Zbigniew Kotulski
Wojciech Szczepiński

Error Analysis
with Applications
in Engineering

Springer

Error Analysis with Applications in Engineering

SOLID MECHANICS AND ITS APPLICATIONS
Volume 169

Aims and Scope of the Series

The fundamental questions arising in mechanics are: *Why?*, *How?*, and *How much?* The aim of this series is to provide lucid accounts written by authoritative researchers giving vision and insight in answering these questions on the subject of mechanics as it relates to solids.

The scope of the series covers the entire spectrum of solid mechanics. Thus it includes the foundation of mechanics; variational formulations; computational mechanics; statics, kinematics and dynamics of rigid and elastic bodies: vibrations of solids and structures; dynamical systems and chaos; the theories of elasticity, plasticity and viscoelasticity; composite materials; rods, beams, shells and membranes; structural control and stability; soils, rocks and geomechanics; fracture; tribology; experimental mechanics; biomechanics and machine design.

The median level of presentation is the first year graduate student. Some texts are monographs defining the current state of the field; others are accessible to final year undergraduates; but essentially the emphasis is on readability and clarity.

Zbigniew Kotulski · Wojciech Szczepiński

Error Analysis
with Applications
in Engineering

Springer

Zbigniew Kotulski
PAN Warszawa
Inst. Fundamental
Technological Research
ul. Pawińskiego 5B
02-106 Warsaw
Poland
zkotulsk@ippt.gov.pl

Wojciech Szczepiński
PAN Warszawa
Inst. Fundamental
Technological Research
ul. Pawińskiego 5B
02-106 Warsaw
Poland

Printed on acid-free paper

Preface

Our intention in preparing this book was to present in as simple a manner as possible those branches of error analysis which find direct applications in solving various problems in engineering practice. The main reason for writing this text was the lack of such an approach in existing books dealing with the error calculus. Most of books are devoted to mathematical statistics and to probability theory. The range of applications is usually limited to the problems of general statistics and to the analysis of errors in various measuring techniques.

Much less attention is paid in these books to two-dimensional and three-dimensional distributions, and almost no attention is given to problems connected with the two-dimensional and three-dimensional vectorial functions of independent random variables. The theory of such vectorial functions finds new applications connected, for example, with analysis of the positioning accuracy of various mechanisms, among them of robot manipulators and automatically controlled earth-moving and loading machines, such as excavators.

Besides basic information concerning classical simple applications of error calculus, a substantial part of the book is devoted to new aspects of more advanced problems along with numerous examples of practical applications in engineering practice. Among others, the Mohr circles representation of tensors is used for transformation of the components of covariance tensors, for determination of linear regression, for analysis of the accuracy of artillery fire, and for analysis of the positioning accuracy of various mechanisms. Methods of determination of the ellipses and ellipsoids of probability concentration have been described in detail, along with examples of practical calculations.

Chapters 1, 2, 3 and 4 contain a presentation of the fundamentals of error calculus: basic characteristics of error distributions, histograms and their various applications, basic continuous distributions of errors and functions of independent random variables. In Chap. 5, two-dimensional distributions of errors are discussed with applications to analysis of the accuracy of artillery fire, to the determination of linear regression for sets of experimental points, and to the calculation of correlation coefficients. Fundamentals of the theory of two-dimensional continuous independent and dependent random variables are also discussed in that chapter. Then the methods of determination of the ellipses of probability concentration for a two-dimensional continuous normal distribution are given.

Chapter 6 deals with the two-dimensional vectorial functions of independent random variables along with practical applications to analysis of the positioning accuracy of mechanisms with two-dimensional movements. The procedure of determination of ellipses of probability concentration is also described.

In Chap. 7, three-dimensional distributions of errors are considered, while Chap. 8 deals with three-dimensional vectorial functions of independent random variables. The theory is illustrated by examples of analysis of the positioning accuracy of robot manipulators. Examples of determining the ellipsoids of probability concentration are presented.

Chapter 9 consists of basic information connected with calculation of probabilities that functions of independent random variables satisfy specific inequalities. Such problems are inherent in the theory of reliability of engineering structures and, on the other hand, are a natural generalization of the traditional error calculus.

The concluding Chap. 10 gives introductory material for future studies in application of more advanced probability theory in engineering. It also contains bibliographical remarks for future studies and extended references.

This book has been written for readers whose main interests are applications of error calculus in various problems of engineering. We have indicated that certain important concepts of that calculus such as, for example, variance and covariance, are notionally analogous to the concepts of inertia moments of plane or solid figures. The standard deviation is analogous to the so-called inertia radius of such figures. The procedure for calculating such values is analogous to that of determination of the centers of gravity of plane or solid figures.

In the first nine chapters much attention is paid to the practical significance of error analysis. However, some additional information concerning its mathematical foundations has been included in this book. It may be omitted by readers who are mainly interested in applications of error calculus.

The authors want to express their thanks to Professors Marek Sokolowski and Richard Hetnarski for their help and discussions during preparation of this book.

Warsaw   Zbigniew Kotulski
                                                                      Wojciech Szczepiński

Contents

Chapter 1
Basic Characteristics of Error Distribution; Histograms

1.1 Introductory Remarks; Histograms

In engineering practice one can distinguish errors of various origins. Most commonly dealt with are:

− measurement errors of a particular parameter (mass, time, length, angle) with respect to the actual magnitude of that parameter,
− error of a particular parameter of a manufactured product with respect to the required nominal magnitude of that parameter,
− positioning error of a robot manipulator with respect to the required position.

The errors may be of systematic and random origin. *Systematic errors* are usually caused by strictly defined factors affecting the result of measurements (e.g., slight warming of the measuring device, faulty vision of the operator of the measuring device, incorrect adjusting of that device). Usually, such systematic errors can be estimated and eliminated.

On the other hand, *random errors* result from the random action of numerous small, sometimes imperceptible, factors connected with the structure of the measuring device, manufacturing machine or a manipulator (e.g., clearances between moving parts). During measurements they may be the result of the subjectivity of readings made by the operator.

These random errors are subject to the error calculus based on probability theory. In general, it is assumed that errors are small with respect to the magnitude measured during the manufacturing process.

Such random errors are analyzed in the error calculus based on the theory of probability. Generally, it is assumed that error is small with respect to the magnitude of the parameter measured during the manufacturing process.

As an exception to this assumption, the example of the center distance error of two intentionally coaxial elements may be mentioned. The expected magnitude of such a center distance equals zero. In such cases, the center distance error should be small compared to the characteristic dimension of the object in question. Other examples of this kind are discussed in Sect. 3.4.2.

