D_{11/2}^+$	$\boldsymbol{\Gamma}_6 + \boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8$	$D_{11/2}^-$	$\boldsymbol{\Gamma}_6 + \boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8$
$D_{13/2}^+$	$\boldsymbol{\Gamma}_6 + 2\boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8$	$D_{13/2}^-$	$2\boldsymbol{\Gamma}_6 + \boldsymbol{\Gamma}_7 + 2\boldsymbol{\Gamma}_8$

Table 26.15. Character table and basis functions of the point group O_h (or m3m) which contains the inversion and consequently has parity as a good quantum number

O_h			E	$\bar{E}$	$8C_3$	$8\bar{C}_3$	$3C_2$ $3\bar{C}_2$	$6C_4$	$6\bar{C}_4$	$6C_{2'}$ $6\bar{C}_{2'}$	I	$\bar{I}$	$8S_6$	$8\bar{S}_6$	$3\sigma_h$ $3\bar{\sigma}_h$	$6S_4$	$6\bar{S}_4$	$6\sigma_d$ $6\bar{\sigma}_d$	Time reversal	Basis functions
A_{1g}	Γ_1	Γ_1^+	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	a	R
A_{2g}	Γ_2	Γ_2^+	1	1	1	1	−1	−1	−1	1	1	1	1	1	1	−1	−1	−1	a	$(x^2-y^2)(y^2-z^2)(z^2-x^2)$
E_g	Γ_{12}	Γ_3^+	2	2	−1	−1	2	0	0	0	2	2	−1	−1	2	0	0	0	a	$(2z^2-x^2-y^2)$, $\sqrt{3}(x^2-y^2)$
F_{1g}	Γ'_{15}	Γ_4^+	3	3	0	0	−1	1	1	−1	3	3	0	0	−1	1	1	−1	a	S_x, S_y, S_z
F_{2g}	Γ'_{25}	Γ_5^+	3	3	0	0	−1	−1	−1	1	3	3	0	0	−1	−1	−1	1	a	yx, xz, xy
A_{1u}	Γ'_1	Γ_1^-	1	1	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1	−1	−1	a	$\Gamma_2^- \times \Gamma_2^+$
A_{2u}	Γ'_2	Γ_2^-	1	1	1	1	1	−1	−1	−1	−1	−1	−1	−1	−1	1	1	1	a	xyz
E_u	Γ'_{12}	Γ_3^-	2	2	−1	−1	2	0	0	0	−2	−2	1	1	−2	0	0	0	a	$\Gamma_3^+ \times \Gamma_2^-$
F_{1u}	Γ_{15}	Γ_4^-	3	3	0	0	−1	1	1	−1	−3	−3	0	0	1	−1	−1	1	a	$\Gamma_5^+ \times \Gamma_1^-$
F_{2u}	Γ_{25}	Γ_5^-	3	3	0	0	−1	−1	−1	1	−3	−3	0	0	1	1	1	−1	a	$\Gamma_5^+ \times \Gamma_1^-$
		Γ_6^+	2	−2	1	−1	0	$\sqrt{2}$	$-\sqrt{2}$	0	2	−2	1	−1	0	$\sqrt{2}$	$-\sqrt{2}$	0	a	$\phi(1/2,-1/2)$, $\phi(1/2,1/2)$
		Γ_7^+	2	−2	1	−1	0	$-\sqrt{2}$	$\sqrt{2}$	0	2	−2	1	−1	0	$-\sqrt{2}$	$\sqrt{2}$	0	a	$\Gamma_6^+ \times \Gamma_2^+$
		Γ_8^+	4	−4	−1	1	0	0	0	0	4	−4	−1	1	0	0	0	0	a	$\phi(3/2,-3/2)$, $\phi(3/2,3/2)$
																				$\phi(3/2,1/2)$, $\phi(3/2,3/2)$
		Γ_6^-	2	−2	1	−1	0	$\sqrt{2}$	$-\sqrt{2}$	0		−2	2	−1	1	0	$-\sqrt{2}$	$\sqrt{2}$	0	c	$\Gamma_6^+ \times \Gamma_1^-$
		Γ_7^-	2	−2	1	−1	0	$-\sqrt{2}$	$\sqrt{2}$	0		−2	2	−1	1	0	$\sqrt{2}$	$-\sqrt{2}$	0	c	$\Gamma_6^+ \times \Gamma_2^-$
		Γ_8^-	4	−4	−1	1	0	0	0	0		−4	4	1	−1	0	0	0	0	c	$\Gamma_8^+ \times \Gamma_1^-$

Table 26.16. The compatibility table of the irreducible representations of the group O_h and the irreducible representations of its proper subgroups

O^h	Γ_1^+	Γ_2^+	Γ_3^+	Γ_4^+	Γ_5^+	Γ_1^-	Γ_2^-	Γ_3^-	Γ_4^-
O	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_1	Γ_2	Γ_3	Γ_4
T_d	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_2	Γ_1	Γ_3	Γ_5
T_h	Γ_1^+	Γ_1^+	$\Gamma_2^+ + \Gamma_3^+$	Γ_4^+	Γ_4^+	Γ_1^-	Γ_1^-	$\Gamma_2^- + \Gamma_3^-$	Γ_4^-
D_{4h}	Γ_1^+	Γ_3^+	$\Gamma_1^+ + \Gamma_3^+$	$\Gamma_2^+ \Gamma_5^+$	$\Gamma_4^+ + \Gamma_5^+$	Γ_1^-	Γ_3^-	$\Gamma_1^- + \Gamma_3^-$	$\Gamma_2^- + \Gamma_5^-$
D_{3d}	Γ_1^+	Γ_2^+	Γ_3^+	$\Gamma_2^+ \Gamma_3^+$	$\Gamma_1^- + \Gamma_3^-$	Γ_1^-	Γ_2^-	Γ_3^-	$\Gamma_2^- + \Gamma_3^-$
$C_{4h} : H(z)$	Γ_1^+	Γ_2^+	$\Gamma_1^+ + \Gamma_2^+$	$\Gamma_1^+ + \Gamma_3^+ + \Gamma_4^+$	$\Gamma_2^+ + \Gamma_3^+ + \Gamma_4^+$	Γ_1^-	Γ_2^-	$\Gamma_1^- + \Gamma_2^-$	$\Gamma_1^- + \Gamma_3^- + \Gamma_4^-$
$C_{2h} : H(v)$	Γ_1^+	Γ_2^+	$\Gamma_1^+ + \Gamma_2^+$	$2\Gamma_2^+ + \Gamma_1^+$	$2\Gamma_1^+ + \Gamma_2^+$	Γ_1^-	Γ_2^-	$\Gamma_1^- + \Gamma_2^-$	$2\Gamma_2^- + \Gamma_1^-$
$C_{31} : H(w)$	Γ_1^+	Γ_1^+	$\Gamma_2^+ + \Gamma_3^+$	$\Gamma_1^+ + \Gamma_2^+ + \Gamma_3^+$	$\Gamma_1^+ + \Gamma_2^+ + \Gamma_3^+$	Γ_1^-	Γ_1^-	$\Gamma_2^- + \Gamma_3^-$	$\Gamma_1^- + \Gamma_2^- + \Gamma_3^-$
$C_{4v} : E(z)$	Γ_1	Γ_3	$\Gamma_1 + \Gamma_3$	$2\Gamma_2 + \Gamma_5$	$2\Gamma_4 + \Gamma_5$	Γ_2	Γ_4	$\Gamma_2 + \Gamma_4$	$\Gamma_1 + \Gamma_5$
$C_{2v} : E(v)$	Γ_1	Γ_2	$\Gamma_1 + \Gamma_2$	$2\Gamma_2 + \Gamma_3 + \Gamma_4$	$2\Gamma_1 + \Gamma_3 + \Gamma_4$	Γ_3	Γ_4	$\Gamma_3 + \Gamma_4$	$\Gamma_1 + \Gamma_2 + \Gamma_4$
$C_{3v} : E(w)$	Γ_1	Γ_2	Γ_3	$2\Gamma_2 + \Gamma_3$	$2\Gamma_1 + \Gamma_3$	Γ_2	Γ_1	Γ_3	$\Gamma_1 + \Gamma_3$

O^h	Γ_5^-	Γ_6^+	Γ_7^+	Γ_8^+	Γ_6^-	Γ_7^-	Γ_8^-
O	Γ_5	Γ_6	Γ_7	Γ_8	Γ_6	Γ_7	Γ_8
T_d	Γ_4	Γ_6	Γ_7	Γ_8	Γ_7	Γ_6	Γ_8
T_h	Γ_4^-	Γ_5^+	Γ_5^+	$\Gamma_6^+ + \Gamma_7^+$	Γ_5^-	Γ_5^-	$\Gamma_6^- + \Gamma_7^-$
D_{4h}	$\Gamma_4^- + \Gamma_5^-$	Γ_6^+	Γ_7^+	$\Gamma_6^+ + \Gamma_7^+$	Γ_6^-	Γ_7^-	$\Gamma_6^- + \Gamma_7^-$
D_{3d}	$\Gamma_1^- + \Gamma_3^-$	Γ_4^+	Γ_4^-	$\Gamma_6^+ + \Gamma_5^- + \Gamma_6^-$	Γ_4^-	Γ_4^-	$\Gamma_4^- + \Gamma_5^- + \Gamma_6^-$
$C_{4h} : H(z)$	$\Gamma_2^- + \Gamma_3^- + \Gamma_4^-$	$\Gamma_5^+ + \Gamma_6^+$	$\Gamma_7^+ + \Gamma_8^+$	$\Gamma_5^+ + \Gamma_6^+ + \Gamma_7^+ + \Gamma_8^+$	$\Gamma_5^- + \Gamma_6^-$	$\Gamma_8^- + \Gamma_7^-$	$\Gamma_5^- + \Gamma_6^- + \Gamma_8^- + \Gamma_7^-$
$C_{2h} : H(v)$	$2\Gamma_1^- + \Gamma_2^-$	$\Gamma_3^+ + \Gamma_4^+$	$\Gamma_3^+ + \Gamma_4^+$	$2\Gamma_3^+ + 2\Gamma_4^+$	$\Gamma_3^- + \Gamma_4^-$	$\Gamma_3^- + \Gamma_4^-$	$2\Gamma_3^- + 2\Gamma_4^-$
$C_{3l} : H(w)$	$\Gamma_1^- + \Gamma_2^- + \Gamma_3^-$	$\Gamma_4^- + \Gamma_5^-$	$\Gamma_4^- + \Gamma_5^-$	$\Gamma_4^+ + \Gamma_5^+ + 2\Gamma_6^+$	$\Gamma_4^- + \Gamma_5^-$	$\Gamma_4^- + \Gamma_5^-$	$2\Gamma_4^+ + \Gamma_5^+ + 2\Gamma_6^+$
$C_{4v} : E(z)$	$\Gamma_3 + \Gamma_5$	Γ_6	Γ_7	$\Gamma_6 + \Gamma_7$	Γ_6	Γ_7	$\Gamma_6 + \Gamma_7$
$C_{2v} : E(v)$	$\Gamma_1 + \Gamma_2 + \Gamma_3$	Γ_5	Γ_5	$2\Gamma_5$	Γ_5	Γ_5	$2\Gamma_5$
$C_{3v} : E(w)$	$\Gamma_2 + \Gamma_3$	Γ_4	Γ_4	$\Gamma_4^+ + \Gamma_5^+ + 2\Gamma_6^+$	Γ_4	Γ_4^+	$\Gamma_4 + \Gamma_5 + \Gamma_6$

