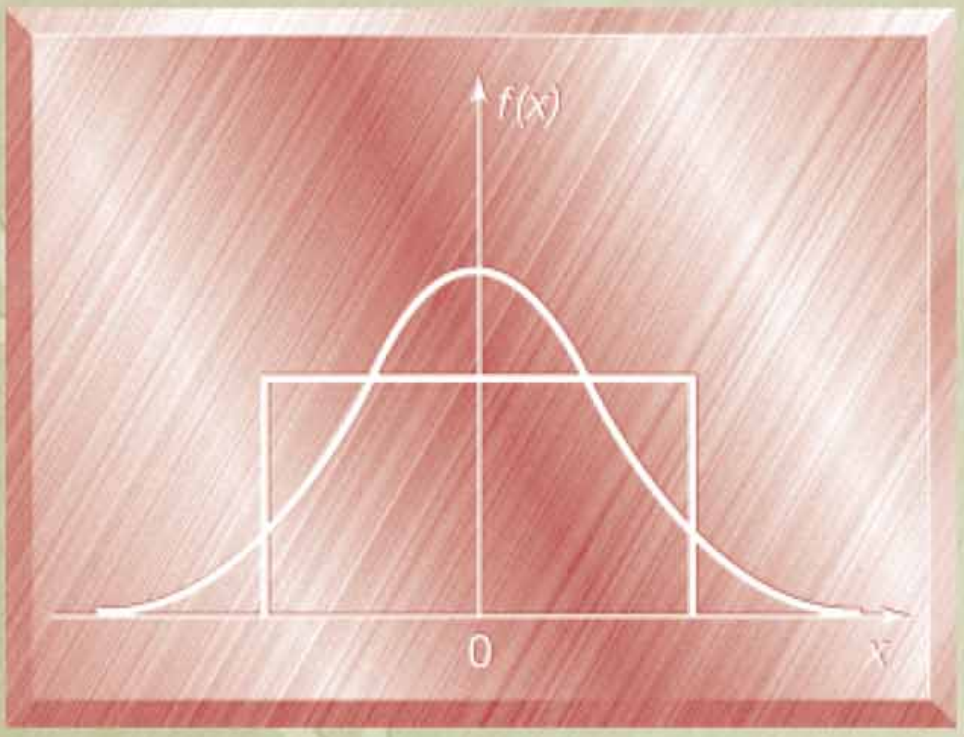
THIRD EDITION

Measurement Errors and Uncertainties

Theory and Practice



Semyon G. Rabinovich

Springer

Measurement Errors and Uncertainties

Third Edition

Semyon G. Rabinovich

Measurement Errors and Uncertainties

Theory and Practice

Third Edition

AIP PRESS

Springer

Semyon G. Rabinovich
142 Manchester Drive
Basking Ridge, NJ 07920
USA

Library of Congress Control Number: 2005923501

ISBN-10: 0-387-25358-0 Printed on acid-free paper.
ISBN-13: 978-0387-25368-9

Second edition translated by M. E. Alferieff.

Printed in the United States of America. (TB/MVY)

9 8 7 6 5 4 3 2 1 SPIN 10994696

springeronline.com

Preface

The major objective of this book is to give methods for estimating errors and uncertainties of real measurements: measurements that are performed in industry, commerce, and experimental research.

This book is needed because the existing theory of measurement errors was historically developed as an abstract mathematical discipline. As a result, this theory allows estimation of uncertainties of some ideal measurements only and is not applicable to most practical cases. In particular, it is not applicable to single measurements. This situation did not bother mathematicians, whereas engineers, not being bold enough to assert that the mathematical theory of errors cannot satisfy their needs, solved their particular problems in one or another ad hoc manner.

Actually, any measurement of a physical quantity is not abstract, but it involves an entirely concrete procedure that is always implemented with concrete technical devices—measuring instruments—under concrete conditions. Therefore, to obtain realistic estimates of measurement uncertainties, mathematical methods must be supplemented with methods that make it possible to take into account data on properties of measuring instruments, the conditions under which measurements are performed, the measurement procedure, and other features of measurements.

The importance of the methods of estimating measurement inaccuracies for practice can scarcely be exaggerated. Indeed, in another stage of planning a measurement or using a measurement result, one must know its error limits or uncertainty. Inaccuracy of a measurement determines its quality and is related to its cost. Reliability of product quality control also depends on accuracy of measurements. Without estimating measurement inaccuracies, one cannot compare measurement results obtained by different authors. Finally, it is now universally recognized that the precision with which any calculation using experimental data is performed must be consistent with the accuracy of these data.

In this book, the entire hierarchy of questions pertaining to measurement errors and uncertainties is studied, a theory of measurement inaccuracy is developed, and specific recommendations are made for solving the basic problems arising in practice. In addition, methods are presented for calculating the errors of measuring instruments. The attention devoted to the properties of measuring instruments,

taking into account their relations with measurement inaccuracies, is one highlight of this book.

This book is a product of my professional scientific experience accumulated over many years of work in instrumentation and metrology. From 1948 to 1964, I was involved in the investigation and development of various electric measuring instruments, including calibrating potentiometers and stabilizers, extremely sensitive dc voltage and current amplifiers, automatic plotters, and so on. This experience gave me a grip in understanding problems arising in real measurements. Then, in 1965, I organized, and until 1980 directed, a laboratory of theoretical metrology. I focused on the analysis and generalization of theoretical problems in metrology. In particular, because I discovered that a rift exists between theory and practice (as mentioned above), I concentrated on the problem of estimating measurement errors and uncertainties. The results achieved during these years formed the foundation of my book *Measurement Errors* [44]. Further work and new results led to the writing of this book.

This book was initially published under the title *Measurement Errors: Theory and Practice* and has since gone through several editions, each reflecting new results that I have obtained. The initial hardware edition recommended a way to calculate the inaccuracy of single measurements. The paperback edition that followed added new treatment of indirect measurements, notably, a way of accounting for dependencies between the components of the uncertainty of indirect measurements. The second edition offered a full analysis of the method of reduction for processing indirect measurement data. The analysis shows the great advantage of this method over the traditional one based on the Taylor's series. In particular, the method of reduction obviates the need for the calculation of correlation coefficients. This development is important because the calculation of the correlation coefficient is one of the most notorious stumbling blocks in estimating the inaccuracy of measurement results.

However, the method of reduction is applicable only to dependent indirect measurements such as the measurement of electrical resistance using a voltmeter and ammeter. For independent indirect measurements, such as the measurement of the density of a solid body, the traditional method with its shortcomings was still inevitable. Only recently did I find a better solution for processing independent indirect measurement data. I called it the method of transformation. This method supplements the method of reduction and thus completes the creation of the new theory of indirect measurements. In addition to removing the need to calculate the correlation coefficient, the new theory allows the construction of the confidence intervals and produces well-grounded estimates of the uncertainty of both types of indirect measurements. This new theory is presented in this third edition of the book.

This edition has 12 chapters. Chapter 1 contains general information on measurements and metrology. Although introductory, the chapter includes some questions that are solved or presented anew. Also partially introductory is Chapter 2, devoted to measuring instruments. However, a large portion of it presents analysis of methods of standardization of the metrological characteristics of measuring instruments, which are important for practice and necessary for estimating measurement

errors and uncertainties. Statistical analysis of errors of several batches of various measuring instruments obtained by standards laboratories is given. The analysis shows that such data are statistically unstable and hence cannot be the basis for obtaining a distribution function of errors of measuring instruments. This important result has influenced the ways in which many problems are covered in this book.

The inaccuracy of measurements always has to be estimated based on indirect data by finding and then summing the elementary components of the inaccuracy. In Chapter 3, a general analysis of elementary errors of measurements is given. Also, the classification of elementary errors is presented and their mathematical models are introduced. Two important methods of constructing a convolution of distribution functions are presented. These methods are necessary for summing elementary errors.

Chapter 4 contains methods of mathematical statistics as applied to idealized multiple measurements. In essence, these methods constitute the classical theory of measurement errors. New to the third edition is the review of modern robust and nonparametric methods of measurement data processing.

In Chapter 5, real direct measurements are considered. It is shown that single measurements should be considered as the basic form of measurement. Various methods for estimating and combining systematic and random errors are considered, and a comparative analysis of these methods is given. Special attention is paid to taking into account the errors of measuring instruments. For instance, it is shown how the uncertainty of a measurement result decreases when more accurate information on the properties of measuring instruments is used. This chapter concludes with a step-by-step procedure for estimating errors and uncertainties of direct measurements.

Chapter 6 presents the new theory of indirect measurements including the method of transformation that is added in this edition. The current edition also expands the examples of indirect measurements to illustrate the new method. These examples are taken out from Chapter 6 and organized into a separate Chapter 7.

In Chapter 8, combined measurements are considered. The well-known least-squares algorithm is described in detail. The new theory of indirect measurements allowed us to eliminate here the category of simultaneous measurements.

Chapter 9 contains methods for combining measurement results. Such methods are necessary in the cases where the same measurand is measured in multiple stages or in different laboratories. Along with the traditional solution, which takes into consideration only random errors, Chapter 9 includes a method taking into account systematic errors as well.

In Chapters 10 and 11, I return to considering measuring instruments. Chapter 10 gives general methods for calculating their total errors that are useful during the development of the instruments. In Chapter 11, calibration methods that tie measuring instruments to corresponding standards are considered.

The current edition also adds Chapter 12 with concluding remarks. This chapter briefly reviews the history of measurement data processing and outlines some current open problems in the theory and practice of measurements. The chapter also discusses two recent documents produced by international standards bodies,

which are of fundamental importance to metrology: The "International Vocabulary of Basic and General Terms in Metrology" [2] and the "Guide to the Expression of Uncertainty in Measurement" [1].

In addition to the new theory of indirect measurements, the third edition contains many clarifications and corrections to the text of the second edition. Also, the list of references is updated.

The book is targeted for practical use and, to this end, includes many concrete examples, many of which illustrate typical problems arising in the practice of measurements.

This book is intended for anyone who is concerned with measurements in any field of science or technology, who designs technological processes and chooses for them instruments having appropriate accuracy, and who designs and tests new measuring devices. I also believe this book will prove useful to many university and college students. Indeed, measurements are of such fundamental importance for modern science and engineering that every engineer and every scientist doing experimental research must know the basics of the theory of measurements and especially how to estimate their accuracy.

In conclusion, I would like to thank Dr. E. Richard Cohen for carefully reading the manuscript of the second edition of this book and for many useful comments.

I would like to also thank Dr. Abram Kagan, now Professor at the University of Maryland, College Park, for the many years of collaboration and friendship. This book benefited from our discussions on various mathematical problems in metrology.

The initial hardback edition of the book was translated by M. E. Alferieff. The additions and changes to the subsequent editions were translated or edited by my son, Dr. Michael Rabinovich. Beyond that, Michael provided support and assistance throughout my work on this book, from editing the book proposal to publishers to discussing new results and the presentation. This book would not be possible without his help.

Basking Ridge, New Jersey Semyon G. Rabinovich

Contents

1
General Information About Measurements

1.1. Basic Concepts and Terms

The theory of measurement errors is a branch of metrology—the science of measurements. In presenting the theory we shall adhere, whenever possible, to the terminology given in the *International Vocabulary of Basic and General Terms of Metrology* [2]. We shall discuss the terms that are most important for this book.

A measurable quantity (briefly—measurand) is a property of phenomena, bodies, or substances that can be defined qualitatively and expressed quantitatively.

The first measurable quantities were probably length, mass, and time, i.e., quantities that people employed in everyday life, and these concepts appeared unconsciously. Later, with the development of science, measurable quantities came to be introduced consciously to study the corresponding laws in physics, chemistry, and biology.

Measurable quantities are also called physical quantities. The principal feature of physical quantities is that they can be measured.

The term *quantity* is used in both the general and the particular sense. It is used in the general sense when referring to the general properties of objects, for example, length, mass, temperature, or electric resistance. It is used in the particular sense when referring to the properties of a specific object: the length of a given rod, the electric resistance of a given segment of wire, and so on.

Measurement is the process of determinating the value of a physical quantity experimentally with the help of special technical means called *measuring instruments.*

The *value of a physical quantity* is the product of a number and a unit adapted for these quantities. It is found as the result of a measurement.

The definitions presented above underscore three features of measurement:

(1) The result of a measurement must always be a concrete denominated number expressed in sanctioned units of measurements. The purpose of measurement is essentially to represent a property of an object by a number.
(2) A measurement is always performed with the help of some measuring instrument; measurement is impossible without measuring instruments.
(3) Measurement is always an experimental procedure.

The *true value of a measurand* is the value of the measured physical quantity, which, being known, would ideally reflect, both qualitatively and quantitatively, the corresponding property of the object.

Measuring instruments are created by humans, and every measurement on the whole is an experimental procedure. Therefore, results of measurements cannot be absolutely accurate. This unavoidable imperfection of measurements is expressed in their inaccuracy. Quantitatively the measurement inaccuracy is characterized by the notion of either limits of error or uncertainty.

We shall use the term *uncertainty* to characterize the inaccuracy of a measurement result, whereas the term *error* is used to characterize the components of the uncertainty. We shall return to these terms many times later in this book.

The *measurement error* is the deviation of the result of measurement from the true value of the measurable quantity, expressed in absolute or relative form.

If A is the true value of the measurable quantity and $\tilde{A}$ is the result of measurement, then the absolute error of measurement is $\zeta = \tilde{A} - A$. This equation is often used as a definition of this term, but by doing that, one narrows the essence of this term.

The error expressed in absolute form is called the absolute measurement error. The error expressed in relative form is called the relative measurement error.

The absolute error is usually identified by the fact that it is expressed in the same units as the measurable quantity.

Absolute error is a physical quantity, and its value may be positive, negative, or even given by an interval that contains that value. One should not confuse the absolute error with the absolute value of that error. For example, the absolute error -0.3 mm has the absolute value 0.3.

The relative error is the error expressed as a fraction of the true value of the measurable quantity $\varepsilon = (\tilde{A} - A)/A$. Relative errors are normally given as percent and sometimes per thousand (denoted by ‰). Very small errors, which are encountered in the most precise measurements, are customarily expressed directly as fractions of the measured quantity.

Uncertainty of measurement is an interval within which a true value of a measurand lies with a given probability. Uncertainty is defined with its limits that are read out from a result of measurement in compliance with the mentioned probability. Like an error, uncertainty can be specified in absolute or relative form. The relation between the terms “error” and “uncertainty” is discussed in more detail in Section 3.1.

Inaccuracy of measurements characterize the imperfection of measurements. A positive characteristic of measurements is their accuracy. The accuracy of a measurement reflects how close the result is to the true value of the measured quantity.

A measurement is all the more accurate the smaller its error is. Absolute errors, however, depend in general on the value of the measured quantity, and for this reason, they are not a suitable quantitative characteristic of measurement accuracy. Relative errors do not have this drawback. For this reason, accuracy can be characterized quantitatively by a number equal to the inverse of the relative error

expressed as a fraction of the measured quantity. For example, if the limits of error of a measurement are $\pm 2 \times 10^{-3}\% = \pm 2 \times 10^{-5}$, then the accuracy of this measurement will be 5×10^4. The accuracy is expressed only as a positive number; that calculation is based on the absolute value of the limits of the measurement error.

Although it is possible to introduce in this manner the quantitative characteristic of accuracy, in practice, accuracy is normally not estimated quantitatively and it is usually characterized indirectly with the help of the measurement error or the uncertainty of measurement.

Other concepts and terms will be explained as they are introduced, and they are given in the Glossary.

1.2. Metrology and the Basic Metrological Problems

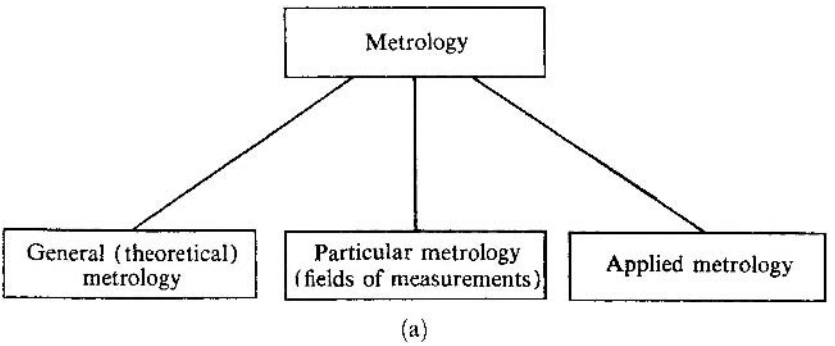
Comparison is an age-old element of human thought, and the process of making comparisons lies at the heart of measurement: Homogeneous quantities characterizing different objects are identified and then compared; one quantity is taken to be the unit of measurement, and all other quantities are compared with it. This process is how the first measures, i.e., objects the size of whose corresponding physical quantity is taken to be unity or a known number of units, arose.

At one time even different cities each had their own units and measures. Then it was necessary to know how measures were related. This problem gave birth to the science of measures—metrology.

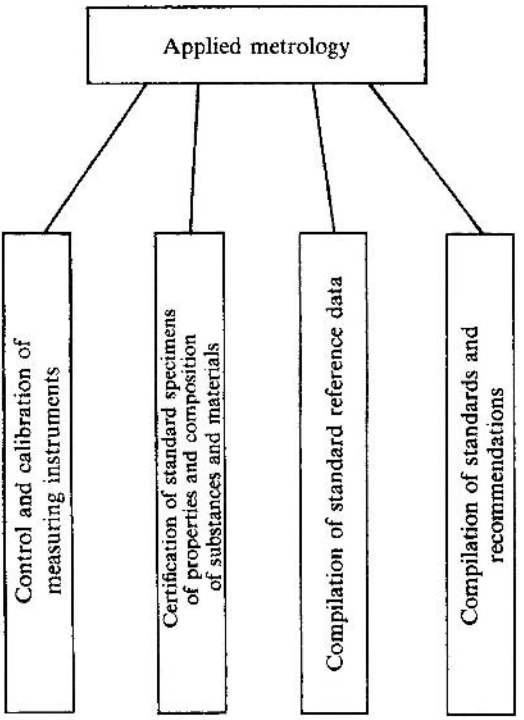
But the content of metrology, as that of most sciences, is not immutable. Especially profound changes started in the second half of the nineteenth century, when industry and science developed rapidly and, in particular, electrical technology and instrument building began. Measurements were no longer a part of production processes and commerce, and they became a powerful means of gaining knowledge—they became a tool of science. The role of measurements has increased especially today, in connection with the rapid development of science and technology in the fields of nuclear, space, electronics, information systems, and so on.

The development of science and technology, intercourse among peoples, and international trade have prompted many countries to adopt the same units of physical quantities. The most important step in this direction was the signing of the Metric Convention [(Treaty of the Meter), 1875]. This act had enormous significance not only with regard to the unification of physical quantities and dissemination of the metric system, but also with regard to unifying measurements throughout the world. The Metric Convention and the institutions created by it—the General Conference on Weights and Measures (CIPM), the International Committee, and the International Bureau of Weights and Measures (BIPM)—continue their important work even now. In 1960, the General Conference adopted the international system of units (SI) [10]. Most countries now use this system.

The content of metrology also changed along with the change in the problems of measurements. Metrology has become the science of measurements. The block diagrams in Fig. 1.1 show the range of questions encompassed by modern metrology.

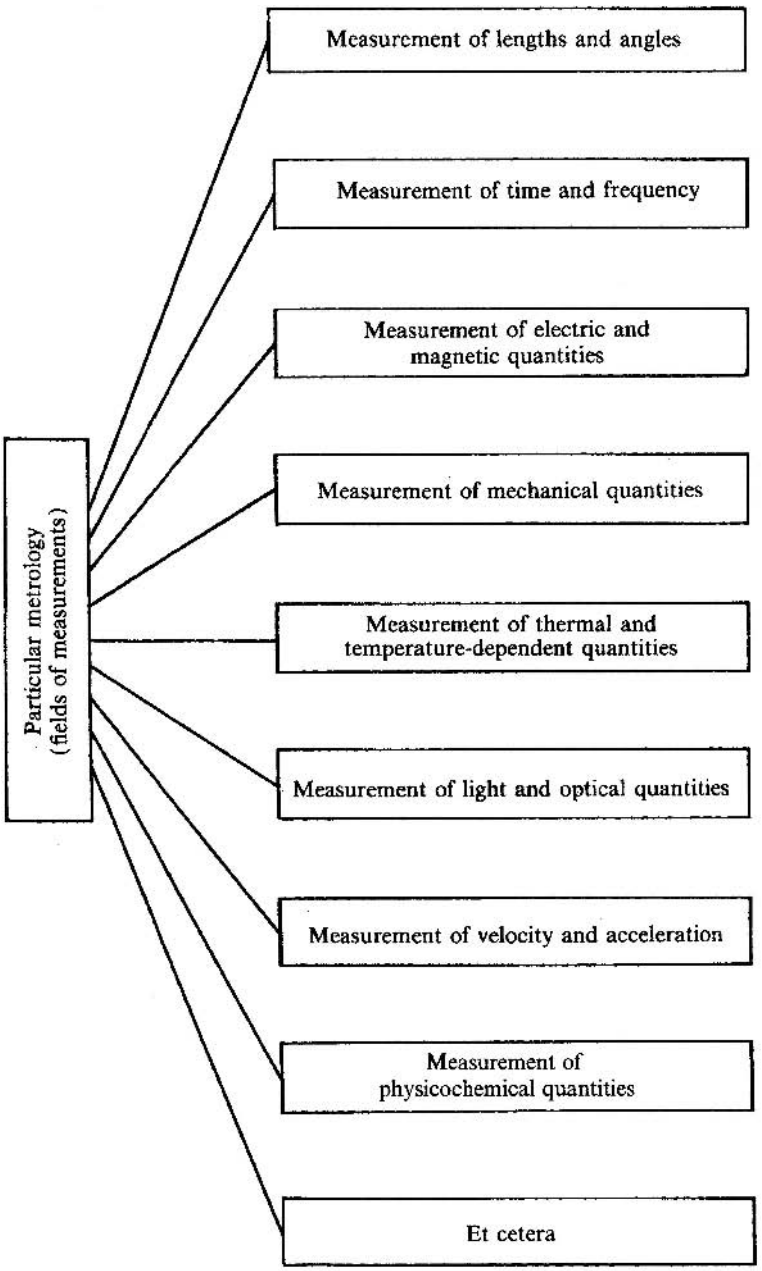


(a)



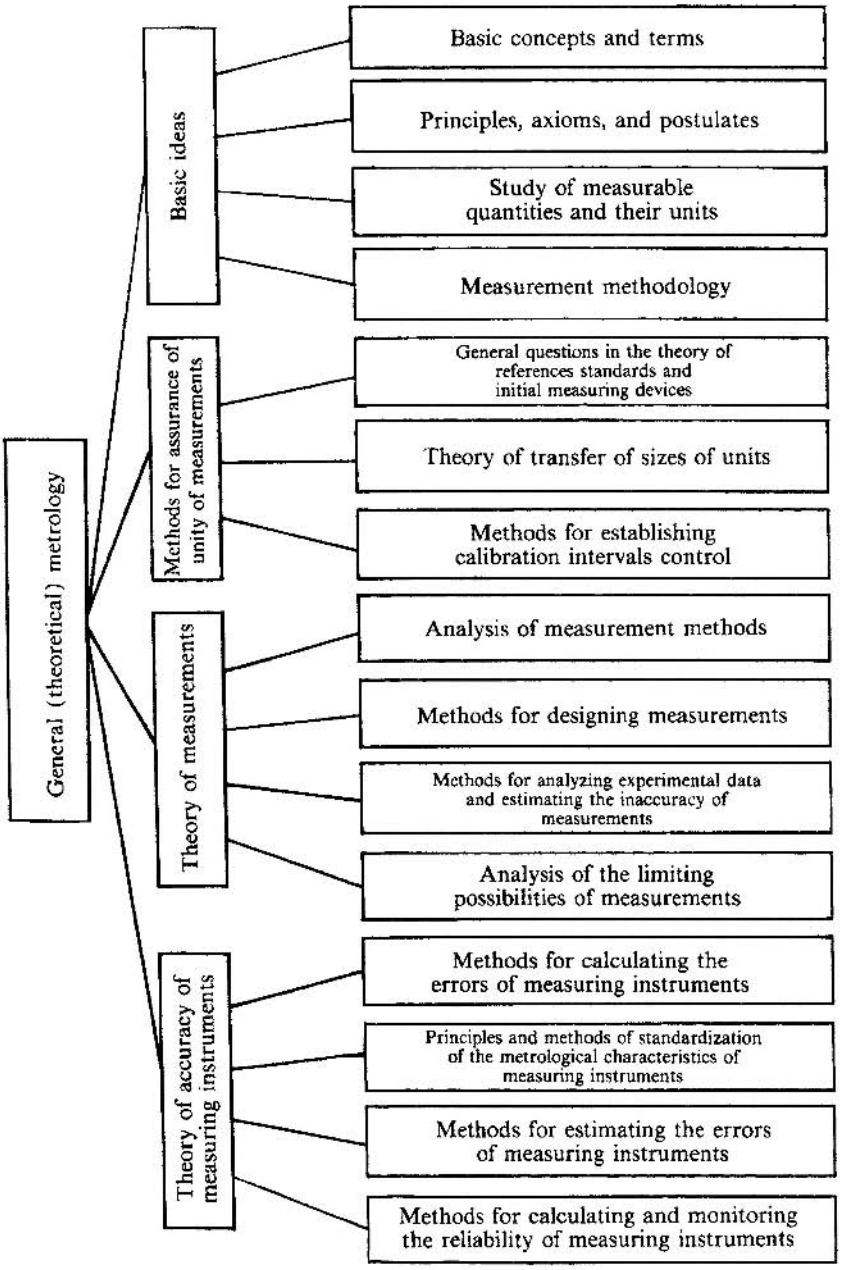
(b)

FIGURE 1.1. Schematic picture of the basic problems of metrology: (a) metrology, (b) applied metrology, (c) particular metrology, and (d) general metrology.



(c)

FIGURE 1.1. (*continued*)



(d)

FIGURE 1.1. (*continued*)

The questions are incorporated into sections and subsections, whose names give an idea of their content. The content of some of them must nonetheless be explained.

(1) The Study of Physical, i.e., Measurable, Quantities and their Units [Fig. 1.1 (d)]

Physical quantities are introduced in different fields of knowledge, in physics, chemistry, biology, and so on. The rules for introducing and classifying them and for forming systems of units and for optimizing these systems cannot be addressed in any of these sciences, and already for this reason, they must be included among the problems addressed in metrology. Moreover, the size of a quantity to be used as a unit of measurement and its determination are also important for measurement accuracy. One need only recall that when the distance between two markings on a platinum-irridium rod was adopted for the meter, for the most accurate measurement of length, the inaccuracy was not less than 10^{-6}. When the meter was later defined as a definite number (1,650,763.73) of wavelengths of krypton-86 radiation in vacuum, this inaccuracy was reduced to $10^{-7} - 10^{-8}$. Now, when the definition of the meter is based on the velocity of light in vacuum, the inaccuracy in measuring length has been reduced by another order of magnitude and it can be reduced even more.

(2) General Theory of Reference Standards and Initial Measuring Devices

The units of physical quantities are materialized; i.e., they are reproduced, with the help of reference standards and initial measuring devices, and for this reason, these measuring devices play an exceptionally important role in the unity of measurements. The reference standard of each unit is unique, and it is physically created based on the laws of specific fields of physics and technology. For this reason, general metrology cannot answer the question of how a reference standard should be constructed. But metrology must determine when a reference standard must be created, and it must establish the criteria for determining when such a reference standard must be a single or group reference standard. In metrology, the theory and methods of comparing reference standards and monitoring their stability as well as methods for expressing errors must also be studied. Practice raises many such purely metrological questions.

(3) Theory of Transfer of the Size of Units into Measurement Practice

In order that the results of all measurements be expressed for established units, all means of measurement (measures, instruments, measuring transducers, measuring systems) must be calibrated with respect to reference standards. This problem cannot, however, be solved directly based on primary reference standards,

i.e., reference standards that reproduce units. It is solved with the help of a system of secondary reference standards, i.e., reference standards that are calibrated with respect to the primary reference standard, and working reference standards, i.e., reference standards that are calibrated with respect to secondary reference standards. Thus the system of reference standards has a hierarchical structure. The entire procedure of calibrating reference standards and, with their help, the working measuring instruments is referred to as transfer of the sizes of units into measurement practice. The final stages of transferring the sizes of units consists of calibration of the scales of the measuring instruments, adjustment of measures, and determination of the actual values of the quantities that are reproduced by them, after which all measuring instruments are checked at the time they are issued and then periodically during use.

In solving these problems, a series of questions arises. For example, how many gradations of accuracy of reference standards are required? How many secondary and working reference standards are required for each level of accuracy? How does the error increase when the size of a unit is transferred from one reference standard to another? How does this error increase from the reference standard to the working measuring instrument? What should be the relation between the accuracy of the reference standard and the measuring instrument that is calibrated (verified) with respect to it? How should complicated measurement systems be checked? Metrology should answer these questions.

The other blocks in the diagram of Fig. 1.1 (d) do not require any explanations. We shall now turn to Fig. 1.1 (a) and focus on the section *particular metrology*, which the fields of measurement comprise. Examples are lineal–angular measurements, measurements of mechanical quantities, measurements of electric and magnetic quantities, and so on. The central problem arising in each field of measurement is the problem of creating conditions under which the measurements of the corresponding physical quantities are unified. For this purpose, in each field of measurement, a system of initial measuring devices—reference standards and standard measures—is created, and methods for calibrating and checking the working measuring instruments are developed. The specific nature of each field of measurement engenders a great many problems characteristic of it. However, many problems that are common to several fields of measurement are encountered. The analysis of such problems and the development of methods for solving them are now problems of general metrology.

Applied metrology, which incorporates the metrological service and legislative metrology, is of great importance for achieving the final goals of metrology as a science. The metrological service checks and calibrates measuring instruments and certifies standards of properties and composition; i.e., it maintains the uniformity of measuring instruments employed in the country. The functions of legislative metrology are to enact laws that would guarantee uniformity of measuring instruments and unity of measurements. Thus, a system of physical quantities and the units, employed in a country, can only be established by means of legislation. The rules giving the right to manufacture measuring instruments and to check the

state of these instruments when they are in use are also established by means of legislation.

We shall now define more accurately some of the expressions and terms mentioned above.

Uniformity of measuring instruments refers to the state of these instruments in which they are all carriers of the established units and their errors and other properties, which are important for the instruments to be used as intended, fall within the established limits.

Unity of measurements refers to a common quality of all measurements performed in a region (in a country, in a group of countries, or in the world) such that the results of measurements are expressed for established units and agree with one another within the limits of estimated error or uncertainties.

Uniformity of measuring instruments is a necessary prerequisite for unity of measurements. But the result of a measurement depends not only on the quality of the measuring instrument employed but also on many other factors, including human factors (if measurement is not automatic). For this reason, unity of measurements in general is the limiting state that must be strived for, but which, as any ideal, is unattainable.

This is a good point at which to discuss the development of reference standards. A reference standard is always a particular measuring device: a measure, instrument, or measuring apparatus. Such measuring devices were initially employed as reference standards arbitrarily by simple volition of the institution responsible for correctness of measurements in the country. However, there is always the danger that a reference standard will be ruined, which can happen because of a natural disaster, fire, and so on. An arbitrarily established reference standard, i.e., a prototype reference standard, cannot be reproduced.

As a result, scientists have for a long time strived to define units of measurement so that the reference standards embodying them could be reproducible. For this, the units of the quantities were defined based on natural phenomena. Thus, the second was defined based on the period of revolution of the Earth around the sun: the meter was defined based on the length of the Parisian meridian, and so on. Scientists hoped that these units would serve "for all time and for all peoples." Historically this stage of development of metrology coincided with the creation of the metric system.

Further investigations revealed, however, that the chosen natural phenomena are not sufficiently unique or are not stable enough. Nonetheless the idea of defining units based on natural phenomena was not questioned. It was only necessary to seek other natural phenomena corresponding to a higher level of knowledge of nature.

It was found that the most stable or even absolutely stable phenomena are characteristic of phenomena studied in quantum physics, and that the physical constants can be employed successfully for purposes of defining units, and the corresponding effects can be employed for realizing reference standards. The meter, the second, and the volt have now been defined in this manner. It can be conjectured that in the

near future, the volt, defined and reproduced based on the Josephson effect, will replace the ampere as the basic electric unit.

The numerical values of the basic physical constants are widely used in various calculations, and therefore, these values must be in concordance with each other. To reach this goal, all values of fundamental physical constants obtained by experiments must be adjusted. The most recent adjustment was carried out in 1998, and the results were published in 1999 [40].

As one can see from the problems with which it is concerned, metrology is an applied science. However, the subject of metrology—measurement—is a tool of both fundamental sciences (physics, chemistry, and biology) and applied disciplines, and it is widely employed in all spheres of industry, in everday life, and in commerce. No other applied science has such a wide range of applications as does metrology.

We shall return once again to particular metrology. A simple list of the fields of measurement shows that the measurable quantities and therefore measurement methods and measuring instruments are extremely diverse. What then do the different fields of measurement have in common? They are united by general or theoretical metrology and, primarily, the general methodology of measurement and the general theory of inaccuracy of measurements. For this reason, the development of these branches of metrology is important for all fields of science and for all spheres of industry that employ measurements. The importance of these branches of metrology is also indicated by the obvious fact that a specialist in one field of measurement can easily adapt to and work in a different field of measurement.

1.3. Initial Points of the Theory of Measurements

Measurements are so common and intuitively understandable that one would think there is no need to identify the prerequisites on which measurements are based. However, a clear understanding of the starting premises is necessary for the development of any science, and for this reason, it is desirable to examine the prerequisites of the theory of measurements.

When some quantity characterizing a specific object is being measured, this object is made to interact with a measuring instrument. Thus to measure the diameter of a rod, the rod is squeezed between the jaws of a vernier caliper; to measure the voltage of an electric circuit, a voltmeter is connected to it; and so on. The indication of the measuring instrument—the sliding calipers, voltmeter, and so on—gives an estimate of the measurable quantity, i.e., the result of the measurement. When necessary, the number of divisions read on the instrument scale is multiplied by a certain factor. In many cases, the result of measurement is found by mathematical analysis of the indications of a instrument or several instruments. For example, the density of solid bodies, the temperature coefficients of the electric resistance of resistors, and many other physical quantities are measured in this manner.

The imperfection of measuring instruments, the inaccuracy with which the sizes of the units are transferred to them, as well as some other factors that we shall study

below, result in the appearance of measurement errors. Measurement errors are in principle unavoidable, because a measurement is an experimental procedure and the true value of the measurable quantity is an abstract concept. As the measurement methods and measuring instruments improve, however, measurement errors decrease.

The introduction of measurable quantities and the establishment of their units is a necessary prerequisite of measurements. Any measurement, however, is always performed on a specific object, and the general definition of the measurable quantity must be specified taking into account the properties of the object and the objective of the measurement. The true value of the measurable quantity is essentially introduced and defined in this manner. Unfortunately, this important preparatory stage of measurements is usually not formulated and not singled out.

To clarify this question, we shall study a simple measurement problem—the measurement of the diameter of a disk. First we shall formulate the problem. The fact that the diameter of a disk is to be measured means that the disk, i.e., the object of study, is a circle. We note that the concepts "circle" and "diameter of a circle" are mathematical, i.e., abstract, concepts. The circle is a representation or model of the given body. The diameter of the circle is the parameter of the model and is a mathematically rigorous definition of the measurable quantity. Now, in accordance with the general definition of the true value of the measurable quantity, it can be stated that the true value of the diameter of the disk is the value of the parameter of the model (diameter of the disk) that reflects, in the quantitative respect, the property of the object of interest to us; the ideal qualitative correspondence must be predetermined by the model.

Let us return to our example. The purpose of the disk permits determining the permissible measurement error and choosing an appropriate measuring instrument. Bringing the object into contact with the measuring instrument, we obtain the result of measurement. But the diameter of the circle is, by definition, invariant under rotation. For this reason, the measurement must be performed in several different directions. If the difference of the results of the measurements are less than the permissible measurement error, then any of the obtained results can be taken as the result of measurement. After the value of the measurable quantity, a concrete number, which is an estimate of the true value of the measurand, has been found, the measurement can be regarded as being completed.

But it may happen that the difference of the measurements in different directions exceeds the permissible error of a given measurement. In this situation, we must state that within the required measurement accuracy, our disk does not have a unique diameter, as does a circle. Therefore, no concrete number can be taken, with prescribed accuracy, as an estimate of the true value of the measurable quantity. Hence, the adopted model does not correspond to the properties of the real object, and the measurement problem has not been correctly formulated.

If the object is a manufactured article and the model is a drawing of the article, then any disparity between them means that the article is defective. If, however, the object is a natural object, then the disparity means that the model is not applicable and it must be reexamined.

Of course, even when measurement of the diameter of the disk is assumed to be possible, in reality, the diameter of the disk is not absolutely identical in different directions. But as long as this inconstancy is negligibly small, we can assume that the circle as a model corresponds to the object and therefore a constant, fixed true value of the measurable quantity exists, and an estimate of the quantity can be found as a result of measurement. Moreover, if the measurement has been performed, we can assume that the true value of the measurable quantity lies somewhere near the obtained estimate and differs from it by not more than the measurement error.

Thus the idealization necessary for constructing a model gives rise to an unavoidable discrepancy between the parameter of the model and the real property of the object. We shall call this discrepancy the threshold discrepancy.

As we saw above, the error caused by the threshold discrepancy between the model and the object must be less than the total measurement error. If, however, this component of the error exceeds the limit of permissible measurement error, then it is impossible to make a measurement with the required accuracy. This result indicates that the model is inadequate. To continue the experiment, if this is permissible for the objective of the measurement, the model must be redefined. Thus, in the example of the measurement of the diameter of a disk, a different model could be a circle circumscribing the disk.

The case studied above is simple, but the features demonstrated for it are present in any measurement, although they are not always so easily and clearly perceived as when measuring lineal dimensions.

The foregoing considerations essentially reduce to three prerequisites:

(a) Some parameter of the model of the object must correspond to a measurable property of the object.
(b) The model of the object must permit the assumption that during the time required to perform the measurement, the parameter of the object, corresponding to the property of the object being measured, is constant.
(c) The error caused by the threshold discrepancy between the model and the object must be less than the permissible measurement error.

Generalizing all three assumptions, we formulate the following principles of metrology: A measurement with fixed accuracy can be performed only, if to a measurable property of the object, it is possible to associate with no less accuracy a determinate parameter of its model.

Any constant parameter is, of course, a determinate parameter. The instantaneous value of a variable (varying) quantity can also be regarded as a determinate parameter.

We note that the parameter of a model of an object introduced in this manner is the true value of the measurable quantity.

The foregoing considerations are fundamental, and they can be represented in the form of postulates of the theory of measurement [44], [50]:

(α) The true value of the measurable quantity exists.
(β) The true value of the measurable quantity is constant.
(γ) The true value cannot be found.

The threshold discrepancy between the model and the object was employed above as a justification of the postulate (γ). However, other unavoidable restrictions also exist on the approximation of the true value of a measurable quantity. For example, the accuracy of measuring instruments is unavoidably limited. For this reason, it is possible to formulate the simple statement: *The result of any measurement always contains an error*.

It was mentioned above that a necessary prerequisite of measurements is the introduction of physical quantities and their units. These questions are not directly related with the problem of estimating measurement errors, and for this reason, they are not studied here. These questions are investigated in several works. We call attention to the book by B.D. Ellis [24] and the work of K.P. Shirokov [48].

At this point we shall discuss some examples of models that are employed for specific measurement problems.

Measurement of the Parameters of Alternating Current

The object of study is an alternating current. The model of the object is a sinusoid

$$i = I_m \sin(\omega t + \varphi),$$

where t is the time and I_m, ω, and φ are the amplitude, the angular frequency, and the initial phase, and they are the parameters of the model.

Each parameter of the model corresponds to some real property of the object and can be a measurable quantity. But, in addition to these quantities (arguments), several other parameters that are functionally related with them are also introduced. These parameters can also be measurable quantities. Some parameters can be introduced in a manner such that by definition they are not related with the "details" of the phenomenon. An example is the effective current

$$I = \sqrt{\frac{1}{T}\int_0^T i^2 dt},$$

where $T = 2\pi/\omega$ is the period of the sinusoid.

A nonsinusoidal current is also characterized by an effective current. However, in developing measuring instruments and describing their properties, the form of the current, i.e., the model of the object of study, must be taken into account.

The discrepancy between the model and the object in this case is expressed as a discrepancy between the sinusoid and the curve of the time dependence of the current strength. In this case, however, only rarely is it possible to discover the discrepancy between the model and the process under study by means of simple repetition of measurements of some parameters. For this reason, the correspondence between the model and the object is checked differently, for example, by measuring the form distortion factor.

The model is usually redefined by replacing one sinusoid by a sum of a certain number of sinusoids.

Measurement of the Parameters of Random Processes

The standard model is a stationary ergodic random process on the time interval T. The constant parameters of the process are the mathematical expectation $E[X]$ and the variance $V[X]$. Suppose that we are interested in $E[X]$. The expectation $E[X]$ can be estimated, for example, with the help of the formula

$$\bar{x} = \left(\frac{\sum_{i=1}^{n} x_i}{n} \right)_T ,$$

where T is the observational time interval, x_i are the estimates of the realization of the random quantity, whose variation in time forms a random process at times $t_i \in T$, and n is the total number of realizations obtained.

Repeated measurements on other realizations of the process can give somewhat different values of $\bar{x}$. The adopted model can be regarded as corresponding to the physical phenomenon under study, if the differences between the obtained estimates of the mathematical expectation of the process are not close to the permissible measurement error. If, however, the difference of the estimates of the measured quantity are close to the error or exceed it, then the model must be redefined, which is most simply done by increasing the observational interval T.

It is interesting to note that the definitions of some parameters seem, at first glance, to permit arbitrary measurement accuracy (if the errors of the measuring instrument are ignored). Examples of such parameters are the parameters of stationary random processes, the parameters of distributions of random quantities, and the average value of the quantity. One would think that to achieve the required accuracy in these cases, it is sufficient to increase the number of observations when performing the measurements. In reality, however, the accuracy of measurement is always limited, and in particular, it is limited by the correspondence between the model and the phenomenon, i.e., by the possibility of assuming that to the given phenomenon, there corresponds a stationary random process or a random quantity with a known distribution.

In the last few years, much has been written about measurements of variable and random quantities, But these quantities, as such, do not have a true value, and for this reason, they cannot be measured.

For a random quantity, it is possible to measure the parameters of its distribution function, which are not random; it is possible to measure the realization of a random quantity. For a variable quantity, it is possible to measure its parameters that are not variable; it is also possible to measure the instantaneous values of a variable quantity.

We shall now discuss in somewhat greater detail the measurement of instantaneous values of quantities. Suppose that we are studying an alternating current, the model of which is a sinusoid with amplitude I_m, angular frequency ω, and initial phase φ. At time t_1 to the instantaneous current, there corresponds in the model the instantaneous value $i_1 = I_m \sin(\omega t_1 + \varphi)$. At a different time, there will

be a different instantaneous value, but at each moment, it has some definite value. Thus to the measurable property of the object, there always corresponds a fixed parameter of its model.

Measurement, however, requires time. The measurable quantity will change during the measurement time, and this will generate a specific error of the given measurement. The objective of the measurement permits setting a level that the measurement error, as well as its component caused by the change in the measurable quantity over the measurement time, must not exceed.

If this condition is satisfied, then the effect of the measurement time can be neglected, and it can be assumed that as a result we obtain an estimate of the measured instantaneous current, i.e., the current strength at a given moment in time.

In the literature, the term measurement of variable quantities usually refers to measurement of instantaneous values, and the expression *measurement of a variable quantity* is imprecise. In the case of measurement of a random quantity, the writer usually has in mind the measurement of a realization of a random quantity.

Physical quantities are divided into active and passive. Active quantities are quantities that can generate measurement signals without any auxiliary sources of energy; i.e., they act on the measuring instruments. Such quantities are the emf, the strength of an electric current, mechanical force, and so on. Passive quantities cannot act on measuring instruments, and for measurements, they must be activated. Examples of passive quantities are mass, inductance, and electric resistance. Mass is usually measured based on the fact that in a gravitational field, a force proportional to the mass acts on the body. Electric resistance is activated by passing an electric current through a resistor.

When measuring passive physical quantities characterizing some objects, the models of the objects are constructed for the active quantities that are formed by activation of passive quantities.

1.4. Classification of Measurements

In metrology there has been a long-standing tradition to distinguish direct, indirect, and combined measurements. In the last few years, metrologists have begun to divide combined measurements into strictly combined measurements and simultaneous measurements [4]. This classification is connected with a definite method used for processing experimental data to find the result of a measurement and to estimate its uncertainty.

In the case of direct measurements, the object of study is made to interact with the measuring instrument, and the value of the measurand is read from the indications of the latter. Sometimes the instrumental readings are multiplied by some factor, corresponding corrections are made in it, and so on.

In the case of indirect measurements, the value of the measurable quantity is found based on a known dependence between this quantity and its arguments. The

arguments are found by means of direct and sometimes indirect or simultaneous or combined measurements. For example, the density of a homogeneous solid body is found as the ratio of the mass of the body to its volume, and the mass and volume of the body are measured directly.

Sometimes direct and indirect measurements are not so easily distinguished. For example, an ac wattmeter has four terminals. The voltage applied to the load is connected to one pair of terminals, whereas the other pair of terminals is connected in series with the load. As is well known, the indications of a wattmeter are proportional to the power consumed by the load. However, the wattmeter does not respond directly to the measured power. Based on the principle of operation of the instrument, measurement of power with the help of a wattmeter would have to be regarded as indirect. In our case, it is important, however, that the value of the measurable quantity can be read directly from the instrument (in this case, the wattmeter). In this sense, a wattmeter is in no way different from an ammeter. For this reason, in this book, it is not necessary to distinguish measurement of power with the help of a wattmeter and measurement of current strength with the help of an ammeter: We shall categorize both cases as direct measurements. In other words, when referring a specific measurement to one or another category, we will ignore the arrangement of the measuring instrument employed.

Simultaneous and combined measurements employ close methods for finding the measurable quantities: In both cases, they are found by solving a system of equations, whose coefficients and separate terms are obtained as a result of measurements (usually direct). In both cases, the method of least squares is usually employed. But the difference lies in that in the case of combined measurements, several quantities of the same kind are measured simultaneously, whereas in the case of simultaneous measurements, quantities of different kinds are measured simultaneously. For example, a measurement in which the electric resistance of a resistor at a temperature of $+20\,^{\circ}\mathrm{C}$ and its temperature coefficients are found based on direct measurements of the resistance and temperature performed at different temperatures is a simultaneous measurement. A measurement in which the masses of separate weights in a set are found based on the known mass of one of them and by comparing the masses of different combinations of weights from the same set is a combined measurement.

Depending on the properties of the object of study, the model adopted for the object, and the definition of the measurable quantity given in the model as well as on the method of measurement and the properties of the measuring instruments, the measurements in each of the categories mentioned above are performed either with single or repeated observations. The method employed for processing the experimental data depends on the number of observations—are many measurements required or are one or two observations sufficient to obtain a measurement? If a measurement is performed with repeated observations, then to obtain a result the observations must be analyzed statistically. These methods are not required in the case of measurements with single observations. For this reason, for us, the number of observations is an important classification criterion.

We shall term measurements performed with single observations *single measurements* and measurements performed with repeated observations *multiple measurements*. An indirect measurement, in which the value of each of the arguments is found as a result of a single measurement, must be regarded as a single measurement.

Combined measurements can be regarded as single measurements, if the number of measurements is equal to the number of unknowns when the measurements are performed, so that each unknown is determined uniquely from the system of equations obtained.

Among combined measurements, it is helpful to single out measurements for which the measurable quantities are related by known equations. For example, in measuring the angles of a planar triangle, it is well known that the sum of all three angles is equal to 180°. This relation makes it possible to measure two angles only, and this is a single and, moreover, direct measurement. If, however, all three angles are measured, then the relation mentioned permits correlating their estimates, using, for example, the method of least squares. In the latter case, this is a combined and multiple measurement.

Measurements are also divided into static and dynamic measurements. Adhering to the concept presented in [49], we shall classify as static those measurements in which, in accordance with the problem posed, the measuring instruments are employed in the static regime and as dynamic those measurements in which the measuring instruments are employed in the dynamic regime.

The static regime of a measuring instrument is a regime in which, based on the function of the instrument, the output signal can be regarded as constant. For example, for an indicating instrument, the signal is constant for a time sufficient to read the instrument. A dynamic regime is a regime in which the output signal changes in time, so that to obtain a result or to estimate its accuracy, this change must be taken into account.

According to these definitions, static measurements include, aside from trivial measurements of length, mass, and so on, measurements of the average and effective (mean-square) values of alternating current by indicating instruments. Dynamic measurements refer to measurements of successive values of a quantity that varies in time (including stochastically). A typical example of such measurements is recording the value of a quantity as a function of time. In this case, it is logical to regard the measurement as consisting not of a single measurement but of many measurements.

Other examples of dynamic measurements are measurement of the magnetic flux by the ballistic method and measurement of the high temperature of an object based on the starting section of the transfer function of a thermocouple put into contact with the object for a short time (the thermocouple would be destroyed if the contact time was long).

Static measurements also include measurements performed with the help of digital indicating instruments. According to the definition of static measurements, for this conclusion, it is not important that during the measurement, the state of the elements in the device changes. The measurement will also remain static when

the indications of the instrument change from time to time, but each indication remains constant for a period of time sufficient for the indication to be read or recorded automatically.

A characteristic property of dynamic measurements is that to obtain results and estimate their accuracy in such measurements, it is necessary to know a complete dynamic characteristic of the measuring instrument: a differential equation, transfer function, and so on. (The dynamic characteristics of measuring instruments will be examined in Chapter 2.)

The classification of measurements as static and dynamic is justified by the difference in the methods employed to process the experimental data. At the present time, however, dynamic measurements as a branch of metrology are still in the formative stage.

The most important characteristic of the quality of a measurement is accuracy. The material base, which ensures the accuracy of numerous measurements performed in the economy, consists of reference standards. The accuracy of any particular measurement is determined by the accuracy of the measuring instruments employed, the method of measurement employed, and sometimes by the skill of the experimenter. However, as the true value of a measurable quantity is always unknown, the errors of measurements must be estimated computationally (theoretically). This problem is solved by different methods and with different accuracy.

In connection with the accuracy of the estimation of a measurement error, we shall distinguish measurements whose errors are estimated before and after the measurement. We shall refer to them as measurements with ante-measurement or a priori estimation of errors and measurements with postmeasurement or a posteriori estimation of errors.

Estimates with ante-measurement estimation of errors must be performed according to an established procedure, included in the calculation of the errors. Measurements of this type include all mass measurements. For this reason, we shall call them mass measurements. Sometimes they are called technical measurements.

Mass measurements are common. Their accuracy is predetermined by the types (brands) of measuring instruments indicated in the procedure, the techniques for using them, as well as the stipulated conditions under which the measurements are to be performed. For this reason, the person performing the measurement is interested only in the result of measurement, and he or she knows nothing about the accuracy beforehand, i.e., whether it is adequate.

A posteriori estimation of errors is characteristic for measurements performed with an objective in mind, when it is important to know the accuracy of each result. For this reason, we shall call such measurements individual measurements.

We shall divide individual measurements, in turn, into two groups: measurements with exact estimation of errors and measurements with approximate estimation of errors.

Measurements with exact estimation of errors are measurements in which the properties of the specific measuring instruments employed are taken into

account. *Measurements with approximate estimation of errors* are measurements in which the specifications of the measuring instruments employed are taken into account.

In both cases, the conditions under which the measurements are performed are taken into account. For this reason, the influence quantities or some of them are often measured; in other cases, they are estimated.

Here we must call attention to a fact whose validity will become obvious from further discussion. Suppose that measurements whose errors are estimated with different accuracy are performed with the same measuring instruments. Despite that the same instruments are employed, the accuracy of the measurements in this case is different. The most accurate result will be the result obtained with exact estimation of the errors. Measurements for which the errors are estimated approximately will in most cases be more accurate than measurements whose errors are estimated beforehand. The results of measurements with ante- and postmeasurement estimation of errors will be only rarely equally accurate.

But when measuring instruments having different accuracy are employed, this result will no longer be the case. For example, measurement of voltage with the help of a potentiometer of accuracy class 0.005, performed as a mass measurement, i.e., with preestimation of errors, will be more accurate than measurement with the help of an indicating voltmeter of class 0.5, performed as an individual measurement with exact estimation of the errors.

In all cases studied above, the objective of the measurements was to obtain an estimate of the true value of the measurable quantity, which, strictly speaking, is the problem of any measurement. However, measurements are often performed during the preliminary study of a phenomenon. We shall call such measurements *preliminary measurements*.

The purpose of preliminary measurements is to determine the conditions under which some indicator of the phenomenon can be observed repeatedly and its regular relations with other properties of the object, systems of objects, or with an external medium can be studied. As the object of scientific investigation of the world is to establish and study regular relations between objects and phenomena, preliminary measurements are important. Thus, it is known that the first task of a scientist who is studying some phenomenon is to determine the conditions under which a given phenomenon can be observed repeatedly in other laboratories and can be checked and confirmed.

Preliminary measurements, as one can see from the concepts presented above, are required to construct a model of an object. For this reason, preliminary measurements are also important for metrology.

Apart from preliminary measurements, for metrological purposes it is also possible to distinguish supplementary measurements. Supplementary measurements are measurements of influence quantities that are performed to determine and make corrections in the results of measurements.

Enormous literature exists on different aspects of measurements. Massey [38] gives an idea of the wide range of questions pertaining to real measurements.

1.5. Classification of Measurements Errors

A measurement of a quantity whose true value is A gives an estimate $\tilde{A}$ of that quantity. The absolute measurement error ζ expresses the difference between $\tilde{A}$ and A : $\zeta = \tilde{A} - A$.

However, this equation cannot be used to find the error of a measurement for the simple reason that the true value of the measurable quantity is always unknown. If the true value was known, then there would be no need for a measurement.

Measurements performed for calibration of measuring instruments are an exception. In this case, the value of the measurable quantity must be known with sufficient accuracy so that it can be used for this purpose instead of the true value of the quantity.

For this reason, measurement errors must be estimated with indirect data.

The necessary components of any measurement are the method of measurement and the measuring instrument; measurements are often performed with the participation of a person. The imperfection of each component of measurement contributes to the measurement error. For this reason, in the general form,

$$\zeta = \zeta_m + \zeta_i + \zeta_p,$$

where ζ is the measurement error, ζ_m is the methodological error, ζ_i is the instrumental error, and ζ_p is the personal error.

Each component of the measurement error can in turn be caused by several factors. Thus, *methodological errors* can arise as a result of an inadequate theory of the phenomena on which the measurement is based and inaccuracy of the relations that are employed to find an estimate of the measurable quantity. In particular, the error caused by the threshold discrepancy between the model of a specific object and the object itself is also a methodological error.

Instrumental measurement errors are caused by the imperfection of measuring instruments. Normally the intrinsic error of measuring instruments, i.e., the error obtained under reference conditions regarded as normal, are distinguished from additional errors, i.e., errors caused by the deviation of the influence quantities from their values under reference conditions. Properties of measuring instruments that cause the instrumental errors will be examined in detail in Chapter 2.

Personal errors: Measurements are often performed by people. Someone reads the indications of instruments, and records the moment at which an image of a filament vanishes on the screen of an optical pyrometer. The individual characteristics of the person performing the measurement give rise to individual errors that are characteristic of that person. They include errors caused by incorrect reading of the tenths graduation of an instrument scale, asymmetric placement of the mark of an optical indicator between two graduation lines, and so on.

Improvement of the reading and regulating mechanisms of measuring instruments has led to the fact that for modern measuring instruments, the personal errors are usually insignificant; for example, for digital instruments, they are virtually nonexistent.

The foregoing classification of measurement errors is based on the cause of the errors.

Another important classification of measurement errors is based on their properties. In this respect, systematic and random errors are distinguished.

A measurement error is said to be *systematic* if it remains constant or changes in a regular fashion in repeated measurements of one and the same quantity. The observed and estimated systematic error is eliminated from measurements by introducing corrections. However, it is impossible to eliminate completely the systematic error in this manner. Some part of the error will remain, and then this residual error will be the systematic component of the measurement error.

To define a random measurement error, imagine that some quantity is measured several times. If there are differences between the results of separate measurements, and these differences cannot be predicted individually and any regularities inherent to them are manifested only in many results, then the error from this scatter of the results is called the *random error*.

The division of measurement errors into systematic and random is important, because these components are manifested differently and different approaches are required to estimate them.

Random errors are discovered by performing measurements of one and the same quantity repeatedly under the same conditions, whereas systematic errors can be discovered experimentally either by comparing a given result with a measurement of the same quantity performed by a different method or by using a more accurate measuring instrument. However, systematic errors are normally estimated by theoretical analysis of the measurement conditions, based on the known properties of a measurand and of measuring instruments. Other specifics of the terms systematic and random errors are discussed in Section 3.2.

The quality of measurements that reflects the closeness of the results of measurements of the same quantity performed under the same conditions is called the *repeatability of measurements*. Good repeatability indicates that the random errors are small.

The quality of measurements that reflects the closeness of the results of measurements of the same quantity performed under different conditions, i.e., in different laboratories (at different locations) and using different equipment, is called the *reproducibility of measurements*. Good reproducibility indicates that both the random and systematic errors are small.

In speaking about errors, we shall also distinguish gross or outlying errors and blunders. We shall call an error *gross* (outlying) if it significantly exceeds the error justified by the conditions of the measurements, the properties of the measuring instrument employed, the method of measurement, and the qualifications of the experimenter. Such measurements can arise, for example, as a result of a sharp, brief change in the grid voltage (if the grid voltage in principle affects the measurements).

Outlying or gross errors in multiple measurements are discovered by statistical methods and are usually eliminated from analysis.

Blunders occur as a result of errors made by the experimenter. Examples are a slip of the pen when writing up the results of observations, an incorrect reading of the

indications of an instrument, and so on. Blunders are discovered by nonstatistical methods, and they must always be eliminated from the analysis.

Measurement errors are also divided into static and dynamic. Static errors were mentioned above. Dynamic errors are caused by the inertial properties of measuring instruments.

If a varying quantity is recorded with the help of a recording device, then the difference between the obtained function and the actual process of change of the recorded quantity in time (taking into account the necessary scale transformations) is the dynamic error of the given dynamic measurement. In this case, it is also a function of time, and the instantaneous dynamic error can be determined for each moment in time.

We shall now study the case when the process is recorded by measuring individual instantaneous values. It is clear that if within the time of a single measurement, the measurable quantity does not change significantly and the instantaneous values of the process are obtained at known times and sufficiently frequently, then the collection of points ultimately obtained gives a picture of the change of the measurement in time with a negligibly small error. Thus, there will be no dynamic error here.

The inertial properties of an instrument can be such, however, that the changes in the measurable quantity over the measurement time will give rise to a definite error in the results of measurements of the instantaneous values. In this case, the obtained collection of instantaneous values will not be coincident with the process of change of the measurable quantity in time, and their difference, exactly as in the case of an analog automatic-plotting instrument, will give the dynamic error. Correspondingly, it is natural to call the *instantaneous dynamic error* the error arising in the measurement of a separate instantaneous quantity as a result of the rate of change of the measurable quantity and the inertial properties of the instrument.

If some isolated instantaneous quantity, for example, the amplitude of a pulse, is measured and the measurement is performed with a special indicating device, then the difference between the shape of the pulse and the shape obtained with a calibrated instrument will give rise to a supplementary error as a result of the measurement. Based on what we have said above, one could call this error a dynamic error. However, the general term "dynamic error" is normally avoided in such situations, and such an error is given a name that indicates its cause. In this example, it is natural to call the error the pulse-shape error. In practice, the pulse shape is characterized by several parameters, and to each parameter, a separate component of the error is associated.

1.6. Principles of Estimation of Measurement Errors and Uncertainties

Measurements are regarded metrologically to be better the lower their uncertainty is. However, measurements must be reproducible, because otherwise they lose their objective character and therefore become meaningless.

A measure of the nonreproducibility of a measurement permitted by the experimenter is the limits of measurement error or uncertainty estimated by the experimenter.

Correctly estimated measurement uncertainty permits comparing the obtained result with the results obtained by other experimenters. The fact that the correctness of a given estimate is later confirmed in a more accurate measurement attests to the high skill of the experimenter.

The validity of the uncertainty calculated for every measurement result is based on the validity of the estimates of errors of this measurement. Therefore, we shall pay attention to the practice of estimating measurement errors.

Errors are customarily estimated based on the following considerations:

(1) The measurement is regarded as more accurate and, in this sense, better, the smaller its relative error is

$$\varepsilon = \frac{\zeta}{A} \approx \frac{\Delta}{\tilde{A}},$$

where $\tilde{A}$ is an estimate of the true value of the measurand A and Δ is an estimate of the limits of measurement error ζ. The relative error is studied because its value, as a rule, does not depend on the value of the measurable quantity. This condition is not true for the absolute error.

(2) The estimate of the measurement error must satisfy the inequality

$$|\zeta| \leq |\Delta|.$$

The meaning of this inequality is as follows. For any measurement, it is, in principle, desirable that no error exists. But despite all efforts $\tilde{A} \neq A$ and correspondingly $\zeta \neq 0$, and in our case, the error can be both greater than and less than zero. In the primary estimate of any error, the limits for ζ, i.e., Δ_1 and Δ_2, are established, so that $\Delta_1 \leq \zeta \leq \Delta_2$. In the calculations, the value of $|\Delta_1|$ or $|\Delta_2|$, whichever is larger, is often used as the estimate for ζ. Most often, $|\Delta_1| = |\Delta_2| = |\Delta|$. Thus, we arrive at the inequality $|\zeta| \leq |\Delta|$.

If a measurement error is mainly random, its limits should be estimated with such a high probability that the above inequality is practically always satisfied.

The second assumption means that the estimate of the measurement error must be an upper estimate. This requirement should be regarded as a principle of error estimation.

It makes a great deal of sense to use the limiting values of an error as estimates of the error. First, humans naturally make comparisons and they establish relations of the type greater than, equal to, and less than. For this reason, limiting errors are easily perceived. Furthermore, measurement results are most easily compared when the limiting errors are known.

It should not be forgotten, however, that the measurement error characterizes the uncertainty of the result of a measurement and the spread and nonreproducibility of the measurement. For this reason, the estimate of error cannot be precise, and it is not required that it be precise. But its uncertainty, if one can say so, should be weighted toward overestimation and not underestimation. In accordance with this

principle, we shall still say that it is better to overestimate than to underestimate an error: In the first case, the quality of the measurement is reduced, whereas in the second case, the entire measurement can be made worthless. Of course, the overestimation should be kept to a minimum.

It should also be kept in mind that the correctness of estimates of measurement errors or uncertainty cannot be checked based on data obtained in a particular measurement. Regarding a given measurement, all obtained experimental data and other reliable information, for example, corrections to the indications of instruments, are employed to find the measurement result, and the error must be estimated with additional information about the properties of the measuring instruments, the conditions of the measurements, and the theory. There is no point in performing a special experiment to check or estimate the measurement error or uncertainty. It would then be necessary to organize in parallel with the given measurement a more accurate measurement of the same measurable quantity. Then the given measurement would be meaningless: Its result would be replaced by the result of the more accurate measurement. The problem of estimating the error in a given measurement would be replaced by the problem of estimating the error of the more accurate measurement; i.e., the basic problem would remain unsolved.

The correctness of estimates of errors and uncertainty is nonetheless checked. It is confirmed either by the successful use of the measurement result for the purpose intended or by the fact that the measurement agrees with the results obtained by other experimenters. As in the case of measurement of physical constants, the correctness of the estimates of uncertainties is sometimes checked with time as a result of improvements in measuring instruments and methods of measurement and increased measurement accuracy.

1.7. Presentation of Results of Measurements; Rules for Rounding Off

If $\tilde{A}$ is the result of a measurement and Δ_U and Δ_L are the upper and lower limits of the error in the measurement, then the result of the measurement and the measurement error can be written in the form

$$\tilde{A}, \quad \Delta_U, \quad \Delta_L, \quad \text{and} \quad \tilde{A}_{\Delta_L}^{\Delta_U}.$$

In the first case, it is convenient to preserve the conventional notation, for example, $\tilde{A} = 1.153$ cm, $\Delta_U = +0.002$ cm, and $\Delta_L = -0.001$ cm. The second case is convenient for technical documentation. For example, $1.153_{-0.001}^{+0.002}$ cm. Often $|\Delta_U| = |\Delta_L| = \Delta$. Then the result and the error are written in the form $\tilde{A} \pm \Delta$.

In the cases above, the inaccuracy of measurement results is expressed as permissible errors. But usually, the inaccuracy is expressed as uncertainty. In this case, the corresponding probability must be given. For uniformity, it is recommended that the probability be given in parentheses after the value of the uncertainty or a symbol of a measurand.

For example, if a measurement gives the value of the voltage, 2.62 V, and the uncertainty of this result, $u = \pm 2\%$, was calculated for the probability 0.95, then the result will be written in the form

$$\tilde{U} = 2.62V, \quad u = \pm 2\%(0.95),$$

or in the more compact form

$$U(0.95) = (2.62 \pm 0.05)V.$$

The remark regarding compactness refers to the method for indicating the value of the probability and is unrelated to the fact that the relative error is given in the first case and the absolute error is given in the second case.

If the confidence probability is not indicated in the measurement result, then the inaccuracy must be assumed to have been estimated without the use of probability methods. For example,

$$U = (2.1 \pm 0.1)V.$$

Although an error estimate obtained without the use of probability methods can be reliable, it cannot be associated with a probability of one or some other value. As a probabilistic model was not employed, the probability cannot be estimated and it should not be indicated. So, in this case again, we have the limits of an error of a measurement.

The form, examined above, for representing measurement errors is desirable for the final result, intended for direct practical application, for example, in quality control problems. In this case, it is usually convenient to express the error in the form of absolute errors. In many cases, however, it is desirable to know not the limiting values of the total measurement error but the characteristics of the random and systematic components separately. Such a representation of the error makes it easier to analyze and determine the reasons for any discrepancy between the results of measurements of one and the same quantity performed under different conditions. Such an analysis is usually necessary in the case of measurements performed for scientific purposes, for example, measurements of physical constants. It is also desirable to record the components separately in those cases when the result of a measurement is to be used for calculations together will other data that are not absolutely precise. For example, for errors of measurements of quantities measured directly in indirect measurements, recording the error in this form makes it possible to estimate more accurately the uncertainty of the result of the indirect measurement.

When the error components are recorded separately, the systematic component is characterized, as a rule, by the limiting values θ_U, θ_L, and θ, if $|\theta_U| = |\theta_L| = \theta$. If these limits are calculated by probabilistic methods, then the probability employed should be indicated in parentheses immediately after the value of the error. For a random error, the standard deviation $S_{\bar{x}}$, determined from the experimental data and the number of observations n are usually indicated. Sometimes the uncertainty and the corresponding probability are given instead of $S_{\bar{x}}$.

For scientific measurements, apart from the above-indicated parameters of the error, it is helpful to describe the basic sources of error together with an estimate

of their contribution to the total measurement uncertainty. For the random error, it is of interest to present the form and parameters of the distribution function of the observations and how it was determined (the method employed for testing the hypothesis regarding the form of the distribution function, the significance level used in this testing, etc.).

The errors in the results of mass measurements are usually not indicated at all, because they are estimated and are known beforehand.

The number of significant figures employed in the number expressing the result of a measurement must correspond to the accuracy of the measurement, which means that the uncertainty of a measurement can be equal to 1 or 2 units in the last figure of the number expressing the result of the measurement. In any case, this uncertainty should not exceed 5 units in the last figure.

As measurement uncertainty determines only the vagueness of the resuits, it need not be known precisely. For this reason, in its final form, the uncertainty is customarily expressed by a number with one or two significant figures. Two figures are retained for the most accurate measurements and if the most significant digit of the number expressing the uncertainty is equal to or less than 3.

It should be noted, however, that in calculations and intermediate computations, depending on the computational operations performed, one or two significant figures more than suggested by the result should be retained so that the roundoff error not distort the results too much.

The numerical value of the result of a measurement must be represented in a manner so that the last decimal digit is of the same rank as its uncertainty. There is no point in including more digits, because this will not reduce the uncertainty of the result. But less digits, which can be obtained by further rounding off the number, would increase the uncertainty and make the result less accurate, and thereby it would make pointless the measures employed in the measurement.

When analyzing the results of observations and recording the results of measurements, the rounding off should be done according to the following rules:

(1) The last digit retained is not changed if the adjacent digit being discarded is less than 5. Extra digits in integers are replaced by 0's, whereas extra digits in decimal fractions are dropped.

Examples. The numerical value of the result of a measurement 85.6342 with an error in the limits ± 0.04 should be rounded off to 85.63. If the error limits are ± 0.012, the same number should be rounded off to 85.634.

Retaining four significant figures, the number 165,245 should be rounded off to 165.2.

(2) The last digit retained is increased by 1 if the adjacent digit being discarded is greater than 5 or if it is 5 and there are digits other than 0 to its right.

Examples. If three significant digits are retained, the number 18,598 is rounded off to 18,6 and the number 152.56 is rounded off to 153.

(3) If the digit being discarded is equal to 5 and the digits to its right are unknown or are equal to 0, then the last retained digit is not changed if it is even and it is increased by 1 if it is odd.

Examples. If two significant digits are retained, the number 10.5 is rounded off to 10 and the number 11.5 is rounded off to 12.

(4) If the decimal fraction in the numerical value of the result of a measurement terminates in 0's, then the 0's are dropped only up to the digit that corresponds to the rank of the numerical value of the error.

The foregoing rules were established by convention, and for calculations performed by humans, they are entirely satisfactory. In the case of calculations performed with the help of computers, however, rounding off depending on the evenness or oddness of the last retained digit [rule (3)] is inconvenient, because it complicates the algorithm. It has been suggested that this rule be dropped and the last retained figure not be changed, irrespective of whether it is even or odd. This suggestion, however, has not been adopted. The main objection is that many such roundoffs of intermediate results can significantly distort the final result.

If the rules presented above are used, then the number of significant figures in the numerical value of the result of a measurement makes it possible to judge approximately the accuracy of a measurement. For this reason, it should be noted that the limiting error, caused by roundoff, is equal to one half the last digit in the numerical value of the result of the measurement, and the measurement error can reach two units in the next-to-last and several units in the last digit.

We shall now estimate the relative roundoff error. Assume, for example, that the result of a measurement is expressed by a number with two significant figures. Then the minimum number will be equal to 10 and the maximum number will be equal to 99. Therefore, the relative roundoff error will be $0.5\% < \varepsilon_2 \leq 5\%$.

If the result of a measurement is expressed by a number with three significant figures, this error will fall in the range $0.05\% < \varepsilon_3 \leq 0.5\%$, and so on.

The error limits obtained above show the effect of roundoff on the measurement error. In addition, these data permit focusing on the minimum number of significant figures necessary to record the result of a measurement with the prescribed accuracy.

When analyzing the results of observations and, especially, when estimating the errors, it is useful to employ methods and formulas of approximate calculations.

1.8. Basic Conventional Notations

We shall employ Latin letters for the measurands. Greek letters will be employed for errors.

The notations employed for errors, their limits, and uncertainties are presented in Table 1.1.

The notations for confidence limits of errors are constructed by adding to the symbol of the corresponding error a subscript α.

We shall distinguish estimates of quantities from the true values by adding a tilde to the corresponding symbol. For example, $\tilde{A}$ is the estimate of the true value of A.

TABLE 1.1. Designations of measurement errors and uncertainties.

Uncertainty	u
Combined uncertainty	u_c
Total uncertainty	u_t
Error	ζ
Limits of error	Δ
Random error	ψ
Limits of random error	Ψ
Conditionally constant error	ϑ
Limits of conditionally constant error	θ
Absolutely constant error	η
Limits of absolutely constant error	H

We shall denote the arithmetic mean with the help of an overbar on the corresponding symbol. For example, $\bar{x}$ is the arithmetic mean of the obtained values of $x_i (i = 1, \ldots, n)$.

In addition, we shall use the following mathematical symbols: $E[X]$ is the mathematical expectation, and $V[X]$ is the variance of a random quantity X.

Of the notation used for specific concepts, we present the following: p is the probability at which an event first occurs, α is the confidence probability, q is the significance level, σ^2 is the variance of a random quantity, σ is the rms or standard deviation, and S^2 and S are the estimations of σ^2 and σ.

2
Measuring Instruments and Their Properties

2.1. Types of Measuring Instruments

Measuring instruments are the technical objects that are specially developed for the purpose of measuring specific physical quantities. A general property of measuring instruments is that their accuracy is standardized.

Measuring instruments are divided into material measures, measuring transducers, indicating instruments, measuring setups, and measuring systems.

A *material measure* is a measuring instrument that reproduces one or more known values of a given physical quantity. Examples of measures are balance weights, measuring resistors, and measuring capacitors.

Single-valued measures, multiple-valued measures, and collections of measures are distinguished. Examples of multiple-valued measures are graduated rulers, measuring tapes, resistance boxes, and so on.

In addition to multiple-valued measures, which reproduce discrete values of quantities, multiple-valued measures exist that continuously reproduce quantities in some range, for example, a measuring capacitor with variable capacitance. Continuous measures are usually less accurate than discrete measures.

When measures are used to perform measurements, the measurands are compared with the known quantities reproduced by the measures. The comparison is made by different methods, but so-called comparators are a specific means that are used to compare quantities. The simplest comparator is the standard equal-armed pan balance.

A *comparator* is a measuring device that makes it possible to compare similar physical quantities and has a known sensitivity.

In some cases, quantities are compared without comparators, by experimenters, with the help of their viewing or listening perceptions.

Thus, when measuring the length of a body with the help of a ruler, the ruler is placed on the body and the observer fixes visually the graduations of the ruler (or fractions of a graduation) at the corresponding points of the body.

A *measuring transducer* is a measuring instrument that converts the measurement signals into a form suitable for transmission, processing, or storage. The

measurement information at the output of a measuring transducer cannot, as a rule, be directly observed by the observer.

It is necessary to distinguish measuring transducers and the transforming elements of a complicated instrument. The former are measuring instruments, and as such, they have standard metrological properties (see below). The latter, on the other hand, do not have an independent metrological significance and are not used separately from the instrument of which they are a part.

Measuring transducers are diverse. Examples are thermocouples, resistance thermometers, measuring shunts, the measuring electrodes of *p*H meters, and so on. Measuring current or voltage transformers and measuring amplifiers are also measuring transducers. But this group of transducers is characterized by the fact that the signals at their inputs and outputs are a physical quantity of the same form, and only the dimension of the quantity changes. For this reason, these measuring transducers are called scaling measuring transducers.

Measuring transducers that convert an analog quantity at the input (the measurand) into a discrete signal at the output are called analog-to-digital converters. Such converters are manufactured in the form of autonomous, i.e., independent measuring instruments, and in the form of units built into other instruments, in particular, in the form of integrated microcircuits. Analog-to-digital converters are a necessary component of digital devices, but they are also employed in monitoring, regulating, and control systems.

An *indicating instrument* is a measuring instrument that is used to convert measurement signals into a form that can be directly perceived by the observer.

Based on the design of the input circuits, indicating instruments are just as diverse as measuring transducers, and it is difficult to survey all of them. Moreover, such a review and even classification are more important for designing instruments than for describing their general properties.

A common feature of all indicating instruments is that they all have readout devices. If these devices are implemented in the form of a scale and an indicating needle, then the indications of the instrument are a continuous function of the measurand. Such instruments are called analog instruments. If the indications of instruments are in a digital form, then such instruments are called digital instruments.

The definition of digital instruments presented above formally includes both automatic digital voltmeters, bridges, and similar instruments and induction meters for measuring electrical energy. In these instruments, however, the measuring transformations are performed in a discrete form, and in the case of induction meters, all measuring transformations of signals occur in an analog form and only the output signal assumes a discrete form. The conversions of measurement information into a discrete form have several specific features. Therefore, only instruments in which the measurement conversions occur in a discrete form are usually considered to be digital instruments.

The indications of digital instruments are easily recorded and are convenient for entering into a computer. In addition, their design usually makes it possible to obtain significantly higher accuracy than analog instruments. Moreover, when

digital instruments are employed, no reading error occurs. However, when analog instruments are used, it is easier to judge trends in the variation of the measurands.

In addition to analog and digital instruments, analog-discrete measuring instruments also exist. In these instruments, the measuring conversions are performed in an analog form, but the readout device is a discrete unit. The readout device has a scale and a glowing strip, whose length changes discretely, which plays the role of the indicator. Sometimes the indicator is a glowing dot that moves along a scale.

Analog-discrete instruments combine the advantages of both analog and digital instruments. Induction meters for measuring electric energy are examples of such hybrid instruments.

In many cases, measuring instruments are designed so that their indications are recorded. Such instruments are said to be *recording instruments.* Data can be recorded in the form of a continuous record of the variation of the measurand in time or in the form of a series of points of this dependence.

Instruments of the first type are called automatic-plotting instruments, and instruments of the second type are called printing instruments. Printing instruments can record the values of a measurand in digital form. Printing instruments give a discrete series of values of the measurand in some interval of time. The continuous record provided by automatic-plotting instruments can be regarded as an infinite series of values of the measurand.

Sometimes measuring instruments are equipped with induction, photooptical, or contact devices and relays for purposes of control or regulation. Such instruments are called regulating instruments. Designers strive to design regulating units so as not to reduce the accuracy of the measuring instrument. However, this is rarely possible.

Measuring instruments also customarily include comparators, mentioned above, for comparing measures, and null indicators, for example, galvanometers. The reason is that a comparator with a collection of measures becomes a comparison measuring instrument, whereas a galvanometer can be used as a highly sensitive indicating instrument.

A *measurement setup* is a collection of functionally and structurally integrated measuring instruments and auxiliary devices that provides efficient organization of the measurements. An example is the potentiometric setup for electric measuring instrument calibration.

A *measuring system* is a collection of functionally unified measuring, computing, and auxiliary means for obtaining measurement information and for converting and processing it to provide the user with information in the required form, introducing it into the control system, or performing logical functions automatically. Modern measuring systems include microprocessors and even entire computers, and apart from processing and providing output of the measurement information, they can control the measurement process.

Finally, systems whose units must, in accordance with the purpose of the system, operate under the same conditions can be distinguished from systems whose units operate under different conditions. We shall call the former *uniform measuring systems* and the latter *nonuniform measuring systems.* This classification makes it

easier to study questions concerning the metrological support of measuring systems and the calculation of the errors of such systems.

2.2. The Concept of an Ideal Instrument: Metrological Characteristics of Measuring Instruments

Any technical object can be described by a collection of characteristics. Measuring instruments are not an exception in this respect. We shall divide all characteristics of measuring instruments into two groups: metrological, which is necessary for using a measuring instrument in the manner intended, and secondary. We shall include in the latter such characteristics as mass, dimensions, and degree of protection from moisture and dust. We shall not discuss characteristics of the secondary group, although sometimes they determine the selection and application of an instrument, because they are not directly related with the measurement accuracy.

By metrological characteristics of a measuring instrument, we mean the characteristics that make it possible to judge the suitability of the instrument for performing measurements in a known range with known accuracy, to obtain a value of the measurand, and to estimate its inaccuracy.

To sort the metrological characteristics, it is helpful to introduce the concept of an ideal instrument. The ideal of any technical object is its design, i.e., its model. For measuring instruments, this is not sufficient, because such a device must contain also an "impression" of the corresponding unit of measurement. The impression of the unit cannot be prepared; it must be obtained from a standard. We shall give several examples.

Gauge block. The ideal is a completely regular parallelepiped, one edge of which is determined exactly for the established units of length.

Measure of constant voltage. The ideal is a source of constant voltage with a value that is known exactly and that is free of any noise at the output.

Measuring transformer. The ideal is a voltage or current transformer with a conversion factor that is known exactly and that has no losses and parasitic noise in the input and output circuits.

Integrating analog-to-digital converter of voltage. The ideal is an instrument with an output voltage U_0 that is related to the input voltage U_x by the dependence

$$U_0 = K_1 \int_{t_1}^{t_2} U_x dt,$$

where $K_1 = \text{const}$ and $\Delta t = t_2 - t_1 = \text{const}$.

As a result, we obtain $U_0 = K_2 U_x$. The voltage U_0 can be quantized, and a code, reflecting the voltage at the input without any distortions (with the exception of the quantization error, which can be made to be small), is obtained at the output of the converter.

Moving-coil ammeter. In this instrument, a constant current flows through a moving coil and forms with the help of a permanent magnet a mechanical moment, which twists a spring to a point of balance. In the process, an indicating needle is deflected by an amount along a scale that is proportional to the current strength. Each value of the current strength corresponds to a definite indication, which is fixed by calibrating the instrument.

Thus, the ideal instrument performs (theoretically) a series of single-valued transformations and after calibration acquires a precise scale in the units of the measurand.

An ideal representation of each type of measuring instruments is formed for a specific model of the objects—carriers of the corresponding physical quantity.

We shall call the metrological characteristics established for ideal measuring instruments of a specific type the *nominal metrological characteristics.* An example of such a characteristic is the nominal value of the measure (10 Ω, 1 kG, etc.), the measurement range of the instrument (0–300 V, 0–1200 °C, etc.), the conversion range of the transducer, the value of the scale factor of the instrument scale, and so on.

The relation between the input and the output signals of instruments and transducers is determined by the transfer function. For instruments, it is fixed by the scale, whereas for measuring transducers, it is determined by a graph or an equation. Either the graph or the equation represents the nominal metrological characteristic if the graph or equation was determined (indicated) before these measuring instruments were developed.

The real characteristics of measuring instruments differ from the nominal characteristics because of fabrication errors and changes occurring in the corresponding properties in time.

An ideal measuring instrument (transducer) would react only to the measured physical quantity or to the parameter of the input signal of interest, and its indication would not depend on the external conditions, the power supply regime, and so on. For a real measuring transducer, as for other types of measuring instruments, these undesirable phenomena occur.

The quantities characterizing the external conditions are called influence quantities.

For some types of measuring instruments, the dependence of the output signal, the indications, or the error from one or another influence quantity can be represented as a functional dependence, called the influence function. The influence function can be expressed in the form of an equation (for example, the temperature dependence of the emf of standard cells) or a graph. In the case of a linear dependence, it is sufficient to give the coefficient of proportionality between the output quantity and the influence quantity. We shall call this coefficient the influence coefficient.

Influence coefficients and functions make it possible to take into account the conditions under which measuring instruments are used by introducing the corresponding corrections. The imperfection of measuring instruments is also manifested because when one and the same quantity is measured repeatedly under

identical conditions, the results can differ somewhat from one another. In this case, it is said that the indications are nonrepeatable.

The inaccuracy of a measuring instrument is usually characterized by its error. We shall explain this concept for the example of an indication measuring instrument. Let the true value of a quantity at the input of the instrument be A_t. The instrument indicates the value A_r. The absolute error of the instrument will be

$$\zeta = A_r - A_t.$$

The nonrepeatability of the indications of the instrument is manifested by the fact that when A_t is measured repeatedly, the indications of the instrument will be somewhat different. For this reason, one can talk about a random component of instrument error. This component is referred to as the repeatability error of a measuring instrument.

The random component of instrument error is normally caused by friction in the supports of a movable part of the instrument and hysteresis phenomena, and its limits are sharp. The limits can be found experimentally if the quantity measured by the instrument varies continuously. The strength of the electric current, the voltage, and other quantities can be varied continuously. Correspondingly, the indications of ammeters and voltmeters can vary continuously. The indications of weighing balances and several other instruments cannot be varied continuously.

For instruments whose indications can vary continuously, the limits of the random error are found by continuously driving the indicator of the instrument up to the same scale marker, first from below and then from above (or vice versa) a marker. We will call *the dead band* the absolute value of the difference of the values of the measurand that are obtained in such a test and that correspond to a given scale marker of the instrument.

The dead band is the length of the range of possible values of the random component of instrument error, and one half of this length is the limiting value of the random.

Figure 2.1 shows graphs of the "input-output" in the presence of (a) only friction, (b) only hysteresis, and (c) friction together with hysteresis. These examples of processes reveal dead bands.

The random error of weighing scales is usually characterized by the standard deviation [12]. This characteristic of an instrument is calculated from the changes produced in the indications of the scales by a load with a known mass; the test is performed at several scale markers, including the limits of the measurement range. One method for performing the tests, and the computational formula for calculating the standard deviation of weighing scales, are presented in [12].