Fig. 1.1 The histogram: the accuracy of an automatic machine manufacturing small rivets
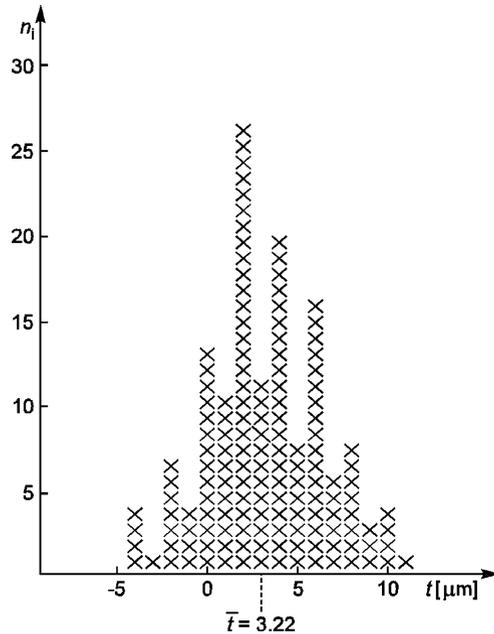
In measuring practice, it is convenient to divide the whole range of possible errors into a number of small sectors and then to include each specific result of measurements in the respective sector. Usually, all such sectors are chosen to have the same length. The final result of such a procedure of measurements may be then represented in the form of a so-called *histogram*. On the horizontal axis of a histogram, the ranges of the consecutive sectors are laid off, while the ordinate of each of the sectors is the so-called frequency n_i of results of the measurements which have been included into that sector, numbered as the i-th sector. Another possibility is to lay off on the ordinate axis the fraction or proportion $w_i = n_i/n$, of the total number of measurements n.

In such a procedure the distribution of errors, being by its nature continuous, is treated as a discrete (quasi-step) distribution. It is implicitly assumed that the errors are of a finite selected number of magnitudes (e.g., equal to the initial magnitude corresponding to a specific sector), instead of assuming all practically possible values of the entire range. The difference between the continuous and step-wise approaches is more visible when the notion of a probability distribution is used.

In Fig. 1.1 is presented an example of a histogram summarizing the results of testing the accuracy of a certain automatic machine that manufactures small rivets for use in the production of light planes. Altogether 149 rivets produced by the machine were tested. As the pivotal parameter, the elevation T of a rivet's conical head over a certain reference plane was measured. The accuracy of measurement was ± 1 μm. In the figure each cross-mark corresponds to the measured elevation of the particular rivet's head included in a specific sector of errors of width equal to 1 μm. The irregularity of results of measurements included in particular sectors
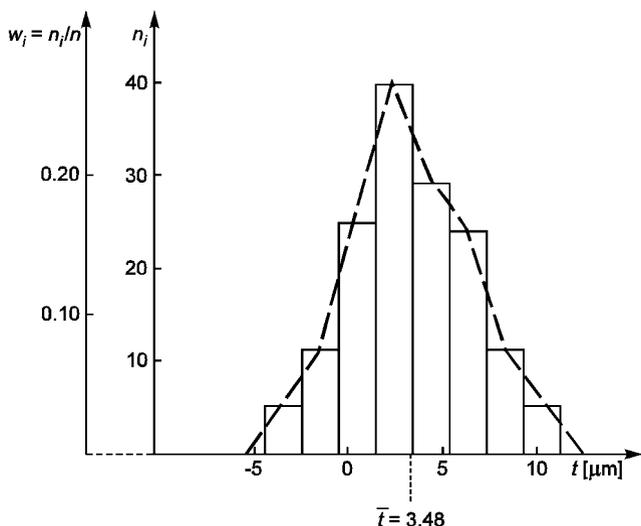
Fig. 1.2 The histogram from Fig. 1.1 with fewer intervals

of the histogram indicates that the range of each sector was chosen too small. The subjectivity in including a particular measured elevation in one of the adjoining sectors plays an important role. When the number of sectors is reduced to eight, such an irregularity does not occur (Fig. 1.2).

On the vertical axis of the histogram one can also lay off the proportion $w_i = n_i/n$ as shown in Fig. 1.2 on its left-hand side. The sum of these proportions in particular sectors must obviously be equal to unity. Then, joining central points on the tops of the columns of the histogram, as shown in Fig. 1.2, we obtain the frequency diagram.

1.2 The Average of a Sample of Measurements

An important parameter in the analysis of the results of a collection (sample) of measurements is the sample average $\overline{x}$. The *sample average* is defined as follows:

$$\overline{x} = \frac{1}{n} \sum_{j=1}^{n} x_j. \tag{1.1}$$

If in the sample that was taken, the result x_j occurred n_j times ($j = 1, 2, \ldots, k$), where $\sum_{j=1}^{k} n_j = n$, then the sample average can be determined using the following equivalent formula:

$$\overline{x} = \frac{1}{n} \sum_{j=1}^{k} x_j n_j. \tag{1.2}$$

In the particular case presented in Fig. 1.2, where the deviations of the measured quantity from its desired value was measured, the average deviation was found to be:

$$\bar{t} = \frac{1}{149} \sum_{j=1}^{8} n_j t_j = 3.48 \; \mu m.$$

This is based on the modified histogram shown in that figure, using the central value of the deviation in each column.

However, if the sample average value is calculated on the basis of the original histogram (shown in Fig. 1.1), it is slightly different:

$$\bar{t} = \frac{1}{149} \sum_{j=1}^{16} n_j t_j = 3.22 \; \mu m.$$

The difference is about 8%.

Such differences in the values of the sample average are not uncommon when the histograms are transformed and formula (1.2) is used. Let us note that the true value of the average, based on the individual values of the measured quantities and calculated with the use of the basic formula (1.1), would be different than the two estimates calculated above.

1.3 Dispersion Measures in Error Analysis

Other important factors in the analysis of results presented in the form of histograms are the dispersion measures of a set of observations or a certain characteristic. The simplest of such measures is the *difference* R between the maximum value x_{max} in the sample and the minimum value x_{min}:

$$R = x_{max} - x_{min}. \tag{1.3}$$

This measure is of a limited practical significance, because it does not give any information concerning the distribution of the quantity in question. The *sample variance* s^2 is the usual measure for estimating a distribution dispersion. By definition, the sample variance is the arithmetical mean value of all squares of deviations of particular values x_j from the average $\bar{x}$ of the entire sample $x_1, \ldots, x_n$. Thus, it is defined by the formula:

$$s^2 = \frac{1}{n} \sum_{j=1}^{n} \left(x_j - \bar{x}\right)^2, \tag{1.4a}$$

or by the equivalent, so-called computing formula:

$$s^2 = \left(\frac{1}{n} \sum_{j=1}^{n} x_j^2\right) - \bar{x}^2. \tag{1.4b}$$

If the particular value x_j is repeated n_j times ($j = 1, 2, \ldots, k$), the sample variance may be defined as

$$s^2 = \frac{1}{n} \sum_{j=1}^{k} (x_j - \bar{x})^2 n_j \tag{1.5a}$$

or in another form

$$s^2 = \left(\frac{1}{n} \sum_{j=1}^{k} x_j^2 n_j \right) - \bar{x}^2. \tag{1.5b}$$

The quantity s is called the *sample standard deviation*.

