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

Preface v

CHAPTER 1
General Information About Measurements 1

1.1. Basic Concepts and Terms 1
1.2. Metrology and the Basic Metrological Problems 3
1.3. Initial Points of the Theory of Measurements 10
1.4. Classification of Measurements 15
1.5. Classification of Measurement Errors 20
1.6. Principles of Estimation of Measurement Errors and Uncertainties 22
1.7. Presentation of Results of Measurements; Rules for Rounding Off 24
1.8. Basic Conventional Notations 28

CHAPTER 2
Measuring Instruments and Their Properties 29

2.1. Types of Measuring Instruments 29
2.2. The Concept of an Ideal Instrument: Metrological Characteristics of Measuring Instruments 32
2.3. Standardization of the Metrological Characteristics of Measuring Instruments 36
2.4. Some Suggestions for Changing Methods of Standardization of Errors of Measuring Instruments and Their Analysis 48
2.5. Dynamic Characteristics of Measuring Instruments and Their Standardization 52
2.6. Statistical Analysis of the Errors of Measuring Instruments Based on Data Provided by Calibration Laboratories 57

CHAPTER 3
Prerequisites for the Analysis of the Inaccuracy of Measurements and for Synthesis of Their Components 61

3.1. Relationship Between Error and Uncertainty 61

3.2. Classification of Elementary Errors 62
3.3. Mathematical Models of Elementary Errors 64
3.4. Methods for Describing Random Quantities 66
3.5. Construction of the Composition of Uniform Distributions 70
3.6. Universal Method for Constructing the Composition of Distributions 74
3.7. Natural Limits of Measurements 81

CHAPTER 4
Statistical Methods for Experimental Data Processing 91

4.1. Requirements for Statistical Estimations 91
4.2. Estimation of the Parameters of the Normal Distribution 92
4.3. Outlying Results 95
4.4. Construction of Confidence Intervals 97
4.5. Methods for Testing Hypotheses About the Form of the Distribution Function of a Random Quantity 101
4.6. Methods for Testing Sample Homogeneity 103
4.7. Trends in Applied Statistics and Experimental Data Processing. 109
4.8. Example: Analysis of Measurement Results in Comparisons of Measures of Mass 112

CHAPTER 5
Direct Measurements 115

5.1. Relation Between Single and Multiple Measurements 115
5.2. Identification and Elimination of Systematic Errors 118
5.3. Estimation of Elementary Errors 124
5.4. Method for Calculating the Errors and Uncertainties of Single Measurements 128
5.5. Example: Calculation of Uncertainty in Voltage Measurements Performed with a Pointer-Type Voltmeter 132
5.6. Methods for Calculating the Uncertainty in Multiple Measurements 138
5.7. Comparison of Different Methods for Combining Systematic and Random Errors 149
5.8. Essential Aspects of the Estimation of Measurement Errors when the Number of Measurements Is Small 153
5.9. General Plan for Estimating Measurement Uncertainty 155

CHAPTER 6
Indirect Measurements 159

6.1. Basic Terms and Classification 159
6.2. Correlation Coefficient and its Calculation 160
6.3. The Traditional Method of Experimental Data Processing 162
6.4. Shortcomings of the Traditional Method 166
6.5. The Method of Reduction 168
6.6. The Method of Transformation 169
6.7. Errors and Uncertainty of Indirect Measurement Results 174

CHAPTER 7
Examples of Measurements and Measurement Data Processing 179

7.1. An Indirect Measurement of the Electrical Resistance of a Resistor 179
7.2. The Measurement of the Density of a Solid Body 182
7.3. The Measurement of Ionization Current by the Compensation Method 189
7.4. The Measurement of Power at High Frequency 192
7.5. The Measurement of Voltage with the Help of a Potentiometer and a Voltage Divider 193
7.6. Calculation of the Uncertainty of the Value of a Compound Resistor 197