Measuring instruments are created to introduce certainty into the phenomena studied and to establish regular relations between the phenomena, and the uncertainty created by the non-single-valuedness of instrument indications interferes with using an instrument in the manner intended. For this reason, the first problem that must be solved when developing a new measuring device is to make its

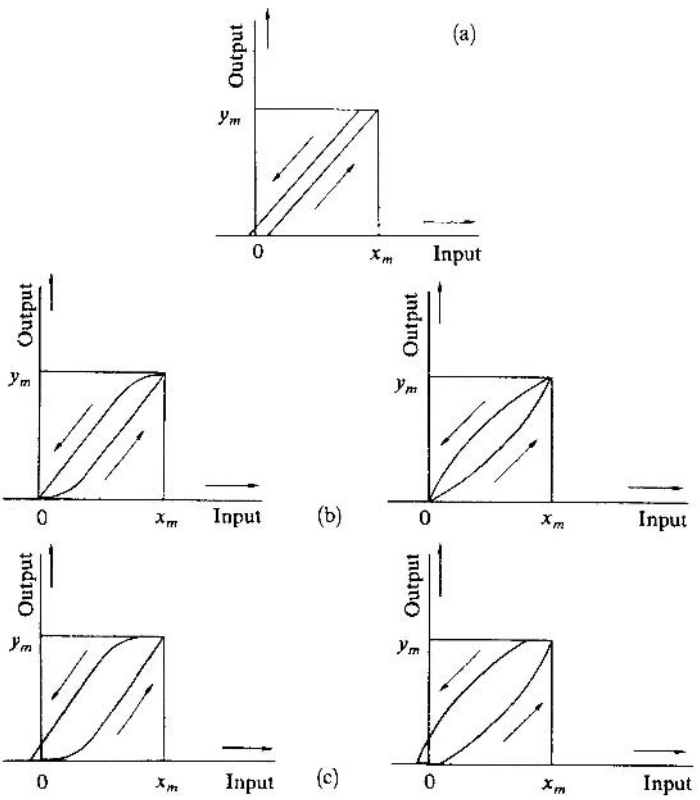


FIGURE 2.1. Dependences between the input and output of the instrument with a dead band. (a) In the presence of a friction, (b) in the presence of hysteresis (two types, for example), and (c) in the presence of a friction and hysteresis (two types mentioned above).

random error insignificant, i.e., either negligibly small compared with other errors or falling within prescribed limits as the limits of admissable errors for measuring devices of the given type.

If the random error is insignificant and the elements determining instrument accuracy are stable, then by calibration, the measuring device can always be "tied" to a corresponding standard and the potential accuracy of the instrument can be realized.

The value of a scale division or the value of a significant figure is the value of the measurand corresponding to the interval between two neighboring markers on the instrument scale or one figure of some digit of a digital readout device.

The *sensitivity* is the ratio of the change in the output value of the measuring instruments to the input value of the quantity that causes the output value to change. The sensitivity can be a nominal metrological characteristic and an actual characteristic of a real instrument.

The *discrimination threshold* is the minimum change in the input signal that causes an appreciable change in the output signal.

The *resolution* is the smallest interval between two distinguishable neighboring discrete values of the output signal.

Instability (of a measuring instrument) is a general term that expresses the change in any property of the measuring instrument in time.

Drift is the change occurring in the output signal (always in the same direction) over a period of time that is significantly longer than the measurement time when using a given measuring instrument.

The drift and the instability do not depend on the input signal or the load, but they can depend on the external conditions. The drift is usually determined in the absence of a signal at the input.

The metrological characteristics of measuring instruments should also include their dynamic characteristics. These characteristics reflect the inertial properties of measuring instruments. It is necessary to know them to correctly choose and use many types of measuring instruments. The dynamical characteristics are examined below in Section 2.5.

The properties of measuring instruments can normally be described based on the characteristics enumerated above. For specific types of measuring instruments, however, additional characteristics are often required. Thus, for the gauge rods, the so-called flatness and polishability are important. For voltmeters, the input resistance is important. We shall not study such characteristics, because they refer only to individual types of measuring instruments.

2.3. Standardization of the Metrological Characteristics of Measuring Instruments

Measuring instruments can only be used as intended when their metrological properties are known. In principle, the metrological properties can be established by two methods. One method is to find the actual characteristics of a specific instrument. In the second method, the nominal metrological characteristics and the permissable deviations of the real characteristics from the nominal characteristics are given.

The first method is laborious, and for this reason, it is used primarily for the most accurate and stable measuring instruments. Thus, the second method is the main method. The nominal characteristics and the permissible deviations from them are given in the technical documentation when measuring instruments are designed, which predetermines the properties of measuring instruments and ensures that they are interchangeable.

In the process of using measuring instruments, checks are made to determine whether the real properties of the devices deviate from the established standards. If one real property deviates from its nominal value by an amount greater than demonstrated by the standards, then the measuring instrument is adjusted, remade, or discarded and no longer used.

Thus, the choice of the nominal characteristics of measuring instruments and the designation of permissable deviations of the real characteristics from them—standardization of the metrological characteristics of measuring instruments—are of great importance for measurement practice. We shall examine the practice of standardization of the metrological characteristics of measuring instruments that has evolved.

Both the production of measuring instruments and the standardization of their characteristics initially arose spontaneously in each country. Later, rules that gave order to this standardization were developed in all countries in which instrument building was highly developed. The recommendations developed at this time by international organizations, primarily Publication 51 of the International Electrotechnical Commission (IEC), were of great importance for the preparation or national standards [8]. We should also mention the International Organization for Standardization (ISO) and the International Organization of Legal Metrology (OIML). The terminological documents are also of great value for this work [2], [4], [7].

We shall now return to the gist of the problem. The significance of nominal metrological characteristics, such as the upper limits of measurement ranges, the nominal values of the measures, the scale factors of instruments and so on, is chosen from a standardized series of values of these characteristics. There is nothing special here. Another task is to standardize the accuracy characteristics, errors, and stability.

Despite the efforts of designers, the real characteristics of measuring instruments depend to some extent on the external conditions. For this reason, some narrow ranges of values of all influence quantities are fixed first, and in this manner, the conditions under which measuring instruments are to be calibrated and checked are determined. These conditions are called reference conditions. The error of measuring instruments under reference conditions is called the *intrinsic error.*

In the standard in [7], this question is solved less formally: The conditions under which the characteristics of measuring instruments depend negligibly on the possible variations of influence quantities are called *reference* conditions. In other words, these are conditions under which the metrological characteristics are practically constant.

This definition of reference conditions seems attractive. I stated the identical idea in [44]. However, I also stated there my doubts in the possibility of implementing the idea. We shall return to this question in the next section.

Thus, the reference conditions of measuring instruments are prescribed and the intrinsic errors of measuring instruments are determined.

In addition to the reference conditions, the normal operating conditions of measuring instruments are also established, i.e., the conditions under which the characteristics of measuring instruments remain within certain limits and the measuring instruments can be employed as intended. Understandably, errors in the normal operating conditions are larger than errors in the reference conditions (i.e., these errors are larger than the intrinsic errors).

When any influence quantity exceeds the normal value or range for a reference condition, the error of the measuring instrument changes. This change is

characterized and standardized by indicating the limit of the permissable additional error, by indicating the highest permissable value of the influence factor of the corresponding influence quantity, or by indicating the limit of the permissable error under the normal operating conditions.

The errors of measuring instruments are expressed not only in the form of absolute and relative errors, adopted for estimating measurement errors, but also in the form *of fiducial errors.* The fiducial error is the ratio of the absolute error of the measuring instrument to some standardizing value—fiducial value. The latter value is established by standards on separate types of measuring instruments. For indicating instruments, for example, the fiducial value is established depending on the characteristic features and character of the scale. The fiducial errors make it possible to compare the accuracy of measuring instruments that have different measurement limits. For example, the accuracy of an ammeter with a measurement limit of 1 A can be compared with that of an ammeter with a measurement limit of 100 A.

In addition, cases when the error of an indicating instrument is expressed in fractions of a graduation are also encountered.

For measuring transducers, the errors can be represented by the errors relative to the input or output.

Figure 2.2 shows the nominal and, let us assume, the real transfer functions of some transducer. The nominal dependence, as done in practice whenever possible, is assumed to be linear. We shall investigate the relationship between the errors of the transducer that are scaled to the input and the output.

We denote the input quantity by x and the output quantity by y. They are related by the relation

$$x = Ky,$$

where K is the nominal transduction constant.

At the point with true values of the quantities x_t and y_t, the true value of the transduction constant will be $K_t = x_t/y_t$. Calculations based on the nominal constant K, however, are given an error.

Let $x_a = Ky_t$ and $y_a = x_t/K$ be determined based on y_t and x_t (see Fig. 2.2). Then the absolute transducer error with respect to the input will be

$$\Delta x = x_a - x_t = (K - K_t)y_t.$$

The error with respect to the output is expressed analogously:

$$\Delta y = y_a - y_t = \left(\frac{1}{K} - \frac{1}{K_t}\right)x_t.$$

We note, first, that Δx and Δy always have different signs: If $(K - K_t) > 0$, then $(1/K - 1/K_t) < 0$.

But this is not the only difference. The quantities x and y can also have different dimensions; i.e., they can be physically different quantities, so that the absolute input and output errors are not comparable. For this reason, we shall study the

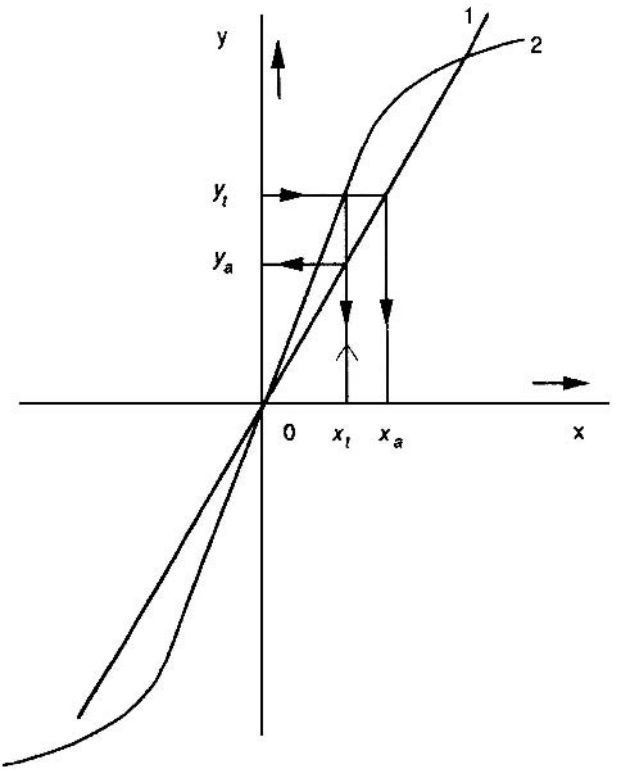


FIGURE 2.2. Nominal (curve 1) and real (curve 2) functions of a measuring transducer.

relative errors:

$$\varepsilon_x = \frac{\Delta x}{x_t} = (K - K_t)\frac{y_t}{x_t} = \frac{K - K_t}{K_t},$$
$$\varepsilon_y = \frac{\Delta y}{y_t} = \frac{(K_t - K)}{K K_t}\frac{x_t}{y_t} = \frac{K_t - K}{K}.$$

As $K_t \neq K$, we have $|\varepsilon_x| \neq |\varepsilon_y|$.

We denote the relative error in the transduction constant at the point (x_t, y_t) as ε_k, where $\varepsilon_k = (K - K_t)/K_t$. Then

$$\frac{\varepsilon_x}{\varepsilon_y} = -(1 + \varepsilon_k).$$

However, $\varepsilon_k \ll 1$, and in practice relative errors with respect to the input and output can be regarded as equal in magnitude.

We must stop to consider how the error of measures is determined: The error of measures is the difference between the nominal value of the measure and the true value of the quantity reproduced by the measure. Indeed, in the case of indicating

instruments, the nominal value of measures is the analog of the indication of the instrument, and the definition given becomes obvious.

It is also interesting that measures that reproduce passive quantities, for example, mass, electric resistance, and so on, have only systematic errors. The error of measures of active quantities (electric voltage, electric current, etc.) can have both systematic and random components. Multiple-valued measures of passive quantities can have random errors from switching elements.

So, when the errors of measuring instruments are standardized, the permissible limits of the intrinsic and all additional errors are prescribed. At the same time, the reference and normal operating conditions are indicated.

Of all forms enumerated above for expressing the errors of measuring instruments, the best is the relative error, because in this case, the indication of the permissible limit of error gives the best idea of the level of measurement accuracy that can be achieved with the given measuring instrument. The relative error, however, usually changes significantly over the measurement range of the instrument, and for this reason, it is difficult to use for standardization.

The absolute error is frequently more convenient than the relative error. In the case of an instrument with a scale, the limit of the permissible absolute error can be standardized with the same numerical value for the entire scale of the instrument. But then it is difficult to compare the accuracies of instruments having different measurement ranges. This difficulty disappears when the fiducial errors are standardized.

In the following discussion, we shall follow primarily [8] and [9].

The limit of the permissible absolute error Δ can be expressed by a single value (neglecting the sign)

$$\Delta = \pm a,$$

in the form of the linear dependence

$$\Delta = \pm(a + bx), \tag{2.1}$$

where x is the nominal value of the measure, the indication of a measuring instrument, or the signal at the input of a measuring transducer, and a and b are constants, or by a different equation,

$$\Delta = f(x).$$

When the latter dependence is complicated, it is given in the form of a table or graph.

The fiducial error γ (in percent) is defined by the formula

$$\gamma = 100\Delta/x_N,$$

where x_N is the fiducial value.

The fiducial value is assumed to be equal to the following:

(i) The value at the end of the instrument scale, if the zero marker falls on the edge or off the scale.

(ii) The span that is a sum of the end values of the instrument scale (neglecting the signs), if the zero marker falls within the scale.
(iii) The nominal value of the measurand, if it has been established.
(iv) The length of the scale, if the scale graduations narrow sharply toward the end of the scale. In this case, the error and the length of the scale are expressed in the same units.

For instruments having a scale that is calibrated in units of a quantity for which a scale with a conventional zero is adopted (for example, in degrees Celsius) the *fiducial value* is assumed to be equal to the difference of the final and starting values of the scale (the measurement range or span).

According to Recommendation 34 of OIML [9], for measuring instruments with a zero marker within the scale, the fiducial value is taken to be equal to the larger (neglecting the sign) of the end values of the indication range of the instrument. According to Publication 51 of IEC [8], for electrical measuring instruments, it can be set equal to the sum of the end values of the scale.

A progressive and correct solution is the one recommended by OIML. Indeed, consider, for example, an ammeter with a scale 100–0–100 A and with a permissible absolute error of 1 A. In this case, the fiducial error of the instrument will be 1% according to OIML and 0.5% according to IEC. But when using this instrument, the possibility of performing a measurement with an error of up to 0.5% cannot be guaranteed for any point of the scale. An error not exceeding 1%, however, can be guaranteed when measuring a current of 100 A under reference conditions. The tendency to choose a fiducial value such that the fiducial error would be close to the relative error of the instrument was observed in the process of improving IEC Publication 51. Thus, in the previous edition of this publication, the fiducial value for instruments without a zeromarker on the scale was taken to be equal to the difference of the end values of the range of the scale, and now it is taken to be equal to the larger of these values (neglecting the sign). Consider, for example, a frequency meter with a scale 45–50–55 Hz and a limit of permissible absolute error of 0.1 Hz. Previously, the fiducial error of the frequency meter was assumed to be equal to 1%, and now it is equal to 0.2%. But when measuring a 50-Hz frequency, its relative error indeed will not exceed 0.2% (under reference conditions), and the 1% error has no relation to the error of frequency measurement with this meter, so that the new edition is more correct.

The next step in this direction was made in Recommendation 34 of OIML. One must hope that in the future IEC will take into account the recommendation of OIML, and the stipulation mentioned regarding electrical measurement instruments in the recommendation of OIML will disappear.

The fiducial error is expressed in percent, but it is not a relative error. As its limit is equal to that of the permissible relative error, the limit of the permissible relative error for each value of the measurand must be calculated according to the formula

$$\delta = \gamma \frac{x_N}{x}.$$

The limit of permissible relative error δ is usually expressed in percent according to the formula

$$\delta = \frac{100\Delta}{x} = \pm c.$$

If the limit of the absolute error Δ is determined by formula (2.1), then the last expression is possible for $a \approx 0$.

For digital instruments, the errors are often standardized in the conventional form $\pm(b + q)$, where b is the relative error in percent and q is some number of figures of the least significant digit of the digital readout device. For example, an instrument with a measurement range of 0–300 mV is assigned the limits of permissible error $\pm(0.5\% + 2)$. The indicator of the instrument has four digits, so that the figure 2 of the least significant digit corresponds to 0.2 mV. Now the limit of the relative error of the instrument when measuring, for example, a voltage of 300 mV can be calculated as follows:

$$\delta = \pm\left(0.5 + \frac{0.2 \times 100}{300}\right) = \pm 0.57\%.$$

Thus, to estimate the limit of permissible error of an instruments, some calculations must be performed. For this reason, although the conventional form gives a clear representation of the components of instrument error, it is inconvenient to use.

A more convenient form is given in Recommendation 34 of OIML: The limit of permissible relative error is expressed by the formula

$$\delta = \pm\left[c + d\left(\frac{x_e}{x} - 1\right)\right], \tag{2.2}$$

where x_e is the end value of the measurement range of the instrument or the input signal of a transducer and c and d are relative quantities.

With the adopted form of formula (2.2), the first term on the right-hand side is the relative error of the instrument at $x = x_e$. The second term in this expression characterizes the increase of the relative error as the indications of the instrument decrease.

Formula (2.2) can be obtained from $\pm(b + q)$ as follows. To the figure q, there corresponds the measurand qD, where D is the value of one figure in the same digit as the figure q, in units of the measurand. In the relative form, it is equal to qD/x. Now the sum of the terms b and qD/x has the following physical meaning: It is the limit of permissible relative error of the instrument.

So

$$\delta = \left(b + \frac{qD}{x}\right).$$

With the help of identity transformation, we obtain

$$\delta = b + \frac{qD}{x} + \frac{qD}{x_e} - \frac{qD}{x_e} = \left(b + \frac{qD}{x_e}\right) + \frac{qD}{x_e}\left(\frac{x_e}{x} - 1\right).$$

Writing

$$c = b + \frac{qD}{x_e}, \qquad d = \frac{qD}{x_e},$$

we obtain formula (2.2).

In application to the example of a digital millivoltmeter studied above, we have

$$\delta = \pm\left[0.57 + 0.07\left(\frac{x_e}{x} - 1\right)\right].$$

It is clear that the last expression is more convenient to use, and in general, it is more informative than the conventional expression.

Note that for standardization, the error limits are established for the total instrument error and not for the separate components. If, however, the instrument has an appreciable random component, then a permissible limit is established separately for it. For example, aside from the limit of the permissible intrinsic error, the limit of the permissible dead band or hysteresis is also established. Sometimes, however, the limits are nonetheless set separately for the systematic and random components. For example, the error of reference standards is customarily given in this manner in Russia.

Additional errors of measuring instruments are standardized by prescribing the limits for each additional error separately. The intervals of variation of the corresponding influence quantities are indicated simultaneously with the limits of the additional errors. The collection of ranges provided for all influence quantities determines the normal operating conditions of the measuring instrument. The limit of permissible additional error is often represented in proportion to the value of the influence quantity or its deviation from the limits of the interval determining the standard values of these quantities. In this case, the corresponding coefficients are standardized. We shall call it the *influence coefficient.*

In the case of measuring instruments the term *variation of indications* is used as well as the term *additional error.* The term variation of indications is used, in particular, for electric measuring instruments [8].

The additional errors arising when the influence quantities are fixed are systematic errors. For different instruments of the same type, however, they can have different values and, what is more, different signs. For this reason, in the overwhelming majority of standards, the limits of additional errors are set both positive and negative with equal numerical values. For example, the change in the indications of an electric measuring instrument caused by a change in the temperature of the surrounding medium should not exceed the limits $\pm 0.5\%$ for each $10\,^{\circ}\mathrm{C}$ change in temperature under normal operating conditions (the numbers here are arbitrary).

If, however, the properties of a measuring device are sufficiently uniform, it is best to standardize the influence function, i.e., to indicate the dependence of the indications of the instruments or output signals of the transducers on the influence quantities and the limits of permissible deviations from each such dependence. If the influence function can be standardized, then it is possible to introduce

corrections to the indications of the instruments and thereby to use the capabilities of the instruments more fully.

It should be emphasized that the properties of only the measuring instruments are standardized with the help of the norms of the additional errors. The actual additional error that can arise in a measurement will depend not only on the properties of the measuring instrument but also on the value of the corresponding influence quantity.

Often a measuring instrument has an electrical signal on its input. This input signal can be characterized by several parameters. One of them reflects the value of a measurand. This parameter is called the *informative parameter:* By measuring its value, we can find the value of the measurand. All other parameters do not have direct connections with the value of the measurand, and they are called *noninformative parameters.*

Measuring instruments are constructed to make them insensitive to all noninformative parameters of the input signal. This result, however, cannot be achieved completely, and in the general case, the effect of the noninformative parameters is only decreased. Furthermore, for all noninformative parameters, it is possible to determine limits such that when the noninformative parameters vary within these limits, the total error of the measuring instrument will change insignificantly, which makes it possible to establish the reference ranges of the values of the noninformative parameters.

If some noninformative parameter falls outside the reference limits, then the error arising is regarded as an *additional error.* The effect of each non-informative parameter is standardized separately, as for influence quantities.

Standardization of the effect of the noninformative parameters and estimation of the errors arising from them are performed based on the same assumptions as those used for taking into account the additional errors caused by the external influence quantities.

The errors introduced by changes in the noninformative parameters of the input signals are occasionally called *dynamic errors.* In the case of several parameters, however, little information is provided. It is more informative to give each error a characteristic name, as is usually done in electric and radio measurements. For example, the change produced in the indications of an ac voltmeter by changes in the frequency of the input signal is called the frequency error. In the case of a voltmeter, for measurements of the peak variable voltages, apart from the frequency errors, the errors caused by changes in the widths of the pulse edges, the decay of the flat part of the pulse, and so on, are taken into account.

The errors caused by deviations of the noninformative parameters of the input signal from the standard values should also be included among the additional errors of a measuring instrument. For example, for a voltmeter in an electromagnetic system, the frequency of the alternating current is one noninformative parameter of the signal. According to Section 1.3, these errors are because one or more parameters of the model do not correspond to the properties of the real object. As these errors are characteristic for the measuring instruments in which they are observed, they are usually given the name of the corresponding parameter of the

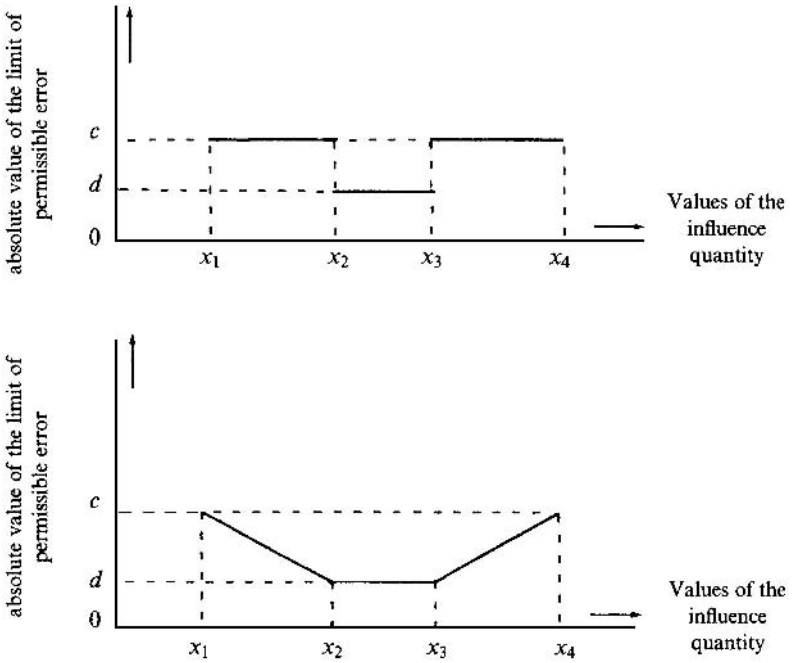


FIGURE 2.3. Variants of standardization of the limits of additional errors of measuring instruments. The interval $(x_3 - x_2)$ corresponds to *reference conditions*; the interval $(x_4 - x_1)$ corresponds to the *normal operating conditions*; d is the limit of permissible intrinsic error; c is the limit of permissible error in the *normal operating conditions*; and $(c - d)$ is the limit of permissible additional error.

model. Thus, in the foregoing example, the model of the signal is a sinusoidal voltage with a fixed parameter (frequency). The corresponding error is called the frequency error.

The basic cases of standardization of additional errors are shown in Fig. 2.3.

Stability of measuring instruments. Stability, like accuracy, is a positive quality of a measuring instrument. Just as the accuracy is characterized by inaccuracy (error, uncertainty), stability is characterized by instability.

Instability is standardized by the limits of permissible variations of the error over a definite period of time or by prescribing different error limits to different "lifetimes" of the instrument after it is calibrated. In addition, limits are sometimes prescribed for the drift of the indications of the instrument; these limits, naturally, are indicated together with the time. It is desirable to standardize drift for the automatic-plotting instruments. But standards for the drift are also helpful for other types of measurement instruments, because it makes it possible to judge how often the indications or the zeros of the instruments must be corrected.

To correct the indications of electric measuring instruments, standard cells or electronic voltage stabilizers are often built into them. For example, weak sources

of stable radioactivity are built into meters for measuring the parameters of radioactive radiations.

It is significant that separate standardization of the drift does not change the standards for the instrumental error; i.e., it gives additional information about the properties of the instruments.

The second method of standardization of instability consists of indicating different standards for the error of the instrument for different periods of time after the instrument is calibrated. For example, a table with the following data is provided in the specifications of some digital instrument:

Time after calibration	24 hour	3 month	1 year	2 years
Temperature	$23 \pm 1\,^{\circ}\mathrm{C}$	$23 \pm 5\,^{\circ}\mathrm{C}$	$23 \pm 5\,^{\circ}\mathrm{C}$	$23 \pm 5\,^{\circ}\mathrm{C}$
Limit of error	$0.01\% + 1$ digit	$0.015\% + 1$ digit	$0.02\% + 1$ digit	$0.03\% + 2$ digit

The limits of error are presented here in the conventional form.

The first method for standardizing instability of instruments is widely used in Russia, and the second method is widely used in the United States. The second method reveals more fully the capabilities of instruments. For example, the limits of error of a digital instrument manufactured in Russia, with the parameters indicated in the table above, would have to be checked once per year $\pm$ ($0.02\% + 1$ digit). The maximum instrument accuracy that can be realized in a short period of time after calibration, although in a more restricted temperature regime, would remain unknown.

Standardization predetermines the properties of measuring instruments and is closely related with the concept of accuracy classes of measuring instruments.

Accuracy classes were initially introduced for indicating electric measuring instruments [8]. Later this concept was also extended to all other types of measuring instruments [9]. Unification of the accuracy requirements of measuring instruments, the methods for determining them, and the notation in general are certainly useful to both the manufacturers of measuring instruments and to users, because it makes it possible to limit, without harming the manufacturers or the users, the list of instruments, and it makes it easier to use and check the instruments. We shall discuss this concept in greater detail.

In [2], the following definition is given for the term accuracy class (the following definition is that close to that given in [8]): The accuracy class is a class of measuring instruments that meet certain metrological requirements that are intended to keep errors within specified limits.

Every accuracy class has conventional notation, established by agreement—the class index—that is presented in [8] and [9].

On the whole, the accuracy class is a generalized characteristic that determines the limits for all errors and standards for all other characteristics of measuring instruments that affect the accuracy of measurements performed with their help.

For measuring instruments whose permissible limits of intrinsic error are expressed in the form of relative or fiducial errors, the following series of numbers, which determine the limits of permissible intrinsic errors and are used for denoting the accuracy classes, was established in [9]:

$$(1,\ 1.5,\ 1.6,\ 2,\ 2.5,\ 3,\ 4,\ 5,\ \text{and}\ 6) \times 10^n,$$

where $n = +1, 0, -1, -2, \ldots$; the numbers 1.6 and 3 can be used, but they are not recommended. For any one value of n, not more than five numbers of this series are allowed. The limit of permissible intrinsic error for each type of measuring instrument is set equal to one number in the indicated series.

Conventional designations of accuracy classes, employed in documentation accompanying measuring instruments, as well as the designations imposed on them, have been developed with the numbers in the indicated series. Of course, this process refers to measuring instruments whose errors are standardized in the form of relative and fiducial errors. Table 2.1 gives examples of the adopted designations of accuracy classes of these measuring instruments.

In those cases when the limits of permissible errors are expressed in the form of absolute errors, the accuracy classes are designated by Latin capital letters or roman numerals.

If formula (2.2) is used to determine the limit of permissible error, then both numbers c and d are introduced into the designation of the accuracy class. These numbers are selected from the series presented above, and in calculating the limits of permissible error for a specific value of x, the result is rounded off so that it would be expressed by not more than two significant figures; the roundoff error should not exceed 5% of the computed value.

The limits of all additional errors and other metrological characteristics of measuring instruments must be related with their accuracy class. In general, it is impossible to establish these relations for all types of measuring instruments simultaneously—measuring instruments are too diverse. For this reason, these relations must be given in the specifications together with the characteristics of specific types of measuring instruments, which the designers formulate.

TABLE 2.1. Designations of accuracy classes.

Form of the expression for the error	Limit of permissible error (examples)	Designation of the accuracy class (for the given example)
Fiducial error, if the fiducial value is expressed in units of the measurand	$\gamma = \pm 1.5\%$	1.5
Fiducial error, if the fiducial value corresponds to the span	$\gamma = \pm 0.5\%$	|0.5|
Relative error, constant	$\delta = \pm 0.5\%$	(0.5)
Relative error, increasing as the measurand decreases	$\delta = \pm\left[0.02 + 0.01\left(\frac{x_e}{x} - 1\right)\right]\%$	0.02/0.01

2.4. Some Suggestions for Changing Methods of Standardization of Errors of Measuring Instruments and Their Analysis

Standardization, i.e., establishment of standards, is basically a volitional act. For this reason, in principle, different suggestions can be made for solving this question, and in the last few years, several new methods for expressing the errors of measuring instruments and for standardizing them have indeed been proposed.

To evaluate these suggestions, it is necessary to determine how well they solve problems for whose sake the properties of measuring instruments are standardized. From what we have said above, it can be concluded that the purpose of standardization of errors of measuring instruments is to solve the following problems:

(1) To ensure that the entire collection of measuring instruments of the same type have the required accuracy and to ensure that they are uniform and interchangeable.
(2) To make sure that it is possible to evaluate the instrumental measurement errors according to established standards for metrological properties of measuring instruments.
(3) To ensure that measuring instruments can be compared with one another according to accuracy.

The first problem is ultimately solved by monitoring new measuring instruments during the manufacturing process and checking periodically the units that are in use. As measuring instruments are employed individually, the standards must be established so that it is possible to check that each sample measuring instrument satisfies these standards.

To solve the second problem successfully, it is desirable to know accurately the properties of measuring instruments. For this reason, the established standards must be as close as possible to the real properties of the measuring instruments. The degree of detail with which the errors of measuring instruments can be described is limited by the instability of the instruments, by the change in their errors in time, as well as by the degree of nonuniformity of the measuring instruments introduced by their construction and manufacturing technology. In addition, the calibration process must be simple. Complicated methods for describing and standardizing the errors of measuring instruments, which lead to laborious and prolonged checks, are nonviable.

Having made these preliminary remarks, we shall now examine the most interesting suggestions.

(1) The calculation of the errors of measuring instruments under real conditions involves summation of the errors and presents several difficulties. For this reason, it has been repeatedly suggested that the reference conditions be extended to absorb all possible values of the influence quantities. One would think that in so doing

the additional errors of measuring instruments would vanish and only the intrinsic error would remain, and all difficulties would be simply resolved.

The actual properties of measuring instruments, however, do not depend on the method by which they are standardized, and they remain unchanged. Suppose that in the usual method of standardization, we have the following:

Δ_0, the limit of permissible intrinsic error; and

Δ_i, the limit of permissible additional error, caused by the change in the ith influence quantity from the standard value to the limit of the range of the given influence quantity ($i = 1, \ldots, n$) for normal operating conditions.

By transferring to a new method of standardization, the manufacturer of the instruments can adopt as the limit of permissible error of the measuring instrument only the arithmetic sum

$$\Delta = \Delta_0 + \sum_{i=1}^{n} \Delta_i .$$

The manufacturer cannot proceed otherwise, because he or she must guarantee that the errors of a given measuring instrument will be less than Δ for any combination of limiting values of the influence quantities. What then can this suggestion give?

From the standpoint of evaluating the measurement errors, it can significantly simplify the procedure. But in exchange, the error is significantly overestimated because under the actual operating conditions of the measuring instrument in the overwhelming majority of the cases, the influence quantities do not all reach their limiting values simultaneously and in the most unfavorable combination. For this reason, even the arithmetic sum of the errors occurring in a specific measurement will be less than Δ and closer to the real value.

With respect to uniformity and interchangeability of measuring instruments of the same type, the suggestion worsens the existing situation, because the same value of Δ can be obtained for different values of the components. Thus, to adopt this suggestion means taking a step backward compared with the present situation.

The foregoing analysis also shows that the definition of reference conditions given in [7] gives the most complete disclosure of the properties of a measuring instrument. In the overwhelming majority of the cases, however, before a measuring instrument can be developed, it is necessary to establish the technical requirements that it must meet. In the process, the reference conditions and the permissible limits of the intrinsic error are determined. During the design process, the investigators and designers strive to satisfy these requirements within some margin. Normally this result is possible, which essentially means that the reference conditions established earlier can be defined more stringently. But if this path is followed, then the reference conditions would have to be redetermined after the sample measuring instruments have been built, and it would be found that they are diverse for different types of measuring instruments, which would

create great difficulties for technical monitoring services and calibrating laboratories. For this reason, on the whole, the reference conditions are best determined by agreement between specialists, and these conditions should be unified as much as possible for different types of measuring instruments. When developing measuring instruments, however, it should be kept in mind that if the intrinsic error is appreciably correlated with one or another influence quantity, then the real properties of the measuring instruments are not completely disclosed by the prescribed standards.

(2) It has been suggested that the integral accuracy index I, calculated according to the formula

$$I = \sqrt{\sum_{i=0}^{n} \varepsilon_i^2},$$

where ε_i is the limit of additional error determined by the ith influence quantity and ε_0 is the limit of intrinsic error, be standardized.

It is clear that one and the same value of I can be obtained for different values of the components. For example, one instrument can have a large temperature error and a small frequency error, whereas the opposite could be true for a different instrument. Ultimately, replacing one instrument by another (of the same type) results in a large error, and this error cannot be estimated beforehand. Therefore, it becomes more difficult to estimate the measurement errors. In addition, uniformity of measuring instruments is not achieved. The conclusion is obvious: The suggestion is not acceptable.

(3) Another suggestion was to characterize the accuracy of instruments by the weighted mean of the permissible relative error, determined according to the formula

$$\delta_c = \int_{x_i}^{x_f} [\varepsilon(x) f(x)] dx,$$

where x_i and x_f are the initial and final (upper) values of the instrument scale, $\varepsilon(x)$ is the relative error of the instrument, and $f(x)$ is the probability distribution of the indications of the instrument.

This suggestion has the drawback that the probability distribution of the indications of instruments is, in general, unknown. More importantly, however, this weighted-mean characteristic, as any other average characteristic, is completely unsuitable for standardizing the properties of measuring instruments, because uniformity of measuring instruments cannot be achieved in this manner. For example, an instrument that has one or two significant error components, and for which other errors are small, can have the same weighted-mean error as an instrument whose errors are approximately the same.

In addition, when using an instrument whose errors are standardized as weighted means, experimenters cannot estimate the error of a specific result they have obtained, because in this method of standardization, the error of the instrument with a fixed indication can in principle be virtually arbitrarily large.

Thus, none of the goals of standardization is achieved with this method of standardization of errors of measuring instruments and this method cannot be used.

(4) It has been repeatedly suggested that the additional error caused by the simultaneous action of all influence quantities be standardized. It can be conjectured that some of them will mutually compensate one another so that it will be possible to use the instruments more fully or, vice versa, the error will be larger than in the case when each influence quantity acts separately.

In practice, however, normally not all influence quantities assume their worst (for us) values simultaneously, and it is impossible to take into account only some influence quantities by standardizing in this manner. Instead of lowering the estimate of the measurement error or increasing its accuracy, the measurement error will increase and it will not be estimated as accurately. In addition, the testing equipment would become much more complicated.

In reality, some additional errors can be correlated with one another, and it would be correct to determine these cases and to standardize the correlation of the corresponding errors. However, I have never encountered in practice such cases of standardization of additional errors. There does not appear to be any need to do so. However, this question deserves a detailed study.

(5) In the former USSR, in 1972, a standard that decisively changed the practice of standardization of errors of measuring instruments was adopted. This standard greatly complicates the standardization of errors of measuring instruments and contains wholly unrealistic requirements. They include, for example, the requirement that the mathematical expectation and the rms deviation of the systematic component of the errors of measuring instruments of each type be standardized. These characteristics must be estimated according to the formulas

$$\bar{x} = \frac{\sum_{i=1}^{m} x_i}{m}, \qquad S = \sqrt{\frac{\sum_{i=1}^{m} (x_i - \bar{x})^2}{m-1}},$$

where m is the number of instruments in a batch and x_i is the systematic error of the ith instrument.

Estimates of these parameters can be calculated by checking instruments in a batch. Let us assume that they satisfy the standards. Does this mean that the systematic error is sufficiently small for all instruments in the batch? Obviously not. In exactly the same way, nothing can be said about a separate instrument, if these estimates do not satisfy the standard. According to this standard, an instrument cannot be rejected, and it cannot be judged satisfactory in the case of verification. Of course, once estimates have been found for a batch of instruments, then the entire batch of instruments can either be discarded or accepted. But this action is absurd: A bad batch can contain several good instruments, and they should not be discarded. Conversely, it is absurd to pass as satisfactory some unsatisfactory instruments simply because they are contained in the batch that has been found to satisfy the standard. Every instrument is used individually, and the standard must

make it possible to determine whether the instrument is good or bad. Parameters pertaining to batches do not meet this requirement, and for this reason, they cannot be used as standards for measuring instruments.

I opposed the adoption of this standard, but the standard was adopted (GOST 8.009-72). In 1984, a new edition of this standard was published. Now the characteristics that we studied above are no longer obligatory, and this eases somewhat the situation of instrument manufacturers: They do not have to standardize these characteristics, and this will not be a violation of the standard.

In conclusion, we shall formulate the basic rules for standardization of errors of measuring instruments:

(i) all properties of a measuring instrument that affect the accuracy of the results of measurements must be standardized;
(ii) every property that is to be standardized should be standardized separately;
(iii) methods of standardization must make it possible to check experimentally, and as simply as possible, how well each sample of a measuring instrument corresponds to the established standards; and
(iv) the standardization must be performed so that measuring instruments can be chosen based on the established standards and so that the measurement error can be estimated.

In some cases, exceptions must be made to these rules. Such an exception is necessary for strip strain gauges that can be glued on an object only once. For this reason, the strain gauges that are checked can no longer be used for measurements, whereas the gauges that are used for measurements usually cannot be checked or calibrated. In this case, it is necessary to resort to regulation of the properties of a collection of strain gauges, such as, for example, the standard deviation of the sensitivity and mathematical expectation of the sensitivity. The sensitivity of a separate strain gauge, which is essentially not a random quantity, is a random quantity in application to a collection of strain gauges. Once the sensitivity x_i of every strain gauge chosen at random from a batch (sample) has been determined, it is possible to construct a statistical tolerance interval, i.e., the interval into which the sensitivity of a prescribed fraction p of the entire collection of strain gauges will fall with a chosen probability a. As $a \neq 1$ and $p \neq 1$, there is a probability that the sensitivity of any given strain gauge falls outside these tolerance limits. For this reason, the user must take special measures that exclude such a case. In particular, several strain gauges, rather than one, should be used.

2.5. Dynamic Characteristics of Measuring Instruments and Their Standardization

The dynamic characteristics of measuring instruments reflect the relation between the change in the output signal and one or another action that produces this change. The most important action is a change in the input signal. In this case, the dynamic characteristic is called the dynamic characteristic for the input signal. Dynamic

characteristics for one or another influence quantity and for a load (for measuring instruments whose output signal is an electric current or voltage) are also studied.

Complete and partial dynamic characteristics are distinguished [28].

The complete dynamic characteristics determine uniquely the change in time of the output signal caused by a change in the input signal or other action. Examples of such characteristics are a differential equation, transfer function, amplitude–and phase–frequency response, the transient response, and the impulse characteristic. These characteristics are essentially equivalent, but the differential equation is still the source characteristic.

A partial dynamic characteristic is a parameter of the full dynamic characteristic or a functional of it. Examples are the response time of the indications of an instrument and the transmission band of a measuring amplifier.

Measuring instruments can most often be regarded as inertial systems of first or second order. If $x(t)$ is the signal at the input of a measuring instrument and $y(t)$ is the corresponding signal at the output, then the relation between them can be expressed with the help of first-order (2.3) or second-order (2.4) differential equations, respectively, which reflect the dynamic properties of the measuring instrument:

$$T y'(t) + y(t) = K x(t), \tag{2.3}$$

$$\frac{1}{\omega_0^2} y''(t) + \frac{2\beta}{\omega_0} y'(t) + y(t) = K x(t). \tag{2.4}$$

The parameters of these equations have specific names: T is the time constant of a first-order device, K is the transduction coefficient in the static state, ω_0 is the angular frequency of free oscillations, and β is the damping ratio.

Equations (2.3) and (2.4) reflect the properties of real devices, and for this reason, they have zero initial conditions: for $t \le 0$, $x(t) = 0$ and $y(t) = 0$, $y'(t) = 0$, and $y''(t) = 0$.

For definiteness, in what follows, we shall study the second-order equation and we shall assume that it describes a moving-coil galvanometer. Then $\omega_0 = 2\pi f_0$, where f_0 is the frequency of free oscillations of the moving part of the galvanometer.

To obtain transfer functions from differential equations, it is first necessary to transfer from signals in the time domain to their Laplace transforms, and then to form their ratio. Thus

$$\mathscr{L}[x(t)] = x(s), \qquad \mathscr{L}[y(t)] = y(s),$$
$$\mathscr{L}[y'(t)] = s y(s), \qquad \mathscr{L}[y''(t)] = s^2 y(s),$$

where s is the Laplace operator.

For the first-order system, we obtain

$$W(s) = \frac{y(s)}{x(s)} = \frac{K}{1 + sT},$$

and for the second-order system, we obtain

$$W(s) = \frac{y(s)}{x(s)} = \frac{K}{\left(1/\omega_0^2\right)s^2 + (2\beta/\omega_0)s + 1}. \tag{2.5}$$

If in the transfer function the operator s is replaced by the complex frequency $j\omega$ ($s = j\omega$), then we obtain the complex frequency response. We shall study the relation between the named characteristics for the example of a second-order system. From (2.4) and (2.5), we obtain

$$W(j\omega) = \frac{K}{\left(1 - \omega^2/\omega_0^2\right) + j2\beta\omega/\omega_0}, \tag{2.6}$$

where $\omega = 2\pi f$ is the running angular frequency.

The complex frequency response is often represented for its real and imaginary parts,

$$W(j\omega) = P(\omega) + jQ(w).$$

In our case,

$$P(\omega) = \frac{K\left(1 - \left(\omega^2/\omega_0^2\right)\right)}{\left(1 - \left(\omega^2/\omega_0^2\right)\right)^2 + 4\beta^2\left(\omega^2/\omega_0^2\right)},$$
$$Q\left(\omega\right) = \frac{2\beta(\omega/\omega_0)K}{\left(1 - \left(\omega^2/\omega_0^2\right)\right)^2 + 4\beta^2\left(\omega^2/\omega_0^2\right)}.$$

The complex frequency response can also be represented in the form

$$W(j\omega) = A(\omega)e^{j\varphi(w)},$$

where $A(\omega)$ is the amplitude-frequency response and $\varphi(\omega)$ is the frequency response of phase. In the case at hand

$$A(\omega) = \sqrt{P^2(\omega) + Q^2(\omega)} = \frac{K}{\sqrt{\left(1 - \left(\omega^2/\omega_0^2\right)\right)^2 + 4\beta^2\left(\omega^2/\omega_0^2\right)}},$$
$$\varphi(\omega) = \arctan\frac{Q(\omega)}{P(\omega)} = -\arctan\frac{2\beta(\omega/\omega_0)}{1 - \left(\omega^2/\omega_0^2\right)}. \tag{2.7}$$

Equations (2.7) have a well-known graphical interpretation.

The transient response is the function $h(t)$ representing the output signal produced by a unit step function $1(t)$ at the input. We recall that the unit step function is a function $x(t)$ satisfying the following conditions: $x(t) = 0$ for $t < 0$ and $x(t) = 1$ for $t \geq 0$. As the input is not periodic, $h(t)$ is calculated with (2.3) or (2.4). Omitting the simple but, unfortunately, complicated calculations, we arrive at the final

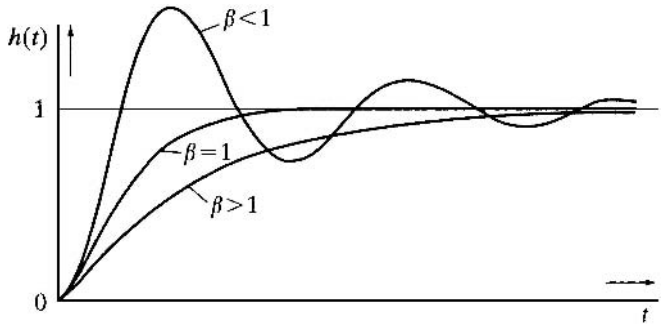


FIGURE 2.4. The transient response of an instrument described by a second-order differential equation; β is the damping ratio.

form of the transient response of the instrument under study:

$$h(t) = \begin{cases} 1 - e^{-\beta\tau} \dfrac{1}{\sqrt{1-\beta^2}} \sin\left(\tau\sqrt{1-\beta^2} + \arctan\dfrac{\sqrt{1-\beta^2}}{\beta}\right) & \text{if } \beta < 1, \\ 1 - e^{-\tau}(\tau + 1) & \text{if } \beta = 1, \\ 1 - e^{-\beta\tau} \dfrac{1}{\sqrt{\beta^2-1}} \sinh\left(\tau\sqrt{\beta^2-1} + \operatorname{arctanh}\dfrac{\sqrt{\beta^2-1}}{\beta}\right) & \text{if } \beta > 1. \end{cases}$$

Here $\tau = \omega_0 t$ and the steady-state value of the output signal is taken to be equal to unity, i.e., $h(t) = y(t)/K$. Thanks to this condition, the formulas above and the graphs corresponding to them, presented in Fig. 2.4, are universal in the sense that they do not depend on the specific values of ω_0 and K.

The impulse characteristic $g(t)$ is found from the transient response in accordance with its definition:

$$g(t) = \frac{dh(t)}{dt}.$$

It should be noted that some types of measuring instruments do not have dynamic characteristics at all: measures of length, weights, vernier calipers, and so on. Some measuring instruments, such as measuring capacitors (measures of capacitance), do not have an independent dynamic characteristic. But when they are connected into an electric circuit, which always has some resistance and sometimes an inductance, the circuit always acquires, together with a capacitance, definite dynamic properties.

Measuring instruments are diverse. Occasionally, to describe adequately their dynamic properties, it is necessary to resort to linear equations of a higher order, nonlinear equations, or equations with distributed parameters. However, complicated equations are used rarely, which is not an accident. After all, measuring instruments are created specially to perform measurements, and their dynamic properties are made to guarantee convenience of use. For example, in designing an automatic plotting instrument, the transient response is made to be short,

approaching the established level monotonically or oscillating insignificantly. In addition, the instrument scale is made to be linear. But when these requirements are met, the dynamic properties of the instrument can be described by one characteristic corresponding to a linear differential equation of order no higher than second.

A differential equation of high order is most often obtained when synthesizing the dynamic characteristics of an instrument based on the dynamic characteristics of its subunits. Thus, for example, calculating the dynamic characteristic of a galvanometric amplifier with a photoelectric converter that converts the angle of rotation of the moving part of the galvanometer into a voltage (current), we formally obtain an equation of third order: The galvanometer gives two orders, and the photoelectric converter gives one order. Such a description of the properties of the amplifier is necessary at the design stage, because otherwise it is impossible to understand why self-excited oscillations sometimes arise in the system. But when the design is completed and reasonable parameters of the subunits are chosen, it is desirable to simplify the description of the dynamic properties. Thus, the amplifier must be regarded as a black box. Analyzing the relation between the input and output, we find that it is described well by a second-order equation. The same result can also be obtained informally, as done, for example, in [42]. Decomposing the dynamic characteristics of all subunits of the amplifier into first-order characteristics and comparing them, we can see that one can be neglected.

Standardization of the dynamic characteristics of measuring instruments is performed for a specific type of instrument. The problem is solved in two stages. First, an appropriate dynamic characteristic must be chosen, after which the nominal dynamic characteristic and the permissible deviations from it must be established. Thus, for recording instruments and universal measuring transducers, one complete dynamic characteristic must be standardized: Without having the complete dynamic characteristic, a user cannot effectively use these instruments.

For indicating instruments, it is sufficient to standardize the response time. In contrast to the complete characteristics, this characteristic is a partial dynamic characteristic. The dynamic error is another form of a partial dynamic characteristic. Standardization of the limits of a permissible dynamic error is convenient for the measuring instruments employed, but it is justified only when the form of the input signals does not change much.

For measuring instruments described by linear first- and second-order differential equations, the coefficients of all terms in the equations can be standardized. In the simplest cases, the time constant is standardized in the case of a first-order differential equation, and the natural frequency and the damping ratio of the oscillations are standardized in the case of a second-order differential equation.

When imposing requirements on the properties of measuring instruments, it is always necessary to keep in mind how compliance will be checked. For dynamic characteristics, the basic difficulties are connected with creating test signals of predetermined (with sufficient accuracy) form, or with recording the input signal with a dynamically more accurate measuring instrument than the measuring instrument whose dynamic properties are being checked.

If test signals with adequate accuracy can be created and the dynamic characteristic is found with the help of the corresponding signal, i.e., a transient response as a response of a unit step function signal and frequency response as a response of a sinusoidal test signal, then in principle the obtained experimental data can be processed without any difficulties.

But sometimes the problem must be solved with a test signal that does not correspond to the signal intended for determining the complete dynamic characteristic. For example, one would think that the problem can be solved given the tracing of signals at the input and output of a measuring instrument. In this case, however, special difficulties arise because small errors in recording the test signal and reading the values of the input and output signals often lead to the fact that the dynamic characteristic obtained based on them do not correspond to the dynamic properties of the measuring instrument and are physically meaningless. Such an unexpected effect is explained because the problem at hand is a so-called improperly posed problem. A great deal of attention is currently being devoted to such problems in mathematics, automatics, geophysics, and other disciplines. Improperly posed problems are solved by the methods of regularization, which essentially consist of the fact that the necessary degree of filtering (smoothing) of the obtained solution is determined based on a priori information about the true solution.

Improperly posed problems in dynamics in application to measurement engineering are reviewed in [28].

A separate problem, which is important for some fields of measurement, is the determination of the dynamic properties of measuring instruments directly when the instruments are being used. An especially important question here is the question of the effect of random noise on the accuracy with which the dynamic characteristics are determined.

This section, then, has been a brief review of the basic aspects of the problem of standardizing and determining the dynamic properties of measuring instruments.

2.6. Statistical Analysis of the Errors of Measuring Instruments Based on Data Provided by Calibration Laboratories

A general characteristic of the errors of the entire population of measuring instruments of a specific type could be their distribution function. I made an attempt to find such functions for several types of measuring instruments. The results of these investigations, which were performed together with T.L. Yakovleva, were published in [43] and [53].

The errors of measuring instruments are determined by calibration, and a decision was made to use the data provided by calibration laboratories. Because it is impossible to obtain the errors of all instruments of a given type that are in use, the use of a sampling method is unavoidable.

TABLE 2.2. The example of main statistical characteristics of errors for six types of measuring instruments.

Type of measuring instrument	Year of calibration	Point of check	Sample size	Moment		Coefficient	
				First	Second central	Skewness	Excess
Э59 ammeter	1974	80 Divisions	160	0.163	0.0074	−0.40	0.56
	1976		160	0.180	0.042	−1.33	4.27
Э59 voltmeter	1974	150 Divisions	120	0.050	0.063	−0.47	−0.29
	1976		108	0.055	0.065	−0.18	0.15
Д 566 wattmeter	1974	150 Divisions	86	0.088	0.024	−0.50	−0.54
	1976		83	0.062	0.021	0.05	0.81
TH-7 thermometer	1975	100 °C	92	−0.658	0.198	0.14	−0.14
	1976		140	−0.454	0.128	0.45	1.57
Standard spring manometer	1973	9.81 kPa	250	0.158	0.012	0.55	0.54
	1976		250	0.128	0.012	0.59	−0.13
P331 resistance measure	1970	100 Ω	400	0.33×10^{-3}	1.6×10^{-2}	0.82	1.08
	1975		400	0.1×10^{-3}	1.2×10^{-2}	0.44	2.02

To establish a property of an entire group (general collection) based on a sample, the samples must be representative. Sample homogeneity is a necessary indicator of representativeness. In the case of two samples, to be sure that the samples are homogeneous, it is necessary to check the hypothesis $H_0 : F_1 = F_2$, where F_1 and F_2 are distribution functions corresponding, respectively, to the first and second samples.

The results of the check, as is well known, depend not only on the error of the measuring instrument being calibrated but also on the error of the standard. For this reason, measuring instruments that are checked with not less than a fivefold margin of accuracy were selected for analysis.

In addition, to ensure that the samples are independent, they were formed either based on data provided by calibration laboratories in different regions of the former USSR or, if data from a single laboratory were used, the data were separated by a significant time interval. The sample size was maintained approximately constant.

We shall discuss [43] first. Table 2.2 gives the basic statistical characteristics of the samples for six types of different instruments. Two samples, obtained at different times, are presented for each of them. For brevity, the data referring to only one numerical scale marker are presented. The arithmetic mean of the values obtained by continuously approaching the marker checked from both sides was taken as the value of the error. The first initial and second central moments are given in the same units in which the value of the point of checking is presented, i.e., in fractions of a scale graduation, in degrees Celsius, and so on. (in the corresponding power). Errors exceeding twice the limit of permissible error were eliminated from the analysis.

The test was made with the help of the Wilcoxon–Mann–Whitney and Siegel–Tukey criteria with a significance level $q = 0.05$. The technique of applying these criteria is described in Chapter 4.

TABLE 2.3. The results of testing the hypothesis of homogeneity for samples of six types of measuring instruments.

Type of measuring instrument	Year of calibration	Point of check	Result of testing the hypothesis based on the criterion of	
			Wilcoxon–Mann–Whitney	Siegel–Tukey
Э59 ammeter	1974	30 Divisions	+	−
	1976	60	0	−
		80	0	−
		100	+	+
Э59 voltmeter	1974	70 Divisions	−	0
	1976	150	+	+
Д 566 wattmeter	1974	70 Divisions	+	+
	1976	150	+	+
TH-7 thermometer	1975	100 °C	0	−
	1976	150 °C	−	+
		200 °C	+	+
Standard spring manometer	1973 1976	9.81 kPa	+	+
P331 resistance measure	1970	10 kΩ	0	−
	1975	100 Ω	0	−
		10 Ω	0	−

The results of the analysis are presented in Table 2.3. Rejection of the hypothesis is indicated by a minus sign, and acceptance is indicated by a plus sign. The symbol 0 means that a test based on the given criterion was not made.

The Wilcoxon–Mann–Whitney and Siegel–Tukey criteria are substantially different: The former is based on comparing averages, and the latter is based on comparing variances. For this reason, it is not surprising that cases when the hypothesis H_0 is rejected according to one criterion but accepted according to the other are encountered. The hypothesis of sample homogeneity must be rejected if even one of the criterion rejects it. Both samples were found to be homogeneous only for the Д 566 wattmeters and standard manometers. For other measuring instruments, the compared samples were often found to be nonhomogeneous. It is interesting that on one scale marker, they can be homogeneous, on another, they are inhomogeneous (Э59 voltmeters and ammeters). TH-7 thermometers had homogeneous samples in one range of measurement and nonhomogeneous in a different range. The calculations were repeated for significance levels of 0.01 and 0.1, but on the whole, the results were the same in both cases.

The experiment described was formulated to check the stability of the distribution functions of the errors, but because in the samples compared, the instruments were not always the same, the result obtained has a different but no less important meaning: It indicates that they are nonhomogeneous. It means that the parameters of one sample are statistically not the same as these parameters of another sample of the same type of measuring instruments. Thus, the results obtained show that samples of measuring instruments are frequently nonhomogeneous with respect to

errors. For this reason, they cannot be used to determine the distribution function of the errors of the corresponding instruments.

This result is indicated also by the results of [53], in which samples obtained based on data provided for Э59 ammeters by four calibration laboratories in different regions of the former USSR were compared. The number of samples was equal to 150–160 everywhere. The errors were recorded at the numerical markers 30, 60, 80, and 100 graduations. The samples were assigned the numbers 1, 2, 3, and 4, and the hypotheses $H_0 : F_1 = F_2$, $F_2 = F_3$, $F_3 = F_4$, and $F_4 = F_2$ were checked. The combinations of samples were arbitrary. The hypothesis testing was based on the Wilcoxon–Mann–Whitney criterion with $q = 0.05$. The analysis showed that we can accept the hyphothesis $H_0 : F_1 = F_2$ only, and only at the marker 100. In all other cases, the hypothesis had to be rejected.

Thus, the sample method does not permit finding the distribution function of the errors of measuring instruments. There are evidently two reasons for this result. The first reason is that the stock of instruments of each type is not constant. On the one hand, new instruments that have just been manufactured are added to it. On the other hand, in the verification, some instruments are rejected, some instruments are replaced, and others are discarded. The ratio of the numbers of old and new instruments is constantly changing. The second reason is that the errors of the instruments unavoidably change with time. Moreover, many instruments are used under different conditions, and the conditions of use affect differently the rate at which the instrumental errors change.

The temporal instability of measuring instruments raises the question of whether the errors of measuring instruments are in general sufficiently stable so that a collection of measuring instruments can be described by some distribution function. At a fixed moment in time, each type of instruments without doubt can be described by distribution function of errors. The other problem is how to find this distribution function. The simple sampling method, as we saw above, is not suitable. But even if the distribution function can be found by some complicated method, after some time, it would have to be redetermined, because the errors, and the composition of the stock of measuring instruments, change. Therefore it must be concluded that the distribution of errors of measuring instruments cannot be found based on the experimental data.

The results presented above were obtained in the former USSR, and instruments manufactured in the former USSR were studied. However, there are no grounds for expecting that instruments manufactured in other countries will have different statistical properties.

3
Prerequisites for the Analysis of the Inaccuracy of Measurements and for Synthesis of Their Components

3.1. Relationship Between Error and Uncertainty

As mentioned, a measurement error cannot be found directly from its definition, i.e., using the definition as an algorithm, because the true value of the measured quantity is unknown. The problem must be solved by performing calculations based on estimates of all components of the measurement inaccuracy. This condition is why the problem of analysis—the identification of the sources and the reasons for the appearance of the measurement errors and estimation of these errors—is so important.

We shall call the smallest of the measurement errors, based on whose estimates the total measurement error or uncertainty of measurement is calculated, the elementary errors.

If in the analysis it is possible to find for some elementary errors concrete specific values, i.e., in the language of mathematical statistics, to find point estimates of these errors, then these components are immediately eliminated by introducing the corresponding corrections. Of course, this is possible only in the case of elementary systematic errors. However, no corrections can make the measurement result absolutely accurate; an uncertainty always remains. In particular, the corrections cannot be absolutely accurate, and after they are introduced, residuals of the corresponding errors remain that have not been eliminated and that later play the role of elementary errors.

Let us turn back once more to the terms *error* and *uncertainty*. For the last two decades or perhaps a bit longer, both were used in the United States with the same meaning.[1] Thus, they were synonyms. But synonyms are not allowed in a proper system of terms, and this situation had to be improved.

One would think article [18] provides the solution to this problem. The main idea of that paper is that the term "measurement error" appears to be used in two different senses. In one sense, in the opinion of the authors of [18], it expresses that the measurement result is different from the true value of the measured quantity,

[1] John R. Taylor. An Introduction to Error Analysis. The Study of Uncertainty in Physical Measurements. Oxford University Press, 1982. Second ed. 1997.

whereas in the other sense, it reflects the uncertainty of the measurement result. For example, in the first case, one would use the expression "the error $+1\%$," whereas in the second case, one would say "the error $\pm 1\%$." To distinguish the meaning of the word "error" in these cases, it is proposed that in the second case, the word "uncertainty" be used instead of the word "error."

To understand the essential significance of this proposition, it is first necessary to check the correctness of the examples presented. The first example is obvious. The second example requires some analysis. To be precise, the expression "the error $\pm 1\%$" means that the measurement error is simultaneously both $+1\%$ and -1%. But this result cannot be, because there can only be one result of a measurement, a fixed numerical value; i.e., this expression is incorrect. In this case, one should say "the error falls within the range $\pm 1\%$" or "the limits of error are $\pm 1\%$." If the correct expression were used, then the contradiction mentioned in [18] would not occur.

Thus, the problem lies not in that the term *measurement error* has two meanings but that this term is not used correctly. Nevertheless, it may seen that this proposal eliminates the synonymy. But as a matter of fact, it replaces the term *error* with the term *uncertainty* because the first case mentioned above is rare; it is used almost only in calibration practice.

A much better solution follows from [3] and [14]. An inaccuracy of measurement results is expressed there with the term *uncertainty*, but every numerical value of uncertainty is accompanied with a corresponding confidence probability. The latter is important, and it is a good reason to have a special term in this case. Of course, a new term could be constructed by adding special adjectives to the root word *error*. For example, there is a term *confidence limits of error* or just confidence error in [4]. But shorter terms are preferable. Therefore, we shall use the term *uncertainty* in the present book with this meaning. Also, we shall use the term *error* for all components of uncertainty, and the term *limits of error* of measurement results for those cases where the cause of inaccuracy of measurement is the intrinsic error of the measuring instrument involved. In other words, the term *limits of error* will be used when a corresponding level of confidence cannot be stated, Thus, this solution keeps both terms giving them different areas of application. Therefore this solution enriches the terminology while the first one demages it.

I would like to note that the second edition of Vocabulary [2] foresees now the term *uncertainty* exactly with the same meaning as was described above.

Thus the imperfection of measurement results can be quantitatively described using two terms: *limits of error* and *uncertainty*. Yet another term is needed to refer to the imperfection of measurements. In this book, the term *inaccuracy* is used for this purpose.

3.2. Classification of Elementary Errors

The classification of measurement errors presented in Chapter 1 also applies, of course, to elementary errors. Continuing the analysis, this classification must be further developed.

Most elementary errors are estimated by analysis, and definite limits are found for them. We shall divide elementary errors that have definite limits into absolutely constant and conditionally constant errors.

By *absolutely constant elementary errors*, we mean errors that, although they have definite limits, remain the same in repeated measurements performed under the same conditions as well as for all measuring instruments of the same type. An example of such an error is the error caused by the inaccuracy in the formula used to determine the quantity being measured, if the limits of this error have been established. Another example is the error of digital thermometers, for which the temperature dependence of the emf of the thermocouple is linearized with the help of a polynomial of a fixed degree. Thus, absolutely constant elementary errors are, based on their properties, purely systematic errors.

By *conditionally constant errors*, we mean errors that have certain limits but can vary within these limits both caused by the nonrepeatability and because of the nonreproducibility of the results. A typical example of such an error is the measurement error because of the intrinsic error of the measuring instrument.

The intrinsic error, by its nature, can be a purely systematic error, but it can also have a random component. For example, for weights, the intrinsic error does not have a random component, but its actual magnitude varies from one weight to another. The intrinsic error of an electric measuring instrument with an indicator needle has both systematic and random components, but on the whole, the intrinsic error has definite limits that are the same for any instrument of a given type.

A conditionally constant error can even be purely random. Examples are the roundoff error in reading the indications of analog instruments and the error caused by the limited resolution of digital instruments.

Thus, a fundamental property of conditionally constant elementary errors is that although they have certain limits, they can vary within these limits.

An elementary error that does not have estimable limits is the common random error.

A random error, as is well known, is estimated after a measurement is performed. The estimate is based on data obtained in the course of the measurements. If the random error is significant, then the measurement is performed many times. The primary characteristic of a random error is usually the standard deviation, which is calculated from the experimental data, and the entire standard deviation, and not its separate components, is estimated directly. For this reason, there is no need to add to the term *random component of the measurement error* or briefly *random measurement error* the additional word *elementary.*

Let us note that the random error of a multiple measurement includes all random components of conditionally constant errors, and therefore, it can happen that random parts of conditionally constant errors in multiple measurements are taken into account twice.

When performing an analysis, however, it is important to distinguish purely random and quasirandom errors. Purely random errors can arise from different reasons. For example, they can arise from noise or small (regarded as permissible)

variations in the influence quantities or the random components of the errors of the measuring equipment.

Quasirandom errors appear in measurements of quantities that are by definition averages when the quantities appearing in the group being averaged are of different size. The difference is not random but is regarded as random and is characterized, just as in the case of a purely random error, by an estimate of the standard deviation.

The error classification studied above is reminiscent of the classification contained in the Guide [1], but the last does not separate absolutely constant errors. There is some difference in terminology also. The Guide avoids the term *error*, and names the components of measurement uncertainty as type A and type B uncertainties. The type A uncertainty is defined there as a component of the measurement uncertainty that is estimated by statistical methods, whereas the type B uncertainty is estimated by nonstatistical methods. These terms are purely arbitrary, and the Vocabulary [2] does not contain them. It is important also that the classification indicator here refers not to the object of classification and to its properties but to the method employed to estimate them, and in general, it is a secondary indicator that follows from something that has not been identified.

The proposed classification does not have this deficiency.

3.3. Mathematical Models of Elementary Errors

A measurement error and uncertainty is calculated based on data of its components; i.e., this is a problem of synthesis, performed mathematically. Correspondingly, elementary errors must be represented by mathematical models. We shall examine all four types of elementary errors from this viewpoint.

Absolutely constant errors. Each such error has a constant value that is the same in any measurement, although it is unknown. Only the limits of these errors are known. A mathematical model of such errors could be a determinate quantity whose magnitude has an interval estimate; i.e., it lies within an interval of known limits. We shall use this model for absolutely constant elementary errors.

We can foresee an objection to this model. Some people think that if the value of the error is unknown, then it can be regarded as a random quantity. However, this is not correct. A model of an object can be constructed only based on what we know about it and not based on what we do not know.

There is another objection. If a determinate model is adopted, then when several absolutely constant errors are summed, their limits must be added arithmetically. This process is equivalent to the assumption that all terms have limiting values and the same sign, which is unlikely. This objection also is invalid. First, the argument "unlikely" is not correct here, because we are not using a probabilistic model. Second, the fact that we do not like the result—the answer seems exaggerated—is also not an argument. In mathematics, precisely the same situation arises in methods of approximate calculations and the limits of errors are added arithmetically.

Fortunately, in a measurement, rarely more than one or two absolutely constant errors exist, and they are, as a rule, insignificant.

Conditionally constant errors. The values of these errors characteristically vary from one measurement to another and from one measuring instrument to another, and they are different under different conditions. In all cases, however, in each such error, the limits of the interval containing any possible realization of such an error remain unchanged.

As a mathematical model of conditionally constant errors, one would like to use a random quantity. For this reason, however, it is necessary to know the probability distribution function corresponding to this random quantity. Best of all, one would like to find this function based on the experimental data. Such an attempt was made for the intrinsic error of measuring instruments. The results of such an investigation were presented in Chapter 2. They showed that the distribution function of the intrinsic error and, of course, the distribution function of the additional errors cannot be found from selective data.

Thus, to adopt a probability model, the form of the distribution function, in this case, must be prescribed. It is well known that among distributions with fixed limits, the uniform distribution has the highest uncertainty (in the sense of information theory). The roundoff error also has known limits, and in mathematics, this error has for a long time been regarded as a random quantity with a uniform probability distribution. For this reason, we shall also assume that the model of conditionally constant errors will be a random quantity with a uniform probability distribution within prescribed limits.

This suggestion was made comparatively a long time ago [41]. At the present time, this model is widely employed in the theory of measurement errors [1], [5], [6].

Purely random errors. Such errors appear in multiple measurements. They are characterized by the standard deviation that is computed from the experimental data.

The form of the distribution function of random errors can, in principle, be found based on the data from each multiple measurement. In practice, however, the number of measurements performed in each experiment is insufficient for this. Every time measurements are performed, it is assumed that the hypothesis of a normal distribution was checked in the preceding experiment. For example, when measures of mass are compared on the standard balances at the D.I. Mendeleev All-Union Scientific-Research Institute of Metrology (USSR), it is assumed that the distribution is normal, but this is not directly checked. True, the results obtained are not inconsistent with the practice so that this assumption is evidently justified.

In general, I have never encountered a case when the normality of the distribution of a random error was checked mathematically and when this has led to misunderstandings. Thus, we shall assume that the mathematical model of random errors is, as a rule, a *normal distributed random quantity.*

Quasirandom errors. As noted above, these errors occur when measuring quantities that are averages by definition, and the value of each separate quantity in the group of quantities being averaged remains constant. These quantities are essentially not random, but in aggregate, they can be regarded as a general collection of quantities, which is possible in accordance with the goal of the measurement, based on agreement of experts. The parameters of the distribution that characterize this distribution should be determined by agreement. Most often the standard deviation is chosen as this parameter.

We shall now discuss the question of interdependence and correlation of elementary errors. Mathematically, it is preferable to regard these errors as correlated quantities, because this approach has great generality. However, such an approach complicates the problem, and most of the time, it is not justified. Under reference conditions, all elementary errors are independent and they are uncorrelated. Exceptions can be encountered in measurements performed under normal operating conditions, especially in the case of indirect measurements and measurements performed with the help of measuring systems, when one and the same influence quantity gives rise to appreciable additional errors in several instruments or components in the measuring channel of the system. An example is a measurement in which a measuring transducer, amplifier, and automatic-plotting instrument is employed. A change in the temperature of the medium can cause these devices to acquire an additional temperature-induced error. Obviously, these additional errors will be interrelated.

3.4. Methods for Describing Random Quantities

Random quantities are studied in the theory of probability, a well-developed field of mathematics. The properties of a random quantity are completely described by the distribution function $F(x)$, which determines the probability that a random quantity X will assume a value less than x:

$$F(x) = P\{X < x\}.$$

The distribution function is a nondecreasing function, defined so that $F(-\infty) = 0$ and $F(+\infty) = 1$.

Together with the distribution function $F(x)$, which is said to be cumulative or integral, the differential function, usually called the probability density $f(x)$, is also widely employed:

$$f(x) = \frac{dF(x)}{dx}.$$

We call attention to the fact that the probability density is a dimensional function:

$$\dim f(x) = \dim \frac{1}{X}.$$

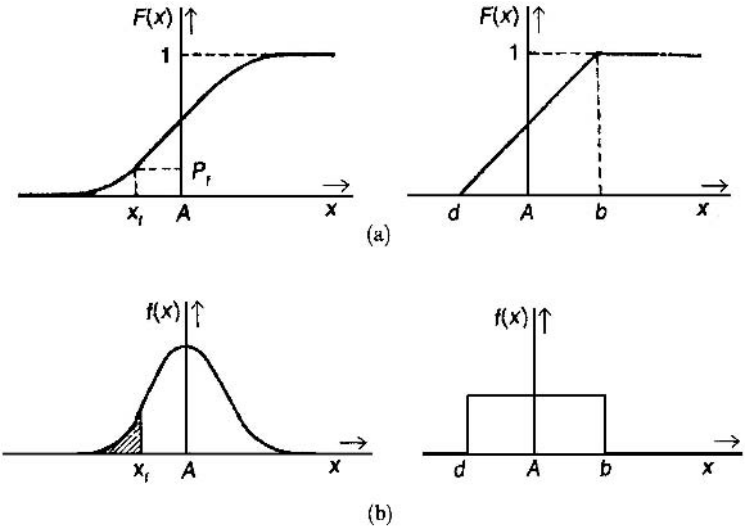


FIGURE 3.1. (a) The probability distribution and (b) the probability density for a normal distribution (on the left) and uniform distribution (on the right) of continuous random quantities.

In the practice of precise measurements one most often deals with normal and uniform distributions. Figure 3.1(a) shows integral functions of these distributions, and Fig. 3.1(b) shows the probability densities of the same distributions.

For the normal distribution, we have

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-A)^2/2\sigma^2},$$
$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(x-A)^2/2\sigma^2} dx, \tag{3.1}$$

The parameter a^2 is the variance, and A is the mathematical expectation of the random quantity.

Calculation of $F(x)$ for some fixed x_f gives the probability $P\{X < x_f\} = P_f$. When the graph of $f(x)$ is used to calculate this probability, it is necessary to find the area under the curve to the left of the point x_f in Fig. 3.1(b).

The normal distribution function obtained by transforming to the random quantity $z = (X - A)/\sigma$ is widely employed in calculations:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \qquad F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-y^2/2} dy. \tag{3.2}$$

Tables of values of the function $\Phi(z)$ defined by the expression

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{0}^{z} e^{-y^2/2} dy \tag{3.3}$$

and called the normalized Gaussian function are often given.

It is obvious that for $z \geq 0$

$$F(z) = 0.5 + \Phi(z).$$

The branch for $z < 0$ is found based on symmetry considerations:

$$F(z) = 0.5 - \Phi(z).$$

A table of the function $\Phi(z)$ is given in the Appendix (Table A.1).

The normal distribution is remarkable in that according to the central limit theorem, a sum of an infinite number of infinitesimal random quantities with an arbitrary distribution has a normal distribution. In practice, the distribution of the sum of a comparatively small number of random quantities already is found to be close to a normal distribution.

The uniform distribution is defined as

$$f(x) = \begin{cases} 0, & x < d, \\ \dfrac{1}{b-d}, & d \leq x \leq b, \\ 0, & x > b, \end{cases} \qquad F(x) = \begin{cases} 0 & x < d, \\ \dfrac{x-d}{b-d}, & d \leq x \leq b, \\ 1, & x > b. \end{cases} \tag{3.4}$$

We shall also use the uniform distribution often.

In addition to continuous random variables, discrete random variables are also encountered in metrology. An example of an integral distribution function and the probability distribution of a discrete random variable are given in Fig. 3.2.

Distribution functions are complete characteristics of random quantities, but they are not always convenient to use in practice. For this reason, random quantities

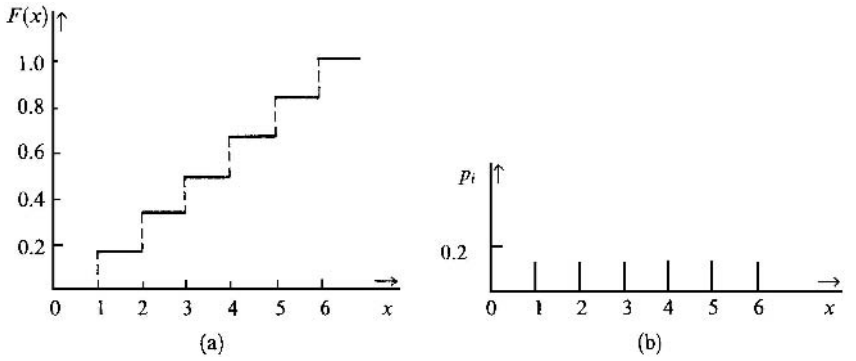


FIGURE 3.2. (a) The probability distribution and (b) the distribution of probabilities of a discrete random quantity.

are also described by their numerical characteristics, and the moments of random quantities are employed.

The initial moments m_k (moments about zero) and central moments μ_k (moments about the mean value) of order k are defined by the formulas

$$\begin{aligned} m_k &= E[X^k] = \int_{-\infty}^{\infty} x^k f(x)dx, \\ m_k &= E[X^k] = \sum_{i=1}^{n} x_i^k p_i. \end{aligned} \tag{3.5}$$

$$\begin{aligned} \mu_k &= E[X - E[X]]^k = \int_{-\infty}^{\infty} (x - E[X])^k f(x)dx, \\ \mu_k &= E[X - E[X]]^k = \sum_{i=1}^{n} (x_i - E[X])^k p_i. \end{aligned} \tag{3.6}$$

In the relations (3.5)–(3.8), the first formulas refer to continuous and the second to discrete random quantities.

Of the initial moments, the first moment ($k = 1$) is most often employed. It gives the mathematical expectation of the random quantity

$$\begin{aligned} m_1 &= E[X] = \int_{-\infty}^{\infty} x f(x)dx, \\ m_1 &= E[X] = \sum_{i=1}^{n} x_i p_i. \end{aligned} \tag{3.7}$$

It is assumed that $\sum_{i=1}^{n} p_i = 1$; i.e., the complete group of events is studied.

Of the central moments, the second moment ($k = 2$) plays an especially important role. It is the variance of the random quantity

$$\begin{aligned} \mu_2 &= \mathrm{V}[X] = E[(X - m_1)^2] = \int_{-\infty}^{\infty} (x - m_1)^2 f(x)\, dx, \\ \mu_2 &= \mathrm{V}[X] = E[(X - m_1)^2] = \sum_{i=1}^{n} (x_i - m_1)^2 p_i. \end{aligned} \tag{3.8}$$

The positive square root of the variance is called the standard deviation of the random quantity

$$\sigma = +\sqrt{V[X]}. \tag{3.9}$$

Correspondingly, $V[X] = \sigma^2$.

The third and fourth central moments are also used in applications. They are used to characterize the symmetry and sharpness of distributions. The symmetry is characterized by the skewness $a = \mu_3/\sigma^3$, and the sharpness is characterized by the excess $e = \mu_4/\sigma^4$. The latter is defined some times as $e' = \mu_4/\sigma^4 - 3$.

The normal distribution is completely characterized by two parameters: $m_1 = A$ and σ. For it, characteristically, $a = 0$ and $e' = 0$. The uniform distribution is also

determined by two parameters: $m_1 = A$ and $l = d - b$. It is well known that

$$m_1 = \frac{d+b}{2}, \qquad V[X] = \frac{(d-b)^2}{12} = \frac{l^2}{12}. \tag{3.10}$$

Instead of l, the quantity $h = l/2$ is often used. Then $V[X] = h^2/3$ and $\sigma(X) = h/\sqrt{3}$.

3.5. Construction of the Composition of Uniform Distributions

So, we have adopted the uniform distribution as a mathematical model of conditionally constant elementary errors. In solving the problem of synthesis of these errors, one must know how to construct the composition of uniform distributions.

The theoretical solution of this problem is well known and is presented, for example, in [52]. For our purposes, it is interesting to clarify the possibility of constructing a simplified solution for the applied problem at hand.

Consider n random quantities x_i $(i = 1, \ldots, n)$, each of which has a uniform distribution centered at zero in the interval $[-\frac{1}{2}; +\frac{1}{2}]$. We introduce the notation $\vartheta = \sum_{i=1}^{n} x_i$. The probability densities of the sum of these random quantities has the form

$$f_n(\vartheta) = \frac{1}{(n-1)^!}\left[\left(\vartheta + \frac{n}{2}\right)^{n-1} - C_n^1\left(\vartheta + \frac{n}{2} - 1\right)^{n-1} + C_n^2\left(\vartheta + \frac{n}{2} - 2\right)^{n-1} + \cdots\right],$$

where the sum must include only those terms in which the additive to ϑ, i.e., $n/2$, $(n/2 - 1)$, and so on, is nonnegative for a given value of n; for example, if $n = 2$, then

$$f_2(\vartheta) = (\vartheta + 1) - 2\vartheta = \begin{cases} 0, & \vartheta \le -1, \\ \vartheta + 1, & -1 < \vartheta \le 0, \\ 1 - \vartheta, & 0 \le \vartheta < 1, \\ 0, & 1 < \vartheta. \end{cases}$$

The probability density of the sum of two terms has the form of a triangle. For $n = 3$, the graph of $f_3(\vartheta)$ consists of three segments of a quadratic parabola and looks very much like the curve of a normal distribution. For $n = 4$, this distribution function is almost indistinguishable from the normal distribution.

Given the equation for the probability density, it is not difficult to find the probability distribution function

$$F_n(\vartheta) = \frac{1}{n!}\left[\left(\vartheta + \frac{n}{2}\right)^n - C_n^1\left(\vartheta + \frac{n}{2} - 1\right)^n + C_n^2\left(\vartheta + \frac{n}{2} - 2\right)^n + \cdots\right]. \tag{3.11}$$

In practice, however, it is desirable to have a simpler and more convenient solution. Such a solution can be found by taking into account that in accordance with the principle of error estimation from above, we are interested in limits $\pm\theta$ for the sum of the components such that the probability $P\{|\vartheta| \le \theta\} > 0.9$.

Bearing the last remark in mind, we shall examine the distribution function $F_n(\vartheta)$ in the extreme intervals $[-n/2, -n/2 + 1]$ and $[n/2 - 1, n/2]$.

For one section, (3.11) assumes the form

$$F_n(\vartheta) = \begin{cases} \dfrac{1}{n!}\left(\vartheta + \dfrac{n}{2}\right)^n & \text{for } -\dfrac{n}{2} < \vartheta - \dfrac{n}{2} + 1, \\ 1 - \dfrac{1}{n!}\left(\vartheta - \dfrac{n}{2}\right)^n & \text{for } \dfrac{n}{2} - 1 < \vartheta < \dfrac{n}{2}. \end{cases}$$

The composition of the distributions is symmetric relative to the ordinate axis.

We shall discuss how to calculate, given the probability distribution, the limits of the confidence interval corresponding to a fixed value α of the confidence probability. The limits of the confidence interval corresponding to α are $\pm\theta_\alpha$.

By definition, the probability that the true value of a quantity lies within the confidence interval $[-\theta, +\theta]$ is α. Therefore, the probability that the quantity does not lie in the confidence interval is $(1 - \alpha)$. If the distribution is symmetric relative to 0 (and we are studying a symmetric distribution), then the probability that the quantity will take on a value less than $-\theta$ will be equal to the probability that it will take on a value greater than $+\theta$. These probabilities are obviously equal to $(1 - \alpha)/2$.

Consider first the left-hand branch of the distribution function. The probability corresponding to the point $-\theta$ [the arguments (points) of the distribution function are also called quantiles of the distribution] is equal to $P\{\vartheta \le -\theta\} = (1 - \alpha)/2$. We shall now consider the right-hand branch. The probability that $\vartheta \le +\theta$ will obviously be equal to $1 - [(1 - \alpha)/2] = (1 + \alpha)/2$.

We shall now return to our problem. Given $F_n(\vartheta)$ and α, we are required to find the quantiles $-\theta$ and $+\theta$. Their absolute values are equal. For this reason, we shall only calculate $-\theta$, and we have the condition

$$P\{\vartheta \le -\theta\} = F_n(-\theta) = \frac{1}{n!}\left(-\theta + \frac{n}{2}\right)^n = \frac{1 - \alpha}{2}, \tag{3.12}$$

from which θ can be calculated. We shall represent in the following form the values of θ found from formula (3.12):

$$\theta_\alpha = k\sqrt{\sum_{i=1}^{n} \theta_i^2}, \tag{3.13}$$

where θ_i is the limit of the range of values of $x_i (-\theta_i \le x_i \le +\theta_i)$, where k is a correction factor.

In the case at hand, $\theta_i = 1/2$ for all $i = 1, \ldots, n$; i.e.,

$$\theta = k\sqrt{n}/2, \qquad k = 2\theta/\sqrt{n}. \tag{3.14}$$

TABLE 3.1. Values of the coefficient k as a function of the number of terms and confidence probability.

	Values of the coefficient k for confidence probability α			
Number of terms n	0.90	0.95	0.99	0.9973
2	0.97	1.10	1.27	1.34
3	0.96	1.12	1.37	1.50
4	*	1.12	1.41	1.58
5	*	*	*	1.64
...	...	...	...	...
∞	0.95	1.13	1.49	1.73

* Cases for which the coefficient k is not calculated, because one interval is not enough for that n.

Formula (3.13) is convenient for calculations, and for this reason, we shall investigate the dependence of the coefficient k on α and n. The calculations are performed as follows. Given α and n, we find θ from (3.12). Next, the correction factor k is found for the given values of α and n from formula (3.13) or (3.14).

For example, let $\alpha = 0.99$ and $n = 4$. Then $(1 - \alpha)/2 = 0.005$. Let us check whether the value of θ corresponding to this probability falls within the left extreme interval $[-2, -1]$. For this reason, we shall find the probability corresponding to the highest value of ϑ in this interval, i.e., $F_4(-1)$:

$$F_4(-1) = \frac{1}{4!}(-1+2)^4 = \frac{1}{1 \times 2 \times 3 \times 4} = 0.041.$$

As $0.005 < 0.041$, the value of θ of interest to us lies in this interval.

Substituting the initial data into (3.12), we find θ:

$$\frac{1}{4!}(-\theta+2)^4 = 0.005, \qquad -\theta + 2 = \sqrt[4]{24 \times 0.005}, \qquad \theta = 1.41.$$

Having found θ, we obtain from formula (3.14):

$$k = \frac{2 \times 1.41}{\sqrt{4}} = 1.41.$$

The values of k for other values of α and n were calculated analogously and are presented in Table 3.1.

The value of k for $n \to \infty$ was found using the fact that by virtue of the central limit theorem, the resulting distribution may be regarded as being normal.