Replacing in the formula (1.4a) all elements x_j by the expressions $x_0 + a_j$, where x_0 is a certain arbitrarily chosen constant value and a_j stands for the deviation of element x_j from x_0, we arrive at another expression for the variance:

$$s^2 = \frac{1}{n} \sum_{j=1}^{n} a_j^2 - \frac{1}{n^2} \left(\sum_{j=1}^{n} a_j \right)^2. \tag{1.6a}$$

Such a formula is useful in certain calculations. Its application makes it unnecessary to know the actual values of x_j. Only the relative variations of their values are needed.

By extension, if n_j ($j = 1, \ldots, k$) is the number of measurements with the same deviation a_j, the last formula can be written in another form:

$$s^2 = \frac{1}{n} \sum_{j=1}^{k} n_j a_j^2 - \frac{1}{n^2} \left(\sum_{j=1}^{k} n_j a_j \right)^2. \tag{1.6b}$$

As an example, the latter formula was used to calculate variance and standard deviation for the particular case shown in Fig. 1.1, where the value $t_0 = 0$ was chosen as the reference point. For this case we have

$$\sum_{j=1}^{16} a_j^2 = 3054, \qquad \sum_{j=1}^{16} a_j = 483$$

and

$$s^2 = \frac{3054}{149} - \left(\frac{476}{149} \right)^2 = 10.28 \ \mu m^2,$$

$$s = 3.21 \ \mu m.$$

Note that if a histogram is altered according to such operations as that shown in Fig. 1.2, in which a modified version of the histogram from Fig. 1.1 is presented, the value of the standard deviation calculated with the use of formula (1.6b) may be

slightly different than that calculated for the original histogram. Using the altered histogram (shown in Fig. 1.2) we obtain

$$s = 3.11 \; \mu m,$$

while for the original histogram (Fig. 1.1), we get a slightly larger value

$$s = 3.21 \; \mu m.$$

Such a consequence of altering the histograms should be kept in mind.

As yet another measure of data dispersion, the *sample average deviation* $\overline{d}$ is sometimes used. This measure is defined as the mean of the absolute values of all deviations of x_i from the sample average $\overline{x}$,

$$\overline{d} = \frac{1}{n} \sum_{j=1}^{n} |x_j - \overline{x}|. \tag{1.7}$$

This measure is used when all the data are automatically registered. For example, measuring the weight of the entire set of n objects, one can directly find the representative (average or expected) weight $\overline{x}$ of a single object in the set. By consecutive measuring of the actual weight of each of them we can obtain the sum in (1.7) by adding consecutively the absolute values of deviations of the weights from the average value.

The average deviation d is for all theoretical distributions of probability directly connected with the standard deviation σ (see Chap. 2). E.g., for the normal distribution we have $\sigma = 1.25d$.

1.4 Cumulative Frequency Distribution

Using the simple example of a histogram representing the elevation of a rivet's head as a random variable (Fig. 1.2), one can explain the notion of the cumulative frequency distribution (empirical), which is a step function assuming for each t belonging to the l-th section on the axis of abscissa of the histogram the value

$$F_l(t) = \sum_{j=1}^{l} w_j, \tag{1.8}$$

where w_j stand for the probabilities (calculated on the basis of the histogram) that the random variable assumes a value included in the j-th section ($j = 1, 2, \ldots, k$). The calculated empirical values of the cumulative frequency distribution for the histogram shown in Fig. 1.2 ($k = 8$) are given in Table 1.1.

The graph of the calculated quasi-stepwise cumulative frequency distribution for the quasi-stepwise random variable from Fig. 1.2 is shown in Fig. 1.3. The level of each step of that graph indicates the number of rivets whose head elevation above

Table 1.1 The empirical cumulative frequency distribution for the example from Fig. 1.2

l	1	2	3	4	5	6	7	8
w_l	0.033	0.073	0.167	0.267	0.193	0.160	0.073	0.033
$\sum_{j=1}^{l} w_j$	0.033	0.106	0.273	0.540	0.733	0.893	0.966	1.000

Fig. 1.3 The cumulative frequency distribution
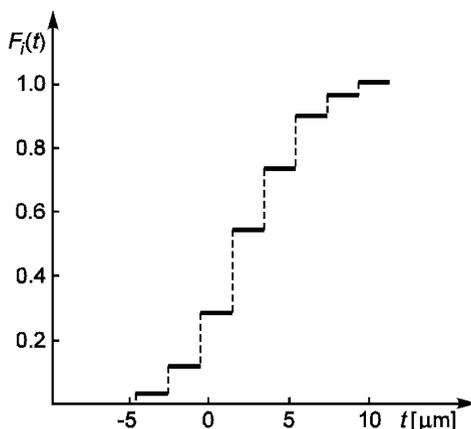
the reference plane is equal to or smaller than the magnitude represented by the right-hand limit point of the section under consideration. The value $F_l(t) = 1$ of the highest step ($l = 8$) indicates that all the measured deviations t were equal to or smaller than the right-hand end point of the last section.

1.5 Examples of Empirical Distributions

In Fig. 1.4 is presented a old and simple example of practical application of histograms for estimation of the accuracy of artillery fire (comp. [2]). Places hit by shells during the test are represented by points in the figure. The field covered by the hit points has been first divided into 22 horizontal strips of equal width. Then the number of hit points in each of the strips have been represented in the form of a histogram shown on the right-hand side of the figure. Next the field was divided into 15 vertical strips and the histogram shown in the lower part of the figure was prepared. Using such histograms, a preliminary estimation of the accuracy of the shooting distance X and of lateral deviation from the shooting direction Y could be found. This problem belongs to the class of two-dimensional distributions. The theory of such problems will be discussed in Chap. 5.