CHAPTER 8
Combined Measurements 201

8.1. General Remarks About the Method of Least Squares 201
8.2. Measurements with Linear Equally Accurate Conditional Equations 203
8.3. Reduction of Linear Unequally Accurate Conditional Equations to Equally Accurate Conditional Equations 205
8.4. Linearization of Nonlinear Conditional Equations 206
8.5. Examples of the Applications of the Method of Least Squares 208
8.6. Determination of the Parameters in Formulas from Empirical Data and Construction of Calibration Curves 213

CHAPTER 9
Combining the Results of Measurements 219

9.1. Introductory Remarks 219
9.2. Theoretical Principles 219
9.3. Effect of the Error of the Weights on the Error of the Weighted Mean 223
9.4. Combining the Results of Measurements in Which the Random Errors Predominate 225
9.5. Combining the Results of Measurements Containing both Systematic and Random Errors 226
9.6. Example: Measurement of the Activity of Nuclides in a Source 233

CHAPTER 10
Calculation of the Errors of Measuring Instruments 237

10.1. The Problems of Calculating Measuring Instrument Errors 237
10.2. Methods for Calculating Instrument Errors 238
10.3. Calculation of the Errors of Electric Balances (Unique Instrument) 249
10.4. Calculation of the Error of ac Voltmeters (Mass-Produced Instrument) 251
10.5. Calculation of the Error of Digital Thermometers (Mass-Produced Instrument) 258

CHAPTER 11
Problems in the Theory of Calibration 263

11.1. Types of Calibration 263

11.2. Estimation of the Errors of Measuring Instruments in Verification 265
11.3. Rejects of Verification and Ways to Reduce Their Number 269
11.4. Calculation of a Necessary Number of Standards 275

CHAPTER 12
Conclusion 283

12.1. Measurement Data Processing: Past, Present, and Future 283
12.2. Remarks on the "International Vocabulary of Basic and General Terms in Metrology" 285
12.3. Drawbacks of the "Guide to the Expression of Uncertainty in Measurement" 286

Appendix 289

Glossary 295

References 299

Index 303

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.