We can write

$$\vartheta = \sum_{i=1}^{n} x_i, \qquad V[\vartheta] = V\left[\sum_{i=1}^{n} x_i\right] = \sum_{i=1}^{n} V[x_i], \qquad E[x_i] = 0.$$

But, as is well known, $V[x_i] = \theta_i^2/3$. Therefore

$$V[\vartheta] = \frac{\sum_{i=1}^{n} \theta_i^2}{3}, \qquad \sigma[\vartheta] = \sqrt{\frac{1}{3}\sum_{i=1}^{n} \theta_i^2}. \tag{3.15}$$

Thus, if $n \to \infty$, we have a random quantity with a normal distribution $N(0, \sigma)$. We shall calculate the absolute value of the limits of the confidence interval from its upper limit $\theta = z_p\sigma$, where z_p is the quantile of the normal distribution corresponding to the probability $p = (1 + \alpha)/2$ (see above). Thus, we obtain

$$\theta = \frac{z_p}{\sqrt{3}}\sqrt{\sum_{i=1}^{n} \theta_i^2}. \tag{3.16}$$

Comparing (3.16) with (3.13), we find

$$\underset{n\to\infty}{k} = \frac{z_p}{\sqrt{3}}.$$

For $\alpha = 0.9973$, we obtain $z_p = 3$ and $k = 1.73$.

Looking at the table obtained in this manner (Table 3.1), it should be noted that the correction factor k has the interesting property that for $\alpha \leq 0.99$, it is virtually independent of the number of terms. We can make use of this property and take for k the average values:

α	0.90	0.95	0.98	0.99
k	0.95	1.10	1.30	1.40

The error caused by using the average values of k, as one can see by comparing them with the exact values given in Table 3.1, does not exceed 10%.

The small effect of the number of terms indicates indirectly that it is not necessary to assume, as was done above, that all θ_i are equal. Thus, if one of the terms θ_l is reduced, then in the limit, instead of n, we obtain $n - 1$ terms. The value of the factor k, however, in the process remains practically unchanged. If, on the other hand, θ_l is gradually increased, then the factor k will decrease.

The dependence of k on the ratio $c = \theta_l/\theta_0$ for $\alpha = 0.99$ is given in Fig. 3.3; θ_0 is the absolute value of the remaining terms, which are assumed to be equal.

The factor k can also be calculated using the formulas approximating the curves presented in Fig. 3.3. For $\alpha = 0.99$ and $n = 4$, this formula is

$$k = 1.45 - 0.05\frac{\theta_l}{\theta_0}. \tag{3.17}$$

Formula (3.16) can be used instead of (3.13) to calculate θ when the number of terms is large. However, as follows from the above-presented estimate of the error of calculations based on formula (3.13), the accuracy cannot be increased by more than 10% (for $\alpha = 0.99$). At the same time, formula (3.13) is also useful

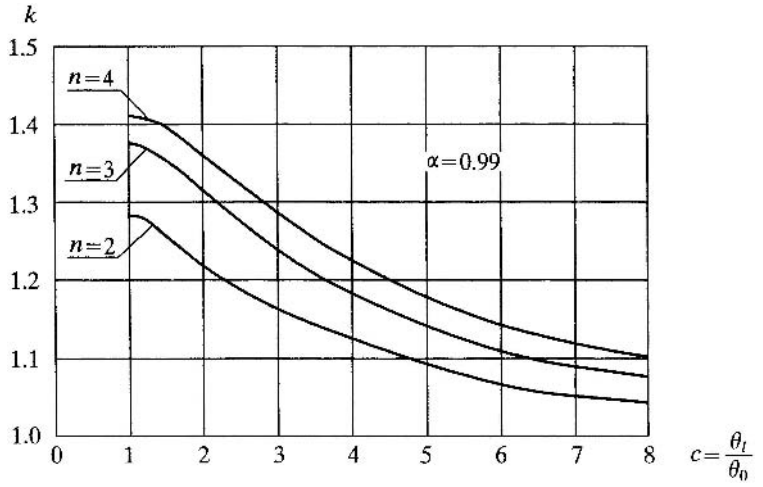


FIGURE 3.3. The coefficient k as a function of the change in one of the terms relative to the other terms ($n = 2, 3, 4$).

for summing a small number of terms. For this reason, for practical calculations, relation (3.13) is preferable.

3.6. Universal Method for Constructing the Composition of Distributions

In the general case, to combine random quantities, it is necessary to construct the composition of the distributions of the terms. If the distribution functions are given analytically, then their composition is found either by direct integration of the derivatives of the functions or by using the characteristic functions, which usually simplifies the solution.

In practice, however, the analytical form of the distribution functions is usually unknown. Based on the experimental data, it is possible only to construct a histogram, and an error is unavoidably made by passing from the histogram to the distribution function. For this reason, we shall study the summation of random quantities whose distribution is given by histograms and not by distribution functions [30].

Suppose that we are required to find the distribution function of the random quantity $\zeta = \zeta_1 + \cdots + \zeta_n$, where ζ_i is a random quantity given by a histogram with m_i intervals in the region of possible values of ζ_i with the limits a_i and b_i.

Thus, the interval

$$[a_i, b_i] = l_{i1} + l_{i2} + \cdots + l_{im_i}, \qquad i = 1, \ldots, n.$$

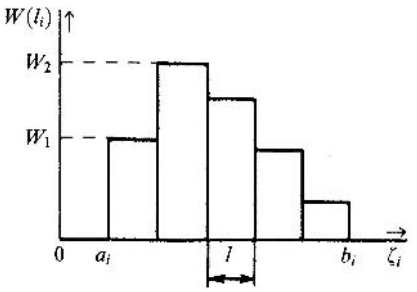


FIGURE 3.4. Histogram of the distribution of some random quantity.

We shall assume that the probability that the random quantity falls within each interval of the histogram is equal to the area of the part of the histogram that corresponds to this interval (the area of the corresponding column of the histogram):

$$P\{\zeta_i \varepsilon\ l_{ik}\} = p_{ik},$$

where $k = 1, \ldots, m_i$ is the number of the interval of the histogram of the distribution of the random quantity ζ_i.

Figure 3.4 shows as an example a histogram with five intervals of equal length $l_i = l$, so that $b_i - a_i = 5l$. For this histogram,

$$p_{i1} = W_1 l, \qquad p_{i2} = W_2 l, \ldots, \qquad p_{i5} = W_5 l,$$

where $W_1, \ldots, W_5$ are the heights of the columns of the histogram; by construction, the area of the entire histogram is equal to unity; i.e., $\sum_{k=1}^{5} p_{ik} = 1$.

We recall that in constructing histograms (which are constructed based on empirical data), the height of the column of each interval is found by dividing the relative frequency with which the values fall within the corresponding interval by the length of this interval. This frequency is an empirically obtained estimate of the probability of the corresponding event.

Next, we shall represent continuous random quantities by discrete random quantities corresponding to them. For this reason we denote by a_{ik} the center of each interval l_{ik} and we introduce a new random quantity η_i, which corresponds to the random quantity ζ_i so that η_i assumes the value a_{ik} with probability p_{ik}. This result is possible, because from what we have said above, it is obvious that

$$\sum_{k=1}^{m_i} p_{ik} = 1 \qquad \text{for all} \qquad i = 1, \ldots, n.$$

It is desirable to represent the obtained data for each random quantity η_i by a table of the following form:

η_i	a_{i1}	a_{i2}	$\cdots$	a_{im_i}
	p_{i1}	p_{i2}	$\cdots$	p_{im_i}

We shall now study the random variable $\eta = \eta_1 + \eta_2 + \cdots + \eta_n$. We obtain all its possible values by sorting through all combinations of the obtained realizations of a_{ik} of the components η_i.

For the calculations, it is convenient to write out the possible values of the random quantities in a single table of the form

$$\begin{array}{llll} \eta_1 & a_{11} & \ldots & a_{1m_1}, \\ \eta_2 & a_{21} & \ldots & a_{2m_2}, \\ & & \ldots & \\ \eta_n & a_{n1} & \ldots & a_{nm_n}. \end{array}$$

Next we calculate the values of the random quantity η that correspond to each possible combination of realizations of the random quantities η_i,

$$\eta_t = a_{1k_1} + a_{2k_2} + \cdots + a_{nk_n}$$

and the corresponding probabilities, which we find from the formula

$$p_t = P\{\eta_1 = a_{1k_1}, \eta_2 = a_{2k_2}, \ldots\} = \prod_{i=1}^{n} p_{ik_i}. \tag{3.18}$$

Adding the probabilities that correspond to one and the same realization $\eta_t = a_t$, we obtain the probability that the random quantity η assumes each possible value from series of $a_1, \ldots, a_N$.

The number of combinations of terms will be $\prod_{i=1}^{n} m_i$, but because among them there are terms whose values are the same,

$$N \le \prod_{i=1}^{n} m_i. \tag{3.19}$$

The obtained data make it possible to construct a step function of the distribution $F_1(x)$ of the random quantity η:

$$F_1(x) = \sum_{t} P\{\eta = a_t\}, \qquad a_t \le x. \tag{3.20}$$

The curve $F_1(x)$ is the first approximation to the distribution function $F(x)$ sought. The obtained step function can be smoothed by the method of linear interpolation as follows.

We find the center of the intervals $[a_t, a_{t+1}]$ with $t = 1, \ldots, N-1$:

$$\beta_t = \frac{a_{t+1} + a_t}{2}. \tag{3.21}$$

From the points β_t, we raise perpendiculars up to the broken line $F_1(x)$. We obtain points with the coordinates $(\beta_t, F_1(x))$ for $t = 1, \ldots, N-1$. To the points obtained, we associate points at which the distribution function assumes the values $F_1(\beta_0) = 0$ and $F_1(\beta_n) = 1$:

$$\beta_0 = \sum_{i=1}^{n} a_i, \qquad \beta_N = \sum_{i=1}^{n} b_i. \tag{3.22}$$

Joining the $N + 1$ points so obtained with straight lines, we obtain the function $F_2(x)$, which is the approximation sought.

The method presented above gives a solution of the problem using all available information and does not introduce any distortions. In the general case, however, $V[\zeta_i] \neq V[\eta_i]$ and the variance of the random quantity with the distribution $F_1(x)$ or $F_2(x)$ can differ from the variance of the random quantity ζ. For this reason, if the terms are independent, the variance of their sum must be calculated in the standard manner using the formula

$$V[\zeta] = V\left[\sum_{i=1}^{n} \zeta_i\right] = \sum_{i=1}^{n} V(\zeta_i).$$

We note that for $n > 5$, the distribution of the sum of terms can be regarded as a normal distribution, which is completely determined by the variance and the mathematical expectation. Both parameters can be easily calculated from the parameters of the terms

$$E[\zeta] = \sum_{i=1}^{n} E[\zeta_i], \qquad V[\zeta] = \sum_{i=1}^{n} V[\zeta_i],$$

and for this reason, the calculations presented above are useful only if the number of terms $n < 5$.

It should also be noted that the method presented above for constructing a composition of distributions is also useful in the case when the distributions of the random quantities are given in analytic form. The smooth curve expressing the density of the distribution of the random quantity ζ_i is replaced by a step curve with m_i steps, in a manner so that the area it bounds, as also the area under the smooth curve, is equal to unity. If the branches of the smooth curve of the distribution function approach the abscissa axis asymptotically, this distribution is replaced by a truncated distribution.

It is also obvious that this method is useful both for the case of discrete quantities ζ_i and for the mixed case.

In general, the method examined above is essentially an algorithm for constructing numerically the composition of distributions and can be easily implemented with the help of computers.

We shall illustrate the method with an example. Let $\zeta = \zeta_1 + \zeta_2$, where ζ_1 has a normal distribution with the density

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-(x-2)^2/2},$$

and ζ_2 has a distribution with a uniform density $f_2(x) = 1/6$.

For a normal distribution with the parameters $A = 2$ and $\sigma = 1$, we shall take the domain of ζ_1 to be $[A - 3\sigma, A + 3\sigma] = [-1, 5]$. We divide this interval into five intervals ($m_1 = 5$), symmetrically arranged relative to the point 2—the mathematical expectation:

$$[-1, 5] = [-1, 0.5] + [0.5, 1.5] + [1.5, 2.5] + [2.5, 3.5] + [3.5, 5].$$

For the random quantity ζ_2, whose domain is the interval $[-3, 3]$ and which has a distribution with a uniform density, we assume $m_2 = 3$:

$$[-3, 3] = [-3, -1] + [-1, 1] + [1, 3].$$

Next we calculate the probability that the random quantities fall into the corresponding intervals. For the normal distribution, we have

$$p_{11} = \int_{-1}^{0.5} \frac{1}{\sqrt{2\pi}} e^{-(x-2)/2}\, dx = 0.067,$$

$$p_{12} = \int_{0.5}^{1.5} \frac{1}{\sqrt{2\pi}} e^{-(x-2)^2/2}\, dx = 0.242,$$

$$p_{13} = \int_{1.5}^{2.5} \frac{1}{\sqrt{2\pi}} e^{-(x-2)^2/2}\, dx = 0.382.$$

In view of the symmetry of the normal distribution

$$p_{14} = p_{12} = 0.242, \qquad p_{15} = p_{11} = 0.067.$$

For the uniform distribution

$$p_{21} = \int_{-3}^{-1} \frac{1}{6}\, dx = \frac{1}{3}, \qquad p_{22} = \int_{-I}^{1} \frac{1}{6}\, dx = \frac{1}{3}, \qquad p_{23} = \int_{1}^{3} \frac{1}{3}\, dx = \frac{1}{3}.$$

Next we find the centers of the constructed intervals:

$$a_{11} = \frac{-1 + 0.5}{2} = -0.25, \qquad a_{12} = \frac{0.5 + 1.5}{2} = 1,$$

$$a_{13} = \frac{1.5 + 2.5}{2} = 2, \qquad a_{14} = \frac{2.5 + 3.5}{2} = 3, \qquad a_{15} = \frac{3.5 + 5}{2} = 4.25,$$

$$a_{21} = \frac{-3 - 1}{2} = -2, \qquad a_{22} = \frac{-1 + 1}{2} = 0, \qquad a_{23} = \frac{1 + 3}{2} = 2.$$

This process determines η_1, which assumes values a_{1k} with probabilities p_{1k}, where $k = 1, \ldots, 5$, and η_2, which assumes values a_{2k} with probabilities p_{2k}, where $k = 1, 2$, and 3. As a result of the calculations we have obtained

$$\eta_1 \begin{cases} a_{1k} & -0.25 & 1 & 2 & 3 & 4.25, \\ p_{1k} & 0.067 & 0.242 & 0.382 & 0.242 & 0.067, \end{cases}$$

$$\eta_2 \begin{cases} a_{2k} & -2 & 0 & 2, \\ p_{2k} & 0.333 & 0.333 & 0.333. \end{cases}$$

Next we transfer to the random quantity $\eta = \eta_1 + \eta_2$. We estimate the number of different terms η from formula (3.19). In our case, $m_1 = 5$, $m_2 = 3$, and $N \le 15$.

We shall represent the values obtained for η_1 and η_2 in the form of a table:

η_1	-0.25	1	2	3	4.25,
η_2	-2.00	0	2	—	—,

We find the values of $\eta = \eta_1 + \eta_2$ with the help of this table. The order of the calculations is explained in Table 3.2.

Next we arrange the values of η_t in increasing order. To each value of η_t there corresponds a unique probability p_t. If one and the same value of η_t is encountered several times, then the probability of this value is taken to be the sum of these probabilities. We obtain

η_t	-2.25	-1	-0.25	0	1	1.75	2
p_t	0.022	0.081	0.022	0.127	0.162	0.022	0.127

η_t	2.25	3	4	4.25	5	6.25
p_t	0.22	0.162	0.127	0.022	0.081	0.022

As $\eta = 1$ and $\eta = 2$ were encountered twice, $N = 13$.

Based on the data obtained, using (3.20), it is not difficult to construct $F_1(x)$. The values of this function in the intervals found are presented in Table 3.3, and the corresponding graph is given in Fig. 3.5 in the form of a stepped line.

We find β_t for $t = 1, \ldots, 12$ from (3.21), and we determine β_0 and β_{13} from (3.22). Using these calculations as well as the data of Table 3.3, we construct the distribution function $F_2(x)$. The function $F_2(x)$ is plotted in Fig. 3.5 as a broken

TABLE 3.2. Data for sorting through variants of sums of the random quantities η_1 and η_2 and the corresponding probabilities.

η	p
$-0.25 - 2 = -2.25$ $-0.25 + 0 = -0.25$ $-0.25 + 2 = 1.75$	$0.067 \times 0.333 = 0.022$
$1 - 2 = -1$ $1 + 0 = 1$ $1 + 2 = 3$	$0.242 \times 0.333 = 0.081$
$2 - 2 = 0$ $2 + 0 = 2$ $2 + 2 = 4$	$0.382 \times 0.333 = 0.127$
$3 - 2 = 1$ $3 + 0 = 3$ $3 + 2 = 5$	$0.242 \times 0.333 = 0.081$
$4.25 - 2 = 2.25$ $4.25 + 0 = 4.25$ $4.25 + 2 = 6.25$	$0.067 \times 0.333 = 0.022$

TABLE 3.3. Data for the stepped approximation to the distribution function of a sum of two random quantities studied.

x		$F_1(x)$
$-\infty$	-2.25	0.000
-2.25	-1.00	0.022
-1.00	-0.25	0.103
-0.25	0.00	0.125
0.00	1.00	0.252
1.00	1.75	0.414
1.75	2.00	0.436
2.00	2.25	0.563
2.25	3.00	0.585
3.00	4.00	0.747
4.00	4.25	0.874
4.25	5.00	0.896
5.00	6.25	0.978
6.25	∞	1.000

line connecting the points $(\beta_t, F_1(\beta_t))$ for $t = 0, \ldots, 13$. The numerical values of $F_2(x)$ for $x = \beta_t$, where $t = 0, \ldots, 13$ are presented in Table 3.4.

As we have already mentioned, the approximation of the limiting distribution function $F(x)$ by $F_2(x)$ can be improved by reducing the subdivisions of the domains of the starting random quantities.

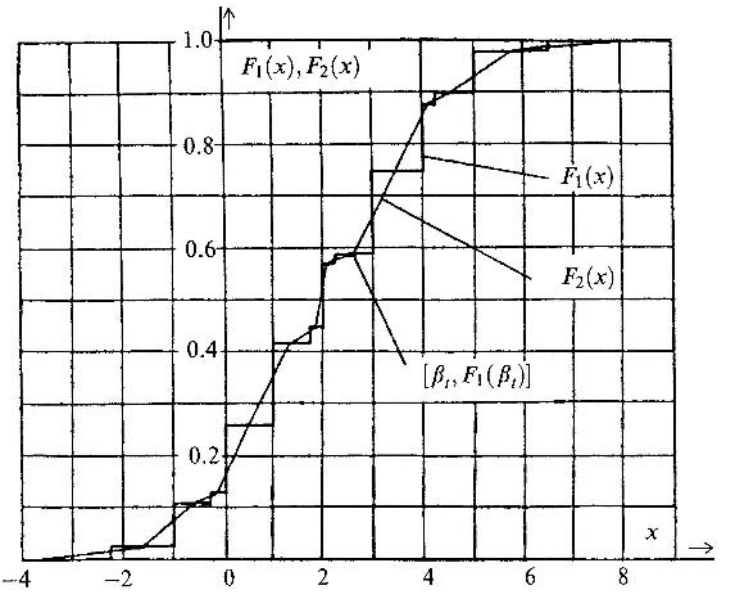


FIGURE 3.5. Step and linear approximations of the distribution function.

TABLE 3.4. Data for the linear approximation to the distribution function of the sum of two random quantities studied.

t	β_t	$F_2(x)$
0	−4.00	0.000
1	−1.62	0.022
2	−0.62	0.103
3	−0.14	0.125
4	0.50	0.252
5	0.14	0.414
6	1.87	0.436
7	2.12	0.563
8	2.62	0.585
9	3.50	0.747
10	4.12	0.874
11	4.62	0.896
12	5.62	0.978
13	8.00	1.000

It is interesting to note that the solution given above makes it possible to find the edges of the distribution function $F_2(x)$ without constructing the entire function. In many cases, this is the main problem.

3.7. Natural Limits of Measurements

For metrology as the science of measurements, it is of fundamental interest to estimate the limiting possibilities of measurements. First, extremely small and extremely large measurable quantities must be estimated. Next, in the case when instantaneous values are measured, the question of the maximum rate of change of the quantity arises. For functionals, such as the effective current, it is important to establish both the maximum and the minimum frequency of the initial process. It is obviously possible to add to this list.

Among all limiting parameters, the lower limits of measurements are of greatest interest, because they are determined by the physical, i.e., natural, limitations. Comparing these limits with the limits that the real measuring instruments permit makes it possible to judge the level of development of measuring instruments and stimulates improvements in their construction. For this reason, we shall confine our attention to the natural limits of measurements.

3.7.1. *Limitations Imposed by Thermal Noise*

Measurements are always accompanied by an interaction of the object of study and the measuring instruments. For this reason, the limiting possibilities of

measurements must be estimated for the measuring instruments together with the object of study—the carrier of the physical quantity.

We shall first study instruments that have an inertial moving system and elastic elements that keep it in a position of equilibrium. Examples are galvanometers, some types of balances, and so on. Such systems are modeled mathematically by (2.3) and (2.4). The moving part of these instruments is continuously bombarded by air molecules (and molecules of liquid, if the oscillations are damped by a liquid). On average, the number of impacts and their effect are the same on all sides of the system. But at any given moment in time, the effect of impacts from one side can be greater than from another side, whereas at the next instant, the situation is reversed. As a result, careful observations reveal continuous oscillations of the movable part of the instrument around the position of equilibrium. These fluctuations limit the possibilities of instruments.

According to a well-known theorem of statistical physics, in the state of thermal equilibrium with a medium at temperature T, to each degree of freedom of the body, there is associated an average energy of fluctuations equal to $\bar{\varepsilon} = kT/2$, where k is Boltzmann's constant. The movable part of an instrument has one degree of freedom. For this reason, the average energy of the fluctuations of the movable part is equal to $\bar{\varepsilon}$. But this energy is equal to the average strain energy of the elastic elements $\bar{P} = W\overline{\alpha^2}/2$, where W is the stiffness of the elastic elements and α is the strain, i.e., the displacement of the moving part. From the equality $\bar{\varepsilon} = \bar{P}$, it follows that

$$\overline{\alpha^2} = \frac{kT}{W}. \tag{3.23}$$

We shall transform this formula, introducing into it the measured quantities. It is obvious that such a transformation cannot be universal: It depends on the type of measured quantity and on the principle of operation of the instrument. Consider, for example, a moving-coil galvanometer. From the relation between the current strength I in the steady-state regime, we have $\alpha W = \Phi I$, where Φ is the magnetic constant of the galvanometer. In addition, the operating parameters β and ω_0 are related to the structural parameters J, P, and W, where J is the moment of inertia of the moving part and P is the damping constant [see (2.3) and, for example, [42]], $\omega_0^2 = W/J$, and $2\beta/\omega_0 = P/W$.

With the help of these relations, (3.23) can be transformed as follows:

$$\overline{I^2} = \frac{kTW}{\Phi^2} = \frac{kT}{\sum R}\frac{W}{P} = \frac{kT}{\sum R}\frac{\omega_0}{2\beta}.$$

Here $\sum R$ is the sum of the resistances of the moving coil of the galvanometer and the external circuit. The damping constant P is related to $\sum R$ by the equation $P = \Phi^2/\sum R$.

The mean-square fluctuations are usually written not for the angular frequency ω_0 but for the period of free oscillations of the system $T_0 = 2\pi/\omega_0$. Then we obtain

$$\overline{I^2} = \frac{\pi kT}{\sum R}\frac{1}{\beta T_0}. \tag{3.24}$$

For the measured voltage E, we obtain analogously

$$\overline{E^2} = \pi kT \frac{\sum R}{\beta T_0}. \tag{3.25}$$

It is well known that for $\beta > 0.8$, $\beta T_0 \approx t_r$ [42], where t_r is the response time of the moving part of the instrument. For this reason, instead of (3.24) and (3.25), we can write

$$\overline{I^2} = \frac{\pi kT}{\sum R} \frac{1}{t_r}, \qquad \overline{E^2} = \pi kT \left(\sum R\right) \frac{1}{t_r}.$$

Analogous arguments for torsion balances give a relation that is similar to (3.24) and (3.25): $\overline{m_x^2} = (kTP/g^2)(\omega_0/2\beta)$, where g is the acceleration of gravity. In general, we can write

$$\overline{x^2} = kTC_x \frac{1}{t_r}, \tag{3.26}$$

where x is the measured quantity and C_x is a constant, determined by the principle of operation and the construction of the instrument. For balances, $C_x = \pi P/g^2$, whereas for a galvanometer,

$$C_I = \frac{\pi}{\sum R}, \qquad C_E = \pi \sum R.$$

The expressions (3.24)–(3.26) show that the mean-square flucuations of the indications of instruments with an inertial moving part are inversely proportional to their response time. But the response time is also the minimum measurement time (approximately). Moreover, it is obvious that the minimum value of the measured quantity is related to $\sqrt{\overline{x^2}}$ by the fixed accuracy of the measurement. For this reason, it follows from expressions (3.24)–(3.26) that

$$x_{\min} = K \frac{1}{\sqrt{t_r}}, \tag{3.27}$$

where $x_{\min}$ is the minimum value of the measured quantity and K is a constant.

The moving part of balances can interact with the medium in only one way: by means of collisions with molecules of the medium (for example, air). For a galvanometer, the situation is different. Apart from mechanical contact with the medium, the moving part of these instruments also interacts with the medium through the electric circuit, in which electron velocity fluctuations occur. Relations (3.24) and (3.25) were derived because the galvanometer interacts with the medium only by means of collisions of air molecules with the moveable part of the galvanometer. But these relations can be derived under the assumption that the entire interaction occurs as a result of fluctuations of the electrons' velocity in the input circuit. In general, in statistical physics, it has been established that for fluctuations of a system, it makes no difference how the system interacts with the medium; only the temperature of the system is important.

Thermal fluctuations in electric circuits are usually calculated with the help of Nyquist's formula:

$$\overline{E^2} = 4kTR\Delta f, \tag{3.28}$$

where $\overline{E^2}$ is the mean-square noise of emf that is brought in a circuit with an active resistance R and Δf is the frequency band in which the noise emf is calculated. Nyquist's formula is often presented as an expression for the spectral density of the square of the noise emf:

$$F_E(f) = 4kTR. \tag{3.29}$$

The spectral density may be assumed to be constant up to very high frequencies, which correspond to the collision frequency of the charge carriers.

Consider one of the best galvanometers—a galvanometer of the type Z_c manufactured by the Kipp Company. For $T_0 = 7$s, $\beta = 1$, $\sum R = 50\Omega$, and $S_u = 1.7 \times 10^7$mm m/V. Therefore, at room temperature ($T = 293$ K), we shall have ($k = 1.38 \times 10^{-23}$ J/K), $\overline{E^2} = \pi \times 1.38 \times 10^{-23} \times 293 \times 50/1 \times 7 = 9.06 \times 10^{-20}$ V^2; i.e., $\bar{E} = 3 \times 10^{-10}$ V.

The galvanometer constant $C_u = 1/S_u = 6 \times 10^8$ V/mm m. If the beam length is taken to be 2 m rather than 1 m, then we find that to a displacement of 1 mm, there corresponds a voltage of about $E_m = 3 \times 10^{-8}$ V. Further increasing of the beam length has no effect, because the effect of the shaking of the ground and the base to which the galvanometer is fastened usually increases correspondingly. The difference between $E_m = 3 \times 10^{-8}$ V and $\bar{E} = 3 \times 10^{-10}$ V is very large, and it is obvious that the galvanometer does not exhibit thermal noise. In precisely the same way, thermal noise is also usually not observed when using balances. The situation is different in the case of instruments with amplifiers. For example, consider an electronic-measuring instrument having an input circuit with resistance R and a wideband amplifier with a large gain. The thermal noise can now be appreciable. An analogous situation exists with electromechanical devices.

The sensitivity of modern galvanometers to shaking has been radically reduced by using taut bands and liquid dampers for damping the transverse oscillations of the moving part, which makes it possible to increase the beam length. But, instead, the rotation of the moving part is indicated photoelectro-optically, which is more efficient. The noise associated with the electronic circuit is suppressed with the help of negative feedback. Such devices can be made to be so sensitive that it is possible to observe thermal noise in their input circuits. To calculate this noise, it is necessary to take into account that the photoelectric amplifier and the instrument at the output have certain inertial properties. We shall examine, as an example, photogalvanometric self-balancing amplifiers [42].

The structural arrangement of the self-balancing amplifier is shown in Fig. 3.6. Block 1 is a galvanometer together with its input circuit; block 2 is a photoelectrooptic transducer—amplifier; block 3 is the feedback block (balancing resistor); block 4 is the output device.

We are interested in the noise at the output of the self-balancing amplifier. As is well known, the spectral density $F(\omega)$ of the noise at the output is related to the

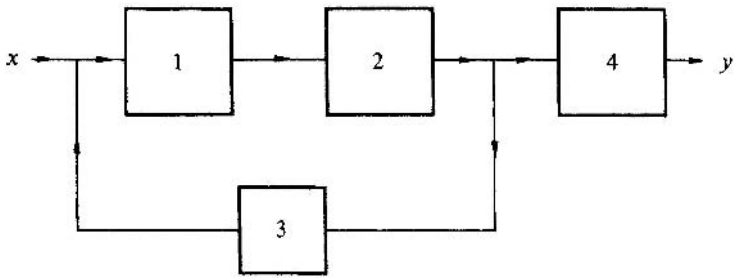


FIGURE 3.6. Structural layout of a photogalvanometric self-balancing amplifier.

spectral density of the noise at the input $F_{E,I}(\omega)$ by the relation

$$F(\omega) = F_{E,I}(\omega)|W(j\omega)|^2, \tag{3.30}$$

where $|W(j\omega)|$ is the modulus of the amplitude-frequency response of the system and ω is the angular frequency. For the angular frequency $\omega = 2\pi f$, formula (3.29), determining the spectral density of the thermal noise at the input of the system, assumes the form

$$F_E(\omega) = \frac{2kTR}{\pi}. \tag{3.31}$$

The self-balancing amplifier's transfer function, which relates the indications of the output instrument and the measured emf, is expressed by the following formula, if the output instrument is calibrated based on the input:

$$W(p) = \frac{1}{[\varkappa p^3 + (1 + 2\beta\varkappa)p^2 + 2\beta p + 1](q^2p^2 + 2\beta_{\text{out}}q + 1)}. \tag{3.32}$$

Here $p = s/\omega_0$, $\varkappa = \tau\omega_0$, $q = \omega_0/\omega_{\text{out}}$, ω_0 is the angular frequency of the characteristic oscillations of the self-balancing amplifier; ω_{out} is the angular frequency of the characteristic oscillations of the output instrument; s is the Laplace operator; and τ is the time constant of the photoelectro-optic transducer amplifier. In addition, β_{out} is the damping ratio of the output instrument. The transfer function (3.32) does not take into account the steady-state residual error of the self-balancing amplifier, because it is always small. It is also assumed that the damping of the output device does not depend on the resistance of its circuit.

The operator p is dimensionless and corresponds to the relative frequency $\eta = \omega/\omega_0$. Substituting $p = j\eta$ into (3.32), we obtain the amplitude-frequency response in complex form. Its absolute value can be represented in the form

$$|W(j\eta)| = \frac{1}{|C_5(j\eta)^5 + C_4(j\eta)^4 + C_3(j\eta)^3 + C_2(j\eta)^2 + C_1(j\eta) + C_0|}.$$

In our case, $C_0 = 1$, $C_1 = 2\beta + 2\beta_{\text{out}}q$, $C_2 = 2\beta\varkappa + 4\beta\beta_{\text{out}}q + q^2 + 1$, $C_3 = \varkappa + 4\beta\beta_{\text{out}}\varkappa q + 2\beta_{\text{out}}q + 2\beta q^2$, $C_4 = 2\beta_{\text{out}}\varkappa q + 2\beta\varkappa q^2 + q^2$, and $C_5 = \varkappa q^2$.

The spectral density of the noise at the input must also be expressed for the relative frequency. As $F_E(\omega)\Delta\omega = F_E(\eta)\Delta\eta$,

$$F_E(\eta) = F_E(\omega)\omega_0 = \frac{2}{\pi}kTR\omega_0,$$

or as $\omega_0 = 2\pi/T_0$, $F_E(\eta) = 4kTR/T_0$.

So, the general expression for the spectral density of the noise in the indications of the photogalvanometric self-balancing amplifier has the form

$$F(\eta) = \frac{4kTR}{T_0}\frac{1}{|C_5(j\eta)^5 + C_4(j\eta)^4 + C_3(j\eta)^3 + C_2(j\eta)^2 + C_1(j\eta) + C_0|^2}. \tag{3.33}$$

Now we can calculate the mean-square fluctuations of the indications of the self-balancing amplifier: $\overline{E^2} = \int_0^\infty F(\eta)\,d\eta$. Integrals of expressions of the type (3.33) for stable systems are well known; they can be found in books on automatic control,[2] and for the specific formula (3.23), the integral is given in [42]

$$\overline{E^2} = \frac{2\pi kT\sum R}{T_0} \times \frac{C_2^2C_5 + C_1C_4^2 - C_2C_3C_4 - C_0C_4C_5}{C_0C_3^2C_4 - C_1C_2C_3C_4 - 2C_0C_1C_4C_5 + C_0^2C_5^2 + C_0C_2C_3C_5 + C_1C_2^2C_5 + C_1^2C_4^2}. \tag{3.34}$$

Formula (3.34) is complicated, but for a specific instrument, all coefficients are simply numbers, so that this formula is easy to use. As shown in [42], Eq. (3.34) yields formulas for all particular cases: an individual galvanometer, a galvanometric amplifier with noninertial transducer–amplifier ($\varkappa = 0$), and a galvanometric amplifier with current output. If measurement of the current strength in the circuit with the resistance R (including the resistance of the galvanometer) and not the emf is studied, then bearing in mind that $I = E/R$, the transfer from (3.34) to the mean-square current does not present any difficulties.

Consider a specific instrument, a НФК-1 photogalvanometric self-balancing nanovoltmeter [42]. The modifications НФК-2 and НФК-3 of this instrument have the same parameters and differ from НФК-1 only by external finishing. One graduation of the instruments is equal to 4×10^{-10} V. The parameters of the instrument are as follows: $T_0 = 0.37$s, $\beta = 12$, $\beta_{\text{out}} = 1$, $\tau = 0.35$s, $\sum R = 11\Omega$, and $T_{\text{out}} = 0.8$ s. Therefore, $\omega_0 = 2\pi/T_0 = 17$ rad/s, $q = 0.8/0.37 = 2.2$, $\varkappa = 0.35 \times 17 = 6$, $C_0 = 1$, $C_1 = 24 + 2 \times 2.2 = 28.4$, $C_2 = 2 \times 12 \times 6 + 4 \times 12 \times 1 \times 2.2 + 2.2^2 + 1 = 255$, $C_3 = 6 + 4 \times 12 \times 6 \times 2.2 + 2 \times 2.2 + 2 \times 12 \times 2.2^2 = 755$, $C_4 = 2 \times 6 \times 2.2 + 2 \times 12 \times 6 \times 2.2^2 + 2.2^2 = 693$, and $C_5 = 6 \times 2.2^2 = 27.6$.

[2] These integrals were first presented in the book *Theory of Servomechanisms* by H.M. James, N.B. Nichols, and R.S. Phillips (McGraw-Hill, New York, 1947).

The cofactor in formula (3.24), consisting of the coefficients $C_0 - C_5$, is equal to 4.2×10^{-2}. At $T = 293$ K, we obtain

$$\overline{E^2} = \frac{2\pi \times 1.38 \times 10^{-23} \times 293 \times 11}{0.37} \times 4.2 \times 10^{-2} = 3.2 \times 10^{-20}\ \text{V}^2,$$

$$\sqrt{\overline{E^2}} = 1.8 \times 10^{-10}\ \text{V}.$$

The experimentally estimated mean-square voltage of the fluctuations of the nanovoltmeter is equal to 2.8×10^{-10} V, which exceeds the theoretical value by only a factor of less than 2. Therefore, it can be considered that the sensitivity of the HΦK instruments has practically reached its theoretical limit.

3.7.2. *Restrictions Imposed by Shot Noise*

An electric current, as is well known, consists of a flow of electrons. When the current becomes very weak, i.e., the number of electrons passing per second through the transverse cross section of the conductor becomes small, it is observed that its number fluctuates randomly. The random oscillation of the current strength arising in this manner is caused by the randomness of the moments at which electrons appear in the electric circuit. The noise arising is called the shot noise or Schottky noise (sometimes this noise is also called generation–recombination noise). Shot noise is characterized by the fact that it is observed only when current flows along the circuit. Thermal noise, however, does not depend on whether a current flows in the circuit; it also occurs in the absence of current.

The mean-square fluctuation produced in the current strength because of shot noise is given by the formula

$$\overline{I_{\text{sn}}^2} = 2eI\Delta f_{\text{sn}}, \tag{3.35}$$

where I is the current flowing in the circuit, e is the electron charge, and Δf_{sn} is the equivalent bandwidth of the shot noise. Formula (3.35) was derived by replacing the real spectrum of the squared current, caused by the shot noise, which has a maximum at zero frequency, by the equivalent spectrum that is uniform in the band Δf_{sn}; in addition, $\Delta f_{\text{sn}} = 1/2t_0$, where t_0 is the average transit time of an electron or, in the case of the generation–recombination noise, the average lifetime of an electron.

The fluctuations of the indications of the measuring instrument connected in a circuit with current I can be calculated based on (3.35) in precisely the same manner as was done for thermal noise. If the transmission band of the device is Δf and $\Delta f \leq \Delta f_{\text{sn}}$, then the mean-square fluctuations produced by shot noise in the indications of the instrument can be estimated from the formula $\overline{I_{\text{sn}}^2} = 2eI\Delta f$.

It is sometimes convenient to estimate the shot noise based on the variance of the number of electrons forming the current I : $\overline{\Delta n^2} = n$. Thus, a current of 1×10^{-14} A is formed by the passage of approximately 6×10^4 electrons per second through the transverse cross section of the conductor. The mean-square

fluctuations of this number of particles is equal to 6×10^4, and the relative value of the mean-square deviation will be

$$\frac{\sqrt{\overline{\Delta n^2}}}{n} = \frac{1}{\sqrt{n}} \approx \frac{1}{2.4 \times 10^2}$$

or 0.4%.

A current of 1×10^{-16} A is equal to approximately only 600 electrons per second, and the mean square of its fluctuations is already equal to 4%.

Shot and thermal noise are independent of one another. For this reason, when both types of noise occur, the variance of the fluctuations of the measuring instrument is calculated as the sum of the variances of the shot and thermal components.

3.7.3. *Estimate of the Minimum Measurable Value*

Shot and thermal noise in the input circuit of a measuring instrument are essentially indicators of the fact that the model of the object does not correspond to the object. In accordance with what was said in Section 1.3, measurement with a prescribed accuracy is possible only if the error caused by this discrepancy (threshold discrepancy) will be less than the permissible prescribed measurement error. But how does one compare these errors?

The measurement error is usually established in the form of a limit of permissible error (absolute or relative). The error caused by natural noise is usually calculated as the mean square of the fluctuations of the indications of the instrument or the mean-square deviation. These are entirely different indicators of the error. The permissible limit of the measurement error must be compared not with the mean-square deviation but with the width of the confidence interval for fluctuations of the indications. The latter width is not so easy to calculate, because the random process at the output of the instrument can hardly be regarded as being white noise. It is clear, however, that for a confidence level of the order of 0.95, the width of the confidence interval will be more than one mean-square deviation on each side of the central line. The limit of permissible error, even for measurement of the smallest values, cannot exceed 50%. The width of the confidence interval can be expressed as $k\sqrt{\overline{x^2}}$ (more accurately, this is the half-width of this band), and the limit of permissible error can be expressed as $0.5x_{\min}$. If it is assumed that these quantities are equal, then we obtain

$$x_{\min} = 2k\sqrt{\overline{x^2}}.$$

If $k = 2$, then $x_{\min} = 4\sqrt{\overline{x^2}}$.

The foregoing arguments, however, are unsatisfactory in certain respects, because not all frequency components of the noise are equivalent to the observer: High-frequency noise can be easily averaged when reading the indications of the instrument, whereas noise of very low frequency can in many cases be neglected. In addition, it was necessary to make an assumption and use the relation between

the measured quantity and the limit of permissible measurement error. Is it possible to approach this question in some other manner?

The mean-square fluctuation of the indications is an integral characteristic, which is not related either to the time or to the frequency. Meanwhile, measurements require some time, so that the temporal characteristics of these fluctuations are of interest. In the theory of random processes, the average frequency of excesses above a prescribed level is used as a temporal characteristic of the process. The physical meaning of this characteristic is clear and simple. For example, let the response time of the instrument, and therefore the measurement time, be equal to 5 s, and let one overshoot per graduation of the instrument occur on average once every 60 s. It is clear that an instrument with such noise can clearly measure a quantity corresponding to one graduation of the instrument scale. But if one overshoot occurred on the average over a time close to the response time of the instrument, then such a small quantity could no longer be measured.

The overshoots above a prescribed level provide a convenient characteristic, one that is easy to estimate experimentally. Mean-square fluctuations are difficult to estimate experimentally. But it is easier to calculate the mean-square fluctuations than the average frequency of overshoots.

We shall now calculate the average frequency of overshoots above a prescribed level for a normal stationary process. Thermal noise corresponds to these conditions. Let σ be the mean-square deviation of the random process and c the threshold levels for determining overshoot values. The average frequency of overshoots $\bar{N}$ can be calculated from the formula

$$\bar{N} = e^{-\gamma^2/2}\sqrt{\frac{1}{\sigma^2}\int_0^\infty f^2 F(f)df},$$

where $\gamma = c/\sigma$ and $F(f)$ is the spectral density of the process.

Consider once again a photogalvanometric self-balancing amplifier. The variance of the indications of the instrument is given by formula (3.34), and the spectral density is given by formula (3.33). The latter formula, however, pertains to the relative frequency η. For this reason, we shall transform the integrand in the expression so that f is replaced by η : $f^2F(f)\,df = (1/4\pi^2) \times \omega^2 F(\omega)\,d\omega = (\omega_0^2/4\pi^2) \times \eta^2 F(\eta)\,d\eta$.

Referring once again to the tables of integrals mentioned above, we find [42]

$$\bar{N} = \frac{1}{T_0}e^{-\gamma^2/2}\Phi, \quad \Phi = \sqrt{\frac{C_0(C_4C_3 - C_2C_5)}{C_0C_4C_5 + C_2C_3C_4 - C_2^2C_5 - C_1C_4^2.}}$$

For the HΦK-1 instrument, $\sqrt{\overline{E^2}} = 1.8 \times 10^{-10}$ V and $c = 4 \times 10^{-10}$ V. Hence, $\gamma = 4 \times 10^{-10}/1.8 \times 10^{-10} = 2.2$. The coefficients $C_0 - C_5$ were presented above, and knowing them, we calculate $\Phi = 0.065$. Substituting the

numerical values into the working formula, we obtain

$$\bar{N} = \frac{1}{0.37} \times 0.09 \times 0.065 = 1.6 \times 10^{-2} \text{ overshoot/s},$$

which means that one overshoot per graduation occurs on average once per minute. The response time of the instrument is 5 s. It is clear that noise does not prevent indicating, with the help of the HΦK-1 instrument, a quantity corresponding to one graduation.

4
Statistical Methods for Experimental Data Processing

4.1. Requirements for Statistical Estimations

The estimates obtained from statistical data must be consistent, unbiased, and efficient.

An estimate $\tilde{A}$ is said to be consistent if, as the number of observations increases, it approaches the true value of the estimated quantity A (it converges in probability to A):

$$\tilde{A}(x_1, \ldots, x_n) \underset{n\to\infty}{\to} A.$$

The estimate of A is said to be unbiased if its mathematical expectation is equal to the true value of the estimated quantity:

$$E[\tilde{A}] = A.$$

In the case when several unbiased estimates can be found, the estimate that has the smallest variance is, naturally, regarded as the best estimate. The smaller the variance of an estimate, the more efficient the estimates.

Methods for finding estimates of a measured quantity and indicators of the quality of the estimates depend on the form of the distribution function of the observations.

For a normal distribution of the observations, the arithmetic mean of the observations, as well as their median can be taken as an estimate of the true value of the measured quantity. The ratio of the variances of these estimates is well known [20]:

$$\sigma_{\bar{x}}^2/\sigma_m^2 = 0.64,$$

where $\sigma_{\bar{x}}^2$ is the variance of the arithmetic mean and σ_m^2 is the variance of the median.

Therefore, the arithmetic mean is a more efficient estimate of A than the median.

In the case of a uniform distribution, the arithmetic mean of the observations or the half-sum of the minimum and maximum values can be taken as an estimate of A:

$$\tilde{A}_1 = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \tilde{A}_2 = \frac{x_{\min} + x_{\max}}{2}.$$

The ratio of the variances of these estimates is also well known [20]:

$$\frac{V[\tilde{A}_1]}{V[\tilde{A}_2]} = \frac{(n+1)(n+2)}{6n}.$$

For $n = 2$, this ratio is equal to unity, and it increases for $n > 2$. Thus, for $n = 10$, it is already equal to 2.2. Therefore, the half-sum of the minimum and maximum values is, in this case, already a more efficient estimate than the arithmetic mean.

4.2. Estimation of the Parameters of the Normal Distribution

If the available data are consistent with the hypothesis that the distribution of the observations is normal, then to describe fully the distribution, the expectation $E[X] = A$ and the variance σ^2 must be estimated.

When the probability density of a random quantity is known, its parameters can be estimated by the method of maximum likelihood. We shall use this method to solve our problem.

The elementary probability of obtaining some result of an observation x_i in the interval $x_i \pm \Delta x_i/2$ is equal to $f_i(x_i, A, \sigma)\Delta x_i$. All observational results are independent. For this reason, the probability of encountering all experimentally obtained observations with $\Delta x_i = \cdots = \Delta x_n$ is equal to

$$P_l = \prod_{i=1}^{n} f_i(x_i, A, \sigma)\Delta x_1 \cdots \Delta x_n.$$

The idea of the method is to take for the estimate of the parameters of the distribution (in our case, these are the parameters A and σ), the values that maximize the probability P_l. The problem is solved, as usual, by equating to zero the partial derivatives of P_l with respect to the parameters being estimated. The constant cofactors do not affect the solution, and for this reason, only the product of the functions f_i is studied; this product is called the likelihood function:

$$L(x_1, \ldots, x_n; A, \sigma) = \prod_{i=1}^{n} f_i(x_1, \ldots, x_n; A, \sigma).$$

We now return to our problem. For the available group of observations $x_1, \ldots, x_n$, the values of the probability density will be

$$f_i(x_i, A, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i - A)^2/2\sigma^2}.$$

Therefore,

$$L = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - A)^2\right).$$

To find the maximum of L, it is convenient to investigate $\ln L$:

$$\ln L = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - A)^2.$$

The maximum of L will occur when $\partial L/\partial A = 0$ and $\partial L/\partial \sigma^2 = 0$:

$$\frac{\partial L}{L\partial A} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - A) = 0,$$

$$\frac{\partial L}{L\partial(\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - A)^2 = 0.$$

From the first equation, we find an estimate for A:

$$\tilde{A} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{4.1}$$

The second equation gives the estimate $\tilde{\sigma}^2 = (1/n)\sum_{i=1}^{n}(x_i - A)^2$. But A is unknown; taking instead of A its estimate $\bar{x}$, we obtain

$$\tilde{\sigma}_*^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

We shall check to see whether the obtained estimates are consistent and unbiased. The mathematical expectation $E(x_i) = A$, because all x_i refer to one and the same distribution. For this reason,

$$E[\tilde{A}] = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = A.$$

Therefore, $\tilde{A}$ is an unbiased estimate of A. It is also a consistent estimate, because as $n \to \infty$, $\tilde{A} \to A$, according to the law of large numbers.

We shall now investigate $\tilde{\sigma}_*^2$. In the formula derived above, the random quantities are x_i and $\bar{x}$. For this reason, we shall rewrite it as follows:

$$\begin{aligned}
\tilde{\sigma}_*^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i - A + A - \bar{x})^2 \\
&= \frac{1}{n}\sum_{i=1}[(x_i - A)^2 - 2(x_i - A)(\bar{x} - A) + (\bar{x} - A)^2] \\
&= \frac{1}{n}\sum_{i=1}^{n}(x_i - A)^2 - \frac{2}{n}\sum_{i=1}^{n}(x_i - A)(\bar{x} - A) + \frac{1}{n}\sum_{i=1}^{n}(\bar{x} - A)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}(x_i - A)^2 - (\bar{x} - A)^2,
\end{aligned}$$

because

$$\frac{1}{n}\sum_{i=1}^{n}(\bar{x}-A)^2=(\bar{x}-A)^2$$

and

$$\frac{2}{n}\sum_{i=1}^{n}(x_i-A)(\bar{x}-A)=\frac{2}{n}(\bar{x}-A)\sum_{i=1}^{n}(x_i-A)=2(\bar{x}-A)^2.$$

We shall find $E[\tilde{\sigma}_*^2]$. For this result, the following relations must be used. By definition, according to (3.8), we have $E(x_i-A)^2=\sigma^2$. Therefore,

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-A)^2\right]=\frac{1}{n}E\left[\sum_{i=1}^{n}(x_i-A)^2\right]=\sigma^2.$$

For the random quantity $\bar{x}$, we can write analogously $E(\bar{x}-A)^2=V[\bar{x}]$. We shall express $V[\bar{x}]$ for $\sigma^2=V[X]$:

$$V[\bar{x}]=V\left[\frac{1}{n}\sum_{i=1}^{n}x_i\right]=\frac{1}{n^2}\sum_{i=1}^{n}V(x_i)=\frac{1}{n}V[X]=\frac{\sigma^2}{n}.$$

Thus,

$$E[\tilde{\sigma}_*^2]=\sigma^2-\frac{\sigma^2}{n}=\frac{n-1}{n}\sigma^2.$$

Therefore, the obtained estimate $\tilde{\sigma}_*^2$ is biased. But as $n\to\infty$, $E[\tilde{\sigma}_*^2]\to\sigma^2$, and therefore, this estimate is consistent.

To correct the estimate, i.e., to make it unbiased, $\tilde{\sigma}_*^2$ must be multiplied by the correction factor $n/(n-1)$. Then we obtain

$$\tilde{\sigma}^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2. \tag{4.2}$$

This estimate is also consistent, but, as one can easily check, it is now unbiased. Some deviation from the maximum of the likelihood function is less important for us than the biasness of the estimate.

The standard deviation of the random quantity X is $\sigma=\sqrt{V[X]}$, and it is not the random quantity. Instead of σ^2 we must use the estimate of the variance from formula (4.2)—a random quantity. The extraction of a square root is a nonlinear procedure; it introduces bias into the estimate $\tilde{\sigma}$. To correct this estimate, a factor k_n, depending on n as follows, is introduced:

n	3	4	5	6	7	10
k_n	1.13	1.08	1.06	1.05	1.04	1.03

So,

$$\tilde{\sigma} = k_n \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}. \tag{4.3}$$

The following formula gives approximately the same result [29]:

$$\tilde{\sigma} = \sqrt{\frac{1}{n-1.5} \sum_{i=1}^{n} (x_i - \bar{x})^2}.$$

We have obtained estimates of the parameters of the normal distribution, but they are also random quantities: When the measurement is repeated, we obtain a different group of observations with different values of $\bar{x}$ and $\tilde{\sigma}$. The spread in these estimates can be characterized by their standard deviations $\sigma(\bar{x})$ and $\sigma(\tilde{\sigma})$. We already obtained above that $V[\bar{x}] = \sigma^2/n$. Therefore,

$$\sigma(\bar{x}) = \sqrt{V[\bar{x}]} = \frac{\sigma}{\sqrt{n}}. \tag{4.4}$$

As instead of $\sigma(\bar{x})$ we shall take $\tilde{\sigma}(\bar{x})$, we obtain the estimate $\tilde{\sigma}(\bar{x})$. Often $\tilde{\sigma}(\bar{x})$ is denoted by the symbols $S_{\bar{x}}$ or $S(\bar{x})$. Neglecting the value of k_n, we arrive at the well-known formula

$$S(\bar{x}) = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n(n-1)}}. \tag{4.5}$$

Uncertainty of the estimate given in (4.5) depends on the number of measurements n and of the confidence probability α. For example, for $n = 25$ and $\alpha = 0.80$, the uncertainty of this estimate is about 20%; for $n = 15$ and $\alpha = 0.80$, it is about 30%. The method of this computation is described in Section 4.4.

As the number of observations is rarely large, the error in the determination of the standard deviation can be significant. In any case, this error is significantly larger than the error from the biasness introduced into the estimate by extraction of the square root (it can be eliminated by the correction factor k_n). For this reason, in practice, this biasness can usually be neglected and the formula

$$S(\bar{x}) = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} \tag{4.6}$$

is used instead of formula (4.3).

4.3. Outlying Results

If in the group of measurement results, one or two differ sharply from the rest, and no slips of the pen, reading errors, and similar blunders have not been found, then

it is necessary to check whether they are extreme events that should be excluded. The problem is solved by statistical methods based on the fact that the distribution to which the group of observations under study refers can be regarded as a normal distribution. The methodology for solving the problem and the computed tables are presented in the standard reference [11].

The solution scheme is as follows. An ordered series $x_1 < x_2 < \cdots < x_n$ is constructed from the obtained results. From all x_i, we calculate $\bar{x}$ and S, and then t given by

$$t = \frac{\max |x_i - \bar{x}|}{S}, \tag{4.7}$$

where S is calculated using (4.6). Obviously, this result will either be

$$t_1 = \frac{\bar{x} - x_1}{S} \tag{4.8a}$$

or

$$t_n = \frac{x_n - \bar{x}}{S}. \tag{4.8b}$$

Table A.3 of the Appendix gives the 0.5, 1, and 5 percentage points t_q of the corresponding unilateral check of the series $x_1, \ldots, x_n$.

In performing the measurements, we cannot foresee whether x_1 or x_n will be checked. Therefore, we are interested in a bilateral check. In this case, the critical value of t_q must be taken from the column of Table A.3 in which the significance level is one half the level we adopted for checking our data.

If the value of $t_{1,n}$ that we calculated from (4.8) is greater than t_q, then the corresponding value of x_1 or x_n must be discarded: The probability of an observation giving $t > t_q$ is small; it is less than or equal to the adopted significance level.

The described procedure is necessary and widely used. But one could say that an "abnormal" observation may actually reflect some unknown feature of the subject under study. Let us consider this issue in more detail. Imagine a measurement in which one such observation occurred. What will an expert performing this measurement do? Obviously, the expert will continue the experiment collecting more observations until he or she either finds the physical reason that explains the abnormality or concludes that this was a random error. In the latter case, after checking with the methods described earlier in this section that this was indeed an outlier, the expert will discard it for the following two reasons:

1. A real measurement as a rule consists of a small number of observations, and the probability of them including more than one outlier is extremely small. Therefore, this outlier cannot be compensated with another one having the opposite sign.
2. Because the outlier deviates significantly from the rest of the results, it skews the average value of the set of data. In other words, it increases the inaccuracy of a measurement.

Thus, if there are no physicals reasons for the outlying result, it must be discarded.

EXAMPLE. Current measurements gave the following data (the current strength in mA): 10.07, 10.08, 10.10, 10.12, 10.13, 10.15, 10.16, 10.17, 10.20, and 10.40. The value 10.40 differs sharply from the other values. We shall check to see whether or not it can be discarded. We shall use the criterion presented, though we do not have the data that would allow us to assume that these observations satisfy the normal distribution:

$$\bar{x} = 10.16\,\text{mA},$$
$$S = 0.094\,\text{mA},$$
$$t = \frac{10.40 - 10.16}{0.094} = 2.55.$$

Let $q = 1\%$. In the column of Table A.3 with the significance level 0.5% for $n = 10$ we find $t_q = 2.48$. Since $2.55 > 2.48$, the observation 10.40 mA can be discarded.

4.4. Construction of Confidence Intervals

Having obtained the estimate $\tilde{A}$, it is of interest to determine by how much it can change in repeated measurements performed under the same conditions. This question is clarified by constructing the confidence interval for the true value of the measured quantity.

The *confidence interval* is the interval that includes, with a prescribed probability called the *confidence probability*, the true value of the measurand.

In our case, the confidence interval can be constructed based on the Chebyshev inequalities [20]:

$$P\{|X - A| \geq t\sigma\} \leq \frac{1}{t^2}.$$

For the random quantity $\bar{x}$, we have

$$P\left\{|\bar{x} - A| \geq \frac{t\sigma}{\sqrt{n}}\right\} \leq \frac{1}{t^2}. \tag{4.9}$$

It is not necessary to know the form of the distribution of the observations, but it is necessary to know $\sigma[X]$. However, the intervals obtained with the help of the Chebyshev inequalities are found to be too large for practice, and they are not used.

If the distribution of the observations can be regarded as normal with a known standard deviation, then the confidence interval is constructed based on the expression

$$P\left\{|\bar{x} - A| \leq z_{\frac{1-\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right\} = \alpha,$$

where $z_{\frac{1-\alpha}{2}}$ is the quantile of the normalized normal distribution, corresponding to the selected confidence probability.

We shall show how to find the value of $z_{\frac{1-\alpha}{2}}$, using the normalized Gaussian function (Table A.1 of the Appendix).

Let $\alpha = 0.95$. With this probability, the interval

$$\left[\bar{x} - z_{\frac{1-\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{1-\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

should include the true value of A. The probability that A falls outside this interval is equal to $1 - \alpha = 0.05$. As the normal distribution is symmetric, the probabilities that A exceeds the upper and lower limits of the interval are the same. Each probability is equal to $(1 - \alpha)/2 = 0.025$. Therefore, the quantile $z_{\frac{1-\alpha}{2}}$ is the quantile corresponding to the probability $p = 0.025$. It is obvious that the probability of the upper limit of this interval is $(1 - 0.025) = 0.975$. It can be calculated as

$$p = 1 - \frac{1-\alpha}{2} = \frac{1+\alpha}{2}.$$

The Gaussian function $\Phi(z)$ is related to the distribution function $F(z)$ by the relation $F(z) = 0.5 + \Phi(z)$.

Therefore, in our example, $\Phi(z) = F(z) - 0.5 = 0.975 - 0.5 = 0.475$. In Table A.1, we find the quantile $z_{0.975} = 1.96$ corresponding to the argument 0.475.

Often, on the other hand, the value of the quantile $z_{\frac{1+\alpha}{2}}$ is given and the corresponding probability α is found. For example, for $z_{\frac{1+\alpha}{2}} = 1$, $\Phi(z) = 0.3413$ and $F(z) = \Phi(z) + 0.5 = 0.841$. Then $(1 + \alpha)/2 = 1 - F(z) = 0.159$ and $\alpha = 0.682$. Analogously, for $z_{\frac{1+\alpha}{2}} = 3$, we find $\Phi(z) = 0.49865$, $F(z) = 0.99865$, $(1 + \alpha)/2 = 0.00135$, and $\alpha = 0.9973$.

In practice, however, the standard deviation is rarely known. Usually we know only its estimate S and, correspondingly, $S_{\bar{x}} = S/\sqrt{n}$. Then the confidence intervals are constructed based on Student's distribution, which is the distribution of the random quantity

$$t = \frac{\bar{x} - A}{S_{\bar{x}}}, \tag{4.10}$$

where $S_{\bar{x}}$ is the estimate of the standard deviation of the arithmetic-mean value $\bar{x}$, calculated from formula (4.5).

The confidence interval $[\bar{x} - t_q S_{\bar{x}}, \bar{x} + t_q S_{\bar{x}}]$ corresponds to the probability

$$P\{|\bar{x} - A| \leq t_q S_{\bar{x}}\} = \alpha,$$

where t_q is the q percent point of Student's distribution; the value of t_q is found from Table A.2 based on the degree of freedom $\nu = n - 1$ and the significance level $q = 1 - \alpha$. (Note: $\nu = n - 1$ since here is one unknown parameter A).

The confidence probability should not be too low, but even for a value that is very high, there are usually not enough reliable starting data available. In measurement practice, the confidence probability is increasingly often set equal to 0.95.

Existing methods make it possible to check the admissibility of the hypothesis that the observations are described by a normal distribution and therefore the hypothesis that Student's distribution is admissible (see Section 4.5). The significance

level q, used for constructing the confidence probability, has to be consistent with the significance level adopted when checking the normality of the distribution, but this problem does not yet have a definite solution.

In practice, confidence intervals are constructed based on Student's distribution, often without checking its admissibility. The fact that, in the process, as a rule, no misunderstandings arise indirectly confirms the opinion stated above that real distributions are truncated distributions that are narrower than normal distributions.

Sometimes confidence intervals are constructed for the standard deviation. So the χ^2 distribution, presented in Table A.4, is employed. The confidence interval, with the limits $(\sqrt{n-1}/\chi_L)\tilde{\sigma}$ and $(\sqrt{n-1}/\chi_U)\tilde{\sigma}$ for the probability,

$$P\left\{\frac{\sqrt{n-1}}{\chi_L}\tilde{\sigma} < \sigma < \frac{\sqrt{n-1}}{\chi_U}\tilde{\sigma}\right\} = \alpha$$

is found as follows. Table A.4 gives the probabilities $P\{\chi^2 > \chi_q^2\}$. The value of χ_U^2 is found from the table for $p_U = (1+\alpha)/2$, and the value of χ_L^2 is found for $p_L = (1-\alpha)/2$.

For example, let $\tilde{\sigma} = 1.2 \times 10^{-5}$ and $n = 10$. Take $\alpha = 0.90$. Then $p_U = (1 + 0.9)/2 = 0.95$ and $p_L = (1 - 0.9)/2 = 0.05$. The degree of freedom $\nu = 10 - 1 = 9$. From Table A.4, we find $\chi_U^2 = 3.325$ and $\chi_L^2 = 16.92$. The confidence interval will be

$$\left[\frac{\sqrt{10-1}}{\sqrt{16.92}} \times 1.2 \times 20^{-5}, \frac{\sqrt{10-1}}{\sqrt{3.325}} \times 1.2 \times 10^{-5}\right];$$

i.e., $[0.88 \times 10^{-5}, 2.0 \times 10^{-5}]$. The confidence probability in this case can be taken to be less than the confidence probability when constructing the confidence interval for the true value of the measured quantity. Often $\alpha = 0.70$ is sufficient.

We return to Chebyshev's inequality, which is attractive because it is not related to the form of the distribution function of the observations. The measured quantity can practically always be estimated by the arithmetic mean (although in the case when the distribution differs from a normal distribution, the estimate will not be the most efficient estimate), and if instead of the standard deviation its estimate is employed, then the limits of the error of this result can be estimated with the help of Chebyshev's inequality.

We shall transform the inequality (4.9) so that it would determine the probability that a deviation of the random quantity from its true value is less than $t\sigma$. The random quantity here is the arithmetic mean $\bar{x}$. After simple transformations, we obtain

$$P\left\{|\bar{x} - A| \le t\frac{\sigma}{\sqrt{n}}\right\} \ge 1 - \frac{1}{t^2}.$$

The variance of the results of measurements can be estimated with formula (4.6).

The coefficient t can be calculated based on a prescribed confidence probability α from the relation $\alpha = 1 - 1/t^2$, which gives

$$t = \frac{1}{\sqrt{1-\alpha}}.$$

If the distribution of the random errors can be assumed to be symmetric relative to A, then the confidence interval can be narrowed somewhat [20], using the inequality

$$P\left\{|\bar{x} - A| \leq t\frac{\sigma}{\sqrt{n}}\right\} \geq 1 - \frac{4}{9}\frac{1}{t^2}.$$

Now

$$t = \frac{2}{3\sqrt{1-\alpha}}.$$

Unfortunately, the confidence intervals constructed in this manner are still only approximate, because the effect of replacing the standard deviation by its estimate is not taken into account. Moreover, as we have already mentioned, the intervals obtained in the process are too wide; i.e., the uncertainty is exaggerated.

Confidence intervals should not be confused with statistical tolerance intervals.

The interval that, with prescribed probability α, contains not less than a prescribed fraction p_0 of the entire collection of values of the random quantity (population) is said to be the *statistical tolerance interval.* Thus, the statistical tolerance interval is the interval for a random quantity, and this distinguishes it in principle from the confidence interval that is constructed to cover the value of a nonrandom quantity.

If, for example, the sensitivity of a group of strain gauges is measured, then the obtained data can be used to find the interval with limits l_1 and l_2 in which, with prescribed probability α, the sensitivity of not less than the fraction p_0 of the entire batch (or the entire collection) of strain gauges of the given type will fail. This is the statistical tolerance interval. Methods for constructing this tolerance interval can be found in books on the theory of probability and mathematical statistics.

One must also guard against confusing the limits of statistical tolerance and confidence intervals with the tolerance range for the size of some parameter. The tolerance or the limits of the tolerance range are, as a rule, determined before the fabrication of a manufactured object, so that the objects for which the value of the parameter of interest falls outside the tolerance range are unacceptable and are discarded. In other words, the limits of the tolerance range are strict limits that are not associated with any probabilistic relations.

The statistical tolerance interval, however, is determined by objects that have already been manufactured, and its limits are calculated so that with a prescribed probability, the parameters of a prescribed fraction of all possible manufactured objects fall within this interval. Thus, the limits of the statistical tolerance interval, as also the limits of the confidence interval, are random quantities, and this is what distinguishes them from the tolerance limits or tolerance that are nonrandom quantities.

4.5. Methods for Testing Hypotheses About the Form of the Distribution Function of a Random Quantity

The problem is usually posed as follows: For a group of measurement results, it is hypothesized that these results can be regarded as realizations of a random quantity with a distribution function having a chosen form. Then this hypothesis is checked by the methods of mathematical statistics and is either accepted or rejected.

For a large number of observations ($n > 50$), Pearson's test (χ^2 test) for grouped observations and the Kolmogorov–Smirnov test for nongrouped observations are regarded as the best tests. These methods are described in many books devoted to the theory of probabilities and statistics. For example, see [20], [47], and [52].

We shall discuss the χ^2 test, and for definiteness, we shall check the data on the corresponding normal distribution.

The idea of this method is to monitor the deviations of the histogram of the experimental data from the histogram with the same number of intervals that is constructed based on the normal distribution. The sum of the squares of the differences of the frequencies over the intervals must not exceed the values of χ^2 for which tables were constructed as a function of the significance level of the test q and the degree of freedom $v = L - 3$, where L is the number of intervals and minus 3 is because the measurement data have two unknown parameters and Pearson's distribution has one.

The calculations are performed as follows:

(1) The arithmetic mean of the observations and an estimate of the standard deviations are calculated.
(2) Measurements are grouped according to intervals. For about 100 measurements, five to nine intervals are normally taken. For each interval, the number of measurements $\tilde{\varphi}_i$ falling within the interval is calculated.
(3) The number of measurements that corresponds to the normal distribution is calculated for each interval. For this reason, the range of data is first centered and standardized.

Let $x_{\min} = a_0$ and $x_{\max} = b_0$, and divide the range $[a_0, b_0]$ into L intervals of length $h_0 = (b_0 - a_0)/L$.

Centering and standardization are then achieved with the formula

$$x_{ic} = \frac{x_{i0} - \bar{x}}{\tilde{\sigma}}.$$

For example, the transformed limits of the range of the data for us will be as follows:

$$a_c = \frac{a_0 - \bar{x}}{\tilde{\sigma}}, \quad b_c = \frac{b_0 - \bar{x}}{\tilde{\sigma}}.$$

The length of the transformed interval $h_c = (b_c - a_c)/L$. Then we mark the limits $\{z_i\}$, $i = 0, 1, \ldots, L$, of all intervals of the transformed range $[a_c, b_c]$:

$$z_0 = a_c, \qquad z_1 = a_c + h_c, \qquad z_2 = a_c + 2h_c, \ldots, z_L = a_c + Lh_c = b_c.$$

Now we calculate the probability that a normally distributed random quantity falls within each interval:

$$p_i = \frac{1}{2\pi} \int_{z_i}^{z_i+1} e^{-x^2/2} dx.$$

After this we calculate the number of measurements that would fall within each interval if the population of measurements is normally distributed:

$$\varphi_i = p_i n.$$

(4) If less than five measurements fall within some interval, then this interval in both histograms is combined with the neighboring interval. Then the degree of freedom $v = L - 3$, where L is the total number of intervals (if the intervals are enlarged, then L is the number of intervals after the enlargement), is determined.

(5) The indicator χ^2 of the difference of frequencies is calculated:

$$\chi^2 = \sum_{i=1}^{L} \chi_i^2, \qquad \chi_i^2 = \frac{(\tilde{\varphi}_i - \varphi_i)^2}{\varphi_i}.$$

(6) The significance level of the test q is chosen. The significance level must be sufficiently small so that the probability of rejecting the correct hypothesis (committing false rejection) would be small. On the other hand, too small a value of q increases the probability adopted for the incorrect hypothesis, i.e., for committing false retention.

From the significance level q and a degree of freedom v in Table A.4, we find the limit of the critical region χ_q^2, so the $P\{\chi^2 > \chi_q^2\} = q$.

The probability that the value obtained for χ^2 exceeds χ_q^2 is equal to q and is small. For this reason, if it turns out that $\chi^2 > \chi_q^2$, then the hypothesis that the distribution is normal is rejected. If $\chi^2 < \chi_q^2$, then the hypothesis that the distribution is normal is accepted.

The smaller the value of q, the larger is the value of χ_q^2 for the same value of v, and the more easily the condition $\chi^2 < \chi_q^2$ is satisfied and the hypothesis being tested is accepted. But, in this case, the probability of committing false retention increases. For this reason, q should not be taken to be less than 0.01. For too large a value of q, as pointed out above, the probability of false rejection increases and, in addition, the sensitivity of the test decreases. For example, for $q = 0.5$ the value of χ^2 may be greater or less than χ_q^2 with equal probability, and therefore it is impossible to accept or reject the hypothesis.

In order to achieve a uniform solution of the problem at hand, it is desirable to standardize the significant levels of q adopted in metrology. To this end, for example, we can try to limit the choice of significant level to the interval $0.01 \leq q \leq 0.1$.

It should be noted that the test examined above makes it possible to check he correspondence between the empirical data and any theoretical distribution, not only

a normal distribution. This test, however, as also with, by the way, other goodness-of-fit tests does not make it possible to establish the form of the distribution of the observations; it only makes it possible to check whether the observations conform to a normal or some other previously selected distribution.

4.6. Methods for Testing Sample Homogeneity

Measurements with large random errors require careful attention. One must make sure that the obtained results are statistically under control, stable, i.e., that the measurement results cluster around the same central value and have the same variance. If the measurement method and the object of investigation have been little studied, then the measurements must be repeated until one is sure that the results are stable [25]. This process determines the duration of the investigation and the required number of measurements.

The stability of measurements is often estimated intuitively based on prolonged observations. However, mathematical methods exist that are useful for solving this problem, so-called methods for testing homogeneity.

A necessary condition is that indications of homogeneity must be present, but this is not sufficient for homogeneity in reality, because groups of measurements can be incorrectly or unsuccessfully chosen.

Figure 4.1 shows the results of measurements of some quantities, presented in the sequence in which they were obtained. Consider three groups of measurements performed in the time intervals $t_2 - t_1$, $t_3 - t_2$, and $t_4 - t_3$. They apparently will be homogeneous. Meanwhile, subsequent measurements would differ significantly from the first measurements, and on the whole, the results obtained from the first

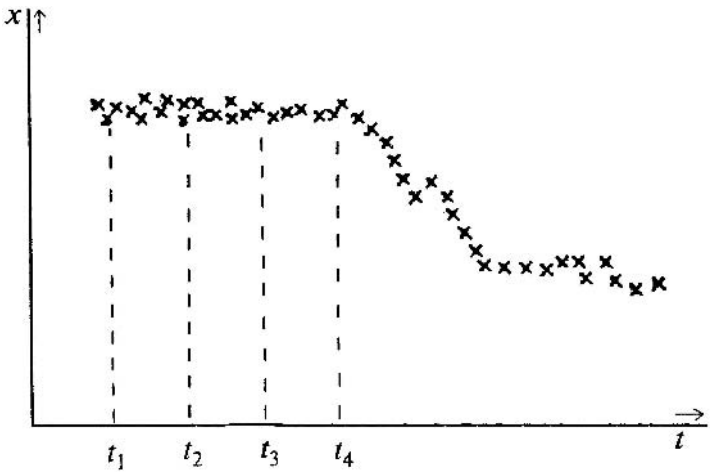


FIGURE 4.1. Example of a sequence of single-measurement results obtained in an unstable measurement.

group of measurements will give a picture of a stable, statistically under control, measurement, which is actually not the case.

The choice of groups for monitoring homogeneity remains a problem for the specialist-experimenter, just as does the problem of separating one group from another. In general, it is best to have of the order of ten measurements in a group (according to [23], from five to ten measurements), and it is better to have several such groups than two groups with a large number of measurements.

Once the groups have been reliably determined to be homogeneous, they can be combined and later regarded as one group of data.

We shall consider first the most common parametric methods for testing homogeneity. These methods are based on the normal distribution of a population. For this reason, each group of data must first be checked for normality.

The admissibility of differences between estimates of the variances is checked with the help of R. Fisher's test in the case of two groups of observations and M. Bartlett's test if there are more than two groups. We shall present both methods.

Let the unbiased estimates of the variances of these groups be S_1^2 and S_2^2, where $S_1^2 > S_2^2$. The number of observations in the groups is n_1 and n_2, so that the degrees of freedom are, respectively, $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$. We form the ratio

$$F = \frac{S_1^2}{S_2^2}.$$

Next, from Tables A.5 and A.6, where the probabilities

$$P\{F > F_q\} = q$$

for different degrees of freedom υ_1 and υ_2 are presented, we choose the value F_q.

The hypothesis is accepted, i.e., estimates of the variances can be regarded as corresponding to one and the same variance, if $F < F_q$. The significance level is equal to $2q$.

Now assume that there are L groups, and for them, unbiased estimates of the variances of groups of observations are known, $S_1^2, \ldots, S_L^2 (L > 2)$, and each group has $\nu_i = n_i - 1$ degrees of freedom; in addition, all $\nu_i > 3$. The test of the hypothesis, that the variances of the groups are equal, is based on the statistic

$$M = N \ln \left(\frac{1}{N} \sum_{i=1}^{L} \nu_i S_i^2 \right) - \sum_{i=1}^{L} \nu_i \ln S_i^2,$$

where $N = \sum_{i=1}^{L} \nu_i$.

If the hypothesis that the variances are equal is correct, then the ratio

$$\chi_1^2 = \frac{M}{1 + \dfrac{1}{3(L-1)} \left(\displaystyle\sum_{i=1}^{L} \frac{1}{\nu_i} - \frac{1}{N} \right)}$$

is distributed approximately as χ^2 with $\nu = L - 1$ degrees of freedom.

Given the significance level q, from Table A.4, we find χ_q^2, such that $P\{\chi^2 > \chi_q^2\} = q$. If the inequality $\chi_1^2 < \chi_q^2$ is satisfied, then differences between the estimates of the variances are admissible.

The admissibility of differences between the arithmetic means is also checked differently in the case of two or more groups of observations. We shall first examine the comparison of the arithmetic means for two groups of observations, when there are many observations, so that each estimate of the variances can be assumed to be equal to its variance.

We denote by $\bar{x}_1$, σ_1^2, and n_1 the data belonging to one group and by $\bar{x}_2$, σ_2^2, and n_2 the data belonging to the other group. We form the difference $\bar{x}_1 - \bar{x}_2$ and estimate its variance:

$$\sigma^2(\bar{x}_1 - \bar{x}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Next, having chosen a definite significance level q, we find $\alpha = 1 - q$, and from Table A.1, we find the argument $z_{\frac{1-\alpha}{2}}$ of the Gaussian function corresponding to the probability $\frac{1-\alpha}{2}$. A difference between the arithmetic means is regarded as acceptable, if

$$|\bar{x}_1 - \bar{x}_2| \leq z_{\frac{1-\alpha}{2}} \sigma(\bar{x}_1 - \bar{x}_2).$$

If the variances of the groups are not known, then the problem can be solved only if both groups have the same variances (the estimates of this variance $\tilde{\sigma}_1^2$ and $\tilde{\sigma}_2^2$ can, naturally, be different). In this case, t is calculated as

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{(n_1 - 1)\tilde{\sigma}_1^2 + (n_2 - 1)\tilde{\sigma}_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}.$$

Next, given the significance level q, from Table A.2 for Student's distribution with $\nu = n_1 + n_2 - 2$ degrees of freedom, we find t_q. A difference between the arithmetic means is regarded as admissible if $t < t_q$.

If the number of groups is large, the admissibility of differences between the arithmetic means is checked with the help of Fisher's test. It is first necessary to check that all groups have the same variance.

Fisher's method consists of comparing estimates of the intergroup variance S_L^2 and the average variance of the groups $\overline{S^2}$:

$$S_L^2 = \frac{1}{L-1} \sum_{i=1}^{L} n_i (\bar{x}_i - \bar{x})^2,$$

where

$$\bar{x} = \frac{\sum_{i=1}^{L} n_1 \bar{x}_i}{N}, \quad N = \sum_{i=1}^{L} n_i$$

(the estimate S_L^2 has $v_1 = L - 1$ degrees of freedom);

$$\bar{S}^2 = \frac{1}{N-L} \sum_{i=1}^{L} \sum_{j=1}^{n_i} (x_{ji} - \bar{x}_i)^2$$

(the degree of freedom $v_2 = N - L$).

Both estimates of the variances have a χ^2 distribution with v_1 and v_2 degrees of freedom, respectively. Their ratio has Fisher's distribution with the same degrees of freedom.

A spread of the arithmetic means is acceptable if $F = S_L^2/\bar{S}^2$ for the selected probability α lies within the interval from F_L to F_U:

$$P\{F_L \le F \le F_U\} = \alpha.$$

The upper limits of Fisher's distribution F_U are presented in Tables A.5 and A.6; the lower limits are found from the relation $F_L = 1/F_U$. If the significance levels in finding F_U and F_L are taken to be the same $q_1 = q_2 = q$, then the common significance level of the test will be $2q$ and

$$\alpha = 1 - 2q.$$

A method for checking the admissibility of the spread in the arithmetic means of the groups when the variances of the groups are different has also been developed, but it is more complicated.

It should be noted that a significant difference between the arithmetic means could indicate that a constant systematic error exists in the observational results of one or another group, as well as that the interesting parameter of the model used to describe the object of investigation is variable. The latter means that the postulate β is not satisfied, and therefore, measurements cannot be performed with the required accuracy.

We shall now discuss nonparametric methods for testing homogeneity. These methods do not require any assumptions about the distribution function of the population and are widely used in mathematical statistics.

Wilcoxon and Mann–Whitney tests. Assume that we have two samples: $\{x_i\}$, $i = 1, \ldots, n$, and $\{y_j\}$, $j = 1, \ldots, n_2$, and let $n_1 \le n_2$. We check the hypothesis H_0 : $F_1 = F_2$, where F_1 and F_2 are the distribution functions of the random quantities X and Y, respectively.

The sequence of steps in checking H_0 is as follows. Both samples are combined, and an ordered series is constructed from $N = n_1 + n_2$ elements; i.e., all x_i and y_j are arranged in increasing order, irrespective of the sample to which one or another value belongs. Next, all terms of the ordered series are enumerated. The order number of a term is called its rank.

When the numerical values of several elements are the same, to each of them a rank equal to the arithmetic mean of the ranks of the corresponding values is assigned.

Next the sum of the ranks of all elements of the smaller sample is calculated. The sum T obtained is then compared with the critical value T_q.

For small values of n_1 and n_2, tables $T_q(n_1, n_2)$ are given in most modern books on statistics. For $n_1, n_2 > 25$, the critical value T_q can be calculated using the normal distribution $N(m_1, \sigma^2)$:

$$T_q = m_1 + z_{1-q}\sigma,$$

where

$$m_1 = \frac{[n_1(N+1)-1]}{2}$$

and

$$\sigma^2 = \frac{[n_1 n_2(N+1)]}{12},$$

z_{1-q} is the quantile of the level $(1-q)$ of the standard normal distribution $N(0, 1)$.

The hypothesis H_0 is rejected with significance level q if $T > T_q$. This is Wilcoxon's test.

A variant of Wilcoxon's test is the Mann–Whitney test. This test is based on calculations of the so-called inversions U, which are more difficult to calculate than the ranks T. But U and T are uniquely related, and the values of U can be found from the values of T:

$$U_{(1)} = T_{(1)} - \frac{n_1(n_1+1)}{2}, \qquad U_{(2)} = n_1 n_2 + \frac{n_2(n_2+1)}{2} - T_{(2)}.$$

The parameter U is compared with the critical value U_q, for which tables or a calculation similar to the one performed above are employed.

Given q, we find from the standard normal distribution $N(0, 1)$ the quantile z_{1-q} and calculate U_q:

$$U_q = \frac{n_1 n_2 - 1}{2} + z_{1-q}\sqrt{\frac{n_1 n_2(N+1)}{12}}.$$

The hypothesis H_0 is rejected if both $U_{(1)}$ and $U_{(2)}$ are greater than U_q.

For the Siegel–Tukey test, as in the case of Wilcoxon's test, two samples x_i and y_j are studied and likewise the hypothesis $H_0 : F_1 = F_2$ is checked. The notation x_i is given to the sample whose number of terms n_1 is less than the number n_2 in the second sample. All $N = n_1 + n_2$ values of x_i and y_j are likewise arranged in increasing order, with an indication that each term belongs to the sequence X or Y. The ranks are assigned as follows: rank 1 to the first term, rank 2 to the last (Nth) term, rank 3 to the $(N-1)$st term, rank 4 to the second term, rank 5 to the third term, rank 6 to the $(N-2)$nd term, and so on. Values that are equal to one another are assigned the average rank. Next, the sums of the ranks R_1 and R_2, referring to the samples $\{x_i\}$ and $\{y_j\}$ and the standardized variable z, defined as

$$z = \frac{\left|R_1 - \dfrac{n_1(N+1)}{2}\right| - 0.5}{\sqrt{\dfrac{n_1 n_2(N+1)}{12}}},$$

TABLE 4.1. The example of rank determination.

Value of the error	Number of instruments with a given error in the sample			Wilcoxon's test		Siegel–Tukey test	
	x	y	$x+y$	Average rank of a given value of the error	Sum of ranks for a given value of the error in the sample x	Average rank of a given value of the error	Sum of ranks for a given value of the error in the sample x
−0.50	1	1	2	1.5	1.5	2.5	2.5
−0.40	3	0	3	4.0	12.0	7.3	22.0
−0.30	3	0	3	7.0	21.0	13.7	41.0
−0.25	1	0	1	9.0	9.0	17.0	17.0
−0.20	13	5	18	18.5	240.5	36.5	474.5
−0.15	2	2	4	29.5	59.0	58.5	117.0
−0.10	10	8	18	40.5	405.0	80.5	805.0
−0.05	3	2	5	52.0	156.0	103.6	310.8
0.00	15	28	43	76.0	1140.0	151.5	2272.5
0.05	5	5	10	102.5	512.5	204.5	1022.5
0.10	26	35	61	138.0	3588.0	573.5	7108.4
0.15	7	4	11	174.0	1218.0	293.5	2054.5
0.20	34	41	75	217.0	7378.0	207.5	7055.0
0.25	1	3	4	256.5	256.5	128.5	128.5
0.30	17	11	28	272.5	4632.5	96.5	1640.5
0.40	13	11	24	298.5	3880.5	44.5	578.5
0.45	1	1	2	311.5	311.5	18.5	18.5
0.50	4	2	6	315.5	1262.0	10.5	42.0
0.60	0	1	1	319.0	0.0	3.0	0.0
0.80	1	0	1	320.0	320.0	2.0	2.0

are calculated. For significance level q, the hypothesis H_0 is rejected if $z > z_{1-q}$, where z_{1-q} is a quantile of order $(1-q)$ of the standard normal distribution $N(0,1)$.

The calculations of R_1 and R_2 can be checked with the help of the relation

$$R_1 + R_2 = \frac{N(N+1)}{2}.$$

The Wilcoxon test is based on comparing the average values of two samples, whereas the Siegel–Tukey test is based on estimates of the variances. For this reason, these two tests supplement one another.

As an example of the application of these tests, Table 4.1 gives data from calculations on a check of the homogeneity of two batches of 160 ammeters for a moving-iron instrument 59 with respect to the error at marker 30 of the graduated scale [43]. The experiment is described in Section 2.6.

Following the recommendations made above, we obtain $T = 25{,}403$. Let $\alpha = 0.05$. Then $z_{0.95} = 1.96$ and

$$T_q = \frac{160 \times 321}{2} - 0.5 + 1.96\sqrt{\frac{160 \times 160 \times 321}{12}} = 27{,}620.$$

As 25,403 < 27,620, the hypothesis that the samples are homogeneous is accepted based on Wilcoxon's test.

Consider now the Siegel–Tukey test. According to the data in the table, $R_1 = 23{,}713$. Analogous calculations give $R_2 = 27{,}647$. Taking $R_1 (R_1 < R_2)$, we obtain

$$z = \frac{\left| 23{,}713 - \dfrac{160 \times 321}{2} \right| - 0.5}{\sqrt{\dfrac{160 \times 160 \times 321}{12}}} = 2.3.$$

We chose $q = 0.05$ and therefore $z_{0.95} = 1.96$. As $z > z_{0.95}$, the hypothesis that the samples are homogeneous is rejected based on the Siegel–Tukey test.