In Fig. 1.5 is presented the histogram of the age of persons who received the D.Sc. degree in the field of engineering sciences in Poland (years 1991...1993). The histogram has three peaks. The reason for this is the fact that the population under investigation is not homogeneous. The first group of scientists represented in

Fig. 1.4 The accuracy of artillery fire, after [2]
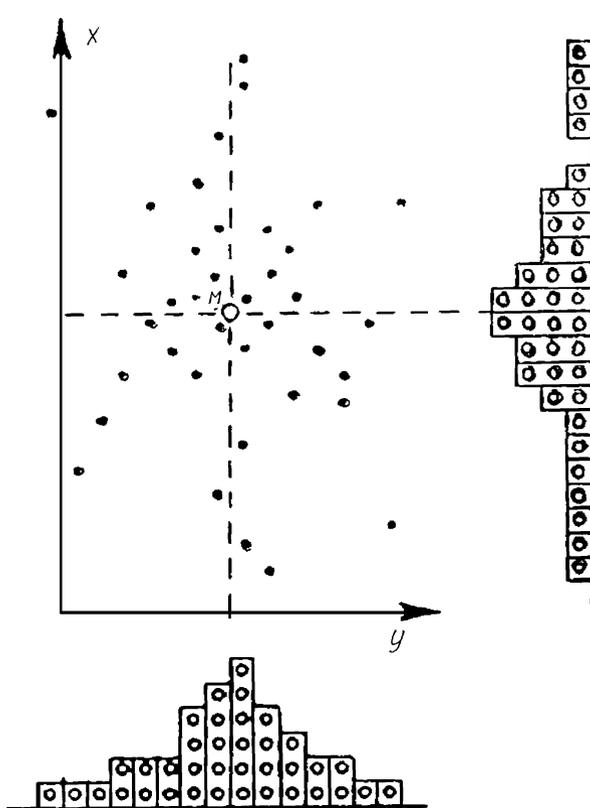
Fig. 1.5 was employed permanently at the universities and in research institutes and systematically received scientific degrees. The second peak corresponds to the individuals who had been awarded the associate professor position without the degree, and in view of new legal regulations were required to prepare their dissertations to maintain their positions. The last peak in Fig. 1.5 reflects the group of engineers involved in practical projects who wanted to finalize their professional career with the second doctor's degree (D.Sc.). As it is seen, even such a simple histogram may contain significant information about the population.

In Fig. 1.6 is presented a histogram of the lengths of 26713 codfish caught in the Baltic Sea in 1952 by Polish fishermen. It is made with a step of 1 cm. It is typical as a picture illustrating a dimension (height, length, thickness, etc.) of some random population: an average value with the maximum value of the histogram in it and the tails (longer or shorter) at the left-hand and the right-hand sides. However, a more detailed analysis of the histogram shows that the displayed population is in fact a sum of two populations: the males and the females of codfish, with different average lengths. (Two very close peaks of the histogram in Fig. 1.6 testify to this fact.)

In certain cases histograms may display significant irregularities of their structure. An example is shown in Fig. 1.7. Two histograms for the distribution of the
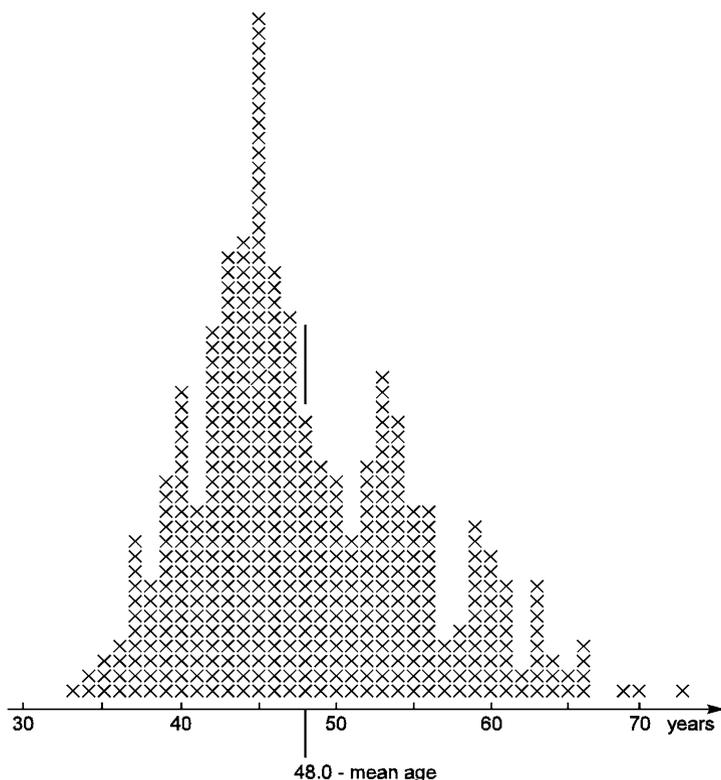
Fig. 1.5 The age of persons who received the D.Sc. degree

dilatation waves velocity in the bedrock below a dam under construction in Poland are presented (comp. [1]). For the two investigated strata the histograms are of different form. The form of one of them is close to the normal distribution, while for the other stratum the histogram has a two-modal form (comp. Sect. 3.2).

As another example the scheme of a histogram shown in Fig. 1.8 may be taken. It represents the distribution of the water level below the ground surface measured during several years of observations. Such measurements were performed, for example, by L. Réthai (comp. [3]).

1.6 Parameters Obtained from the Measured Data and Their Theoretical Values

In this chapter we have presented several examples of statistical measurements; we have calculated mean values and variances of the obtained results and plotted histograms on the basis of the measured data. Let us recall two of the examples.
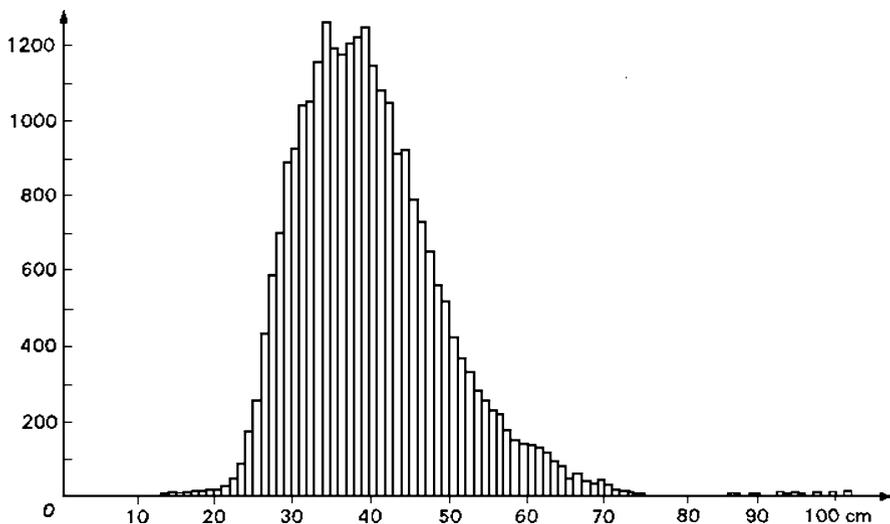
Fig. 1.6 Length of codfish caught in Baltic Sea in 1952 by Polish fishermen, see [4]

Consider at the beginning the problem illustrated by Fig. 1.5, concerning researchers obtaining the D.Sc. degree in engineering sciences. The histogram represents here all possible results of the considered phenomenon (all persons who obtained the degree are represented by crosses in the histogram). Thus, calculating e.g., the mean age of the population (according to (1.1)) or its variance (using (1.4a)), we obtain the exact values of these quantities.