26.7 Problems

1. Calculate the symmetries of the states which can be reached by one- and two-photon transitions in point groups O_h and T_d.
2. Calculate the possible symmetries of the excitons with principal quantum number $n_B = 1, 2$ and 3 in the point groups O_h, T_d and C_{6v}.
3. Calculate the symmetries of the ground state ($\boldsymbol{\Gamma}_1$ envelope) of D^0X, A^0X and D^+X centers in O_h, T_d, and C_{6v}, assuming that the radius of these complexes is so large that the carriers feel the full symmetry of the lattice and not only the nearest neighbors.
4. SnO_2 is a direct gap semiconductor with the band extrema at the $\boldsymbol{\Gamma}$-point. It crystallizes in the point group D_{4h}. Which symmetries do you expect for the highest valence and the lowest conduction band. The lowest band-to-band transition is dipole forbidden. Why? Do you except dipole-allowed exciton transitions? Do you think that they will have a big oscillator strength?
5. Rock salt, diamond and Cu_2O all have point group O_h (i.e., a face-centered-cubic lattice), but different crystal structures. Inspect the crystal structures. Find the primitive unit cell and the nonprimitive cubic one. Parity is a good quantum number. The highest valence band is formed in Cu_2O from the $3d$ levels of Cu^+ and the lowest conduction bands result from the Cu^+ $4s$ and $4p$ levels. Which symmetries do you expect at the $\boldsymbol{\Gamma}$-point and which exciton states with S envelope? Are they dipole allowed or forbidden?

References to Chap. 26

[60H1] V. Heine, Group Theory in Quantum Mechanics, Pergamon, Oxford (1960)
[63K1] G.F. Koster, et al., eds., Properties of the Thirty-Two Point Groups, MIT, Cambridge (1963)
[64K1] R.S. Knox and A. Gold, Symmetry in the Solid State, Benjamin, New York (1964)
[64T1] W.A. Tinkham, Group Theory in Quantum Mechanics, McGraw-Hill, New York (1964)
[67S1] H.W. Streitwolf, Gruppentheorie in der Festkörperphysik, Akad. Verl. Leipzig (1967)
[74B1] G.L. Bir and G.E. Pikus, Symmetry and Strain-Induced Effects in Semiconductors, Wiley, New York (1974)
[76C1] K. Cho, Phys. Rev. B **14**, 4463 (1976)
[77G1] O. Goede, phys. stat. sol. (b) **81**, 235 (1977)
[85H1] B. Hönerlage, et al., Phys. Rep. **124**, 161 (1985)
[88L1] W. Ludwig and C. Falter, Symmetries in Physics, Springer Ser. Solid-State Sci. Vol. **64**, Springer, Berlin, Heidelberg (1988)
[90I1] T. Inui, Y. Tanabe and Y. Onadera, Group Theory and its Applications in Physics, Springer Series on Solid State Sciences **78**, Springer, Berlin (1990)

[92I1] International Tables for Crystallography Vol. A Space Groups 3^{rd} ed., T. Hahn, Kluver, Dordrecht (1992)
[93J1] S. Jorda, U. Rössler and D. Broido, Phys. Rev. B **48**, 1669 (1993)
[94H1] Th. Hahn and H. Wondratschek, Symmetry of Crystals, Lecture Notes to the Summer School in Gjuletchitsa, Bulgaria, Heron Press Ltd., Sophia, Bulgaria (1994)
[96I1] E.L. Ivchenko, Yu.A. Kaminski and U. Rössler, Phys. Rev. B **54**, R2315 (1996)
[96Y1] P.Y. Yu and M. Cardona, Fundamental of Semiconductor, Springer, Berlin, Heidelberg (1996)
[97I1] E.L. Ivchenko and G.E. Pikus, Superlattices and other Heterostructures, 2nd ed. Springer, Berlin (1997)
[00S1] M. Schmidt, et al., Appl. Phys. Lett. **77**, 85 (2000)
[01L1] Landolt-Börnstein, New Series, Group III, Vol. **41**a1 Part 2, U. Rössler, ed., Springer, Berlin, Heidelberg (2001)
[02M1] A.A. Maksimov, et al., phys. stat. sol. **229**, 35 (2002)
[03T1] E. Tsitisishvili, NATO Science Series II, Vol. **90**, p 357, Kluver, Dordrecht (2003)

27

Semiconductor Bloch Equations

By R. v. Baltz

The electrodynamic description of matter requires constitutive equations that relate the induced charge ρ and current density $\boldsymbol{j}$ of the semiconductor (or, equivalently, the polarization $\boldsymbol{P}$, with $\boldsymbol{j} = \dot{\boldsymbol{P}}$ and $\rho = -\mathrm{div}\boldsymbol{P}$) to the electromagnetic fields $\boldsymbol{E}, \boldsymbol{B}$. Generic models in this respect are the Lorentz-oscillator and the Drude free-carrier model of linear optics that have already been extensively used in Chaps. 4, 5, and 11–13. The description of nonlinear properties of matter, on the other hand, mostly uses a power series expansion of the polarization in terms of the electrical field (see Chap. 19). Such an expansion, however, is inappropriate under resonant or near-resonant conditions. In some cases new solutions may even arise "spontaneously" above a critical light field and can lead to second harmonic generation although a power expansion (including even-order terms with respect to the light field) does not exist. Therefore, a realistic description of semiconductor optics requires the proper dependence on the light field, the inclusion of the valence conduction band continuum states, exciton effects, as well as band-filling dynamics. These phenomena are consistently described by the *Semiconductor Bloch-Equations* (SBEs), which have become the standard model for semiconductor optics. [1] In this approach the semiconductor is treated quantum mechanically leads to a set of coupled nonlinear differential equations for the polarization and the electron/hole distribution functions (supplemented by higher order correlation functions, which will be omitted here). The polarization acts as a source term in the (classical) Maxwell equations as discussed in Chap. 2. In this sense, the SBEs are a semiclassical theory. It successfully covers linear as well as nonlinear phenomena such as pump-probe, four-wave-mixing, or photon echo experiments (see Chaps. 20– 25).

[1] The main idea of the SBEs originates from Bloch's seminal work on the theory of (nuclear) spin resonance [46B1] and Haken's theory of the laser [70H1]. For semiconductors, the band-edge equations were first studied by Stahl [84S1] and subsequently developed, by Haug and his coworker and colleagues [94H1] and others, into a powerful tool for semiconductor theory.

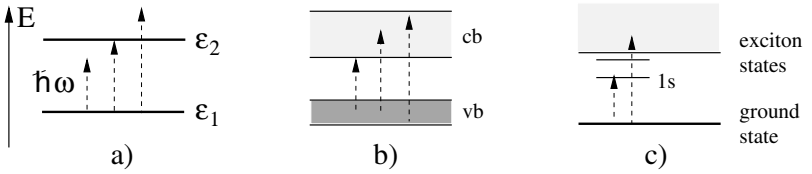


Fig. 27.1. Route for a qualitative derivation of the the semiconductor Bloch equations. (**a**) Two-level system (resonant and nonresonant excitation) (**b**) Noninteracting valence and conduction bands viewed as a collection of TLSs with different transition energies (**c**) Two interacting bands including some bound exciton states

The SBEs are of considerable complexity in derivation and application, therefore, we shall give only a "pedestrian version" of their derivation and some selected applications. Details can be found in Haug and Koch's textbook [94H1]. For a recent comprehensive presentation of the SBEs see, e.g., Schäfer and Wegener's book [02S1]. We approach the problem in three steps as sketched in Fig. 27.1

1. First we study the dynamics of atoms near resonance in the two-level approximation and derive the optical Bloch equations. In this formulation, damping processes can also be included on a phenomenological level.
2. Second, we generalize this result for a two-band model of a direct semiconductor omitting the Coulomb interaction between electrons and holes. In this description the semiconductor is sketched as a collection of noninteracting two-level systems (TLSs).
3. Finally, we add the Coulomb interaction between the electrons and holes, which includes exciton and screening effects. This will lead us to the SBEs in their simplest form, including excitons or the transition to an electron hole plasma.