4.7. Trends in Applied Statistics and Experimental Data Processing

One problem addressed in statistics has to do with accurately modeling measurement data. A normal distribution is often used for this purpose. But sometimes it is desirable to have heavier tails than this distribution allows. Mathematics offers a convenient solution to add weight to the tails of a distribution called a *contaminated scheme*: A normal distribution is contaminated by adding another distribution, which is usually also a normal distribution. The general form of this combined distribution is:

$$F(x) = (1 - \epsilon)\Phi(x) + \epsilon H(x),$$

where $\Phi(x)$ is the basic normal distribution having variance σ_b^2, $H(x)$ is the contaminating normal distribution having variance σ_c^2, and ϵ and $(1 - \epsilon)$ are the weights of the terms above. Also, $\sigma_c^2 \gg \sigma_b^2$ and $\epsilon \ll 1$, so that the contaminating distribution mostly affects only the tails of the basic distribution.

The contaminated scheme is simple and convenient. Nonetheless, it has been mostly used in theoretical studies.

Another statistical problem is to decrease the influence of the form of the distribution function used to model the data on the measurement result. The distribution function by its nature is a mathematical concept. It is used in measurements as a theoretical model for a set of measurements. As always, a complete conformance between the model and the real set of data is impossible. Therefore, different models can be chosen for the same data. A small difference between the models may lead to significantly different estimation of the measurand. A solution to this problem was offered by so-called *robust estimations* [16, 32]. Among the earliest known robust estimations, the most popular are the truncated means, the Winsor's means, and the weighted means [32]. These methods assume that measurement results are arranged in an ordered series; i.e.,

$$x_1 \leq x_2 \leq \ldots \leq x_n.$$

- **The Truncated Means.** Given the ordered series above, the method of truncated means discards k values from the left and the right ends of this series. The number k is obtained as $k = \lfloor np \rfloor$, where $0 < p < 0.5$ and the notation $\lfloor np \rfloor$ means that k is the greatest integer number that is equal to or smaller than np. The rest of the series provides the robust estimate of the measurand by the formula

$$\tilde{A}_T = \frac{1}{n-2k} \sum_{i=k+1}^{n-k} x_i .$$

  Note that the truncating procedure is similar to the usual practice of eliminating the outlying result from the sample, which is described in Section 4.3.
- **The Winsor's Means.** Rather than discarding extreme items in the ordered series, the Winsor's method replaces them with the neighboring items. The robust estimate of the measurand is calculated by the formula:

$$\tilde{A}_W = \frac{1}{n} \left\{ \sum_{i=k+1}^{n-(k+1)} x_i + (k+1)(x_{k+1} + x_{n-k}) \right\} .$$

- **The Weighted Means.** The weighted means method obtains a robust estimate by computing a linear combination of the measurement data. There are numerous variations in this method [16, 31]. Here we present one such variation, which uses the weighted average of the median of the series and two items symmetrically located around the median in the series [32].

  Median M is determined by the formula:

$$M = \begin{cases} x_{k+1} & \text{if } n = 2k+1; \\ \frac{1}{2}(x_k + x_{k+1}) & \text{if } n = 2k. \end{cases}$$

  The robust estimate of the mean according to this method is then given by the following formula:

$$\tilde{A}_C = (1-2\epsilon)M + 2\epsilon \frac{(x_l + x_{n-l+1})}{2},$$

  where $(1 - 2\epsilon)$ and 2ϵ are the weights, $\epsilon \ll 1$, and l and $(n - l + 1)$ are the positions of the two symmetrical items chosen for the estimation.

Numerous other robust estimates were also proposed [16]. Thus, it is not clear which method to choose for a given measurement. This difficulty was addressed by Hogg as follows [31]. His method takes advantage of the natural assumption that all density distributions are symmetrical, the assumption on which all other robust estimates are based anyway. Symmetrical distributions can be characterized by one parameter—the excess e (see Section 3.4):

$$e = \frac{\mu_4}{\sigma^4} .$$

Hogg proposed to divide all distributions into several classes depending on the value of e, in such a way that for all distributions in the same class, the mean value can be calculated with the same formula. Thus, the estimate of the measurand for each class will not depend on the distribution function. The estimate of the excess

TABLE 4.2. Classes of distribution functions and formulas for estimation of their mean values after Hogg.

Distribution Class	æ	Formula for the measurand estimation
A	$æ < 2$	$\tilde{A}_a = \frac{1}{2}(x_1 + x_n)$
B	$2 < æ < 4$	$\tilde{A}_b = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
C	$4 < æ < 5.5$	$\tilde{A}_c = \frac{1}{n - 2\lfloor n/4 \rfloor} \sum_{i=\lfloor n/4 \rfloor + 1}^{n - \lfloor n/4 \rfloor} x_i$
D	$5.5 < æ$	$\tilde{A}_d = M$

e is found from the formula:

$$æ = \frac{\sum_{i=1}^{n}(x_i - \tilde{A})^4}{nS^4}.$$

The price this method pays for the robust estimate is the loss in the efficiency of the estimate. Therefore, a desired solution would find a compromise between the number of classes and the loss of the efficiency. Hogg studies the system of four classes named classes A, B, C, and D. The range of values of æ for each class and the corresponding formulas for estimating the mean value of the data are given in Table 4.2.

Hogg found that the four classes he proposed lead to loss in efficiency of no more then 20%, which is acceptable.

Another system of classes was proposed later [39]. This system contains only three classes, which are also determined by the values of æ. These classes and the corresponding formulas for the estimation of the mean are shown in Table 4.3.

As one can see, the formulas in Table 4.3 are the same as those used in the Hogg system: Class 1 uses the same formula as Class D, Class 2 as Class B, and Class 3 as Class A.

The estimations of variances of robust estimates are calculated in a common way, but constructing confidence intervals presents a difficult problem that is generally not discussed in the robust estimates literature. The simplest solution is to construct these intervals using a nonparametric method. In this method, the confidence interval is defined by the two items located symmetrically about the median in the ordered series.

TABLE 4.3. Classes of distribution functions and formulas for estimation of their average values after Mechanikov.

Distribution Class	æ	Formula for the measurand estimation
1	$4 < æ$	$\tilde{A}_{1m} = M$
2	$2.5 < æ < 4$	$\tilde{A}_{2m} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
3	$1.8 < æ < 2.5$	$\tilde{A}_{3m} = \frac{x_1 + x_n}{2}$

Let the confidence interval be $[x_l, x_r]$, where x_l and x_r are the left and right symmetrical items in the ordered series. The numbers l and r are found as follows [29]: l is $\lfloor 1/2(n+1-z_{\frac{1+\alpha}{2}}\sqrt{n}\rfloor$, and r is $\lceil 1/2(n+1-z_{\frac{1+\alpha}{2}}\sqrt{n}\rceil$,[1] where α is the confidence probability and $z_{\frac{1+\alpha}{2}}$ is the corresponding quantile of the standard normal distribution.

For example, for the ordered series size $n = 49$ and $\alpha = 0.95$, $A \approx M = x_{25}$ and $l = 19$ and $r = 31$. The confidence interval is $[x_{19}, x_{31}]$.

The inverse calculation was proposed in [39]. Here, we first choose the items in the ordered series as the confidence interval boundaries and then calculate the corresponding confidence probability according the formula:

$$P\{x_l \le A \le x_r\} = \frac{1}{2^n}\sum_{i=l}^{n-r} C_n^i.$$

In particular, for

$$k = 2, \qquad P\{x_2 < A < x_{n-1}\} = 1 - \frac{n+1}{2^{n-1}},$$

$$k = 3, \qquad P\{x_3 < A < x_{n-2}\} = 1 - \frac{n^2+n+2}{2^n}.$$

For $k > 3$, the working formulas become much more complicated. But for $k = 4$ and 5, approximate relations, presented in paper [39] mentioned above, can be used:

$$k = 4, \qquad P\{x_4 < A < x_{n-3}\} \approx 1 - \frac{0.17n^3}{2^{n-1}},$$

$$k = 5, \qquad P\{x_5 < A < x_{n-4}\} \approx 1 - \frac{0.037n^4}{2^{n-1}},$$

Nonparametric methods are widely used in statistical analysis. However, to construct confidence intervals, they require many more observations than parametrical methods. Unfortunately, it is not clear how to construct confidence intervals for robust estimates using parametric methods.

4.8. Example: Analysis of Measurement Results in Comparisons of Measures of Mass

Let us consider a measurement that determines the real value of a 1-kg mass measure by comparing it with the reference standard measure of mass with the same nominal value. Column 1 of Table 4.4 presents the measurement results obtained from this comparison.

Column 2 gives the values of $x_{i0} = (x_i - 999.998\,000) \times 10^6$, and columns 3 and 4 gives the results of auxiliary calculations.

[1] As usual, $\lfloor x \rfloor$ denotes the greatest integer equal to or smaller than x and $\lceil x \rceil$ stands for the smallest integer equal to or greater than x.

TABLE 4.4. For calculation of the statistical characteristics based on the results of comparison of measures of mass.

x_i g	$x_{i0} \times 10^{-6}$g	$x_{i0} - \bar{x}_{i0} \times 10^{-6}$g	$(x_{i0} - \bar{x}_{i0})^2 \times 10^{-12}$g^2
999.998 738	738	+17	289
999.998 699	699	−22	484
999.998 700	700	−21	441
999.998 743	743	+22	484
999.998 724	724	+3	9
999.998 737	737	+16	256
999.998 715	715	−6	36
999.998 738	738	+17	289
999.998 703	703	−18	324
999.998 713	713	−8	64
Sum	7210	0	2676

The measurement was performed by the methods of precise weighing, which eliminated the error caused by the fact that the arms of the balances were not equal. Thus, it can be assumed that there are no systematic errors.

The results of measurements in column 1 can be assumed to be discrete random independent quantities $\{x_i\}$, $i = 1, \ldots, n$ and $n = 10$, and therefore, we can apply the method of moments to find the solution. The probability of all x_i is the same and is equal to $p_i = 1/n$. To simplify the computations, column 2 presents only the varying last three digits of x_i, denoted as x_{i0}. Following the Equation 3.7 from Section 3.4, we compute the first initial moment:

$$m_1 = \sum_{i=1}^{n} x_{i0} p_i = \frac{1}{n} \sum_{i=1}^{n} x_{i0}.$$

Thus, we obtain the mean value $\bar{x}_{i0} = \frac{1}{10} \cdot 7210 \times 10^{-6} = 721 \times 10^{-6} g$, and the estimate of the value of the mass is

$$\bar{x} = 999.998000 + \bar{x}_{i0} = 999.998721 g.$$

Using Equation 3.8, the second central moment is

$$\mu_2 = \sum_{i=1}^{n} (x_{i0} - \bar{x}_0)^2 p_i.$$

After applying the known correction multiple $n/(n-1)$ (see Section 4.2), we obtain the unbiased estimate of the variance:

$$S^2(x_i) = \frac{1}{n-1} \sum_{i=1}^{n} (x_{i0} - \bar{x}_0)^2.$$

Hence, the standard deviation is

$$S(x_{i0}) = S(x_i) = \sqrt{S^2(x_{i0})} = \sqrt{\frac{2676}{9} \times 10^{-12}} = 17 \times 10^{-6} g.$$

An estimate of the standard deviation of the obtained real value of the mass measure is

$$S_{\bar{x}} = \frac{17 \times 10^{-6}}{\sqrt{10}} = 5 \times 10^{-6} g.$$

The sample shown in column 2 does not look like it corresponds to a normal distribution. But the distribution of the mean values in accordance with the central limit theorem is asymptotically normal, which allows us to find the uncertainty of the result.

Let $\alpha = 0.95$; then, using Student's distribution (Table A.2), we find the coefficient t_q. The degree of freedom $\nu = 10 - 1 = 9$ and $q = 1 - \alpha = 0.05$. Therefore, $t_{0.05} = 2.26$. In accordance with formula (4. l0), we obtain the uncertainty of measurement result:

$$u = 2.26 \times 5 \times 10^{-6} = 11 \times 10^{-6} \text{g (0.95)}.$$

Thus, with the confidence probability $\alpha = 0.95$, the mass m of the measure studied lies in the interval

$$999.998\,710 \text{ g} \leq m \leq 999.998\,732 \text{ g}.$$

The result obtained can be written more compactly as

$$m(0.95) = (999.998\,721 \pm 11 \times 10^{-6}) \text{ g}.$$

If the data above were processed by the nonparametric methods, the estimate of the measurand would be practically the same but its uncertainty would be twice as wide.

5
Direct Measurements

5.1. Relation Between Single and Multiple Measurements

The classical theory of measurement errors is constructed based on the well-developed statistical methods and pertains to multiple measurements. In practice, however, the overwhelming majority of measurements are single measurements, and however strange it may seem, for this class of measurements, there is no accepted method for estimating errors.

In searching for a solid method for estimating errors in single measurements, it is first necessary to establish the relation between single and multiple measurements.

At first glance, it seems natural to regard single measurements as a particular case of multiple measurements, when the number of measurements is equal to 1. Formally this is correct, but it does not give anything, because statistical methods do not work when $n = 1$. In addition, the question of when one measurement is sufficient remains open. In the approach examined, to answer this question—and this is the fundamental question—it is first necessary to perform a multiple measurement, and then, analyzing the results, to decide whether a single measurement was possible. But such an answer is in general meaningless: A multiple measurement has already been performed, and nothing is gained by knowing, for example, that in a given case, one measurement would have sufficed. Admittedly, it can be objected that such an analysis will make it possible not to make multiple measurements when future such measurements are performed. Indeed, that is what is done, but only when preliminary measurements are performed, i.e., in scientific investigations when some new object is studied. This is not done in practical measurements.

When it is necessary to measure, for example, the voltage of some source with a given accuracy, a voltmeter with suitable accuracy is chosen and the measurement is performed. If, however, the numbers on the voltmeter indicator dance about, then it is impossible to perform a measurement with the prescribed accuracy, and the measurement problem must be reexamined rather than performing a multiple measurement.

For practical applications, we can state the opinion that single measurements are well founded by experience, concentrated in the construction of the corresponding measuring instruments, and that measuring instruments are manufactured so that single measurements could be performed.

From the foregoing assertion a completely different point of view follows regarding the relationship between single and multiple measurements, namely, that single measurements are the primary, basic form of measurement, whereas multiple measurements are derived from single measurements. Multiple measurements are performed when necessary, based on the formulation of the measuring problem. It is interesting that these problems are known beforehand; they can even be enumerated. Namely, multiple measurements are performed as follows:

(a) When investigating a new phenomenon or a new object and relationships between the quantities characterizing the object, as well as their connection with other physical quantities, are being determined, or briefly, when preliminary measurements, according to the classification given in Chapter 1, are performed.
(b) When measuring the average value of some parameter, corresponding to the goal of the measuring problem.
(c) When the effect of random errors of measuring instruments must be reduced.
(d) In measurements for which measuring instruments have not yet been developed.

Of the four cases presented above, the first is typical for investigations in science and the third is typical for calibration practice.

There is another point of view, namely, that any measurement must be a multiple measurement, because otherwise it is impossible to judge the measurement process and its stability and to estimate its inaccuracy.

We cannot agree with this opinion. First, it contradicts practice. Second, it also does not withstand fundamental analysis. Imagine that one and the same constant quantity is measured simultaneously with the help of multiple and single methods of measurement. In both cases, the measurements are performed with the same instrument whose response time is t_r.

In Fig. 5.1(a), the dots represent the results of single measurements comprising a multiple measurement, and, Fig. 5.1(b) shows a photorecording of the indications of the analog instrument in a single measurement.

A single measurement makes it possible to obtain the value of the measurand immediately after the response time t_r of the measuring instrument. If it is desirable to check the stability of the measurement, then the observation must be continued. The process of measurement is stable if the readings of the instrument over a chosen time ΔT do not change appreciably.

The reading of the instrument gives the estimation $\tilde{A}$ of a measurand A. Of course, $\tilde{A} \neq A$. Methods for calculating errors and uncertainty of the results of single measurements are given later in this chapter. Thus, in this case, a single measurement is sufficient to obtain the measurement result and to estimate its

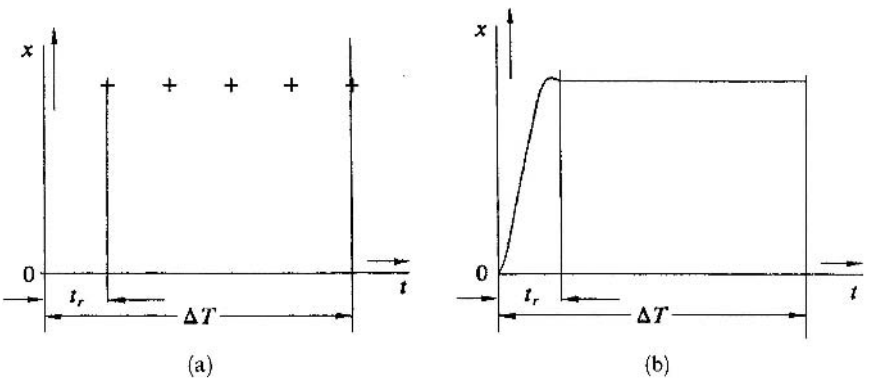


FIGURE 5.1. Results of measurements in the case of (a) a multiple measurement and in (b) a single measurement (a motion picture film of the indications of the measuring instrument).

inaccuracy. As to the stability of the measurement process, a single measurement allows one to make a better judgment than a multiple measurement because the latter represents only separate moments of the process, whereas the former gives the whole continuous picture.

The foregoing example does not say that a single measurement is better than a multiple measurement. It says only that a multiple measurement should not be performed when a single measurement can be performed. But when a multiple measurement is necessary, a single measurement is useless, and in this case and in this sense, a multiple measurement is better than a single measurement.

So, single measurements must be regarded as independent and the basic form of measurement. Correspondingly, the problem of developing methods for estimating the errors of single measurements must be regarded as an independent and important problem of the theory of measurements.

This is a good point at which to discuss another aspect of the question at hand. In many fields of measurements, modern measuring digital instruments can operate so fast that over the time allotted for measurements, say, 1 s, hundreds of measurements can be performed. Carrying out these measurements and averaging their results, we employ all of the time allotted for measurement, and, thanks to this, we reduce correspondingly the effect of interference and noise.

Consider now an analog instrument having the same accuracy as a fast measuring device, but with a response time equal to the time allotted to the measurement, i.e., in our case, 1s. From the time constant of the instrument, the effect of interference and noise will be suppressed to the same degree as for discrete averaging in the first case; i.e., we shall obtain the same result.

In other words, the measurement time is of fundamental importance, and there is no significance in how the interference and noise are filtered—in the discrete or analog form—over this time. In practice, discrete averaging is often more convenient, because in this case, the averaging time can be easily changed.

5.2. Identification and Elimination of Systematic Errors

Taking into account and eliminating systematic errors is an important problem in every accurate measurement. In the theory of errors, however, little attention has been devoted to systematic errors.

In most books on methods of data processing, the question of systematic errors is either neglected or it is stipulated that these errors are assumed to have been eliminated.

In reality, however, systematic errors cannot be completely eliminated; some unexcluded residuals always remain. These residuals must be taken into account to estimate the limits of the unexcluded systematic error of the result, which determine its systematic error.

In addition, many measurements are performed without special measures taken to eliminate systematic errors, because either it is known a priori that they are small or the measurement conditions make it impossible to eliminate them. For example, suppose the mass of a body is being measured and corrections are not made for the balances employed either because the corrections are small or because the actual values of the masses are not known (only the limits of their errors are known).

Sometimes the unexcluded residuals of the systematic errors are assumed to be random errors based on the fact that their values are unknown. We cannot agree with this point of view. When classifying errors as systematic or random, attention should be focused on their properties rather than on whether their values are known.

For example, suppose that the resistance of a resistor is being measured and a correction is made for the influence of the temperature. The systematic error would be eliminated if we knew exactly the temperature coefficient of the resistor and the temperature. We know both quantities with limited accuracy, and for this reason, we cannot completely eliminate this error. An unexcluded residual of the error will remain. It can be small or large; this we can and should estimate, but its real value remains unknown. Nonetheless, this residual error has a definite value, which remains the same when the measurement is repeated under the same conditions, and for this reason, it is a systematic error.

Errors that have been eliminated are no longer errors. For this reason, as we have already mentioned, the systematic error in a measurement also should be understood to be the unexcluded residual of the systematic error, if it cannot be neglected.

The error in a measurement can be both systematic and random, but after the measurement has already been performed, the measurement error becomes a systematic error. Indeed, the result of a measurement has a definite numerical value, and its difference from the true value of the measured quantity is also constant. Even if the entire error in a measurement was random, for a measurement result, it becomes systematic; i.e., it seemingly freezes.

We shall now discuss the classification of systematic errors. We shall base our discussion on the work of M.F. Malikov, and following this work, we shall distinguish systematic errors according to their sources and properties [35].

The sources of systematic errors can be three components of the measurement: the method of measurement, the measuring instrument, and the experimenter. Correspondingly, methodological, instrumental, and personal systematic errors are customarily distinguished.

Methodological errors arise from imperfections of the method of measurement and from the limited accuracy of the formulas used to describe the phenomena on which the measurement is based. We shall classify as methodological errors the errors arising as a result of the influence of the measuring instrument on the object whose property is being measured.

For example, the moving-coil voltmeter draws current from the measurement circuit. Because of the voltage drop on the internal resistance of the source of the voltage being measured, the voltage on the terminals of the voltmeter will be less than the measured value. The indications of the voltmeter, however, are proportional to the voltage on its terminals. The error that arises—a methodological error—should be insignificant if the measurement is performed properly.

A methodological error can also arise in connection with the use of the measuring instrument. The gain of a voltage amplifier is determined by measuring the voltages at the input and the output. If these voltages are measured successively using the same voltmeter, as is often done in practice, then, aside from the voltmeter error, the measurement error will include the error from the uncontrollable change in voltage at the amplifier input over the time during which the voltmeter is switched on and the voltage at the output is measured (or vice versa). This error does not arise when two voltmeters are employed. When the measurement is performed using one voltmeter, however, the effect of the voltmeter error decreases.

We note that the error from the threshold discrepancy between the model and the object is also a methodological error.

Instrumental systematic errors are errors caused by imperfections of the measuring instrument. Classic examples of such errors are errors of a measuring instrument that are caused by imprecise calibration of the instrument scale and the error of a resistive voltage divider from the inaccurate adjustment of the resistances of its resistors.

Another group of such errors is additional and dynamic errors. These errors also depend on the imperfections of the measuring instruments, but they are caused by influence quantities and noninformative parameters of the input signal as well as by the change in the input signal in time. Most often the additional and dynamic errors are systematic errors. When the influence quantities and the forms of the input signal are unstable, however, they can become random errors.

Setup errors, i.e., errors arising from the arrangement of the measuring instruments and their effect on one another, are also instrumental errors.

Personal systematic errors are systematic errors connected with the individual characteristics of the observer.

We shall discuss the errors in the reading of the indications of indicating instruments. Such errors were investigated by H. Bäkström [17]. Although real reading devices were simulated in this work by blanks with lines, depicting the edge of a

scale graduation and the indicator of the instrument, drawn on them, the results obtained are plausible.

The results of the investigation consist of the following.

The systematic errors made by every observer when estimating tenths of the graduation of an instrument scale can reach 0.1 graduations and are much larger than the random errors. These systematic errors are manifested by the fact that for different tenths of a graduation, different observers characteristically make estimates with different frequencies, and in addition, the distribution characteristic of every observer remains constant for a long period of time. Thus, one observer refers, more often than one would think, indications to the lines forming the edges of graduation, and to the value 0.5 of a graduation. Another observer refers indications to the values 0.4 and 0.6 of a graduation. A third observer prefers 0.2 and 0.8 graduations, and so on. Tenths of a graduation, which are arranged symmetrically in the space between scale markers, are estimated with the same frequency.

The error in estimation of tenths of graduations depends on the thickness of the markers—the lines forming the scale. The optimal thickness of these markers is 0.1 of the length of a graduation. The length of a graduation significantly affects the error in reading tenths of a graduation. Instrument scales for which tenths of a graduation can be read are usually made so that the length of a graduation is equal to about 1 mm (not less than 0.7 mm and not more than 1.2 mm).

On the whole, for a random observer, the distribution of systematic errors in the readings of tenths of a graduation can be assumed to be uniform with limits of ± 0.1 graduations.

It is interesting that the components of the random error are usually not singled out, because the random error in a measurement is, as a rule, estimated from the experimental data and the entire error is measured at once, whereas the systematic error is measured by components.

Constant systematic errors are distinguished from regularly varying systematic errors. The latter errors, in turn, are subdivided into progressing and periodic errors and errors that vary according to a complicated law.

A constant systematic error is an error that remains constant and for this reason is repeated in each observation or measurement. For example, such an error will be present in measurements performed using the same material measures that have a systematic error: balances, resistors, and so on. The personal errors made by experienced experimenters can also be classified as constant (for inexperienced experimenters, they are usually of a random character).

Progressing errors are errors that increase or decrease throughout the measurement time. Such errors are caused, for example, by the change in the working current of a potentiometer from the voltage drop of the storage battery powering it.

Periodic errors are errors that vary with a definite period.

In the general case, a systematic error can vary according to a complicated aperiodic law.

The discovery of systematic errors is a complicated problem. It is especially difficult to discover a constant systematic error. To solve the problem, in this case,

several measurements (at least two) should be performed by fundamentally different methods. This method is ultimately decisive. It is often realized by comparing the results of measurements of one and the same quantity that were obtained by different experimenters in different laboratories.

It is easier to discover variable systematic errors, which can be done with the help of statistical methods, correlation, and regression analysis. But nonmathematical possibilities also should not be avoided. Thus, in the process of performing a measurement, it is helpful to employ a graph on which the results of the measurements are plotted in the sequence in which they were obtained. The overall arrangement of the points obtained makes it possible to discover the presence of a systematic change in the results of observations without mathematical analysis. The human capability of perceiving such regularities is widely employed in metrology, although this capability has apparently still not been thoroughly studied.

If a regular change in observational results has been found and it is known that the measured quantity did not change in the process, then this indicates the presence of a regularly varying systematic error.

It is also helpful to measure the same quantity using two different instruments (methods) or to measure periodically a known quantity instead of the unknown quantity.

If the presence of a systematic error has been discovered, then it can usually be estimated and eliminated. In precise measurements, however, this often presents great difficulties and is not always possible.

In most fields of measurements, the most important sources of systematic errors are known and measurement methods have been developed that eliminate the appearance of such errors or prevent them from affecting the result of a measurement. In other words, systematic errors are eliminated not by mathematical analysis of experimental data but rather by the use of appropriate measurement methods. The analysis of measurement methods and the systematization and generalization of measurement methods are important problems, but they fall outside the scope of this book, which is devoted to the problem of analysis of experimental data. For this reason, we shall confine our attention to a brief review of the most widely disseminated general methods for studying such problems.

Elimination of Constant Systematic Errors

Method of replacement. This method gives the most complete solution of the problem. It is a version of the method of comparison, when the comparison is made by replacing the measured quantity by a known quantity and in a manner so that in the process no changes occur in the state and operation of all measuring instruments is employed.

Consider, for example, weighing performed by Borda's method. The method is designed to eliminate the error from the inequality of the arms of the balances. Let x be the measured mass, P the mass of the balancing weights, and l_1 and l_2 the lengths of the arms of the balances. The measurement is performed as follows. First the weighed body is placed in one pan of the balances and the balances are

balanced with the help of a weight with mass T. Then

$$x = \frac{l_2}{l_1} T.$$

Next, the mass x is removed and a mass P that once again balances the pans is placed in the empty pan:

$$P = \frac{l_2}{l_1} T.$$

As the right-hand sides of both equations are the same, the left sides are also equal to one another, i.e., $x = P$, and the fact that $l_1 \neq l_2$ has no effect on the result.

The resistance of a resistor can be measured in an analogous manner with the help of a sensitive but inaccurate bridge and an accurate magazine of resistances. Several other quantities can be measured analogously.

Method of contraposition. This measurement method is a version of the comparison method. The measurement is performed with two observations, and it is performed so that the reason for the constant error would affect the results of observations differently but in a known, regular fashion.

An example of this method is Gauss's method of weighing. First the weight being weighed is balanced by balance weights P_1. Using the notation of the preceding example, we have

$$x = \frac{l_2}{l_1} P_1.$$

Next the unknown weight is placed into the pan that held before the balancing weights, and the two loads are once again balanced. Now we have

$$x = \frac{l_1}{l_2} P_2.$$

We now eliminate the ratio l_2/l_1 from these two equalities and find

$$x = \sqrt{P_1 P_2}.$$

The sign method of error compensation. This method involves two measurements performed so that the constant systematic error would appear with different signs in each measurement.

For example, consider the measurement of an emf x with the help of a dc potentiometer that has a parasitic thermo-emf. One measurement gives E_1. Next the polarity of the measured emf is reversed, the direction of the working current in the potentiometer is changed, and once again the measured emf is balanced. This, process gives E_2. If the thermo-emf gives the error ϑ and $E_1 = x + \vartheta$, then $E_2 = x - \vartheta$. From here

$$x = \frac{E_1 + E_2}{2}.$$

Elimination of progressing systematic errors. The simplest, but most frequent case of a progressing error is an error that varies linearly, for example, in proportion to the time.

An example of such an error is the error in the measurement of voltage with the help of a potentiometer, if the voltage of the storage battery, generating the working current, drops appreciably. Formally, if it is known that the working current of the potentiometer varies linearly in time, then to eliminate the error arising, it is sufficient to perform two observations at times after the working current along the standard cell is regulated. Let

$$E_1 = x + Kt_1, \qquad E_2 = x + Kt_2,$$

where t_1 and t_2 are the time intervals between regulation of the working current and the observations, K is the coefficient of proportionality between the measurement error and the time, and E_1 and E_2 are the results of the observations. From here

$$x = \frac{E_1 t_2 - E_2 t_1}{t_2 - t_1}.$$

For accurate measurements, however, it is best to use a somewhat more complicated method of symmetric observations. In this method, several observations are performed equally separated in time and then the arithmetic means of the symmetric observations are calculated. Theoretically, these averages must be equal, and this makes it possible to control the course of the experiment and to eliminate these errors.

When the errors vary according to more complicated laws, the methods for eliminating the errors become more complicated, but the problem can always be solved if these laws are known. If, however, the law is so complicated that it is pointless or impossible to find it, then the systematic errors can be reduced to random or quasirandom errors, which requires a series of observations, arranged in a manner so that the observational errors would be as diverse as possible and look like random errors. However, this technique is not as effective as finding the error and eliminating it directly.

The methods listed above do not exhaust all possibilities for eliminating systematic errors. Thus, to eliminate the systematic error of a measuring instrument from the result of a measurement, the measurement can be performed not by one but rather by several instruments simultaneously (if the errors of the instruments are uncorrelated). Taking for the result of the measurement a definite combination of indications of all instruments, we can make their systematic errors, which are different for the different instruments, compensate one another somehow, and the error of the result obtained will be less than for an individual instrument. In this case, the systematic errors of the instruments can be regarded as a realization of a random quantity.

In those cases when for the measured quantity several exact relations between it and other quantities are known, these relations can be used to reduce the measurement error. For example, if the angles of a plane triangle are measured, then the fact that their sum is equal to 180° must be taken into account.

5.3. Estimation of Elementary Errors

It is difficult to describe in a generalized form a method for estimating elementary errors, because these errors are by their nature extremely diverse. The general rules for solving this problem can nonetheless be formulated.

To estimate measurement errors, it is first necessary to determine their possible sources. If it is known that some corrections will be introduced (or corrections have been introduced), then the errors in determining the corrections must be included among the elementary errors.

All elementary measurement errors must be estimated in the same manner, i.e., in the form of either absolute or relative errors. Relative errors are usually more convenient for a posteriori error estimation, and absolute errors are more convenient for a priori error estimation. However, the tradition of each field of measurement should be kept in mind. Thus, for lineal–angular measurement, absolute errors are preferred, whereas for measurements of electromagnetic quantities, relative errors are preferred.

An unavoidable elementary error in any measurement is the intrinsic error of the measuring instrument. If the limits of this error are given in the form of absolute or relative errors, then conversions are not required and these limits are the limits of the given elementary error. But often the limits of intrinsic error of a measuring instrument are given in the form of a fiducial error, i.e., as a percentage of the fiducial value. The conversion into relative error is made using the formula

$$\delta_{\text{in}} = \gamma \frac{x_N}{x},$$

where δ_{in} is the limit of the intrinsic error in relative form, γ is the limit of the fiducial error, x_N is the fiducial value, and x is the value of the measurand. Conversion into the form of absolute errors is done according to the formula

$$\Delta_{\text{in}} = \delta_{\text{in}} x = \gamma x_N.$$

Often the environmental conditions, characterized by the temperature, pressure, humidity, vibrations, and so on, affect the result of a measurement. Each influence quantity, in principle, engenders its elementary error. To estimate it, it is first necessary to estimate the possible value of the corresponding influence quantity and then compare it with the limits of the range of values of this quantity concerning the reference condition. If the influence quantity falls outside the limits of reference values, then it engenders a corresponding additional error; this error is also an elementary error.

Consider an error of the temperature. Let the temperature of the medium exceed its reference values by ΔT. If $T_1 \le \Delta T \le T_2$ and the limit of the additional error for the interval $[T_1, T_2]$ has the same absolute value, then this limit is the limit of the given additional error. If, however, for this interval, the limiting value of the temperature coefficient is given, then the limits of temperature error are calculated according to the formula

$$\delta_T = \pm w_T \Delta T,$$

where δ_T is the limit of additional temperature error and w_T is the limiting value of the absolute value of the temperature coefficient of the instrument.

In the general case, the dependence of the limit of additional error δ_i or Δ_i on the deviations of the influence quantity outside the limits of its reference values can be given in the form of a graph or expressed analytically (see Chapter 2). If the influence function of some influence quantity is indicated in the specifications provided by the manufacturer of the instrument, then a deviation of this quantity outside the limits of its reference values can be taken into account by the corresponding correction. In the process, the elementary error decreases significantly, even if the influence function is given with a large margin of error.

Suppose, for example, instead of the limiting value of the temperature coefficient $w_T = \pm b/T$ the influence function $w'_T = (1 \pm \varepsilon)b/T$ is given. To calculate the correction, one must know the specific value of the temperature during the measurement and therefore its deviation ΔT from the reference value. Then the additional temperature error will be

$$\zeta_T = (1 \pm \varepsilon)b\frac{\Delta T}{T}.$$

We eliminate the constant part of this error with the help of the correction

$$c = -\frac{b\Delta T}{T}.$$

There then remains the temperature error

$$\delta'_T = \pm\varepsilon\frac{b\Delta T}{T}.$$

If the influence function is given comparatively inaccurately, for example, $\varepsilon = 0.2$ (20%), then the temperature error even in this case decreases by a factor of 4–6:

$$\frac{\delta_T}{\delta'_T} = \frac{1 \pm 0.2}{0.2} = 4 \text{ or } 6.$$

It should also be kept in mind that if the influence quantity is estimated with an appreciable error, then this error must also be taken into account when calculating the corresponding additional error.

In many cases, the input signal in a measurement is a function of time and therefore the measurement result may have a dynamic error.

Several peculiarities in estimating dynamic errors exist. The most important of these peculiarities must be discussed.

First, it should be noted that although for a long time now the dynamic errors have been taken into account in particular situations, the general theory of estimation of dynamic errors of measurements, as the theory of dynamic measurements in general, is still in the formative stage. In [49], an attempt is made to formulate the basic concepts of the theory of dynamic measurements. In studying methods for estimating dynamic errors below, we shall adhere to the concepts presented in [49].

Next, we have to mention that measuring instruments do not have dynamic errors but they may have additional errors in the dynamic regime. These errors are a special type of elementary errors of measurements.

A typical example of a measurement for which the dynamic error is significant is a measurement in which a time-varying signal is recorded. In this case, in accordance with the general definition of absolute error, the dynamic error can be written as

$$\zeta_d(t) = \frac{y(t)}{K} - x(t), \tag{5.1}$$

where $\zeta_d(t)$ is the dynamic error; $x(t)$ and $y(t)$ are the signals at the input and output, respectively, of the measuring instrument; and K is the transduction constant.

The relation between the signals at the input and the output of the measuring instrument can be represented by an operator equation

$$y = Bx, \tag{5.2}$$

where B is the operator of the recording instrument.

The operator expresses in general form the entire aggregate of dynamic properties of a measuring instrument. These properties depend on the particular action with respect to which they are studied. Thus, the dynamic properties with respect to a variable influence quantity or interference that does not act at the input of the measuring instrument can be different from the dynamic properties with respect to the input signal. In (5.2), the operator B pertains to the input signal.

When measuring instruments are constructed, the transduction constant is usually made to be independent of the strength of the input. Then the measuring instruments can be described by a linear model, and as a rule, linear models can have lumped parameters.

Substituting Eq. (5.2), Eq. (5.1) can be represented in the operator form

$$\zeta_d = \left(\frac{B}{K} - I\right)x,$$

where I is the identity operator, $Ix \equiv x$.

The input and output signals vary in time, and therefore, the dynamic error is a function of time.

One would think that given the output signal $y(t)$ and the operator of the measuring instrument, it is possible using (5.2) to find the input signal $x(t)$ and then, using formula (5.1), to find the dynamic error. This is difficult to do, however, because the operator of the measuring instrument is usually not known accurately enough.

Sometimes the problem can be solved without knowing the operator of the instrument at all. Suppose that we have an instrument and a record of some process realized with its help. We now disconnect the instrument from the process being studied and connect to the instrument a circuit in which we can control an analogous process. An example of such a device is a standard-signal generator (if an electric measuring instrument is employed). Next we record a signal at the input of the

instrument such that at the output we obtain a process that is identical to the initially recorded process. When the records are identical enough, the input processes are also identical. Therefore, a we have found the input signal, and comparison of it (after multiplication by K in accordance with (5.1)) with the given record can produce an estimate of the dynamic error.

In measurement practice, efforts are always made to use measuring instruments whose output signals would be close to conform to the goals of the measurement) in form to the input signal. But in those cases when such measuring instruments are not available and existing measuring instruments must be used, despite the distortions created by them, the reconstruction of the form (keeping the parameters unchanged) of the input signal becomes an important method for increasing measurement accuracy. It should be noted, however, that reconstruction of the form of the input signal presents great difficulties, which are associated with the fact that this problem is a so-called improperly posed problem (in the terminology of J. Hadamard); i.e., in this problen, the solution does not depend continuously on the initial data, which means that when there are small errors in the specification of the dynamic characteristic of the measuring instrument and the reading of the values of the output signal, the error in determining the input signal can be so large that the solution obtained is physically meaningless.

Physically the main idea of the problem of improper formulation in application to reconstruction of the form of the input signal consists of the following. Ultimately the spectral composition of the output signal of a measuring instrument always decreases in intensity as the frequency increases. The amplitude-frequency response of a measuring instrument (which, naturally, is a stable system) at high frequencies also approaches the frequency axis. Thus, it is required to find, based on two functions with decreasing spectra, a tnird function (the input signal) that provides a unique relation between them. At low and medium (for the given functions) frequencies, where the intensity of the spectra is high, the signal sought can be determined reliably, and the unavoidable errors in the initial data and the computational procedures operate in the normal fashion; i.e., they distort the solution without destroying its physical meaning. At high frequencies, the intensities of the spectra drop to such an extent that their effect on the solution is comparable with that of errors in the initial data. The effect of these errors can be so large that the true solution is completely suppressed. Time-domain distortions usually have the form of rapidly oscillating functions, whose amplitude is often several orders of magnitude greater than the true solution.

Methods for solving improperly posed problems (methods of regularization) are under active development in mathematics, mathematical physics, geophysics, and the theory of automatic control. A list of the most important publications on this question relevant to metrology is given in [28].

The essential feature of methods of regularization consists of filtering out the distortions based on a priori information about the true solution. The main question is how to establish the optimal degree of filtering to filter out noise without distorting the true solution. Different methods of regularization require different volumes and forms of a priori information.

Dynamic errors are most often estimated when the choice of a recording instrument is being made. The problem is solved in the following way.

The worst form of input process is chosen and expressed analytically. One complete dynamic characteristic of the recording instrument is assumed to be known. Then it is possible to compute the corresponding output process. A superposition of these output and input processes gives the dynamic error of the expected measurement. If the absolute value of dynamic error lies within the permissible limits, then this recording instrument can be used for the measurement.

But it is inconvenient to work with an error as a function of time. For this reason, efforts are usually made to describe the dynamic error, when recording data, by a parameter that assumes a single value for the entire function. Most often, the error having the maximum absolute value or its standard deviation is used.

It should be noted that the computational scheme presented above can be modified for different measurement problems. Thus, a shift of the output signal in time relative to the input signal is often possible. In this case, the signals can be artificially arranged to minimize the error.

Despite the difficulty of estimating dynamic errors, the dynamic error is an elementary error. In those cases when the dynamic error is represented by its components, these components are regarded as elementary measurement errors.

We shall use the symbols presented in Table 1.1 to designate elementary errors. If an elementary error has both systematic and random components, we shall designate it with the symbol used for the dominant component.

5.4. Method for Calculating the Errors and Uncertainties of Single Measurements

Once the errors of a single measurement have been analyzed, we have an estimate of the limits of all elementary errors of the measurement. We now proceed to the problem of synthesis. First, we single out absolutely constant errors, if they exist, and write out estimates of their limits H_f:

$$|\eta_f| \le H_f \qquad \text{or} \qquad H_{fl} \le \eta_f \le H_{fr},$$

where $f = 1, \ldots, k$ and k is rarely greater than 2.

The remaining elementary errors are conditionally constant:

$$|\vartheta_i| \le \theta_i, \qquad i = 1, \ldots, n.$$

Above we modeled conditionally constant errors by a random quantity with a uniform probability distribution. For direct measurements, in the overwhelming majority of cases, elementary errors can be assumed to be independent of one another. Starting from this fact, we calculate the limiting value of the sum of all

conditionally constant errors. Thus, we use formula (3.13):

$$\theta_\alpha = k \sqrt{\sum_{i=1}^{n} \theta_i^2}. \tag{5.3}$$

With a confidence probability $\alpha = 0.99$ and $n \leq 4$, it could turn out that $\theta_\alpha > \sum_{i=1}^{n} \theta_i$. But it is obvious that this result cannot happen. In this case, it is possible to take

$$\theta = \sum_{i=1}^{n} \theta_i.$$

It would, of course, be more correct to take a more accurate value of the coefficient k from the curves presented in Fig. 3.3.

There arises, however, the question of how well founded the choice $\alpha = 0.99$ is. In most cases, this limit does not correspond to the reliability of the initial data, and the limit $\alpha = 0.95$ is more appropriate. For $\alpha = 0.95$, formula (5.3) assumes the form (see tables on pages 72 and 73)

$$\theta_{0.95} = 1.1 \sqrt{\sum_{i=1}^{n} \theta_i^2}. \tag{5.4}$$

In this case, $\theta_\alpha < \sum_{i=1}^{n} \theta_i$. We shall show this result by investigating the last inequality. First, let $n = 2$ with $\theta_1 \leq \theta_2$ and consider the inequality $1.1 \sqrt{\theta_1^2 + \theta_2^2} < (\theta_1 + \theta_2)$ It is not difficult to verify that the inequality holds if $\theta_1/\theta_2 > 0.11$. This condition corresponds to practice, because an elementary error that is about ten times smaller than any other elementary error can be neglected.

Consider now the three terms $\theta_3 > \theta_2 > \theta_1$. Introducing $T = \theta_3 + \theta_2$, we obtain the identity

$$1.1\sqrt{T^2 + \theta_1^2 - 2\theta_3\theta_2} < (T + \theta_1).$$

The term $2\theta_3\theta_2 > 0$, and conditions of the inequality, will be stronger if this term is dropped. Then, corresponding to the case we have just studied with two terms, we obtain

$$\frac{\theta_1}{\theta_2 + \theta_3} > 0.11.$$

It is obvious that this inequality holds easier than for two components. On the whole, as the number of terms increases, the inequality is more easily satisfied.

It could happen that m of the n conditionally constant errors have un-symmetric limits:

$$\theta_{jl} \leq \vartheta_j \leq \theta_{jr}, \qquad j = 1, \ldots, n,$$

where θ_{jl} is the left-hand limit and θ_{jr} is the right-hand limit.

For calculations, unsymmetric limits must be represented as symmetric limits with a shift by a_j, where

$$a_j = \frac{\theta_{jl} + \theta_{jr}}{2}.$$

The limits of an interval that is symmetric relative to a_j are calculated according to the formula

$$\theta_j = \frac{\theta_{jr} - \theta_{jl}}{2}.$$

Next, the limits of the error must be calculated from the following formulas instead of (5.3):

$$\begin{aligned} \theta_{r,\alpha} &= \sum_{j=1}^{m} a_j + k\sqrt{\sum_{i=1}^{n-m} \theta_i^2 + \sum_{j=1}^{m} \theta_j^2}, \\ \theta_{l,\alpha} &= \sum_{j=1}^{m} a_j + k\sqrt{\sum_{i=1}^{n-m} \theta_i^a + \sum_{j=1}^{m} \theta_j^2}. \end{aligned} \tag{5.5}$$

The absolutely constant elementary errors must now be taken into account. As the probabilistic model is not appropriate for them, their limits must be summed arithmetically with the limits, calculating according to (5.6), of the sum of the conditionally constant components θ_r and θ_l or θ:

$$\begin{aligned} \Delta_{r,\alpha} &= \sum_{f=1}^{k} H_{fr} + \theta_{r,\alpha}, \\ \Delta_{l,\alpha} &= \sum_{f=1}^{k} H_{fl} - \theta_{l,\alpha}, \end{aligned} \tag{5.6}$$

In the foregoing calculation, the conditionally constant elementary errors were modeled by a random quantity with a uniform probability distribution. However, elementary errors, which appear in the resulting error after some transformation, are encountered. An example is the mismatch error in radioelectronic measurements. The elementary error here is the phase shift Δ_{φ_f}. As in the case of other elementary errors, for this error, the limits of the phase shift are estimated, and it is assumed that the phase shift is uniformly distributed within these limits. But the error in the result contains not Δ_{φ_f}, but rather $\cos \Delta_{\varphi_f}$. When Δ_{φ_f} is distributed uniformly, he quantity $\cos \Delta_{\varphi_f}$ has the so-called arccosine distribution.

When transformed elementary errors are present, their composition with the other errors must be constructed according to adopted mathematical methods. The universal numerical method described in Section 3.6 is convenient.

The distribution of the sum of elementary errors, for each of which a uniform distribution is adopted, can be regarded as a normal distribution, if $n > 4$ and they all have approximately the same limits. If, however, $n \leq 4$ or the limits of

the elementary errors are substantially different, then the composition of these distributions must also be constructed.

We now return to the case when the elementary error has unsymmetric limits. The transformation of these limits to a symmetric form with a shift by a_j, creates the temptation to introduce into the measurement result a correction corresponding to the shift a_j. One must be decisively cautioned against doing this: Information about errors is too unreliable to use for correcting the result of a measurement.

In the calculations performed above, it was assumed that the elementary errors are independent of one another. In some cases, this assumption is not justified. An example is the case when several measuring devices, connected into a measuring system or forming a measurement channel, are used to perform a measurement and some influence quantity exceeds the limits of reference conditions for these instruments. In this case, the measuring devices will acquire additional errors and they could be dependent, and it must be taken into consideration. The method of such calculations relates to the indirect measurements, and therefore, it is presented in Section 6.6.

From n elementary errors, it is possible to single out $2m$ dependent errors. Then (5.3) must be transformed into the following form:

$$\theta = k\sqrt{\sum_{i=1}^{n-m}\theta_i^2 + \sum_{\mu=1}^{m}(\theta_{\mu_1} \pm \theta_{\mu_2})^2}, \tag{5.7}$$

where θ_{μ_1} and θ_{μ_2} are the limits of pairs of dependent elementary errors with their signs.

The method presented above for calculating errors is equally suitable for a priori and a posteriori estimation, because at the synthesis stage, there is no difference between these cases.

In conclusion, we shall discuss the formula

$$\theta = \sqrt{\sum_{i=1}^{n}\theta_i^2},$$

which is often used in practice. This formula can be obtained under the assumption that the elementary errors have a normal distribution and their limiting values were calculated for one and the same confidence probability. Let σ_i be the standard deviation of the ith elementary error and $\theta_i = z_\mathrm{p}\sigma_i$, where z_p is the quantile determined according to a normal distribution and one and the same confidence probability for all i. It is obvious that

$$\sigma^2 = \sum_{i=1}^{n}\sigma_i^2.$$

We multiply both sides of this equality by z_p^2:

$$z_\mathrm{p}^2\sigma^2 = \sum_{i=1}^{n} z_\mathrm{p}^2\sigma_i^2.$$

But $\theta = z_p\sigma$ and $\theta_i = z_p\sigma_i$ From here we obtain the formula

$$\theta = \sqrt{\sum_{i=1}^{n} \theta_i^2}. \tag{5.8}$$

However, the limits of elementary errors are not estimated by probabilistic methods, and a probability cannot be assigned to them. Moreover, there are no grounds for using a normal distribution as the mathematical model of elementary errors.

The formula presented above can be interpreted differently, namely, as a, particular case of formula (5.3) with $k = 1$[45]. The value $k = 1$ corresponds to a confidence probability of 0.916, which explains the commonly held opinion that this formula somewhat underestimates the error, i.e., that this formula is not reliable enough.

On the other hand, this formula is widely used in practice. Its wide dissemination can be regarded as an indirect but practical confirmation of the fact that a uniformly distributed random quantity can be used as a model of conditionally constant elementary errors.

Sometimes the elementary components are summed according to the formula

$$\Delta = \sum_{f=1}^{k} H_f + \sum_{i=1}^{n} \theta_i.$$

Such summation, however, means that all elementary errors are assumed to be absolutely constant. This situation is rare. If, however, it is agreed that conditionally constant errors are also present, then the arithmetic summation means that all elementary errors simultaneously assume their limiting values, and with the same sign. This coincidence is unlikely. Although this formula satisfies the principle of estimating the upper limit of errors, it is used less and less, and primarily only for obtaining a rough estimate a single measurement error or in the extreme case, as done above on page 129 simply to eliminate an erroneous estimate.

5.5. Example: Calculation of Uncertainty in Voltage Measurements Performed with a Pointer-Type Voltmeter

We shall study several examples of the application of a class 1.0 pointer-type dc voltmeter with the following characteristics:

(i) The upper limits of measurement ranges are 3, 7.5, 15, 30, and so on, up to 300 V.
(ii) The scale of the instrument has 75 graduations and starts at the 0 marker.

(iii) The limits of permissible intrinsic error are $\pm 1.0\%$ of a span (it is a fiducial error).
(iv) Full deflection of the pointer corresponds to a current of 15×10^{-6} A $\pm 1\%$.
(v) Under the reference conditions, the temperature is equal to $+20 \pm 5\,^{\circ}$C and the measurements are performed with the instrument positioned horizontally. In this case, we shall ignore all other influence quantities; we shall assume that they are identical to their normal reference values.

Additional errors. A deviation of the temperature from the reference value causes the indications of the instrument to change by not more than $\pm 1.0\%$ for each $10\,^{\circ}$C change in temperature. Inclination of the instrument by 5° from the horizontal position changes the indications by not more than $\pm 1\%$ of the measurement limit.

5.5.1. *A priori Estimation*

Suppose that some piece of equipment is to be monitored by measuring the voltage on several resistors. The equivalent output resistance (the source resistance) of the equipment in one case is equal to about $10\,\text{k}\Omega$ and in all other cases does not exceed $1\,\text{k}\Omega$. The temperature of the medium can change from $+10\,^{\circ}$C to $+25\,^{\circ}$C The slope relative to the horizontal position does not exceed 5°.

We are required to estimate the measurement uncertainty. The uncertainty most be expressed in the relative form.

Before the measurement, the value of the measured quantity is unknown. It will supposedly be less than 3 V. The overlapping of limits in the voltmeter is equal to $3/7.5 = 0.4$ and $7.5/15 = 0.5$, after which these numbers repeat. Thus, the indication of the instrument drops below $0.4 - 0.5$ of the upper limit of measurement, then the range of measurement must be switched. Developing this point of view, we shall assume that if the measured voltage is less than $0.4 \times 3\ \text{V} = 1.2\ \text{V}$, then a different voltmeter must be used.

In the range 1.2–3 V, the largest relative error will occur when a voltage of the order of 1.2 V is being measured. The error will have to be estimated for this worst case.

The sources of error are as follows:

(1) the intrinsic error of the voltmeter;
(2) the reading error;
(3) the temperature error;
(4) the error introduced by the inclination of the instrument; and
(5) the error from the limited input resistance of the voltmeter.

We shall estimate these errors.

(1) Intrinsic error: Its limits will be

$$\theta_{\text{in}} = \pm 1\% \times \frac{1}{0.4} = \pm 2.5\%, \quad |\theta_{\text{in}}| = 2.5\%.$$

(2) Reading error: This error does not exceed 0.25 of a graduation. When measuring 1.2 V at the limit 3 V, this gives

$$\theta_\alpha = \pm 0.25 \times \frac{3 \times 100}{75 \times 1.2} = \pm 0.83\%, \quad |\theta_\alpha| = 0.83\%.$$

(3) Additional temperature error: The maximum deviation of the temperature from the normal value is $(20 - 5) - 10 = 5\,^\circ\mathrm{C}$. For this reason

$$\theta_T = \pm 1\% \times \frac{5}{10} = \pm 0.5\%, \qquad |\theta_T| = 0.5\%.$$

(4) The additional error introduced by the 5° inclination of the instrument when measuring 1.2 V will be

$$\theta_l = \pm 1\% \times \frac{3}{1.2} = \pm 2.5\%, \qquad |\theta_l| = 2.5\%.$$

(5) The errors $\{\varepsilon_R\}$ from the limited input resistance of the voltmeter are as follows. The input resistance of the voltmeter at the limit 3 V is

$$R_V = \frac{3}{1.5 \times 10^{-5}} = 2 \times 10^5\ \Omega.$$

The worst case occurs with the outside resistance $R'_{\mathrm{or}} = 10\,\mathrm{k}\Omega$.

The indications of the voltmeter correspond to the voltage on its terminals. This voltage U is less than the emf E in the circuit:

$$U = \frac{R_V}{R_V + R_{\mathrm{or}}} E.$$

The error is

$$\varepsilon_R = \frac{U - E}{E} = \frac{-R_{\mathrm{or}}}{R_V + R_{\mathrm{or}}}$$

For $R'_{\mathrm{or}} = 10\,\mathrm{k}\Omega$

$$\varepsilon'_R = \frac{-10 \times 10^3}{10 \times 10^3 + 2 \times 10^5} \times 100 = -4.8\%$$

If the outside resistance is $1\,\mathrm{k}\Omega$, then $\varepsilon''_R = -0.5\%$.

The errors $\{\varepsilon_R\}$ are absolutely constant for each unit being monitored. The remaining errors are conditionally constant.

Let us now add all conditionally constant errors. We shall use (5.4), and we shall assume that $\alpha = 0.95$:

$$\Delta_{\%,0.95} = 1.1\sqrt{2.5^2 + 0.83^2 + 0.5^2 + 0.25^2} = 4\%.$$

We now take into account the absolutely constant error. Its limits are

$$H_{Rl} = -4.8\%, \qquad H_{Rr} = -0.5\%,$$

but they are not known accurately enough to eliminate them by introducing the correction. Therefore, in accordance with (5.6), we obtain

$$\Delta_{r,0.95} = -0.5 + 4 = +3.5\%, \qquad \Delta_{l,0.95} = -4.8 - 4.0 = -8.8\%$$

Thus, the error of the planned measurement will not exceed $\sim 10\%$.

5.5.2. *An Approximate a posteriori Estimation*

We shall now estimate the measurement error in the example examined above, assuming that the measurement has already been made. The significant difference from the foregoing case is that now we have an estimate of the measured quantity. Let the indication of the voltmeter in the case $R'_{\text{or}} = 10\,\text{k}\Omega$ be 62.3 graduations. Hence, the voltmeter is indicated

$$U = 62.3\frac{3}{75} = 2.492\ \text{V}.$$

Suppose we found out that $R'_{\text{or}} = 10\,\text{k}\Omega \pm 0.5\%$. The error ε'_R was calculated above: $\varepsilon'_R = -4.8\%$. Now we can introduce the correction C'_R:

$$C'_R = +4.8 \times 10^{-2} \times 2.492 = +0.120\ \text{V}.$$

Taking the corrections into account, we obtain

$$U' = U + C'_R = 2.612\ \text{V}.$$

The errors of the corrections are determined by the errors of the available values of the resistances R_V and R_{or}. We shall establish the relation between them.

$$C'_R = -\varepsilon'_R U = \frac{R_{\text{or}}}{R_{\text{or}} + R_V}U = \frac{R_{\text{or}}}{R_{\text{or}} + R_V} \times \frac{R_V}{R_{\text{or}} + R_V}E = \frac{R_{\text{or}}/R_V}{(1 + R_{\text{or}}/R_V)^2}E.$$

To simplify the notation, let $x = R_{\text{or}}/R_V$. Then

$$C'_R = \frac{x}{(1+x)^2}E.$$

We now construct the differential relations:

$$dx = \frac{1}{R_V}dR_{\text{or}} - \frac{R_{\text{or}}}{R_V^2}dR_V = x\left(\frac{dR_{\text{or}}}{R_{\text{or}}} - \frac{dR_V}{R_V}\right),$$

$$dC'_R = E\left(\frac{dx}{(1+x)^2} - \frac{2x(1+x)dx}{(1+x)^4}\right) = E\frac{1-x}{(1+x)^3}dx,$$

$$dC'_R = E\frac{x(1-x)}{(1+x)^3}\left(\frac{dR_{\text{or}}}{R_{\text{or}}} - \frac{dR_V}{R_V}\right).$$

In the relative form, transforming from differentials to increments, we obtain

$$\varepsilon_c \frac{\Delta C'_R}{C'_R} = \frac{1-x}{1+x}\left(\frac{\Delta R_{\text{or}}}{R_{\text{or}}} - \frac{\Delta R_V}{R_V}\right).$$

As ΔR_{or} and ΔR_V are independent, we shall regard each component of error of the correction as an elementary error of measurement. Obviously, both components are conditionally constant:

$$\theta_{C1} = \left(\frac{1-x}{1+x}\right)\theta_{R_{\text{or}}}, \qquad \theta_{C2} = \left(\frac{1-x}{1+x}\right)\theta_{R_v}.$$

The limits of error of the internal resistance of the voltmeter are determined by the voltmeter class and are equal to $\pm 1\%$. Therefore, because $x = 5 \times 10^{-2}$,

$$|\theta_{C2}| = \left(\frac{1-x}{1+x}\right) 1\% = 0.9 \times 1\% = 0.9\%.$$

The limits of the error in determining the input resistance of our apparatus (the outside resistance for voltmeter) are equal to $\pm 0.5\%$. Therefore,

$$|\theta_{C1}| = \left(\frac{1-x}{1+x}\right) 0.5\% = 0.9 \times 0.5\% = 0.45\%.$$

The limits of the remaining errors are as follows:

$$\begin{aligned} |\theta_{\text{in}}| &= 1\% \times 75/62 = 1.2\%, \\ |\theta_{\alpha}| &= \frac{0.25 \times 100}{62} = 0.4\% \\ |\theta_{T}| &= 0.5\%, \\ |\theta_{l}| &= 1\% \times 75/62 = 1.2\%. \end{aligned}$$

These elementary errors can be assumed to be conditionally constant. According to formula (5.4), for $\alpha = 0.95$, we obtain

$$\Delta_{0.95} = 1.1\sqrt{0.9^2 + 0.45^2 + 1.2^2 + 0.4^2 + 0.5^2 + 1.2^2} = 2.3\%.$$

When the result of the measurement is written in accordance with its uncertainty, only three significant figures can be retained:

$$\tilde{U}' = 2.16\ \text{V}, \qquad \Delta = \pm 2.3\%\ (\alpha = 0.95),$$
$$U'(0.95) = 2.16\ \text{V} \pm 2.3\%, \quad \text{or} \qquad U'(0.95) = (2.16 \pm 0.06)\ \text{V}.$$

5.5.3. *An Accurate a posteriori Estimation*

The largest elementary errors were $\theta_{C2}, \theta_{\text{in}}$, and θ_l. How can they be reduced? The first two can be reduced by taking into account the individual properties of the voltmeter. If the voltmeter has a fresh table of corrections, then this can be done. Assume that at the limit 3 V on marker 60, the correction is equal to $+0.3$ graduations, whereas at marker 70, it is equal to $+ 0.2$ graduations. It can then be assumed that the correction to the indication at 62.3 graduations is also equal to $+0.3$ graduations. Therefore,

$$C_{\text{in}} = +0.3 \times \frac{3}{75} = +0.012\,\text{V}.$$

Taking this correction into account, the voltmeter gives

$$U'' = 2.492 + 0.012 = 2.504\,\text{V}.$$

We shall assume that the limits of error in determining the correction, i.e., the calibration errors, are known and are equal to $\pm 0.2\%$. Converting to the indication of the instrument, we obtain $|\theta'_{\text{in}}| = 0.2 \times 75/62 = 0.24\%$.

With this correction, we have eliminated the systematic component of the error of the voltmeter. The random component, however, remains, and it must be taken into account. The dead band, according to the indicating electric measurement instruments, can reach a value coinciding with the class designation of the instrument. In our case, this value is 1% of 3 V. The random error does not exceed half the dead band. Thus, the limits of random error are equal to

$$|\Psi| = 0.5 \times 1\% \times \frac{75}{62} = 0.6\%.$$

The distribution of the random error in our case, when its limits have been estimated, can be assumed to be uniform, as also the distributions of other conditionally constant elementary errors.

The input resistance of the voltmeter can be measured. Assume that this measurement has been done, and $R_V = 201.7\,\text{k}\Omega \pm 0.2\%$. Then

$$\varepsilon_R = \frac{10 \times 10^3 \times 100}{(10 + 201.7) \times 10^3} = -4.72\%.$$

The correction will be

$$C_R = +4.72 \times 10^{-2} \times 2.504 = +0.118\,\text{V}.$$

Taking the correction C_R into account, we obtain

$$U'' = 2.504 + 0.118 = 2.622\,\text{V}.$$

The limits of error because the input resistance of the voltmeter is not known exactly will now become smaller:

$$|\theta'_{C2}| = 0.9 \times 0.2\% = 0.18\%, \quad |\theta_{C1}| = 0.45\%.$$

The error θ_l can be reduced by taking greater care in positioning the instrument horizontally. Assume that the deviation from the horizontal position does not exceed $\pm 2^\circ$. Then

$$|\theta'_l| = 1 \times 2/5 \times 75/62 = 0.48\%.$$

The temperature error and the reading error will remain the same.

Let us calculate the uncertainty again for $\alpha = 0.95$:

$$\Delta_{0.95} = 1.1\sqrt{0.24^2 + 0.6^2 + 0.18^2 + 0.45^2 + 0.48^2 + 0.5^2 + 0.4^2} = 1.3\%.$$

We now write the result of the measurement as follows:

$$\tilde{U}'' = 2.62\ \text{V}, \quad \Delta = \pm 1.3\%\ (\alpha = 0.95), \quad \text{or} \quad U''(0.95) = 2.62\ \text{V} \pm 1.3\%.$$

The example examined above shows clearly how the measurement uncertainty decreases as one transfers from a priori to a posteriori estimation and then from approximate to accurate error estimation.

5.6. Methods for Calculating the Uncertainty in Multiple Measurements

Multiple measurements are a classic object of mathematical statistics and the theory of measurement errors. Under certain restrictions on the starting data, mathematical statistics give elegant methods for analyzing observations and for estimating measurement errors.

Unfortunately, the restrictions required by mathematics are not often satisfied in practice. Then these methods cannot be used, and practical methods for solving the problems must be developed. But even in this case, the methods of mathematical statistics provide a point of reference and a theoretical foundation.

In this chapter, the principal mathematical methods for solving the problem mentioned above are presented. The situation corresponding to direct multiple measurements, free of systematic errors, i.e., having only random errors, is studied. Under this restriction, mathematical methods can be fully employed.

A separate result, i.e., a separate value of the random error, cannot be predicted. But a large collection of random errors of some measurement satifies definite laws. These laws are statistical (probabilistic). They are established and proved in metrology based on the methods of mathematical statistics and the theory of probability.

The problem can be solved best if the distribution function of the observations is available. In practice, however, distribution functions are, as a rule, unavailable.

If the random character of the observational results is caused by the measurement errors, it is usually assumed that the observations have a normal distribution. The computational results based on this assumption do not, as a rule, lead to contradictions, which probably happens for two reasons. First, the measurement errors consist of many components. According to the central limit theorem, this leads, in the limit, to the normal distribution. In addition, the measurements for which accuracy is important are performed under controlled conditions, as a result of which the distributions of their errors turn out to be bounded. For this reason, their approximation by a normal distribution, for which the random quantity can take on arbitrary values, leads to wider confidence intervals than the intervals that would be obtained if the true distribution were known.

There are examples, however, when the observational results do not correspond to the normal distribution. In addition, when the measured quantity is an average value, the distribution of the observations can have any form. For this reason, the hypothesis that the distribution of the observations is normal must, in principle, be checked.

The methods of statistical calculations for observations that are distributed normally have been well developed and the required tables have been constructed. If, however, the hypothesis that the distribution is normal must be rejected, then the statistical analysis of the observations becomes much more complicated. Mathematicians have been working to find, if not better than, at least satisfactory estimates for parameters of distributions whose form has not been precisely established [16, 31, 32].

Random and quasirandom errors of multiple measurements are always estimated based on the experimental data obtained in the course of the measurements, i.e., a posteriori.

As was shown in Chapter 1, despite the existence of random errors, a measured quantity can only be a quantity that is defined in a model as nonrandom and constant. The problem is to find from the experimentally obtained data the best estimate of the measured quantity.

The mathematical–statistical methods studied in Chapter 4 form the theoretical foundation for estimating measured quantities and their errors in the case of multiple direct measurements. These methods can also be regarded as practical methods, if the systematic components of the measurement error are negligible compared with the random or quasirandom component.

In the general case, both the systematic and the random components of the error must be estimated. The random error can be estimated only a posteriori; the systematic error, however, can also be estimated a priori.

Consider first the case when the measurements are repeated to reduce the random errors. Having n single measurements, we obtain $\{x_i\}$, $i = 1, \ldots, n$, where $x_i = A + \zeta_i$ and $\zeta_i = \psi_i + \vartheta_{ij}$; i.e., the error has both random and systematic components.

By repeating the measurements, we obtain information about the random error. Information about the systematic error cannot be extracted from the measurements. To estimate this error, it is necessary to know the properties of the measuring instrument employed, the method of measurement, and the conditions under which the measurements are performed.

It is important to mention that random components of all conditionally constant errors now become a part of the random error of the measurement. Thus, the remaining parts of conditionally constant errors in multiple measurements are purely systematic errors. But in various measurements of the same measurand performed by the same method, the values of these errors can vary.

Assume that the systematic error of the result of each observation (single measurement) is known. Then, introducing the corrections $C_i = -\vartheta_i$, we obtain a group of corrected measurement results

$$x_i = A + \psi_i.$$

Our problem is to find the estimate $A = f(x_i)$. A mathematically well-founded solution, which is unbiased, consistent, and efficient, can be found if the form of the distribution of x_i is known. Measurement errors can often be assumed to have a normal distribution. The measurement results also have the same distribution. In principle, it is possible to check whether the data obtained conformed to a normal distribution (see the methods presented in Chapter 4). Admittedly, this process requires many measurements; in practice, enough measurements to make such a check are rarely performed, and it is usually simply assumed that the distribution is normal.

For a normal distribution, as shown in Section 4.2, the arithmetic mean is the optimal estimate of the center of the distribution X. As noted above, the arithmetic

mean of the measurements is an unbiased, consistent, and efficient estimate of the true value of the measured quantity only if the observations have a normal distribution. Irrespective of the form of the distribution of the measurement errors, however, the arithmetic mean has two important properties.

(1) The sum of the deviations from the arithmetic mean is equal to 0. Let $x_i, \ldots, x_n$ be a group of observational results whose arithmetic mean is $\bar{x}$. We construct the differences $x_i - \bar{x}$ for all $i = 1, \ldots, n$ and find their sum:

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x}.$$

As $\sum_{i=1}^{n} x_i = n\bar{x}$ and $\sum_{i=1}^{n} \bar{x} = n\bar{x}$,

$$\sum_{i=1}^{n}(x_i - \bar{x}) = 0.$$

This property of the arithmetic mean can be used to check the calculations.

(2) The sum of the squares of the deviations from the arithmetic mean is minimum. Consider the function

$$Q = \sum_{i=1}^{n}(x_i - \tilde{A})^2.$$

We shall find A to minimize Q. For this reason, we find

$$\frac{dQ}{d\tilde{A}} = -2\sum_{i=1}^{n}(x_i - \tilde{A})$$

and set $dQ/d\tilde{A} = 0$; hence, we obtain

$$\sum_{i=1}^{n}(x_i - \tilde{A}) = 0, \qquad \sum_{i=1}^{n} x_i = n\tilde{A}, \qquad \text{and} \qquad \tilde{A} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

As $dQ/d\tilde{A} < 0$ if $\tilde{A}_1 < \bar{x}$ and $dQ/d\tilde{A} > 0$ if $\tilde{A}_2 > \bar{x}$, for $\tilde{A} = \bar{x}$, we have the minimum of Q.[1]

Although the sum of the squares of the deviations from the arithmetic mean is minimum, this only means that in the class of estimates that are a linear function of the measurement results, the arithmetic mean is the most efficient estimate of

[1] Dr. E.R. Cohen has shown me another way to get this result:

$$Q = \sum_{i=1}^{n}(x_i - \tilde{A})^2 = \sum_{i=1}^{n}(x_i - \bar{x} + \bar{x} - \tilde{A})^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 + 2(\bar{x} - \tilde{A})\sum_{i=1}^{n}(x_i - \bar{x}) + \sum_{i=1}^{n}(x - \tilde{A})^2.$$

The second term is equal to zero, and the third one is always positive or equal to zero if $\bar{x} = \tilde{A}$. Therefore, the choice $\tilde{A} = \bar{x}$ gives the minimal value for Q.

the measured quantity. This estimate becomes absolutely efficient if the errors are distributed normally. For other distributions, as pointed out in Chapter 3, estimates exist that are more efficient than the arithmetic mean. Obviously, these estimates are no longer a linear function of the measurement results.

Thus, for the estimate of the measured quantity, we have

$$\tilde{A} = \frac{\sum_{i=1}^{n} x_i}{n}. \tag{5.9}$$

Because of random errors, the measurement results are also random quantities; if another series of measurements is performed, then the new arithmetic mean obtained will differ somewhat from the previously found estimate.

The spread of the arithmetic means is characterized either by the variance of the arithmetic means or by the standard deviation. In accordance with (4.5), the standard deviation of the arithmetic mean is estimated from the experimental data as follows:

$$S_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n(n-1)}}. \tag{5.10}$$

In addition, for A, it is possible to construct the confidence interval, determining the confidence limits of the random error in the measurement results. The confidence interval is determined by the inequalities

$$\tilde{A} - \Psi_\alpha \le A \le \tilde{A} + \Psi_\alpha,$$

where $\Psi_\alpha = t_q S_{\bar{x}}$, and t_q is the q percent point of Student's distribution and depends on the confidence probability α and the number of degrees of freedom $\nu = n - 1$ (see Table A.2).

Therefore, the random error ψ with probability equal to the confidence probability α has the limits $\pm\Psi_\alpha$:

$$\Psi_\alpha = t_q S_{\bar{x}}. \tag{5.11}$$

As one can see from what was said above, the random errors and confidence limits of these errors can be estimated from the data obtained as a result of measurements.

The situation is different in the case of the systematic errors. The biasness, characterized by the systematic errors, of the result of a measurement can be estimated either with the help of more accurate means of measurement or based on indirect data, including data on the metrological properties of the measuring instrument employed to perform the measurements. The first case is pointless; the more accurate measurement would replace the less accurate measurement. The problem of estimating the systematic error would remain, except that now it would pertain to the more accurate result. For this reason, the second case is the main one.

In the case of multiple measurements, the most common variant is the one when the most important systematic errors are eliminated with the help of corrections.

Then the errors in determining the corresponding corrections must be taken into account instead of the eliminated errors.

An important characteristic of multiple measurements is that the random components of the elementary errors are manifested in multiple measurements and contribute to the random error of the result. For this reason, when estimating elementary errors, it is desirable to neglect their random components. We shall return to this question at the end of the section.

By summing the elementary errors freed of random components, we obtain the limits of the systematic error of the result of measurement. The method of summation is presented in Section 5.4.

Thus, we obtain an estimate $S_{\bar{x}}$ of the standard deviation of the random error of the result of measurement with a known number of measurements n and the estimate of the limits of the systematic component θ.

In some cases, this separate estimate of the components of the uncertainty of measurement is sufficient. This is the case, for example, if the result of measurement is to be used for calculations together with other data, for which separate estimates of the components of uncertainty are also known or if the result of a measurement is to be compared with the results of other measurements, for which the uncertainty components are determined separately.

Often, however, it is necessary to find the total uncertainty of a measurement, including both the random and the systematic components. Not too long ago, specialists on accurate measurements objected to this formulation of the problem. They said that systematic and random errors are of a different nature, and for this reason, they cannot be added. In 1965, I still listened to these assertions. However, most people disagreed with this reasoning. Indeed, in a completed measurement, these components are physically indistinguishable; i.e., physically they add: By measuring A, we obtain $\tilde{A}$. The difference $\tilde{A} - A$ contains both systematic and randon components. When analyzing the error, theoretically or experimentally, we decompose this error into its components: systematic and random. Obviously, the inverse problem of adding the components is legitimate.

To solve the problem, we shall take into account only the conditionally constant and random errors. Regarding both types of errors as random quantities, to combine them, we must construct the composition of the corresponding distributions. Unfortunately, this is too difficult in practice. For this reason, the uncertainty of a measurement is sometimes calculated according to the formula

$$u = \theta + \Psi_\alpha,$$

where θ is the limit of the systematic error and $\Psi_\alpha = t_q S_{\bar{x}}$ is the confidence limit of the random error.

This formula is simple, but it is clear that it gives an obviously overestimated estimate. A more plausible solution can be found by the following method [41], [44].

In the general form, the error of a measurement result has three components:

$$\zeta = \eta + \vartheta + \psi.$$

Therefore, the variance of measurement result is

$$V[\zeta] = V[\vartheta] + V[\psi].$$

$V[\zeta]$ has only two terms because $V[\eta] = 0$.

Estimates of $V[\vartheta]$ and $V[\psi]$ can be found using formulas (3.15) and (5.10). Denote them S_ϑ^2 and $S_{\bar{x}}^2$. Denote also the combined variance $V[\zeta] = S_c^2$, where S_c is the combined standard deviation. Then the combined standard deviation S_c is

$$S_c = \sqrt{S_\vartheta^2 + S_{\bar{x}}^2}. \tag{5.12}$$

Given S_c, the uncertainty of the measurement result could be calculated from the formula

$$u_c = t_c S_c, \tag{5.13}$$

if the coefficient t_c was known; unfortunately, this coefficient is unknown.

As the initial data, i.e., the data on the components of the uncertainty, are not known accurately, an approximate estimate of the coefficient t_c can be used. In [41], following formula was proposed for making such an estimate:

$$t_c = \frac{\Psi_\alpha + \theta_\alpha}{S_{\bar{x}} + S_\vartheta}.$$

This formula was constructed based on the following considerations. The coefficient t_q, determining the ratio of the confidence limit and the standard deviation of the random error, is determined by Student's distribution and is known. Given estimates for θ_α and S_ϑ, it can be assumed that the analogous coefficient

$$t_\vartheta = \theta_\alpha / S_\vartheta$$

for the systematic error is also known.

It is natural to assume that the coefficient sought t_c is some function of t_q and t_ϑ, which corresponds to the same probability. The weighted mean of t_q and t_ϑ for the weights $S_\vartheta/(S_{\bar{x}} + S_\vartheta)$ and $S_{\bar{x}}/(S_{\bar{x}} + S_\vartheta)$, respectively, was taken for this function, which results in the proposed formula

$$t_c = \frac{t_q S_{\bar{x}} + t_\vartheta S_\vartheta}{S_{\bar{x}} + S_\vartheta} = \frac{\Psi_\alpha + \theta_\alpha}{S_{\bar{x}} + S_\vartheta}. \tag{5.14}$$

To use this formula, its accuracy must be estimated. The extreme cases are those when the systematic error has a normal or uniform distribution. The distribution of the random error of the arithmetic mean may be assumed to be asymptotically normal.

If both terms have a normal distribution, then $t_q = t_\vartheta$, and as follows from formula (5.14), $t_c = t_q$. As the composition of normal distributions gives a normal distribution, the obtained value of t_c is exact.

For the second case, the results of calculations based on the approximate formula (5.14) must be compared with the results obtained from the exactly constructed composition of normal and uniform distributions.

TABLE 5.1. Characteristic quantiles for the composition of centered normal and uniform distributions.

h/σ	0.50	1.0	2.0	3.0	4.0	5.0	6.0	8.0	10
$z_{0.95}$ ($\alpha = 0.90$)	1.71	1.90	2.49	3.22	4.00	4.81	5.65	7.34	9.10
$z_{0.975}$ ($\alpha = 0.95$)	2.04	2.25	2.90	3.67	4.49	5.34	6.22	8.00	9.81
$z_{0.995}$ ($\alpha = 0.99$)	2.68	2.94	3.66	4.49	5.36	6.26	7.17	9.02	10.90

An expression for the distribution density of the sum of two unknown centered random quantities, one of which has a normal distribution and the other has a uniform distribution, is known from the theory of probability:

$$f(z) = \frac{1}{2h} \int_{-h}^{h} \frac{1}{\sigma\sqrt{2\pi}} e^{-(z-y)^2/2\sigma^2} dy,$$

where h is equal to one half the interval in which the random quantity y is distributed uniformly.

Setting $\sigma = 1$ and transforming to the probability distribution function, we find

$$\underset{z>0}{F(z)} = 0.5 + \frac{1}{2h\sqrt{2\pi}} \int_0^z \int_{-h}^{h} e^{-(z-y)^2/2} dy\, dz.$$

The starting distributions are symmetric relative to 0. For this reason, the density of the resulting distribution also has this property. We must find the limit of the confidence interval corresponding to the probability α. For this reason, it is sufficient to find either the quantile z_r of the level r or the quantile of the level $1 - r$, because $|z_r| = |z_{1-r}|$. As $\alpha = 1 - 2r$, $r = (1 - \alpha)/2$. Obviously, $r < 0.5$ and $z_r > 0$.

Table 5.1 gives values of z_{1-r} calculated using the presented formula for $\alpha = 0.90$, 0.95, and 0.99.

The relative error introduced by the use of the approximate formula (5.14) will be

$$\delta = \frac{u_c - z_{1-r}}{z_{1-r}}.$$

The comparison should be made with $S_c = \sigma_c$, because $S_{\bar{x}} = \sigma = 1$. In so doing,

$$\sigma_c = \sqrt{\sigma^2 + \frac{h^2}{3}} = \sigma\sqrt{1 + \frac{1}{3}\left(\frac{h}{\sigma}\right)^2}.$$

For this reason, introducing the coefficient $t_r = z_{1-r}/\sigma_c$, we obtain

$$\delta = \frac{t_c - t_r}{t_r}.$$

The coefficient t_r depends only on the configuration of the resulting distribution, i.e., on the ratio of h and σ, and not on their absolute values. For this reason, a series of values of this coefficient can be calculated from the data in Table 5.1. These values are presented in Table 5.2.

TABLE 5.2. Values of the combined standard deviation σ_c and of the coefficient t_r as a function of the parameters of the normal and uniform distributions.

h/σ	0.5	1	2	3	4	5	6	8	10
σ_c ($\sigma = 1$)	1.04	1.15	1.53	2.00	2.52	3.06	3.51	4.72	5.85
t_r $\alpha = 0.90$	1.65	1.64	1.63	1.61	1.59	1.58	1.57	1.56	1.55
t_r $\alpha = 0.95$	1.96	1.95	1.90	1.84	1.78	1.75	1.72	1.69	1.67
t_r $\alpha = 0.99$	2.57	2.54	2.40	2.24	2.13	2.05	1.99	1.91	1.86

We shall now once again turn to the approximate formula (5.14). The limits of the confidence interval, which are determined based on the uniform distribution, give θ. As $r = (h - \theta)/2h$ and $r = (1 - \alpha)/2$,

$$\theta = (1 - 2r)h = \alpha h. \tag{5.15}$$

The limit of the confidence interval for a normal distribution with the same confidence probability will be

$$\Psi_\alpha = z_{\frac{1-\alpha}{2}}\sigma$$

where $z_{\frac{1-\alpha}{2}}$ is the quantile of the standard normal distribution that corresponds to the probability α.

Formula (5.14) assumes the form

$$t_c = \frac{z_{\frac{1-\alpha}{2}}\sigma + \alpha h}{\sigma + h/\sqrt{3}}.$$

The values of t_c, calculated for the same ratios h/σ and confidence probabilities as were used for calculating t_r, are presented in Table 5.3.

The errors δ calculated based on the data given in Tables 5.2 and 5.3 are summarized in Table 5.4.

Thus, comparing the results of exact calculations with the results of calculations performed using the approximate formula (5.14) shows that the errors from the use of the approximate formula are in all cases negative and their absolute magnitude does not exceed 12% for $\alpha = 0.99$, 6% for $\alpha = 0.95$ and 2% for $\alpha = 0.90$, which shows that formula (5.14) can be used.

It should also be noticed that the error under study decreases as the distribution of the systematic errors approaches the normal distribution.

TABLE 5.3. Values of the coefficient t_c as a function of the parameters of the normal and uniform distributions.

h/σ	0.5	1	2	3	4	5	6	8	10
t_{1c} ($\alpha = 0.90$)	1.63	1.61	1.60	1.59	1.58	1.58	1.58	1.57	1.57
t_{2c} ($\alpha = 0.95$)	1.89	1.84	1.79	1.76	1.74	1.73	1.72	1.70	1.69
t_{3c} ($\alpha = 0.99$)	2.38	2.26	2.11	2.03	1.97	1.94	1.91	1.87	1.84

TABLE 5.4. Deviations of the coefficient t_c from t_r (in %).

h/σ	0.5	1	2	3	4	5	6	8	10
δ_1 ($\alpha = 0.90$)	−1.2	−1.9	−1.8	−1.1	−0.6	0.0	0.8	0.6	1.2
δ_2 ($\alpha = 0.95$)	−3.6	−5.5	−5.7	−4.1	−2.2	−1.3	0.0	0.5	1.0
δ_3 ($\alpha = 0.99$)	−7.4	−11.0	−12.1	−9.4	−7.3	−5.5	−4.0	−2.2	−1.1

The scheme presented above for estimating the uncertainty of a measurement, which contains both random and systematic components, is a general scheme. In different particular problem, it can be substantially modified.

We also discuss the question of when the systematic or random component of the uncertainty can be neglected.

Figure 5.2 shows plots, constructed based on the data in Table 5.1, of $z_{1-r}(h)$, which correspond to the composition of normal and uniform distributions and to the uniform distributions. Comparing the curves 1 with the values of z_{1-r} of a normal distribution (the points for $h = 0$), we can find the error introduced by neglecting the systematic component. It is assumed that this error does not exceed 15%,

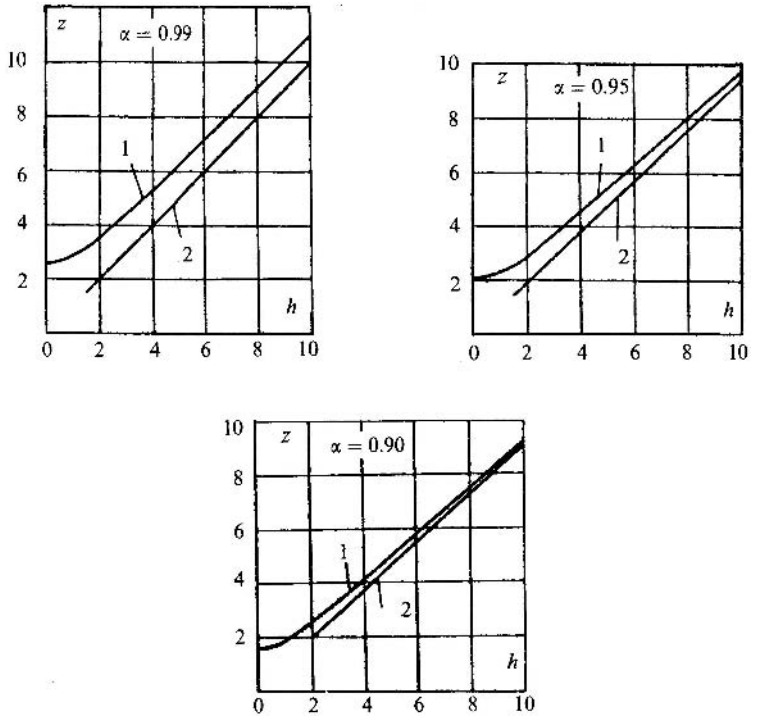


FIGURE 5.2. Quantiles of the levels 0.99, 0.95, and 0.90 for a composition of the normal and uniform distributions (curves, 1) and for uniform distribution (straight lines, 2).

then we obtain the limiting values of the ratio h/σ. Depending on the confidence probability, these ratios are as follows:

α	0.90	0.95	0.99
h/σ	1.2	1.1	1.1

If the random component is neglected, then the error arising is determined by the difference of the ordinates of the exact curve 1 and the straight line 2 for fixed h. For the same 15% error, we obtain the condition under which the random component can be neglected:

α	0.90	0.95	0.99
h/σ	3	4	7

It is obvious that when any component is neglected, the overall error decreases.