We are faced with a different problem considering the example of rivets' head elevation. The rivets are made by an automatic machine, each with certain accuracy. The data used for preparation of the histogram was obtained by sampling an inspection lot. Knowing the data, we want to calculate the parameters (the mean value and the variance) which enable us to characterize the properties of the automatic machine and the parameters of all the rivets produced by the machine in the past and all which will be produced in the future. This means that having a finite set of measurements, we are looking for the mean value and the variance of the random variables describing all the possible results of measurements. Using expressions (1.1) for the sample average and (1.4a) for the sample variance, we deal with the random variables, depending on the choice of the inspection lot, which is now a random event. The quantities calculated in this procedure are not the exact values of the parameters searched for, because they are calculated on the basis of some limited set of the measured data (the sample) and not on the basis of all possible values of the parameters (the whole population). Therefore, they are called the *estimators* of the parameters, and the procedure of calculation is called parametric estimation.

All the information about the random variable is contained in its probability density. Before we define this function (in the next chapter), let us remark that the histogram presented in Fig. 1.1, obtained on the basis of a finite set of measurements, is a certain approximation of the probability density function of the inaccuracy of the

Fig. 1.7 The distribution of the dilatation waves velocity in the bedrock below a certain dam under construction in Poland, see [1]
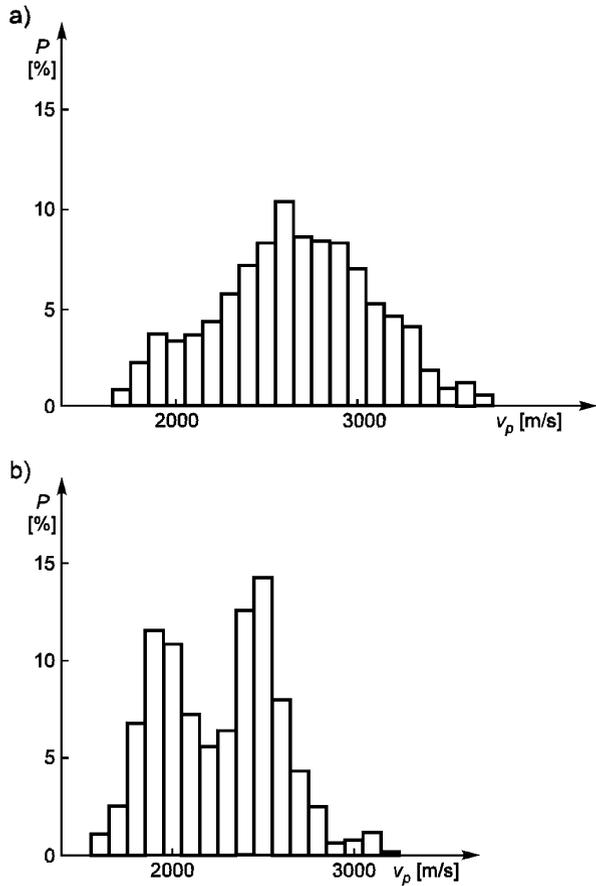
rivets' head elevation. In other words, the histogram is the estimator of the probability density and the procedure of obtaining such approximations is a non-parametric estimation (non-parametric, because its result is not a parameter but a complete function).

In further considerations we shall distinguish two situations: one, when considering the parameters of a certain distribution, like the mean value and the variance, we calculate them from a given distribution (e.g., using the probability density function) and another, when we deal with estimators of these values (calculated from the measurements). More about estimators and estimation is provided in Chap. 10.

The introductory considerations of this chapter show not only the origin of the error calculus; they also indicate that in some advanced problems, if we want to obtain precise answers concerning the accuracy of technical devices, we must apply mathematical tools in formulation of the tasks and in the analysis procedures. The following chapters present such an approach to error calculus.

Fig. 1.8 The distribution of the water level below the ground surface measured during several years of observations, see [3]

GROUND SURFACE

$W_i = n_i/n \cdot 100\%$
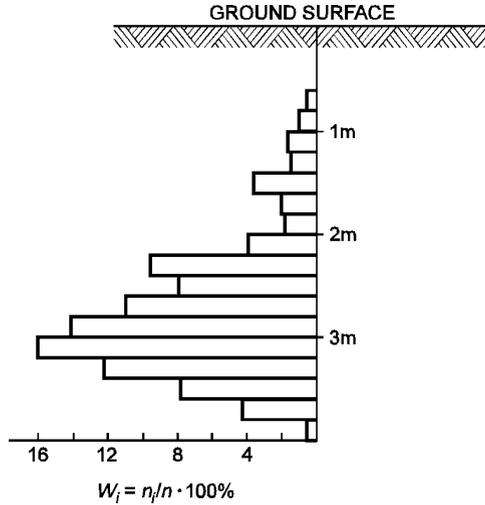
Problem 1.1 We change spacing on the scale of lengths of fish of the data presented in the histogram from Fig. 1.6 to 5 cm and obtain the following table of lengths:

Plot the histogram for the data given in Table 1.2. Change the spacing to a scale 20 cm and plot the histogram for the new data.

Problem 1.2 Using the histograms from Problem 1.1, plot the corresponding cumulative frequency distributions.

Problem 1.3 Calculate the average value, the variance and the standard deviation for both data sets studied in Problems 1.1 and 1.2. Observe the difference between the calculated values of parameters.

Table 1.2 The empirical cumulative frequency distribution for the example from Fig. 1.6

Interval	No.	Interval	No.
1−5	0	51–55	1700
6−10	0	56–60	950
11−15	25	61–65	650
16−20	50	66–70	225
21−25	275	71–75	125
26−30	2900	76–80	100
31−35	5300	81–85	25
36−40	6000	86–90	50
41−45	5100	91–95	25
46−50	3200	96–100	25

Problem 1.4 For the data studied in Problems 1.1–1.3 calculate the average deviation $\overline{d}$. Compare the result with the sample standard deviation s of the corresponding variable.

References

1. Bestyński, Z.: Geophysical properties of rocks. In: K. Thiel (ed.), Physico-Chemical Properties and Modeling of Rock Massifs of East-Carpathian Mountains, Hydrotechnics, No. 19, Institute of Hydro-Engineering, Polish Academy of Sciences (1995) (in Polish)
2. Hänert, L.: Geschütz und Schuß. Eine Einführung in die Geschützmechanik und Ballistik. Springer, Berlin (1928)
3. Rèthai, L.: Forecasting various ground water levels. Proc. 7th ICASP, Aachen, 1975.
4. Puchalski, T.: Statistics. Course of Fundamental Problems. Polish Scientific Editors, Warsaw (1977) (in Polish)

Chapter 2
Random Variables and Probability; Normal Distribution

2.1 Probability and Random Variables

We say that a phenomenon is random if, on the basis of our best knowledge, we cannot exactly predict its result. What really happens is only one of many possibilities. In every day life we meet random results of lotteries, a non-predictable dispersion of gun shots at a target or a random travel time through a crowded city.