27.1 Dynamics of a Two-Level System[2]

Near the resonance of an atomic transition, $\omega \approx \omega_0 = (\epsilon_2 - \epsilon_1)/\hbar$, we may only retain the pair of nearly resonant stationary states $|1\rangle$ and $|2\rangle$ with energies ϵ_1 and ϵ_2, $(\epsilon_2 > \epsilon_1)$, between which the transition occurs (see Fig. 27.1a). In this two-level approximation, the wave function

$$|\,\psi(t)\rangle = c_1(t)|1\rangle + c_2(t)|2\rangle \tag{27.1}$$

[2] The dynamics of a TLS is very well presented in many texts, in particular Allen and Eberly's classic book on optical resonance and two-level atoms [87A1] or the textbooks *The Feynman Lectures*, Vol. III, in connection with the Ammonia maser [64F1] and *Licht und Materie* by Haken [70H1]. Note the different enumeration of states.

is described by coefficients c_1, c_2, which may be arranged as a column vector $\boldsymbol{c}$, the time-dependence of which is governed by the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \widehat{H}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}. \tag{27.2}$$

In addition, we assume that the optical transition between $|1\rangle$ and $|2\rangle$ is dipole-allowed and the light is polarized parallel to the z-axis. As a result, the Hamiltonian reads

$$\widehat{H} = \widehat{H}_0 - \widehat{d}E_z(t), \qquad \widehat{H}_0 = \begin{pmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_2 \end{pmatrix}, \qquad \widehat{d} = \begin{pmatrix} 0 & d^* \\ d & 0 \end{pmatrix}. \tag{27.3}$$

$\widehat{H}_0$ and $\widehat{d}$ denote the Hamiltonian of the isolated atom and the electric dipole operator with $d = \langle 2| - ez|1\rangle$, respectively (see also Sect. 3.2.2).

From $c_1(t), c_2(t)$ the induced dipole moment $d(t)$ and the population inversion $I(t)$ of the TLS are determined by

$$d(t) = \mathbf{c}^\dagger \widehat{d}^* \mathbf{c} = 2\,\mathrm{Re}\left[d^*\,\mathcal{P}(t)\right], \tag{27.4}$$

$$I(t) = |c_2(t)|^2 - |c_1(t)|^2, \quad \text{and} \tag{27.5}$$

$$\mathcal{P}(t) = c_1^*(t)c_2(t), \tag{27.6}$$

where $\mathcal{P}(t)$ is the (dimensionless) complex dipole moment.[3] ϵ_1, ϵ_2 and d are considered as parameters of this model. There are at least three ways to tackle the problem. These will be discussed in the next sections.

27.1.1 Wave-Function Description

We start with a pure state described by (27.1) and factor off the time dependence of the state vector without an external field (interaction picture),

$$c_j(t) = a_j(t)e^{-\imath\epsilon_j t/\hbar}, \quad j = 1, 2, \tag{27.7}$$

where $a_j(t)$ obey the differential equations

$$\dot{a}_1(t) = i\frac{d^*}{\hbar}\,E_z(t)\,e^{-i\omega_0 t}a_2(t), \tag{27.8}$$

$$\dot{a}_2(t) = i\frac{d}{\hbar}\,E_z(t)\,e^{+i\omega_0 t}a_1(t). \tag{27.9}$$

Equations (27.8,27.9) are linear and first order. Nevertheless, no analytical solution is achievable. However, for a monochromatic field with frequency ω near resonance, $\omega \approx \omega_0$, the product of $E_z(t) = E_0\cos(\omega t)$ and $e^{\pm i\omega_0 t}$, respectively, contain a term that is almost constant and another that oscillates

[3] Haug and Koch [94H1] use a reverse notation concerning $\boldsymbol{P}$ and $\mathcal{P}$.

rapidly. The first term drives the system coherently whereas the fast oscillating term almost averages to zero, which will be neglected in the following. For reasons that will become obvious later, this approximation is called the rotating wave approximation (RWA). Within the RWA, (27.8,27.9) become

$$\dot{a}_1(t) = i\frac{\omega_R^*}{2}\, e^{+i\delta t} a_2(t), \tag{27.10}$$

$$\dot{a}_2(t) = i\frac{\omega_R}{2}\, e^{-i\delta t} a_1(t), \tag{27.11}$$

which (besides ω and ω_0) contain two characteristic frequencies

- $\omega_R = d\,E_0/\hbar$, the *Rabi frequency* (real, for simplicity) and
- $\delta = \omega - \omega_0$, which is called the "detuning" frequency.

By insertion, (27.10,27.11) lead to an equation of motion for a harmonic oscillator with "imaginary damping"

$$\ddot{a}_1(t) - i\delta\dot{a}_1(t) + (\frac{\omega_R}{2})^2 a_1(t) = 0\,, \tag{27.12}$$

which can be solved by the standard exponential ansatz. For example, at resonance, $\delta = 0$, and initial conditions $a_1(0) = 1$, $a_2(0) = 0$, we have

$$a_1(t) = \cos(\frac{\omega_R}{2}t), \quad a_2(t) = i\sin(\frac{\omega_R}{2}t). \tag{27.13}$$

Hence, the probability of finding the TLS in the excited state, $|c_2(t)|^2 = |a_2(t)|^2$, oscillates with the Rabi frequency ω_R. The same holds true for the population inversion, (see Fig. 27.2). At time $t_1 = \pi/\omega_R$ the atom is in the excited state and at $2\pi/\omega_R$ it is back in the ground state. This is called *Rabi flopping*. For detuned fields, the solution of (27.12) displays incomplete Rabi flopping where the period becomes shorter and the amplitude is smaller than when on resonance. See also Sect. 23.2.2.3.

A numerical example may be found, e.g., in Boyd [92B1]. For atomic sodium the parameters for the $3s-3p$ transition are: $d = 2.5a_Be$, $\lambda_0 = 589\,\mathrm{nm}$

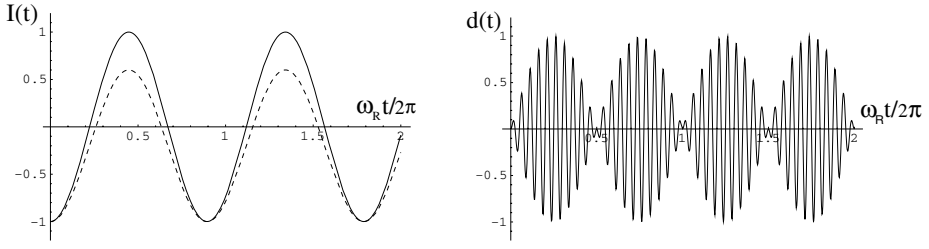


Fig. 27.2. Rabi oscillations of the inversion (*left*) and induced dipole moment (*right*). Resonant excitation (*solid line*), detuned excitation at $\omega = 0.5\omega_0$ (*dashed line*), $\omega_R = 0.05\omega_0$

(a_B is the Bohr radius). For an intensity of $127\,\mathrm{W\,cm^{-2}}$ (which can be easily realized by focusing a low power laser beam) the Rabi frequency $\omega_R/2\pi \approx 1\,\mathrm{GHz}$ becomes larger than the natural line width and oscillations will appear. For further examples see, e.g., Di Bartolo's article [05B1].

27.1.2 Polarization and Inversion as State Variables

Instead of c_1, c_2 one can likewise use the inversion $I(t)$ and the complex polarization $\mathcal{P}(t)$ themselves as state variables which – remarkably – obey a closed set of first order differential equations. Using (27.5,27.6) together with (27.8,27.9) we obtain (even without RWA)

$$\left[\frac{\mathrm{d}}{\mathrm{d}t} + i\omega_0\right] \mathcal{P}(t) = -i\omega_R(t)\, I(t) + \dot{\mathcal{P}}_{\mathrm{rel}}, \tag{27.14}$$

$$\frac{dI(t)}{dt} = -4\,\mathrm{Im}\,\left[\omega_R(t)\mathcal{P}^*(t)\right] + \dot{I}_{rel}, \tag{27.15}$$

$$\omega_R(t) = \frac{d}{\hbar} E_z(t). \tag{27.16}$$

$\omega_R(t)$ is the instantaneous Rabi frequency. (In the following, d as well as $\omega_R(t)$, can be assumed to be real).

The advantage of this description with respect to (27.8,27.9) is the possibility to include damping (i.e., incoherent interactions with a "bath") in a phenomenological way just by adding relaxation terms, which is not possible to do for (27.8,27.9)[4]:

$$\dot{\mathcal{P}}_{rel} = -\frac{\mathcal{P}}{T_2}, \quad \dot{I}_{rel} = -\frac{I(t) - I_{\mathrm{eq}}}{T_1}. \tag{27.17}$$

T_1 and T_2 determine the population (or energy) and phase relaxation of the TLS and they are also called longitudinal and transverse relaxation times. (The notation will become clear at the end of the next section). As I is the square of an amplitude, we may expect $T_2 = 2T_1$. However, there are always phase disturbing processes that do not contribute to energy relaxation so that in general $T_2 \leq 2T_1$. In semiconductors, one finds frequently that $T_2 \ll 2T_1$. I_{eq} is the equilibrium value of the inversion in the absence of the driving field, e.g., at zero temperature $I_{\mathrm{eq}} = -1$. See also Sect. 23.1.

Without damping, there is a conserved quantity,

$$4\,|\mathcal{P}(t)|^2 + I^2(t) = \mathrm{const}\,, \tag{27.18}$$

which may be used to eliminate the inversion from (27.14). Its origin becomes obvious from a remarkable analogy between any TLS and a spin-1/2 system in a magnetic field as discussed in the following subsection.

[4] A microscopic treatment of relaxation processes requires a density matrix approach or equivalent formulations. $I(t)$ contains the diagonal elements of the density operator, $\rho_{jj} (= |c_j|^2)$, whereas $\mathcal{P}(t)$ probes the off-diagonal element $\rho_{12} (= c_1^* c_2)$.