Thus, if $h/\sigma < 1$, then the systematic error can be confidently neglected, and if $h/\sigma > 7$, then the random component can be neglected.

Admittedly, we do not know the exact values of the parameters h and σ; we know only their estimates θ and S. For this reason, to be rigorous, the upper limit of the confidence interval for σ should be used instead of σ when determining whether the random component of the error can be neglected, and the lower limit should be used when determining whether the systematic component can be neglected.

It should be noted that in general, the number of measurements should be chosen so that the random error of the arithmetic mean would be negligible compared with the systematic error.

In summing random and systematic errors, we neglected the absolutely constant systematic error. If this error is present, then its limits must be added arithmetically to the obtained estimate.

The method studied above pertains to measurements whose errors are estimated exactly. For measurements whose errors are estimated approximately, the properties of the measuring instruments employed are taken into account based on the specifications.

We must now make a remark regarding the problem of taking into account the intrinsic error of instruments in the case of multiple measurements. The point of doing this is to avoid taking into account the random component of this error twice. The possibility of such an overestimation is concealed in the fact that, on the one hand, the spread in the results of single measurements reflects all sources of random error, including also the random component of the intrinsic error of instruments, whereas on the other hand, when estimating the elementary component contributed by the intrinsic error of the instrument, this component also enters the calculation as part of the intrinsic error.

To avoid double counting the random component of the intrinsic error of the instrument, it must be eliminated from the intrinsic error when estimating the limits of the corresponding elementary error; it is more difficult to remove it from

the spreads in the results of single measurements or the parameter characterizing them.

In the case when digital instruments are employed, the problem can be solved comparatively simply. For this reason, it is sufficient to drop into the calculation the second term $\pm(b+q)$ in the formula or, if the error is expressed according to formula (2.2), replace δ by $\delta' = c - d$.

In the case of analog instruments, this problem must be solved depending on the properties of the particular instruments. For electric measuring instruments, for example, the random component of the intrinsic error is standardized together with the intrinsic error by prescribing the permissible limits of the dead band. As pointed out above, the limit of random error is equal to one half the dead band. When the corresponding elementary error of a multiple measurement is estimated, the limit of intrinsic error of the instrument must be reduced by one half the limit of the dead band.

Naturally, the elementary error need not be reduced, but this random component can be taken into account when calculating the variance or the standard deviation of the measurement result. For this reason, the variance of the random component must be calculated from the random component of the intrinsic error of the instrument, and then its value must be subtracted from the estimate of the variance of the measurement result.

We shall return to the problem of combining the systematic and random components of the uncertainty. This problem is significantly simplified, if the systematic component has many components and it can be assumed that it has a normal distribution. For this reason, in practice, it is sufficient that the systematic component consist of five or more elementary errors with approximately the same limits. If the number of measurements exceeds 20, then when calculating the limits of the systematic and random components, the quantile coefficient will be the same for the same confidence probability. We shall have $\Psi_\alpha = t S_{\bar{x}}$ and $\theta_\alpha = t S_\vartheta$. As before, $S_c^2 = S_\vartheta^2 + S_{\bar{x}}^2$ or $t^2 S_c^2 = t^2 S_\vartheta^2 + t^2 S_{\bar{x}}^2$. From here, it follows that

$$t S_c = \sqrt{(t S_\vartheta)^2 + (t S_{\bar{x}})^2}$$

or because $u_c = t S_c$,

$$u_c = \sqrt{\theta_\alpha^2 + \Psi_\alpha^2}. \tag{5.16}$$

Clearly, this uncertainty corresponds to the same confidence probability for which Ψ_α and θ_α were calculated.

Now we shall discuss the absolutely constant elementary errors. Their limits are added to the sum of the other errors using the same procedure as that studied above for single measurements:

$$u_t = \sum_{f=1}^{k} H_f + u_c, \tag{5.17}$$

where $\{H_f\}$, $f = 1, ..., k$, are the limits of the absolutely constant elementary errors.

We note that if some elementary errors have unsymmetric limits, then they are represented by symmetric limits after being shifted relative to the result of measurement by a_j. The calculations are performed by the same method as in the case of single measurements. We recall once again that the biases a_j cannot be compensated by introducing corrections: The error estimates are too unreliable to change the measurement result.

We shall now make some notes regarding multiple measurements with quasirandom errors. The main problem here is how to construct the composition of the distributions of the quasirandom and systematic errors. The distribution of the quasirandom error is represented by a histogram, which is constructed from the experimental data. The distribution of the systematic error, however, is most often obtained in an analytic form, i.e., in the form of an equation. The latter equation, however, is easily transformed into a histogram, after which the problem of constructing the composition of two histograms is solved by the numerical method studied in Chapter 3.

Finally, we note that when measuring average quantities (which is precisely when quasirandom errors can appear), the systematic errors are often negligibly small, which in general eliminates me problem of combining errors.

5.7. Comparison of Different Methods for Combining Systematic and Random Errors

The foregoing method for combining systematic and random errors is not the only method.

(1) The U.S. National Institute of Standards and Technology (NIST) gives in [21] the formula

$$u = \theta + \Psi_\alpha, \tag{5.18}$$

where $\theta = \sqrt{\sum_{i=1}^{m} \theta_i^2}$, if $\{\theta_i\}$, $i = 1, \ldots, m$, are independent systematic components, and $\theta = \sum_{i=1}^{m} \theta_i$, if they are dependent, and $\Psi_\alpha = t_q S_{\bar{x}}$.

First, [21] is a working document of NIST, an organization that is especially interested in the problems of checking and calibrating measuring instruments. Measurements in this case can have absolutely constant elementary errors more often than in other cases. If such components predominate, then formula (5.18) is justified. But this method of calculating errors cannot be extended to arbitrary measurements, because in most cases, it results in overestimation of the uncertainty. These formulas are also mentioned in [3].

It is necessary to note that NIST issued in l994 Guidelines [51], where combined uncertainty is calculated in accordance with the Guide [1].

(2) The standard reference [3] and the manual [14] preceding it give two different formulas for calculating the uncertainties with confidence probabilities of 0.95 and 0.99:

$$u_{c,0.99} = \theta + t_{0.95} S_{\bar{x}}, \qquad u_{c,0.95} = \sqrt{\theta^2 + (t_{0.95} S_{\bar{x}})^2}.$$

The coefficient $t_{0.95}$ is chosen according to Student's distribution in both cases for the confidence probability $0.95(q = 0.05)$ and $\nu = n - 1$.

The formulas appear to be strange. They are not related with probabilistic relations, whereas they have a stated confidence probability of 0.99 or 0.95 to the result.

(3) Before the Guide [l] was published, the working group ISO/TAG4/ WG3 issued several drafts of it. The Fourth Draft contained a new method for measurement uncertainty calculation.[1] This method was reflected in [19].

The elementary systematic errors are regarded as uniformly distributed random quantities. However, the limit of their sum is calculated with the formula $\theta = \sqrt{\sum_{i=1}^{n} \theta_i^2}$, i.e., without using the indicated model.

The systematic and random errors are combined with a formula that is essentially the same as (5.13). The only difference lies in the coefficient t_c. Here the coefficient is found from Student's distribution corresponding to the selected confidence probability and the effective degree of freedom ν_{eff}. The following formula is given to calculate ν_{eff}:

$$\frac{S_c^4}{\nu_{\text{eff}}} = \frac{S_{\bar{x}}^4}{\nu} + \sum_{i=1}^{m} \left(\frac{\theta_i^2}{3}\right)^2 .$$

It is assumed here that the random component of uncertainty has a degree of freedom $\nu = n - 1$, and each component of the systematic error has a degree of freedom equal to one. However, the notion of a degree of freedom is not applicable to random variables with a fully defined distribution function. Therefore it is impossible to assume that a quantity with uniform distribution within given limits has a degree of freedom equal to one (or to any other finite number). Thus, the formula under discussion is not mathematically grounded.

(4) The Guide's [l] method. In general, this method is the same method that was described in the Fourth Draft (and in other drafts). But the shortcoming of the Fourth Draft is avoided here as suming the coefficient t_c to be constant: $t_c' = 2$ for $\alpha = 0.95$ and $t_c'' = 3$ for $\alpha = 0.99$.

(5) Method proposed in this book. The resulting formulas are (5.13) and (5.14).

We shall compare all methods enumerated above for summing the systematic and random errors in two numerical examples.

Suppose that as a result of some measurement the following indicators of its errors were obtained:

$$S_{\bar{x}} = 1, \qquad n = 16, \qquad \theta = 3.$$

Suppose also that the random errors have a normal distribution and that the systematic errors have a uniform distribution. Then for the exact solution we can take

[1] "Guide to the Expression of Uncertainty in Measurement; Fourth Draft." July 13, 1990. ISO/TAG4/WG3.

the confidence limits presented in Table 5. 1. As usual, we shall take $\alpha_1 = 0.95$ and $\alpha_2 = 0.99$. Then

$$u_{T,0.99} = 4.49, \qquad u_{T,0.95} = 3.67.$$

Here there is an error: We assumed that $S_{\bar{x}} = \sigma_{\bar{x}}$. But for $n = 16$, this error is not significant, and we shall neglect it.

We shall present the computational results obtained using all of the methods examined above.

(1) *Reference* [21]. The coefficients of Student's distribution with $\nu = n - 1 = 15$ and the indicated values of the confidence probabilities will be as follows:

$$t_{0.99}(15) = 2.95, \qquad t_{0.95}(15) = 2.13,$$
$$\Psi_{0.99} = 2.95 \times 1 = 2.95, \qquad \Psi_{0.95} = 2.13 \times 1 = 2.13.$$

Therefore, $u_{1,0.99} = 3 + 2.95 = 5.95$ and $u_{1,0.95} = 3 + 2.13 = 5.13$.

(2) *Reference* [3]. We shall make use of the calculations that were just performed:

$$u_{2,0.99} = 3 + 2.13 \times 1 = 5.13, \qquad u_{2,0.95} = \sqrt{3^2 + (2.13)^2} = 3.68.$$

(3) *Reference* [19].

$$S_{\vartheta}^2 = 9/3 = 3, \qquad S_{\vartheta} = 1.73,$$
$$S_c^2 = 1 + 3 = 4, \qquad S_c = 2.$$

We shall calculate the effective number of degrees of freedom:

$$\frac{4^2}{\nu_{\text{eff}}} = \frac{1}{15} + 3^2, \qquad \frac{16}{\nu_{\text{eff}}} = 9.07, \qquad \text{and} \qquad \nu_{\text{eff}} = 2.$$

Next, we find from Student's distribution $t_{3,0.99} = 9.9$ and $t_{3,0.95} = 4.3$. Correspondingly, we obtain

$$u_{3,0.99} = 9.9 \times 2 = 19.8, \qquad u_{3,0.95} = 4.3 \times 2 = 8.6.$$

(4) *Reference* [1]. We have, in this case, $S_c = \sqrt{S_{\bar{x}}^2 + S_{\vartheta}^2} = \sqrt{1 + 3} = 2.0$. Because $t_{0.99} = 3$ and $t_{0.95} = 2$, we obtain

$$u_{4,0.99} = 3 \cdot 2 = 6, \qquad u_{4,0.95} = 2 \cdot 2 = 4.$$

(5) *The formulas* (5.13) and (5.14) *give* $S_{\vartheta} = 1.73$ *and* $S_c = 2.0$.

$$t_{5,0.99} = \frac{2.95 \times 1 + 0.99 \times 3}{1 + 1.73} = \frac{5.92}{2.73} = 2.17,$$

$$t_{5,0.95} = \frac{2.13 \times 1 + 0.95 \times 3}{1 + 1.73} = \frac{4.98}{2.73} = 1.82,$$

$$u_{5,0.99} = 2.17 \times 2 = 4.34, \qquad u_{5,0.95} = 1.82 \times 2 = 3.64.$$

TABLE 5.5. Errors of different methods of uncertainty calculation, for example, with $\theta = 3$, $\delta_x = 1$, $n = 16$.

	$(u_i - u_T)/u_T \times 100\%$	
Method of computation	$\alpha = 0.99$	$\alpha = 0.95$
1	32	39.0
2	14	0.3
3	340	132.0
4	34	6.0
5	3	0.8

TABLE 5.6. Errors of different methods of uncertainty calculation, for example, with $\theta = 0.5$, $\delta_x = l$, $n = 16$.

	$(u_i - u_T)/u_T \times 100\%$	
Method of computation	$\alpha = 0.99$	$\alpha = 0.95$
1	29	30
2	2	7
3	13	8
4	12	2
5	4	3

We shall compare the estimated uncertainties with the exact values $u_{T,0.99}$ and $u_{T,0.95}$ initially presented for the corresponding confidence intervals. The results are summarized in Table 5.5. The errors for the case $\theta = 0.5$ and the previous values $S_{\bar{x}} = 1$ and $n = 16$ were calculated analogously. The results are presented in Table 5.6. In comparison with the previous example, method 4 and especially method 3 in this case show a significant decrease in error. It is not surprising because now the systematic component is insignificant relative to the random component.

The examples presented show the following:

(a) As expected, the method of [19] cannot be used when the systematic error is significant.[1]
(b) The method from the standard reference [3], irrespective of the remarks made above, gave in both examples satisfactory results.

[1] The shortcomings of this method were discussed in the report "The U.S.A. and the U.S.S.R. Standards for Measurement Uncertainty" given by S. Rabinovich at the Measurement Science Conference in Anaheim, CA, January 31 and February 1, 1991.

(c) The method of [21], as expected, gave in the examples studied estimates that were too high.
(d) Method 4 is good if the systematic component is small relative to the random component.
(e) The formulas (5.13) and (5.14) gave the best results in both examples.

Examples are not, of course, proof, but they nonetheless illustrate well the consideration stated above.

5.8. Essential Aspects of the Estimation of Measurement Errors when the Number of Measurements Is Small

In practice, measurements are often performed with a very small number of observations, for example, two or three observations. This amount is not enough for statistical analysis, and such measurements cannot be called multiple measurements. At the same time, they are also not single measurements.

What is the point of this seemingly strange choice of the number of measurements? Analysis shows that this small amount checks the suitability of the model selected for the object of study. For example, when the diameter of a shaft is being measured, it is measured at several locations along the shaft and in different directions. If the conditions under which the measurements are performed are sufficiently stable and the model (cylinder) corresponds to the object (shaft), then the differences between the single-measurement results should be small, and in any case less than twice the error of a single measurement. If the difference between the measurements is large, then it is pointless to use the selected measurement method. If the difference between the measurements is small, then the question arises of what should be regarded as the result of a measurement and what is its error.

In this case, in principle, the result of any single measurement can be used as the result of the measurement. But because the measurements have already been performed, it would still be nice to use them somehow. For this reason, the arithmetic mean of the observational results is used as the result of measurement. In this respect, this case is no different from the case of multiple measurements. However, the situation is different with regard to the measurement uncertainty. Here two cases must be borne in mind.

(1) The difference between the observational results is insignificant, i.e., three or more times less than the limit of error. This result means that the model corresponds well to the object and that the random component of the error of the observational results is small. Regarding systematic errors, because the conditions under which the observations are performed are constant and the measuring instruments are not changed, the systematic error will be the same for all measurements.

For this reason, in this case, the measurement error is also the error of a single measurement.

(2) The difference between the observational results is significant. Let us see how the measurement error can be estimated in this case.

The difference between the largest and the smallest values of n observations is called in mathematical statistics the range R_n of the observations in a sample. The distribution function of the range is tabulated for normally distributed random quantities. Assuming that the random error of a single measurement has a normal distribution, we can write

$$P\{R_n \le a\} = \alpha,$$

where a is the limit chosen for R_n. Next, assume that we have three single measurements and that $\alpha = 0.95$.

From the table,[1] we find, for $\alpha = 0.95$ and $n = 3$, that $a = 3.3\sigma$ or $\sigma = a/3.3$.

Let $a = b\Delta$, where Δ is the limit of error of a single measurement. But for $\alpha = 0.95$, the limit of random error can be estimated as

$$\Psi_1 = 1.96\sigma = \frac{1.96}{3.3}a = 0.594b\Delta$$

Let the systematic error consist of more than four components. Then

$$\Delta = \sqrt{\theta_2 + \Psi_1^2}, \qquad \theta = \sqrt{\Delta^2 - \Psi_1^2} = \Delta\sqrt{1 - 0.353b^2}.$$

The radicand must be greater than 0. Then the maximum value of the coefficient is $b_m = 1.7$.

We shall now consider the error of the arithmetic mean. The standard deviation in this case will be $\sigma_{\bar{x}} = \sigma/\sqrt{n} = \sigma/\sqrt{3}$. From here, we find the limiting value of the random error:

$$\Psi_{0.95} = t_{0.95}\sigma_{\bar{x}} = \frac{\Psi_1}{\sqrt{3}} = \frac{0.594}{\sqrt{3}}b\Delta.$$

The limiting value of the systematic error has already been estimated:

$$\theta = \Delta\sqrt{1 - 0.353b^2}.$$

For the total measurement error, we obtain from here the expression

$$\Delta_{\bar{x}} = \sqrt{\theta^2 + \Psi_\alpha^2} = \Delta\sqrt{(1 - 0.353b^2) + \left(\frac{0.594b}{\sqrt{3}}\right)^2} = \Delta\sqrt{1 - 0.235b^2}.$$

[1] See, for example, A. Hald, *Statistical Theory With Engineering Applications* (Wiley, New York, 1952).

Now we can estimate the decrease in the error of the arithmetic mean compared with the error of a single measurement with $n = 3$ measurements: $\mu = \Delta\bar{x}/\Delta = \sqrt{1 - 0.235b^2}$

$$\begin{aligned} b &= 0.5, \qquad \mu = 0.97, \\ b &= 1.0, \qquad \mu = 0.87, \\ b &= 1.5, \qquad \mu = 0.69, \\ b &= 1.7, \qquad \mu = 0.57. \end{aligned}$$

The case $b = 1.7$ means that the entire error of a single measurement is determined by the random component. In this case, the measurement should be designed as a multiple measurement.

For $b = 1.0$, we have $\mu = 0.87$; i.e., the error in the measurement result is equal to the error in a single measurement.

Thus, if we assume that the range, i.e., the largest difference between the measurements, for $n = 3$ can reach the limit of permissible error of a single measurement, then the error of the arithmetic mean will be approximated the same as the error of a single measurement. In addition, requiring that the limit of error of a single measurement be the smallest fraction of the range will not significantly reduce the estimate of the error of the arithmetic mean.

5.9. General Plan for Estimating Measurement Uncertainty

The purpose of this section is to give, without getting into details, an overall plan for estimating the uncertainties of direct measurements. This plan should help the reader concentrate on the essential points of each step in me solution of the problem.

1. Analyze the initial data.

1.1. Study the measurement problem. For this, one must first get an idea of the object whose parameter is being measured, the purpose of the measurement, and the required measurement accuracy. In connection with these questions, it is necessary to determine a model of the object and to try to check that the parameter to be measured (the measured quantity) corresponds to the required measurement accuracy. Next, it is necessary to write out the physical quantities characterizing the surrounding environment and affecting the size of the measured parameter, to estimate their nominal values and range of variation, and to determine how these quantities must be measured, if the measurement is being planned, or were measured, if the measurement has already been performed.

1.2. Establish which of the metrological properties of the measuring instruments chosen for the measurement, or already employed in performing the measurement, are important for the given measurement.

2. Prepare the data for the calculations.

2.1. Compile a list of all possible elementary errors in the given measurement.

2.2. Estimate the limits of all elementary errors. Express them in the form of absolute or relative errors and scale to the input of the measuring apparatus, instrument, or channel.

2.3. Determine whether it is useful to introduce corrections and the possibility of obtaining point estimates of the corresponding elementary errors necessary for this. This question must first be resolved for the dominant elementary errors. Determine the corrections to be made. Estimate the limits of inaccuracy of each correction, and add them to the list of elementary errors.

2.4. Check the independence of the elementary errors. If the errors ε_1 and ε_2 from different causes depend on some third physical quantity, then these errors will be dependent. To eliminate this dependence, it is often sufficient to introduce a new elementary error that reflects the effect of this third quantity on the result of measurement. Then, instead of ε_1 and ε_2, we shall have new elementary errors ε_1' and ε_2', which can now be regarded as being independent.

2.5. Divide all elementary errors into conditionally and absolutely constant errors, and single out those errors whose limits differ in absolute magnitude, i.e., are unsymmetric relative to the result of measurement. If the measurement is multiple, then it is necessary to determine whether its error is purely random or quasirandom. Estimate the confidence limits of this error.

2.6. Estimate the quantities necessary for calculating the additional errors. For this reason, it is desirable to measure these quantities; this process is necessary if one intends to introduce the corresponding corrections. In the case when the limiting value of the additional error is estimated, it is usually sufficient to have the limiting value of the influential quantity.

2.7. Estimate the possible change in the intrinsic error of the instruments over the time period since they were calibrated. If there are grounds for assuming that the intrinsic error could have exceeded permissible values, then such instruments must be rechecked before performing the measurement and, if necessary, adjusted or recalibrated.

3. Calculate the result of measurement.

3.1. In the case of single measurements, the result of a measurement is often obtained directly from the indication of the measuring instrument, and no calculations are required for this. Sometimes, however, the indication of an instrument in units of the scale graduations must be multiplied by the scale factor, corrections must be introduced, and other nonspecific calculations must be performed. In the case of multiple measurements, the arithmetic mean is usually taken as the result of measurement. However, a different algorithm, determined by the definition of the measured quantity, can also be used. The corrections, if they are the same for all single measurements, can be introduced in the arithmetic mean and not in the result of each measurement.

3.2. A priori estimation of error or uncertainty is usually made for the least favorable case. If multiple measurement is planned, then the possible value of the

standard deviation is taken based on recommendations of experts. The methods for performing the calculations were presented in this chapter.

3.3. A posteriori estimation of error or uncertainty is performed using the methods presented in this chapter.

3.4. The form in which the results of measurement are presented and their error or uncertainty were presented in Chapter 1.

6
Indirect Measurements

6.1. Basic Terms and Classification

Indirect measurement is a measurement in which the value of the unknown quantity sought is calculated using measurements of other quantities related to the measurand by some known relation. We shall call these other quantities *measurement arguments* or, briefly, *arguments*.

In an indirect measurement, the true value of a measurand A is related to the true values of arguments $A_j (j = 1, \ldots, N)$ by a known function f. This relationship can be represented in a general form as

$$A = f(A_1, \ldots, A_N). \tag{6.1}$$

This equation is called a *measurement equation*. The specific forms of measurement equations can be considered as mathematical models of specific indirect measurements.

Various classifications of indirect measurement are possible. We shall limit ourselves to classifications that will be useful for our purposes.

From the perspective of conducting a measurement, we shall distinguish *single* and *multiple indirect measurements*. In single measurements, all arguments are measured once. In multiple measurement, all arguments are measured several times.

According to the form of the functional dependency (6.1), we shall distinguish *linear* and *nonlinear indirect measurements.* In the case of a linear indirect measurement, the measurement equation has the form

$$A = b_0 + \sum_{j=1}^{N} b_j A_j, \tag{6.2}$$

where $\{b_j\}$ are constant coefficients. Nonlinear indirect measurements are diverse, and therefore, it is impossible to represent all of them with one general form of measurement equation.

The physics of the processes of indirect measurements gives us another important classification criterion. To motivate this classification, let us compare the

accurate measurement of the density of a solid, and the measurement of the temperature coefficient of the electrical resistance of a resistor.

To measure the density of a solid, its mass and volume should be measured independently, with consistent accuracy.

In the temperature coefficient measurement, the resistance of the resistor and its temperature are measured simultaneously, which means that the measurements of these arguments are not independent. Thus, we shall distinguish *dependent* and *independent indirect measurements*.

Indirect measurements, just like direct measurements, are divided into static and dynamic. Static indirect measurements can be different depending on the properties of the measured arguments. If the measured arguments can be regarded as being constant in time, then the indirectly measured quantity is also constant; i.e., we have the usual static situation.

However, the measured quantity can also be constant when the arguments vary. For example, suppose we are measuring the resistance of a resistor by the ammeter and voltmeter method, and the voltage of the source changes in time according to the conditions of the measurement. Although the measured arguments change, the measured quantity remains unchanged.

To obtain the correct result in the case under study, the arguments must be measured with instruments such that the arguments do not change significantly over the time interval during which the indications of the instruments settle down.

6.2. Correlation Coefficient and its Calculation

The traditional methods for estimating the uncertainty of indirect measurements, as pointed out in the Guide [1], include the calculation of the correlation coefficient. Later in this book, we shall develop a new theory, which obviates any need for the correlation coefficient. However, given the traditional importance of the correlation coefficient and a great deal of confusion in metrology with its calculation,[1] it makes sense to describe here a clear and correct procedure for practical calculation of the correlation coefficient.

The mathematical foundation and methods of the correlation coefficient calculations can be found in many books on the Theory of Probability and Mathematical Statistics. I refer to *Mathematical Statistics* by B.L. van der Waerden [52].

Consider two random quantities X and Y with mathematical expectations equal to zero ($E[X] = 0$ and $E[Y] = 0$) and finite variances. Denote their variances as $V[X] = \sigma_X^2$ and $V[Y] = \sigma_Y^2$.

The variance of a random quantity $Z = X + Y$ can be calculated using the equation

$$V[Z] = E[(X + Y)^2] = E[X^2] + E[Y^2] + 2E[XY]. \tag{6.3}$$

The last term $E[XY]$ is named *second mixed moment or covariance.*

[1] I agree with R.H. Dieck that "probably one of the most misunderstood and misused statistics is the correlation coefficient" [22].

The covariance divided by the square root of the product of variances $\sigma_X^2\sigma_Y^2$ gives the correlation coefficient ρ_{XY}:

$$\rho_{XY} = \frac{E[XY]}{\sigma_X\sigma_Y}.$$

The value of the correlation coefficient always lies within $[-1, +1]$, and if $|\rho_{XY}| = 1$, then there is a linear functional dependency between X and Y. When $\rho_{XY} = 0$, X and Y are uncorrelated. But it does not mean they are independent. Otherwise, when $0 < |\rho_{XY}| < 1$, the nature of the dependency between X and Y cannot be determined unambiguously: It can be stochastic as well as functional nonlinear dependency. Therefore, in the last case, if the knowledge about the nature of the dependency between X and Y is required, it can only be obtained based on physical properties of the problem rather than inferred mathematically.

From the above formulas, we obtain

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 + 2\rho_{XY}\sigma_X\sigma_Y. \tag{6.4}$$

In practice, we have to work not with the exact values of parameters of random quantities but with their estimates. So, instead of variances $\sigma_Z^2, \sigma_X^2, \sigma_Y^2$ and the correlation coefficient ρ_{XY}, we have to use their estimates S_Z^2, S_X^2, S_Y^2 (we will also use interchangeably $S^2(A)$ to denote an estimate of the variance of random quantity A), and r_{XY}. If n is the number of measured pairs (x_i, y_i) of random quantities X and $Y (i = 1, \ldots, n)$, and $\bar{x}$ and $\bar{y}$ are averages over n observed values of X and Y, then

$$S_X^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}, \quad S_Y^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}.$$

The estimate of $E[XY]$ will be

$$m_{XY} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}.$$

Then, $r_{XY} = m_{XY}/S_X S_Y$.

Thus, the formulas for calculations are

$$r_{XY} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)S_X S_Y}, \tag{6.5}$$

$$S_Z^2 = S_X^2 + S_Y^2 + 2r_{XY}S_X S_Y. \tag{6.6}$$

The estimates of the variances of the average values $\bar{x}$ and $\bar{y}$ are known to be

$$S_{\bar{x}}^2 = \frac{S_X^2}{n} \quad \text{and} \quad S_{\bar{y}}^2 = \frac{S_Y^2}{n}.$$

Then, by dividing (6.6) by n, we obtain the estimate of the variance of the mean value of Z:

$$S_{\bar{Z}}^2 = S_{\bar{x}}^2 + S_{\bar{y}}^2 + 2r_{XY}S_{\bar{x}}S_{\bar{y}}. \tag{6.7}$$

The correlation coefficient estimation here is the same as in (6.5). One can also use $S_{\bar{x}}$ and $S_{\bar{y}}$ for the calculation of the correlation coefficient estimation using the fact that $S_X S_Y = nS_{\bar{x}}S_{\bar{y}}$. Then, (6.5) will change to the following:

$$r_{XY} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n(n-1)S_{\bar{x}}S_{\bar{y}}}. \tag{6.8}$$

It is necessary to stress that the number of realizations of X and Y (e.g., the number of measurements of X and Y) must be the same. Moreover, each pair of these realizations must be obtained under the same conditions, for example, at the same time, at the same temperature, using measuring instruments with the same dynamic characteristics (in the case of dynamic measurements), and so on.

The theory of correlations says that realizations x_i and y_i must belong to the same event i. A clear illustration of this statement is given by the classic example of the accuracy analysis of firing practice. Here, each event is one shot. Each shot i is described by a pair of values x_i and y_i that express the deviation of the bullet from the center of the target in orthogonal coordinates. In the case of an indirect measurement, one event is the set of matched measurement results of all arguments. This event corresponds to a point in the multidimensional space with arguments as coordinates. We shall call this point a *measurement vector*.

In the above-mentioned example of the measurement of the temperature coefficient of the electrical resistance of a resistor, each pair of measurements of the resistance and temperature is a measurement vector.

6.3. The Traditional Method of Experimental Data Processing

The processing of experimental data obtained in a measurement consists of two steps. In the first step, we estimate the value of the measurand, and in the second step, we calculate the inaccuracy of this estimate.

In an indirect measurement, the first step traditionally is based on the assumption that the estimate $\tilde{A}$ of the measurand A can be obtained by substitution $\tilde{A}_j$ for A_j in (6.1):

$$\tilde{A} = f(\tilde{A}_1, \ldots, \tilde{A}_N). \tag{6.9}$$

The second step, also traditionally, is solved by expansion of the function (6.1) in a Taylor series.

Usually the Taylor series is written in the form of an approximate value of the given function, which is brought to its true value with the help of corrections. We want, however, to work with errors rather than with corrections. Thus, we shall therefore write the series in such a form that the approximate value of the function is expressed by adding something to its true value.

To simplify further calculation, suppose that the number of arguments $N = 2$. Then we have the Taylor series in the form:

$$f(\tilde{A}_1, \tilde{A}_2) = f(A_1, A_2) + \left(\frac{\partial}{\partial A_1}\zeta_1 + \frac{\partial}{\partial A_2}\zeta_2\right) f(A_1, A_2) + \frac{1}{2!}\left(\frac{\partial}{\partial A_1}\zeta_1 + \frac{\partial}{\partial A_2}\zeta_2\right)^2 f(A_1, A_2) + \cdots + \frac{1}{m!}\left(\frac{\partial}{\partial A_1}\zeta_1 + \frac{\partial}{\partial A_2}\zeta_2\right)^m f(A_1, A_2) + R_{m+1}, \quad (6.10)$$

where $\tilde{A}_1 = A_1 + \zeta_1$, $\tilde{A}_2 = A_2 + \zeta_2$ (ζ_1 and ζ_2, the errors of $\tilde{A}_1$ and $\tilde{A}_2$), and R_{m+1} is the remainder term.

The remainder term can be expressed in the Lagrange form:

$$R_{m+1} = \frac{1}{(m+1)!}\left(\frac{\partial}{\partial A_1}\zeta_1 + \frac{\partial}{\partial A_2}\zeta_2\right)^{m+1} f(A_1 + v_1\zeta_1, A_2 + v_2\zeta_2), \quad (6.11)$$

where $0 < v_{1,2} < 1$.

If the indirect measurement is linear, all terms, except the linear one, are equal to zero.

The general form of the error of an indirect measurement is

$$\zeta = \tilde{A} - A = f(\tilde{A}_1, \tilde{A}_2) - f(A_1, A_2).$$

Turning to the Taylor series, one obtains

$$\zeta = \left(\frac{\partial}{\partial A_1}\zeta_1 + \frac{\partial}{\partial A_2}\zeta_2\right) f(A_1, A_2) + \frac{1}{2}\left(\frac{\partial}{\partial A_1}\zeta_1 + \frac{\partial}{\partial A_2}\zeta_2\right)^2 f(A_1, A_2) + \cdots + R_{m+1}. \quad (6.12)$$

In practice, however, only the first linear term is used for error calculations:

$$\zeta = \frac{\partial f}{\partial A_1}\zeta_1 + \frac{\partial f}{\partial A_2}\zeta_2.$$

The partial derivatives above are customarily called *influence coefficients.* We shall denote them as follows:

$$w_j = \frac{\partial f}{\partial A_j}, \qquad j = 1, \ldots, N.$$

Now the above equation can be written in the general form:

$$\zeta = \sum_{j=1}^{N} w_j \zeta_j. \tag{6.13}$$

Note that all partial derivatives are calculated for the estimates $\tilde{A}_1$, $\tilde{A}_2$ because the true values A_1, A_2 are unknown.

Putting aside for a while absolutely constant errors, we can write

$$\zeta_j = \vartheta_j + \psi_j,$$

where ϑ_j and ψ_j are conditionally constant and random components of the error, respectively. So, (6.13) takes the form:

$$\zeta = \sum_{j=1}^{N} w_j \vartheta_j + \sum_{j=1}^{N} w_j \psi_j. \tag{6.14}$$

The last formula says that, in indirect measurements, not only the systematic error consists of components, but so also does the random error.

We shall consider in this section the random errors only, which means there are no systematic errors in the argument estimation, i.e., that $E|\zeta_1| = 0$ and $E|\zeta_2| = 0$. We shall take these errors into account later, in Section 6.7.

The most important characteristic of a random error is its variance. In accordance with the mathematical definition of the variance, we obtain from (6.13), for $N = 2$,

$$V[\zeta] = E[(w_1\zeta_1 + w_2\zeta_2)^2] = w_1^2 E\left[\zeta_1^2\right] + w_2^2 E\left[\zeta_2^2\right] + 2w_1 w_2 E[\zeta_1 \times \zeta_2].$$

This formula is different from (6.3) only in the notations. Therefore, one can write

$$\sigma^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2\rho_{1,2} w_1 w_2 \sigma_1 \sigma_2, \tag{6.15}$$

where

$$\sigma^2 = V[\zeta] = E[\zeta^2], \quad \sigma_1^2 = E\left[\zeta_1^2\right],$$
$$\sigma_2^2 = E\left[\zeta_2^2\right], \quad \text{and} \quad \rho_{1,2} = \frac{E[\zeta_1 \times \zeta_2]}{\sigma_1 \sigma_2}.$$

We should like to point out that the variance of a random error of the measurement result is identical to the variance of the measurement result:

$$V[\zeta] = V[\tilde{A}].$$

Also note that (6.15) has three items, which corresponds to the case when $N = 2$. When $N = 3$, we shall have six items. So, with the number of arguments increasing, the complexity of calculations increases rapidly.

In (6.15), the values of variance σ_j^2 and correlation coefficient $\rho_{k,l}$ are unknown and, in practice, their estimations S_j^2 and $r_{k,l}$ are used instead. Taking into account this substitution and assuming the general case of N arguments, (6.15) becomes

$$S^2(\tilde{A}) = S_r^2 = \sum_{j=1}^{N} w_j^2 S^2(\tilde{A}_j) + 2\sum_{k<l} r_{k,l} w_k w_l S(\tilde{A}_k) S(\tilde{A}_l). \tag{6.16}$$

The notation S_r^2 is introduced to stress that this estimate of variance reflects the random errors only.

For estimating the variance and correlation coefficient, we have the formulas

$$\left.\begin{aligned} S_j^2 = S^2(\tilde{A}_j) &= \frac{1}{n(n-1)} \sum_{i=1}^{n} (x_{ji} - \bar{x}_j)^2, \\ r_{k,l} &= \frac{\sum_{i=1}^{n} (x_{ki} - \bar{x}_k)(x_{li} - \bar{x}_l)}{n(n-1) S(\tilde{A}_k) S(\tilde{A}_l)}. \end{aligned}\right\} \tag{6.17}$$

Here, n is the number of measurement vectors. The fact that realizations x_{ki} and x_{pi} have the same subscript i means that these realizations must be taken from the same vector i.

If measurements of all arguments are independent, i.e., $\rho_{k,l} = 0$, then (6.16) is simplified:

$$S_r^2 = \sum_{j=1}^{N} w_j^2 S^2(\tilde{A}_j).$$

This equation gives the following expression for the standard deviation:

$$S_r = \sqrt{w_1^2 S^2(\tilde{A}_1) + \cdots + w_N^2 S^2(\tilde{A}_N)}. \tag{6.18}$$

The last two formulas are often called the *error propagation formulas*, although in reality they express the propagation of variances.

Although (6.18) was deduced for the random errors only, it have a wide use as universal formula for the summation of all kinds of errors. This way of error calculation has even a specific name: the square-root sum method.

The next problem is to calculate the confidence interval for the true value of the measurand. This problem can be solved for linear indirect measurements. For nonlinear measurements, it cannot be solved because in general it is impossible to form the error distribution for these measurements.

Consider the linear indirect measurement. The spread in the observations from the random error in the measurement of each argument can usually be regarded as a normally distributed random quantity. But even if one of these distributions must be assumed to be different from a normal distribution, the distribution of the arithmetic mean, in practice, can still be regarded as being normal. The random error in the result of an indirect measurement, which is determined by adding the random errors in the measurements of the arguments, can be regarded with even greater justification to be a normally distributed random quantity.

If the number of measurements of each argument is greater than 25–30, then after choosing the confidence probability α, the limit of the confidence interval will be

$$|\Psi_\alpha| = z_{\frac{1-\alpha}{2}} S_r,$$

where $z_{\frac{1-\alpha}{2}}$ is the quantile of the standardized normal distribution. Its value can also be found in the table of Student's distribution, in the line with the infinite number of degrees of freedom.

A difficulty arises when the number of observations is less than 25–30. In this case, in principle, one could use Student's distribution, but an exact expression for the degrees of freedom is not known. An approximate solution, which gives an estimate of the degrees of freedom, called the effective number of degrees of freedom, is given by the well-known Welch–Satterthwaite formula [3],

$$\nu_{\text{eff}} = \frac{\left(\sum_{i=1}^{m} b_i^2 S_i^2(\tilde{A}_i)\right)^2}{\sum_{i=1}^{m} \frac{b_i^4 S_i^4(\tilde{A}_i)}{\nu_i}}, \tag{6.19}$$

where $\nu_i = n_i - 1$. The limit Ψ_α in this case can be calculated as

$$|\Psi_\alpha| = t_q S_r,$$

where t_q is found from Student's distribution table for the degree of freedom ν_{eff} and the chosen confidence probability $\alpha = 1 - q$.

We have mentioned above that confidence intervals in general cannot be calculated for nonlinear indirect measurements. Therefore, in practice uncertainty is often calculated by the square-root sum formula

$$u_t = \sqrt{\sum_{j=1}^{N} w_j^2 u_j^2}, \tag{6.20}$$

where u_j is the uncertainty of the measurement of jth argument and w_j is the influence coefficient. But the square-root sum formula is correct for summing variances, not intervals. Therefore, obtained in this way, it is difficult to call this result a confidence interval or uncertainty. But (6.20) is correct only if measurement errors of all arguments after Taylor series expansion are normally distributed and all uncertainties u_j have been calculated for the same confidence probability. The proof of this statement is obvious from the discussion of formula (5.8).

The next problem is how to calculate the systematic error of an indirect measurement result, and how to combine it with the random error to obtain the uncertainty of the indirect measurement result. A reasonable solution of this problem will be discussed below in Section 6.7.

6.4. Shortcomings of the Traditional Method

The traditional method has been used for a long time. But as the analysis presented in [46] showed, it has a series of shortcomings.

First, for a nonlinear function f

$$E[f(X_1, \ldots, X_N)] \neq f(E[X_1], \ldots, E[X_N]),$$

where $X_1, \ldots, X_N$ are random quantities. Therefore the estimate of the measurand given by (6.9) is incorrect when the measurement equation is nonlinear. Let us evaluate this incorrectness.

Go back to (6.10) and now retain not only the first term but the second one also. Again, assuming $N = 2$ for simplicity, we get

$$\zeta = \left(\frac{\partial f}{\partial A_1}\zeta_1 + \frac{\partial f}{\partial A_2}\zeta_2\right) + \frac{1}{2}\left(\frac{\partial}{\partial A_1}\zeta_1 + \frac{\partial}{\partial A_2}\zeta_2\right)^2 f(A_1, A_2).$$

Assume, as before, ζ_1 and ζ_2 to be free from systematic errors: $E[\zeta_1] = 0$ and $E[\zeta_2] = 0$. Then the mathematical expectation of the first term is equal to zero:

$$E\left[\left(\frac{\partial f}{\partial A_1}\zeta_1 + \frac{\partial f}{\partial A_2}\zeta_2\right)\right] = w_1 E[\zeta_1] + w_2 E[\zeta_2] = 0.$$

But the variances of the errors ζ_1 and ζ_2

$$V[\zeta_1] = \sigma_1^2 > 0 \qquad \text{and} \qquad V[\zeta_2] = \sigma_2^2 > 0,$$

and therefore the mathematical expectation of the second term is not equal to zero. Indeed

$$\begin{aligned} E[\zeta] &= E\left[\frac{1}{2}\left(\frac{\partial}{\partial A_1}\zeta_1 + \frac{\partial}{\partial A_2}\zeta_2\right)^2 f(A_1, A_2)\right] \\ &= \frac{1}{2}\frac{\partial^2 f}{\partial A_1^2} E\left[\zeta_1^2\right] + \frac{1}{2}\frac{\partial^2 f}{\partial A_2^2} E\left[\zeta_2^2\right] + \frac{\partial f}{\partial A_1}\cdot\frac{\partial f}{\partial A_2} E[\zeta_1 \times \zeta_2] \\ &= \frac{1}{2}\frac{\partial^2 f}{\partial A_1^2}\sigma_1^2 + \frac{1}{2}\frac{\partial^2 f}{\partial A_2^2}\sigma_2^2 + \frac{\partial f}{\partial A_1}\cdot\frac{\partial f}{\partial A_2}\rho_{1,2}\sigma_1\sigma_2. \end{aligned} \tag{6.21}$$

As $\sigma_1^2 > 0$, $\sigma_2^2 > 0$ and $|\rho_{1,2}| < 1$, $E[\zeta] = B \neq 0$.

Thus, for nonlinear indirect measurements, the estimate of the measurand given by the traditional method is biased! The bias of the measurement result can be reduced by correction C:

$$C = -B.$$

But even after correction, the estimate of a measurand will not be exact because it takes into account only two terms, whereas the Taylor series may have an infinite number of terms.

This is the first deficiency of the traditional theory of indirect measurements. It must be considered as an essential disadvantage for it affects the results of measurements.

The second deficiency is as follows. The estimate of the variance of the measurement result, given by (6.16), is imperfect because it was derived using only one linear term in the Taylor series. In other words, the traditional theory does not use all of the information contained in the results of measurements of arguments.

The next disadvantage of the traditional theory is the problem of the confidence intervals. As a matter of fact, this theory does not provide solid methods to construct

the confidence intervals for the true value of a measurand in the case of nonlinear measurement equation.

An additional problem is the above-mentioned problem of estimating correlation coefficients that are a part of the traditional method.

6.5. The Method of Reduction

The essence of the method of reduction is as follows. Assume that $x_{1i}, x_{2i}, \ldots, x_{Ni}$ are measurement results of arguments from a measurement vector i. Recall that a measurement vector compiles measurements of all arguments performed under the same conditions and at the same time. Each dependent indirect measurement always consists of a definite number of measurement vectors.

So, let n be the number of measurement vectors obtained. These vectors can be represented as a set:

$$\{x_{1i}, x_{2i}, \ldots, x_{Ni}\}, \qquad i = 1, \ldots, n.$$

Substituting the results from the ith vector into the measurement equation, we obtain the ith value of the measurand. Denote it by y_i. This transformation is obviously justified because it reflects the physical relationship between a measurand and measurement arguments.

In the same way, n measurement vectors give us a set of n values of the measurand:

$$\{y_i\}, \qquad i = 1, \ldots, n.$$

This set does not differ from a set of data obtained by direct measurements of the measurand A. Hence, we can now use all simple and well-understood methods of direct measurements, which immediately provides an estimate of the measurand

$$\tilde{A} = \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i, \tag{6.22}$$

and an estimate of the variance

$$S^2(\tilde{A}) = S_r^2 = \frac{1}{n(n-1)} \sum_{i=1}^{n} (y_i - \bar{y})^2. \tag{6.23}$$

The method of reduction also solves the problem of the calculation of confidence intervals, because we now have the set of n values of the measurand. The limits of a confidence interval and therefore the uncertainty of the measurement result are

$$u = t_q S(\tilde{A}), \tag{6.24}$$

where t_q is found from Student's distribution for the chosen confidence probability and obtained exact number of degrees of freedom

$$\nu = n - 1.$$

One might think that the method of reduction imposes special requirements for performing the measurement, namely that the measurements of arguments be performed so that the results can be represented as a number of measurement vectors. However, the traditional method imposes this requirement as well. Indeed, if we have a dependent indirect measurement, all arguments must be measured under the same conditions for the traditional method also, because, otherwise, it is impossible to calculate the correlation coefficients and therefore impossible to estimate the variance of the measurement result.

Thus, the method of reduction has some important advantages over the traditionally used method:

(1) It produces an unbiased estimate of the measurand.
(2) It uses all of the information obtained in the course of the measurement for estimation of the variance of the indirect measurement results.
(3) It gets rid of the correlation coefficient in the measurement uncertainty calculations.
(4) It gives the exact number of degrees of freedom and allows us to calculate the confidence intervals for the true value of the measurand.

The listed advantages lead us to conclude that the method of reduction is the preferable method for all kinds of dependent indirect measurements.

It is important to emphasize here that the independent indirect measurements do not need correlation coefficients. As the method of reduction eliminates the need for correlation coefficients in the case of dependent indirect measurements, *the concept of the correlation coefficient is no longer necessary in measurement data processing.*

To conclude, I would like to note that I first proposed this method of reduction approximately in 1970. It found immediate application in national and international comparisons of standards of unit radium mass and in measurements of other radioactive quantities carried at All-Union State Research Institute of Metrology named after D.I. Mendeleev in the former Soviet Union. With the reports of these measurements, the information about the Method of Reduction spread outside that Institute and outside the country. The first formal publication appeared in 1975 [34]. By now this method has became well known; it is mentioned in the Guide [1] under the name “Approach 2” and with a note that this approach is preferable to “Approach 1” (which is the traditional method).

6.6. The Method of Transformation

The method of reduction described in Section 6.5 replaces the traditional method for processing data obtained from indirect measurements. Unfortunately, that method is inapplicable to independent indirect measurements and therefore there was no alternative to the traditional method for that type of measurements. But the traditional method satisfies neither theory nor practice of measurements because it does not provide a grounded way to construct confidence intervals and

therefore to calculate uncertainty of a measurement result. Section 6.4 discusses the drawbacks of this method in more detail. Consequently, we are presenting a new method here, which we call the *method of transformation* and which is free of the above drawbacks.

The essence of the method of transformation can be understood intuitively if one considers a black box with the arguments as its input and values of the measurand as its output. The black box transforms every observed value of each argument into the corresponding value of the measurand. The obtained set of measurand values provides the basis for the estimate of the measurand along with its uncertainty for a chosen confidence probability.

Turning to a more detailed description, let A_j, $j = 1, \ldots, N$ be the arguments of an independent indirect measurement of a measurand x:

$$x = f(A_1, \ldots, A_N). \tag{6.25}$$

First, we obtain the estimates of all arguments, usually as mean values $\bar{A}_j$. Next, we substitute all but one arguments in (6.25) with their estimates. The remaining argument will be denoted as A_v and considered as a variable quantity. Let $A_{v,i}$ $i = 1, \ldots, \ldots, n_v$ be the values of the argument A_v. Each value $A_{v,i}$ of the variable A_v, together with estimates of all other arguments, produces one value of the measurand. Thus, the (6.25) can be presented in the new form

$$x_{v,i} = f(\bar{A}_1, \ldots, \bar{A}_{v-1}, A_{v,i}, \bar{A}_{v+1}, \ldots, \bar{A}_N), \quad i = 1, \ldots, n_v. \tag{6.26}$$

This formula can be written also as

$$x_{v,i} = C_v \Phi_v(A_{v,i}), \tag{6.27}$$

where

$$C_v = \Psi_v(\bar{A}_j), \quad j \neq v. \tag{6.28}$$

Coefficient C_v is determined by estimates $\bar{A}_j$ and therefore is constant for all values $A_{v,i}$. Similarly, function Φ_v is the same for all $A_{v,i}$. In this way, a set of n_v measurements of A_v is transformed into the set of the corresponding values of the measurand $\{x_{v,i}\}$, $i = 1, \ldots, n_v$. The same calculations are performed for each argument, producing N sets of values of the measurand. The problem now is how to use these sets to estimate the measurand and its uncertainty.

Let us start with the analysis of error in the calculated coefficient C_v. In the performed measurement of x (which consists of multiple measurements of arguments), this coefficient is constant. Therefore, its variance is zero: $V[C_v] = 0$. But because $\bar{A}_j \neq A_j$, the value of C_v is not absolutely accurate. Imagine now a repetition of the same measurement of x (perhaps in another laboratory). The new estimates $\bar{A}'_j$ would be slightly different from $\bar{A}_j$. Therefore, the coefficient C'_v would also be not the same as C_v. Thus, the error in $x_{v,i}$ caused by the inaccuracy of C_v is a conditionally constant systematic error: It is constant in the performed measurement but may vary in other possible measurements of the same measurand. Having the estimates of the variances $S^2(\bar{A}_j)$, the change in C_v can be characterized by the variance $S^2(C_v)$ over possible measurements of x.

We can now turn to the calculations. The estimate of the measurand, obtained from a set $\{x_{v,i}\}$, is

$$\bar{x}_v = \frac{1}{n_v} \sum_{i=1}^{n_v} x_{v,i}. \tag{6.29}$$

For convenience, let us introduce a new variable $y_{v,i}$ to replace $\Phi(A_{v,i})$ in (6.27): $y_{v,i} = \Phi(A_{v,i})$. Then, (6.27) can be rewritten as

$$x_{v,i} = C_v y_{v,i}.$$

We obtain from this equation:

$$\zeta(x_{v,i}) = w_y \zeta(y_{v,i}) + w_c \zeta(c_v),$$

where $\zeta(x_{v,i})$, $\zeta(y_{v,i})$, and $\zeta(C_v)$ are the errors of, respectively, $x_{v,i}$, $y_{v,i}$, and C_v, and w_y and w_c are the influence coefficients of $y_{v,i}$ and C_v. This equation provides the dependency between the variances:

$$S^2(x_{v,i}) = w_y^2 S^2(y_{v,i}) + w_c^2 S^2(C_v). \tag{6.30}$$

Denote the terms in Eq. (6.30) as

$$S_{1,i}^2 = w_y^2 S^2(y_{v,i}); \quad S_2^2 = w_c^2 S^2(C_v).$$

Coefficient C_v in set $\{x_{v,i}\}$ is constant as discussed earlier. Therefore, for the performed measurement, $S_2^2 = 0$. Then, $S^2(x_{v,i}) = S_{1,i}^2$.

The above calculation means that $S_{1,i}^2$ express the spread of $x_{v,i}$ in the performed measurement, and hence, it can be computed in the usual way:

$$S_{1,i}^2 = S^2(x_{v,i}) = \frac{1}{n_v - 1} \sum_{i=1}^{n_v} (x_{v,i} - \bar{x}_v)^2.$$

Then, the variance of the estimate $\bar{x}_v$ can be found as

$$S^2(\bar{x}_v) = S_1^2 = \frac{1}{n_v(n_v - 1)} \sum_{i=1}^{n_v} (x_{v,i} - \bar{x}_v)^2. \tag{6.31}$$

Now we can turn to accounting for the variations of C_v in other possible measurements of the same measurand, which will result in a nonzero value of $S^2(C_v)$. If we knew $S^2(C_v)$, we could then compute $S_2^2 = w_c^2 S^2(C_v)$ and estimate the combined variance $S_c^2(\bar{x}_v)$:

$$S_c^2(\bar{x}_v) = S_1^2 + S_2^2. \tag{6.32}$$

Without specifying a concrete function Ψ in (6.28), we cannot provide a general formula for $S^2(C_v)$. However, in any specific measurement, function Ψ is known, and $S^2(C_v)$ can be calculated. An example of such calculation is presented in Section 7.2. Thus, we consider $S_c^2(\bar{x}_v)$ to be known. We will only note here that these calculations are simpler if the estimates of the variances are expressed in the relative form.

Recall that the above calculation was performed for one set $\{x_{v,i}\}$ obtained by assuming that one argument A_v in (6.25) was a variable. By repeating this calculation for each argument A_j, we obtain the estimates $\bar{x}_j$ and $S_c^2(\bar{x}_j)$ for j, $j = 1, \ldots, N$. From these estimates, we now need to find the overall estimate of the measurand and its uncertainty.

For the estimate of the measurand, we can take the weighted mean $\bar{\bar{x}}$ of all $\bar{x}_j$ and compute it according to the recommendation in Chapter 9:

$$\bar{\bar{x}} = \sum_{j=1}^{N} g_j \bar{x}_j. \tag{6.33}$$

The weights g_j are found as

$$g_j = \frac{g'_j}{\sum_{j=1}^{N} g'_j} \quad and \quad g'_j = \frac{1}{S_c^2(\bar{x}_j)}. \tag{6.34}$$

The estimate of the variance of $\bar{\bar{x}}$ is

$$S^2(\bar{\bar{x}}) = \frac{1}{\sum_{j=1}^{N} \frac{1}{S_c^2(\bar{x}_j)}}. \tag{6.35}$$

Having $S^2(\bar{\bar{x}})$, we can calculate the confidence interval for the true value of the measurand. Let K be the total number of elements in all sets $\{x_{j,i}\}$: $K = \sum_{j=1}^{N} n_j$. Then, the number of degrees of freedom is $\nu = K - 1$. Note that this value is accurate. The distribution function of $\bar{\bar{x}}$ can be assumed to be normal because it is a result of the linear sum of many items. So, the Student's distribution with inputs ν and chosen confidence probability α provides the value of t_q, which allows us to calculate the confidence interval and hence the total uncertainty u_t of measurement result as

$$u_t = t_q S(\bar{\bar{x}}). \tag{6.36}$$

A detailed example of using this method is presented in Section 7.2.

In most cases, the measurement equation can be transformed into the following form:

$$x = f_1(A_1) \oplus f_2(A_2) \oplus \ldots \oplus f_N(A_N),$$

where $\oplus$ denotes an arbitrary arithmetic operation (addition, subtraction, multiplication, or division), and the terms in the right-hand side are ordered according to the order of computation (that is, the left-most operation is applied first to the first two terms, then the second operation is applied to the result of the first operation and the third term, and so on). Then, the calculations for the indirect measurement processing can be simplified by a series of successive argument substitutions. Each step of this process substitutes a pair of arguments with one new argument. After $(N - 2)$ steps, the original equation with N arguments will be transformed into an equivalent measurement equation having only two arguments. The processing at each step, as well as handling of the final equation, use the same simple

calculations based on the method of transformation for a measurement with two arguments.

To illustrate the main idea of this method, consider an indirect measurement with four arguments:

$$x = f_1(A_1) \oplus f_2(A_2) \oplus f_3(A_3) \oplus f_4(A_4).$$

We start by substituting the first two arguments, A_1 and A_2. To this end, we replace the corresponding terms with a new argument $B' = f_1(A_1) \oplus f_2(A_2)$. The measurement equation now becomes

$$x = B' \oplus f_3(A_3) \oplus f_4(A_4).$$

We now apply the method of transformation to the expression for B' above. (An example of applying the method of transformation to an indirect measurement with two arguments is described below in Section 7.2.) According to this method, we use the measurement data for arguments A_1 and A_2 to obtain the data set for B', $\{B'_i\}$, $i = \overline{1, (n_1 + n_2)}$, and from it the estimate $\tilde{B}'$ and its standard deviation, to be used in the next step.

Continuing the substitution process, we substitute the first pair of arguments in the equation that resulted from the previous step, B' and A_3, with a new argument $B'' = B' \oplus f_3(A_3)$. The measurement equation will now become

$$x = B'' \oplus f_4(A_4).$$

Similar to the first step, we use the data set for B', its estimate and standard deviation (from the previous step), as well as the measurement data for A_3, to produce the set $\{B''_i\}$, $i = \overline{1, (n_1 + n_2 + n_3)}$ for argument B'', its estimate $\tilde{B}''$ and its standard deviation.

The measurement equation produced by the last step contains only two arguments. Using the data set and estimate for B'' and the measurement data for A_4, we can now obtain the data set for the measurand x, $\{x_k\}$, $k = 1, \ldots, \sum_{j=1}^{N} n_j$. This last set, along with the standard deviation of B'', allows us to obtain the estimate of the measurand and its uncertainty.

The above calculations concern the random error of the indirect measurement. But a real measurement can also be distorted by a bias (if function f is not linear) and systematic errors. One must check for the presence of the bias and correct it if necessary as described in Section 6.4. The method of accounting for systematic errors is presented later in Section 6.7.

To conclude, the method of transformation provides the uncertainty of the results of independent indirect measurements that is well-grounded and more accurate than what was possible with the traditional method. The method of transformation compliments the method of reduction, which solved the same problem for the dependent indirect measurements. Thus, these two methods together allow well-grounded processing of data of any multiple indirect measurement.

6.7. Errors and Uncertainty of Indirect Measurement Results

Uncertainty of an indirect measurement result is calculated on the basis of the estimation of its components, i.e., systematic and random errors.

Random errors are characterized by their variance whose estimate is given by (6.16) and (6.23) for dependent indirect measurements and by (6.18) and (6.30) for independent measurements.

Systematic errors are not apparent in the process of measurements, and therefore, they must be evaluated, taking into account the possible causes of them: first, the systematic errors in the measurements of arguments. The calculations for estimating these errors are the same for the two types of indirect measurements. They are the same also for multiple and single measurements.

The relationship between the measurement errors of arguments and the error of the indirect measurement is represented by (6.13). This equation reflects the transformation of the errors in measurements of arguments into the error of an indirect measurement.

In addition to the error from the measurement errors of arguments, the indirect measurements have an additional source of error. It is an inaccuracy of the measurement equation. The next example will illustrate this error.

Suppose that we are required to measure the area of a plot of land that is depicted by a rectangle on a sketch. Here the rectangle is the model of the object. Its area is $S_m = ab$, where a and b are the lengths of the sides of the rectangle.

The discrepancies between the model and the object can in this case be expressed by the fact that the angle between the sides will not be exactly 90°, that the opposite sides of the section will not be precisely identical, and that the lines bounding the area will not be strictly straight. Each discrepancy can be estimated quantitatively and then the error introduced by it can be calculated. It is usually obvious beforehand which source of error will be most important.

Suppose that in our example the most important source of error is that the angle between adjoining sides differs from 90° by β, as shown in Fig. 6.1. Then the area of the plot would have to be calculated according to the formula $S_t = ab \cos \beta$.

Therefore the error from the threshold discrepancy in this case will be

$$S_1 - S_2 = ab(1 - \cos \beta).$$

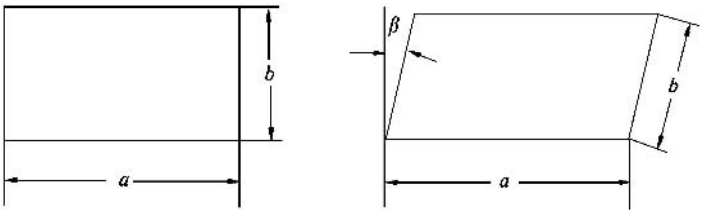


FIGURE 6.1. Rectangle and parallelogram as a model of the plot of land.

The admissible angle β_a must be estimated from the required accuracy in determining the area of the pylot of land. If $\beta \geq \beta_a$, then the model must be redefined and the measured quantity must be defined differently. Correspondingly, we shall obtain a different formula for calculating the measured area.

We should note that the inaccuracy of the measurement equation, or the threshold discrepancy between the model of an object to be studied, and the object is a methodological error and it is an absolutely constant systematic error.

The random errors of indirect measurements were analyzed previously in this chapter. Let us now begin the analysis of the systematic errors of indirect measurements.

The general approach to the problem of the estimation of systematic errors is similar to the one developed for direct measurements. Still, indirect measurements have some specifics. One difference has to do with the existence influence coefficients w_j. Usually their values are calculated by substituting the estimates of arguments for their true values. In other cases, these coefficients are found from special experiments. In all cases, they are obtained with some errors. These errors can be avoided if the measurement equation has the form

$$A = A_1^k A_2^l \dots A_m^n. \tag{6.37}$$

In this case, the influence coefficients are determined by the expressions

$$\begin{aligned} w_1 &= \frac{\partial A}{\partial A_1} = k A_1^{k-1} A_2^l \dots A_N^n, \\ w_2 &= \frac{\partial A}{\partial A_2} = A_1^k l A_2^{l-1} \dots A_N^n, \\ &\dots \\ w_N &= \frac{\partial A}{\partial A_N} = A_1^k A_2^l \dots n A_N^{n-1}. \end{aligned}$$

The absolute error is determined by formula (6.13). We shall now transfer from the absolute error to the relative error:

$$\begin{aligned} \varepsilon = \frac{\tilde{A} - A}{A} = \frac{k A_1^{k-1} A_2^l \dots A_N^n}{A} \zeta_1 + \frac{l A_1^k A_2^{l-1} \dots A_N^n}{A} \zeta_2 \\ + \dots + \frac{A_1^k A_2^l \dots n A_N^{n-1}}{A} \zeta_N. \end{aligned}$$

Substituting formula (6.37) for A, we obtain

$$\varepsilon = k \frac{\zeta_1}{A_1} + l \frac{\zeta_2}{A_2} + \dots + n \frac{\zeta_N}{A_N}.$$

Thus the influence coefficients for the relative errors in the measurements of the arguments are equal to the powers of the corresponding arguments: $w_1' = k$, $w_2' = t$, $w_N' = n$. The coefficients $k, l \dots n$ are known exactly a priori, so that the error noted above does not arise.

This result can be obtained without use of (6.13), in other words, without application of Taylor series. Indeed, let us take the logarithm of the equation (6.37). It will be

$$\ln A = k \ln A_1 + l \ln A_2 + \dots + n \ln A_N.$$

The differential of it is

$$\frac{dA}{A} = k\frac{dA_1}{A_1} + l\frac{dA_2}{A_2} + \cdots + n\frac{dA_N}{A_N}.$$

Measurement errors are small. Therefore the differentials can be replaced by increments after that the last equation gets the same form that was obtained above.

So, relative form of errors provides the uncertainty calculations with exact values of influence coefficients. This is another advantage of expressing the measurement errors in the form of relative errors.

The systematic error of the measurement of each argument consists of elementary components. As always, they can be divided into two categories: absolutely and conditionally constant errors.

Absolutely constant errors are deterministic quantities. However, we cannot find their exact values and can only estimate their limits. These limits are estimated differently in every specific case. In general, these estimations are based on the experience of the person performing the measurement. Usually, there are very few such errors and they are small. But it is necessary to keep them in mind.

One example of absolutely constant errors is the error in a measurement equation considered above. Another example is the linearization error of the standard characteristic of a thermocouple.

Conditionally constant errors can be computed using the first term of (6.14):

$$\vartheta_{cc} = \sum_{j=1}^{N} w_j \vartheta_j,$$

where ϑ_{cc} is the conditionally constant error of an indirect measurement.

This formula can be represented in the form

$$\vartheta_{cc} = \sum_{j=1}^{N} \sum_{i=1}^{k_j} w_j \vartheta_{ji}, \tag{6.38}$$

where k_j is the number of conditionally constant errors in the measurement of the jth argument.

It is necessary to note that the obtained formula is based on Taylor's series. The shortcomings of this approach have been discussed already in this chapter. But unlike in the case of random errors, there is no alternative to this method in evaluation of systematic errors.

Now we shall turn from analysis to synthesis. As was discussed in Chapter 5, it is possible to consider all conditionally constant errors as random quantities with a uniform distribution. But there is one peculiarity now: In the case of dependent indirect measurements, some elementary errors in the measurements of different arguments are caused by the same influence quantity. When such a quantity grows, some of these errors can grow also, the rest of them go in the opposite direction.

For example, assume that two measuring instruments used in an indirect measurement have temperature errors. When the temperature changes, these errors will also change, and both of them can change either in the same direction or in opposite directions.

So, the sum of the elementary errors caused by the same influence quantity can be obtained by simply adding all error values retaining their signs.

Taking into consideration this peculiarity, (6.38) becomes

$$\vartheta_{cc} = w_1 \sum_{i=1}^{k_1-m_1} \vartheta_{1i} + \cdots + w_N \sum_{i=1}^{k_N-m_N} \vartheta_{Ni} + (w_1\vartheta_{1t} \pm w_2\vartheta_{2t} \pm \cdots) + (w_1\vartheta_{1p} \pm w_2\vartheta_{2p} \pm \cdots), \tag{6.39}$$

where m_j is the number of the elementary components of the measurement error of the jth arguments, which are caused by the same influence quantity as some components of the measurement errors of other arguments; indexes t, p denote these influence quantities.

Equation (6.39) allows one to obtain the expression of the variance σ_s^2 of the conditionally constant error of the measurement result:

$$\sigma_{cc}^2 = w_1^2 \sum_{i=1}^{k_1-m_1} \sigma_{1i}^2 + \cdots + w_N^2 \sum_{i=1}^{k_N-m_N} \sigma_{Ni}^2 + (w_1\sigma_{1t} \pm w_2\sigma_{2t} \pm \cdots)^2 + (w_1\sigma_{1p} \pm w_2\sigma_{2p} \pm \cdots)^2 + \cdots. \tag{6.40}$$

The last two terms appeared as a consequence of (6.4) for $\rho_{1,2} = \pm 1$. Indeed, if $\rho_{1,2} = \pm 1$, then (6.4) has the form

$$\sigma_2^2 = \sigma_x^2 + \sigma_y^2 \pm 2\sigma_x\sigma_y = (\sigma_x \pm \sigma_y)^2.$$

Recall that the variance σ^2 and the limits θ of a random quantity having a uniform distribution function related by the formula are

$$\sigma^2 = \frac{1}{3}\theta^2.$$

Therefore, knowing the limits θ_{ji} of all conditionally constant elementary errors, we also know the estimates of their variances. Inserting them into (6.40), we obtain

$$S_{cc}^2 = \frac{1}{3}\left[\sum_{j=1}^{N-m}\sum_{i=1}^{m} w_j^2\theta_{ji}^2 + \sum_{\mu=1}^{m}(w_k\theta_{k\mu} \pm w_l\theta_{l\mu} \pm \cdots)^2\right], \tag{6.41}$$

where m is the number of influence quantities that affect two or more measurements of the arguments.

The confidence limits of the conditionally constant error of an indirect measurement can be calculated using the same method that was discussed in Chapter 5. The only difference is that now we have to account for influence coefficients. So, from (5.3), we get

$$\theta_{c\alpha} = k\sqrt{\sum_{j=1}^{N-m}\sum_{i=1}^{m} w_j^2\theta_{ji}^2 + \sum_{\mu=1}^{m}(w_k\theta_{R\mu} \pm w_l\theta_{l\mu} \pm \cdots)^2}, \tag{6.42}$$

where $\theta_{c\alpha}$ is the confidence limit of a conditionally constant error. The values of k are given in Section 3.5. In particular, for the probability $\alpha = 0.95$, $k = 1.1$.

If $m = 0$, or the indirect measurement is performed under reference conditions, then (6.42) has the form

$$\theta_{c\alpha} = k\sqrt{\sum_{j=1}^{N} w_j^2\theta_j^2}. \tag{6.43}$$

I would like to recall here that the random error of a multiple measurement includes all random components of conditionally constant errors of this measurement. Thus, the remains of conditionally constant errors in multiple measurements are purely systematic errors.

Now let us return to the absolutely constant errors. Summarizing their limits, we obtain the limits H of the absolutely constant error of the result of an indirect measurement:

$$H = H_e + \sum_{j=1}^{N} w_j \sum_{i=1}^{gj} H_{ji}, \tag{6.44}$$

where H_e are the limits of an error of the measurement equation; H_{ji} are the limits of the ith absolutely constant component of the measurement error of the jth argument; and q_j is the number of absolutely constant components of the measurement error of the jth argument.

Thus, we have the estimate of the variance of conditionally constant errors S_{cc}^2 and the limits of the absolutely constant error H. We also have the estimate of the variance of random errors S_r^2. These findings are exactly the same as those used for the uncertainty calculation in Chapter 5 for direct measurements. Therefore, in the same way, we can now calculate the uncertainty of indirect measurements. The resulting formulas are repeated below.

The combined standard deviation S_c can be calculated using (5.12):

$$S_c = \sqrt{S_{cc}^2 + S_r^2}. \tag{6.45}$$

The combined uncertainty can be found from (5.13)

$$u_c = t_c S_c, \tag{6.46}$$

and the coefficient t_c is calculated by (5.14):

$$t_c = \frac{\theta_{c\alpha} + t_q S_r}{S_{cc}^2 + S_r}. \tag{6.47}$$

In the case when $\theta_{c\alpha}$ consists of many approximately equal components (five or more), its distribution is approximately normal and therefore we can assume $t_c = 1.96$ for $\alpha = 0.95$ and $t_c = 2.58$ for $\alpha = 0.99$.

Taking into account the limit of the absolutely constant error, we obtain the total uncertainty u_t of the measurement result:

$$u_t = H + u_c. \tag{6.48}$$

7
Examples of Measurements and Measurement Data Processing

7.1. An Indirect Measurement of the Electrical Resistance of a Resistor

Consider the measurement of electrical resistance using an ammeter and a voltmeter. Equation (6.1) now turns into $R = U/I$, where R is the electrical resistance of the resistor, U is the voltage drop on the resistor, and I is the strength of the current. This measurement is an example of dependent indirect measurements. Indeed, the value of I depends on the value of U. Thus, we can use in this example both the traditional method and the method of reduction. Let us use each in turn and compare the calculations and results. The connections of the instruments and the resistor are shown in Fig. 7.1. Assume that the measurement was performed under reference conditions for the instruments, and that the input resistance of the voltmeter is so high that its influence on the accuracy of the measurement can be neglected.

The results of measurements of the strength of current and voltage are given in Table 7.1. In accordance with the note given in Section 6.2, all results presented in the table were obtained in pairs, which means that the results with the same subscript belong to the same measurement vector.

Using the values of $n\bar{U}$ and $n\bar{I}$ given in Table 7.1 (columns 2 and 3, the last row), we obtain the estimate for R:

$$\tilde{R} = \frac{n\bar{U}}{n\bar{I}} = \frac{66.002}{0.659\,97} = 100.0075 \approx 100.01\,\Omega.$$

Deviding the sums given in the colomns 2 and 3 by 11, we obtain the estimates of I and V:

$$\bar{I} = 0.059\,997 \approx 0.060\,00\,\mathrm{A}, \qquad \bar{U} = 6.000\,18 \approx 6.0002\,\mathrm{V}.$$

Now we must calculate the variance and the standard deviation of this result.

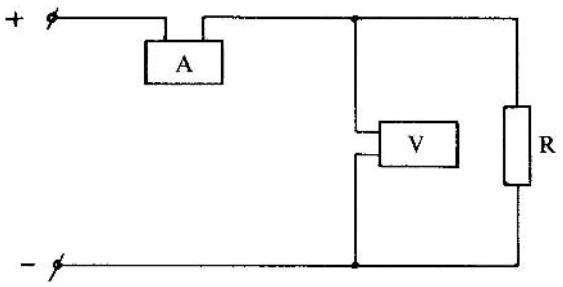


FIGURE 7.1. The connections for the indirect measurement of an electrical resistance.

First, we will estimate the variances of $\bar{I}$, $\bar{U}$, their standard deviations, and the correlation coefficient. Using (6.17), we obtain

$$S^2(\bar{I}) = \frac{\sum_{i=1}^{n}(I_i - \bar{I})^2}{n(n-1)} = \frac{74.19 \times 10^{-10}}{11 \times 10} = 0.674 \times 10^{-10}\,\mathrm{A}^2,$$

$$S^2(\bar{U}) = \frac{\sum_{i=1}^{n}(U_i - \bar{U})^2}{n(n-1)} = \frac{63.61 \times 10^{-6}}{11 \times 10} = 0.578 \times 10^{-6}\,\mathrm{V}^2.$$

The estimates of standard deviations are

$$S(\bar{I}) = 0.82 \times 10^{-5}\,\mathrm{A}, \qquad S(\bar{U}) = 0.76 \times 10^{-3}\,\mathrm{V}.$$

The estimate of the correlation coefficient is

$$r_{I,U} = \frac{\sum_{i=1}^{n}(I_i - \bar{I})(U_i - \bar{U})}{n(n-1)S(I)S(U)} = \frac{29.6 \times 10^{-8}}{110 \times 0.82 \times 10^{-5} \times 0.76 \times 10^{-3}} = 0.43.$$

TABLE 7.1. Data processing for indirect measurement of electrical resistance using the traditional method.

Num.	I_i A	U_i V	$(I_i - \bar{I})$ $\times 10^{-5}$ A	$(I_i - \bar{I})^2$ $\times 10^{-10}$ A^2	$(U_i - \bar{U})$ $\times 10^{-3}$ V	$(U_i - \bar{U})^2$ $\times 10^{-6}$ V^2	$(I_i - \bar{I})(U_i - \bar{U})$ $\times 10^{-8}$AV
1	2	3	4	5	6	7	8
1	0.059 96	6.003	−3.7	13.69	+2.82	7.95	−10.4
2	0.060 01	6.001	+1.3	1.69	+0.82	0.67	+1.1
3	0.059 98	5.998	−1.7	2.89	−2.18	4.75	+3.7
4	0.060 03	6.001	+3.3	10.89	+0.82	0.67	+2.7
5	0.060 01	5.997	+1.3	1.69	−3.18	10.11	−4.1
6	0.059 98	5.999	−1.7	2.89	−1.18	1.39	+2.0
7	0.060 03	6.004	+3.3	10.89	+3.82	14.59	+12.6
8	0.059 95	5.997	−4.7	22.09	−3.18	10.11	+14.9
9	0.060 02	6.001	+2.3	5.29	+0.82	0.67	+1.9
10	0.060 01	6.003	+1.3	1.69	+2.82	7.95	+3.7
11	0.059 99	5.998	−0.7	0.49	−2.18	4.75	+1.5
Sum	0.659 97	66.002		74.19		63.61	+29.6

It is interesting to note that this value is statistically insignificant. Indeed, applying a standard method of [20] and [52], we can check the hypothesis $H_0 : \rho_{I,\,U} = 0$ against $H_1 : \rho_{I,\,U} \neq 0$. The degree of freedom here is $\nu = 11 - 2 = 9$, and we will take the significance level to be $q = 0.05$ as usual, which gives the critical values $t_q = 2.26$ and $r_q = t_q/\sqrt{t_q^2 + \nu} = 0.60$. Because $0.43 < 0.60$, we must accept H_0 and conclude that the obtained value $r_{I,\,U} = 0.43$ is not significant, which means that, when the number of measurements n increases the estimation $r_{I,\,U}$ of the correlation coefficient will in general decrease. However, it does not mean that the value of $r_{I,\,U}$ obtained for a *specific sample* can be neglected. On the contrary, it must be always taken into consideration when calculating the estimation of variance for that sample.

In our example, inserting the obtained values into (6.16), we can calculate the desired estimation $S(\tilde{R})$. But before that we have to calculate the influence coefficients. They are

$$w_1 = \frac{\partial R}{\partial U} = \frac{1}{I}, \qquad w_2 = \frac{\partial R}{\partial I} = -\frac{U}{I^2},$$

$$\begin{aligned} S^2(\tilde{R}) &= \left(\frac{\bar{U}}{\bar{I}^2}\right)^2 \times S^2(\bar{I}) + \frac{1}{\bar{I}^2} \times S^2(\bar{U}) - r_{I,\,U}\frac{\bar{U}}{I^2} \times \frac{1}{I} \times S(\bar{I})S(\bar{U}) \\ &= \left(\frac{6}{36 \times 10^{-4}}\right)^2 \times 0.674 \times 10^{-10} + \frac{1}{36 \times 10^{-4}} \times 0.578 \times 10^{-6} \\ &\quad - 2 \times 0.43 \times \frac{6}{36 \times 10^{-4}} \times \frac{1}{6 \times 10^{-2}} \times 0.82 \times 10^{-5} \times 0.76 \times 10^{-3} \\ &= 1.87 \times 10^{-4} + 1.61 \times 10^{-4} - 1.49 \times 10^{-4} \\ &= 1.99 \times 10^{-4}\,\Omega^2, \end{aligned}$$

and

$$S(\tilde{R}) = \sqrt{S^2(R)} = 1.41 \times 10^{-2}\,\Omega.$$

We now turn to the *method of reduction.*

The initial data from Table 7.1 are repeated in columns 2 and 3 of Table 7.2. The calculated values of $R_i (i = 1, \ldots, 11)$ are given in column 4. Treating these values as if they were obtained by direct measurements, we obtain immediately the estimate of R as

$$\bar{R} = \frac{1}{n}\sum_{i=1}^{n} R_i = 100.0075 \approx 100.01\,\Omega$$

and the estimates of its variance and standard deviation as

$$S^2(\bar{R}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}(R_i - \bar{R})^2 = \frac{2.184 \times 10^{-2}}{11 \times 10} = 1.99 \times 10^{-4}\,\Omega^2,$$

$$S(\bar{R}) = 1.41 \times 10^{-2}\,\Omega.$$

TABLE 7.2. Data processing for indirect measurement of electrical resistance using the method of reduction.

Num.	I_i A	U_i V	R_i Ω	$(R_i - \bar{R})$ Ω	$(R_i - \bar{R})^2$ $\times 10^{-2}\ \Omega^2$
1	2	3	4	5	6
1	0.059 96	6.003	100.117	+0.109	1.188
2	0.060 01	6.001	100.000	−0.002	0.000
3	0.059 98	5.998	100.000	−0.002	0.000
4	0.060 03	6.001	99.967	−0.041	0.168
5	0.060 01	5.997	99.933	−0.075	0.562
6	0.059 98	5.999	100.017	+0.009	0.008
7	0.060 03	6.004	100.017	+0.009	0.008
8	0.059 95	5.997	100.033	+0.025	0.0625
9	0.060 02	6.001	99.983	−0.025	0.0625
10	0.060 01	6.003	100.033	+0.025	0.0625
11	0.059 99	5.998	99.983	−0.025	0.0625
Sum			1100.083		2.184

The comparison shows the coincidence of $\tilde{R}$ and $\bar{R}$ and $S(\tilde{R})$ and $S(\bar{R})$, which could have been predicted because it is a case of an accurate measurement. At the same time, it is a measurement with a simple measurement equation. Even in this simple case, the calculations with the method of reduction are much simpler than with the traditional method.

Further calculations for this example are of little interest and are therefore not presented here.

7.2. The Measurement of the Density of a Solid Body

The accurate measurement of the density of a solid body can serve as an example of a multiple nonlinear independent indirect measurement. Either the traditional method or the new method of transformation could be used to process the measurement data in this case. We will use both of them to highlight the benefits of the latter. We will begin with the traditional method.

The density of a solid body is given by the formula

$$\rho = m/V,$$

where m is the mass of the body and V is the volume of the body. In the experiment considered, the mass of the body was measured by methods of precise weighing using a collection of standard weights whose errors did not exceed 0.01 mg. The volume of the body was determined by the method of hydrostatic weighing using the same set of weights.

The results of measurements are presented in Table 7.3 in the first and fourth columns.

TABLE 7.3. The results of measurements of the density of a solid body and data from initial processing.

Mass of body $m_i \times 10^{-3}$ kg	$(m_i - \bar{m})$ $\times 10^{-7}$ kg	$(m_i - \bar{m})^2$ $\times 10^{-14}$ kg^2	Volume of body $V_i \times 10^{-6}$ m^3	$(V_i - \bar{V})$ $\times 10^{-10}$ m^3	$(V_i - \bar{V})^2$ $\times 10^{-20}$ m^6
1	2	3	4	5	6
252.9119	−1	1	195.3799	+1	1
252.9133	+13	169	195.3830	+32	1024
252.9151	+31	961	195.3790	−8	64
252.9130	+10	100	195.3819	+21	441
252.9109	−11	121	195.3795	−3	9
252.9094	−26	676	195.3788	−10	100
252.9113	−7	49	195.3792	−6	36
252.9115	−5	25	195.3794	−4	16
252.9119	−1	1	195.3791	−7	49
252.9115	−5	25	195.3791	−7	49
252.9118	−2	4	195.3794	−4	16

The difference between the observational results is explained by the random error of the balances. As follows from the data presented, this error is so much larger than the systematic errors in the masses of the weights that these errors can be neglected.

As the mass of the solid body and its volume are constants, to estimate the density of the solid, the mass, and volume of the solid must be estimated with the required accuracy and their ratio must be formed. For this reason, we find the average values of the observational results and estimates of the standard deviations for the groups of measurements:

$$\bar{m} = 252.9120 \times 10^{-3}\,\text{kg}, \qquad \bar{V} = 195.3798 \times 10^{-6}\,\text{m}^3,$$

$$S^2(m_i) = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(m_i - \bar{m})^2 = \frac{2132 \times 10^{-14}}{10} = 213.2 \times 10^{-14}\,\text{kg}^2,$$

$$S^2(V_i) = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(V_i - \bar{V})^2 = \frac{1805 \times 10^{-20}}{10} = 180.5 \times 10^{-20}\,\text{m}^6.$$

The estimates of the variances, in the relative form, are equal to

$$S_r^2(m_i) = \frac{213 \times 10^{-14}}{(252.9 \times 10^{-3})^2} = 3.32 \times 10^{-11},$$

$$S_r^2(V_i) = \frac{180 \times 10^{-20}}{(195.4 \times 10^{-6})^2} = 4.74 \times 10^{-11}.$$

The estimate of the measured quantity is

$$\tilde{\rho} = \frac{\bar{m}}{\bar{V}} = \frac{252.9120 \times 10^{-3}}{195.3798 \times 10^{-6}} = 1.294\,463 \times 10^3\,\text{kg/m}^3.$$

To calculate the uncertainty of the result we use here the usual method of linearization, but it is necessary to check that only the first term from the Taylor series is enough. For this reason, it is necessary to estimate the remainder R_2 according to formula (6.11):

$$R_2 = \frac{1}{2}\left[\frac{\partial^2\rho}{\partial m^2}(\Delta m)^2 + \frac{\partial^2\rho}{\partial V^2}(\Delta V)^2 + 2\frac{\partial^2\rho}{\partial m\,\partial V}\Delta m\,\Delta V\right],$$

$$\frac{\partial\rho}{\partial m} = \frac{1}{V}, \qquad \frac{\partial^2\rho}{\partial m^2} = 0,$$

$$\frac{\partial\rho}{\partial V} = -\frac{m}{V^2}, \qquad \frac{\partial^2\rho}{\partial V^2} = \frac{2m}{V^3},$$

$$\frac{\partial\rho}{\partial m\,\partial V} = -\frac{1}{V^2}.$$

We shall calculate the partial derivatives at the point with the coordinates $\bar{m}$ and $\bar{V}$, so that the errors Δm and ΔV are relatively insignificant. We obtain

$$R_2 = \frac{\bar{m}}{\bar{V}^3}(\Delta V)^2 - \frac{1}{\bar{V}^2}\Delta m\,\Delta V = \frac{\bar{m}}{\bar{V}}\left(\frac{\Delta V}{\bar{V}}\right)^2 - \frac{\bar{m}}{\bar{V}}\frac{\Delta m}{\bar{m}}\frac{\Delta V}{\bar{V}} = \tilde{\rho}\frac{\Delta V}{\bar{V}}\left(\frac{\Delta V}{\bar{V}} - \frac{\Delta m}{\bar{m}}\right)$$

or

$$\frac{R_2}{\tilde{\rho}} = \frac{\Delta V}{\bar{V}}\left(\frac{\Delta V}{\bar{V}} - \frac{\Delta m}{\bar{m}}\right).$$

For ΔV and Δm, we take the largest deviations from the average values observed in the experiment:

$$\Delta V = 32 \times 10^{-10}\,\mathrm{m}^3, \qquad \Delta m = 31 \times 10^{-7}\,\mathrm{kg}.$$

The relative errors are equal to

$$\frac{\Delta V}{\bar{V}} = \frac{32 \times 20^{-10}}{195.4 \times 10^{-6}} = 1.64 \times 10^{-5},$$

$$\frac{\Delta m}{\bar{m}} = \frac{31 \times 10^{-7}}{252.9 \times 10^{-3}} = 1.22 \times 10^{-5}.$$

As the errors are random, the sign of the error should not be specified. We obtain

$$R_2/\tilde{\rho} = 1.64 \times 10^{-5}(1.64 + 1.22) \times 10^{-5} = 4.7 \times 10^{-10}.$$

This error is so much smaller than the errors associated with $\Delta V/\bar{V}$ and $\Delta m/\bar{m}$ it is obvious that linearization is possible.