An intuitively understandable idea of random phenomena can be formalized by the concept of random events and probability. Using formal definitions, we say that the *random event (a collection of sample points)* is a result of some random phenomena, and its *probability* is the chance that this phenomenon will occur, expressed by a number from the interval [0, 1].

Example 2.1 (An unbiased coin flipping) During an experiment of a single flip of an unbiased coin, two results are possible: the occurrence of heads and the occurrence of tails. Both results are random events. The probability of heads and the probability of tails are equal and they are 1/2.

In this example we can say that the set of all possible results of the experiment has two elements (occurrence of heads and occurrence of tails). We interpret the probabilities of occurrence of these elementary events in the following way: if we repeat flipping the coin a sufficient number of times, then the number of occurrences of heads (or, equivalently, the number of occurrences of tails) divided by the number of flips will tend to 1/2. This is the so-called frequency interpretation of the probability.

Example 2.2 (Dice casting) During an experiment of a single cast of an unbiased die, six results are possible: the occurrence of a face with $n = 1, 2, 3, 4, 5$, or 6 spots. Then the set of the results (elementary events, sample points) contains six elements. The probability of each event (the occurrence of a face with n spots) equals 1/6. This means that if the number of casts tends to infinity, then the following ratio:

$$\frac{\text{number of casts, when } n \text{ spots occurred}}{\text{total number of casts}}$$

tends to $1/6$, for $n = 1, 2, 3, 4, 5$, and 6.

The expected outcome of the experiment described in Example 2.2 may be more complicated than the occurrence of a fixed number of spots. For example, we can ask: What is the probability that the single cast results in a face with an even number of spots? What is the probability that we will see a face which has more than 4 spots? Of course, we can easily deduce that in the first case the probability is $1/2$ and in the second one $1/3$.

The above examples show that it is conceptually easy to define an event and the probability of an event if the number of possible outcomes of the experiment (e.g., coin flipping or dice casting) is finite and the outcome of each result is equally probable. In such a case, the probability of some event is defined as the frequency of occurrences of this event when the number of experiments tends to infinity.

In some situations, we can introduce another definition of probability. If the set of results of an experiment is infinite but it is contained in some set on a plane (alternatively: in 3-dimensional space, on a straight line, etc.), then the probability has a geometrical interpretation. The probability of a certain outcome of an experiment is the ratio of the area of the subset corresponding to these results, to the area of the set corresponding to all possible results of the experiment. The geometrical definition of probability has some limitations: the results of the experiment must be located in a bounded set on the plane and, moreover, they must be evenly distributed over this set.

The definitions of an event and the probability of an event used today have their origin in measure theory. The fundamental object of probability theory is the probability space. A *probability space* is defined by the triad (Ω, F, P), where Ω is the *sample space* containing all elementary events (sample points), F is the σ-algebra of Borel subsets of the sample space Ω containing all possible *events* (elementary and compound), and P is a (probability) measure defined on F.

We will now comment on the above definitions. Elementary events ω (being elements of the sample space Ω) are results of some experiment, mutually excluding each other; this means that only one elementary event can be the result of the experiment. Generally, (compound) events in an experiment are elements of the σ-algebra F. Occurrence of an event A can be the result of several elementary events; knowing the outcome of an experiment we are able to decide if the event A occurred. The probability measure P or, simply, the probability, has the property that it is equal to 1 for the certain event (the whole sample space Ω, that is, the event that the experiment had some outcome). Certainly, the probability of the impossible event (the empty set $\emptyset$) is zero.

Example 2.3 (An unbiased coin flipping, continuation) The probability space for the experiment of a single fair coin flip is (Ω, F, P), where: the sample space Ω is the following 2-element set:

$$\Omega = (\{\text{heads}\}, \{\text{tails}\});$$

the σ-algebra F consists of four elements: the empty set $\emptyset$, two 1-element sets, and the whole sample space Ω:

$$F = (\emptyset, \{\text{heads}\}, \{\text{tails}\}, \Omega);$$

the probability P is defined as:

$$P(\{\text{heads}\}) = \frac{1}{2}, \qquad P(\{\text{tails}\}) = \frac{1}{2}.$$

Example 2.4 (Dice casting, continuation) In the experiment of a single balanced die cast, the probability space is the following: the sample space Ω has 6 elements:

$$\Omega = (\{1 \text{ spot}\}, \{2 \text{ spots}\}, \{3 \text{ spots}\}, \{4 \text{ spots}\}, \{5 \text{ spots}\}, \{6 \text{ spots}\});$$

the σ-algebra F consists of the following elements: the empty set $\emptyset$, all subsets of Ω containing 1, 2, 3, 4, and 5 elements and the whole sample space Ω:

$$F = \begin{pmatrix} \emptyset, \ 6 \text{ one-element subsets, } 15 \text{ two-element subsets,} \\ 20 \text{ three-element subsets, } 15 \text{ four-element subsets,} \\ 6 \text{ five-element subsets, } \Omega \end{pmatrix};$$

the probability P is defined as:

$$P(\{1 \text{ spot}\}) = P(\{2 \text{ spots}\}) = P(\{3 \text{ spots}\})$$
$$= P(\{4 \text{ spots}\}) = P(\{5 \text{ spots}\}) = P(\{6 \text{ spots}\}) = \frac{1}{6}.$$

The concept of randomness and probability presented here identifies events with subsets of the sample space Ω, which are elements of the σ-algebra F. Therefore, we are able to perform on these events the operations analogous to the operations of set theory. For two events $A, B \in F$, we can define the union $A \cup B$ (A or B happens), intersection $A \cap B$ (A and B occur simultaneously), difference $A \backslash B$ (A occurs but B does not), etc. Probability, as we mentioned, is a measure; it has the following properties:

$$0 \le P(A) \le 1, \tag{2.1}$$

$$P(\emptyset) = 0, \qquad P(\Omega) = 1, \tag{2.2}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B), \tag{2.3}$$

and for a countable number of disjoint events A_j:

$$P\left(\bigcup_j A_j\right) = \sum_j P(A_i). \tag{2.4}$$

In probability theory, it is very important to know the relationship between events: their dependence or independence. We say that two events A and B are *independent* if their probabilities satisfy the following condition:

$$P(A \cap B) = P(A) P(B), \tag{2.5}$$

which means that the probability of the simultaneous occurrence of both events is equal to the product of probabilities of their separate occurrence. If condition (2.5) is not satisfied, the events A and B are *dependent*.

To know to what extent the events A and B are dependent, we can use the *conditional probability* $P(A|B)$, which is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{2.6}$$

The quantity $P(A|B)$, which is the probability of A conditioned on B, we understand to be the probability of occurrence of A under the condition that B has occurred.