27.1.3 Pseudo-Spin Formulation

The Hamiltonian of a single spin-1/2 in a magnetic field $\boldsymbol{B}$ reads

$$\widehat{H} = -g\frac{\mu_B}{\hbar}\,\widehat{S}\boldsymbol{B} = -\mu_B \boldsymbol{B}\widehat{\sigma}, \tag{27.19}$$

where $\widehat{S} = \frac{\hbar}{2}\widehat{\sigma}$ is the spin vector operator in terms of Pauli matrices $\widehat{\sigma}_x$, $\widehat{\sigma}_y$, $\widehat{\sigma}_z$, $g = 2$ is the g-factor of the electron, and μ_B is the Bohr magneton, respectively [50H1].

When comparing (27.3) with (27.19) we conclude that any TLS is dynamically equivalent to a spin-1/2 system (zero energy such that $\epsilon_2 = -\epsilon_1$). The level-splitting between the ground state and the excited state of the atom corresponds to a constant magnetic field in z-direction, whereas the light field is equivalent to an oscillatory magnetic field in the x-direction. The expectation value of the (pseudo-) spin operator is closely related to the complex dipole moment and the inversion:

$$S_1 = <\widehat{\sigma}_x> = c_1^* c_2 + c_1 c_2^* \quad = 2\,\mathrm{Re}\,\mathcal{P}(t), \tag{27.20}$$

$$S_2 = <\widehat{\sigma}_y> = -ic_1^* c_2 + ic_1 c_2^* = 2\,\mathrm{Im}\,\mathcal{P}(t), \tag{27.21}$$

$$S_3 = <\widehat{\sigma}_z> = \mid c_1 \mid^2 - \mid c_2 \mid^2 \; = -I(t). \tag{27.22}$$

The vector $\boldsymbol{S} = (S_1, S_2, S_3)$ obeys the Bloch equation

$$\frac{\mathrm{d}\boldsymbol{S}(t)}{\mathrm{d}t} = \boldsymbol{\Omega} \times \boldsymbol{S}(t) + \dot{\boldsymbol{S}}_{rel}, \qquad \boldsymbol{\Omega} = (-2\omega_R(t), 0, -\omega_0)\ , \tag{27.23}$$

which describes a rotation of $\boldsymbol{S}$ around a vector $\boldsymbol{\Omega}$ at each instant of time. These equations are identical with the classical equations of motion of a spherical top in an external field (yet with fixed magnitude of the angular momentum) (see Fig. 27.3). The relaxation term $\dot{\boldsymbol{\Omega}}_{sc}$ is of similar structure as (27.17).

In RWA, the rotation vector $\boldsymbol{\Omega} = (-2\omega_R \cos\omega t, 0, -\omega_0)$, is replaced by

$$\boldsymbol{\Omega}_{RWA} = (-\omega_R \cos\omega t, \omega_R \sin\omega t, -\omega_0) \tag{27.24}$$

so that the notation eventually becomes obvious. The linearly polarized electromagnetic wave is decomposed into two counter-rotating circular (pseudo-magnetic) waves and only the resonant component is retained, which rotates with the same chirality as the spin vector.

In the absence of relaxation, the length of the Bloch vector is conserved, $|\boldsymbol{S}| = 1$ [which is equivalent to (27.18)] and its motion can be nicely visualized (see Fig. 27.3). We consider two limiting cases.

- For $\boldsymbol{E}(t) = 0$, $\boldsymbol{S}$ rotates on a cone around the z-axis, which is called the *Larmor precession*

$$\boldsymbol{S}_{\mathrm{L}}(t) = \left(s_0 \sin\omega_0 t, s_0 \cos\omega_0 t, 1 - s_0^2\right), \quad 0 \le |s_0| \le 1. \tag{27.25}$$

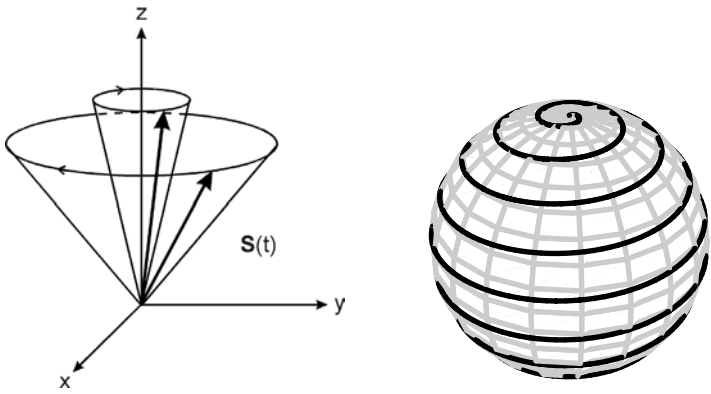


Fig. 27.3. Larmor precession of the Bloch vector in a constant magnetic field $\boldsymbol{B} \parallel z$ (left) and Rabi flopping due to an additional superimposed *ac* field along the 1-direction (right) (without relaxation)

- If the system is excited at a resonance from initial state $\boldsymbol{S}(0) = (0, 0, 1)$ the angle of the Larmor cone increases with time and the Bloch vector performs a spiral motion on the unit sphere from the north to the south pole and vice versa (Rabi flopping). In RWA, we have

$$\boldsymbol{S}_{\mathrm{R}}(t) = (\sin\omega_{\mathrm{R}}t \sin\omega_0 t, \sin\omega_{\mathrm{R}}t \cos\omega_0 t, \cos\omega_{\mathrm{R}}t)\,. \tag{27.26}$$

For Rabi frequencies comparable with or even larger than the atomic transition frequency, the regime of extreme nonlinear optics is reached. Obviously, the two-level approximation is no longer well justified in this regime but, nevertheless, it seems to describe the physics fairly well. For an overview of the amazing complexity and beauty of the nonlinear optical response see Sect. 19.3 and the paper by Tritschler et al. [03T1].

27.1.4 Linear Response of a Two Level System

We are looking for a solution of (27.14,27.15) in a linear approximation with respect to the light field and zero temperature. In this case $I(t) = -1$ in (27.14) so the problem of finding a solution reduces to

$$\left[\frac{\mathrm{d}}{\mathrm{d}t} + (\imath\omega_0 + \gamma)\right] \mathcal{P}(t) = -i\frac{d}{\hbar} E_z(t), \tag{27.27}$$

where $\gamma = 1/T_2$. For $\mathcal{P}(-\infty) = 0$ the solution reads

$$\mathcal{P}(t) = -i\frac{d}{\hbar} \int_{-\infty}^{t} e^{-(\imath\omega_0+\gamma)(t-t')} E_z(t')\,\mathrm{d}t'. \tag{27.28}$$

In particular, for a monochromatic field the dipole moment (27.4) becomes

$$d(t) = \mathrm{Re}\, \alpha(\omega) E_0\, e^{-\imath\omega t}, \tag{27.29}$$

$$\alpha(\omega) = \frac{|d|^2}{\hbar} \left[\frac{1}{\omega_0 - (\omega + \imath\gamma)} + \frac{1}{\omega_0 + (\omega + \imath\gamma)} \right]. \tag{27.30}$$

$\alpha(\omega)$ is the polarizability of the TLS. Remarkably, this result is almost identical with the polarizability of a classical harmonic oscillator (e.g., a particle with mass m and charge e bound to a spring with $D = m\,\omega_0^2$, see Chap. 4).

$$\alpha_{cl}(\omega) = \frac{\frac{e^2}{m}}{\omega_0^2 - \omega^2 - 2\imath\gamma\omega} \tag{27.31}$$

In contrast to (27.30), the resonance frequency of a classical oscillator shifts with increasing damping to lower values $\sqrt{\omega_0^2 - \gamma_2^2}$.

27.2 Optical Bloch Equations

As sketched in Fig. 27.4 the generalization of the optical Bloch equations to the case of a two-band semiconductor is straightforward. Neglecting the momentum of the absorbed photon, the optical transitions are vertical in $\boldsymbol{k}$-space. Under homogeneous conditions and neglect of scattering processes (between different $\boldsymbol{k}$-states in the same band) a two-band semiconductor is just an assembly of uncoupled TLSs and resembles the case of a inhomo-

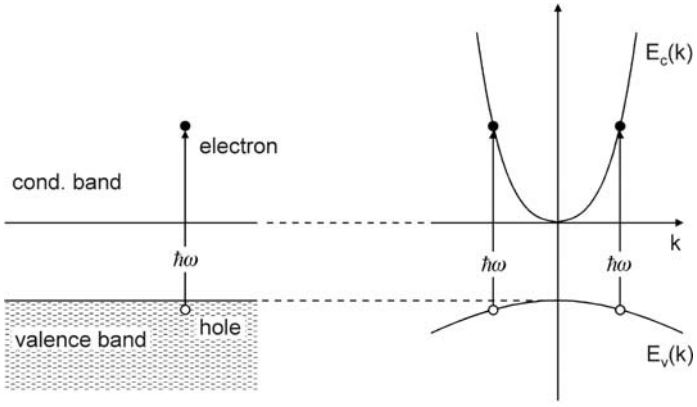


Fig. 27.4. Two noninteracting bands viewed as a collection of TLSs with different transition energies

geneously broadened line problem in atomic physics. The equations of motion are:

$$i\hbar\frac{\partial\mathcal{P}(\boldsymbol{k},t)}{\partial t} = \Big[E_c(\boldsymbol{k}) - E_v(\boldsymbol{k})\Big]\mathcal{P}(\boldsymbol{k},t) + \boldsymbol{d}_{cv}(\boldsymbol{k})\boldsymbol{E}(t)\Big[n_c(\boldsymbol{k},t) - n_v(\boldsymbol{k},t)\Big] + i\hbar\dot{\mathcal{P}}_{rel}, \tag{27.32}$$

$$\frac{\partial n_c(\boldsymbol{k},t)}{\partial t} = -\,2\,\mathrm{Im}\,\Big[\boldsymbol{d}_{cv}(\boldsymbol{k})\boldsymbol{E}(t)\mathcal{P}^*(\boldsymbol{k},t)\Big] + \dot{n}_c^{rel}, \tag{27.33}$$

$$\frac{\partial n_v(\boldsymbol{k},t)}{\partial t} = +\,2\,\mathrm{Im}\,\Big[\boldsymbol{d}_{cv}(\boldsymbol{k})\boldsymbol{E}(t)\mathcal{P}^*(\boldsymbol{k},t)\Big] + \dot{n}_v^{rel}. \tag{27.34}$$

These are the optical Bloch equations. n_c and n_v are the mean occupation numbers (per spin) of the conduction and valence band Bloch states. In addition, particle conservation requires $n_c + n_v = 1$. In a phenomenological description, relaxation may be included analogously to (27.17). For a qualitative treatment and allowed (dipole) transitions, the $\boldsymbol{k}$-dependence of the interband dipole matrix element $\boldsymbol{d}_{\mathrm{cv}}(\boldsymbol{k}) \propto \langle c, \boldsymbol{k}\,|\widehat{\boldsymbol{p}}|\,v, \boldsymbol{k}\,\rangle$ may often be neglected near the band edge. The $\boldsymbol{k}$-dependence is only relevant for forbidden transitions or if the variation of the polarization over the whole Brillouin-zone is needed. Note that these equations are uncoupled with respect to the wave number $\boldsymbol{k}$.