We shall estimate the standard deviation in the relative form using formula (6.18)

$$S_r(\tilde{\rho}) = \sqrt{\frac{S_r^2(m_i)}{n_1} + \frac{S_r^2(V_i)}{n_2}} = \sqrt{\frac{3.32 \times 10^{-11} + 4.74 \times 10^{-11}}{11}} = 2.7 \times 10^{-6}.$$

Here the influence coefficients are $k = 1$ and $l = -1$ (see p. 174). For percentages, $S_{\%}(\tilde{\rho}) = 2.7 \times 10^{-4}\%$.

In units of density, we obtain

$$S(\tilde{\rho}) = 2.7 \times 10^{-6} \times 1.294 \times 10^{3} = 3.5 \times 10^{-3}\,\mathrm{kg/m^3}.$$

We shall now find the uncertainty of the result. Based on formula (6.14), we can write

$$\psi_\rho = \psi_m - \psi_V.$$

As the errors are random, the minus sign need not be included. We shall find the uncertainty of the components. We shall take the confidence probability $\alpha = 0.95$. In addition, we have $\nu = 10$. Then we can find from Student's distribution $t_q = 2.23$.

In the relative form, the uncertainties are

$$u_{m\%} = 100 t_q \frac{S_r(\bar{m})}{\sqrt{n_1}} = 100 \times 2.23 \sqrt{\frac{3.32 \times 10^{-11}}{11}} = 3.88 \times 10^{-4}\%,$$

$$u_{v\%} = 100 t_q \frac{S_r(\bar{V})}{\sqrt{n_2}} = 100 \times 2.23 \sqrt{\frac{4.74 \times 10^{-11}}{11}} = 4.64 \times 10^{-4}\%.$$

In accordance with (6.20), we can find the value of combined uncertainty

$$u_{c\%} = \sqrt{u_{m\%}^2 + u_{v\%}^2} = 10^{-4}\sqrt{3.88^2 + 4.64^2} = 6.0 \times 10^{-4}\%.$$

So, the combined relative uncertainty of the measurement result is

$$u_{\rho\%} = 6 \times 10^{-4}\%.$$

The above calculations are similar to those recommended by the *Guide* [1] (Approach 1). Both methods are based on transforming the measurement equation into a linear form by applying a Taylor series. The difference is only in the last step of uncertainty calculation, which uses the square-root sum formula (6.20) in the above calculations and Student distribution in the *Guide*. Let us reproduce the *Guide*'s calculation.

As in the case of linear indirect measurements, the effective number of degrees of freedom given by formula (6.19) is calculated. We have already obtained the values of all terms in this formula. Note that in our case $n_1 = n_2 = n = 11$ and $\nu_1 = \nu_2 = \nu = n - 1$.

Thus,

$$\nu_{\mathrm{eff}} = \frac{\left(\dfrac{S_r^2(m_i)}{n} + \dfrac{S_r^2(V_i)}{n}\right)^2}{\dfrac{S_r^4(m_i)}{n^2\nu} + \dfrac{S_r^4(V_i)}{n^2\nu}} = \frac{(3.32 + 4.74)^2 \times 10^{-22}}{\left(\dfrac{3.32^2}{10} + \dfrac{4.74^2}{10}\right) \times 10^{-22}} = \frac{64.96}{33.49} \times 10 = 19.$$

For $\nu_{\mathrm{eff}} = 19$ and $\alpha = 0.95$ we find from a table of Student's distribution $t_q = 2.10$. From here

$$u'_{\rho\%} = 2.1 \times 2.7 \times 10^{-6} \times 100 = 5.7 \times 10^{-4}\%.$$

TABLE 7.4. The set of $\rho_{m,i}$ and data derived from this set.

i	$\rho_{m,i} \times 10^3\, kg/m^3$	$(\rho_{m,i} - \bar{\rho}m) \times 10^{-3} kg/m^3$	$(\rho_{m,i} - \bar{\rho}m)^2 \times 10^{-6}\, (kg/m^3)^2$
1	1.2944626	−0.3	0.1
2	1.2944698	+6.9	47.6
3	1.2944790	+16.1	259.2
4	1.2944682	+5.3	28.1
5	1.2944575	−5.4	29.2
6	1.2944498	−12.1	146.4
7	1.2944595	−3.4	11.6
8	1.2944605	−2.4	5.8
9	1.2944626	−0.3	0.1
10	1.2944605	−2.4	5.8
11	1.2944620	−0.9	0.8

This value is quite close to that obtained by the square-root sum formula but it is also not well-grounded because, after the measurement equation was linearized, the distribution function of errors of volume measurements of the solid body is unknown.

Let us now turn to the method of transformation described in Section 6.6. We start with the same measurement data that were used in the traditional method and that are presented in Table 7.3. The measurement equation still remain as before:

$$\rho = \frac{m}{V}.$$

Let us first substitute argument V by its estimate $\bar{V}$, retaining m as a variable argument. Then, for each value m_i of argument m (6.26) gives $\rho_{m,i} = \frac{m_i}{\bar{V}}$. Comparing with (6.27) and (6.28), we conclude that the constant coefficient here is $C_m = \Psi_m(\bar{V}) = \frac{1}{\bar{V}}$ and $\Phi(m_i) = m_i, i = 1, \ldots, n_m$. The mean value $\bar{V}$ was calculated earlier when we applied the traditional method. Eleven observations m_i are given in column 1 of Table 7.3. Division m_i by $\bar{V}$ provides the set of individual values $\rho_{m,i}$ shown in column 2 of Table 7.4.

From this set, we obtain

$$\bar{\rho}_m = \frac{1}{n_m} \sum_{i=1}^{n_m} \rho_{m,i} = 1.2944629 \times 10^3 kg/m^3$$

and

$$S_1^2(\bar{\rho}_m) = \frac{1}{n_m(n_m-1)} \sum_{i=1}^{n_m} (\rho_{m,i} - \bar{\rho}_m)^2 = \frac{534.7 \times 10^{-6}}{11 \times 10} = 4.86 \times 10^{-6} (kg/m^3)^2.$$

The relative form of the above variance is

$$S_{1r}^2(\bar{\rho}_m) = \frac{S_1^2(\bar{\rho}_m)}{\bar{\rho}_m^2} = \frac{4.86 \times 10^{-6}}{1.29^2 \times 10^6} = 2.93 \times 10^{-12}.$$

Now the function $\Psi_m = \frac{1}{\bar{V}}$ is known, and thus we can obtain the estimate of the variance of the conditionally constant systematic error $S^2(C_m)$ using the Taylor's series:

$$S^2(C_m) = w_v^2 S^2(\bar{V})$$

where w_v is the influence coefficient of $\bar{V}$. In the relative form of the above variance, the influence coefficient w_v becomes $w_v' = l = -1$, and $S_r^2(C_m) = S_r(\bar{V})$. Therefore, following (6.30), we can calculate $S_{2r}^2(\bar{\rho}m)$:

$$S_{2r}^2(\bar{\rho}_m) = S_r^2(C_m) = S_r^2(\bar{V}).$$

The estimate $S_r^2(V_i)$ was already calculated in the course of applying the traditional method:

$$S_r^2(V_i) = 4.74 \times 10^{-11}.$$

Because $n_v = 11$, we have

$$S_r^2(\bar{V}) = \frac{4.74 \times 10^{-11}}{11} = 4.31 \times 10^{-12}.$$

Thus, $S_{2r}^2(\bar{\rho}m) = 4.31 \times 10^{-12}$ and the estimate of the combined variance of $\bar{\rho}_m$ is

$$S_{rc}^2(\bar{\rho}_m) = S_{1r}^2(\bar{\rho}_m) + S_{2r}^2(\bar{\rho}_m) = (2.93 + 4.31)10^{-12} = 7.24 \times 10^{-12}.$$

Now we substitute m with its estimate $\bar{m}$ and take V as a variable, thus obtaining $\rho_{V,i} = \frac{\bar{m}}{V_i}$ for each value V_i. Here, $C_V = \Psi_V(\bar{m}) = \bar{m}$ and $\Phi(V_i) = \frac{1}{V_i}$, $i = 1, \ldots, n_V$. The mean value $\bar{m}$ we already calculated when applying the traditional method. Thus, we can compute the set $\rho_{V,i}$, which is shown in column 2 of Table 7.5.

From the data in this table, we obtain

$$\bar{\rho}_V = \frac{1}{n_V} \sum_{i=1}^{n_V} \rho_{V,i} = 1.2944613 \times 10^3 kg/m^3$$

TABLE 7.5. The set of $\rho_{V,i}$ and data derived from this set.

i	$\rho_{V,i} \times 10^3 kg/m^3$	$(\rho_{V,i} - \bar{\rho}_V) \times 10^{-3} kg/m^3$	$(\rho_{V,i} - \bar{\rho}_V)^2 \times 10^{-6} (kg/m^3)^2$
1	1.2944606	−0.7	0.5
2	1.2944419	−19.4	376.4
3	1.2944684	+7.1	50.4
4	1.2944474	−13.9	193.2
5	1.2944633	+1.9	3.6
6	1.2944679	+6.6	43.6
7	1.2944653	+4.0	16.0
8	1.2944640	+2.7	7.3
9	1.2944659	+4.6	21.2
10	1.2944659	+4.6	21.2
11	1.2944640	+2.7	7.3

and

$$S_1^2(\bar{\rho}_V) = \frac{1}{n_V(n_V-1)}\sum_{i=1}^{n_V}(\rho_{V,i}-\bar{\rho}_V)^2 = \frac{740.7\times10^{-6}}{11\times10} = 6.73\times10^{-6}(kg/m^3)^2.$$

In the relative form, the above variance becomes

$$S_{1r}^2(\bar{\rho}_V) = \frac{6.73\times10^{-6}}{1.29^2\times10^6} = 4.04\times10^{-12}.$$

Let us turn to component $S_{2r}^2(\bar{\rho}_V)$. Now, $C_V = \bar{m}$, and therefore $S_{2r}^2(\bar{\rho}_V) = S_r^2(C_V) = S^2(\bar{m})$. The value of $S_r^2(m_i) = 3.32\times10^{-11}$ was found before in the application of the traditional method. Thus,

$$S_{2r}^2(\bar{\rho}_V) = \frac{3.32\times10^{-11}}{11} = 3.02\times10^{-12},$$

and the combined variance is

$$S_{rc}^2(\bar{\rho}_V) = S_{1r}^2(\bar{\rho}_V) + S_{2r}^2(\bar{\rho}_V) = (4.05+3.02)\times10^{-12} = 7.07\times10^{-12}.$$

At this point, we have obtained the results for two sets of data, $\{\rho_{m,i}\}$ and $\{\rho_{V,i}\}$, and we are ready to calculate the overall result of the measurement. First, the weights g_j must be found. Following (6.34), we obtain

$$g'_m = \frac{1}{S_{rc}^2(\bar{\rho}m)} = \frac{1}{7.24\times10^{-12}} = 0.138\times10^{12}$$

and

$$g'_V = \frac{1}{S_{rc}^2(\bar{\rho}_V)} = \frac{1}{7.07\times10^{-12}} = 0.141\times10^{12}.$$

Thus,

$$g_m = \frac{g'_m}{g'_m+g'_V} = 0.495$$

and

$$g_V = \frac{g'_V}{g'_m+g'_V} = 0.505.$$

Note that the above weights add to 1.

Then, in accordance with (6.33), the estimate of the measurand is

$$\bar{\bar{\rho}} = g_m\bar{\rho}_m + g_V\bar{\rho}_V = 1.294462\times10^3 kg/m^3.$$

The estimate of the total variance is

$$S_{rt}^2(\bar{\bar{\rho}}) = \frac{1}{\frac{1}{S_{rc}^2(\bar{\rho}m)}+\frac{1}{S_{rc}^2(\bar{\rho}_V)}} = 3.6\times10^{-12}.$$

The standard deviation is

$$S_{rt}(\bar{\bar{\rho}}) = \sqrt{S_{rt}^2} = 1.9\times10^{-6},$$

or, in percents, $S_{rt} = 1.9\times10^{-4}\%$.

Now we can calculate the uncertainty. The number of degrees of freedom is $\nu = K - 1 = 21$. For confidence probability $\alpha = 0.95$, the percentile point of Student's distribution is $t_q = 2.08$. Therefore,

$$u_{t\%} = 2.08 \cdot 1.9 \times 10^{-4} \approx 4 \times 10^{-4}\%.$$

The comparison of the obtained result with the result of the traditional method shows that the estimates of the measurand are the same (the difference in one last digit is absolutely insignificant). But uncertainty obtained by the method of transformation is about 30% smaller than that produced by the traditional method. But what is much more important is that, unlike the traditional method, the uncertainty obtained by the method of transformation corresponds to the chosen probability and is well-grounded.[1]

7.3. The Measurement of Ionization Current by the Compensation Method

Accurate measurements of weak currents, generated, for example, by γ rays from standards of unit radium mass, are performed by the compensation method using an electrometer. The measured strength of current I is defined by the

[1] It is easy to see that the mean value of $\rho_{m,i}$ (see Table 7.4) is equal to

$$\bar{\rho}_m = \frac{1}{n}\sum_{i=1}^{n} \rho_{m,i} = \frac{n\bar{m}}{n\bar{V}} = \frac{\bar{m}}{\bar{V}}.$$

Less obviously, we can also show that the mean value of $\rho_{V,i}$ (see Table 7.5) is equal to the same ratio. Indeed, we have

$$\bar{\rho}_V = \frac{1}{n}\sum_{i=1}^{n} \rho_{V,i} = \frac{\bar{m}}{n}\sum_{i=1}^{n} \frac{1}{V_i}.$$

An individual measurement of the volume can be represented in the form $V_i = \bar{V} + v_i = \bar{V}(1 + \epsilon_i)$, where v_i is an absolute error and ϵ_i is the relative form of it. In accordance with the rules of approximate calculation,

$$\frac{1}{V_i} \simeq \frac{1}{\bar{V}_i}(1 - \epsilon_i).$$

As one can see from Table 7.3, the highest absolute value of v_i is $32 \times 10^{-10}\, m^3$. Thus, $\epsilon_i \sim 1.6 \times 10^{-3}\%$, and the above equality can be considered precise. Therefore,

$$\sum_{i=1}^{n} \frac{1}{V_i} = \sum_{i=1}^{n} \frac{1}{\bar{V}_i}(1 - \epsilon_i) = \frac{n}{\bar{V}} - \frac{1}{\bar{V}_i}\sum_{i=1}^{n} \epsilon_i.$$

As it was shown in Section 5.6, $\sum_{i=1}^{n} \epsilon_i = 0$. Thus, $\sum_{i=1}^{n} \frac{1}{V_i} = \frac{n}{\bar{V}}$, and therefore,

$$\bar{\rho}_V = \frac{\bar{m}}{n}\sum_{i=1}^{n} \frac{1}{V_i} = \frac{\bar{m}}{\bar{V}}$$

as was to be shown. Unfortunately, proving this result in general form seems difficult.

expression

$$I = CU/\tau,$$

where C is the capacitance of the capacitor, with whose help the ionization current is compensated; U is the initial voltage on the capacitor; and τ is the compensation time.

As U and τ are dependent, it is a dependent measurement.

We shall examine the measurement of ionization current on the special apparatus described in [33]. A capacitor, whose capacitance $C = 4006.3$ pF is known to within 0.005%, is employed. The voltage on the capacitor is established with the help of a class 0.1 voltmeter with a measurement range of 0–15 V. The time is measured with a timer whose scale is divided into tenths of a second.

The measurement is performed by making repeated observations. Each time the same indication of the voltmeter $U = 7$ V is established and the compensation time is measured. The results of 27 observations are given in the first column of Table 7.6.

TABLE 7.6. The results of measurements of the ionization current and data from initial processing.

τ s	I_i $\times 10^{-10}$ A	$(I_i - \bar{I})$ $\times 10^{-14}$ A	$(I_i - \bar{I})^2$ $\times 10^{-28}$ A^2
74.4	3.7694	7	49
74.6	3.7593	−94	8836
74.3	3.7745	58	3364
74.6	3.7593	−94	8836
74.4	3.7694	7	49
74.4	3.7694	7	49
74.4	3.7694	7	49
74.4	3.7694	7	49
74.4	3.7694	7	49
74.3	3.7745	58	3364
74.5	3.7643	−44	1936
74.4	3.7694	7	49
74.5	3.7643	−44	1936
74.4	3.7694	7	49
74.6	3.7593	−94	8836
74.2	3.7705	18	324
74.5	3.7643	−44	1936
74.3	3.7745	58	3364
74.4	3.7694	7	49
74.4	3.7694	7	49
74.5	3.7643	−44	1936
74.5	3.7643	−44	1936
74.3	3.7745	58	3364
74.3	3.7745	58	3364
74.3	3.7745	58	3364
74.4	3.7694	7	49
74.5	3.7643	−44	1936

The largest difference between the obtained values of the compensation time is equal to 0.4 s; i.e., the deviations from the average reach 0.25%. What can explain this spread? Obviously, the systematic errors of the measuring instrument employed here have nothing to do with the error. We shall estimate the random error of the instruments and their role.

According to established standards, the dead band of an electric measuring instrument must not exceed the limit of the intrinsic error permissible for it. This limit in indicating a voltage of 7 V is $\delta U = 0.1 \times (15/7) = 0.21\%$.

For this reason, when a voltage of 7 V is set on the voltmeter, voltages that can differ from the average by not more than one half the dead band, i.e., 0.1%, will be obtained on the capacitor.

The timer has virtually no random error.

The compensation time is a function of the voltage on the capacitor, and the spread in the voltage on the capacitor is accompanied by the same spread in the compensation time. The obtained spread of 0.25% is larger than expected. Therefore, the observed phenomenon should have a different reason.

In the experiment under study, the background current could be the reason. This current adds to the measured current and is indistinguishable from it. But it is known that the background current can be assumed to be a stationary process over a time interval shorter than that required to perform a measurement, which makes it possible to eliminate the background current by measuring during the experiment the average background current and subtracting it from the value obtained for the ionization current. For this reason, however, the ionization current must also be measured as an average current.

Essentially, in this manner, the model of the phenomenon was redefined and a new definition of the specific measured quantity was given.

Our measurement is a dependent measurement. For this reason, we shall use the method of reduction; for each value of τ_i ($i = 1, \ldots, 27$), we find the corresponding current I_i and then calculate, for the entire group of values, the average value, giving an estimate of the measured quantity.

The values of I_i are presented in the second column of Table 7.6. The average value $\bar{I} = 3.7687 \times 10^{-10}$ A.

We shall now estimate the errors. First we shall find an estimate of the standard deviation of the measurement result. As

$$\sum_{i=1}^{27} (I_i - \bar{I})^2 = 59\,171 \times 10^{-28},$$

we have

$$S(\bar{I}) = \sqrt{\frac{\sum_{i=1}^{27} (I_i - \bar{I})^2}{27 \times 26}} = 9.2 \times 10^{-14}\,\text{A}.$$

In the relative form, $S_{\%}(\bar{I}) = 0.027\%$.

We shall now estimate the conditionally constant errors. The influence coefficients (in the relative form) w_C, w_U, and w_τ of the errors of the estimates of the arguments $C, U,$ and τ are equal to $w_C = 1$, $w_U = 1$, and $w_\tau = -1$. Therefore the existing conditionally constant errors are related by the relation $\vartheta_{I\%} = \vartheta_{C\%} + \vartheta_{U\%} - \vartheta_{\tau\%}$.

For each elementary error, we estimate its limit $|\vartheta_i| \le \theta_i$.

The limit of the total error of the voltmeter (neglecting the sign) is equal to 0.21%. As the voltmeter also has a random error, we shall take $\theta_U = 0.15\%$. For the capacitor, $\theta_C = 0.05\%$ is given. The limit of the error of the timer is equal to the value of one graduation, i.e., $\theta_\tau = 0.1 \times 100/74 = 0.135\%$.

Turning to formula (6.42) and setting $\alpha = 0.95$, we obtain

$$\theta_{I\%} = k\sqrt{\theta_C^2 + \theta_U^2 + \theta_\tau^2}$$
$$= 1.1\sqrt{(5 \times 10^{-2})^2 + (13.5 \times 10^{-2})^2 + (15 \times 10^{-2})^2} = 0.23\% \quad (\alpha = 0.95).$$

The average background current is usually equal to $(0.5 - 1) \times 10^{-12}$ A. It can be measured to within 5%. With respect to the measured ionization current, this error is equal to 0.013%, and it can obviously be neglected.

If during the measurement the average background current was equal to $\bar{I}_b = 0.75 \times 10^{-12}$ A, then

$$\tilde{I} = \bar{I} - \bar{I}_b = 3.7612 \times 10^{-10}\,\text{A}.$$

So

$$\tilde{I} = 3.7612 \times 10^{-10}\,\text{A}, \qquad S(\tilde{I}) = 9 \times 10^{-14}\,\text{A} \quad (n = 27),$$
$$\theta_I = 8.7 \times 10^{-13}\,\text{A} \quad (\alpha = 0.95).$$

As $\theta/S > 7$, the random error can be neglected. After rounding off, we obtain finally

$$I = (3.761 \pm 0.009) \times 10^{-10}\,\text{A} \quad (\alpha = 0.95).$$

7.4. The Measurement of Power at High Frequency

As an example of a single independent indirect measurement, we shall study the measurement of the power generated by a high-frequency current in a resistor according to the formula $P = I^2R$,where P is the power measured, I is the effective current, and R is the active resistance of the resistor.

Measurements of the current and resistance give estimates of their values $\tilde{I}$ and $\tilde{R}$ and the limits of the relative errors $\delta I = 0.5\%$ and $\delta R = 1\%$.

The errors of measurements of arguments are given in the relative form. Therefore the influence coefficients are $w'_I = 2$ and $w'_R = 1$.

The limit of the error of the result must be found differently depending on what is known about the errors δI and δR. Even if these errors are determined only

by the properties of the instruments employed, several situations are possible. We shall study the most typical ones.

(1) The measurement is performed under reference conditions, the measuring instruments have been recently checked, and the working standards were at least five times more accurate than the instruments. In this case, it can be assumed that the limits of error of the measurements of the current and resistance δI and δR are reliable. Assuming that within the estimated limits the actual errors ε_I and ε_R are distributed uniformly over the set of instruments, the uncertainty of the measurement result can be found according to the formula

$$u_1(\alpha) = k\sqrt{4(\delta I)^2 + (\delta R)^2}.$$

For $\alpha = 0.95$, $k = 1.1$, and we obtain

$$u_1(0.95) = 1.1\sqrt{4 \times 0.25 + 1} = 1.5\%.$$

(2) The measurement is performed under the normal operating conditions appropriate for the instruments employed and the estimates of the errors δI and δR are determined based on several components. For this reason, the errors of argument measurements can be assumed to have normal distribution. If their limits correspond to the same probability α.

Then the uncertainty of the result of measurement can be found from the formula (6.20).

$$u_2(\alpha) = \sqrt{4(\delta I)^2 + (\delta R)^2} = 1.4\%.$$

This uncertainty estimate is just as reliable as the estimates of its components and corresponds to the same confidence probability.

(3) Another situation will occur if we do not have information about the origin of the limits of errors δI and δR. In this case, and because there are only two components of the measurement error, we shall sum them arithmetically:

$$\delta P = 2\delta I + \delta R = 2\%.$$

7.5. The Measurement of Voltage with the Help of a Potentiometer and a Voltage Divider

The measurement of voltage with the help of a potentiometer is a direct measurement. However, when the errors of the potentiometers and the errors of the standard cell are standardized separately, and when measurements are performed using a voltage divider, the error of the result of such a measurement is estimated by methods that are specifically designed for indirect measurements.

We shall study the case of single measurements with accurate estimation of errors, for example, measurement of the voltage with the help of a class 0.005 P309 potentiometer, a class 0.005 standard cell, and a class 0.005 P35 voltage divider. These instruments were manufactured in the former USSR.

It is well known that when working with such potentiometers, at first the potentiometrical current I_p is adjusted in the circuit with accurate resistors so that the voltage drop on the section of the circuit with the resistance R_{sc} would balance the emf of the standard cell U_{sc}. In this case,

$$I_p = U_{sc}/R_{sc}.$$

Next, the standard cell is disconnected and the measured voltage U_p is connected to the potentiometer circuit. By switching the potentiometer, a fraction of the resistors of the potentiometer is introduced into the comparison circuit such that the voltage drop on their resistance R_p would compensate U_p; i.e., $U_p = I_p R_p$. Then

$$U_p = \frac{R_p}{R_{sc}} U_{sc},$$

and knowing the emf of the standard cell and the ratio R_p/R_{sc}, we find U_p.

The indications of the potentiometer are proportional to R_p, but the error of the potentiometer is determined not by the errors of the resistances R_p and R_{sc}, but by the error of the ratio R_p/R_{sc}. The uncertainty associated with the operations of comparing the voltages can be neglected, because the smoothness of the unit controlling the potentiometer and the sensitivity of the zero indicator were designed so that this condition would be satisfied.

The potentiometer has six decades and a built-in self-balancing amplifier. The limit of permissible error as a function of the measured voltage U is calculated using the formula (given in the manufacturer's documentation)

$$\Delta U = \pm(50U + 0.04) \times 10^{-6} \text{ V}.$$

The error of the potentiometer does not exceed the indicated limits if the ambient air temperature ranges from $+15$ to $+30$ °C and differs by not more than 2.5 °C from the temperature at which the measuring resistors of the potentiometer were adjusted (the P309 potentiometer has built-in calibration and adjusting systems).

The emf of the class 0.005 standard cell can be determined with an error of ± 10 μV. The effect of the temperature is taken into account with the help of a well-known formula, which describes accurately the temperature dependence of the emf.

Assume that in a measurement of one and the same voltage, performed using a voltage divider whose voltage division ratio was set equal to 1:10, the following potentiometer indications were obtained:

$$x_1 = 1.256\,316 \text{ V}, \qquad x_2 = 1.256\,321 \text{ V}, \qquad x_3 = 1.256\,318 \text{ V}.$$

The limit of permissible error of the potentiometer in this case is

$$\Delta U = \pm 50 \times 1.26 \times 10^{-6} = \pm 63 \text{ μV}.$$

For this reason, the difference of 5 μV between the results of the three observations presented above can be regarded as resulting from the random error of the

measurement, whose magnitude is acceptable. In the calculation, therefore, any results obtained or their average value can be used.

In the process of adjusting the measuring resistors, which is done before the measurement, the corrections of the higher order decades were estimated. We shall introduce them into the indications of the potentiometer.

Let the correction for indication "12" of the decade "$\times 100$ mV" equal $+15 \times 10^{-6}$ V, and the correction of the indication "5" of the decade "$\times 10$ mV" equal -3×10^{-6} V. The corrections for the other decades are so small that they are no longer of interest. Each correction is determined with an error of $\pm 5 \times 10^{-8}$ V. The error of the potentiometer corresponding to the indications of the remaining decades that are 0.0063 V falls within the limits determined in accordance with the formula given above and are equal to

$$\Delta U = \pm(50 \times 0.0063 + 0.04) \times 10^{-6} = \pm 0.4 \times 10^{-6} \text{ V}.$$

In addition, it is necessary to take into account the possible change in the air temperature in the room. If this change falls within permissible limits, then according to the specifications of the potentiometer, the error can change approximately by $\frac{1}{4}$ of the permissible limit, i.e., by 16 μV.

We shall take for the result the average value of the observations performed, correcting it by the amount $C = (15 - 3) \times 10^{-6} = 12 \times 10^{-6}$ μV:

$$U_p = \bar{x} = 1.256\,318 + 0.000\,012 = 1.256\,330 \text{ V}.$$

The errors of the potentiometer, which enter into this result, are

$$\theta_1 = \pm 16 \times 10^{-6} \text{ V}, \qquad \theta_2 = \pm 0.4 \times 10^{-6} \text{ V}.$$

The error in determining the corrections and the error θ_2 can be neglected.

Thus, the limits of error of the potentiometer are equal to θ_1:

$$\theta_p = \theta_1 = \pm 16 \times 10^{-6} \text{ V}.$$

Next, we must estimate the errors from the standard cell and the voltage divider. Assuming that the division coefficient of the voltage divider is equal to K_d, the measured voltage is determined from formula $U_x = K_d U_p$, where

$$U_p = \frac{R_p}{R_{\text{sc}}} U_{\text{sc}},$$

and for this reason, we can write the measurement equation in the form:

$$U_x = K_d \frac{R_p}{R_{\text{sc}}} U_{\text{sc}}.$$

The error of the voltage divider can reach $5 \times 10^{-3}\%$. But the real division coefficient of the divider can be found and taken into account, which is precisely what we must do in the case at hand. In the given measurement, $K_d = 10.0003$ and the error in determining K_d falls within the range $\pm 2 \times 10^{-3}\%$.

The emf of the standard cell is taken into account with the help of the special decades of the potentiometer. The discrepancy between the real value of the emf

of the standard cell and the value exhibited on the potentiometer falls within the limits of error in determining the emf of the standard cell ($\pm 10\,\mu$V).

We estimate the measured voltage U_x as

$$\tilde{U}_x = K_d U_p = 10.0003 \times 1.256\,330 = 12.563\,68 \text{ V}.$$

To estimate the measurement error, we shall use the standard trick. First, we shall take the logarithm of the obtained measurement equation. Then we find the differentials of both sides of the equation, and neglecting errors that are second-order infinitesimals, we replace the differentials by the increments. This process gives

$$\frac{\Delta U_x}{U_x} = \frac{\Delta K_d}{K_d} + \frac{\Delta(R_p/R_{\text{sc}})}{R_p/R_{\text{sc}}} + \frac{\Delta U_{\text{sc}}}{U_{\text{sc}}}.$$

For the terms, we have only estimates of the limits, and not the values of the errors. For this reason, we shall estimate the limits of the measurement error. So, we can use formula (5.3). First, all components must be represented in the form of relative errors. The relative error of the potentiometer, more accurately, its limits in percent, will be

$$\theta_{p\%} = \frac{100\theta_p}{U_p} = \frac{100\Delta(R_p/R_{\text{sc}})}{R_p/R_{\text{sc}}} = \pm\frac{100 \times 16 \times 10^{-6}}{1.26} = \pm 1.3 \times 10^{-3}\%.$$

The limits of the relative error of the voltage divider were estimated directly as $\theta_{K\%} = \pm 2 \times 10^{-3}\%$. The limits of error in determining the emf of the standard cell in the form of a relative error will be

$$\theta_{\text{sc}\%} = \pm\frac{100 \times 10 \times 10^{-6}}{1.018} = \pm 1 \times 10^{-3}\%.$$

We now find the limit of the measurement error according to formula (5.3):

$$\theta_{\alpha\%} = k\sqrt{1.3^2 + 2^2 + 1^2} \times 10^{-3} = k \times 2.6 \times 10^{-3}\%.$$

Let $\alpha = 0.95$. Then $k = 1.1$ and

$$\theta_{0.95\%} = 1.1 \times 2.6 \times 10^{-3} = 2.9 \times 10^{-3} \approx 3 \times 10^{-3}\%.$$

Finally, we must check the number of significant figures in the result of measurement. For this reason, we shall put the limits $\theta_{\%}$ in the form of absolute errors

$$\theta_{0.95} = \pm 2.9 \times 10^{-3} \times 10^{-2} \times 12.6 = \pm 37 \times 10^{-5} \text{ V}.$$

As the measurement is accurate, the error in the result of measurement is expressed by two significant figures and no extra figures are in the result obtained. The final result is

$$U_x = (12.563\,68 \pm 0.000\,37) \text{ V} \quad (\alpha = 0.95).$$

If the measurement was performed with approximate estimation of the errors, then the errors of all components would have to be set equal to $5 \times 10^{-3}\%$ and the

limit of the measurement error would be

$$\theta'_{0.95\%} = 1.1 \times 10^{-3}\sqrt{3 \times 5^2} = 0.01\%.$$

Then $\theta' = \pm 0.0013$ V and the result of measurement would have to be written with fewer significant figures:

$$U'_x = (12.5637 \pm 0.0013)\ \text{V} \quad (\alpha = 0.95).$$

Here two significant figures are retained in the numerical value of the measurement error because the value of the most significant digit is less than 3.

7.6. Calculation of the Uncertainty of the Value of a Compound Resistor

We shall study the case in which 12 resistors with three different nominal resistances are connected in series:

$$R_\Sigma = 2R_1 + 4R_2 + 6R_3.$$

This equation is a particular case of the dependence (6.2) with $N = 3, b_1 = 2, b_2 = 4$, and $b_3 = 6$.

For resistors each having a nominal resistance R_i the limits of permissible errors $\theta_{i\alpha}$ are known:

i	R_i Ω	$\theta_{i\alpha}$ Ω
1	100.00	0.03
2	10.00	0.02
3	1.00	0.01

For the case at hand, we shall assume that the distribution of the actual resistance over the collection of resistors having the same nominal resistance is normal and truncated at the probability $\alpha = 0.98$.

The nominal resistance of our compound resistor, according to relation (6.2), is equal to

$$R_\Sigma = 2 \times 100.00 + 4 \times 10.00 + 6 \times 1.00 = 246.00\ \Omega.$$

We shall find the uncertainty in the value of the compound resistor for $\alpha = 0.98$ using formula (6.20) having in mind that now $w_j = b_i$ and $u_j = \theta_i$:

$$u = \sqrt{\sum_{i=1}^{3} b_i^2 \theta_i^2} = \sqrt{2^2 \times 0.03^2 + 4^2 \times 0.02^2 + 6^2 \times 0.01^2} = 0.11\ \Omega.$$

Rounding off, we obtain $u = 0.1\,\Omega$. Finally, taking into account the required number of significant figures, we can write

$$R_\Sigma = 246.0 \pm 0.1\,\Omega, \qquad \alpha = 0.98.$$

If it must be assumed that the real resistances of the resistors are distributed uniformly, then the uncertainty must be calculated using formula (6.43); in which case, we would obtain for the same confidence probability $\alpha = 0.98$

$$u' = ku = 1.3 \times 0.11 = 0.14\,\Omega.$$

The difference between u' and u for many cases is significant.

We shall consider the variant when N resistors with the same nominal resistance and the same tolerance are connected in series:

$$R_\Sigma = NR, \qquad \theta_R = \text{const}.$$

For example, $R = 100\,\Omega$, $\theta_{R\%} = 0.5\%$, and $N = 10$. What is the error of a compound resistor?

The problem is not as simple as it looks, because its solution depends on the technology employed to fabricate the resistors. Suppose that resistors fabricated at different times using different equipment are connected together. In this case, their errors are independent, and the error of each resistor can be regarded as a realization of a uniformly distributed random quantity. Then, according to formula (6.43), we have

$$\theta_\alpha = k\sqrt{\sum_{i=1}^{N}\theta_R^2} = k\theta_R\sqrt{N}.$$

It is convenient to transform this formula so that it would contain the relative errors. For this reason, we shall divide both sides of the equation by $R_\Sigma = NR$ and write $100 \times \theta/R_\Sigma = \theta_\%$, $100 \times \theta_R/R = \theta_{R\%}$. Then we obtain

$$\theta_\% = (k/\sqrt{N})\theta_{R\%}.$$

In the case at hand, we must focus on a high confidence probability. Let $\alpha = 0.99$. Then $k = 1.4$ and

$$\theta_{0.99\%} = 1.4/\sqrt{10} \times 0.5 = 0.2\%;$$

i.e., the accuracy of a compound resistor is higher than that of a single resistor. However, the increase in accuracy is limited by the accuracy of the measuring instrument used in checking the resistance of the resistors.

Sometimes a compound resistor can be made up from resistors whose resistance was adjusted individually by the same operator with the help of the same measuring instrument, which is most often possible especially when accurate resistors are fabricated. The actual errors in all resistors must become approximately equal to one another, and because they are all, essentially, systematic, the relative error of the compound resistor will become the same as that of the separate resistors.

Expanding on this example, we can move on to the case when some additive quantity is measured in several identical applications of the same measuring instrument, for example, measurement of the length of a flat body with the help of a short ruler. It is obvious from the foregoing discussion that the relative systematic component of the error of such a length measurement will be equal to the relative systematic error of the ruler. The random component, however, must be estimated for a concrete measurement.

8
Combined Measurements

8.1. General Remarks About the Method of Least Squares

Combined measurements, as pointed out in Chapter 1, are measurements performed so that the number of equations relating the measured quantities is larger than the number of the latter. Because of measurement errors, it is impossible to find values of the unknowns such that all equations would be satisfied. Under these conditions, the estimated values of the unknowns usually are found with the help of the method of least squares.

The method of least squares is a widely employed computational technique that makes it possible to eliminate the nonuniqueness of experimental data. This method is easily implemented with the help of computers, and good least-squares software is available.

There is an extensive literature on the method of least squares, and it has been well studied. It is known that this method does not always give results that satisfy the criteria of optimality of estimation theory. Nevertheless, the method of least squares is widely employed, because in general, it is simple, and the biasness of the estimates obtained is usually not significant.

An alternative to the least-squares method is the method of minimizing the sum of absolute deviations. This method provides even more visual results than the first one. Nevertheless, the application of this method is not a problem now, it is seldom used.

In the first edition of this book, the classification of measurements contained combined and simultaneous categories of measurements.

An example of simultaneous measurements is finding the parameters of the equation

$$R = R_{20} + a(t - 20) + b(t - 20)^2,$$

which expresses the temperature dependence of an accurate measuring resistor.

By measuring simultaneously R (the resistance of the resistor) and t (the temperature of the resistor) and by varying the temperature, we obtain several equations, from which it is necessary to find R_{20}—the resistance of the resistor at $t = 20\,^{\circ}\mathrm{C}$—and the temperature coefficients a and b.

But, in accordance with the new results in the theory of indirect measurements, this example is an example of dependent indirect measurements, and the method of reduction is pertinent to it. So, generalizing the discussed example, we can say that the category of simultaneous measurements is not necessary, and this category was eliminated from this book.

We shall discuss the method of least squares, because it is the main computational method used for combined measurements, and to use this method knowingly, it is necessary to know its basic ideas.

We can write the basic equation in the general form

$$F_0(A, B, C, \ldots, x, y, z, \ldots) = l, \tag{8.1}$$

where x, y, z, and l are known coefficients and directly measured quantities, and A, B, and C are the unknowns to be determined.

Substituting the experimentally obtained numerical values of x_i, y_i, and z_i into (8.1), we obtain a series of equations of the form .

$$F_i(A, B, C, \ldots, x_i, y_i, z_i) = l_i, \tag{8.2}$$

which contains only the unknown quantities A, B, and C to be found and the numerical coefficients or numbers.

The quantities sought are found by solving the obtained equations simultaneously.

An example of a combined measurement is finding the capacitances of two capacitors from the measurements of the capacitance of each one of them separately, as well as when the capacitors are connected in parallel and in series. Each measurement is performed with one observation, but ultimately, we shall have four equations for two unknowns:

$$C_1 = x_1, \qquad C_2 = x_2, \qquad C_1 + C_2 = x_3, \qquad \frac{C_1 C_2}{C_1 + C_2} = x_4.$$

Substituting into these equations the experimentally found values of x_i, we obtain a system of equations analogous to (8.2).

As we have already pointed out, the number of equations in the system (8.2) is greater than the number of unknowns, and because of measurement errors, it is impossible to find values of the measured quantities such that all equations would be satisfied simultaneously, even if they are equations known exactly. For this reason, (8.2), in contrast to normal mathematical equations, are said to be *conditional equations*. When the values of the unknowns found by some method are substituted into the conditional equations (8.2), for the reasons mentioned, we obtain

$$F_i(\tilde{A}, \tilde{B}, \tilde{C}, \ldots) - l_i = v_i \neq 0.$$

The quantities v_i are called residuals. The solution of the conditional equation that minimizes the sum of the squares of the residuals is generally recognized. This proposition was first published by Legendre and is called Legendre's principle. He implemented this principle by the method that is now called the method of least squares.

8.2. Measurements with Linear Equally Accurate Conditional Equations

To simplify the formulas, we shall consider the case of three unknowns. Let the system of conditional equations have the form

$$Ax_i + By_i + Cz_i = l_i \quad (i = 1, \ldots, n,\ n > 3), \tag{8.3}$$

where A, B, and C are the unknowns to be determined, and x_i, y_i, z_i, and l_i are the results of the ith series of measurements and known coefficients.

In the general case, the number of unknowns $m < n$; if $m = n$, then the system of conditional equations can be solved uniquely, although the obtained results are burdened with errors.

If some estimates of the measured quantities $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ are substituted into (8.3), then we obtain the residuals

$$v_i = \tilde{A}x_i + \tilde{B}y_i + \tilde{C}z_i - l_i.$$

We shall find estimates of $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ from the conditions

$$Q = \sum_{i=1}^{n} v_i^2 = \min.$$

For this condition to be satisfied, it is necessary that

$$\frac{\partial Q}{\partial \tilde{A}} = \frac{\partial Q}{\partial \tilde{B}} = \frac{\partial Q}{\partial \tilde{C}} = 0.$$

We shall find these particular derivatives and equate them to 0:

$$\frac{\partial Q}{\partial \tilde{A}} = 2\sum_{i=1}^{n}(\tilde{A}x_i + \tilde{B}y_i + \tilde{C}z_i - l_i)x_i = 0,$$

$$\frac{\partial Q}{\partial \tilde{B}} = 2\sum_{i=1}^{n}(\tilde{A}x_i + \tilde{B}y_i + \tilde{C}z_i - l_i)y_i = 0,$$

$$\frac{\partial Q}{\partial \tilde{C}} = 2\sum_{i=1}^{n}(\tilde{A}x_i + \tilde{B}y_i + \tilde{C}z_i - l_i)z_i = 0.$$

From here we obtain a system of so-called normal equations:

$$\begin{aligned}
\tilde{A}\sum_{i=1}^{n} x_i^2 + \tilde{B}\sum_{i=1}^{n} x_i y_i + \tilde{C}\sum_{i=1}^{n} x_i z_i &= \sum_{i=1}^{n} x_i l_i,\\
\tilde{A}\sum_{i=1}^{n} y_i x_i + \tilde{B}\sum_{i=1}^{n} y_i^2 + \tilde{C}\sum_{i=1}^{n} y_i z_i &= \sum_{i=1}^{n} y_i l_i,\\
\tilde{A}\sum_{i=1}^{n} z_i x_i + \tilde{B}\sum_{i=1}^{n} z_i y_i + \tilde{C}\sum_{i=1}^{n} z_i^2 &= \sum_{i=1}^{n} z_i l_i.
\end{aligned}$$

The normal equations are often written using Gauss's notation:

$$\sum_{i=1}^{n} x_i^2 = [xx], \qquad \sum_{i=1}^{n} x_i y_i = [xy], \qquad \text{and so on.}$$

It is obvious that

$$\sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} y_i x_i \qquad \text{and therefore} \qquad [xy] = [yx].$$

In Gauss's notation, the normal equations assume the simpler form

$$\begin{aligned}
[xx]\tilde{A} + [xy]\tilde{B} + [xz]\tilde{C} &= [xl],\\
[xy]\tilde{A} + [yy]\tilde{B} + [yz]\tilde{C} &= [yl],\\
[xz]\tilde{A} + [yz]\tilde{B} + [zz]\tilde{C} &= [zl].
\end{aligned} \tag{8.4}$$

We call attention to two obvious but important properties of the matrix of coefficients of the unknowns in the system of equations (8.4):

(1) The matrix of these coefficients is symmetric relative to the main diagonal.
(2) All elements on the main diagonal are positive.

These properties are general. They do not depend on the number of unknowns, but in this example, they are shown in application to the case with three unknowns.

The number of normal equations is equal to the number of unknowns, and solving these equations by known methods we obtain estimates of the measured quantities. The solution can be written most compactly with the help of the determinants

$$\tilde{A} = \frac{D_x}{D}, \qquad \tilde{B} = \frac{D_y}{D}, \qquad \tilde{C} = \frac{D_z}{D}, \tag{8.5}$$

where

$$D = \begin{vmatrix} [xx] & [xy] & [xz] \\ [yx] & [yy] & [yz] \\ [zx] & [zy] & [zz] \end{vmatrix}.$$

The determinant D_x is obtained from the principal determinant D of the system by replacing the column with the coefficients of the unknown $\tilde{A}$ with the column

of free terms:

$$D_x = \begin{vmatrix} [xl] & [xy] & [xz] \\ [yl] & [yy] & [yz] \\ [zl] & [zy] & [zz] \end{vmatrix}.$$

The determinants D_y and D_z are found analogously, i.e., by replacing the second and third columns, respectively, with the indicated column.

Now we must estimate the errors of the obtained results. The estimate of the variance of the conditional equations is calculated from the formula

$$S^2 = \frac{\sum_{i=1}^{n} v_i^2}{n-m}, \tag{8.6}$$

where v_i is the residual of the ith conditional equation. Then the estimates of the variances of the values found for the unknowns can be calculated using the formulas

$$S^2(\tilde{A}) = \frac{D_{11}}{D} S^2, \qquad S^2(\tilde{B}) = \frac{D_{22}}{D} S^2, \qquad S^2(\tilde{C}) = \frac{D_{33}}{D} S^2, \tag{8.7}$$

where D_{11}, D_{22}, and D_{33} are the algebraic complements of the elements $[xx]$, $[yy]$, and $[zz]$ of the determinant D, respectively (they are obtained by removing from the matrix of the determinant D the column and row whose intersection is the given element).

The confidence intervals for the true values of the measured quantities are constructed based on Student's distribution. In this case, the number of degrees of freedom for all measured quantities is equal to $\nu = n - m$.

8.3. Reduction of Linear Unequally Accurate Conditional Equations to Equally Accurate Conditional Equations

In Section 8.2, we studied the case in which all conditional equations had the same variance. Such conditional equations are said to be equally accurate. In practice, there can be cases in which the conditional equations have different variances, which usually happens if equations reflecting the measurements performed under different conditions are added to the system of equations. For example, if in calibrating a collection of weights, special measures are not taken, because of the different loading of the weights, the weighing errors will be different for different combinations of weights. Correspondingly, the conditional equations will not be equally accurate either.

For unequally accurate conditional equations, the most likely set of values of the unknowns $A, B, C, \ldots$ will be obtained if the expression

$$Q = \sum_{i=1}^{n} g_i v_i^2,$$

where g_i is the weight of the ith conditional equation, is minimized.

The introduction of weights is equivalent to multiplying the conditional equations by $\sqrt{g_i}$. Finally, the cofactors g_i will appear in the coefficients of the unknowns in the normal equations.

Thus, the first equation of the system of normal equations (8.4) will assume the form

$$[gxx]\tilde{A} + [gxy]\tilde{B} + [gxz]\tilde{C} + [gxl] = 0.$$

All remaining equations will change analogously. Each coefficient in the equation is a sum of terms of the form

$$[gxy] = g_1 x_1 y_1 + g_2 x_2 y_2 + \cdots + g_n x_n y_n.$$

The weights of the conditional equations are found from the conditions

$$\sum_{i=1}^{n} g_i = 1,$$
$$g_1 : g_2 : \cdots : g_n = \tfrac{1}{\sigma_1^2} : \tfrac{1}{\sigma_2^2} : \cdots : \tfrac{1}{\sigma_n^2}.$$

Therefore, to solve the problem, it is necessary to know the variance of the conditional equations. If the weights have been determined (or chosen), then after the transformations presented above, the further solution of the problem proceeds in the manner described in Section 8.2, and finally we obtain estimates of the measured quantities and their rms deviations. However, the weights are usually determined approximately.

8.4. Linearization of Nonlinear Conditional Equations

For several fundamental reasons, the method of least squares has been developed only for linear conditional equations. For this reason, nonlinear conditional equations must be put into a linear form.

The general method for doing this task is based on the assumption that the incompatibility of the conditional equations is small; i.e., their residuals are small. Then, taking from the system of conditional equations as many equations as there are unknowns and solving them, we find the initial estimates of the unknowns A_0, B_0, C_0. Next, assuming that

$$A = A_0 + a, \qquad B = B_0 + b, \qquad C = C_0 + c,$$

and substituting these expressions into the conditional equations, we expand the conditional equations in series. Let

$$F_i(A, B, C) = l_i.$$

Then retaining only terms with the first powers of the corrections a, b, and c, we obtain

$$f_i(A_0, B_0, C_0) - l_i + \left(\frac{\partial f_i}{\partial A}\right)_0 a + \left(\frac{\partial f_i}{\partial B}\right)_0 b + \left(\frac{\partial f_i}{\partial C}\right)_0 c = 0.$$

We find the partial derivatives by differentiating the functions $f_i(A, B, C)$ with respect to A, B, and C, respectively, and then we substitute A_0, B_0, and C_0 into the obtained formulas and find their numerical values. In addition,

$$f_i(A_0, B_0, C_0) - l_i = \lambda_i \neq 0.$$

Thus, we obtain a system of linear conditional equations for a, b, and c. The solution of this system gives their estimates and standard deviations. Then

$$\tilde{A} = A_0 + \tilde{a}, \qquad \tilde{B} = B_0 + \tilde{b}, \qquad C = C_0 + \tilde{c}.$$

As A_0, B_0, and C_0 are nonrandom quantities, $S^2(\tilde{A}) = S^2(\tilde{a})$, $S^2(\tilde{B}) = S^2(\tilde{b})$, and so on.

In principle, once $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ have been obtained, the second approximation can be constructed.

In addition to the foregoing method of linearization of the conditional equations, the method of substitutions is employed. Thus, if, for example, the conditional equation has the form

$$y_i = x_i \sin A + z_i e^{-2B},$$

where x, y, and z are directly measured quantities, and A and B must be determined, then the substitution

$$U = \sin A, \qquad E = e^{-2B},$$

can be made.

Then we obtain the linear conditional equation

$$y_i = x_i U + z_i E.$$

The solution of these equations gives $\tilde{U}$ and $\tilde{E}$ and estimates of their variances that can then be used to find the required quantities A and B.

The method of substitutions is convenient, but it is not always applicable. In principle, one can imagine one other general method for solving a system of equations when the number of equations is greater than the number of unknowns. This method is as follows.

Take from the available conditional equations a group of equations such that their number is equal to the number of unknowns. Such a group gives a definite value for each unknown.

Next, replacing in turn the equations in the group by each of the other equations that were not in the group, we obtain other values of the same unknowns. Irrespective of the method used to combine the equations, all possible combinations of equations must be sorted through, and for each combination, the values of the unknowns must be found. As a result of such calculations, obtain for each unknown a group of values that can be regarded as the group of observations obtained with direct measurements.

All values in the group are equivalent, but, unfortunately, they are not independent. This result presents difficulties in estimating the variances of the values obtained for the unknowns.

8.5. Examples of the Applications of the Method of Least Squares

The examples studied below are presented to demonstrate the computational technique as well as the physical meaning of the method, and for this reason, they were chosen so that the calculations would be as simple as possible. The initial data for the examples are taken from [35].

EXAMPLE 1. Determine the angles of a trihedral prism. Each angle is measured three times. The measurements of all angles are equally accurate. The results of all single measurements are as follows:

$$\begin{aligned} x_1 &= 89^\circ\,55', \qquad y_1 = 45^\circ\,5', \qquad z_1 = 44^\circ\,57', \\ x_2 &= 89^\circ\,59', \qquad y_2 = 45^\circ\,6', \qquad z_2 = 44^\circ\,55', \\ x_3 &= 89^\circ\,57', \qquad y_3 = 45^\circ\,5', \qquad z_3 = 44^\circ\,58'. \end{aligned}$$

If each angle is found as the arithmetic mean of the corresponding observations, then we obtain

$$A_0 = 89^\circ\,57', \qquad B_0 = 45^\circ\,5.33', \qquad C_0 = 44^\circ\,56.67'.$$

The sum of the angles of the traingle must satisfy the condition

$$A + B + C = 180^\circ.$$

We obtain $A_0 + B_0 + C_0 = 179^\circ\,59'$. This discrepancy is the result of measurement errors. The values of A_0, B_0, and C_0 obtained must be changed so that the exactly known condition is satisfied.

The relations between the unknowns that must be satisfied exactly are called constraints.

In [35], the problem is solved by the method of least squares. However, in this case, the method of least squares must be regarded only as a procedure leading to a unique answer. Its application here can be justified because the residual is equal to only $180^\circ - 179^\circ\,59' = 1'$, so that the changes in the directly obtained values

of the angles should be insignificant. Therefore, these changes need not be found with high accuracy. We now proceed to the solution of the problem.

If we have n conditional equations, m unknowns, and k constraints, and $n > m - k$ and $m > k$, then k unknowns can be eliminated from the conditional equations by expressing these unknowns for the remaining unknowns. Next, using the method of least square, we find the values of $m - k$ unknowns and the estimates of standard deviations of these estimates. We obtain the remaining k unknowns using the constraint equations. To find their standard deviations, strictly speaking, another cycle of calculations with the conditional equations, in which the previously excluded unknowns are retained and the other unknowns are excluded, must be performed. This repeated calculation is not often performed, because any conclusion about the standard deviation of the previously excluded unknowns can be made using the estimate of the standard deviation of the other unknowns.

Let us return to our problem. To simplify the calculations, we shall assume that

$$A = A_0 + a, \qquad B = B_0 + b, \qquad C = C_0 + c,$$

and we shall find the values of the corrections a, b, and c.

The system of conditional equations transforms into the following system:

$$\begin{array}{lll} a_1 = -2', & b_1 = -0.33', & c_1 = +0.33', \\ a_2 = +2', & b_2 = +0.67', & c_2 = -1.67', \\ a_3 = 0, & b_3 = -0.33', & c_3 = +1.33'. \end{array}$$

The constraint equation will assume the form

$$A_0 + a + B_0 + b + C_0 + c = 180^\circ.$$

Therefore

$$a + b + c = 180^\circ - 179^\circ\, 59' = 1'.$$

We exclude c from the conditional equations using the relation

$$c = 1' - a - b,$$

and in each equation, we indicate both unknowns. We obtain the following system of conditional equations:

$$\begin{array}{lll} 1 \times \tilde{a} + 0 \times \tilde{b} = -2', & 0 \times \tilde{a} + 1 \times \tilde{b} = -0.33', & 1 \times \tilde{a} + 1 \times \tilde{b} = +0.67', \\ 1 \times \tilde{a} + 0 \times \tilde{b} = +2', & 0 \times \tilde{a} + 1 \times \tilde{b} = 0.67', & 1 \times \tilde{a} + 1 \times \tilde{b} = +2.67', \\ 1 \times \tilde{a} + 0 \times \tilde{b} = 0', & 0 \times \tilde{a} + 1 \times \tilde{b} = -0.33', & 1 \times \tilde{a} + 1 \times \tilde{b} = -0.33'. \end{array}$$

We now construct the system of normal equations. Its general form will be

$$[xx]\tilde{a} + [xy]\tilde{b} = [xl],$$
$$[xy]\tilde{a} + [yy]\tilde{b} = [yl],$$

Here we obtain:

$$\begin{aligned}
[xx] &= 1+1+1+1+1+1 = 6,\\
[xy] &= 1+1+1 = 3,\\
[yy] &= 1+1+1+1+1+1 = 6,\\
[xl] &= -2' + 2' + 0.67' + 2.67' - 0.33' = +3',\\
[yl] &= -0.33' + 0.67' - 0.33' + 0.67' + 2.67' - 0.33' = +3'.
\end{aligned}$$

Therefore, the normal equations will assume the form

$$6\tilde{a} + 3\tilde{b} = 3', \qquad 3\tilde{a} + 6\tilde{b} = 3'.$$

In accordance with the relations (8.5), we calculate

$$D = \begin{vmatrix} 6 & 3 \\ 3 & 6 \end{vmatrix} = 36 - 9 = 27,$$

$$D_a = \begin{vmatrix} 3' & 3 \\ 3' & 6 \end{vmatrix} = 18' - 9' = 9'.$$

$$D_b = \begin{vmatrix} 6 & 3' \\ 3 & 3' \end{vmatrix} = 18' - 9' = 9',$$

and we find

$$\tilde{a} = \tilde{b} = 9'/27 = 0.33'.$$

Therefore, $\tilde{c} = 0.33'$ also.

Substituting the estimates obtained into the conditional equations, we calculate the residuals:

$$\begin{aligned}
v_i &= 2.33', \qquad & v_4 &= 0.67', \qquad & v_7 &= 0,\\
v_2 &= 1.67', \qquad & v_5 &= -0.33', \qquad & v_8 &= 2',\\
v_3 &= 0.33', \qquad & v_6 &= 0.67', \qquad & v_9 &= -1'.
\end{aligned}$$

From formula (8.6), we calculate an estimate of the variance of the equations

$$S^2 = \frac{\sum\limits_{i=1}^{n} v_i^2}{n-m+k} = \frac{\sum\limits_{i=1}^{9} v_i^2}{9-2} = \frac{14.34}{7} = 2.05.$$

Now $D_{11} = 6$, $D_{22} = 6$ and formulas (8.7) give

$$S^2(\tilde{a}) = S^2(\tilde{b}) = \frac{6}{27} \times 2.05 = 0.456, \qquad S(\tilde{a}) = S(\tilde{b}) = 0.675.$$

As the conditional equations are equally accurate and the estimates $\tilde{a}$, $\tilde{b}$, and $\tilde{c}$ are equal to one another, the repeated calculations need not be performed, and we can write immediately $S(\tilde{c}) = 0.675'$. Finally, we obtain $\tilde{A} = 89^\circ\,57.33'$, $\tilde{B} = 45^\circ\,5.67'$, $C = 44^\circ\,57.00'$, and $S(\tilde{A}) = S(\tilde{B}) = S(\tilde{C}) = 0.68'$.

We construct the confidence interval for each angle based on Student's distribution. The number of degrees of freedom in this case is equal to $9 - 2 = 7$, and for

$\alpha = 0.95$, Student's coefficient $t_{0.95} = 2.36$. Therefore, $\Delta_{0.95} = 2.36 \times 0.68' = 1.6'$. Thus, we obtain finally

$$A(0.95) = 89^\circ\,57.3' \pm 1.6', \qquad B(0.95) = 45^\circ\,5.7' \pm 1.6',$$
$$C(0.95) = 44^\circ\,57.0' \pm 1.6'.$$

EXAMPLE 2. We shall study the example, which was presented at the beginning of this chapter, of combined measurements of the capacitance of two capacitors. The results of the direct measurement are as follows:

$$x_1 = 0.2071\,\mu\text{F}, \qquad x_2 = 0.2056\,\mu\text{F},$$
$$x_1 + x_2 = 0.4111\,\mu\text{F}, \qquad \frac{x_1 x_2}{x_1 + x_3} = 0.1035\,\mu\text{F}.$$

The last equation is nonlinear. We expand it in a Taylor series, for which we first find the partial derivatives

$$\frac{\partial f}{\partial C_1} = \frac{C_2(C_1 + C_2) - C_1 C_2}{(C_1 + C_2)^2} = \frac{C_2^2}{(C_1 + C_2)^2}$$

and analogously

$$\frac{\partial f}{\partial C_2} = \frac{C_1^2}{(C_1 + C_2)^2}$$

As $C_1 \approx x_1$ and $C_2 \approx x_2$, we can write

$$C_1 = 0.2070 + e_1, \qquad C_2 = 0.2060 + e_2.$$

We make the expansion for the point with the corrdinates $C_{10} = 0.2070$ and $C_{20} = 0.2060$. We obtain

$$\frac{C_{10} C_{20}}{C_{10} + C_{20}} = 0.103\,25,$$
$$\frac{\partial f}{\partial C_1} = \frac{0.206^2}{(0.207 + 0.206)^2} = 0.249,$$
$$\frac{\partial f}{\partial C_2} = \frac{0.207^2}{(0.207 + 0.206)^2} = 0.251.$$

We find the conditional equations, setting $x_1 = C_1$ and $x_2 = C_2$:

$$1 \times e_1 + 0 \times e_2 = 0.0001,$$
$$0 \times e_1 + 1 \times e_2 = -0.0004,$$
$$1 \times e_1 + 1 \times e_2 = -0.0019,$$
$$0.249 e_1 + 0.251 e_2 = 0.000\,25.$$

We now calculate the coefficients of the normal equations

$$[xx] = 1 + 1 + 0.249^2 = 2.062, \qquad [xy] = 1 + 0.249 \times 0.251 = 1.0625,$$
$$[yy] = 1 + 1 + 0.251^2 = 2.063, \qquad [xl] = -0.0004 - 0.0019 + 0.249 \times 0.00025 = -0.001\,738,$$
$$[yl] = -0.0004 - 0.0019 + 0.251 \times 0.00025 = -0.002\,237.$$

The normal equations will be

$$2.062e_1 + 1.0625e_2 = -0.001\,738,$$
$$1.0625e_1 + 2.063e_2 = -0.002\,237.$$

We now find the unknowns e_1 and e_2. According to (8.5), we calculate

$$D = \begin{vmatrix} 2.062 & 1.0625 \\ 1.0625 & 2.063 \end{vmatrix} = 3.125,$$

$$D_x = \begin{vmatrix} -0.001\,738 & 1.0625 \\ -0.002\,237 & 2.063 \end{vmatrix} = -0.001\,22,$$

$$D_y = \begin{vmatrix} 2.062 & -0.001\,738 \\ 1.0625 & -0.002\,237 \end{vmatrix} = -0.002\,75.$$

From here we find

$$e_1 = \frac{D_x}{D} = -0.000\,39, \qquad e_2 = \frac{D_y}{D} = -0.000\,88.$$

Therefore,

$$\tilde{C}_1 = 0.2070 - 0.000\,39 = 0.206\,61\ \mu\text{F},$$
$$\tilde{C}_2 = 0.2060 = 0.000\,88 = 0.205\,12\ \mu\text{F}.$$

We find the residuals of the conditional equations by substituting the estimates obtained for the unknowns into the conditional equations:

$$v_1 = 0.000\,49, \qquad v_3 = -0.000\,63,$$
$$v_2 = 0.000\,58, \qquad v_4 = 0.000\,48.$$

Now we can calculate from formula (8.6) an estimate of the variance of the conditional equations:

$$S^2 = \frac{\sum_{i=2}^{4} v_i^2}{4-2} = \frac{120 \times 10^{-8}}{2} = 6 \times 10^{-7}.$$

The algebraic complements of the determinant D will be $D_{11} = 2.063$ and $D_{22} = 2.062$. As $D_{11} \approx D_{22}$,

$$S^2(\tilde{C}_1) = S^2(\tilde{C}_2) = \frac{D_{11}}{D} S^2 = \frac{2.063}{3.125} \times 6 \times 10^{-7} = 4 \times 10^{-7},$$
$$S(\tilde{C}_1) = S(\tilde{C}_2) = 6.3 \times 10^{-4}\ \mu\text{F}.$$

The method, examined above, for measuring the capacitances of the capacitors was apparently chosen to reduce somewhat the systematic error of the measurement, which is different at different points of the measurement range; to reduce the random component of the error, it would be sufficient to multiply measure each capacitance.

8.6. Determination of the Parameters in Formulas from Empirical Data and Construction of Calibration Curves

The purpose of almost any investigation in natural science is to find regularities in the phenomena in the material world, and measurements are the characteristic method that give objective data for achieving this goal.

It is desirable to represent the regular correlations determined between physical quantities based on measurements in an analytic form, i.e., in the form of formulas. The initial form of the formulas is usually established based on an unformalized analysis of the collection of data obtained. One important prerequisite of analysis is the assumption that the dependence sought must be expressed by a smooth curve; physical laws usually correspond to smooth curves. Once the form of the formula is chosen, its parameters are then found by an interpolation approximation of the empirical data by the formula obtained, and this is most often done by the method of least squares.

This problem is of great importance, and many mathematical and applied works are devoted to it. We shall discuss some aspects of the solution of this problem that are connected with the application of the method of least squares. The application of this method is based on the assumption that the criterion for the optimal choice of the parameter sought can be assumed to be that the sum of squares of the deviations of the empirical data from the curve obtained is minimized. This assumption is often justified, but not always. For example, sometimes the curve must be drawn so that it exactly passes through all prescribed points, which is natural, if the coordinates of the points mentioned are given as exact coordinates. The problem is solved by the methods of the interpolation approximation, and it is known that the degree of the interpolation polynomial will be only one less than the number of fixed points.

Sometimes the maximum deviation of the experimental data from the curve is minimized. As we have pointed out, however, most often the sum of the squares of the indicated deviations is minimized by the method of least squares. For this purpose, all values obtained for the quantities (in physically justified combinations) are substituted successively into the chosen formula. Ultimately, a system of conditional equations, from which the normal equations are constructed, is obtained; the solution of these equations gives the values sought for the parameters.

Next, substituting the values obtained for the parameters into the conditional equations, the residuals of these equations can be found and the standard deviation of the conditional equations can be estimated from them.

It is significant that in this case, the standard deviation of the conditional equations is determined not only by the measurement errors but also by the imperfect structure of the formula chosen to describe the dependence sought. For example, it is well known that the temperature dependence of the electric resistance of many metals is reminiscent of a parabola. In engineering, however, it is often

found that some definition section of this dependence can be approximated by a linear function. The inaccuracy of the chosen formula, naturally, is reflected in the standard deviation of the conditional equations. Even if all experimental data were free of any errors, the standard deviation would still be different from 0. For this reason, in this case, the standard deviation characterizes not only the error of the conditional equations, but also that the empirical formula adopted does not correspond to the true relation between the quantities.

In connection with what we have said above, the estimates, obtained by the method described above, of the variances of the determined parameters of the empirical formulas become conditional in the sense that they characterize not only the random spread in the experimental data, as usual, but also the uncertainty of the approximation, which is nonrandom.

It should be noted that if the empirical formula can be assumed to be linear, then the parameters of this formula can also be determined by the method of correlation and regression analysis. They also make it possible to construct the confidence intervals for the parameters and the confidence band for the approximating straight line. In this case, however, methods for indirect measurements are also effective.

Everything said above is completely relevant to the problem of constructing calibration curves of measuring transducers and instruments.

We shall discuss the problem of constructing linear calibration curves, which are most often encountered in practice.

Thus, the relation between a quantity y at the output of a transducer and the quantity x at the input of the transducer must be expressed by the dependence

$$y = a + bx. \tag{8.8}$$

When calibrating the transducer, the values of $\{x_i\}, i = 1, \ldots, n$, in the range $[x_{\min}, x_{\max}]$ are given and the corresponding values $\{y_i\}$ are found.

Using this data, we have to estimate the coefficients a and b. Let us start with the least-squares method.

Relation (8.8) gives a system of conditional equations

$$\tilde{b}_{xi} + \tilde{a} = y_i + v_i.$$

The residuals v_i are determined by the relation

$$v_i = \tilde{b}x_i + \tilde{a} - y_i.$$

Following the least-squares scheme presented above, we obtain the system of normal equations

$$\tilde{b}\sum_{i=1}^{n} x_i^2 + \tilde{a}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i, \qquad \tilde{b}\sum_{i=1}^{n} x_i + n\tilde{a} = \sum_{i=1}^{n} y_i. \tag{8.9}$$

The principal determinant of the system (8.9) will be

$$D = \begin{vmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{vmatrix} = n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2 .$$

The determinant D_x is given by

$$D_x = \begin{vmatrix} \sum_{i=1}^{n} x_i y_i & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} y_i & n \end{vmatrix} = n \sum_{i=1}^{n} (x_i y_i) - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i .$$

From here we find an estimate of the coefficient b:

$$\tilde{b} = \frac{D_x}{D} = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2} = \frac{\sum_{i=1}^{n} x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2} .$$

It is not difficult to show that

$$\sum_{i=1}^{n} x_i y_i - n \bar{x} \bar{y} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \tag{8.10}$$

and that

$$\sum_{i=1}^{n} x_i^2 - n \bar{x}^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 . \tag{8.11}$$

Then the expression for $\tilde{b}$ assumes the simpler form

$$\tilde{b} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} . \tag{8.12}$$

The determinant D_y is given by

$$D_y = \begin{vmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} y_i \end{vmatrix} = n \bar{y} \sum_{i=1}^{n} x_i^2 - n \bar{x} \sum_{i=1}^{n} x_i y_i .$$

Therefore

$$\tilde{a} = \frac{D_y}{D} = \frac{n \bar{y} \sum_{i=1}^{n} x_i^2 - n \bar{x} \sum_{i=1}^{n} x_i y_i}{n \sum_{i=1}^{n} x_i^2 - n^2 (\bar{x})^2} .$$

Using the identite (8.11), we put the estimate $\bar{a}$ into the form

$$\tilde{a} = \frac{\bar{y}\sum_{i=1}^{n} x_i^2 - \bar{x}\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}. \tag{8.13}$$

Relations (8.12) and (8.13) solve the problem; i.e., they determine the calibration curve

$$y = \tilde{a} + \tilde{b}x. \tag{8.14}$$

To evaluate the uncertainty of the calibration curve, the characteristics of each specific problem must be carefully analyzed.

From the experimental data and the obtained estimates $\tilde{a}$ and $\tilde{b}$, we find the residuals of the conditional equations

$$v_i = \tilde{a} + \tilde{b}x_i - y_i.$$

Next, according to the general scheme of the least-squares method, we calculate the estimate of variance of the conditional equations using formula (8.6),

$$S^2 = \frac{\sum_{i=1}^{n} v_i^2}{n-2},$$

and estimates of the variances of $\tilde{a}$ and $\tilde{b}$ using formulas (8.7). After this process, we find the confidence limits Δ_a and Δ_b, which essentially solves the problem. As pointed out above, the confidence limits are constructed based on Student's distribution with $n-2$ degrees of freedom in our case, because for us two parameters are being determined.

When working with measuring transducers the dependence $x = f(y)$ and not $y = \varphi(x)$ is required. Obviously, there is no difficulty in marking the transformation

$$x = (y - \tilde{a})/\tilde{b}.$$

If the calibration curve is plotted, then there is no need to convert it; it can simply be used in the "reverse" direction.

The methods of confluent analysis in application to the problem of constructing linear dependence are reviewed in [29].

Now we shall discuss how the theory of indirect measurements can be applied to the problem of constructing a linear calibration curve. Consider that we have an estimate $\tilde{a}$ of the coefficient a. Then (8.8) can be transformed into the form.

$$b = \frac{y - \tilde{a}}{x}.$$

This equation can be considered the measurement equation for the indirect measurement of the measurand b using the measuring arguments x and y. Because the values of y depend on the values of x, it is a dependent indirect measurement.

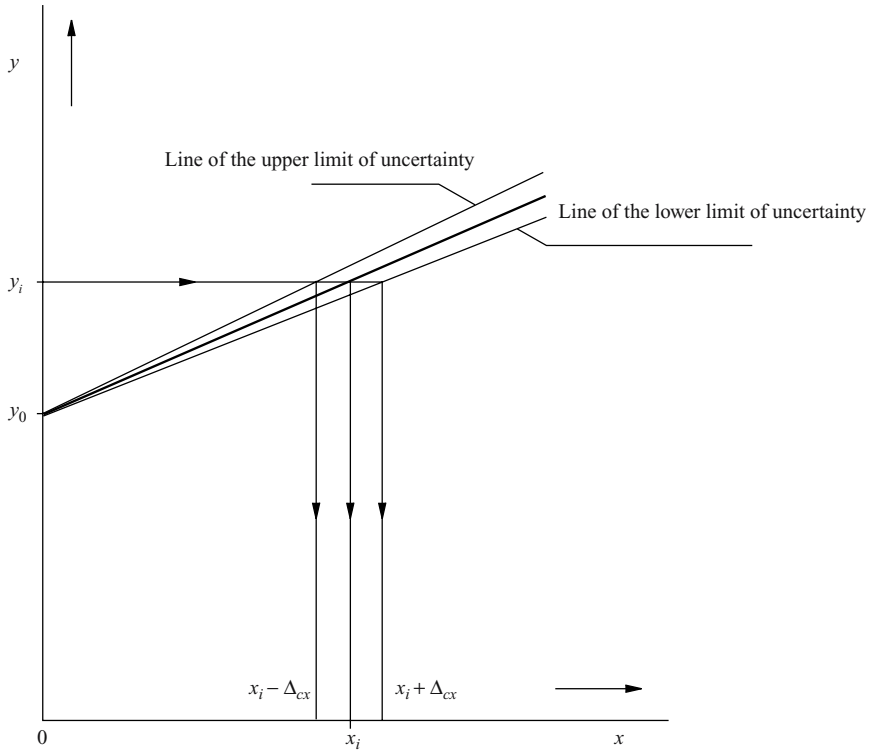


FIGURE 8.1. A calibration line and its band of uncertainty.

The estimate $\tilde{a}$ is easy to find: For $x_0 = 0$, (8.8) gives

$$\tilde{a} = y_0.$$

Calibration provides us with n pairs of x_i, y_i. Using the method of reduction, we transform this set of $\{x_i, y_i\}$ into a set of b_i:

$$\{b_i\}, \qquad i = 1, \ldots, n,$$

which allows us to get

$$\tilde{b} = \bar{b}, \qquad S(\bar{b})$$

and the band of confidence errors of $\bar{b}$.

Figure 8.1 shows an example of a calibration line found using the method of reduction and the band of its uncertainty.

Calibrating a transducer that we believe has a linear calibration curve, we have to check its linearity. We consider the calibration curve linear if its parameter b is constant or varies in defined permissible limits. These limits are given in the specifications of each type of transducer.

The total inaccuracy of a transducer is considered to be equal to the inaccuracy of the estimate $\tilde{b}$, which is defined by the standard deviation $S(\bar{b})$ and uncertainties of measurements of arguments x and y. The methods presented in Chapter 6 allow us to perform these calculations. The accuracy of the estimate a is, as a rule, much bigger than the accuracy of $\tilde{b}$ because the value $x_0 = 0$ does not need to be measured and the value y_0 can be adjusted. Thus, the inaccuracy of $\tilde{a}$ can often be neglected.

9
Combining the Results of Measurements

9.1. Introductory Remarks

Measurements of the same quantity are often performed in different laboratories and, therefore, under different conditions and by different methods. Sometimes there arises the problem of combining the data obtained to find the most accurate estimate of the measured quantity.

In many cases, in the investigation of new phenomena, measurements of the corresponding quantities take a great deal of time. By collecting into groups measurements performed over a limited time, intermediate estimates of the measured quantity can be obtained in the course of the measurements. It is natural to find the final result of a measurement by combining the intermediate results.

The examples presented show that the problem of combining the results of measurements is of great significance for metrology. At the same time, it is important to distinguish situations in which one is justified in combining results from those in which one is not justified in doing so. It is pointless to combine results of measurements in which quantities in the essence of different dimension were measured.

It should be noted that when comparing results of measurements, the data analysis is often performed based on the intuition of the experimenters without using formalized procedures. It is interesting that in the process, as a rule, the correct conclusions are drawn. On the one hand, this indicates that modern measuring instruments are of high quality and on the other hand that the experimenters, who by estimating the errors could determine all sources of error and exhibited reasonable care, were highly qualified.

9.2. Theoretical Principles

The following problem has a mathematically rigorous solution. Consider L groups of measurements of the same quantity A. Estimates of the measured quantity $\bar{x}_1, \ldots, \bar{x}_L$ were made from the measurements of each group, and

$$E[\bar{x}_1] = \cdots = E[\bar{x}_L] = A.$$

The variances of the measurements in each group $\sigma_1^2, \ldots, \sigma_L^2$ and the number of measurements in each group $n_1, \ldots, n_L$ are known.

The problem is to find an estimate of the measured quantity based on data from all groups of measurements. This estimate is denoted as $\bar{\bar{x}}$ and is called the combined average or weighted mean.

We shall seek $\bar{\bar{x}}$ as a linear function of $\bar{x}_j$; i.e., as their weighted mean,

$$\bar{\bar{x}} = \sum_{t=1}^{L} g_j \bar{x}_j. \tag{9.1}$$

Therefore the problem reduces to finding the weights g_j. As $E[\bar{\bar{x}}_j] = A$ and $E[\bar{x}] = A$, we obtain from (9.1)

$$E[\bar{\bar{x}}] = E\left[\sum_{j=1}^{L} g_j \bar{x}_j\right] = \sum_{j=1}^{L} g_j E[\bar{x}_j], \quad A = A\sum_{j=1}^{L} g_j.$$

Therefore,

$$\sum_{j=1}^{L} g_i = 1.$$

Next, we require that $\bar{\bar{x}}$ be an efficient estimate of A; i.e., $V[\bar{\bar{x}}]$ must be minimum. For this reason we find an expression for $V[\bar{\bar{x}}]$, using the formula

$$V[\bar{\bar{x}}] = V\left[\sum_{j=1}^{L} g_j \bar{x}_j\right] = \sum_{j=1}^{L} g_j^2 V[\bar{x}_j] = g_1^2\sigma^2(\bar{x}_1) + g_2^2\sigma^2(\bar{x}_2) + \cdots + g_L^2\sigma^2(\bar{x}_L). \tag{9.2}$$

Using the condition $\sum_{j=1}^{L} g_j = 1$, we write $g_L = 1 - g_1 - g_2 - \cdots - g_{L-1}$. We shall now find the condition under which $V[\bar{\bar{x}}]$ has a minimum. For this reason, we differentiate (9.2) with respect to g_j and equate the derivatives to 0. As we have $L - 1$ unknowns, we take $L - 1$ derivatives:

$$\begin{aligned}
2g_1\sigma^2(\bar{x}_1) - 2(1 - g_1 - g_2 - \cdots - g_{L-1})\sigma^2(\bar{x}_L) &= 0,\\
2g_2\sigma^2(\bar{x}_2) - 2(1 - g_1 - g_2 - \cdots - g_{L-1})\sigma^2(\bar{x}_L) &= 0,\\
&\cdots\\
2g_{L-1}\sigma^2(\bar{x}_{L-1}) - 2(1 - g_1 - g_2 - \cdots - g_{L-1})\sigma^2(\bar{x}_L) &= 0,
\end{aligned}$$

As the second term is identical in each equation, we obtain

$$g_1\sigma^2(\bar{x}_1) = g_2\sigma^2(\bar{x}_2) = \cdots = g_L\sigma^2(\bar{x}_L).$$

The transfer from g_{L-1} to g_L is made because the elimination of g_L was not dictated by some fundamental considerations, and instead of g_L, a weighting coefficient with any number could have been taken.

Thus, we have found a second condition that the weights of the arithmetic means of the groups of measurements must satisfy:

$$g_1 : g_2 : \cdots : g_L = \frac{1}{\sigma^2(\bar{x}_1)} : \frac{1}{\sigma^2(\bar{x}_2)} : \cdots : \frac{1}{\sigma^2(\bar{x}_L)}. \tag{9.3}$$

To find weight g_j, it is necessary to know either the variances of the arithmetic means or the ratio of the variances. If we have the variances $\sigma^2(\bar{x}_j)$, then we can set $g'_i = 1/\sigma^2(\bar{x}_j)$. We then obtain

$$g_j = \frac{g'_j}{\sum\limits_{j=1}^{L} g'_j}. \tag{9.4}$$

As the weights are nonrandom quantities, it is not difficult to determine the variance for $\bar{\bar{x}}$. According to relation (9.2), we have

$$V[\bar{\bar{x}}] = \sum_{j=1}^{L} g_j^2 V[\bar{x}_j] = \frac{\sum\limits_{j=1}^{L} (g'_j)^2 V[\bar{x}_j]}{\left(\sum\limits_{j=1}^{L} g'_j\right)^2} = \frac{\sum\limits_{j=1}^{L} \left(\frac{1}{\sigma^2(\bar{x}_j)}\right)^2 \sigma^2(\bar{x}_j)}{\left(\sum\limits_{j=1}^{L} \frac{1}{\sigma^2(\bar{x}_j)}\right)^2} = \frac{1}{\sum\limits_{j=1}^{L} \frac{1}{\sigma^2(\bar{x}_j)}}. \tag{9.5}$$

The relation (9.3) makes it possible to obtain the exact weights g_j if the variances $\sigma^2(\bar{x}_j)$ are not known but only their ratios are known. In this case, having estimates of the variances of the arithmetic means of the groups instead of their values, an expression can be derived for the estimate of the variance of the weighted mean:

$$S^2(\bar{\bar{x}}) = \frac{1}{N-1}\left(\sum_{j=1}^{L} g_j \frac{n_j - 1}{n_j} S_j^2 + \sum_{j=1}^{L} g_j (\bar{x}_j - \bar{\bar{x}})^2\right). \tag{9.6}$$

The particular case when the variances of the measurements are the same for all groups but the number of observations in the groups is different is of interest. In this case, we can set $g'_j = n_j$. Then the weights of the arithmetic means will be

$$g_j = n_j/N, \tag{9.7}$$

where $N = \sum_{j=1}^{L} n_j$ and the relation (9.6) will assume the form

$$S^2(\bar{\bar{x}}) = \frac{1}{N(N-1)}\left(\sum_{j=1}^{L} (n_j - 1) S_j^2 + \sum_{j=1}^{L} n_j (\bar{x}_j - \bar{\bar{x}})^2\right). \tag{9.8}$$

This result can also be obtained directly, combining the measurements of all groups into one large group of measurements.

The number of measurements in the combined group is $N = \sum_{j=1}^{L} n_j$.

If the measurements are collected according to groups, then the combined average will be

$$\bar{\bar{x}} = \frac{\sum_{j=1}^{L}\sum_{i=1}^{n_j} x_{ji}}{N}.$$

Let us expand the numerator, which gives

$$\begin{aligned}\bar{\bar{x}} &= \frac{(x_{11}+x_{12}+\cdots+x_{1n_1})+(x_{21}+x_{22}+\cdots+x_{2n_2})+\cdots}{N}\\ &= \frac{n_1\bar{x}_1+n_2\bar{x}_2+\cdots+n_L\bar{x}_L}{N} = \sum_{j=1}^{L} g_j\bar{x}_j,\end{aligned}$$

where $g_j = n_j/N$ is the weight of the jth arithmetic mean.

The aggregate average $\bar{\bar{x}}$, for this reason, is also called the weighted mean. The estimate of standard deviation of the weighted mean can be estimated by regarding the weighted mean as the average of the large group of combined measurements:

$$S^2(\bar{\bar{x}}) = \frac{\sum_{k=1}^{N}(x_k-\bar{\bar{x}})^2}{N(N-1)}.$$

We gather the terms in the numerator

$$S^2(\bar{\bar{x}}) = \frac{\sum_{j=1}^{L}\sum_{i=1}^{n_j}(x_{ji}-\bar{\bar{x}})^2}{N(N-1)}$$

and perform simple transformations of the numerator to simplify the calculations:

$$\begin{aligned}\sum_{j=1}^{L}\sum_{i=1}^{n_j}(x_{ji}-\bar{\bar{x}})^2 &= \sum_{j=1}^{L}\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j+\bar{x}_j-\bar{\bar{x}})^2\\ &= \sum_{j=1}^{L}\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)^2 + 2\sum_{j=1}^{L}\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)(\bar{x}_j-\bar{\bar{x}})\\ &\quad + \sum_{j=1}^{L}\sum_{i=1}^{n_j}(\bar{x}_j-\bar{\bar{x}})^2.\end{aligned}$$

The second term in the last expression is equal to zero, because by virtue of the properties of the arithmetic mean, $\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)=0$. For this reason,

$$S^2(\bar{\bar{x}}) = \frac{1}{N(N-1)}\left(\sum_{j=1}^{L}\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_j)^2 + \sum_{j=1}^{L}\sum_{i=1}^{n_j}(\bar{x}_j-\bar{\bar{x}})^2\right).$$

Note that

$$\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)^2=(n_j-1)S_j^2,\quad \sum_{i=1}(\bar{x}_j-\bar{\bar{x}})^2=n_j(\bar{x}_j-\bar{\bar{x}})^2.$$

Then, retaining the summation over groups, we obtain

$$S^2(\bar{\bar{x}})=\frac{1}{N(N-1)}\left[\sum_{j=1}^{L}(n_j-1)S_j^2+\sum_{j=1}^{L}n_j(\bar{x}_j-\bar{\bar{x}})^2\right].$$

The first term in the formula obtained characterizes the spread in the measurements in groups, and the second term characterizes the spread of the arithmetic means of the groups.

9.3. Effect of the Error of the Weights on the Error of the Weighted Mean

Looking at the general form of the formula determining the weighted mean, one would think, because the weights g_j and the weighted values of $\bar{x}_j$ appear in it symmetrically, that the weights must be found with the same accuracy as $\bar{x}_j$. In practice, however, the weights are usually expressed by numbers with one or two significant figures. How is the uncertainty of the weights reflected in the error of the weighted mean?

We shall regard weight g_i in (9.1) to be fixed, constant values. In addition, as usual, we shall assume that $\sum_{j=1}^{L}g_j=1$. This equality is also satisfied for the inaccurately determined weighting coefficients, i.e., for $\tilde{g}_j$. Therefore,

$$\sum_{j=1}^{L}\Delta g_j=0,$$

where Δg_j is the error in determining the coefficient g_j.