Using formula (2.5) in (2.6), we see that if events A and B are independent then

$$P(A|B) = P(A). \tag{2.7}$$

The concept of the conditional probability is strongly related to the definition of *complete probability*. If we have some sequence of mutually excluding events $B_j,\ j = 1, 2, \ldots, n,\ B_k \cap B_l = \emptyset$ for $k \neq l$, satisfying additionally $\bigcup_j B_j = \Omega$, then the probability of any event A can be represented as

$$P(A) = \sum_j P(A|B_j) P(B_j). \tag{2.8}$$

The last equation enables us to calculate the probability of some event A if we know its probability under some additional conditions, that is, if we know that some event B_j has occurred.

Example 2.5 (Dice casting, continuation) Consider the experiment of the die single cast and define two events: A, the outcome is a face with an even number of spots, and B, the face with a number of spots greater than 4. We can verify whether these two events are independent. Using the elementary events defined in Example 2.4 we find that the events are: $A = (\{2 \text{ spots}\}, \{4 \text{ spots}\}, \{6 \text{ spots}\})$, $B = (\{5 \text{ spots}\}, \{6 \text{ spots}\})$, and their probabilities are: $P(A) = \frac{1}{2},\ P(B) = \frac{1}{3}$.

The intersection of the events is: $A \cap B = (\{6 \text{ spots}\})$, and the probability of intersection, $P(A \cap B) = \frac{1}{6}$. It is seen that the events A and B satisfy condition (2.5), that is, they are independent.

If we replace the event B with a new one: B_1—the number of spots is greater than 5 (that is, $B_1 = (\{6 \text{ spots}\})$ and $P(B_1) = \frac{1}{6}$), then the intersection of the events is $A \cap B_1 = (\{6 \text{ spots}\})$ and it is seen that the events A and B_1 are dependent, because $P(A)P(B_1) = \frac{1}{12}$ and $P(A \cap B_1) = \frac{1}{6}$, so the condition (2.5) is not satisfied.

The description of results of experiments or observations of random phenomena in terms really existing in these processes is very complicated. To make the modeling of the processes more convenient we can introduce the concept of a random variable.

The real-valued function $X(\omega)$ defined on the sample space Ω of random events ω is called a *random variable* if a pre-image[1] A of every interval of real numbers of the form $I = (-\infty, x)$ is a random event (an element of the σ-algebra F).

Probability P describing properties of random events can also describe random variables. It is transferred from the σ-algebra of events to the space of real-valued random variables by pre-images of the intervals I:

$$P(I) = P(\omega \text{ such that } X(\omega) < x). \tag{2.9}$$

For a given sample space we can consider various random variables. Our choice depends on the purpose of the modeling.

Example 2.6 (An unbiased coin flipping, continuation)

(a) Consider the experiment of a symmetric coin single flip. Assign number 1 to the outcome of heads and number -1 to the outcome of tails. Such a random variable may be used for description of a random walk on a straight line. We start from $x = 0$ and repeat the coin flipping. If the outcome is heads then we add 1 to x, if tails, we subtract 1. After every trial the value of x is greater by 1 or smaller by 1 than the value in the previous step. We repeat the trial many times obtaining the x-coordinate of the walking particle in every step (see, e.g., [11]).
(b) Consider the same experiment. We assign number 1 to heads and number 0 to tails. Repeating the trials many times and writing down the obtained numbers we generate random numbers in binary notation.

Analogously to the events, we can define independence of random variables. We will say that two random variables X and Y (defined on the probability space (Ω, F, P)) are *independent* if for all $x_1 \leq x_2$ and $y_1 \leq y_2$ the events of the form $\{\omega : x_1 \leq X(\omega) < x_2\}$ and $\{\omega : y_1 \leq Y(\omega) < y_2\}$ are independent.

The theorem concerning the complete probability (2.8) makes it possible to apply in many technical problems the so-called *conditioning technique*. This method is based on the procedure of decomposition of the initial complicated problem into a number of tasks easy to solve when we assume certain conditions to be satisfied with a certain probability. Then the simplified problems are solved and, finally, the general non-conditioned solution is obtained by averaging of the set of solutions with respect to the assumed probability distribution. Such a technique lets us calculate the parameters (e.g., moments) or distributions of random variables in various

---

[1] Assume, we have a function $X : \Omega \rightarrow R$ and let A be a subset of the set of real numbers R. The pre-image (or inverse image) A for the function X is a set $B \subset \Omega$, containing all the elements $\omega \in \Omega$ such that their image belongs to A, which means $X(\omega) \in A$. In such a case we write: $B = X^{-1}(A)$.

engineering problems. The reader can find more about this technique in the papers [12, 13] or the textbook [20].

2.2 The Cumulative Distribution Function; the Probability Density Function

Most problems of the error calculus arising in technological applications concern the analysis of random variables with continuous distributions. Random variables of such a nature may assume any value from a certain range. The *cumulative distribution function* (or: *probability distribution function*) $F(x)$ of any one-dimensional random variable X is defined by the expression[2]:

$$F(x) = P(X < x),\tag{2.10}$$

which means that the cumulative distribution function is defined as a function, the value of which for a given x is equal to the probability of an event that the random variable X is smaller than the number x.

The cumulative distribution function is defined for all real numbers and it is a non-decreasing, continuous on the left, function. Moreover, for x tending to minus infinity and plus infinity, it satisfies the following conditions:

$$F(-\infty) = 0, \qquad F(\infty) = 1.\tag{2.11}$$

The probability distribution function can be applied to the calculation of probabilities of the events related to the random variable X. For instance, the probability of an event that a random variable X belongs to the interval $[x_1, x_2)$ can be expressed by means of the probability distribution function (see Fig. 2.1):

$$P(x_1 \leq X < x_2) = F(x_2) - F(x_1).\tag{2.12}$$
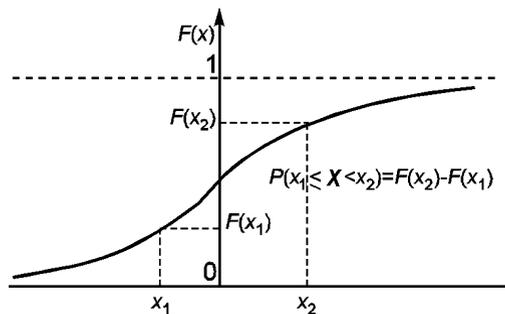
Fig. 2.1 The cumulative distribution function

---

[2]We shall denote random variables by capital letters X, Y, while their values, being numbers, will be denoted by small letters x, y. This does not refer to cases when a random variable in a particular formula has a physical meaning and is usually denoted by a small letter. $P(A)$ is the probability of an event A.