From $\mathcal{P}(\boldsymbol{k},t)$ the electronic polarization of the semiconductor can be obtained from

$$\boldsymbol{P}(t) = \frac{2}{V}\,\mathrm{Re}\sum_{\boldsymbol{k},s}\boldsymbol{d}_{cv}^*(\boldsymbol{k})\,\mathcal{P}(\boldsymbol{k},t), \tag{27.35}$$

where V denotes volume of the sample, which eventually drops out when performing the sum over wavenumbers in terms of the integral

$$\frac{1}{V}\sum_{\boldsymbol{k},s}\ldots = 2\frac{1}{(2\pi)^d}\int\ldots d^d\boldsymbol{k}, \tag{27.36}$$

where $d = 1, 2, 3$ specifies the space-dimension and the factor 2 arises from spin summation.

27.2.1 Optical Susceptibility: Interband Transitions

Next, we discuss the linear response result for the optical susceptibility of an undoped semiconductor at low temperatures and neglecting crystal anisotropy. Here, n_c and n_v can be replaced by the (equilibrium) Fermi functions $f_c(k) \approx 0$, $f_v(k) \approx 1$ for the electrons in the conduction and valence bands. As $\boldsymbol{k}$ is merely a parameter in this approximation, the required solution of (27.32) can be found by the ansatz

$$\mathcal{P}(\boldsymbol{k},t) = Q(\boldsymbol{k},t)\exp\Big[-\frac{i}{\hbar}\left(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k})\right)t\Big] \tag{27.37}$$

with a simple integration of $Q(\boldsymbol{k},t)$. Eventually, we read off the electrical susceptibility $\chi(\omega)$ from the Fourier components of $\boldsymbol{P}(t)$ and $\boldsymbol{E}(t)$ (see Prob. 3):

$$\boldsymbol{P}(\omega) = \epsilon_0 \chi(\omega)\, \boldsymbol{E}(\omega), \tag{27.38}$$

$$\chi(\omega) = \frac{1}{V\epsilon_0} \sum_{k,s} |d_{cv}(\boldsymbol{k})|^2 \left\{ \frac{f_v(\boldsymbol{k}) - f_c(\boldsymbol{k})}{E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) - \hbar(\omega + i\eta)} + \right.$$
$$\left. + \frac{f_v(\boldsymbol{k}) - f_c(\boldsymbol{k})}{E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) + \hbar(\omega + i\eta)} \right\}, \quad \eta \to +0. \tag{27.39}$$

For parabolic bands and $\boldsymbol{k}$-independent dipole matrix elements d_{cv} (along the field direction) the absorptive part of the susceptibility is proportional to the joint density of states of the valence and conduction bands. In the $d = 1, 2, 3$ dimensions, the (joint) density of states, and hence the absorption, rises as $\mathrm{Im}\ \chi(\omega) \propto [\hbar\omega - E_g]^{(d-2)/2}$ (see also Sect. 7.7 and 8.9).

27.3 Semiconductor Bloch Equations

We are close to the summit of our tour towards SBEs. Two features have not yet been taken into account:

- There is a change in Coulomb and exchange energy of the interacting many-electron (ground) state when exciting an electron to the conduction band and leaving a hole behind. The main part of this interaction turns out to be an attractive Coulomb potential, which adds to the external electrical field (also see Chap. 9).
- With increasing band filling there is a renormalization of the electron/hole band energy by the electron/hole Coulomb interaction (also see Chap. 21).

$$i\hbar \frac{\partial \mathcal{P}(\boldsymbol{k},t)}{\partial t} = \Big[E_g + E_e(\boldsymbol{k}) + E_h(\boldsymbol{k}) \Big] \mathcal{P}(\boldsymbol{k},t)$$
$$+ \Big[n_e(\boldsymbol{k},t) + n_h(\boldsymbol{k},t) - 1 \Big] \hbar\Omega_{\mathrm{R}}(\boldsymbol{k},t) + i\hbar \dot{\mathcal{P}}_{rel}, \tag{27.40}$$

$$\frac{\partial n_c(\boldsymbol{k},t)}{\partial t} = -2\,\mathrm{Im} \left\{ \Omega_{\mathrm{R}} \mathcal{P}^*(\boldsymbol{k},t) \right\} + \dot{n}_c^{rel}, \tag{27.41}$$

$$\frac{\partial n_h(\boldsymbol{k},t)}{\partial t} = -2\,\mathrm{Im} \left\{ \Omega_{\mathrm{R}} \mathcal{P}^*(\boldsymbol{k},t) \right\} + \dot{n}_h^{rel} \tag{27.42}$$

For convenience, the change in population of the valence band is formulated within the hole picture (as indicated by the index h):

$$n_v(\boldsymbol{k},t) = 1 - n_h(\boldsymbol{k},t), \qquad E_v(\boldsymbol{k},t) = -E_g - E_h(\boldsymbol{k},t). \tag{27.43}$$

The factor $n_e(\boldsymbol{k},t) + n_h(\boldsymbol{k},t) - 1$ is again the population inversion at $\boldsymbol{k}$. Its influence on the optical absorption is often denoted as phase space filling or

Pauli-blocking. The term $-2\mathrm{Im}\{\Omega_R \mathcal{P}^*(\boldsymbol{k},t)\}$ in (27.41, 27.42) describes the generation of electron hole pairs by the absorption of light.

$E_e(\boldsymbol{k},t)$ and $E_h(\boldsymbol{k},t)$ are the electron/hole (Hartree-Fock) energies including the interaction with other electrons/holes. For parabolic bands these are:

$$E_j(\boldsymbol{k},t) = \frac{\hbar^2\boldsymbol{k}^2}{2m_j} - \frac{1}{V}\sum_{\boldsymbol{q}\ (\neq\boldsymbol{k})} V(\boldsymbol{k}-\boldsymbol{q})\, n_j(\boldsymbol{q},t), \quad j = e,h. \tag{27.44}$$

Note, $m_h > 0$.

$\Omega_{\mathrm{R}}(\boldsymbol{k},t)$ denotes the generalized Rabi frequency (function)

$$\hbar\Omega_{\mathrm{R}}(\boldsymbol{k},t) = \boldsymbol{d}_{\mathrm{cv}}(\boldsymbol{k})\boldsymbol{E}(t) + \frac{1}{V}\sum_{\boldsymbol{q}\ (\neq\boldsymbol{k})} V(\boldsymbol{k}-\boldsymbol{q})\, \mathcal{P}(\boldsymbol{q},t), \tag{27.45}$$

where $V(\boldsymbol{q})$ is the Fourier transform of the (repulsive) Coulomb potential screened by a background dielectric constant ϵ_b

$$V(\boldsymbol{q}) = \frac{e^2}{\epsilon_0\epsilon_b\, q^2}, \qquad V(\boldsymbol{r}) = \frac{e^2}{4\pi\epsilon_0\epsilon_b\, r}. \tag{27.46}$$

Going beyond the time-dependent Hartree-Fock approximation is not easy and requires higher-order correlation functions. Physically, this leads to a further renormalization of interactions and energies, which are described by four, six, or more particle (e.g., biexcitons) amplitudes, and a microscopic formulation of the relaxation terms (dephasing of the polarization, collisions of electrons and holes) [88Z1, 94H1, 96H1]. An important mechanism in this respect is the dynamical screening of the (e–e, h–h, e–h) Coulomb interactions by the excited carriers in terms of a (Lindhard-) dielectric function $\epsilon_\ell(\boldsymbol{q},\omega)$. Such contributions are of increasing importance for high excitation and ultrashort pulses. Moreover, in addition to the purely electronic interactions there are other scattering mechanisms, such as carrier-phonon scattering or scattering of the carriers by impurities and interface roughness in spatially confined systems. For very short times, memory effects of the scattering processes come into play so that the scattering integrals in the Boltzmann equation for n_e, n_h become non-Markovian.

Further refinements of the theory may also include some details of the realistic semiconductor band-structure (e.g., more than two bands, heavy hole-light hole splitting, band-warping, etc.) or spatial dispersion. This, however, requires a huge numerical effort. At this level one is generally left with an approximate numerical treatment and the problem is considered to be understood (or solved) if these results agree well with the experimental findings.