Assuming that the exact value of the weighted mean is y, we estimate the error of its estimate:

$$\Delta y=\sum_{j=1}^{L}\tilde{g}_j\bar{x}_j-\sum_{j=1}^{L}g_j\bar{x}_j=\sum_{j=1}^{L}\Delta g_j\bar{x}_j.$$

We shall express Δg_1 for the other errors:

$$\Delta g_1=-(\Delta g_2+\cdots+\Delta g_L)$$

and substitute it into the expression for Δy:

$$\Delta y=(\bar{x}_2-\bar{x}_1)\Delta g_2+(\bar{x}_3-\bar{x}_1)\Delta g_3+\cdots+(\bar{x}_L-\bar{x}_1)\Delta g_L$$

or in the form of relative error

$$\frac{\Delta y}{y} = \frac{g_2(\bar{x}_2 - \bar{x}_1)\dfrac{\Delta g_2}{g_2} + \cdots + g_L(\bar{x}_L - \bar{x}_1)\dfrac{\Delta g_L}{g_L}}{\sum\limits_{j=1}^{L} g_j \bar{x}_j}.$$

The errors of the weights $\Delta g_j/g_j$ are unknown. But let us assume that we can estimate their limits and let $\Delta g/g$ be the largest absolute value of these limits.

Replacing all relative errors $\Delta g_j/g_j$ by $\Delta g/g$, we obtain

$$\frac{\Delta y}{y} \leq \frac{\Delta g}{g} \frac{|g_2(\bar{x}_2 - \bar{x}_1) + g_3(\bar{x}_3 - \bar{x}_1) + \cdots + g_L(\bar{x}_L - \bar{x}_1)|}{\sum\limits_{j=1}^{L} g_j \bar{x}_j}.$$

The numerator on the right-hand side of the inequality can be put into the following form:

$$\begin{aligned} &g_2(\bar{x}_2 - \bar{x}_1) + g_3(\bar{x}_3 - \bar{x}_1) + \cdots + g_L(\bar{x}_L - \bar{x}_1) \\ &\quad = g_2\bar{x}_2 + g_3\bar{x}_3 + \cdots + g_L\bar{x}_L - (g_2 + g_3 + \cdots + g_L)\bar{x}_1. \end{aligned}$$

But $g_2 + g_3 + \cdots + g_2 = 1 - g_1$, so that

$$g_2(\bar{x}_2 - \bar{x}_1) + g_3(\bar{x}_3 - \bar{x}_1) + \cdots + g_L(\bar{x}_L - \bar{x}_1) = \sum_{j=1}^{L} g_j \bar{x}_j - \bar{x}_1 = y - \bar{x}_1.$$

Thus,

$$\frac{\Delta y}{y} \leq \frac{\Delta g}{g} \frac{|y - \bar{x}_1|}{y}.$$

It is obvious that if the entire derivation is repeated, but in so doing the error not in the coefficient g_1 but in some other weight is eliminated, then a weighted value other than $\bar{x}_1$ will appear on the right-hand side of the inequality. Therefore, the result obtained can be represented in the form

$$\frac{\Delta \bar{\bar{x}}}{\bar{\bar{x}}} \leq \frac{\Delta g}{g} \frac{|\bar{\bar{x}} - \bar{x}_j|}{\bar{\bar{x}}}.$$

The obtained inequality shows that the error introduced into the weighted mean as a result of the error of the weights is many times smaller than the latter error. The cofactor $|\bar{\bar{x}} - \bar{x}_j|/\bar{\bar{x}}$ can be assumed to be of the same order of magnitude as the relative error of the terms. Thus, if this error is of the order of 0.01, then the error introduced into the weighted mean as a result of the error of the weights will be at least 100 times smaller than the latter.

9.4. Combining the Results of Measurements in Which the Random Errors Predominate

We shall study a variant that is possible in the case of multiple measurements with negligibly small systematic errors. Each result being combined in this case is usually the arithmetic mean of measurements, and the differences between them are explained by the random spread of the measurements in the groups. However, it must be verified that the true value of the measured quantity is the same for all groups. This problem is solved by the methods presented in Chapter 4. If it cannot be assumed that the same quantity is measured in all cases, then it is pointless to combine the measurements into groups.

If the unification of the groups is justified, then it is necessary to check the hypothesis that the variances of the measurements in the groups are equal. Methods for solving this problem are also presented in Chapter 4.

In the case in which the variances of the groups can be assumed to be equal, the weights for each result are calculated from formula (9.7), the combined average is calculated from formula (9.1), and the variance of the combined average can be determined from formula (9.8).

When the variances of the groups cannot be taken to be equal to one another and the variances and their ratios are unknown, the weights are sometimes found by substituting their estimates, instead of the variances, into formula (9.4). The variance of the weighted mean obtained is estimated by substituting into formula (9.5) estimates of the variances of the combined quantities; i.e., from the formula,

$$S^2(\bar{\bar{x}}) = \frac{1}{\sum\limits_{k=1}^{L} \frac{1}{S^2(\bar{x}_k)}}.$$

The estimates of the variances are random quantities, and the weights obtained based on them are random quantities. However, in the case in which the observations are normally distributed, the weighted mean remains an unbiased estimate of the measured quantity. The error in estimating the standard deviation, obtained based on the formula presented for estimating the variance of the weighted mean, does not exceed 10% already with two groups of observations consisting of more than nine observations.

EXAMPLE. The mass of some body is being measured. In one experiment, the value $\tilde{m}_1 = 409.52$ g is obtained as the arithmetic mean of 15 measurements. The variance of the group of measurements is estimated to be $S_1^2 = 0.1\ \text{g}^2$. In a different experiment, the value $\tilde{m}_2 = 409.44$ g was obtained with $n_2 = 10$ and $S_2^2 = 0.03\ \text{g}^2$. It is known that the systematic errors of the measurements are negligible small, and the measurements can be assumed to be normally distributed. It is necessary to estimate the mass of the body using data from both experiments and to estimate the variance of the result.

We shall first determine whether the unification is justified, i.e., whether an inadmissible difference exists between the estimates of the measured quantity:

$$S^2(\bar{x}_1) = \frac{S_1^2}{n_1} = \frac{0.1}{15} = 0.0067, \quad S^2(\bar{x}_2) = \frac{0.03}{10} = 0.003,$$
$$S^2(\bar{x}_1 - \bar{x}_2) = S^2(\bar{x}_1) + S^2(\bar{x}_2) = 0.0097,$$
$$S(\bar{x}_1 - \bar{x}_2) = 0.098,$$
$$\bar{x}_1 - \bar{x}_2 = \tilde{m}_1 - \tilde{m}_2 = 0.08.$$

As $\bar{x}_1 - \bar{x}_2 < S(\bar{x}_1 - \bar{x}_2)$, the unification is possible.

We shall check whether both groups of observations have the same variance (see Section 4.6):

$$F = S_1^2/S_2^2 = 0.1 : 0.03 = 3.3.$$

The degrees of freedom are $\nu_1 = 14$ and $\nu_2 = 9$. We shall assume that the significance level is 2%. In addition, $q = 0.01$ and $F_q = 5$ (see Table A.5). As $F < F_q$, it can be assumed that the variances of the groups are equal.

We shall now find the weights of the arithmetic means. According to (9.7), we have $g_1 = 15/25 = 0.6$ and $g_2 = 10/25 = 0.4$. The weighted mean is $\bar{\bar{m}} = 0.6 \times 409.52 + 0.4 \times 409.44 = 409.49$ g. Now we find $S(\bar{\bar{m}})$. In accordance with formula (9.8), we have

$$S^2(\bar{\bar{m}}) = \frac{1}{25 \times 24}(14 \times 0.1 + 9 \times 0.03 + 15 \times 0.03^2 + 10 \times 0.05^2)$$
$$= 28 \times 10^{-4}\,\mathrm{g}^2,$$
$$S(\bar{\bar{m}}) = 5.3 \times 10^{-2}\,\mathrm{g}.$$

If in addition to estimating the standard deviation it is also necessary to find the uncertainty, then to use Student's distribution, the effective number of degrees of freedom must be found using formula (6.19).

9.5. Combining the Results of Measurements Containing both Systematic and Random Errors

Let us assume that a quantity A is measured by several methods. Each method gives the result x_j ($j = 1, \ldots, L$) with error ζ_j:

$$x_j = A + \zeta_j.$$

To combine in a well-founded manner the series of values of x and obtain a more accurate estimate of the measured quantity, one must have certain information about the errors ζ_j. We shall start from the condition that none of the measurements have absolutely constant systematic errors. However, this assumption must be checked. If it is not true, then the problem posed cannot be solved.

The error ζ_j is the sum of realizations of the conditionally constant ϑ_j and random ψ_j errors: $\zeta_j = \vartheta_j + \psi_j$.

Having in mind a possible set of results with each method of measurement, the unknown error ϑ_j of a concrete realization of the jth method of measurement can be regarded as the realization of a random quantity. Usually the limits θ_j of ϑ_j are estimated, and they are regarded as symmetric relative to the true value of the measured quantity: $|\vartheta_j| \leq \theta_j$ and $E[\vartheta_j] = 0$.

The random error ψ_j is assumed to be a centered quantity; i.e., $E[\psi_j] = 0$.

Thus, when there are no absolutely constant errors, for example, methodological errors, we can write $E[x_j] = A$.

As follows from the theory of combining of the results of measurements, the weights are determined by the variances of these results. In our case, we can write

$$V[x_j] = V[\vartheta_j] + V[\psi_j].$$

Therefore, given the variances $V[\vartheta_j]$ and $V[\psi_j]$, the problem can be solved exactly and uniquely. Unfortunately, the variances are always unknown, and their estimates must be employed. To estimate the variances of conditionally constant errors of each method of measurement, we shall use the assumption that the errors are uniformly distributed within the estimated limits. Although the weights need not be found with high accuracy, it still casts some doubt on whether the weighted mean is a better estimate of the measured quantity than the combined results. As a result, in metrology, great care is taken in combining results of measurements.

Based on what was said above, when the results of measurements must be combined, it is always necessary to check the agreement between the starting data and the obtained results. If some contradiction is discovered, for example, the combined average falls outside the permissible limits of some term, then the reason for this must be determined and the contradiction must be eliminated. Sometimes this is difficult to do, and special experiments must be performed.

Great care must be exercised in combining the results of measurements because in this case information about the errors is employed to refine the result of the measurement and not to characterize its uncertainty, as is usually done.

It can happen, however, that the weighted mean is a natural estimate of the measured quantity. An example is the accurate measurement of the activity of a source of α particles. To increase the accuracy, the activity is measured at different distances from the source to the detector and with different diaphragms. The measured activity remains the same. However, the estimates of the activity obtained with different diaphragms are found to differ somewhat from one another. Their errors are also different. For this reason, when assigning weights for the obtained estimates of the measured quantity, in this example, one must start from estimates of the variances of the total error of the measurement results being combined. This process, undoubtedly, will lead to a more correct result than if all measurements were assumed to have equal weights or weights were assigned taking into account only the random errors.

So, in the case at hand, the weights of the measurements being combined should be calculated using the formula

$$g_j = \frac{\dfrac{1}{S^2(\vartheta_j) + S^2(\psi_j)}}{\displaystyle\sum_{j=1}^{L} \frac{1}{S^2(\vartheta_j) + S^2(\psi_j)}}, \tag{9.9}$$

where $S^2(\vartheta_j)$ is an estimate of the variance of the possible set of conditionally constant errors of the jth measurement result and $S^2(\psi_j)$ is an estimate of the variance of the random error of the same measurement result.

We shall now estimate the errors of the combined average. In solving this problem, because the errors of the weights are insignificant (see Section 9.3), we shall assume that the weights of the combined measurement results are known exactly.

In the case of multiple measurements, one must have for each result x_j an estimate of the limits of the systematic error θ_j and an estimate of the standard deviation S_j of the random error. Then the corresponding indicators of accuracy of the combined average will be

$$\theta_\alpha(\bar{\bar{x}}) = k\sqrt{\sum_{j=1}^{L} g_j^2\theta_j^2},$$

$$S(\bar{\bar{x}}) = \sqrt{\sum_{j=1}^{L} g_j^2 S_j^2}.$$

The confidence limits of the total error of the combined average can be found based on the estimates obtained for $\theta(\bar{\bar{x}})$ and $S(\bar{\bar{x}})$. The method for solving this problem was examined in detail in Chapter 5.

In the case of single measurements, one usually knows only the estimates of the limits of the errors of the measurements being combined, i.e., $\Delta_j (j = 1, \ldots, L)$. Based on available information about the form of the distribution of the possible sets of actual errors of each measurement result, it is necessary to transfer from the limits Δ_j of the errors to estimates of the variances of these errors. Once the variances have been obtained, it is not difficult to find the weights of the measurement results being combined. Next, the uncertainty of the weighted mean can be calculated using the scheme developed for linear indirect measurements (see Chapter 6).

We shall discuss some particular cases of single measurements. We shall examine the measurement of one quantity with several instruments.

Let the random errors of the instruments be small compared with the limit of permissible errors, and let the permissible errors be the same and equal to Δ for all instruments.

Let the indications of the instruments be $x_1, \ldots, x_n$ and the actual errors in the indications be $\zeta_1, \ldots, \zeta_q$ ($|\zeta_i| < \Delta$). Then we can write

$$A = x_1 - \zeta_1, \ldots, A = x_n - \zeta_n. \tag{9.10}$$

The natural intuitive estimate of the true value of the measured quantity, in the case in which several instruments of equal accuracy are used to perform the measurements, is the arithmetic mean of the instrumental indications:

$$\tilde{A} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

It has been proved mathematically that in the class of linear estimates, this estimate is the best.

We must estimate the error in the result obtained. Adding the left and right sides of (9.10) and dividing them by n,we obtain

$$A = \frac{1}{n} \sum_{i=1}^{n} x_i - \frac{1}{n} \sum_{i=1}^{n} \zeta_i.$$

We do not know the real errors of the instruments. We know only that $|\zeta_i| \leq \Delta$ for all $i = 1, \ldots, n$.

To find the limits of the sum of the random quantities ζ_i (and their errors over a set of instruments of a given type can be assumed to be random quantities), it is necessary to known their distribution functions. As pointed out above, these functions cannot be found from the experimental data. However, it can often be assumed that the errors of complicated instruments have symmetric distributions. The mathematical expectation of the distribution is close to the errors of working standards employed to calibrate these instruments. To a first approximation, we shall assume that $E[\zeta_i] = 0$.

For example, if the errors of the instruments are distributed uniformly, then according to formula (5.3)

$$\theta_{\alpha 1} = k \sqrt{\sum_{i=1}^{n} \Delta^2} = k\Delta\sqrt{n}.$$

From here, the uncertainty of the estimate $\tilde{A}$ will be

$$\Delta_{\alpha 1} = \frac{\theta_{\alpha 1}}{n} = \frac{k\Delta}{\sqrt{n}}.$$

If the errors of the instruments are assumed to have a normal distribution and $\Delta = z_\alpha \sigma$, then

$$\theta_{\alpha 2} = z_\alpha \sqrt{\sum_{i=1}^{n} \sigma_i^2} = z_\alpha \sigma \sqrt{n} = \Delta\sqrt{n}.$$

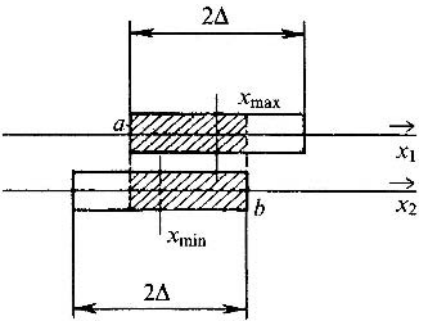


FIGURE 9.1. The highest and lowest indications of the group of instruments and the intervals of their possible errors.

Then

$$\Delta_{\alpha 2} = \frac{\theta_{\alpha 2}}{n} = \frac{\Delta}{\sqrt{n}}.$$

This estimate corresponds to the same probability α, which was used to establish the limit of permissible error Δ. Comparing $\Delta_{\alpha 1}$ and $\Delta_{\alpha 2}$ shows that they differ only by the factor k, which, depending on the confidence probability, can range from 1.1 ($\alpha = 0.95$) to 1.4 ($\alpha = 0.99$). As expected, the number of instruments plays the main role. Five to ten instruments are required to reduce the error by a factor of 2 or 3. But we have to stress here that a real improvement of the errors is limited by the errors of working standards employed to calibrate these instruments.

The problem can also be solved as follows. We choose the maximum and minimum indications of the instruments: x_{max} and x_{min}. We verify that

$$(x_{max} - x_{min}) \leq 2\Delta. \tag{9.11}$$

If inequality (9.11) is not satisfied, then one of the instruments used to perform the measurement has an inadmissably large error or the variation of some influence quantities is too large. The reason for this phenomenon must be determined and eliminated; i.e., inequality (9.11) must be satisfied.

Figure 9.1 illustrates the indications x_{max} and x_{min}, and the intervals corresponding to the limits of permissible errors $\pm\Delta$ are marked off.

The true value of the measured quantity must lie in the section of the tolerance field that belongs simultaneously to the instrument with indication x_{max} and the instrument with indication x_{min}. In the figure, this section is hatched. Its boundaries determine more accurately the tolerance field.

It is natural to take for the estimate of the measured quantity the center of the interval $x_{max} - x_{min}$, which is found from

$$\tilde{A} = \frac{x_{max} + x_{min}}{2}. \tag{9.12}$$

The coordinate of the point a, determining the left-hand boundary of the error in the result, will be $x_a = x_{\max} - \Delta$. The coordinate of the point b, which determines the right-hand limit of error, is equal to $x_b = x_{\min} + \Delta$.

Therefore, the limit of error Δ_1 of the more accurate result is

$$\Delta_1 = |x_b - \tilde{A}| = |\tilde{A} - x_a|$$

or

$$\Delta_1 = \left| x_{\min} + \Delta - \frac{x_{\max} + x_{\min}}{2} \right| = \left| \Delta - \frac{x_{\max} - x_{\min}}{2} \right|.$$

It is easy to see that if the limit $x_{\max} - x_{\min} = 2\Delta$, the error Δ_1 is formally equal to zero. It is clear, however, that the minimum value of Δ_1 cannot be less than the error of the working standard used to calibrate the instruments employed.

It is interesting to note that the estimate, based on (9.12), of the measured quantity mathematically gives the best approximation when the errors ϑ_i are distributed uniformly over the interval $[-\Delta, +\Delta]$.

Based on the foregoing arguments, it can be shown that a well-known assumption of metrology is valid: When measuring instruments having different accuracy are used in parallel, the accuracy of the result is determined by the most accurate measuring instrument.

For example, assume that the voltage of some source was measured simultaneously with three voltmeters having different accuracy but the same upper limit of the measurement range 15 V. The measurements were performed under reference conditions. The following results were obtained.

(1) Class 0.5 voltmeter: $U_1 = 10.05$ V; the limit of permissible intrinsic error $\Delta_1 = 0.075$ V.
(2) Class 1.0 voltmeter: $U_2 = 9.9$ V; the limit of permissible intrinsic error $\Delta_2 = 0.15$ V.
(3) Class 2.5 voltmeter: $U_3 = 9.7$ V, the limit of permissible intrinsic error $\Delta_3 = 0.375$ V.

As the measurements were performed under reference conditions, we shall assume that the limits of permissible intrinsic error of the instruments are equal to the limits of the errors of measurement.

Assume that the errors of the instruments of each type have a uniform distribution. Then

$$\sigma_i = \Delta_i / \sqrt{3}.$$

We shall find the weights of the results. As the upper limit of the measurement range is the same for all instruments, the calculation can be performed based on the limits of fiducial error of the instruments:

$$g_1' = \frac{1}{\Delta_1^2} = \frac{1}{0.25} = 4, \quad g_2' = \frac{1}{\Delta_2^2} = 1, \quad g_3' = \frac{1}{\Delta_3^2} = \frac{1}{6.25} = 0.16.$$

From here,

$$g_1 = \frac{g_1'}{\sum_{i=1}^{3} g_i'} = \frac{4}{5.16} = 0.77,$$

$$g_2 = \frac{g_2'}{\sum_{i=1}^{3} g_i'} = \frac{0.20}{5.16} = 0.20, \quad g_3 = \frac{g_3'}{\sum_{i=1}^{3} g_i'} = \frac{0.16}{5.16} = 0.03.$$

Now we find the weighted mean

$$\tilde{U} = \sum_{i=1}^{3} g_i U_i = 0.77 \times 10.05 + 0.2 \times 9.9 + 0.03 \times 9.7 = 10.01 \text{ V}.$$

The confidence limits of the error in the weighted mean can be found from formula (6.43)

$$\begin{aligned} \Delta\tilde{U} &= k\sqrt{\sum_{i=1}^{3} g_i^2 \Delta_i^2} \\ &= k\sqrt{0.77^2(7.5 \times 10^{-2})^2 + 0.2^2(15 \times 10^{-2})^2 + 0.03^2 \times 0.375^2} \\ &= k\sqrt{(33 + 9 + 1.3) \times 10^{-4}} = 0.066k. \end{aligned}$$

Assuming, as usual, $\alpha = 0.95$, we take $k = 1.1$ and find $\Delta\tilde{U} = 0.07$ V.

In Fig. 9.2, the indications of all three instruments are plotted and the limits of permissible error of the instruments are marked. The value obtained for the weighted mean is also indicated there. This value remained in the tolerance field of the most accurate result, but it was shifted somewhat in the direction of indications of the less accurate instruments; this is natural. The limits of error of the result decreased insignificantly compared with the error of the most accurate term.

If all distributions were assumed to be normal distributions, truncated at the same level by discarding instruments whose error exceeds $z_\alpha \sigma$ ($z_\alpha = \text{const}$), then the weights would not change and the weighted mean would have the same value as we found above. Only the estimate of the error limits would change, because it must now be calculated from formula (6.43) with $k = 1$.

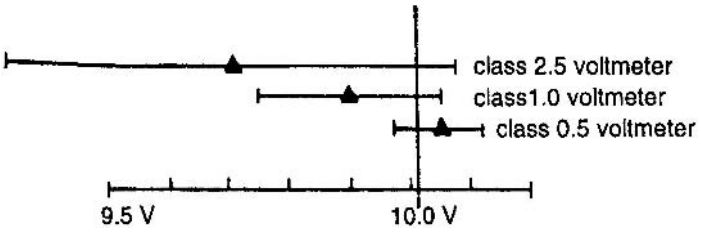


FIGURE 9.2. Indications of the instruments and the intervals of their possible errors.

We obtain $(\Delta\tilde{U})' = 0.066$ V. However, the difference between $\Delta\tilde{U}$ and $(\Delta\tilde{U})'$ is insignificant.

Let us see what would happen if different instruments had different distributions. For example, assume that the class 0.5 and 1.0 instruments have a uniform distribution, that the class 2.5 instruments have a truncated normal distribution, and that $z_\alpha = 2.6$. Then

$$\sigma_1 = \frac{0.075}{\sqrt{3}} = 0.043, \qquad g_1' = \frac{1}{\sigma_1^2} = 306,$$
$$\sigma_2 = \frac{0.15}{\sqrt{3}} = 0.087, \qquad g_2' = \frac{1}{\sigma_2^2} = 78,$$
$$\sigma_3 = \frac{0.375}{2.6} = 0.144, \qquad g_3' = \frac{1}{\sigma_3^2} = 42.$$

Therefore,

$$g_1 = \frac{g_1'}{\sum_{i=1}^{3} g_i'} = \frac{306}{426} = 0.72,$$
$$g_2 = \frac{g_2'}{\sum_{i=1}^{3} g_i'} = \frac{78}{426} = 0.18,$$
$$g_3 = \frac{g_3'}{\sum_{i=1}^{3} g_i'} = \frac{42}{426} = 0.10.$$

From here,

$$\tilde{U}_1 = 0.72 \times 10.05 + 0.18 \times 9.9 + 0.1 \times 9.7 = 9.99\,\text{V}.$$

The values obtained for $\tilde{U}$ and $\tilde{U}_1$ are very close, which indicates that a significant change in the form of the distribution functions in this case does not appreciably affect the result.

The foregoing example could also have been solved by a graphical-analytic method, similar to the method used to solve the problem of combining the indications of equally accurate instruments. Now, however, the relation determining $\tilde{A}$ must contain the weights of the terms. In accordance with the foregoing considerations, these weights can be taken to be inversely proportional to the squares of the limits of permissible errors of the instruments.

9.6. Example: Measurement of the Activity of Nuclides in a Source

We shall examine the measurement of the activity of nuclides by the method of absolute counting of α particles emitted by the source in a small solid angle. The

TABLE 9.1. The results of measurements of the activity of nuclides by different geometric factors.

Number of group i	Source-detector distance (mm)	Diaphragm radius (mm)	Estimate of measured quantity $x_i \times 10^5$	Estimates of the standard deviations	
				Random errors of the result (%)	Systematic errors of the result (%)
1	97.500	20.017	1.65197	0.08	0.52
2	97.500	12.502	1.65316	0.10	0.52
3	397.464	30.008	1.66785	0.16	0.22
4	198.000	20.017	1.66562	0.30	0.42
5	198.000	30.008	1.66014	0.08	0.42

measured activity is determined from the formula

$$A = GN_0\eta,$$

where G is the geometric factor of the apparatus, N_0 is the α-particle counting rate, and η is the α-particle detection efficiency [15].

The geometric factor depends on the diameter of the source, the distance between the source and the detector, as well as the diameter of the diaphragm, and it is calculated from measurements of these quantities. In the course of a measurement, G does not change, so that errors of G create a systematic error of measurement of the activity A. Measurements of the numbers of α particles, however, have random errors.

To reduce the error arising from the error of the geometric factor, the measurements were performed for different values of this factor (by changing the distance between the source and detector and the diameter of the diaphragm). All measurements were performed using the same source ^{239}Pu. Table 9.1 gives the five combinations of the geometric parameters studied. In each case, 50 measurements were performed, and estimates of the measured quantity and the parameters of their errors, which are also presented in Table 9.1, were calculated. The rms deviations of the systematic errors of the results were calculated from the estimated limiting values of all components under the assumption that they can be regarded as centered uniformly distributed random quantities.

The data in Table 9.1 show, first of all, that the systematic errors are much larger than the random errors, so that the number of measurements in the groups was sufficient. The observed difference between the obtained values of the activity of the nuclides in the groups can be explained by their different systematic errors.

In the example studied, the same quantity was measured in all cases. For this reason, here the weighted mean is a well-founded estimate of the measured quantity. Based on the considerations presented in Section 9.5, we shall use formula (9.9) to calculate the weights. First, we shall calculate an estimate of the combined variance

$$S_c^2(\bar{x}_i) = S^2(\psi_i) + S^2(\vartheta_i).$$

The results of the calculations are given in Table 9.2.

TABLE 9.2. The estimates of combined variances and weights of measurement results by different geometric factors.

Number of group	Estimate of combined variance $S_c^2(\bar{x}_i)$	Weight g_i
1	0.28	0.12
2	0.28	0.12
3	0.07	0.46
4	0.27	0.12
5	0.18	0.18

As an example, we shall calculate g_1:

$$g_1 = \frac{\dfrac{1}{0.28}}{\dfrac{1}{0.28} + \dfrac{1}{0.28} + \dfrac{1}{0.07} + \dfrac{1}{0.27} + \dfrac{1}{0.18}} = \frac{3.57}{30.7} = 0.12.$$

Now we find the weighted mean: $\tilde{A} = \sum_{i=1}^{5} g_i \bar{x}_i = 1.6625 \times 10^5$.

It remains to estimate the error in the value obtained. We shall calculate an estimate of the variance of the weighted mean with the help of the formula

$$S^2(\tilde{A}) = \frac{1}{\sum\limits_{i=1}^{5} [1/S_c^2(\bar{x}_i)]} = \frac{1}{30.7} = 0.033.$$

From here $S_{\tilde{A}} = 0.18\%$. As the error of the weighted mean is determined by the systematic component, it is best presented in the form of limits (in this case, confidence limits). For the estimated value of the variance, the limits are calculated for the normal distribution. In this case, this process is all the more justified because, as we have seen, the error of the weighted mean consists primarily of five terms. Even if all terms were uniformly distributed, the distribution of their composition can be regarded as a normal distribution.

For the standard confidence probability $\alpha = 0.95$, $z_{\frac{1-\alpha}{2}} = 1.96$ and $u_{\%0.95} = 1.96 \times 0.18 = 0.35\%$. In the form of absolute uncertainty, we obtain $u_{c0.95} = 0.006 \times 10^5$.

The result of the measurement can be given in the form

$$A(0.95) = (1.662 \pm 0.006) \times 10^5.$$

One can see that in this example the simple arithmetic mean, equal to 1.660×10^5, of the estimates obtained for the measured quantity does not differ significantly from the weighted mean. This agreement, however, is purely accidental. In cases similar to the one examined above, the weighted mean, of course, is a better founded estimate of the measurand than the simple arithmetic mean.

10
Calculation of the Errors of Measuring Instruments

10.1. The Problems of Calculating Measuring Instrument Errors

Measuring instruments are extremely diverse, but because they are used for a common purpose, a general theory of their errors exists. The central problem of this theory is to calculate the intrinsic error of measuring instruments, which is their most important metrological characteristic. The calculation of the additional errors, caused by controlled changes in the influence quantities, depends on the arrangement of a measuring instrument. For this reason, the calculation of these metrological characteristics falls within the purview of the theory of measuring instruments with a particular principle of operation.

In general form, a measuring instrument can be considered as several functionally related units that transform an input signal into an output signal. During the manufacturing process, a desirable shape of functional dependence between these signals (a transfer function) is first obtained by adjusting some units. Then, each instrument is graduated or calibrated. In essence, the purpose of these operations is to fix and to represent the obtained transfer function of the instrument by means of a scale or a graph or an equation.

No matter how accurate these operations are, the resulting instrument will have some errors for the following reasons:

(a) inaccuracy in fixing the transfer function, that is, inaccuracy in constructing the scale or the graph or the equation;
(b) imperfection of the reading device of the measuring instrument;
(c) variations of influence quantities (within limits of reference conditions);
(d) drifting of some properties of the measuring instrument units with time.

Each problem contributes a component to the intrinsic error of a measuring instrument.

The errors of measuring instruments under normal operating conditions (i.e., when the influence quantities deviate from their reference values or when they exceed the limits of the range of reference values), and their calculation based on

known properties of the measuring instruments and the conditions of measurement, are regarded as an integral part of the problem of estimating measurement errors.

The problem of estimating the resulting error can be formulated for a separate instrument or a collection of instruments of a definite type. In the first case, the problem consists of estimating the error of a particular instrument from the known parameters of the components of this instrument. In so doing, one can find either the errors of the instrument on definite segments of the instrument scale or an estimate of the limits of error of the given instrument. These problems must be solved when designing unique measuring instruments and when performing an elementwise calibration.

In the second case, i.e., for a collection of measuring instruments, the problem is formulated differently. The limits of error of instruments can be estimated based on the properties of the components of the instrument (direct problem). But most often the limits of instrument error are prescribed and it is required to find the percentage of instruments whose error will fall within these limits (inverse problem).

Each problem admittedly, can be formulated, with some modifications, for any type of measuring instrument—for standards, measuring transducers, or measuring instruments and systems—only if its errors are caused by deviations from the nominal values of the parameters of the components of the measuring instrument.

10.2. Methods for Calculating Instrument Errors

We shall examine both methods successively, i.e., the direct and inverse problems, referring to collections of measuring instruments.

In the general form, the output signal y of an instrument is related to the informative parameter of the input signal A,the parameters x_i of the components of the instrument ($i = 1, \ldots, n$), noise, and other factors giving rise to errors $z_j (j = 1, \ldots, m)$, by the relation

$$y = f(A, x_i, z_j). \tag{10.1}$$

For each parameter, shall establish the nominal value, i.e., the value for which the measuring instruments would not have an error. The deviations of the real properties of the components from the nominal properties result in the instrument error. We shall call conventionally the deviation from the nominal values of the parameters of the components the errors of the components, and we shall assume that they are expressed in the form of relative errors

$$\varepsilon_i = \frac{x_{ir} - x_i}{x_i}, \tag{10.2}$$

where x_i is the nominal value and x_{ir} is the real value of the parameter of the component.

The effect of the error of each component on the instrument error is determined in the manner studied in Chapter 6 for determining the influence coefficients of the measurement error of the arguments on the error of an indirect measurement.

For relative errors, we can write

$$\frac{\Delta y}{y} = \frac{1}{y}\frac{\partial f}{\partial x_i}\Delta x_i = \frac{1}{y}\frac{\partial f}{\partial x_i}\varepsilon_i x_i.$$

From here, the influence coefficient of the error x_i will be, in relative form,

$$W_i = \frac{\Delta y/y}{\varepsilon_i} = \frac{\partial f}{\partial x_i}\frac{x_i}{y}. \tag{10.3}$$

We shall express the influence coefficients for sources of additive errors, because they cannot be represented as a deviation from some nominal values, in the standard form

$$w_j = \frac{\partial f}{\partial_{z_j}}. \tag{10.4}$$

In what follows we shall refer, somewhat arbitrarily, to the factors responsible for the additive errors as noise.

The absolute error of an instrument at indication y is determined by the relation

$$\zeta = y\sum_{i=1}^{n} W_i\varepsilon_i + \sum_{j=1}^{m} w_j z_j. \tag{10.5}$$

We find (10.1) and the influence coefficients W_i and w_i based on the structural layout of the instrument. Having derived (10.5), we no longer need the structural layout of the instrument, and we need study only the components of the error. In (10.5), the error components are referred to the output of the instrument.

We shall now study in greater detail the direct problem, i.e., the problem of estimating instrument error. We have in mind estimation of errors at any point in the range of indication. If the point where the instrument error is maximum is known, then in many cases, only the error for this point need be calculated. Often this point is the end point y_f of the instrument scale, because in this case, the multiplicative components of the error are maximum.

The calculations are most conveniently performed for the relative and not for the absolute errors. For $y = y_f$, the relative instrument error, as follows from formula (10.5), will be

$$\varepsilon_f = \sum_{i=1}^{n} W_i\varepsilon_i + \frac{1}{y_f}\sum_{j=1}^{m} w_j z_j. \tag{10.6}$$

We shall divide all errors appearing on the right-hand side of formulas (10.5) or (10.6) into systematic and random. If some term in formula (10.5) has both systematic and random components, then we shall separate them and replace such a term in formula (10.5) with two terms. In so doing, the systematic components of the errors are assumed, as always, to have a fixed value for every sample of the instrument or for each component of the instrument.

The systematic components form the systematic instrument error, and the random components form the random instrument error.

The random instrument error is manifested differently in each application of the instrument. For example, if we want to estimate the largest error of the instrument, then we must add to the estimate of the limits of its systematic error the maximum random error.

It should be noted that in the theoretical description of the random instrument error, this error is regarded as a random quantity, and it is most often assumed to have a normal distribution. Such a model admits the possibility of errors of any size, and it becomes unclear how to find the limiting random instrument error in the model.

I would like to stress that different situations arise when measurement errors and the errors of measuring instruments are estimated. In the first case, the random errors have already been realized, and for this reason, the random and systematic components can be summed statistically. In the second case, we are estimating the largest error of the instrument that can be manifested in any future experiment, and for this reason, the components must be added arithmetically.

The systematic errors of a collection of instrument components of the same type can be regarded as a set of realizations of a random quantity. This quantity is described statistically, for example, by a histogram. For components having a systematic error, it is not difficult to construct the histogram of the distribution of systematic errors. For components with a random error, such a description becomes more complicated, because it becomes two dimensional: The realization depends on both the sample of the component and the realization of the random error of this component in a given experiment. But all components of the same type can usually be assumed to have the same distribution of the random error, so that the differences of their errors are determined only by the change in some parameter of this distribution. This characteristic can be taken into account, and one-dimensional distributions of the corresponding parameter and systematic error of the component can be studied instead of the two-dimensional distribution of the random quantity.

The summation of random quantities involves the construction of the composition of their distribution functions.

If the instrument consists of many components and there are many terms in formulas (10.5) or (10.6), then the composition of the error components will give, as is well known, a close-to-normal distribution, which makes it possible to simplify the solution, because it is not difficult to find the parameters of the resulting distribution:

$$\begin{aligned} E[\varepsilon_f] &= \sum_{i=1}^{n} W_i E[\varepsilon_i] + \frac{1}{y_f} \sum_{j=1}^{m} w_j E[z_j], \\ V[\varepsilon_f] &= \sum_{i=1}^{n} W_i^2 V[\varepsilon_i] + \frac{1}{y_f^2} \sum_{j=1}^{m} w_j^2 V[z_j], \end{aligned} \tag{10.7}$$

Using the estimates of the mathematical expectations of the error components and their variances, the estimate of the mathematical expectation and variance of the resulting normal distribution can be calculated from the relations presented. This

problem is solved first for systematic errors. Taking the percentage of instruments whose error must be less than the computed estimate as the confidence probability α, we find the corresponding limit of systematic error:

$$\begin{aligned} \theta_r &= \tilde{E}[\varepsilon_f] + z_\alpha \tilde{\sigma}(\varepsilon_f), \\ \theta_l &= \tilde{E}[\varepsilon_f] - z_\alpha \tilde{\sigma}(\varepsilon_f). \end{aligned} \tag{10.8}$$

Next we estimate the practically limiting random error. Usually the number of terms here is very small, and this error is not calculated at all but it is estimated based on the experimental data. If we find an estimate of the standard deviation of the resulting random error, then for the practically limiting value, we take $\Psi_\alpha = t_q \tilde{\sigma}(y)$, where $q = 1 - \alpha$, and we find t_q from a table of Student's distribution taking into account the degree of freedom from the experimental data.

Next, in accordance with what was said above, we find the practically limiting errors of the instruments:

$$\Delta_r = \theta_r + \Psi_\alpha, \qquad \Delta_l = \theta_l + \Psi_\alpha. \tag{10.9}$$

Usually $E[\varepsilon_f] = 0$ and $|\theta_l| = |\theta_r|$, so that $|\Delta_l| = |\Delta_r|$.

When the number of terms is small, the problem must be solved by constructing a composition of the distributions of the terms.

It should be noted that in the general case, the probability adopted for calculating the limits of systematic error may not be equal to the probability corresponding to the practically limiting random error. Both probabilities should be indicated.

If the terms are given by their permissible limits and no data favor one distribution, then the corresponding errors are best assumed to be uniformly distributed. In this case, the confidence limit of the systematic instrument error can be found from formula (6.42), transformed somewhat:

$$\theta_\alpha = k \sqrt{\sum_{i=1}^{n} (W_i \delta_i)^2 + \frac{1}{y^2} \sum_{j=1}^{m} (w_j \Delta_j)^2}, \tag{10.10}$$

where the values of the coefficient k are presented in Table 3.1; $\delta_i (i = 1, \ldots, n)$ are the limits of permissible systematic errors (of the instrument components) forming the multiplicative component of the instrument error; and $\Delta_j (j = 1, \ldots, m)$ is the same, but for errors that make up the additive instrument error.

The practically limiting random error $\Psi_\alpha(y)$ is found in the same manner as in other cases, and the total instrument error is calculated in the same way. As systematic instrument errors are assumed to be random quantities, the confidence probability α used to calculate the limits of the systematic error indicates the relative number of instruments whose systematic errors do not exceed these limits.

Often the errors of the instrument components are given by their permissible limits, including both systematic and random error components. In this case, one can proceed in two ways. The error of each instrument component can be divided into separate components based on experiment, after which the problem is solved according to the scheme presented above. But the total errors of the

instrument components can be assumed, without separating the random error, to be uniformly distributed within prescribed limits and they can be added statistically. In the absence of data for separating errors into components, the second method is preferable.

The calculations of the errors are repeated for a series of indications of the instrument. In the process, the confidence probabilities, one of which was used to determine the limits of systematic error of a set of instruments and the other was used to determine the practically limiting random error of the instruments, should be kept constant. From the data obtained, it is possible to construct a graph of the limiting instrument error as a function of instrument indication. The error can be expressed as an absolute or relative error, but absolute error is usually more convenient.

We now consider the second problem: The limit of permissible instrument error is prescribed, and it is required to estimate the probability of encountering an instrument with an error less than this limit.

We shall outline the general scheme of the solution. The probability p_g of encountering an instrument whose error does not exceed the permissible limit is equal to

$$p_g = 1 - (p_r + p_l), \tag{10.11}$$

where p_r and p_l are the probabilities of encountering an instrument whose error exceeds the upper limit and drops below the lower limit, respectively.

We can write

$$p_r = P\{\zeta \geq (\Delta - \Psi_\alpha)\}, \qquad p_l = P\{\zeta \leq -(\Delta - \Psi_\alpha)\},$$

where Δ is the limit of permissible instrument error and Ψ_α is the practically limiting random error.

To solve the problem, it is necessary to know the distribution function of the systematic instrument errors over the entire collection of instruments.

But the practically limiting random error cannot always be assumed to be the same for all instruments; it is usually different for different instruments. To obtain a more accurate solution of the problem, it is necessary to find the distribution of the practically limiting random instrument error. An example of a possible density of such a distribution is presented in Fig 10.1. The probabilities p_l and p_r must

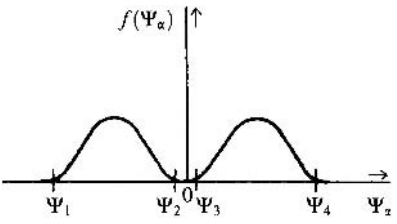


FIGURE 10.1. Possible form of the probability density of the practically limiting random instrument error.

now be calculated from the formulas

$$p_l = K_l \int_{\Psi_1}^{\Psi_2} P\{\zeta < -(\Delta - \Psi_\alpha)\} f(\Psi_\alpha)\, d\Psi,$$
$$p_r = K_r \int_{\Psi_3}^{\Psi_4} P\{\zeta > (\Delta - \Psi_\alpha)\} f(\Psi_\alpha)\, d\Psi, \tag{10.12}$$

in which the values of the probability density $f(\Psi_\alpha)$ play the role of weights and K_l and K_r are normalization factors. If the probability density is $f(\Psi_\alpha)$ and it is symmetric relative to the ordinate axis, then $p_l = p_r$ and $K_l = K_r = 2$. In the general case, however,

$$K_l = \frac{1}{\int_{\Psi_1}^{\Psi_2} f(\Psi_\alpha)\, d\Psi}, \qquad K_r = \frac{1}{\int_{\Psi_3}^{\Psi_4} f(\Psi_\alpha)\, d\Psi}.$$

The coefficients K_l and K_r were introduced in connection with the fact that by construction of the distribution functions, the area under the entire curve of the probability density is equal to unity, and we require that the area under each branch of the curve be equal to unity (for $\Psi_\alpha < 0$ and for $\Psi_\alpha > 0$).

It is not difficult to derive formulas (10.12) if several values of Ψ_α and the percentage of cases when each selected value occurs are given. Having found p_l for each Ψ_α, it is natural to add them, weighting each one by a weight proportional to the percentage of times it is encountered. From here, extrapolating to a continuous distribution of Ψ_α, we arrive at formulas (10.12)

Thus, we can find the probability for manufacturing a measuring instrument whose error is less than a fixed limit, if it has a unique output signal. Examples of such measuring instruments are single-valued measures.

Much more often, however, measuring instruments have a definite range of measurement. In this case, the probability of getting a good instrument is equal to the probability that the error of the selected instrument over the entire measurement range is less than the prescribed limit. How does one estimate this probability?

One would think that for a pointer-type instrument the probability p_{gi} can be calculated for each marker of the instrument scale and p_g can be found by multiplying the probabilities together. However, one cannot proceed in this manner, because the errors at different points of the scale are not independent of one another, primarily because of the multiplicative component. In addition, such calculations would be too laborious, because the instrument scale often has 100 to 150 markers.

We shall examine a different method for solving the problem. For definiteness, we shall consider a pointer-type instrument with one measurement range. The method described above makes it possible to find all components of the instrument error at any marker of the instrument scale. We shall calculate them for the final value of the scale y_f. The random component usually varies insignificantly along the instrument scale. For this reason, once the practically limiting random error

for the final value of the scale has been estimated, we can assume that we have it for the entire scale of the instrument.

We shall assume at first that the practically limiting random error, equal to Ψ_α, is the same for all instruments studied.

Before summing the systematic error components, we separate them into additive and multiplicative components, after which we add them separately. The addition is performed by statistical methods. As a result, we obtain the probability density of the additive and multiplicative error components for $y = y_f$, i.e., for the final value of the scale.

However, the instrument can have an inadmissibly large error at any marker of the scale, and this must be taken into account.

The systematic error ϑ_y of an arbitrarily chosen instrument at the point y of the instrument scale consists of the multiplicative component ϑ_{my} and the additive component ϑ_{ay}:

$$\vartheta_y = \vartheta_{my} + \vartheta_{ay}.$$

In addition,

$$|\vartheta_{my}| \leq \theta_{my},$$

where $\theta_{my} = \theta_{mf}\, y/y_f$ and θ_{mf} is the largest multiplicative component of the error (neglecting the sign) at $y = y_f$.

An example of the change in the multiplicative errors of an instrument along the instrument scale is shown in Fig. 10.2, together with the probability density of this error at $y = y_f$.

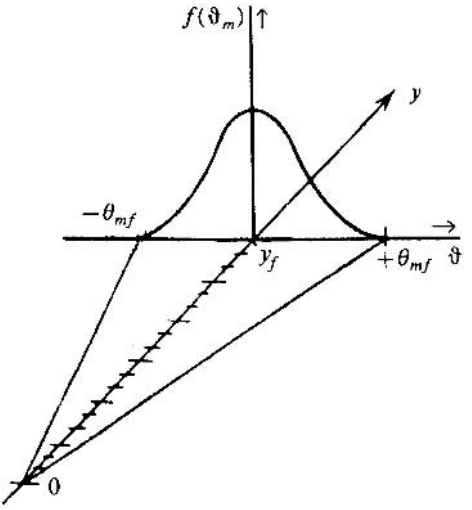


FIGURE 10.2. Example of the representation of multiplicative instrument errors together with their probability density.

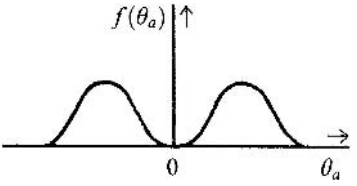


FIGURE 10.3. Possible form of the probability density of the extremal additive instrument errors.

We recall that the multiplicative error increases from the beginning to the end of the scale in proportion to the indications of the instrument.

The additive components of the error, however, vary along the scale in a random manner, but so that $|\theta_{ay}| \leq \theta_{ay}$, where $\theta_{ay} = \varphi(y)$ is the largest additive component of the error (neglecting the sign) at the point y of the scale. For this reason, the additive component must be regarded as a two-dimensional random quantity: It changes in each section of the scale (at each scale marker) as well as along the scale.

For our problem, it is best to not study the entire collection of the additive components of systematic instrument errors, but rather only the collection of the largest and smallest errors, chosen separately for each instrument. Statistically (over the set of instruments), these extremal additive components of the systematic error θ_a are determined by two distribution functions, shown in Figs. 10.3 and 10.4.

The graph presented in Fig. 10 .3 permits finding the probability of the extremal values of θ_a, and the graph presented in Fig. 10.4 permits finding the probability of encountering an extremal value (positive or negative) on one or another section of the scale of the instrument.

Given these dependences, we can find the probability that the error of the manufactured instrument will be less than a predetermined limit at any point of its scale. The solution is obtained by numerical methods. The scheme of the calculations is as follows:

(1) We transfer from the continuous distributions $f(\vartheta_m)$ and $f(\theta_a)$ to discrete distributions. For this reason, the ranges of possible values of the components $[-\theta_{mf}, +\theta_{mf}]$ and $[-\theta_{af}, +\theta_{af}]$ and divided into several intervals so that each interval can be replaced by the average error on it $\theta_{mi} (i = 1, \ldots, h)$ and $\theta_{aj} (j = 1, \ldots, t)$.

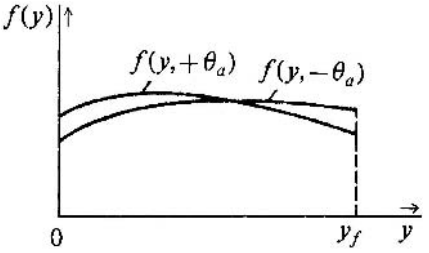


FIGURE 10.4. Possible form of the probability density of the extremal positive and negative additive instrument errors along the scale.

We set the probability that each average will appear in the error interval equal to the area under the curve of the probability density of the corresponding error on this interval.

Thus, we obtain a series of multiplicative and a series of additive errors and the corresponding probabilities:

$$\begin{aligned} &\theta_{mi}, \quad p_{mi}, \qquad i = q, \ldots, h, \\ &\theta_{aj}, \quad p_{aj}, \qquad j = 1, \ldots, t. \end{aligned}$$

We note that the multiplicative errors θ_{mi} correspond to $y = y_f$ and the additive errors θ_{ai} are independent of the position on the scale.

(2) We find the section of the scales on which inadmissibly large errors can appear.

Inadmissably large errors, by definition, are errors that satisfy the inequalities

$$\zeta_l \le -\Delta, \qquad \zeta_r \ge +\Delta. \tag{10.13}$$

We take a pair of components θ_{aj} and θ_{mi} and find a point on the scale of the instrument y_{ij} such that for $y > y_{ij}$, one inequality (10.13) can be satisfied. It is best to study at the same time only positive or only negative components.

Based on the foregoing arguments regarding the separation of the random errors, the inequalities (10.13) will assume the form $\theta_{aj} + \vartheta_{myi} \ge \Delta - \Psi_\alpha$ for positive θ_{aj} and θ_{mi} and the form $\theta_{aj} + \vartheta_{myi} \le -\Delta + \Psi_\alpha$ for negative components. Here

$$\vartheta_{myi} = \theta_{mi}\, y_i / y_f.$$

The solution of the inequalities gives

$$\begin{aligned} y_{ij} &\ge \frac{\Delta - \Psi_\alpha - \theta_{aj}}{\theta_{mi}} y_f, \quad \text{if} \begin{cases} \theta_{aj} > 0, \\ \theta_{mi} > 0, \end{cases} \\ y_{ij} &\ge \frac{\Delta - \Psi_\alpha - \theta_{aj}}{|\theta_{mi}|} y_f, \quad \text{if} \begin{cases} \theta_{aj} > 0, \\ \theta_{mi} < 0, \end{cases} \end{aligned} \tag{10.14}$$

Thus, the section of the scale where inadmissibly large errors can appear for each pair of components is $y_f - y_{ij}$.

(3) We shall calculate the probability that an instrument with an error less than the prescribed limit is manufactured.

The probability that the selected pair of components θ_{mi} and θ_{aj} appears simultaneously is equal to (because they are independent)

$$p = p_{mi}\, p_{aj}. \tag{10.15}$$

The probability that the selected additive component θ_{aj} will appear in the section of the scale $y_f - y_{ij}$ is determined based on one of the two functions $f(y)$ and is equal to

$$p = \int_{y_{ij}}^{y_f} f(y)\, dy.$$

From here, we find the probability that an instrument with an inadmissibly large error for chosen θ_{mi} and θ_{aj} will be manufactured:

$$p_{ij} = p_{mi}\, p_{aj} \int_{y_{ij}}^{y_f} f(y)\, dy. \tag{10.16}$$

To each pair of components, a unique probability p_{ij} is associated. The calculation must be performed separately for positive and negative pairs of components, because in the computational scheme presented above, the left and right-hand branches of the distribution function of the resulting systematic instrument error are taken into account separately.

In the cases where the functions $f(\vartheta_m)$ or $f(\theta_a)$ are unsymmetric, or the functions $f(y)_{+\theta_a}$ and $f(y)_{-\theta_a}$ are not equal, the calculations of p_{ij} must be repeated because different probabilities now correspond to the positive and negative errors with the same absolute value. But if the functions are symmetric and $f(y)_{+\theta_a} = f(y)_{-\theta_a}$, then the calculations can be shortened by setting

$$p_{ij} = 2p_{mi}\, p_{aj} \int_{y_f - y_{ij}}^{yf} f(y)\, dy, \tag{10.17}$$

where the indices i and j now enumerate the positive and negative errors with the same magnitude.

As a result of the calculation, we find

$$p_l = \sum_{\theta<0} p_{ij}, \qquad p_r = \sum_{\theta>0} p_{ij},$$

and then the probability p_g,which we seek, of manufacturing an instrument with an error less than the prescribed limit is

$$p_g = 1 - (p_l + p_r).$$

We obtained the answer for the case in which the practically limiting random error Ψ_α is the same for all instruments. Often, however as we have already noted, it is necessary to take into account that different instruments can have different random errors, i.e., different Ψ_α.

To solve this problem, formulas (10.12) must be used. It is first necessary to establish several discrete values of Ψ_s and the corresponding probabilities p_s. Next, using the scheme presented above, p_{gs} can be found for each Ψ_s. Averaging the obtained probabilities with the weights p_s, we obtain the solution.

$$p_g = \sum_s p_s\, p_{gs}. \tag{10.18}$$

Thus, it is possible to find the probability of manufacturing an instrument with an error less than the prescribed limit if the instrument has one measurement range.

The solution of this problem for instruments with many measuring ranges is, in principle, the same as the solution presented above.

Instruments with voltage dividers, shunts, measuring transformers, and similar devices with a variable transfer coefficient have several measurement ranges. Every instrument has one sample of such devices of one type or another.

One would think that to estimate the probability of getting an instrument with an inadmissibly large error, it is necessary to have the distribution function of the errors of all instrument components. Given these functions, all possible combinations of errors in one instrument must then be sorted through, and the particular combinations that give errors that are less (or greater) than a prescribed limit are then selected. This process is possible but complicated.

The problem simplifies significantly if the distribution of the highest and lowest errors of the instrument components (voltage dividers or shunts) that give the instruments several measurement ranges are studied separately. We shall call these devices multivalued devices.

Each multivalued unit (a particular sample of the unit) can be described by only two errors with the largest absolute values: positive and negative errors. A set of units of one type will then correspond to the distribution function of the largest and smallest errors.

The largest (smallest) error of each instrument consists of the largest (smallest) errors of its components. Correspondingly, the distribution function of the largest (smallest) instrument errors can be constructed according to analogous functions of the components. For multivalued units, this is the distribution function of the largest and smallest errors over the set of units. A single-valued unit has one error distributed over the set of units. After the distribution function of the largest errors of the instruments has been determined, we find the probability sought. An example of the described calculation is presented in Section 10.4.

The foregoing solution contains one fundamental inconsistency. The essence of this inconsistency is that instrument errors must be calculated when the instruments are developed, and the calculation is based on data on the physically nonexistent units of these instruments. It is possible to get out of this difficulty by focusing on units of analogous instruments that are already being produced. The distribution functions of the parameters of such units can be estimated. Of course, experts can introduce into these data certain corrections so they can extrapolate them to the parameters of the units being designed.

In conclusion, it should be added that the accuracy of the calculations in which continuous distributions are replaced by discrete distributions depends on the number of discrete intervals, which can be made very large, if computers are employed. However, the probability of manufacturing an instrument with an error less than the prescribed limit need not be found with high accuracy.

If, however, the starting data are represented in the form of histograms, then in the solution under study, all information contained in them is employed and the computational uncertainty is determined primarily by the uncertainty of the histograms.

10.3. Calculation of the Errors of Electric Balances (Unique Instrument)

Electric balances are an instrument in which the force of interaction of the moving and nonmoving coils, through which the same constant current flows, is balanced by the force of gravity of the weights. Standards of the unit of electric current strength (ampere) have been developed in the United States and the former USSR based on this principle. (Now the ampere is reproduced with higher accuracy using standards of the volt and ohm. Therefore, the electric balance lost its value.)

We shall calculate the errors of Soviet-made electric balances. We shall use the data presented in [27].

The current strength at the point of equilibrium of the electric balance is determined by the expression

$$I = \sqrt{mg/F}, \tag{10.19}$$

where m is the mass of the balancing weight, g is the acceleration of gravity, and F is the constant of the electric balance.

The constant of the electric balance is equal to the derivative of the mutual inductance of the two coils (mobile and immobile) with respect to the vertical displacement of the mobile coil and is calculated from their geometric dimensions.

The difference between the value of the current strength calculated using formula (10.19) and its true value, i.e., the error of the electric balance, is determined by the uncertainty in the quantities entering into this formula as well as by the effect of the field of the wires carrying the current to the mobile coil. These sources of error create the systematic error of the electric balance.

The equilibration of the balance, however, is also accompanied by random errors, which are caused by friction in the supports of the cross arm of the balance, fluctuations of the ambient air temperature, changes in the external magnetic field, effect of air flows, and some other factors.

The systematic error of the electric balance must be estimated by a computational method; it cannot be determined experimentally (as long as one is not concerned with comparing the national standard of the unit of current strength with the standard of this unit in other countries). It is, however, virtually impossible to calculate the parameters of the random errors, but they can be estimated based on the experimental data. For our balances, the relative standard deviation of the current strength is $S_{\text{rel}} = 2 \times 10^{-6}$.

The uncertainty of the quantities entering into formula (10.19) is characterized by the following data. For the mass of the balancing load, the relative error falls within the limits $\pm 1.25 \times 10^{-6}$, and for the acceleration of gravity, the error falls within the limits $\pm 4 \times 10^{-6}$. (At the present time, this error can be significantly smaller.)

The error in the constant of the electric balance is in turn caused by several factors. Table 10.1 gives the limits of the errors introduced into the constant of the electric balance by each factor [27].

TABLE 10.1. Limits of the components of the errors $\varepsilon(F)$.

Source of error	Limits of error δF in the constant $\times 10^{-6}$
Uncertainty in the measurement of the radial dimensions:	
Immobile coil $\delta F(r_{\text{im}})$	± 3
First part of the moving coil $\delta F(r_{m1})$	± 3
Second part of the moving coil $\delta F(r_{m2})$	± 2
Uncertainty in the measurement of the axial dimensions:	
Immobile coil $\delta F(\ell_{\text{im}})$	± 2
First part of the moving coil $\delta F(\ell_{\text{m1}})$	± 1.3
Second part of the moving coil $\delta F(\ell_{\text{m2}})$	± 0.7
Deviation of the coils from the cylindrical shape $\delta F(R)$	± 2

We shall find the influence coefficients of the relative errors of the measurements of the mass $\varepsilon(m)$, acceleration of gravity $\varepsilon(g)$, and the calculation of the constant of the electric balance $\varepsilon(F)$ hi accordance with formula (10.3).

We represent expression (10.19) in the form of the product of the arguments:

$$I = m^{1/2} g^{1/2} F^{-1/2}.$$

As shown in Section 6.6, in this case, the influence coefficients are equal to the powers of the corresponding arguments, i.e.,

$$w_m = \tfrac{1}{2}, \qquad w_g = \tfrac{1}{2}, \qquad w_F = -\tfrac{1}{2}.$$

Aside from the enumerated and estimated components of the error, it is also necessary to take into account the error mentioned above because of the influence of the field generated by the wires conducting the current to the mobile coil. Experiments show that this field creates an additional force on the mobile part that falls within $\pm 2 \times 10^{-6}$ times the nominal strength of the interaction of the coils.

As the influence coefficient of the force of interaction (mq) is $w_{mq} = \frac{1}{2}$, this error has the same influence coefficient $w_H = \frac{1}{2}$.

According to formula (10.6), the total systematic error of the electric balance (in the relative form) will be

$$\varepsilon_\Sigma = w_H \varepsilon(H) + w_m \varepsilon(m) + w_g \varepsilon(g) + w_F \varepsilon(F),$$

where $|\varepsilon(H)| \le 2 \times 10^{-6}$, $|\varepsilon(m)| \le 1.25 \times 10^{-6}$, $|\varepsilon(g)| \le 4 \times 10^{-6}$, and a series of components was given for the error $\varepsilon(F)$.

All components of the error ε_Σ are determined by their limits. For this reason, we shall use formula (10.10) and find the confidence systematic error of the electric balance. We shall take $\alpha = 0.95$ and $k = 1.1$. Then

$$\theta_{0.95} = 1.1\sqrt{\left(\tfrac{1}{2}\right)^2 (2^2 + 1.25^2 + 4^2 + 2 \times 3^2 + 3 \times 2^2 + 1.3^2 + 0.7^2) \times 10^{-12}}$$
$$= 1.1 \times 10^{-6}\sqrt{13.5} = 4 \times 10^{-6}.$$

The practically limiting random error can be estimated if we know S_{rel} and have some idea about the form of the distribution of the experimental data. If it can

be assumed that the data correspond to a normal distribution and the confidence probability is also taken to be 0.95, then $\Psi_{0.95} = 1.96S_{\text{rel}} = 4 \times 10^{-6}$. Then the practically limiting error of the electric balance in the single-balancing regime will be, according to formula (10.9),

$$\delta I(0.95) = \theta_\alpha + \Psi_\alpha = 8 \times 10^{-6}.$$

When measuring with the help of electric balances the emf of standard cells, several balancings can be performed and the error in the result can be reduced by averaging the data obtained.

We underscore that the result obtained pertains to a specific sample of electric balances, because it was obtained using data on the parameters of the components of this instrument.

10.4. Calculation of the Error of ac Voltmeters (Mass-Produced Instrument)

We shall study the inverse problem. The limit of permissible instrument error as the fiducial error is given, and it is required to calculate the percentage of instruments satisfying this requirement when they are manufactured using a technology with a prescribed degree of development.

We shall study the voltmeter in a ferrodynamic system. Figure 10.5 shows the block diagram of this instrument. This diagram was constructed in accordance with the theory of instruments of this system. Figure 10.6 shows a graphical representation of the instrument scale.

Block 1 converts the measured voltage U_x into the current of strength

$$I = U_x/R,$$

where R is the resistance of the input circuit of the voltmeter.

The current I is converted with the help of block 2 into a torque

$$M_t = KI^2,$$

where K is the electrodynamic constant of the instrument.

Block 3 generates a countertorque

$$M_c = W\alpha,$$

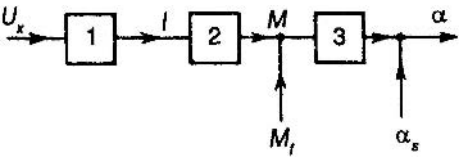


FIGURE 10.5. Block diagram of voltmeter.

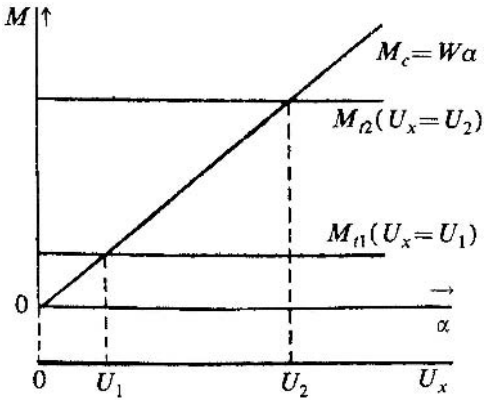


FIGURE 10.6. Graphical construction of the voltmeter scale.

where W is the stiffness of the spring and α is the angle of rotation of the moving part.

When the moving part is in a position of equilibrium, $M_t = M_c$, and from here,

$$\alpha = \frac{K}{WR^2} U_x^2. \tag{10.20}$$

When the instrument is manufactured, the particular combination of the parameters K, W, and R that is realized in the instrument is fixed by adjusting the instrument and calibrating its scale. For this reason, instrument errors will arise only as a result of changes in the stiffness of the spring and the input resistance relative to their values at the moment of regulation. The constant K, however, is virtually unchanged and does not give rise to any errors.

Regarding the parameters W and R, formula (10.20) is exact and permits funding the instrument error introduced by changes in these parameters. Structurally it is identical to formula (6.37). For this reason, the values of the influence coefficients for the relative changes in the stiffness W and resistance R can be written immediately:

$$w_W = -1 \qquad \text{and} \qquad w_R = -2.$$

In addition to the instability of the parameters of the blocks, error can also appear from the friction in the supports of the moving part and the uncertainty of the scale. The sources of these errors are indicated in Fig. 10.5. As these errors are additive, it is best to express them as absolute errors. We shall express them in units of the angle of rotation of the moving part.

Friction introduces a random error (M_f is the friction moment). It is customarily described as a dead band, i.e., by the difference of the indications of the instrument that is obtained by approaching continuously from the right and left a particular marker on the scale. The largest dead band is determined (see, for example, the standards for electric measuring instruments). The largest random

error from friction is equal to one half the dead band. Therefore, the limits Ψ_t of this error are also known. It is assumed that the random errors of each instrument are uniformly distributed within these limits. The limits, however, can be different for different instruments.

The error in the scale of the instrument α_s for each scale marker of a particular instrument is a systematic error. But this error varies from one marker to another. It also varies from one instrument to another. For each particular instrument, it is possible to find the largest error of the scale. It can be assumed that this error is encountered with equal probability on any scale marker. The set of instruments is characterized by the distribution of these largest scale errors.

So, the components of the instrument errors are as follows:

(a) the error from the variation of spring stiffness $\vartheta_1 = -\varepsilon_W$;
(b) the error from the variation of the input resistance $\vartheta_2 = -2\varepsilon_R$;
(c) the error from the friction $\psi = \alpha_f$; and
(d) the error from scale inaccuracies $\vartheta_3 = \alpha_s$.

The errors ϑ_1 and ϑ_2 are multiplicative and are expressed as a percentage; the errors ϑ_3 and ψ are additive and are expressed in units of the angle of rotation of the moving part of the voltmeter (in degrees); i.e., they are referred to the output. For this reason, the instrument error scaled to the output is given by the relation

$$\zeta_\alpha = \frac{\alpha}{100}(\vartheta_1 + \vartheta_2) + \vartheta_3 + \psi, \tag{10.21}$$

where α is the angle of rotation of the moving part of the instrument, which corresponds to its indication U_x for which the error is calculated.

Let us assume that the limits of intrinsic error are given and they are $\pm 1\%$ (as the fiducial errors). Next, we assume that the data given in Table 10.2 are known for

TABLE 10.2. Starting data on the sources of systematic instrument errors.

Source of error or error	Interval of probability distribution		Frequency of occurrence of the interval
	Left limit	Right limit	
Relative change in spring stiffness ε_W	-0.3%	-0.2%	0.2
	-0.2%	-0.1%	0.5
	-0.1%	0.0%	0.3
Relative change in resistance ε_R	-0.3%	-0.1%	0.2
	-0.1%	$+0.1\%$	0.2
	$+0.1\%$	$+0.3\%$	0.6
	-0.6°	-0.2°	0.5
Absolute error of the instrument scale α_s	-0.2°	$+0.2^\circ$	0.0
	$+0.2^\circ$	$+0.6^\circ$	0.5

TABLE 10.3. Description of histograms of distributions of multiplicative components of the instrument error.

Error	Interval of error distribution (%)		Frequency of occurrence of interval
	Left limit	Right limit	
ϑ_1	+0.2	+0.3	0.2
	+0.1	+0.2	0.5
	0.0	+0.1	0.3
ϑ_2	+0.2	+0.6	0.2
	−0.2	+0.2	0.2
	−0.6	−0.2	0.6

each component of the error. These data characterize the degree of development of the manufacturing technology.

We shall also assume that in 30% of the instruments, the practically largest dead band does not exceed 0.8°, and that in 70% of the instruments, it does not exceed 0.4°. Thus, the random error falls within the limits $\Psi_{\alpha 1} = \pm 0.4^\circ$ for 30% of the instruments and within the limits $\Psi_{\alpha 2} = \pm 0.2^\circ$ for 70% of the instruments. Given these data, we must find the probability that the fiducial error of the voltmeters falls within the limits $\Delta = \pm 1\%$.

The fiducial error must be put into the form of an absolute error. We shall express it in degrees of rotation of the moving part, which can be done with the help of a graph similar to that presented in Fig. 10.6. Assume that in our case, the limit of permissible error in degrees $\Delta = 1^\circ$ (neglecting the sign).

Focusing on formula (10.21), we shall first find the composition of the multiplicative errors ϑ_1 and ϑ_2. Using the data presented in Table 10.2 and the influence coefficients found, it is not difficult to describe the histograms of the distributions of these errors. These descriptions are given in Table 10.3.

It is convenient to solve the problem by the method of sorting, described in Section 3.6. For this reason, the histograms must be replaced by discrete distributions. To each interval, an error is assigned equal to the center of the interval. The probability of the appearance of this error is assumed to be equal to the frequency of this interval.

Let the error ϑ_1 be represented by the discrete random quantity η_1 and the error ϑ_2 by the discrete random quantity η_2. We obtain the following:

$$\begin{array}{lrrr} \eta_1 & +0.25 & +0.15 & +0.05, \\ p_1 & 0.20 & 0.50 & 0.30, \\ \eta_2 & +0.40 & 0.00 & -0.40, \\ p_2 & 0.20 & 0.20 & 0.60. \end{array}$$

The random quantity $\eta = \eta_1 + \eta_2$ corresponds to the error $\vartheta_1 + \vartheta_2$. Its realizations are presented in Table 10.4.

TABLE 10.4. Discrete representation of the distribution of the multiplicative instrument error.

Number	$\eta = \eta_1 + \eta_2$	$p = p_1 p_2$
1	$+0.25 + 0.4 = +0.65$	0.04
2	$+0.25 + 0.0 = +0.25$	0.04
3	$+0.25 - 0.4 = -0.15$	0.12
4	$+0.15 + 0.4 = +0.55$	0.10
5	$+0.15 + 0.0 = +0.15$	0.10
6	$+0.15 - 0.4 = -0.25$	0.30
7	$+0.05 + 0.4 = +0.45$	0.06
8	$+0.05 + 0.0 = +0.05$	0.06
9	$+0.05 - 0.4 = -0.35$	0.18

The limiting values of the total error $\eta_{\min} = -0.6\%$ and $\eta_{\max} = +0.9\%$ (see Table 10.3); these errors correspond to probabilities of 0 and 1, respectively. The probability distribution is constructed based on the obtained data. The numerical values are summarized in Table 10.5.

Based on these data we construct a step curve as a first approximation to the distribution function that is sought for the multiplicative error of the instruments, after which the function is smoothed by the method of linear approximation. The distribution function so obtained is presented in Fig. 10.7.

We shall now express the multiplicative error in the form of absolute error—as fractions of the angle of rotation of the mobile part. We shall find the largest error, i.e., the error corresponding to maximum deflection. We assume that $\alpha_{\max} = 100^\circ$. Then the numerical values of the error $\vartheta_m = (\vartheta_1 + \vartheta_2)\ (\alpha_{\max}/100)$ will be equal to the values given in Table 10.5. Using these data and the graph in Fig. 10.7, we construct a histogram of the multiplicative error of the instrument with

TABLE 10.5. Table of the computed values of the probability distribution of multiplicative instrument error.

η	-0.6	-0.35	-0.25	-0.15	$+0.05$
p	0	0.18	0.30	0.12	0.06
Σ_p	0	0.18	0.48	0.60	0.66

η	$+0.15$	$+0.25$	0.45	0.55	0.65	0.9
p	0.10	0.04	0.06	0.10	0.04	0.0
Σ_p	0.76	0.80	0.86	0.96	1.00	1.0

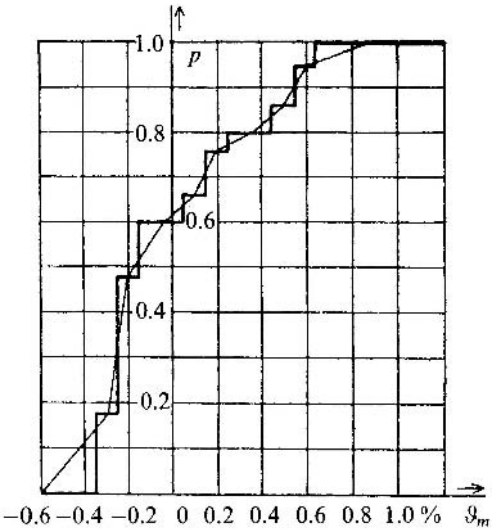


FIGURE 10.7. Step and linear approximations of the distribution function of the multiplicative errors of voltmeters.

maximum angle of rotation of the mobile part. The data for this histogram are as follows:

Interval number i	1	2	3	4	5
Limits of interval for the error ϑ_m (degrees)					
left	+0.60	+0.30	+0.00	−0.30	−0.60
right	+0.90	+0.60	+0.30	0.00	−0.30
Average value of θ_{mi}	+0.75	+0.45	+0.15	−0.15	−0.45
Probability of falling within the interval p_{mi}	0.05	0.15	0.20	0.42	0.18

We obtain, analogously, from the data in Table 10.2, the average values of the intervals of the distribution of the largest scale errors of the instruments and the corresponding probabilities:

Interval number j	1	2	3
Average value θ_{aj} of the error in the interval	+0.4	0	−0.4
Probability of falling within the interval p_{aj}	0.5	0	0.5

In accordance with formulas (10.14), we find the sections of the scale where the instrument can be rejected. In our case, $\Delta_1 = \Delta - \Psi_{\alpha 1} = 0.6^\circ$ and $\Delta_2 = \Delta - \Psi_{\alpha 2} = 0.8^\circ$ (for 30% and 70% of the instruments, respectively).

For $\theta_{aj} > 0, \theta_{mi} > 0$, and $\Delta_1 = 0.6^\circ$, we obtain

$$y_{11} = \frac{0.6 - 0.4}{0.75} y_f = 0.27 y_f, \qquad y_{21} = \frac{0.6 - 0.4}{0.45} y_f = 0.45 y_f,$$
$$y_{12} = \frac{0.6}{0.75} y_f = 0.8 y_f.$$

The remaining combinations give $y_{ij} > y_f$, which means that it is impossible to obtain an inadmissibly large error.

We shall assume that the distribution function of the additive errors of the instruments along the scale uniform and identical for positive and negative errors: $f(y) = 1/y_f$. Then for each scale section studied, we obtain

$$p'_{ij} = \int_{y_{ij}}^{y_f} \frac{1}{y_f}\, dy = \left(1 - \frac{y_{ij}}{y_f}\right).$$

Therefore,

$$p'_{11} = (1 - 0.27) = 0.73, \qquad p'_{21} = (1 - 0.45) = 0.55,$$
$$p'_{12} = (1 - 0.80) = 0.20.$$

For $\theta_{aj} < 0, \theta_{mi} < 0$, and Δ_1, we obtain

$$y_{53} = \frac{0.6 - 0.4}{0.45} y_f = 0.45 y_f, \quad p'_{53} = 0.55.$$

From here, we find the probability that an instrument is rejected for each combination of instrument components:

$$p_{11} = 0.73 \times 0.05 \times 0.5 = 0.018,$$
$$p_{21} = 0.55 \times 0.15 \times 0.5 = 0.041,$$
$$p_{12} = 0,$$
$$p_{53} = 0.55 \times 0.18 \times 0.5 = 0.050.$$

Therefore,

$$p_l = p_{53} = 0.050, \qquad p_r = p_{11} + p_{21} + p_{12} = 0.059.$$

The probability of manufacturing a high-quality instrument is

$$P_{g1} = 1 - (0.050 + 0.059) = 0.89.$$

Analogous calculations for $\Delta_2 = 0.8$ give

$$y_{11} = \frac{0.8 - 0.4}{0.75} y_f = 0.53 y_f, \qquad p'_{11} = 0.47,$$
$$y_{21} = \frac{0.8 - 0.4}{0.45} y_f = 0.89 y_f, \qquad p'_{21} = 0.11,$$
$$y_{53} = \frac{0.8 - 0.4}{0.45} y_f = 0.89 y_f, \qquad p'_{53} = 0.11;$$

$$p_{11} = 0.47 \times 0.05 \times 0.5 = 0.012,$$
$$p_{21} = 0.11 \times 0.15 \times 0.5 = 0.008,$$
$$p_{53} = 0.11 \times 0.18 \times 0.5 = 0.010.$$

Now $p_l = 0.010$ and $p_r = 0.020$, and $p_{g2} = 0.97$.

The weighted-mean probability of manufacturing an instrument whose error is less than the prescribed limit is equal to

$$p_g = 0.3p_{g1} + 0.7p_{g2} = 0.3 \times 0.89 + 0.7 \times 0.97 = 0.95.$$

Therefore, for the properties of the instrument components and the scale fabrication quality presented above, approximately 95% of the instruments will have a fiducial error not exceeding 1%.

This calculation was performed for reference conditions and determines the limits of intrinsic instrument error.

For a prescribed limit of instrument error, the obtained percentage of rejections can serve as a basis for increasing the quality requirement for one or another of the instrument components, improving the technology used to fabricate the components, and so on. The limits of permissible errors of all components can be calculated uniquely if the weights are assigned for their errors. These weights are apparently difficult to determine objectively, and sometimes it is impossible to do so. For this reason, the main method is to estimate the percentage of rejects and select specifications of the instrument components so that the percentage of rejects is acceptable.

10.5. Calculation of the Error of Digital Thermometers (Mass-Produced Instrument)

Digital thermometers are usually constructed according to a scheme in which the digital-analog integrator converts the emf of the thermocouple into a corresponding voltage, after which this voltage is converted into a proportional time interval and thus into the indication of the instrument. This process is explained by the graphs presented in Fig. 10.8.

The graph in Fig. 10.8(a) refers to an analog-digital integrator. The integration time is maintained strictly constant, and the slope of the straight lines is proportional to the emf at the input of the integrator (in accordance with the principle of operation of the integrator). For this reason, the voltage at the output of the integrator U_i is proportional to the emf of the thermocouple.

The graph in Fig. 10.8(b) shows how the voltage U_i is converted into a proportional time interval Δt_i; this is equivalent to conversion into a number—the indication of the instrument.

As the temperature dependence of the thermocouple emf is known, the indications of the instrument give the measured temperature.

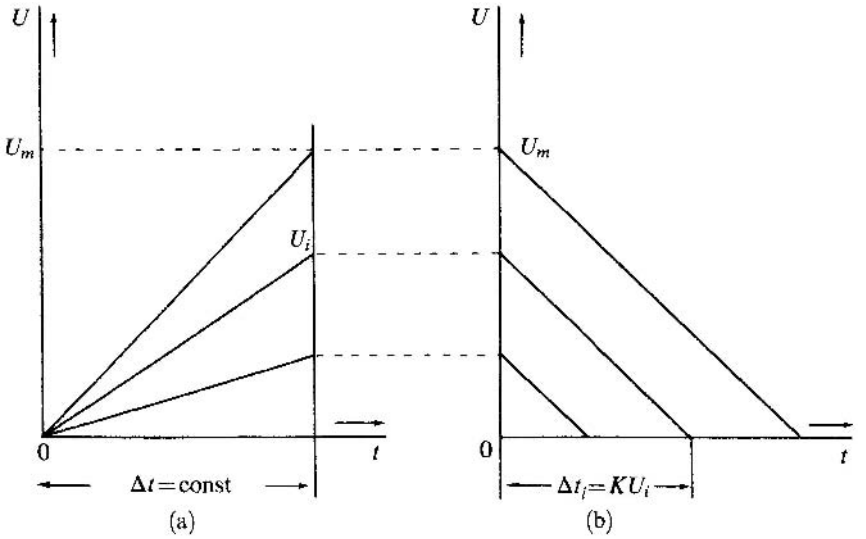


FIGURE 10.8. Graphs explaining the principle of operation of a digital integrating instrument.

Two special procedures are realized in the process of the conversions. One is linearization of the temperature dependence of the thermocouple emf. The other is compensation of the effect of a deviation of the temperature at the so-called cold ends of the thermocouple from the reference level. The latter is assumed to be 0 °C, the melting point of ice.

The specifications provided by manufacturers of thermocouple thermometers still do not give the user clear indications of the temperature measurement accuracy that can be achieved with the instrument. One would think that the well-known Fluke Company, which introduced the concept of total accuracy of digital thermometers [13], would have filled in this omission. The catalog of the firm shows, however, that it gives the thermometer error without taking into account the thermocouple error; i.e., it does not complete the solution of the problem.

We shall study the calculation of the error of digital thermometers based on data on the accuracy of their components and the calibration accuracy. We shall take a thermometer with a thermocouple of type J and a measurement range of 0–750 °C. The instrument can be used in the range $+25 \pm 10$ °C, and after calibration, it can be used for one year. We shall consider the direct problem.

Focusing on [13], we shall assume that the following data are known (the numbers for the calculation here are arbitrary):

1. Linearity: The limits of instrument error caused by deviation from linearity of the characteristic $\theta_L = \pm 0.1$ °C.
2. Reference junction

2.1. The limits of instrument error caused by the effect of a deviation of the temperature by 10 °C from the normal temperature (+25 °C) on this circuit $\theta_{TJ} = \pm 0.2\,^{\circ}\mathrm{C}$.
2.2. The limits of instrument error caused by the change in the parameters of this circuit over a period of one year (instability), $\theta_{SJ} = \pm 0.2\,^{\circ}\mathrm{C}$.

3. Reference voltage.
3.1. The limits of instrument error caused by the effect of a deviation of the temperature by 10 °C from the normal temperature on this circuit, $\theta_{TU} = \pm 0.5\,^{\circ}\mathrm{C}$.
3.2. The limits of instrument error caused by a change in the parameters of this circuit over a period of one year (instability), $\theta_{SU} = \pm 0.2\,^{\circ}\mathrm{C}$.

4. Correspondence to NBS (NIST) data. The limits of instrument error caused by inaccurate linearization of the standard characteristic of the thermocouple, $\theta_{\mathrm{sc}} = \pm 0.15\,^{\circ}\mathrm{C}$.
5. The limits of instrument error caused by the discreteness of the indications, $\Psi_D = \pm 0.5\,^{\circ}\mathrm{C}$.
6. The limits of instrument calibration error, $\theta_C = \pm 0.15\,^{\circ}\mathrm{C}$.

The absolutely constant error, i.e., the error that is the same for all instruments of a given type, will be the error caused by inconsistency with the NBS (NIST) data.

The random error will be the error introduced by the discreteness of the instrument indications. All other errors must be regarded as conditionally constant.

We shall regard conditionally constant errors as uniformly distributed random quantities, as has already been assumed above. Their characteristic feature, in our case, is that some of them depend on one another. Taking this dependence into account, the total conditionally constant error of the thermometers must be calculated using the following formula (for probability $\alpha = 0.95$):

$$\theta_1 = 1.1\sqrt{\theta_L^2 + (\theta_{TJ} + \theta_{TU})^2 + (\theta_{SJ} + \theta_{SU})^2 + \theta_C^2} = 1.1\sqrt{1.01} = 1.1\,^{\circ}\mathrm{C}.$$

The limits of absolutely constant error must be summed with the limits θ_1 arithmetically. We obtain

$$\theta_2 = \theta_{\mathrm{sc}} + \theta_1 = 0.15 + 1.1 = 1.25\,^{\circ}\mathrm{C}.$$

The limits of random error must also be taken into account by arithmetic summation, because we are estimating errors of the instrument in every use of the instrument in the future:

$$\Delta = \theta_2 + \Psi_D = 1.25 + 0.5 = 1.75\,^{\circ}\mathrm{C}.$$

Although Δ is expressed in units of the measured quantity, this is still not the total error of the instrument as a thermometer, because it does not include the thermocouple error, and without the thermocouple, the instrument cannot operate as a thermometer.

According to ANSI Standard MC 96.1, the limits of error of thermocouples of type J in the temperature range 0–750 °C are ±2.2 °C or ±0.75%, whichever is

greater. It is easy to calculate that up to 293 °C, the limits of error will be ±2.2 °C, after which they must be calculated using the relation $\pm 0.75 \times 100 \times T_x$, where T_x is the measured temperature. At $T_x = 750\,^{\circ}\mathrm{C}$, this error will fall within the limits ±5.6 °C.

The thermocouple errors must be regarded as conditionally constant, and for this reason, they must be taken into account when calculating θ_1. Taking this component into account, we obtain $\theta_1' = 1.1\sqrt{1.01 + 5.6^2} = 1.1\sqrt{32.4} = 6.3\,^{\circ}\mathrm{C}$.

After this result, we find the absolute value of the limits of the total instrument error:

$$\Delta' = \theta_1' + \theta_{\mathrm{sc}} + \Psi_D = 6.3 + 0.15 + 0.5 = 6.95 \approx 7\,^{\circ}\mathrm{C}.$$

If this limit is represented in the form of fiducial error, we obtain $\gamma = \pm 7\,^{\circ}\mathrm{C}/750\,^{\circ}\mathrm{C} \times 100 \approx \pm 1\%$.

In this case, this value of the fiducial error will be identical to the limit of relative error in the range 300–750 °C. Under this temperature range, the limit of relative error starts to increase. Thus, the limits of instrument error found in the range 300–750 °C are also an estimate of the limits of minimum error in measuring the temperature that one can count on when using a thermometer consisting of a type J thermocouple and the digital indicator under study. In other words, this estimate of the error characterizes the maximum temperature measurement accuracy expected for this thermometer.

Characterizing the inaccuracy of the indicator of the thermometer is not only insufficient for the user, but it can even mislead the user, because its error is expressed in units of measurement of temperature, whereas it characterizes only the indicator of the thermometer. The difference is significant: $\Delta = 1.75\,^{\circ}\mathrm{C}$, whereas $\Delta' = 7\,^{\circ}\mathrm{C}$.

11
Problems in the Theory of Calibration

11.1. Types of Calibration

Every country wishes to have trustworthy measurements of all physical quantities. One of the most important arrangements to achieve this goal is to have a system for keeping errors of all measuring instruments within permissible limits. Therefore, all measuring instruments in use are periodically checked. In the process, working standards are used either to verify that the errors of the measuring instruments being checked do not exceed their limits or the measuring instruments are recalibrated.

The general term for the above procedures is *calibration.* But one should distinguish between a real calibration and a simplified calibration.

Real calibration results in the determination of a relation between the indications of a measuring instrument and the corresponding true values of a measurand. This relation can be expressed in the form of a table, a graph, or a function. It can also be expressed in the form of the table of corrections to the indications of the measuring instrument.

The simplified calibration (also called verification) simply reveals whether the errors of a measuring instrument exceed their specified limits.