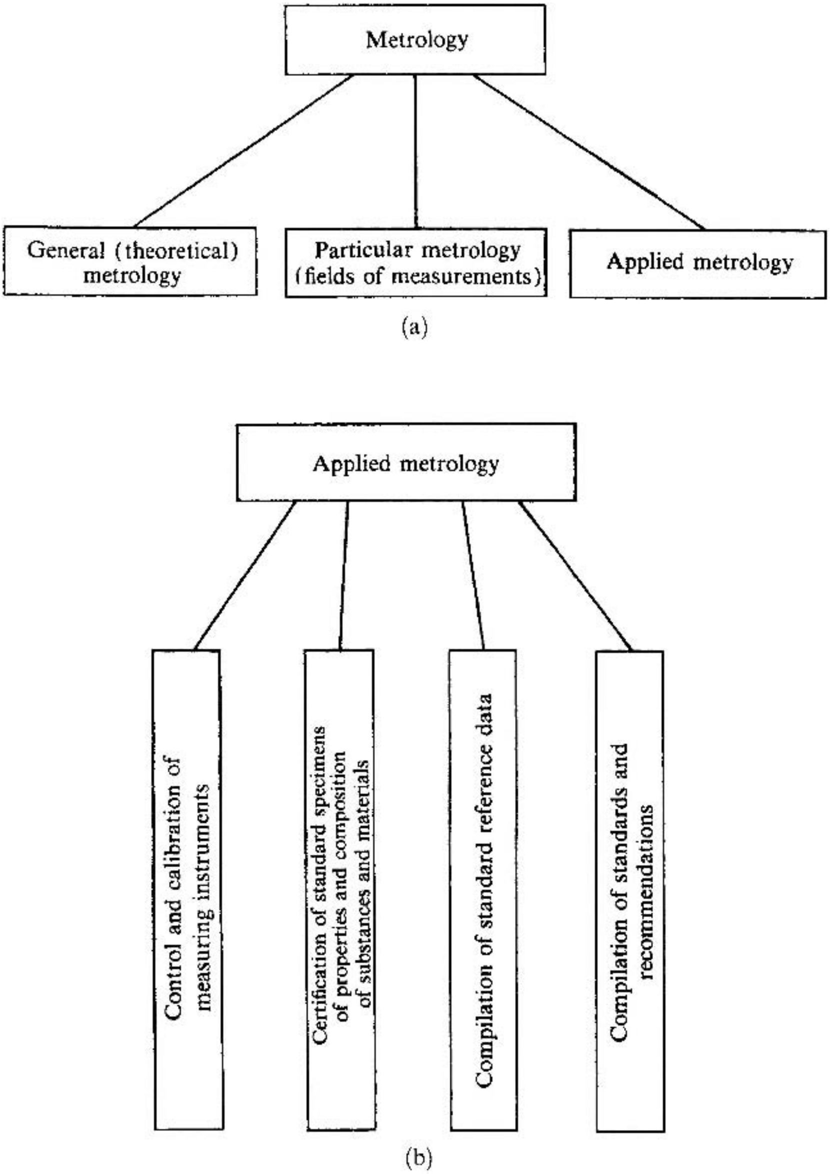
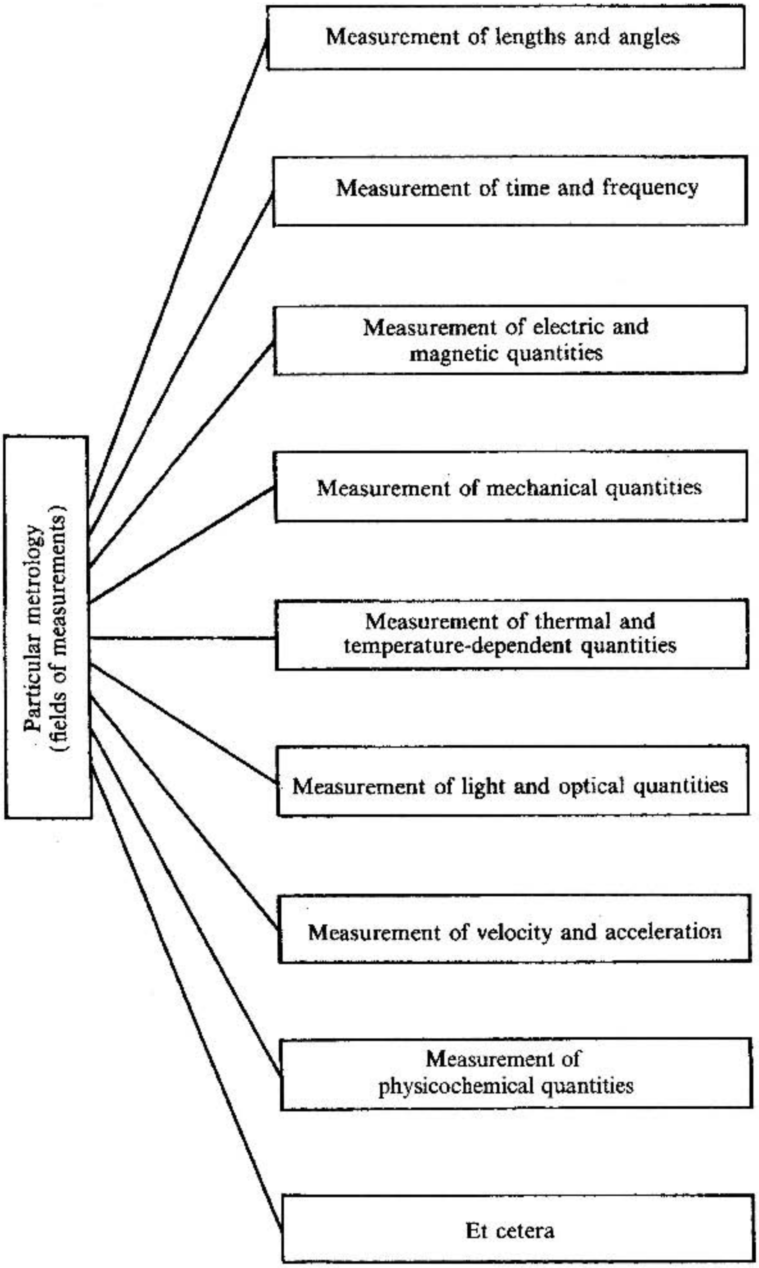
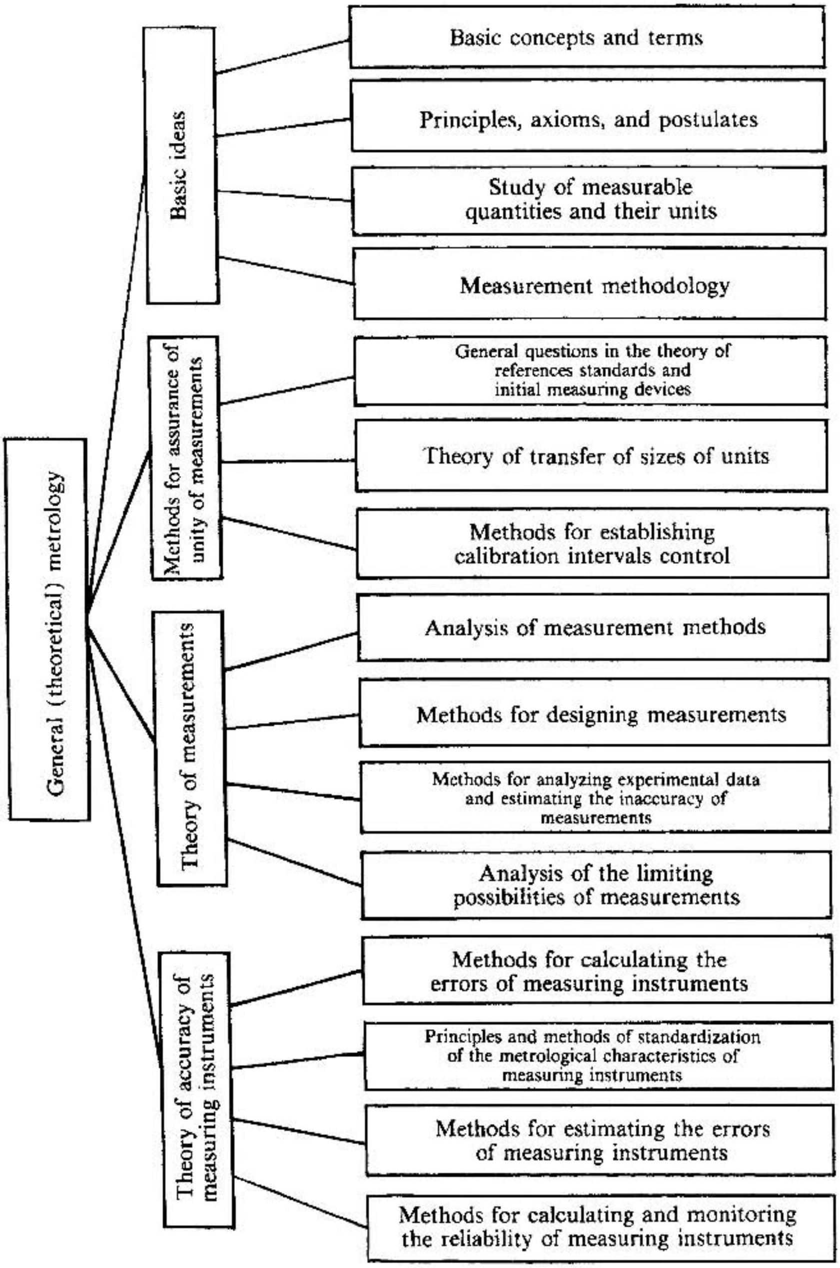


FIGURE 1.1. Schematic picture of the basic problems of metrology: (a) metrology, (b) applied metrology, (c) particular metrology, and (d) general metrology.



(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.