27.3.1 Optical Susceptibility: Excitons

To demonstrate the simplicity as well as the potential of SBEs (e.g., as compared with Elliott's evaluation of the Kubo formula [57E1]) we discuss the

linear optical susceptibility of the interacting electron-hole system including exciton states. At low excitation and zero temperature we have $n_v = n_h \approx 0$ so that (27.40-27.42) reduce to:

$$i\hbar\frac{\partial\mathcal{P}(\boldsymbol{k},t)}{\partial t} = \Big[E_g + \frac{\hbar^2\boldsymbol{k}^2}{2\mu_r}\Big]\mathcal{P}(\boldsymbol{k},t) - \frac{1}{V}\sum_{q(\neq\boldsymbol{k})} V(\boldsymbol{k}-\boldsymbol{q})\mathcal{P}(\boldsymbol{q},t) - \boldsymbol{d}_{cv}\boldsymbol{E}(t). \quad (27.47)$$

μ_r denotes the reduced electron hole mass. As opposed to the optical Bloch equations, (27.32-27.34), $\boldsymbol{k}$ is no longer just a parameter; different $\boldsymbol{k}$s are coupled by $V(\boldsymbol{k}-\boldsymbol{q})$. However, this interaction term is a convolution type and, by Fourier-transformation, it reveals as the (inhomogeneous) Schrödinger equation of the hydrogen atom

$$i\hbar\frac{\partial\mathcal{P}(\boldsymbol{r},t)}{\partial t} = \Big[E_g - \frac{\hbar^2}{2\mu_r}\Delta\Big]\mathcal{P}(\boldsymbol{r},t) - V(\boldsymbol{r})\mathcal{P}(\boldsymbol{r},t) - \boldsymbol{d}_{cv}\,\boldsymbol{E}(t)\,\delta(\boldsymbol{r}). \quad (27.48)$$

Equation (27.48) is also known as the Wannier-Equation. $\mathcal{P}(\boldsymbol{r},t)$ plays the part of the exciton (envelope) wave function. The $\boldsymbol{k}$-dependence of the interband matrix element $\boldsymbol{d}_{cv}(\boldsymbol{k})$ has been neglected.

For $\boldsymbol{E} = 0$ the stationary states are well known from standard texts on quantum mechanics

$$\mathcal{P}_\alpha(\boldsymbol{r},t) = e^{-i\frac{E_\alpha}{\hbar}t}\,\Psi_\alpha(\boldsymbol{r}), \quad E_\alpha = -\frac{Ry^*}{n^2}, \quad Ry^* = Ry\,\frac{\mu_r/m_0}{\epsilon_b^2}. \quad (27.49)$$

Ry $= 13.6\ldots$ eV is the hydrogenic Rydberg energy.[5] Quantum numbers $\alpha = (n,l,m)$ have their usual meaning and, formally, n labels the discrete, as well as the continuum, states. In a semiconductor, these (bound) hydrogenic states are called *excitons*, which have been extensively discussed in previous chapters, (in particular see Chaps. 9, 13 and the references cited therein).

For $\boldsymbol{E}(t) \neq 0$ (parallel to the z-direction) we seek the solution of $\mathcal{P}(\boldsymbol{r},t)$ in terms of the complete set of the stationary states as given by (27.49),

$$\mathcal{P}(\boldsymbol{r},t) = \sum_\alpha Q_\alpha(t)\mathcal{P}_\alpha(\boldsymbol{r},t), \quad (27.50)$$

where $Q_\alpha(t)$ can be found by the simple integration

$$Q_\alpha(t) = i\frac{d_{\rm cv}}{\hbar}\Psi^*_\alpha(\boldsymbol{r}=0)\int_{-\infty}^{t} E_z(t')\,e^{i\frac{E_\alpha}{\hbar}t'}e^{\eta t'}\,{\rm d}t', \quad \eta\to 0^+. \quad (27.51)$$

$\exp(\eta t)$ has been added to switch on the light at $t=-\infty$ adiabatically. Obviously, the (spatially constant) light-field creates only those exciton states that have a nonvanishing wave function at the origin; these are the s-states.[6]

[5] If the excitonic Rydberg Ry* is smaller than the LO-phonon energy, ϵ_b is given by the static dielectric constant ϵ_s and otherwise by ϵ_∞ (see also Sect. 9.2).

[6] This result originates from our approximation that $\boldsymbol{d}_{\rm cv}(\boldsymbol{k}) = {\rm const}$. If the dipole element vanishes linearly with $\boldsymbol{k}$ at the band edge, the coupling is solely to p-states. A prominent example is Cu_2O (see Sect. 13.2.1.2 and Prob. 4.)

Inserting (27.50) in (27.36) and using the completeness relation of the stationary states (27.49), we finally obtain the electrical susceptibility (isotropic approximation, $d_{\rm cv}$ along $\boldsymbol{E}$)

$$\chi(\omega) = 2\frac{|d_{\rm cv}|^2}{\epsilon_0} \sum_\alpha |\Psi_\alpha(\boldsymbol{r}=0)|^2 \left\{ \frac{1}{E_g + E_\alpha - \hbar(\omega + i\eta)} + \right.$$
$$\left. + \frac{1}{E_g + E_\alpha + \hbar(\omega + i\eta)} \right\}. \tag{27.52}$$

In a bulk semiconductor the absorptive part consists of a series of delta-functions at the bound exciton s-states with rapidly decreasing oscillator strength $\propto n^{-3}$ and a continuum part [94H1]

$$\text{Im}\ \chi(\omega) \propto \sum_{n=1}^{\infty} \frac{4\pi}{n^3} \delta(\Delta + \frac{1}{n^2}) + \theta(\Delta)\, \frac{\pi \exp\left(\frac{\pi}{\sqrt{\Delta}}\right)}{\sinh(\frac{\pi}{\sqrt{\Delta}})}. \tag{27.53}$$

$\Delta = (\hbar\omega - E_g)/\text{Ry}^*$ denotes the normalized photon energy. Close to the ionization continuum, $\Delta \to 0$, the absorption increases stepwise in striking distinction to the square root law for noninteracting bands (cf. (27.39)). Thus, the attractive Coulomb interaction not only creates bound states below the gap but leads to a pronounced enhancement of the absorption above the gap (the so-called Sommerfeld-enhancement). Equation (27.53) was first derived by Elliott [57E1] but in a much more elaborate way. The corresponding real part can be calculated from the Kramers–Kronig relations (see Chap. 6).

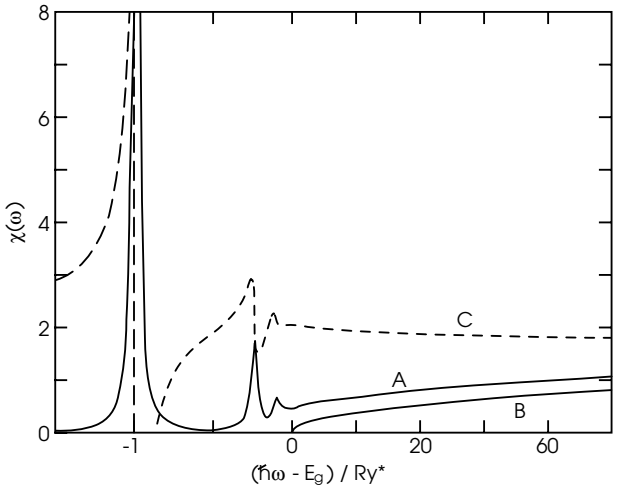


Fig. 27.5. Electrical susceptibility of a direct semiconductor near the gap energy. Im χ (A) and Re χ (B) of the interacting electron hole system, (27.53). Im χ (B) of the noninteracting system, (27.39). Note the different frequency scale above and below the gap. An appropriate broadening was included phenomenologically [87S1]

The optical absorption spectrum bulk GaAs is displayed in Fig. 27.5, which gives a lively impression of the importance of the Coulomb interaction and exciton states near the band gap. In the very best samples, exciton lines up to $n = 3$ can be resolved in this material (see Chap. 13).

27.4 Some Selected Coherent Processes

An important class of optical experiments involves two or three short pulses impinging from different directions and at different times on the sample (see Fig. 27.6 and Sects. 23.2.1.1 and 25.2.2.3). From the analysis of such studies on the dynamics and coherence properties, such as T_1, the T_2 times are obtained. In terms of nonlinear optics, the processes in question are of third order with respect to the light

$$P_j = \chi^{(3)}_{jklm} E_k E_l E_m . \tag{27.54}$$

The products are in fact convolutions with respect to time, and j, k, l, m denote Cartesian components of the (total) electrical light field. One of these pulses, shown in Fig. 27.6, may have a twofold effect. The following sections discuss some typical examples.

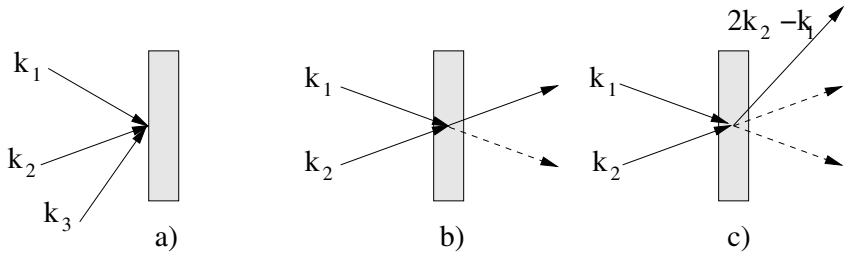


Fig. 27.6. (**a**) Sketch of some third-order nonlinear processes. (**b**) Pump-probe (**c**) Four-wave mixing experiment

27.4.1 Pump-Probe Experiments

In a pump-probe experiment the first pulse (pump) pre-excites the system. After a delay time τ the linear optical (absorption, reflection) response with a second pulse (probe) is measured. As a reference, we give the result for the optical Bloch equations, where a $\delta(t)$ shaped pulse #1 induces the following change in the optical absorption

$$\delta\alpha(\omega;\tau,I_1) = \alpha(\omega;\tau,I_1) - \alpha_0(\omega) \propto -I_1\, e^{\tau/T_1}\, \alpha_0(\omega)\ . \tag{27.55}$$

Hence, pump-probe experiments reflect the influence of the excited carriers and their relaxation on the optical properties. $\alpha_0(\omega)$ is the absorption coefficient of the system in equilibrium. In Fig. 27.6b the electrical field of pulse #1 enters twice because $I_1 \propto E_1^2$.