Essentially, verification is a specific case of quality control, much like quality control in manufacturing. And because it is quality control, verification results do have some rejects.

In addition, a check of an entire set of elements is distinguished from a check of a single element. In a complete check, the error of the checked measuring instrument is determined as a whole, whereas in the case of an elementwise check, the errors of the elements of the measuring instrument being checked are determined. A complete check is always preferable; such a check gives the most reliable solution to the problem. In some cases, however, a complete check is impossible to perform and one must resort to an elementwise check.

In an elementwise check, the error of the measuring instrument being checked is calculated by means of the same methods that were examined in Section 10.2 for solving the direct problem of calculating the errors of a measuring instrument from the errors of its components. The data required for the calculation are obtained by measuring the parameters of the components of the measuring instrument being

checked. Usually, however, this problem is solved completely only once, and in so doing, the standards for the errors of the components are determined. In the future, when a check is performed, only the parameters of the components and the serviceability of the measuring instrument are checked. If the parameters of the components satisfy the standards established for them, then the error of the measuring instrument checked in this manner falls within the established limits.

Elementwise calibration is often employed to check the measuring systems when the entire system cannot be delivered to a standard laboratory and the laboratory does not have necessary working standards that could be transported to the system's site.

The standardization of the metrological properties of the units of a system does not present any difficulties, and the units must be checked by standard methods.

As a rule, the operation of measuring systems cannot be interrupted and interruption for checking is inadmissible. For this reason, in most cases, systems are assembled with a redundant set of units, so that units that are removed for checking could be replaced with units that are known to be serviceable. During the regular check of the system, the units are once again interchanged.

When a system is checked, however, in addition to checking the units, it is also necessary to check the serviceability of the system as a whole. The methods for solving this problem depend on the arrangement of the system, and it is hardly possible to make general recommendations here. For example, the following procedure can be used for a system with a temperature measuring channel.

After the serviceability of all units of the system has been checked, we note some indication of the instrument at the output of the system. Assume that the indication is $+470\ ^\circ$C. Then we find from the nominal calibration characteristic of the primary measuring transducer the output signal that should be observed for the given value of the measured quantity. Thus, if a platinum-rhodium-platinum thermocouple was used as the measuring transducer, then when a temperature of $+470\ ^\circ$C is measured, the emf at the output of the thermocouple must be equal to 3.916 mV. Next, disconnecting the wires from the thermocouple and connecting them to a voltage exactly equal to the nominal output signal of the thermocouple, we once again note the indication of the system. If it remains the same or has changed within the limits of permissible error of the thermocouple and voltmeter, then the system is serviceable.

Of course, this method of checking will miss the case in which the error of the thermocouple is greater than the permissible error and the same is true for the voltmeter, and these errors mutually cancel. However, this result can happen only rarely. Moreover, such a combination of errors is in reality permissible for the system.

At the same time, this method of checking permits evaluating at the same time the state of the thermocouple. The error of the thermocouple usually makes the largest contribution to the error of the measuring system. For this reason, the difference of the indications—observed at the moment of the check and obtained after the thermocouple is disconnected and its output signal is replaced by a nominal

signal—must be less than the limit of permissible error of the thermocouples; it is permissible for this limit to be exceeded by an insignificant amount, which is determined by the accuracy of the system.

The foregoing method of checking the metrological state of measuring systems based on the use of redundant units is also promising in application to many other complicated modern measuring devices, which are technically difficult or impossible to transport to a metrological organization, as well as to devices whose operation cannot be interrupted.

11.2. Estimation of the Errors of Measuring Instruments in Verification

The error ζ of a measuring instrument is defined by the formula

$$\zeta = A_c - A,$$

where A_c is the indication of the instrument being checked, the nominal value of the standard, and so on, and A is the true value of the measured quantity, the quantity reproduced by the standard being checked, and so on.

The true value A is always unknown. If instead of the true value the corresponding indication of a working standard A_r is used, then instead of ζ, we obtain

$$\zeta' = A_c - A_r. \tag{11.1}$$

To estimate the error ζ by ζ', the difference $\zeta' - \zeta$ must be small. The error of the working standard is

$$\gamma = A_r - A.$$

For this reason,

$$\zeta - \zeta' = \gamma. \tag{11.2}$$

Most often, it is known only that the error of the working standard does not exceed the limit Δ_s established for it. Then

$$|\zeta' - \zeta| \leq \Delta_s.$$

In the relative form, the error of the error ζ, as follows from the expression

$$\varepsilon = \frac{|\zeta' - \zeta|}{\zeta},$$

depends on the error ζ and increases as ζ decreases.

It is natural to estimate this relative error as

$$\tilde{\varepsilon} = \Delta_s / \zeta'.$$

When a working standard is chosen, the limit Δ of permissible error of the measuring instrument being checked usually serves as the starting point. In this

case, the ratio

$$k = \Delta_s/\Delta$$

comes into play.

The relative error of the error can be expressed in terms of k:

$$\tilde{\varepsilon} = k\frac{\Delta}{\zeta'}. \tag{11.3}$$

For example, for $k = 0.1$, the error $\zeta' \approx 0.3\Delta$ is estimated with a relative error reaching 30%.

The errors need not be estimated accurately, but the error in estimating errors does not exceed 30%. When the errors exceed this limit, and taking into account the instability of the measuring instruments, the estimates obtained rapidly become meaningless.

In practice, the value $k = 0.3$ is often used. Then, as follows from the relation (11.3), an error of only $\zeta \approx \Delta$ can be estimated with an error not exceeding 30%.

It is interesting to extend the foregoing arguments to the case in which the measuring instruments have significant random errors.

Random errors cause the indications of instruments to be non-single-valued, and they make it difficult both to check and use instruments. If, for example, when checking a pointer-type instrument one need only check whether its errors do not exceed the limit established for them, then the measurements must be repeated several times and the largest errors must be found. In many fields of measurement, the input to an instrument can be varied continuously. In such cases, to determine the largest error at each scale marker checked, two measurements are often sufficient: one by approaching the scale marker from below and the other by approaching from above.

We shall examine a check in which the same quantity is measured simultaneously with a working standard and the instrument being checked. Let y denote the indications of the working standard and x those of the instrument being checked. The difference of the indications of the two devices is

$$z = x - y. \tag{11.4}$$

In the general case,

$$x_i = A + \vartheta_x + \psi_{xi}, \qquad y_i = A + \vartheta_y + \psi_{yi}, \tag{11.5}$$

where ϑ_x and ϑ_y are the systematic errors and ψ_{xi} and ψ_{yi} are the random errors of the instruments in the ith check.

For the random instrument errors, we have

$$E[\psi_{xi}] = 0, \qquad E[\psi_{yi}] = 0.$$

Assume that to find the corrections, the indications of the instruments are averaged. Using relations (11.4) and (11.5), we obtain

$$\frac{\sum\limits_{i=1}^{n} z_i}{n} = \vartheta_x + \frac{\sum\limits_{i=1}^{n} \psi_{xi}}{n} - \left(\vartheta_y + \frac{\sum\limits_{i=1}^{n} \psi_{yi}}{n} \right).$$

For a sufficiently large number of observations the effect of the random errors of the instrument being checked becomes insignificant. Hence,

$$\frac{\sum\limits_{i=1}^{n} \psi_{xi}}{n} \ll \frac{\sum\limits_{i=1}^{n} zi}{n}.$$

Assuming that the error of the working standard is also small, we obtain the answer

$$\tilde{C} = -\bar{z}.$$

The obtained estimate was found with an error not less than the systematic error of the working standard. This error can be estimated by the method described in Section 5.6. The required number of observations can be found with the help of the criterion presented in the same section. For this reason, $S(\bar{z})$ must be compared with the limit Δ_s of permissible error of the working standard:

$$\frac{\Delta_s}{S(\bar{z})} \geq 7.$$

From here,

$$n \approx 7 \frac{\sqrt{\sum\limits_{i=1}^{n} (z_i - \bar{z})^2}}{\Delta_s}.$$

If, for the working standard, the limit of systematic error θ_s is known

$$|\vartheta_y| \leq \theta_s,$$

then in the relations presented, θ_s must be substituted for Δ_s.

The checking method studied above is convenient for analysis, but in practice, it is avoided, because it is difficult to read accurately fractions of a graduation on the scale of the instrument being checked. The results presented for this method of checking are, however, general. In particular, they will also be valid for the main method of verification, in which the indicator of the instrument being checked is set every time on the scale marker being checked, and the corresponding real value of the measured quantity is found based on the indications of a working standard. The same is true for digital instruments.

If the check is made with a reduced number of measurements, thanks to the smooth approach from both sides of the same scale marker of the instrument being

checked, then the estimate of the correction is found based on the arithmetic mean of the two estimates obtained for the error.

Even though the correction was estimated by averaging the indications of instruments, it can then be introduced into each separate indication. After the correction has been introduced, it can be assumed that for the random errors, $E[\psi] = 0$.

When the random errors are significant, it is sometimes desirable to estimate the variance or the standard deviation of this error. If the check is made by measuring an unchanged and known quantity, then (11.4) is valid and it is obvious that

$$S(x_i) = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

and formally there are no difficulties in solving the problem; however, it is difficult to obtain readings x_i that are accurate enough.

If, however, in each observation, the indicator of the instrument being checked is set on the same scale marker, then from the experiment, we will not obtain the data required to solve the problem. In this case, according to (11.1),

$$\zeta_i' = A_c - y_i', \tag{11.6}$$

and $A_c = \text{const}$. Therefore,

$$V[\zeta_i'] = V[y_i']$$

and the estimate obtained for the variance in accordance with this relation depends on both the random error of the instrument being checked and the random error of the working standard. To solve this problem, it is necessary to have an estimate of the standard deviation of the working standard $S(y)$.

Let the same quantity having the true value A be provided for both instruments. Then $A = x - \zeta = y - \gamma$, and therefore,

$$\zeta - \gamma = x - y.$$

From here, based on relations (11.2) and (11.4), we obtain

$$\zeta_i' = z_i,$$

where $z_i = x_i - y_i$.

Therefore, for the indications obtained in accordance with formula (11.6), we have

$$V[\zeta_i'] = V[y_i'] = V[x_i] + V[y_i].$$

Correspondingly,

$$S^2(y') = S^2(x) + S^2(y).$$

Knowing $S^2(y)$, and having $S^2(y')$ based on the experimental data, we find

$$S^2(x) = S^2(y') - S^2(y). \tag{11.7}$$

Relation (11.7) indicates that it is desirable to know the standard deviation for the working standard. However, measuring instruments, for which the standard deviations and, for example, the limits of error of the corrections are known, are difficult to use for checking measuring instruments that are to be made for single measurements. In such cases, for the working standard, it would be helpful to know, in addition to the characteristics mentioned above, the limits of total error. For this reason, it is not necessary to determine separately both the components and the limits of total error in each check. Obviously, the relation between the total error and its components for each type of measuring instrument is the same. For this reason, if such a relation is established in the course of the investigations, performed, for example, while certifying working standards, then in the future, when measuring instruments are routinely checked, it could be sufficient to determine only part of the errors under study.

11.3. Rejects of Verification and Ways to Reduce Their Number

Because of the errors of working standards, some fraction of serviceable instruments, i.e., instruments whose errors do not exceed the limits established for them, is rejected in a verification—false rejection—and some fraction of instruments that are in reality unserviceable are accepted—false retention. This situation is typical for monitoring production quality, and just as with quality control, here a probabilistic analysis of the procedure is interesting.

Suppose that the same quantity is measured simultaneously by a working standard and the instrument being checked. As pointed out above, for analysis, such a scheme is simpler than other schemes, but this is not reflected in the generality of the obtained results. In accordance with the conditions of the experiment, we have

$$A = x - \zeta = y - \gamma,$$

where x and y are the indications of the checked and working standard and ζ and γ are the errors of the checked and working standard. From here,

$$z = x - y = \zeta - \gamma. \tag{11.8}$$

We are required to show that $|\zeta| \leq \Delta$, where Δ is the limit of permissible error of the checked instrument. From the experimental data, we can find z; we shall assume that if $|z| \leq \Delta$, then the checked instrument is serviceable, and if $|z| > \Delta$, then it is not serviceable.

To perform probabilistic analysis in this way, it is necessary to know the probability distribution for the errors of the checked and standard instruments. Let us suppose we know them.

The probability of a false rejection is

$$p_1 = P\{|\zeta - \gamma| > \Delta|_{|\zeta| \leq \Delta}\},$$

and the probability of a false retention is

$$p_2 = P\{|\zeta - \gamma| \le \Delta|_{|\zeta|>\Delta}\}.$$

A false rejection is obtained for $|\zeta| \le \Delta$ when $|\zeta - \gamma| > \Delta$, i.e.,

$$\zeta - \gamma > \Delta, \qquad \zeta - \gamma < -\Delta,$$

or

$$\gamma < \zeta - \Delta, \qquad \gamma > \zeta + \Delta.$$

If the probability distribution of the errors of the checked and working standard are $f(\zeta)$ and $\varphi(\gamma)$, respectively, then

$$p_1 = \int_{-\Delta}^{\Delta} f(\zeta)\left(\int_{-\infty}^{\zeta-\Delta} \varphi(\gamma)\,d\gamma + \int_{\zeta+\Delta}^{+\infty} \varphi(\gamma)\,d\gamma\right)d\zeta.$$

A false retention is possible when $|\zeta| > \Delta$, i.e., when $\zeta > +\Delta$ and $\zeta < -\Delta$. In this case, $|\zeta - \gamma| \le \Delta$, i.e.,

$$\zeta - \gamma \le \Delta, \qquad \zeta - \gamma \ge -\Delta.$$

From here, $\zeta - \Delta \le \gamma \le \zeta + \Delta$. Therefore,

$$p_2 = \int_{-\infty}^{-\Delta} f(\zeta)\left(\int_{\zeta-\Delta}^{\zeta+\Delta} \varphi(\gamma)\,d\gamma\right)d\zeta + \int_{\Delta}^{+\infty} f(\zeta)\left(\int_{\zeta-\Delta}^{\zeta+\Delta} \varphi(\gamma)\,d\gamma\right)d\zeta.$$

Thus, if the probability densities and their parameters are known, then the corresponding values of p_1 and p_2 can be calculated, and their dependence on the relations between the limits of the permissible errors of the standard and the checked instruments can be traced.

If, in addition, cost considerations are added, then, one would think, the problem of choosing this relation can be solved uniquely. In reality, when the accuracy of working standards is increased, the cost of the check increases also. A rejection also has a certain cost. Therefore, by varying the limits of error of working standards, it is possible to find the minimum losses and this variant is regarded as optimal.

The mathematical relations for solving the problem can easily be derived. Unfortunately, however, in the general case, it is impossible to estimate the losses from the use of instruments whose errors exceed the established limits. In general, it is difficult to express in terms of money the often significant economic effect of increasing measurement accuracy. For this reason, it is only in exceptional cases that economic criteria can be used to justify the choice of the relation between the limits of permissible error of the working standard and the checked instruments.

In addition, as has already been pointed out above, the fundamental problem is to determine the probability distribution of the errors of the instruments. The results, presented in Chapter 2, of the statistical analysis of data from a check of a series of instruments showed that the sample data are unstable. Therefore, the distribution function of the instrument errors cannot be found from these data. However, there are no other data; they simply cannot be obtained anywhere.

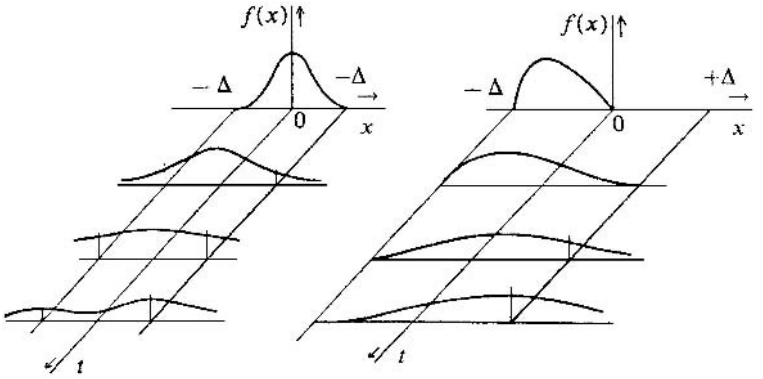


FIGURE 11.1. Examples of possible changes in the probability densities of the errors of measuring instruments in time.

Moreover, the fact that the sampling data are unstable could mean that the distribution functions of the errors of the instruments change in time. There are definite reasons for this supposition.

Suppose that the errors of a set of measuring instruments of some type, at the moment they are manufactured, have a truncated normal distribution with zero mean. For measures (measuring resistors, shunts, weights, etc.), a too large error of the same sign results in certain rejection. This is taken into account when manufacturing measures and, as a result, the distribution of the intrinsic errors of measures is usually unsymmetric. For example, if when a weight is manufactured its mass is found to be even slightly less than the nominal mass, then the weight is discarded. Figure 11.1 shows both variants of the distributions.

Instrument errors change in the course of use. Usually the errors only increase. In those cases in which, as in the case of weights, the direction of the change of the errors is known beforehand, and this is taken into account by the rules of manufacturing, the errors can at first be reduced, but then they will still increase. Correspondingly, changes in the instrument errors deform the distribution functions of the errors. This process, however, does not occur only spontaneously. At the time of routine checks, measuring instruments whose errors exceed the established limits are discarded. Figure 11.1 shows the approximate general picture of the changes occurring in the probability distribution in time. The process ultimately terminates when the measuring instruments under study no longer exist: either their errors exceed the established limits or they are no longer serviceable for other reasons.

The actual picture is still more complicated, because the stock of measuring instruments of each type can also change periodically as a result of the appearance of new measuring instruments.

It should be noted that the properties of measuring instruments, such as influence functions and influence factors of different influence quantities, as a rule, do not change with time, and for this reason, there is a much better foundation for describing them with the help of a probabilistic model.

The foregoing considerations show that the probabilities of rejection in a check must be calculated carefully.

It is nonetheless of interest to analyze purely abstract situations and to examine a series of models to cast light on the general laws. Such an analysis has been performed by several authors.

E. F. Dolinskiĭ obtained the following results under the assumption that the errors of standard and checked instruments have normal distributions [23]:

(i) Rejection in a check depends primarily on the relation between the limit of permissible error of the checked instruments and the standard deviation of the errors of these instruments; as the standard deviation decreases, the number of instruments rejected in a check decreases.

(ii) The relation between the errors of the checked and standard instruments (between the permissible limits of their errors or between the standard deviations of these errors) affects the number of rejections in a check much less than do the properties of the distribution of the errors of the checked instruments.

Digressing from the statistical instability of the distribution of errors of checked instruments, it should be noted that this approach toward describing the quality of a check has a fundamental drawback. Assume that we have the distribution function of the errors of the checked instrument; i.e., we know the error of the entire collection of instruments. For each specific batch of instruments, however, the number rejected will depend on how many of the instruments in the batch are unserviceable. Therefore, the probability of rejection is not a good indicator of the checking effectiveness, because an indicator of the checking effectiveness should not depend on whether the number of bad instruments being checked is large or small.

If batches of instruments were checked, then one could talk about distribution functions for each batch and correspondingly about rejections for each batch and the average number of rejections. But instruments are checked separately or in small batches (several instruments at a time), so that in this approach, one cannot talk about distribution functions.

This contradiction can be resolved by resorting to some conditionally chosen distributions. To obtain an estimate of the highest probability of false retention, one can, for example, take the distribution of errors of bad instruments, i.e., instruments whose error exceeds the permissible limits. As in practice bad instruments are not the only instruments that are checked, it is clear that in reality the probability of false retention will always be less than the value obtained by this method.

Analogously, to estimate the upper limit of the probability of false rejection, one can take some distribution of errors that do not exceed the limits of permissible errors.

When the problem is solved in this manner, the problem of choosing the form of the worst distributions arises. This question cannot be solved objectively, and many variants can be proposed. Examples of such test distributions are as follows: for "bad" instruments—the symmetric distribution, constructed from positive and negative branches of the normal distribution, separated by 2Δ, with standard deviation $\sigma = \Delta/\sqrt{3}$; for "good" instruments—the uniform distribution with the

limits $\pm\Delta$. The distribution of the errors of working standards, out of caution, should be taken as uniform with permissible limits of $\pm\Delta_s$.

However, this supposition cannot be justified, and for this reason, one cannot insist on it.

Thus, based on the widely used checking method examined above, it is impossible to find a sufficiently convincing method for choosing in a well-founded manner the relation between the errors of the standard and the checked instruments. For this reason, in practice, this question is solved by a volitional method by standardizing the critical relation between the limits of permissible errors. Thus, in electric measuring techniques, it is assumed that the error of working standards must not be more than one fifth the limit of permissible error of the checked instruments. In standards based on electronic instruments, the accuracy requirements for working standards are not as stringent: This ratio is usually equal to 3. Other ratios (for example, 1:10) are rarely encountered.

The ratios 1:10 and 1:5 usually are not objectionable, but it is often technically difficult to realize them. The ratio of 1:3, however, is always criticized as being inadequate.

Is it possible to choose a different rule for singling out unserviceable instruments to avoid the difficulties connected with justifying on the basis of probability the choice of the ratio between the errors of standard and checked instruments? This problem can, in principle, be solved as follows.

Based on the definition, a serviceable instrument is an instrument for which $|x - A| \leq \Delta$ and an instrument is unserviceable if $|x - A| > \Delta$.

Analogous inequalities are also valid for a working standard: $|y - A| \leq \Delta_s$, if the instrument is serviceable and $|y - A| > \Delta_s$ if it is not serviceable.

For $x > A$, for a serviceable instrument, $x - A \leq \Delta$. But $y - \Delta_s \leq A \leq y + \Delta_s$. For this reason, replacing A by $y - \Delta_s$, we obtain for a serviceable instrument,

$$x - y \leq \Delta - \Delta_s. \tag{11.9}$$

Analogously, for $x < A$, for a serviceable instrument,

$$x - y \geq -(\Delta - \Delta_s). \tag{11.10}$$

Repeating the calculations for an unserviceable instrument, it is not difficult to derive the corresponding inequalities:

$$x - y > \Delta + \Delta_s, \tag{11.11}$$

$$x - y < -(\Delta + \Delta_s). \tag{11.12}$$

Figure 11.2 graphically depicts the foregoing relations. Let the scale of the checked instrument be the abscissa axis. On the ordinate axis, we mark the points $+\Delta$ and $-\Delta$, and around each of these points, we mark the points displaced from them by $+\Delta_s$ and $-\Delta_s$. If Δ and Δ_s remain the same for the entire scale of the instrument, then we draw from the marked points on the ordinate axis straight lines parallel to the abscissa axis.

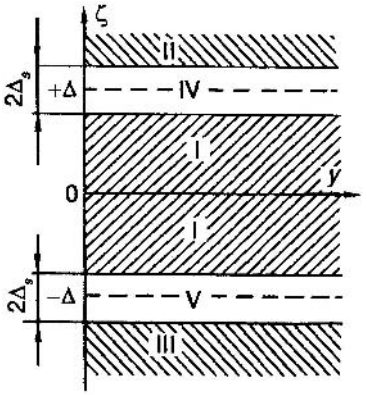


FIGURE 11.2. Zones of definite serviceability (I), definite rejection (II and III), and uncertainty (IV and V) when verification of measuring instruments with the limit Δ of permissible error based on a working standard whose limit of permissible error is Δ_s.

Region I corresponds to inequalities (11.9) and (11.10). The instrument for which the differences $x - y$ fall within this region are definitely serviceable irrespective of the ratio of the errors of the standard and checked instruments.

Inequalities (11.11) and (11.12) correspond to regions II and III. The instruments for which the differences $x - y$ fall within the regions II or III are definitely unserviceable.

Some checked instruments can have errors such that

$$\Delta - \Delta_s < |x - y| < \Delta + \Delta_s.$$

These errors correspond to regions IV and V in Fig. 11.2. Such instruments essentially cannot be either rejected or judged to be serviceable, because in reality, they include both serviceable and unserviceable instruments. If they are assumed to be serviceable, then the user will get some unserviceable instruments. This rejection can harm the user. If, however, all such doubtful instruments are rejected, then in reality, some serviceable instruments will be rejected. For instruments that are doubtful when they are manufactured or when they are checked after servicing, it is best that they be judged unserviceable. This tactic is helpful for anyone using instruments and forces the manufacturers to use more accurate standard instruments, but this is not always possible in regular checks.

In those cases in which the percentage of doubtful instruments is significant and the instruments are expensive and difficult to fix, it is best to check them again. Here several variants are possible. One variant is to recheck the doubtful instruments with the help of more accurate working standards.

In those cases in which more accurate instruments cannot be used for one reason or another, the check can also be made with the help of other samples of working standards that are rated at the same accuracy as those used in the initial check. As different working standards have somewhat different errors, the results of comparing the checked instruments with them will be somewhat different. As a results, some doubtful instruments will be judged absolutely serviceable and some will be confidently rejected.

The best method is to increase the accuracy of the working standard. However, the question then arises as to how much the accuracy of the standard instruments should be increased.

If there are no technical limitations, then the accuracy of the working standard can be increased until the instrument being checked can be judged as being either serviceable or unserviceable. If, however, the limit of permissible error of the standard instrument becomes 5–10 times less than the limit of permissible error of the checked instrument, then the accuracy of the working standard should not be increased further: The errors of instruments are usually not stable enough to be estimated with high accuracy.

For a five- to ten-fold difference in the errors, the error of working standard can usually always be neglected. This practice can be justified because in this case, the probability of rejection is always low (because the zone of uncertainty is narrow, the percentage of instruments incorrectly judged to be serviceable or rejected is always low) and because only those instruments whose errors do not differ much from the limit established for them can be incorrectly judged as serviceable.

Rejection of instruments in checks is eliminated completely if instead of verification the instruments are recalibrated. The accuracy of the new calibration characteristic can be almost equal to the accuracy of the working standard, which makes this method extremely attractive.

The drawback of this method is that the new calibration characteristic is most often constructed with the help of a table of corrections to the old calibration characteristic, which is not convenient for using the instrument. More importantly, in this method, the stability of the instrument is concealed. The possessor of the instrument must accumulate calibration results and analyze them. Analysis makes it possible to judge the stability of an instrument, how often the instrument should be recalibrated, and thereby the desirability of continued use of the instrument, if it must be calibrated often.

11.4. Calculation of a Necessary Number of Standards

Calibration, testing, and verification are metrological operations, with whose help the dimensions of decreed units of physical quantities are transferred to all measuring instruments. The units, however, are reproduced with the help of reference standards.

Reference standards are not created for all units. The circumstances under which reference standards need to be created deserve discussion.

First, we note that reference standards are always necessary for the units of the basic quantities. The question of whether it is desirable to create reference standards pertains only to the units of derived quantities. Derived physical quantities include quantities measured only by indirect methods, for example, an area. It is clear that reference standards are not required for the units of such quantities.

But reference standards are also not always required for the units of quantities measured by direct methods. A reference standard is not necessary if the

instrument used to measure a given quantity can be checked with adequate accuracy and efficiency with the help of working standards for other quantities. For example, to check tachometers, it is sufficient to have a device for rotating the shaft of the tachometer and a stroboscopic timer; a standard is not required in this case.

When a reference standard is created, the reproduction of the unit is centralized. On the one hand, this result complicates the measures that must be taken to ensure uniformity of measuring instruments, because some standard measuring instruments (at a minimum, the working reference standards) must be compared with the reference standard. On the other hand, it is found that usually complicated indirect measurements, with whose help the starting primary setups are certified when the reproduction of units is not centralized, are possible only for certifying one measuring instrument—the primary reference standard.

The question of whether a reference standard should be created is answered by comparing these contradictory factors. Thus, the solution of this question is based on technical and economic considerations, which is why it is difficult to solve the problem.

It should also be noted that in many cases, indirect measurements, which are required for reproducing the units of a derived quantity, do not provide the necessary accuracy. The creation of a reference standard, i.e., centralization of reproduction of the unit, makes it possible in this case to achieve greater uniformity of the measuring instrument than in the absence of a reference standard, because when the size of the unit reproduced with the help of the reference standard is transferred, the systematic error of the standard can be neglected. This circumstance is often exploited, although it is possible that the most accurate measuring instruments in one country can have a significant systematic error compared with the analogous measuring instruments in another country. Comparing reference standards from different countries makes it possible to avoid misunderstandings that can arise because of this.

The sizes of the units reproduced with the help of primary reference standards are transferred to the working standards with the help of a system of standards. The metrological coordination of standards, their relation with the working instruments, and the principles of the methods of comparison employed in Russia are customarily represented with the help of so-called checking or calibration schemes.

Physical standards are divided into ranks. The number of a rank indicates the number of steps included in transferring the size of a unit from the primary reference standard to a given working standard.

One of the most difficult questions arising in the construction of checking schemes is the question of how many ranks of standards should be provided. As the number of ranks increases, the error with which the size of a unit is transferred by the working measuring instrument increases. For this reason, to obtain high accuracy, the number of ranks of standards should be reduced to a minimum.

The higher the accuracy of standards, the more expensive they are. In addition, more accurate measurements are usually more difficult to perform. Increasing the number of ranks makes it possible to have less accurate standards together with

more accurate standards and makes the entire system of transferring the size of a unit more economical. For this reason, in the fields of measurement where there is a large margin in the accuracy of reference standards, the number of ranks of standards can be equal to the number of gradations of accuracy of the working measuring instruments.

Enlarging the stock of standards and increasing the number of ranks, i.e., the number of gradations of accuracy, make the work of calibration laboratories more difficult and involve a certain cost. At the same time, the operations of calibration of measuring instruments usually become more efficient. One would think that it is possible to find an economically optimal number of ranks of the checking scheme. This process, however, requires information about the dependence of the cost of the equipment and labor on the accuracy. This information is usually not available. For this reason, in practice, the optimal checking schemes cannot be determined.

Checking schemes are usually constructed when reference standards and working standards are partially already available, and it is only necessary to arrange them in a hierarchical order. In this case, it can be assumed that the number of working measuring instruments, the frequency with which they must be calibrated, and the permissible number of annual comparisons of the most accurate measuring instruments with the reference standard are known. In addition, the time required to calibrate one sample of each type of measuring instrument and working standard, or the limiting number of calibrations permitted by the reference standards within a prescribed period of time, can be estimated. This information makes it possible to find the minimum necessary number of ranks of the checking scheme. The problem can be solved by the method of successive approximations.

In the general case, checking schemes can be assumed to have the structure shown in Fig. 11.3. We shall first study the case when the checking scheme has only one vertical (i.e., it does not have the branches 2 and 3 shown in Fig. 11.3).

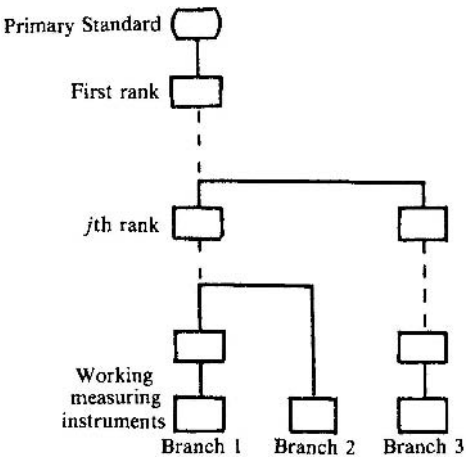


FIGURE 11.3. Typical structure of checking schemes.

If the jth rank has N_j standards, then the maximum number of standards in the rank $(j+1)$ will be

$$N_{j+1} = N_j \frac{\eta_j T_{j+1}}{t_{j+1}}, \tag{11.13}$$

where η_j is the utilization factor of the standards of rank j, T_{j+1} is a time equal to the time interval between calibrations of the measuring instrument of rank $j+1$, and t_{j+1} is the time necessary to calibrate one measuring instrument in the rank $(j+1)$.

When calculating the coefficients η_j, the utilization time of the measuring instrument must be compared with the calendar time, and the losses of working time must be taken into account. For example, if some apparatus is used eight hours per day and one hour is required for preparation and termination, and preventative maintenance, servicing, and checking reduce the working time by 10%, then

$$\eta = \frac{8-1}{24} \times 0.9 = 0.2625.$$

Transferring, in accordance with the checking scheme, from the primary reference standard ($j = 0$, $N_0 = 1$) to the working measuring instrument, we determine the maximum number of standards of each rank and then the number of working measuring instruments N_m guaranteed by calibration:

$$N_m = N_0 N_1 \dots N_{m-1} = \prod_{j=0}^{m-1} \eta_j \frac{T_{j+1}}{t_{j+1}}, \tag{11.14}$$

where m is the total number of steps in transferring the size of a unit from the reference standard to the working measuring instrument, inclusively.

To solve the problem, we first choose some number or ranks $j_0 = m-1$. In principle, it is possible to start with the minimum number of ranks $j_0 = 1$. For given j_0, we find $N_m^{(0)}$. If $N_m^{(0)}$ is less than the number of working measuring instruments that need to be calibrated, then either the number of ranks or η_j—the utilization factor of the standards—and the efficiency of the calibration operations must be increased; i.e., the time $t_j (j = 0, 1)$ must be reduced. If the problem cannot be solved in this manner, then the number of ranks must be increased.

We calculate $N_m^{(1)}$ for the new number of ranks. As soon as the value of $N_m^{(i)}$ obtained is greater than the number of working measuring instruments that must be calibrated, the number of ranks can be regarded as sufficient.

For a more general checking scheme (Fig. 11.3), the possibilities of each branch of the checking scheme must be checked. It is best to start the calculation with the branch adjoining a standard of high rank, and to perform the calculation in the reverse order as compared with the method examined above; i.e., it is best to start from a fixed number of working measuring instruments in each class that must be checked along a given branch. We shall use formula (11.13) to calculate the number of working standards N_j for the lth branch that is required to calibrate N_{j+1} measuring instruments. For this reason, it is useful to write formula (11.13)

in the form

$$N_{jl} = N_{j+1,l} \frac{t_{j+1,l}}{\eta_{jl} T_{j+1,l}}.$$

The calculation proceeds up to standards servicing several branches of the checking scheme. In the process, the number of these standards that is required for the lth branch is determined. Then, for the rest of the checking scheme, $N_j - N_{jl}$ standards remain of the given rank. The number of measuring instruments that must be calibrated according to the vertical of the scheme (branch 1 in Fig. 11.3) is found using formula (11.14), taking into account the branching losses:

$$N_m^{(1)} = N_0(N_1 - N_{12}) \dots (N_j - N_{jl}) \dots N_{m-1}.$$

The number $N_m^{(1)}$ obtained must be greater than the prescribed number of measuring instruments that must be calibrated along this branch. This condition is necessary for all branches. If it is not satisfied, then the number of ranks must be increased. As the number of ranks increases, the efficiency of the checking network, represented by the checking scheme, increases rapidly. The checking schemes employed have the maximum number five of ranks of standards, even for the most developed fields of measurement.

The relations presented above pertained to the simplest case, when at each step of transfer of the size of the unit, the period of time between calibrations and the calibration time were the same for all measuring instruments. In reality, these time intervals can be different for different measuring instruments. Taking this into account makes the calculations more complicated, but it does not change their essential features.

It is necessary to transfer from different time intervals between calibrations to one conditionally standard T_{cs} time interval and to find the number of measuring instruments of each type T_k^{cs} that must be checked within this period. This process is done with the help of the obvious formula

$$N_k^{\text{cs}} = N_k \frac{T_{\text{cs}}}{T_k}.$$

Next, it is necessary to find the average time t_j^{av} required to check one measuring instrument for each step of the checking scheme:

$$t_j^{\text{av}} = \frac{\sum_{k=1}^{n} t_k N_k^{\text{cs}}}{\sum_{k=1}^{n} N_k^{\text{cs}}}. \tag{11.15}$$

Here n is the number of different types of measuring instruments at the jth step of the checking scheme.

We shall give a numerical example. Suppose it is required to organize a calibration of instruments of types A and B and the following data are given:

(1) *Instruments of type* A: $N_A = 3 \times 10^4$; the time interval between calibrations $T_{A1} = 1$ yr for $N_{A1} = 2.5 \times 10^4$ and $T_{A2} = 0.5$ yr for $N_{A2} = 5 \times 10^3$; the calibration time $t_A = 5$ h.
(2) *Instruments of type* B: $N_B = 10^5$; $T_B = 1$ yr; the calibration time $t_B = 2$ h.
(3) *Primary reference standard*: Four comparisons per year are permitted; the frequency of the calibration of the most accurate measuring instruments, which can be working standards of rank 1, is 2 yr; i.e., $T_1 = 2$ yr; for them, $\eta_1 = 0.25$. For measuring instruments that can be working standards of rank 2, $T_2 = 2$ yr, $t_2 = 40$ h, and $\eta_2 = 0.25$.

The possible number of first-rank standards is

$$N_1 = N_0 f T_1 = 8,$$

because $N_0 = 1$, $f = 4$ is the maximum number of comparisons with a reference standard per year, and $T_1 = 2$.

It is obvious that eight standards are not enough to check 130,000 working instruments. We shall check to see if three ranks of standards are sufficient for this.

As the time interval between checks is different for different instruments, we introduce the conditionally standard time interval between checks $T_{cs} = 1$ yr and find the number of instruments that must be checked within this time period. Conversion is necessary only for instruments of type A with $T_{A2} = 0.5$yr:

$$N_{A2}^{cs} = N_{A2} \frac{T_{cs}}{T_{A2}} = 5 \times 10^3 \times \frac{1}{0.5} = 10 \times 10^3.$$

Therefore,

$$\sum_{k=A,B} N_k^{cs} = N_{AB} = N_{A1} + N_{A2}^{cs} + N_B = 135 \times 10^3$$

instruments must be calibrated within the time T_{cs}.

Different amounts of time are required to check instruments of types A and B. We shall find the average checking time t_w^{av} of these working instruments. In accordance with formula (11.15),

$$t_w^{av} = \frac{(N_{A1} + N_{A2}^{cs}) t_A + N_B t_B}{N_{AB}} = \frac{35 \times 10^3 \times 5 + 100 \times 10^3 \times 2}{135 \times 10^3} = 2.78 \text{ h}.$$

Now, using formula (11.13), we shall find the required number of second-rank standards:

$$N_2^{(1)} = \frac{N_{AB} t_w^{av}}{\eta_2 T_{cs}} = \frac{135 \times 10^3 \times 2.78}{0.25 \times 6 \times 10^3} = 250.$$

Here it was assumed that $T_{cs} = 250 \times 24 = 6 \times 10^3$ h.

It remains to verify that all working standards of the second rank can be checked. For this reason, we calculate from formula (11.13) the maximum possible number

of standards of this rank:

$$N_2 = N_1 \frac{\eta_1 T_2}{t_2} = 8 \times \frac{0.25 \times 2 \times 6 \times 10^3}{40} = 600.$$

As $N_2 > N_2^{(1)}$, in this case, two ranks of working standards are sufficient.

With the help of calculations similar to those presented in the foregoing example, it is possible to choose in a well-founded manner the structure of a checking scheme and to estimate the required number of working standards of each rank.

In calculating the checking scheme we did not take into account explicitly the accuracy of the measuring instruments. However, the contemplated scheme must be metrologically realizable, which means that the discrepancy between the accuracy of the primary reference standard and the accuracy of the working measuring instruments must make it possible to insert between them the required number of ranks of working standards. The problem facing instrument makers and metrologists is to provide the combination of accuracy and efficiency required to implement the checking scheme by designing working standards and reference standards.

Checking schemes usually have extra calibration possibilities, which makes it possible to distribute reference and working standards to limit their transport, to maximize the efficiency of calibration laboratories, and to take into account other practical considerations.

12
Conclusion

12.1. Measurement Data Processing: Past, Present, and Future

Once upon a time physicists believed that normal distribution was a mathematically proven rule for random phenomena, whereas mathematicians considered it to be a natural law discovered by physicists. This belief was reflected in the measurement data processing. Namely, the calculated variance of a sample was taken for the variance of the corresponding normal distribution. Then the empirical histogram was changed to be equal to the histogram from that normal distribution. This procedure was called "smoothing out the frequencies." However, in the beginning of last century, researchers realized that experimental data should not be distorted and this procedure was abandoned.

During the next century, mathematical statistics was developing rapidly and became widely used in various fields of science and industry. This same development also happened to measurement data processing, which became dominated by mathematical statistical methods. As a result, the science of measurement data processing was limited to direct multiple measurements, and stayed clear from systematic errors. This state of affairs can be clearly seen by examining books of that period, for example, '*Data Analysis for Scientists and Engineers*' by S. Meyer, '*Data Reduction and Error Analysis for Physical Sciences*' by Ph. Bevington and D. Robinson, and many others. Even books addressed specifically to practitioners, such as the recently published book by I. Gertsbakh [26] and excellent books by J. Mandel [36, 37], remained within the above confines. Because this purely mathematical theory found practical applications, even if in a restricted case of random errors in multiple measurements, this theory obtained the status of the classical theory of measurement data processing. However, it was not sufficient in many practical situations. In particular, every practitioner knew that in addition to random errors, there are systematic errors, and the overall uncertainty of the measurement result combines both of these components. But the classical theory ignored this fact and, furthermore, considered it incorrect to combine these two components. There were other practical problems ignored by the classical theory.

As a result, those who encountered these problems in their practice resorted to ad hoc and often incorrect methods. For example, in the case of single measurements, the measurement errors were often equated to the intrinsic errors of the measuring devices used, which is wrong. To account for systematic errors in multiple measurements, people often simply added them to the random errors, which overestimated the inaccuracy of the result.

Starting from 1970s, in various publications, I was trying to solve the practical problems that were not served by the classical theory. One of the first publications in this area was the article [41], which led in 1976 to a standard for methods of measurement data processing including combining random and systematic errors [5]. An American standard containing a solution to this problem appeared in 1985 [3], followed by a British Recommendation in 1986 [6]. Finally, in 1995, the International Organization for Standardization published the Guide to the Expression of Uncertainty in Measurement.[1]

The problem of combining random and systematic errors is just one example of the limitations of the classical theory. Toward the end of last century, these limitations became obvious, and a new theory started to take shape. This theory, which we can call the physical theory of measurement data processing, does not obviate but subsumes the classical theory, and augments it with methods for processing single measurements, accounting for errors of measuring devices, combining random and systematic errors, and other practical problems. The physical theory also considers the foundational issues of measurements. This book offers systematic treatment of the physical theory and in this way defines this new discipline.

At the same time, this book obviously does not exhaust this subject, and several problems still await their solutions. We list some of these gaps below.

- The theory of single measurements requires further development, especially in regard to accounting for the errors of measuring instruments. A complicating factor in this problem is a large variety of measuring instrument types for which suitable techniques must be developed.
- Although the diversity of measuring instruments prohibits the development of the general theory of their design, it is possible and necessary to develop a general theory of accuracy of measuring instruments. The accuracy is the common aspect that unites these devices. This book takes an initial step toward such a theory, but much more work is required.
- A large and common class of measurements involving recording instruments (such as analog or digital automatic plotters, XY-recorders, etc.) came to be known as dynamic measurements [28,49]. There are many open problems in dynamic measurements; among them is an attractive problem to find the form and parameters of an input signal having the recorded output signal and knowing the dynamic properties of the recorder. Modern computers make solving this problem feasible.

[1] Although this Guide provides a method for combining systematic and random errors, it has its drawbacks, which are discussed in detail in Section 12.3.

- The errors and uncertainty of measurements are always estimated in an indirect way, and the calculations include some assumptions. However, the correctness of these assumptions, and the validity of the resulting estimates, has never been experimentally checked. Filling this gap is therefore extremely important. A general approach to this problem would involve measuring the same measurand in parallel by different methods, with one method being around ten times more accurate than the other, and then comparing the measurement results and the calculated uncertainties.

Although this list of problems is subjective and incomplete, it suffices to show that the physical theory of measurement data processing is a live discipline still under development.

12.2. Remarks on the "International Vocabulary of Basic and General Terms in Metrology"

At the end of last century, the International Organization for Standardization published an important document, the "International Vocabulary of Basic and General Terms in Metrology" (VIM) [2]. This document, prepared under the leadership of BIPM, will be studied and used throughout the world. Two terms defined in the Vocabulary, *error* and *uncertainty*, have a particular bearing on the present book, and it is appropriate to discuss them here.

The definition of error (VIM, Section 3.10) says that it is the "result of measurement minus a true value of the measurand." There is also a note: "Since a true value cannot be determined, in practice a conventional true value is used." Unfortunately, the above description cannot be considered a definition because it does not explain the meaning of the term, but it attempts to provide an algorithm for its calculation. As a matter of fact, the algorithm is unrealistic: In addition to the true value being unknown, its replacement by the conventional true value is impossible because no measurement has such a value. Indeed, the conventional true value can be attributed to a reference standard and used in calibration procedures; it allows one to estimate the concrete value of the error of the measuring instrument under calibration, but not the error of a measurement.

I consider "error" to be properly defined as *a deviation of the result of measurement from the true value of the measurand*. Then the note about the conventional true value should be replaced by the following: Because the true value of a measurand is always unknown, the accuracy of a measurement is characterized by the *limits of error* that are calculated in an indirect way. If these calculations involve a probabilistic model, then the term *uncertainty of a measurement* is used in place of the limits of error. The latter definition of the term "error" is given in [4] and is similar to the definition in [7].

The definition of uncertainty from VIM, Section 3.9 is provided with a note saying that uncertainty "may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence."

This note creates ambiguity that is unacceptable in scientific terminology. Indeed, what is the uncertainty, a standard deviation or a confidence interval?

This ambiguity was not accidental; it arose because it was not known how to construct confidence intervals for many cases of indirect measurements. The methods of reduction and transformation described in this book allows the construction of the confidence intervals in all cases of indirect measurements, and therefore, the ambiguity can be eliminated. Thus, "uncertainty" would be properly defined as *an interval within which a true value of a measurand lies with the given probability.* This definition can be accompanied with a note that *uncertainty is defined with its limits and corresponding confidence probability; the limits of uncertainty are read out from the result of a measurement.*

12.3. Drawbacks of the "Guide to the Expression of Uncertainty in Measurement"

Another important document published by ISO and prepared under the leadership of BIPM is the "Guide to the Expression of Uncertainty in Measurement" (usually referred to as simply "the Guide") [1]. The goal of the Guide is unification of methods of measurement uncertainty estimation and its presentation. The need for such a document is obvious.

In accordance with the Guide, the uncertainty of a measurement must reflect both random and systematic errors in the measurement. Both of these components are characterized by estimates of their variances. The sum of them gives the combined standard deviation. The measurement uncertainty is obtained by multiplying the combined standard deviation by a coverage factor (recommended to be 2 or 3). The coverage factors 2 or 3 correspond to confidence probability 0.95 or 0.99, respectively. The above procedure is correct and in essence is the same as given in this book. The difference, besides terminology, is that the Guide chooses the coverage factors arbitrarily, whereas this book provides a way to calculate them.

Let us turn now to the shortcomings of the Guide. The basic philosophy of the Guide is that the concept of the true value of a measurand is not needed because it is the same as the value of the measurand (See Annexes B.2.3, D.3.5, etc.). But this statement is in contradiction with VIM. In accordance to VIM, Section 1.18, the value of a measurand is the denominate number, which is the product of the unit of measurement and a number. This value can be found as the result of the measurement, whereas the true value is a purely theoretical concept and cannot be obtained (see VIM, Section 1.19). Thus, the meanings of the true value of a measurand and the value of a measurand are different and the idea to replace the first one by the second is incorrect. Furthermore, although the true value cannot be expressed in numbers, it is necessary to in practice for defining and verifying the model of the object under study. We refer the reader back to Section 1.3, where we discussed this issue in detail.

The elimination of the term *true value* in the Guide was motivated by the desire to eliminate the term '*error*'. The idea to drop "error" and always replace it with

"uncertainty" is not new and has its roots in the article [18]. However, as we discussed in Section 3.1, this idea is misguided because it is based on the incorrect use of "error" in [18]. Note that the unconditional replacement of "error" with "uncertainty" also contradicts the second edition of the VIM, which defines both terms in a similar manner to this book.

Another confusion in the Guide has to do with uncertainty. The Guide introduces the terms '*standard uncertainty*' and combined uncertainty. However, these terms are redundant: the "standard uncertainty" is simply the standard deviation, and the "combined uncertainty" is the combined standard deviation. There is no reason to rename the existing terms.

Further, the Guide introduces two new terms, *type A and type B evaluation of uncertainty* (Chapter "Definitions," Sections 2.3.2 and 2.3.3). In accordance with the definitions of these terms, they characterize the methods of uncertainty estimation. But they are used in the Guide to denote *components* of expanded uncertainty. (Indeed, the Guide describes how to combine uncertainty of type A and type B, and obviously methods cannot be combined.) Beyond that, terms type A and type B uncertainty are not expressive. The meaning of these terms (as components of uncertainty) is clearer and more appropriately conveyed by the terms *systematic and random errors*, which are included in VIM. But as we mentioned earlier, the Guide specifically avoids the term "error" and its derivatives, and this forced it to create the artificial terms "type A" and "type B uncertainty." Note that unlike systematic and random errors, these terms are not included in VIM.

In summary, the above definitions in the Guide contradict the philosophy and terminology of VIM. Besides, the Guide does not properly reflect some modern results. For example, although it briefly mentions the method of reduction (as a second approach), it does not point out its main advantage, which is the elimination of the correlation coefficient from uncertainty calculations. Thus, the Guide must be revised.

Appendix

TABLE A.1. Values of the normalized Gaussian function $\Phi(z) = 1\sqrt{2\pi}\int_0^z e^{-y^2/2}dy$.

z	0	1	2	3	4	5	6	7	8	9
0.0	0.000 00	0.003 99	0.007 98	0.011 97	0.015 95	0.019 94	0.023 92	0.027 90	0.031 88	0.035 86
0.1	0.039 83	0.043 80	0.047 76	0.051 72	0.055 67	0.059 62	0.063 56	0.067 49	0.071 42	0.075 35
0.2	0.079 26	0.083 17	0.087 06	0.090 95	0.094 83	0.098 71	0.102 57	0.106 42	0.110 26	0.114 09
0.3	0.117 91	0.121 72	0.125 52	0.129 30	0.133 07	0.136 83	0.140 58	0.144 31	0.148 03	0.151 73
0.4	0.155 42	0.159 10	0.162 76	0.166 40	0.170 03	0.173 64	0.177 24	0.180 82	0.184 39	0.187 93
0.5	0.191 46	0.194 97	0.198 47	0.201 94	0.205 40	0.208 84	0.212 26	0.215 66	0.219 04	0.222 40
0.6	0.225 75	0.229 07	0.232 37	0.235 65	0.238 91	0.242 15	0.245 37	0.248 57	0.251 75	0.254 90
0.7	0.258 04	0.261 15	0.264 24	0.267 30	0.270 35	0.273 37	0.276 37	0.279 35	0.282 30	0.285 24
0.8	0.288 14	0.291 03	0.293 89	0.296 73	0.299 55	0.302 34	0.305 11	0.307 85	0.310 57	0.313 27
0.9	0.315 94	0.318 59	0.321 21	0.323 81	0.326 39	0.328 94	0.331 47	0.333 98	0.336 46	0.338 91
1.0	0.341 34	0.343 75	0.346 14	0.348 50	0.350 83	0.353 14	0.355 43	0.357 69	0.359 93	0.362 14
1.1	0.364 33	0.366 50	0.368 64	0.370 76	0.372 86	0.374 93	0.376 98	0.379 00	0.381 00	0.382 98
1.2	0.384 93	0.386 86	0.388 77	0.390 65	0.392 51	0.394 35	0.396 17	0.397 96	0.399 73	0.401 47
1.3	0.403 20	0.404 90	0.406 58	0.408 24	0.409 88	0.411 49	0.413 09	0.414 66	0.416 21	0.417 74
1.4	0.419 24	0.420 73	0.422 20	0.423 64	0.425 07	0.426 47	0.427 86	0.429 22	0.430 56	0.431 89
1.5	0.433 19	0.434 48	0.435 74	0.436 99	0.438 22	0.439 43	0.440 62	0.441 79	0.442 95	0.444 08
1.6	0.445 20	0.446 30	0.447 38	0.448 45	0.449 50	0.450 53	0.451 54	0.452 54	0.453 52	0.454 49
1.7	0.455 43	0.456 37	0.457 28	0.458 18	0.459 07	0.459 94	0.460 80	0.461 64	0.462 46	0.463 27
1.8	0.464 07	0.464 85	0.465 62	0.466 38	0.467 12	0.467 84	0.468 56	0.469 26	0.469 95	0.470 62
1.9	0.471 28	0.471 93	0.472 57	0.473 20	0.473 81	0.474 41	0.475 00	0.475 58	0.476 15	0.476 70
2.0	0.477 25	0.477 78	0.478 31	0.478 82	0.479 32	0.479 82	0.480 30	0.480 77	0.481 24	0.481 69
2.1	0.482 14	0.482 57	0.483 00	0.483 41	0.483 82	0.484 22	0.484 61	0.485 00	0.485 37	0.485 74
2.2	0.486 10	0.486 45	0.486 79	0.487 13	0.487 45	0.487 78	0.488 09	0.488 40	0.488 70	0.488 99
2.3	0.489 28	0.489 56	0.489 83	0.490 10	0.490 36	0.490 61	0.490 86	0.491 11	0.491 34	0.491 58
2.4	0.491 80	0.492 02	0.492 24	0.492 45	0.492 66	0.492 86	0.493 05	0.493 24	0.493 43	0.493 61
2.5	0.493 79	0.493 96	0.494 13	0.494 30	0.494 46	0.494 61	0.494 77	0.494 92	0.495 06	0.495 20
2.6	0.495 34	0.495 47	0.495 60	0.495 73	0.495 85	0.495 98	0.496 09	0.496 21	0.496 32	0.496 43
2.7	0.496 53	0.496 64	0.496 74	0.496 83	0.496 93	0.497 02	0.497 11	0.497 20	0.497 28	0.497 36
2.8	0.497 44	0.497 52	0.497 60	0.497 67	0.497 74	0.497 81	0.497 88	0.497 95	0.498 01	0.498 07
2.9	0.498 13	0.498 19	0.498 25	0.498 31	0.498 36	0.498 41	0.498 46	0.498 51	0.498 56	0.498 61

Note: The values of $\Phi(z)$ for $z = 3.0{-}4.5$ are as follows:

3.0	0.498 65	3.4	0.499 66	3.8	0.499 93
3.1	0.499 03	3.5	0.499 77	3.9	0.499 95
3.2	0.499 31	3.6	0.499 84	4.0	0.499 968
3.3	0.499 52	3.7	0.499 89	4.5	0.499 997

TABLE A.2. Percentile points of Student's distribution.

Number of degrees of freedom ν	Significance level $q = (1-\alpha) \times 100\ (\%)$		
	10	5	1
1	6.31	12.71	63.66
2	2.92	4.30	9.92
3	2.35	3.18	5.84
4	2.13	2.78	4.60
5	2.02	2.57	4.03
6	1.94	2.45	3.71
7	1.90	2.36	3.50
8	1.86	2.31	3.36
9	1.83	2.26	3.25
10	1.81	2.23	3.17
12	1.78	2.18	3.06
14	1.76	2.14	2.98
16	1.75	2.12	2.92
18	1.73	2.10	2.88
20	1.72	2.09	2.84
22	1.72	2.07	2.82
24	1.71	2.06	2.80
26	1.71	2.06	2.78
28	1.70	2.05	2.76
30	1.70	2.04	2.75
∞	1.64	1.96	2.58

TABLE A.3. Critical values of the distribution of $T_n = (x_n - \bar{x})/S$ or $T_1 = (\bar{x} - x_i)/S$ (with unilateral check).

Number of observations, n	Upper 0.5% significance level	Upper 1% significance level	Upper 5% significance level
3	1.155	1.155	1.153
4	1.496	1.492	1.463
5	1.764	1.749	1.672
6	1.973	1.944	1.822
7	2.139	2.097	1.938
8	2.274	2.221	2.032
9	2.387	2.323	2.110
10	2.482	2.410	2.176
11	2.564	2.485	2.234
12	2.636	2.550	2.285
13	2.699	2.607	2.331
14	2.755	2.659	2.371
15	2.806	2.705	2.409
16	2.852	2.747	2.443
17	2.894	2.785	2.475
18	2.932	2.821	2.504
19	2.968	2.854	2.532
20	3.001	2.884	2.557
21	3.031	2.912	2.580
22	3.060	2.939	2.603
23	3.087	2.963	2.624
24	3.112	2.987	2.644
25	3.135	3.009	2.663
26	3.157	3.029	2.681
27	3.178	3.049	2.698
28	3.199	3.068	2.714
29	3.218	3.085	2.730
30	3.236	3.103	2.745

TABLE A.4. Percentile points of the χ^2 distribution $P\{\chi^2 > \chi_q^2\}$.

Number of degrees of freedom v	Significance level q (%)									
	99	95	90	80	70	30	20	10	5	1
1	0.000 16	0.003 93	0.0158	0.0642	0.148	1.074	1.642	2.706	3.841	6.635
2	0.0201	0.103	0.211	0.446	0.713	2.408	3.219	4.605	5.991	9.210
3	0.115	0.352	0.584	1.005	1.424	3.665	4.642	6.251	7.815	11.345
4	0.297	0.711	1.064	1.649	2.195	4.878	5.989	7.779	9.488	13.277
5	0.554	1.145	1.610	2.343	3.000	6.064	7.289	9.236	11.070	15.086
6	0.872	1.635	2.204	3.070	3.828	7.231	8.558	10.645	12.592	16.812
7	1.239	2.167	2.833	3.822	4.671	8.383	9.803	12.017	14.067	18.475
8	1.646	2.733	3.490	4.594	5.527	9.524	11.030	13.362	15.507	20.090
9	2.088	3.325	4.168	5.380	6.393	10.656	12.242	14.684	16.919	21.666
10	2.558	3.940	4.865	6.179	7.267	11.781	13.442	15.987	18.307	23.209
11	3.053	4.575	5.578	6.989	8.148	12.899	14.631	17.275	19.675	24.725
12	3.571	5.226	6.304	7.807	9.034	14.011	15.812	18.549	21.026	26.217
13	4.107	5.892	7.042	8.634	9.926	15.119	16.985	19.812	22.362	27.688
14	4.660	6.571	7.790	9.467	10.821	16.222	18.151	21.064	23.685	29.141
15	5.229	7.261	8.547	10.307	11.721	17.322	19.311	22.307	24.996	30.578
16	5.812	7.962	9.312	11.152	12.624	18.418	20.465	23.542	26.296	32.000
17	6.408	8.672	10.085	12.002	13.531	19.511	21.615	24.769	27.587	33.409
18	7.015	9.390	10.865	12.857	14.440	20.601	22.760	25.989	28.869	34.805
19	7.633	10.117	11.651	13.716	15.352	21.689	23.900	27.204	30.144	36.191
20	8.260	10.851	12.443	14.578	16.266	22.775	25.038	28.412	31.410	37.566
21	8.897	11.591	13.240	15.445	17.182	23.858	26.171	29.615	32.671	38.932
22	9.542	12.338	14.041	16.314	18.101	24.939	27.301	30.813	33.924	40.289
23	10.196	13.091	14.848	17.187	19.021	26.018	28.429	32.007	35.172	41.638
24	10.856	13.848	15.659	18.062	19.943	27.096	29.553	33.196	36.415	42.980
25	11.524	14.611	16.473	18.940	20.867	28.172	30.675	34.382	37.652	44.314
26	12.198	15.379	17.292	19.820	21.792	29.246	31.795	35.563	38.885	45.642
27	12.879	16.151	18.114	20.703	22.719	30.319	32.912	36.741	40.113	46.963
28	13.565	16.928	18.939	21.588	23.647	31.391	34.027	37.916	41.337	48.278
29	14.256	17.708	19.768	22.475	24.577	32.461	35.139	39.087	42.557	49.588
30	14.953	18.493	20.599	23.364	25.508	33.530	36.250	40.256	43.773	50.892

TABLE A.5. Values of the upper 1% of points of the distribution $F_{0.01} = S_1^2/S_2^2$.

Number of degrees of freedom											
	ν_1										
ν_2	2	3	4	5	6	8	12	16	24	50	∞
2	99.00	99.17	99.25	99.30	99.33	99.36	99.42	99.44	99.46	99.48	99.50
3	30.81	29.46	28.71	28.24	27.91	27.49	27.05	26.83	26.60	26.35	26.12
4	18.00	16.69	15.98	15.52	15.21	14.80	14.37	14.15	13.93	13.69	13.46
5	13.27	12.06	11.39	10.97	10.67	10.29	9.89	9.68	9.47	9.24	9.02
6	10.92	9.78	9.15	8.75	8.47	8.10	7.72	7.52	7.31	7.09	6.88
7	9.55	8.45	7.85	7.46	7.19	6.84	6.47	6.27	6.07	5.85	5.65
8	8.65	7.59	7.01	6.63	6.37	6.03	5.67	5.48	5.28	5.06	4.86
9	8.02	6.99	6.42	6.06	5.80	5.47	5.11	4.92	4.73	4.51	4.31
10	7.56	6.55	5.99	5.64	5.39	5.06	4.71	4.52	4.33	4.12	3.91
11	7.20	6.22	5.67	5.32	5.07	4.74	4.40	4.21	4.02	3.80	3.60
12	6.93	5.95	5.41	5.06	4.82	4.50	4.16	3.98	3.78	3.56	3.36
13	6.70	5.74	5.20	4.86	4.62	4.30	3.96	3.78	3.59	3.37	3.16
14	6.51	5.56	5.03	4.69	4.46	4.14	3.80	3.62	3.43	3.21	3.00
15	6.36	5.42	4.89	4.56	4.32	4.00	3.67	3.48	3.29	3.07	2.87
16	6.23	5.29	4.77	4.44	4.20	3.89	3.55	3.37	3.18	2.96	2.75
17	6.11	5.18	4.67	4.34	4.10	3.79	3.45	3.27	3.08	2.86	2.65
18	6.01	5.09	4.58	4.25	4.01	3.71	3.37	3.20	3.00	2.79	2.57
19	5.93	5.01	4.50	4.17	3.94	3.63	3.30	3.12	2.92	2.70	2.49
20	5.85	4.94	4.43	4.10	3.87	3.56	3.23	3.05	2.86	2.63	2.42
21	5.78	4.87	4.37	4.04	3.81	3.51	3.17	2.99	2.80	2.58	2.36
22	5.72	4.82	4.31	3.99	3.76	3.45	3.12	2.94	2.75	2.53	2.31
23	5.66	4.76	4.26	3.94	3.71	3.41	3.07	2.89	2.70	2.48	2.26
24	5.61	4.72	4.22	3.90	3.67	3.36	3.03	2.85	2.66	2.44	2.21
25	5.57	4.68	4.18	3.86	3.63	3.32	2.99	2.81	2.62	2.40	2.17
26	5.53	4.64	4.14	3.82	3.59	3.29	2.96	2.78	2.58	2.36	2.13
27	5.49	4.60	4.11	3.78	3.56	3.26	2.93	2.74	2.55	2.33	2.10
28	5.45	4.57	4.07	3.75	3.53	3.23	2.90	2.71	2.52	2.30	2.06
29	5.42	4.54	4.04	3.73	3.50	3.20	2.87	2.68	2.49	2.27	2.03
30	5.39	4.51	4.02	3.70	3.47	3.17	2.84	2.66	2.47	2.24	2.01
35	5.27	4.40	3.91	3.59	3.37	3.07	2.74	2.56	2.37	2.13	1.90
40	5.18	4.31	3.83	3.51	3.29	2.99	2.66	2.48	2.29	2.05	1.80
45	5.11	4.25	3.77	3.45	3.23	2.94	2.61	2.43	2.23	1.99	1.75
50	5.06	4.20	3.72	3.41	3.19	2.89	2.56	2.38	2.18	1.94	1.68
60	4.98	4.13	3.65	3.34	3.12	2.82	2.50	2.32	2.12	1.87	1.60
70	4.92	4.07	3.60	3.29	3.07	2.78	2.45	2.28	2.07	1.82	1.53
80	4.88	4.04	3.56	3.26	3.04	2.74	2.42	2.24	2.03	1.78	1.49
90	4.85	4.01	3.53	3.23	3.01	2.72	2.39	2.21	2.00	1.75	1.45
100	4.82	3.98	3.51	3.21	2.99	2.69	2.37	2.19	1.98	1.73	1.43
125	4.78	3.94	3.47	3.17	2.95	2.66	2.33	2.15	1.94	1.69	1.37
∞	4.60	3.78	3.32	3.02	2.80	2.51	2.18	1.99	1.79	1.52	1.00

TABLE A.6. Values of the upper 5% of points of the distribution $F_{0.05} = S_1^2/S_2^2$.

Number of degrees of freedom											
	ν_1										
ν_2	2	3	4	5	6	8	12	16	24	50	∞
2	19.00	19.16	19.25	19.30	19.33	19.37	19.41	19.43	19.45	19.47	19.50
3	9.55	9.28	9.12	9.01	8.94	8.84	8.74	8.69	8.64	8.58	8.53
4	6.94	6.59	6.39	6.26	6.16	6.04	5.91	5.84	5.77	5.70	5.63
5	5.79	5.41	5.19	5.05	4.95	4.82	4.68	4.60	4.53	4.44	4.36
6	5.14	4.76	4.53	4.39	4.28	4.15	4.00	3.92	3.84	3.75	3.67
7	4.74	4.35	4.12	3.97	3.87	3.73	3.57	3.49	3.41	3.32	3.23
8	4.46	4.07	3.84	3.69	3.58	3.44	3.28	3.20	3.12	3.03	2.93
9	4.26	3.86	3.63	3.48	3.37	3.23	3.07	2.98	2.90	2.80	2.71
10	4.10	3.71	3.48	3.33	3.22	3.07	2.91	2.82	2.74	2.64	2.54
11	3.98	3.59	3.36	3.20	3.09	2.95	2.79	2.70	2.61	2.50	2.40
12	3.88	3.49	3.26	3.11	3.00	2.85	2.69	2.60	2.50	2.40	2.30
13	3.80	3.41	3.18	3.02	2.92	2.77	2.60	2.51	2.42	2.32	2.21
14	3.74	3.34	3.11	2.96	2.85	2.70	2.53	2.44	2.35	2.24	2.13
15	3.68	3.29	3.06	2.90	2.79	2.64	2.48	2.39	2.29	2.18	2.07
16	3.63	3.24	3.01	2.85	2.74	2.59	2.42	2.33	2.24	2.13	2.01
17	3.59	3.20	2.96	2.81	2.70	2.55	2.38	2.29	2.19	2.08	1.96
18	3.55	3.16	2.93	2.77	2.66	2.51	2.34	2.25	2.15	2.04	1.92
19	3.52	3.13	2.90	2.74	2.63	2.48	2.31	2.21	2.11	2.00	1.88
20	3.49	3.10	2.87	2.71	2.60	2.45	2.28	2.18	2.08	1.96	1.64
21	3.47	3.07	2.84	2.68	2.57	2.42	2.25	2.15	2.05	1.93	1.81
22	3.44	3.05	2.82	2.66	2.55	2.40	2.23	2.13	2.03	1.91	1.78
23	3.42	3.03	2.80	2.64	2.53	2.38	2.20	2.11	2.00	1.88	1.76
24	3.40	3.01	2.78	2.62	2.51	2.36	2.18	2.09	1.98	1.86	1.73
25	3.38	2.99	2.76	2.60	2.49	2.34	2.16	2.07	1.96	1.84	1.71
26	3.37	2.98	2.74	2.59	2.47	2.32	2.15	2.05	1.95	1.82	1.69
27	3.35	2.96	2.73	2.57	2.46	2.30	2.13	2.03	1.93	1.80	1.67
28	3.34	2.95	2.71	2.56	2.44	2.29	2.12	2.02	1.91	1.78	1.65
29	3.33	2.93	2.70	2.54	2.43	2.28	2.10	2.00	1.90	1.77	1.64
30	3.32	2.92	2.69	2.53	2.42	2.27	2.09	1.99	1.89	1.76	1.62
35	3.26	2.87	2.64	2.48	2.37	2.22	2.04	1.94	1.83	1.70	1.57
40	3.23	2.84	2.61	2.45	2.34	2.18	2.00	1.90	1.79	1.66	1.51
45	3.21	2.81	2.58	2.42	2.31	2.15	1.97	1.87	1.76	1.63	1.48
50	3.18	2.79	2.56	2.40	2.29	2.13	1.95	1.85	1.74	1.60	1.44
60	3.15	2.76	2.52	2.37	2.25	2.10	1.92	1.81	1.70	1.56	1.39
70	3.13	2.74	2.50	2.35	2.23	2.07	1.89	1.79	1.67	1.53	1.35
80	3.11	2.72	2.49	2.33	2.21	2.06	1.88	1.77	1.65	1.51	1.32
90	3.10	2.71	2.47	2.32	2.20	2.04	1.86	1.76	1.64	1.49	1.30
100	3.09	2.70	2.46	2.30	2.19	2.03	1.85	1.75	1.63	1.48	1.28
125	3.07	2.68	2.44	2.29	2.17	2.01	1.83	1.72	1.60	1.45	1.25
∞	2.99	2.60	2.37	2.21	2.09	1.94	1.75	1.64	1.52	1.35	1.00

Glossary

Absolute error (of a measuring instrument): The difference between a value of a measurand obtained by a measuring instrument and the true value of this measurand. *Note*: The absolute error of a material measure is the difference between the nominal value of this measure and the true value of a quantity that was reproduced by this measure.

Absolutely constant elementary error: An elementary error that remains the same value in repeated measurements performed under the same conditions using an arbitrarily chosen measuring instrument of a given type. The value of an absolutely constant error is unknown, but its limits can be estimated.

Accuracy class: A class of measuring instruments that meets certain metrological requirements that are intended to keep errors within specified limits (Ref. 2, no. 5.19).

Accuracy of measurement: A qualitative expression of the closeness of the result of a measurement to the true value of the measurand.

Accuracy of a measuring instrument: The ability of a measuring instrument to produce measurements whose results are close to the true value of a measurand.

Additional error of measuring instruments: The difference between an error of a measuring instrument when the value of one influence quantity exceeds its reference value and the error of the measuring instrument under reference condition.

Calibration: The set of operations that establish, under specified conditions, the relationship between values indicated by a measuring instrument or measuring system or values represented by material measure and the corresponding known values of a measurand (Ref. 2, no. 6.11). Results of a calibration may be presented in the form of a calibration curve, or a statement that the errors of an instrument exceed or do not exceed certain limits.

Combined measurement: Measurements of several quantities of the same kind based on results of direct measurements of different combinations of them (Ref. 4, no. 4.4). The typical sign of combined measurement is that the number of unknown quantities is less than the number of performed measurements. *Example*: Combined measurement is performed when the masses of separate weights from one set are found using the known value of one weight.

Conditionally constant elementary error (of a measurement): An elementary error that varies in repeated measurements performed under the same conditions or with different specimens of measuring instruments of the same type having certain limits. These limits can be calculated or estimated.

Dead band: Maximum interval through which a stimulus may be changed in either direction without producing a change in response of a measuring instrument (Ref. 2, no. 5.13).

Drift: A slow variation with time at an output of a measuring instrument that is independent of a stimulus.

Elementary error (of a measurement): A component of error or uncertainty of a measurement associated with a single source of inaccuracy of the measurement.

Error (of a measurement): The deviation of the result of a measurement from the true value of the measurand expressed in absolute or relative form.

Fiducial error: A ratio of absolute error of a measuring instrument and a value specified for this instrument. The specified value is called the fiducial value. It may be, for example, the span or the upper limit of the nominal range of the measuring instrument. Fiducial error is expressed as a percentage.

Inaccuracy (of a measurement): A qualitative characteristic of the degree of deviation of a measurement result from the true value of the measurand. Quantitatively, inaccuracy can be characterized either as a measurement error or as a measurement uncertainty.

Indicating instrument: A measuring instrument that displays the value of a measurand.

Indirect measurement: A measurement in which the value of the measurand is calculated using measurements of other quantities related to the measurand by some known relation. We shall call these other quantities *measured arguments* or, briefly, *arguments.*

Influence coefficient: A factor that is multiplied by a value of the variation of the influence quantity relative to its reference condition limits gives an additional error.

Influence function: A metrological characteristic of the measuring instrument expressing the relationship between the measuring instrument errors and an influence quantity.

Informative parameter (of an input signal): A parameter of an input process at a measuring instrument that reflects the value of a measurand.

Intrinsic error: The error of a measuring instrument, determined under reference conditions (Ref. 2, no. 5.24).

Limits of permissible error (of a measuring instrument): Extreme values of an error permitted by specification, regulations, and so on, for a given measuring instrument (Ref. 2, no. 5.21).

Material measure: A measuring instrument that reproduces a physical quantity with known value.

Measurand: A particular physical quantity subject to measurement.

Measurement: The set of experimental operations that are performed using special technical products (measuring instruments) for the purpose of determinating the value of a physical quantity (Ref. 4, no. 4.1).

Measurement vector: A set of matched measurements of all arguments defining an indirect measurement.

Measuring instrument: A technical object developed for the purpose of measurements (Ref. 4, no. 5.1). The measuring instrument has standardized metrological characteristics.

Measuring standard: A measuring instrument intended to materialize and/or conserve a unit of a physical quantity to transmit its value to other measuring instruments (Ref. 4, no. 10.1).

Measuring system: A complete set of measuring instruments and supplementary equipment assembled for obtaining measurement results in required form and for inputting data to a control system.

Measuring transducer: A measuring instrument that converts the measurement signals into a form suitable for transmission or processing. The signals at the output of a measuring transducer cannot be directly observed.

Metrological characteristic of a measuring instrument: A characteristic of a measuring instrument that is necessary to judge the suitability of the instrument for performing measurements in a known range or that is necessary to estimate the inaccuracy of measurement results.

Metrology: Science of measurement (Ref. 2, no. 2.2). Metrology is an applied science. It includes knowledge of measurements of any kind of physical quantities and with any level of accuracy.

Normal operating conditions: Conditions of use of measuring instruments giving the ranges of the influence quantities within which a measuring instrument is designed to operate and for which the metrological characteristics of this instrument lie within specified limits.

Primary standard: A measurement standard that has the highest accuracy in a country.

Random error (of a measurement): A component of the inaccuracy of a measurement that, in the course of several measurements of the same measurand under the same conditions, varies in an unpredictable way.

Reference conditions: A complete set of values of influence quantities standardized for specific types of measuring instruments to ensure a maximum accuracy of the performing measurements and to calibrate these instruments.

Reference standard: A standard, generally having the highest metrological quality available at a given location or in a given organization, from which measurements made there are derived (Ref. 2, no. 6.6).

Relative error: Absolute error divided by a true value of the measurand (Ref. 2, no. 3.12). The measurement result substitutes the true value, in practice.

Repeatability of a measurement: The closeness of agreement among several consecutive measurements for the same measurand performed, under the same operating conditions with the same measuring instruments, over a short period of time.

Reproducibility of a measurement: The closeness of agreement among repeated measurements for the same measurand performed in different locations, under different operating conditions, or over a long period of time.

Resolution: The smallest interval between two adjacent values of the output signal of a measuring instrument that can be distinguished.

Response time: The time interval between the instant when a measuring instrument gets a stimulus and the instant when the response reaches and remains within specified limits of its final steady value.

Result of measurement: The value of a measurand obtained by measurement (Ref. 4, no. 8.18). The measurement result is expressed as a product of a number and a proper unit. For example, 1.25 m is the value of the length of a body; here 1.25 is the number, and m (meter) is the unit.

Secondary standard: A measurement standard that obtains the value of a unit from the primary standard.

Sensitivity: The change in the response of a measuring instrument divided by the corresponding change in the stimulus (Ref. 2, no. 5.10).

Span: The absolute value of the difference between the two limits of a nominal range of a measuring instrument (Ref. 2, no. 5.02). Span is expressed in a unit of the measured quantity.

Example: A range is -15 to $+15$ V; the span is 30 V.

Systematic error (of measurement): A component of the inaccuracy of measurement that, in the course of several measurements of the same measurand, remains constant or varies in a predictable way.

True value: The value of a measurand that being known would ideally reflect both qualitatively and quantitatively the corresponding property of an object (Ref. 4, no. 2.5).

Uncertainty of measurement: An interval within which a true value of a measurand lies with given probability. Uncertainty is defined with its limits and corresponding confidence probability. The limits are read out from a result of measurement. Uncertainty can be expressed in absolute or relative form.

Verification: A kind of calibration that reveals whether the errors of a measuring instrument lie within their specified limits.

Working standard: A measurement standard that is used to calibrate measuring instruments.

References

Standards and Recommendations

1. Guide to the Expression of Uncertainty in Measurement, ISO (1995).
2. International Vocabulary of Basic and General Terms in Metrology, 2nd ed., ISO (1993).
3. Measurement Uncertainty, American National Standard ANSI/ASME PTC 19.1 (1985).
4. Metrology, Terms and Definitions, State Standard of the Soviet Union, GOST 16263-70, Moscow (1970).
5. Direct Measurements with Multiple Observations, Methods of Processing the Results of Observations: Basic principles, State Standard of the Soviet Union, GOST 8.207–76, Moscow (1976).
6. National Measurement Accreditation Service, B 3003, English National Physical Laboratory, Teddington (1986).
7. Process Instrumentation Terminology, American National Standard ANSI/ISA S51.1 (1979).
8. Publication 51, PT 1-84 (1984), Direct Acting Indicating Analogue Electrical Measuring Instruments and Their Accessories, 4th ed., Bureau Central de la Comission Electrotechnique Internationale, Geneva (1984).
9. Recommandation Internationale 34, Classes de precision des Instruments de Mesurage, Organisation International de Metrologie Legale, Paris (1974).
10. Standard Practice for Use of the International System of units (SI), American Society for Testing and Materials E 380–91 (1991).
11. Standard Practice for Dealing with Outlying Observations, American Society for Testing and Materials, E 178–80 (Reapproved 1989).
12. Standard Practice for the Evaluation of Single-Pan Mechanical Balances, American Society for Testing and Materials, E 319–85 (1985).
13. Technical Data Fluke, Application Information B0005, Total Instrument Accuracy for Digital Thermometers and Data Loggers (formerly AB-29).

Books and Articles

14. R.B. Abernethy et al., *Measurement Uncertainty Handbook* (Instrument Society of America 1980).
15. V.Ya. Alekseev, F.M. Karavaev et al., "Estimate of the measured quantity from results of measurements with different systematic errors," *Metrologia* (Supplement to the journal *Izmeritel'naya Technika*[1]), No. 1, 18–24 (1978).

16. D.F. Andrews, P.J. Bickel, F.R. Hampel, P.J. Huber, W.H. Rogers, and J.W. Tukey, *Robust Estimates of Location* (Princeton University Press, Princeton, NJ, 1972).
17. H. Bäckström, "Uber die dezimalgleichung beim ablesen von scalen," *Zeitschrift fur Instrumentenkunde* (1930–32).
18. I.E. Burns, P.J. Campion, and A. Williams, "Error and uncertainty," *Metrologia* **9**, 101–104 (1973).
19. E.R. Cohen, "Uncertainty and error in physical measurements," Enrico Fermi Summer School, Italian Physical Society. *Proceedings of the International School of Physics, "Enrico Fermi"*. Course CX, edited by L. Crovini and T.J. Quinn. Villa Marigola, June 27–July, 7, 1989. *Metrology at the Frontiers of Physics and Technology* (North-Holland, Amsterdam, 1992).
20. H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ, 1946).
21. C. Croarkin, *Measurement Assurance Programs,* Part II: *Development and Implementation,* NBS Special Publication 676–11 (U.S. GPO, Washington, DC, 1985).
22. R.H. Dieck, *Measurement Uncertainty. Methods and Applications,* ISA (1992).
23. E.F. Dolinskii, *Analysis of the Results of Measurements* [in Russian], 2nd ed. (Izd. standartov, Moscow, 1973).
24. B.D. Ellis, *Basic Concepts of Measurement* (Cambridge University Press, Cambridge, 1966).
25. C. Eisenhart, "Realistic evaluation of the precision and accuracy of instrument calibration systems," *J. Res. Natl. Bur. Stand* C **67**, 161–187 (1963).
26. I. Gertsbakh, *Measurement Theory for Engineers* (Springer-Verlag, Berlin, 2003).
27. S.V. Gorbatsevich, "Analysis of errors in measurement of the emf of reference standard normal elements on electric balances," in *Investigations in the Field of Electric Measurements,* Proceedings of the All-Union Scientific-Research Institute of Metrology, No. 38 (98) [in Russian] (Standartgiz, Leningrad, 1959), pp. 5–20.
28. V.A. Granovskii, *Dynamic Measurements, Principles of Metrological Support* [in Russian] (Energoatomizdat, Leningrad, 1984).
29. V.A. Granovskii and T.N. Siraya, *Methods of Processing of the Measurements of Experimental Data* [in Russian] (Energoatomizdat, Leningrad division, Leningrad, 1990).
30. I.A. Grachev and S.G. Rabinovich, "Approximate method for constructing the distribution function of the composite of several distributions," *Izmeritel'naya Technika*,[1] No. 1, 8–11 (1968).
31. R.V. Hogg, "Some observations on robust estimation," *Journal of the American Statistical Association*, December 1967, Vol. 62 Issue 320, p. 1179–1186.
32. A.M. Kagan and Yu.V. Linnik, "On some problems of mathematical statistics in metrology," in *Fifty Years of Metric Reform in the USSR,* Proceedings of the Metrological Institute of the USSR [in Russian], D.I. Mendeleev All-Union Scientific-Research Institutes of Metrology, No. 123 (183) (Leningrad, 1970), pp. 39–47.
33. F.M. Karavaev, *Measurements of the Activity of Nuclides* [in Russian] (Izd. standartov, Moscow, 1972).
34. Zn.F. Kudryshova and S.G. Rabinovich, "Methods of experimental data processing in indirect measurements," *in Methods of Experimental Data Processing in Measurements*,

[1] The journal *Izmeritel'naya Technika* was published for many years in the United States in English under the title *Measurement Techniques.*

Proceedings of the Metrological Institutes of the USSR, D.I. Mendeleev All-Union Scientific Institute of Metrology, No. 172 (232), (Energia, Leningrad, 1975).
35. M.F. Malikov, *Foundations of Metrology* [in Russian] (Committee on Measures and Measuring Devices at the Council of Ministers of the USSR, Moscow, 1949).
36. J. Mandel, *The Statistical Analysis of Experimental Data* (National Bureau of Standards, Interscience Publishers, New York, 1964).
37. J. Mandel, *Evaluation and Control of Measurements* (Marcel Dekker, New York, 1991).
38. B.S. Massey, *Measures in Science and Engineering, Their Expression, Relation and Interpretation* (Wiley, New York, 1986).
39. A.I. Mechanikov, "Methods for immediate processing of the results of observations with the help of ordering statistics," *Metrologia* (Supplement to *Izmeritel'naya Technika*[1], No. 4, 47–58 (1972).
40. P.J. Mohr and B.N. Taylor, "CODATA recommended values of the fundamental physical constants," *J. Phys. Chem. Ref. Data* **28**, 1713–1852 (1999).
41. S.G. Rabinovich, "Method for calculating the measurement errors," *Metrologia* (Supplement to *Izmeritel'naya Technika*[1]), No. 1, 3–12 (1970).
42. S.G. Rabinovich, *Galvanometric Self-Balancing Instruments* [in Russian] (Izd. standartov, Moscow, 1972).
43. S.G. Rabinovich and T.L. Yakovleva, "Analysis of the temporal stability of the distributions of the errors of measuring instruments," *Metrologia* (Supplement to *Izmeritel'naya Technika*[1]*)*, No. 7, 8–15 (1977).
44. S.G. Rabinovich, *Measurement Errors* [in Russian] (Energia, Leningrad, 1978).
45. S.G. Rabinovich, "Methods of calculation of measurement errors by their components," *The 6th International Conference of the Israel Society for Quality Assurance.* Tel-Aviv, Nov. 1986, Paper and abstracts, 4.3.4, pp. 1–4.
46. S.G. Rabinovich, "An efficient calculation for indirect measurements and a new approach to the theory of indirect measurements," *Proceedings of the Measurement Science Conference*, Anaheim, California, January 25 and 26, 1996.
47. L. Sachs, *Applied Statistics: A Handbook of Techniques* (Springer-Verlag, New York, 1984).
48. K.P. Shirokov, *Basic Concepts of Metrology, Proceedings of the Metrological Institutes of the USSR* [in Russian], D.I. Mendeleev All-Union Scientific-Research Institute of Metrology No. 130 (190) (Izd. standartov, M.-L., 1972).
49. K.P. Shirokov, V.O. Arutyunov et al., "Basic concepts of the theory of dynamic measurements," *Proceedings of the 8th IMEKO Congress* (Akademiai Kiado, Budapest, 1980).
50. V.M. Sviridenko, "Logical-gnoseological aspect of the problem of the accuracy of measurements," *Izmeritel'naya Technika*[1]*,* No. *5*, 6–8 (1971).
51. B.N. Taylor, and C.E. Kuyatt. "Guidelines for evaluating and expressing the uncertainty of NIST measurement results," NIST Technical Note 1297, 1994 edition.
52. B.L. van der Waerden, *Mathematical Statistics* (Springer-Verlag, New York, 1969).
53. T.L. Yakovleva, "About statistical homogeneity of the samples of measuring instrument errors," *Metrologiya* (Supplement to *Izmeritel'naya Technika*[1]), No. 2, 19–23 (1979).

Index

N