27.4.2 Four-Wave Mixing Experiments

In a wave mixing experiment the nonlinear response is studied in a so-called background free direction, which is different from those of the incident pulses and their transmission or reflection, which, of course, is experimentally favorable (see also Chap. 25). The first pulse excites a coherent polarization whose radiation field interferes with the second pulse forming an interference pattern with wave vector $\boldsymbol{k}_1 - \boldsymbol{k}_2$. This pattern, in turn, modulates the (linear) optical properties of the system and, thus, diffracts off the second pulse in direction $2\boldsymbol{k}_2 - \boldsymbol{k}_1$ (other directions are less efficient). This phenomenon is called (degenerate) four-wave mixing (FWM) as four waves are involved: #1, #2 (counted twice!) and the wave emitted in the direction $2\boldsymbol{k}_2 - \boldsymbol{k}_1$ (see Fig. 27.6c).

For a general overview of nonlinear coherent processes in semiconductors, see Haug and Koch [94H1], Shah [96S1], or Schäfer and Wegener [02S1]. A review on the coherent spectroscopy of semiconductor microcavities has been given by Lyssenko et al. [05L1]. Experimental results for typical semiconductor systems are discussed in Chap. 23.

27.4.3 Photon Echo

The analogy of a TLS with the spin-1/2 problem offers an unexpected insight into an interesting phenomenon called *photon echo*. Here, we examine the rather marvelous notion that not all macroscopic decay processes are irreversible.[7] In a photon echo experiment, many TLS are usually involved and because of different local environments these have individually slightly different transition frequencies (inhomogeneous line broadening, spectral width parameterized by $1/T_2^*$). Hence, the macroscopic (magnetization or polarization) signal contains many contributions with slightly different frequencies that add up almost to zero in a rather short time ($\approx T_2^*$), yet the individual systems are still oscillating with a fixed phase relation with respect to the exciting pulse and the energy stored in the moments being out of phase can be recovered in a coherent fashion. Instead of waiting for the extremely unlikely Poincaré recurrence of the initial state, a second pulse after time T is used, which – loosely speaking – causes a time reversal operation followed by an echo of the initial pulse after total time $2T$.

[7] This technique was developed by Hahn [50H1] for nuclear spin systems and it plays an important role in measuring the T_2-time. For a survey and thorough discussion we refer readers to Chap. 9 of Allen and Eberly [87A1] and to Chap. 23 of this book for experimental results.

To uncover the echo phenomenon it is convenient to describe the dynamics of the Bloch vector in a frame rotating with the frequency of the light (not the atomic transition frequency!) around the 3-axis:

$$R_1(t) = S_1(t)\cos\omega t - S_2(t)\sin\omega t, \tag{27.56}$$
$$R_2(t) = S_1(t)\sin\omega t + S_2(t)\cos\omega t, \tag{27.57}$$
$$R_3(t) = S_3(t). \tag{27.58}$$

(In complex notation the $1,2$ components can be summarized by $R = R_1 + iR_2$, $S = S_1 + iS_2$, with $R = Se^{i\omega t}$.) In this frame the equations of motion read

$$\dot{R}_1(t) = -\delta R_2(t) - \frac{R_1(t)}{T_2}, \tag{27.59}$$
$$\dot{R}_2(t) = +\delta R_1(t) + \omega_R R_3(t) - \frac{R_2(t)}{T_2}, \tag{27.60}$$
$$\dot{R}_3(t) = -\omega_R R_2(t) - \frac{R_3(t) - R_3^{eq}}{T_1}. \tag{27.61}$$

These linear differential equations have constant coefficients so that the solution can be found by an exponential ansatz. In particular, for $\omega_R = 0$ the Bloch vector performs a (damped) Larmor-precession around the 3-axis:

$$R_1(t) = [R_1(0)\cos\delta t - R_2(0)\sin\delta t]\; e^{-\frac{t}{T_2}}, \tag{27.62}$$
$$R_2(t) = [R_1(0)\sin\delta t + R_2(0)\cos\delta t]\; e^{-\frac{t}{T_2}}, \tag{27.63}$$
$$R_3(t) = R_3^{eq} + [R_3(0) - R_3^{eq}]\; e^{-\frac{t}{T_1}}. \tag{27.64}$$

At resonance, $\delta = 0$, but by neglecting damping a light pulse causes a Rabi rotation of the Bloch vector around the 1-axis

$$R_1(t) = +R_1(0), \tag{27.65}$$
$$R_2(t) = +R_2(0)\cos\omega_R t + R_3(0)\sin\omega_R t, \tag{27.66}$$
$$R_3(t) = -R_2(0)\sin\omega_R t + R_3(0)\cos\omega_R t. \tag{27.67}$$

In the discussion of a photon echo experiment four periods have to be distinguished (see Fig. (27.7)). To simplify matters, we shall assume that the pulse duration is short with respect to T_1, T_2, T_2^* and intense, $\omega_R T_2^* \gg 1$, so that the influence of damping and detuning can be safely neglected during the pulses.

1. All TLSs start from the same initial state $\mathbf{R}^{(0)} = (0,0,1)$. Then the ensemble of atoms is polarized by an initial light pulse (duration τ_1), which, according to (27.65–27.67), leads to a common Bloch vector $\mathbf{R}^{(1)} = (u,v,w)$. For a so-called $\pi/2$-pulse, i.e., $\phi_1 = \omega_R\tau_1 = \pi/2$, the polarization (magnetization) is "tipped" to the 2-direction: $\mathbf{R}^{(1)} = (0,1,0)$ (see Fig. 27.7a).
2. After the first light pulse, the individual Bloch vectors $\mathbf{R}^{(2)}$ precess according to (27.62–27.64). Because of their slightly different frequencies the individual dipole moments get out of phase (relative to each other) and

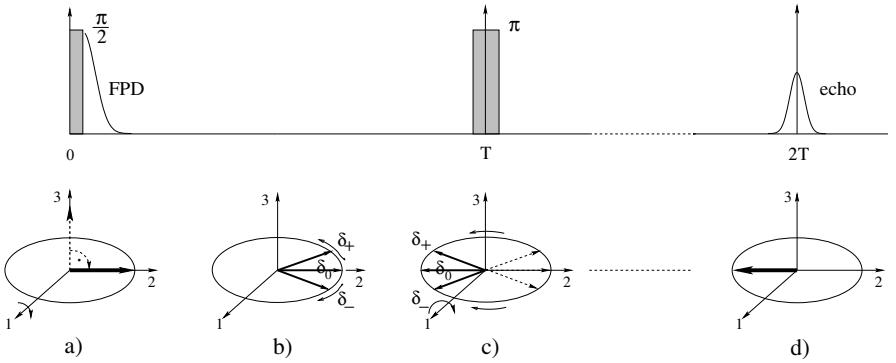


Fig. 27.7. Principle of a photon echo experiment. *Upper part*: Standard pulse sequence (*dashed blocks*, exaggerated duration) and resulting polarization (*solid lines*). *Lower part*: Dynamics of the Bloch vector in the rotating frame. (**a**) Alignment of the individual Bloch vectors along the 2-direction by a $\pi/2$-rotation around the 1-axis. (**b**) Precession of the individual Bloch vectors in the 1-2 plane as determined by their various detunings leading to free polarization decay (FPD). (**c**) π-rotation around the 1-axis transforms $R_2 \to -R_2$ but leaves R_1 unchanged. (**d**) Reproduction of the initial state at time T after the π-pulse (For simplicity only three Bloch vectors are displayed with detunings $\delta_+ > 0$, $\delta_0 = 0$, $\delta_- < 0$)

add to zero in a time T_2^*, which is much shorter than T_2. This dephasing phenomenon is called *free polarization decay* (FPD) and is depicted in Fig. 27.7b. (Since we are in a rotating frame only the phase deviations relative to $\exp(i\omega t)$ appear in this figure.)

3. After time T a second light pulse is applied (duration τ_2, phase $\phi_2 = \omega_R \tau_2$), which, according to (27.65, 27.66), tips the polarization to $\mathbf{R}^{(3)}$.
4. After the second pulse the Bloch vectors $\mathbf{R}^{(4)}$ again rotate freely around the 3-axis. As a result, we obtain for the 2-component

$$R_2^{(4)}(t) = \Big\{ u\left[\cos\delta T \sin\delta t + \cos\phi_2 \sin\delta T \cos\delta t\right] \tag{27.68}$$

$$-v\left[\sin\delta T \sin\delta t - \cos\phi_2 \cos\delta T \cos\delta t\right] \Big\} e^{-\frac{t+T}{T_2}} \tag{27.69}$$

$$+ \sin\phi_2 \cos\delta t \left[1 + (w-1)\, e^{-T/T_1}\right] e^{-T/T_2}, \tag{27.70}$$

where the time t counts from the end of the second pulse. For $\phi_2 = \pi$ the terms in the [..] brackets combine to $\cos\delta(T-t)$ and $\sin\delta(T-t)$ and all individual Bloch vectors are again in phase at time $2T$ after the first pulse and add up to a macroscopic polarization. This causes the emission of a light pulse, the photon echo.

The resurrected free polarization signal has the magic quality of something coming from nothing. However, this resurrection is only possible for times comparable with the T_2 time. For larger times, the intensity of the photon

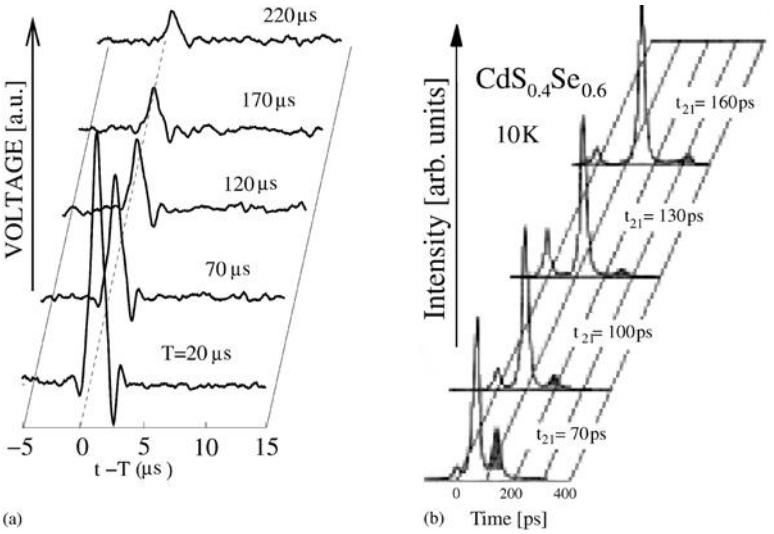


Fig. 27.8. Experimental spin echo for protons in water ($\omega/2\pi = 95$MHz, duration of π-pulse: 1.4μs, $T_2 \approx 75\mu$s, $T_2^* \approx 1\mu$s, $T_1 \approx 10$s, room temperature) [03F1] (*left*) and photon echo in a II-VI semiconductor compound [92S1](*right*)

echo decays as the square of $\exp(-2T/T_2)$, i.e., $\exp(-4T/T_2)$. T_2 is also called the *decoherence or dephasing time*.

It is interesting that the existence of the echo is not attributed to the π-character of the second pulse as it is frequently imputed. When decomposing the products of trigonometric functions $\cos\delta T \cos\delta t$, etc., in terms of sum and differences, we realize that there is a contribution to the polarization of the form:

$$R_2^{(4)} = -\frac{1-\cos\phi_2}{2}\Big\{u\sin\left[\delta(T-t)\right] + v\cos\left[\delta(T-t)\right]\Big\}e^{-\frac{t+T}{T_2}} + \ldots \quad (27.71)$$

Thus, a second pulse of *any duration* will induce an echo. However, its intensity is largest for $\phi_2 = \pi$ (and $\phi_1 = \pi/2$). (Terms omitted in (27.71) indicate nonecho terms).

Although spin and photon echo phenomena have the same theoretical roots, there are profound experimental differences.

- In semiconductors, photon echoes are generated with light pulses that correspond to phase shifts much smaller than π. $\Phi = \pi$ would lead to a complete band inversion, which in turn would lead to a collapse of the semiconductor band structure. This regime is hardly accessible and would generate additional fast dephasing processes.
- Spin echoes are generated with rf or microwave fields with wavelengths much larger than the dimensions of the specimen. The situation is opposite for photon echoes in semiconductors where standard experimental

techniques, e.g., four-wave mixing, uses pulses traveling in different directions so that the echo signal can be spatially separated from the exciting pulses (details are discussed in Chaps. 23 and 25).

A popular analogy of the spin echo phenomenon and a group of runners in a stadium as discussed in Chap. 9 of Allen and Eberly's book [87A1] and in Sect. 23.2 of this book. (However, it should be kept in mind that π-rotation around the 1-axis at time T causes a reconfiguration of the pulse, analogously to optical phase conjugation, rather than a reversal of time of the individual Bloch vectors.)

Experimentally, photon echo measurements are often used as a tool to determine the optical dephasing rate T_2^{-1}. In semiconductor systems, however, the (dynamically screened) interactions of the electron hole pairs may lead to significant deviations from the atomic case and the analysis of experimental results requires a detailed numerical study of the SBEs, which was first undertaken by Lindberg et al. [92L1]. For an overview and details see Haug and Koch [94H1] and Schäfer and Wegener [02S1]. An easily readable review of spin echo and photon echo in semiconductors, as well as an overview on related echo phenomena, was recently published by Meier and Thomas [03M1].

A preliminary version of this chapter was published in the proceedings of an Erice Summer School [98B1].

27.5 Problems

1. Solve (27.12) for arbitrary detuning and calculate the wave function coefficients $c_j(t)$ (or $a_j(t)$), inversion $I(t)$, and complex polarization $\mathcal{P}(t)$. Use the initial conditions $c_1(0) = 1$ and $c_2(0) = 0$.
2. At resonance there are stationary states of the coupled TLS electrical field, (27.12), with time-independent probabilities $|c_j(t)|^2$=const. Find these "dynamic Stark-states."
3. Use the ansatz proposed by (27.37) to solve (27.32) for a monochromatic field and derive the interband susceptibility $\chi(\omega)$ given by (27.39). Hint: In the absence of damping, the light field must be switched on adiabatically at $t = -\infty$, $\boldsymbol{E}(t) = E_0 \exp(\eta t) \cos(\omega t)$, $\eta \to +0$. Otherwise the response contains spurious contributions from the physically irrelevant initial state.
4. If the transitions near the band-edge are forbidden (in first order) the dipole matrix element behaves as $\boldsymbol{d}_{cv}(\boldsymbol{k}) \propto \boldsymbol{k}$. Show that the optical absorption is now linked to exciton p-states.
   Hints: For $r \to 0$ hydrogenic wave functions behave as $\Psi_{n,l,m}(r, \theta, \phi) \propto r^l$. Under Fourier transformation: grad $\to \imath \boldsymbol{k}$.
5. Find the general solution of the Bloch equations for the photon echo problem (27.59–27.61). Express the integration constants in terms of the Bloch vector at $t = 0$, $\mathbf{R}(0)$.
   Hint: The solution can be obtained by an exponential ansatz with $\mathbf{R}(t) = \boldsymbol{\rho} \exp(\lambda t)$, where $\boldsymbol{\rho}$ is a time-independent 3-component vector.

6. Consider a photon echo with pulses of equal duration, i.e. $\Phi = \Phi_1 = \Phi_2$. Show that the echo signal has a maximum for $\Phi = 2\pi/3$ and even a bit stronger than that for the $\pi/2$-π sequence.

References to Chap. 27

[46B1] F. Bloch, Phys. Rev. **70**, 460 and 474 (1946).

[50H1] E. L. Hahn, Phys. Rev. **80**, 580 (1950).
See also: *Nonlinear Spectroscopy of Solids, Advances and Applications*, B. Di Bartolo ed., Nato ASI Series B, Vol. 339, page 75 (Plenum, New York and London 1994).

[57E1] R. J. Elliott, Phys. Rev. **108**, 1384 (1957). See also: *Polarons and Excitons*, edited by C. G. Kuper and G. D. Whitfield (Oliver and Boyd, Edinburgh and London 1963).

[64F1] R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics*, Vol. III (Addison-Wesley, Reading, Massachusetts, Palo Alto, London 1964).

[70H1] H. Haken, Handbuch der Physik, S. Flügge ed., Vol. XXV/2c, *Light and Matter 1c* (Springer, Berlin, Heidelberg, New York 1970).

[79H1] H. Haken, *Licht und Materie*, Vol. 2 (Bibliographisches Institut, Mannheim, Wien, Zürich 1979).

[84S1] A. Stahl, Solid State Commun. **49**, 91 (1984).

[87A1] L. Allen, J. H. Eberly, *Optical Resonance and Two-Level Atoms* (Dover, New York 1987).

[87S1] A. Stahl, I. Balslev, *Electrodynamics of the Semiconductor Band Edge*, Springer Tracts in Modern Physics, Vol.110 (Springer, Berlin 1987).

[88Z1] R. Zimmermann, *Many Particle Theory of Highly Excited Semiconductors*, (Teubner, Leipzig 1988).

[92B1] R. W. Boyd, *Nonlinear Optics* (Academic Press, Boston 1992).

[92L1] M. Lindberg, R. Binder, S. W. Koch, Phys. Rev. A **45**, 1865 (1992).

[92S1] U. Siegner, D. Weber, E. O. Göbel, D. Bennhardt, V. Heukeroth, R. Saleh, S. D. Baranovskii, P. Thomas, H. Schwab, C. Klingshirn, J. M. Hvam, V. G. Lyssenko, Phys. Rev. **B** 46, 4564 (1992).

[94H1] H. Haug, S. W. Koch, *Quantum Theory of the Optical and Electronic Processes of Semiconductors* (World Scientific, Singapore, New Jersey, London, Hong Kong 1994).

[96H1] H. Haug, A-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors* (Springer, Berlin, Heidelberg 1996).

[96S1] J. Shah, *Ultrafast Spectroscopy of Semiconductor and Semiconductor Nanostructures*, Springer Series in Solid State Sciences, **115**, (1996).

[98B1] R. v. Baltz, in: *Ultrafast Dynamics of Quantum Systems: Physical Processes and Spectroscopic Techniques*, B. Di Bartolo ed., Nato ASI series B, Vol. 372 (Plenum, New York and London 1998).

[02S1] W. Schäfer, M. Wegener, *Semiconductor Optics and Transport Phenomena*, Advanced Texts in Physics (Springer, Berlin 2002).

[03F1] F. Fujara, Ch. Köhler, private communication. Institut für Festkörperphysik, Technische Universität Darmstadt.

[03M1] T. Meier, P. Thomas, *Echoes in Festkörpern*, Physik Journal, März issue, p 53, Wiley-VCH (2003).
[03T1] T. Tritschler, O. D. Mücke, M. Wegener, Phys. Rev. A **68**, 33404 (2003).
[05B1] B. Di Bartolo, in: *Investigating Extreme Physical Conditions with Advanced Optical Techniques*, B. Di Bartolo ed., Nato ASI series II (Kluwer, Dordrecht, Boston, London, to be published).
[05L1] V. Lyssenko, D. Meinhold, B. Rosam, K. Leo, D. Birkedal, J. Erland, M. van der Poel, J. Hvam, *Coherent Spectroscopy of Semiconductor Microcavities*, in Ref. [05B1].

The final problem

Find all errors in this book both in physics and in printing, collect all aspects which can be improved and all important phenomena which have been omitted and solve all problems. Send a corresponding listing to the author and help him by doing so to improve the next editions of this book – if there are any. Many thank inadvance for your help.

Actually, the author himself is working hard on parts of this problem, trying to reduce the number of mistakes and misprints from one edition to the next, but possibly introducing some new ones at the same time. A dear colleague, Prof. Dr. B. Di Bartolo at Boston College often states, "In science a case is never closed." The author generally agrees with this statement and especially as it applies to this "final problem".

Subject Index