If the random variable X is *discrete*, that is, if it takes values from a finite (or countable) set: $\{x_j, j = 1, 2, \ldots, N\}$ (or $\{x_j, j = 1, 2, \ldots\}$), then the cumulative distribution function is discontinuous at these points and its jumps are equal to p_j. Moreover, the following equality holds:

$$P(X = x_j) = p_j. \tag{2.13}$$

Over the intervals of continuity, $x \in [x_j, x_{j+1})$, the cumulative distribution function $F(x)$ of the discrete random variable X is constant and equal to $F(x) = \sum_{k=1}^{j} p_k = F_j$. An example of the cumulative distribution function of some discrete random variable is presented in Fig. 1.3.

The cumulative distribution function of a random variable with a continuous distribution (the *continuous random variable*) may be expressed in the form of the integral

$$F(x) = \int_{-\infty}^{x} f(\xi)\, d\xi. \tag{2.14}$$

Function $f(x)$ in (2.14) is referred to as the *probability density function* (or simply *probability density*) of a random variable X. If the cumulative frequency distribution $F(x)$ has a derivative at any point x, then such a derivative represents the density

$$f(x) = F'(x). \tag{2.15}$$

Since the cumulative distribution function describes the normalized probability measure (the probability of the certain event equals 1, which means that $P(-\infty < X < \infty) = 1$) and is a non-decreasing function, the probability density function $f(x)$ has the following two properties:

$$A = \int_{-\infty}^{\infty} f(x)\, dx = 1 \tag{2.16}$$

and

$$f(x) \geq 0. \tag{2.17}$$

Thus, the area A between the graph of function $f(x)$ and the horizontal axis x of the random variable is equal to unity.

The probability of any event that the variable X lies in the interval $[x_1, x_2)$, which is, that it will have the value $P(x_1 \leq X < x_2)$, is defined by the following:[3]

$$P(x_1 \leq X < x_2) = \int_{x_1}^{x_2} f(x)\, dx. \tag{2.18}$$

The relation (2.18) is presented graphically in Fig. 2.2.

---

[3] For continuous distributions the probability that a random variable is located in a closed interval is the same as in an interval closed on one side or as in an open interval. In (2.18) we decided to choose an option of the interval closed on the left-hand side.

Fig. 2.2 The probability
density function
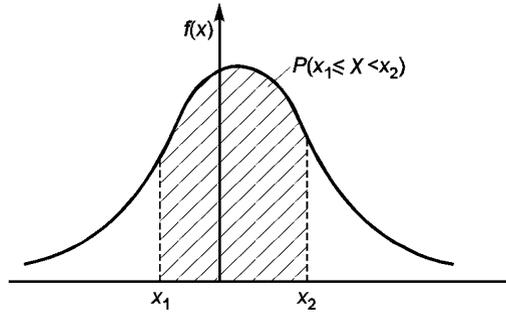
Fig. 2.3 The quantile
function, see (2.19)
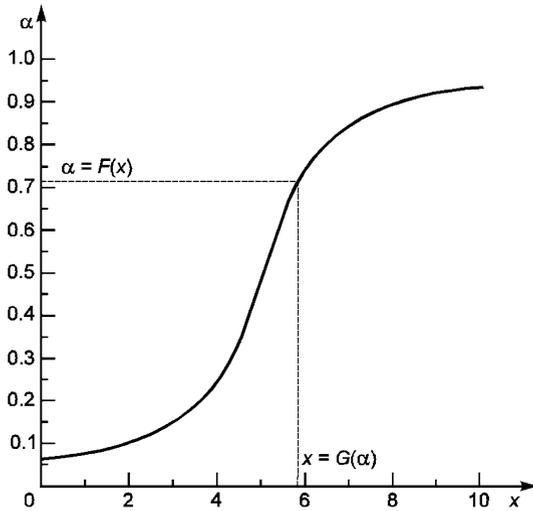
The cumulative distribution function and the probability density function are not the only functions characterizing a random variable. In some situations the inverse distribution function $G(\alpha)$, sometimes called the quantile function, is more convenient. For a given cumulative distribution function $F(x)$, the *quantile function* is defined as a function satisfying the following conditions:

$$x = G(\alpha) = G(F(x)),$$
$$P(X \le G(\alpha)) = \alpha. \tag{2.19}$$

This mutual relation between $F(x)$ and $G(\alpha)$ is shown graphically in Fig. 2.3.

In some applications of inspection theory and reliability theory, and also in some problems of mathematical statistics, *the survival function $S(x)$* is useful. It is defined as the probability that the random variable X is greater than or equal to x:

$$S(x) = P(X \ge x) = 1 - F(x). \tag{2.20}$$

More definitions of functions describing the properties of distributions of random variables can be found in handbooks dealing with probability theory or mathematical statistics (see, e.g., [6, 7]).

2.3 Moments

Moments play an important role in the error calculus, particularly when multidimensional problems are considered. For the one-dimensional random distributions discussed in this chapter, the expressions for moments take simple forms.

The *first-order moment* with respect to the line perpendicular to the x-axis and crossing it at $x = 0$ is defined by the formula

$$m = \int_{-\infty}^{\infty} x f(x) \, dx. \tag{2.21}$$

Assuming such a value $\overline{x}$ that the equality

$$A\overline{x} = m$$

holds true, we obtain, remembering that $A = 1$ (comp. (2.16)), the formula

$$\overline{x} = \int_{-\infty}^{\infty} x f(x) \, dx. \tag{2.22}$$

Using (2.22) one can calculate the average value $\overline{x}$. In other words, $\overline{x}$ represents the abscissa of the gravity center of the area between the graph of the probability density function and the x-axis. The moment m may be interpreted as the statical moment of that field with respect to the $x = 0$ straight line.

The *second-order moment* is the quantity J defined as

$$J = \int_{-\infty}^{\infty} (x - \overline{x})^2 f(x) \, dx. \tag{2.23}$$

Such a moment calculated with respect to the straight line $x = \overline{x}$ is called the *central second order moment*.

Assuming now a quantity σ^2 such that the equality

$$A\sigma^2 = J$$

holds true, we get, still remembering that $A = 1$, the relation

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \overline{x})^2 f(x) \, dx, \tag{2.24}$$

where σ^2 is the *variance* of the distribution $f(x)$. The square root of the variance, denoted by σ, represents the *standard deviation of the distribution*.

Note that the quantity J given by formula (2.23) is, in terms used in engineering applications, the central inertia moment of the area between the graph of the function $f(x)$ and the x-axis. Using such an interpretation it is seen that the standard deviation represents the so-called inertia radius of that field.

Of practical significance is also the *average deviation d* defined as

$$d = \int_{-\infty}^{\infty} |x - \overline{x}| \, f(x) \, dx. \tag{2.25}$$

The concept of the average value $\overline{x}$ may be generalized; in this way we obtain moments of order n, $n = 0, 1, 2, 3, \ldots$ (called the *ordinary moments* of n-th order), defined as:

$$m_n = \overline{x^n} = \int_{-\infty}^{\infty} x^n f(x) \, dx. \tag{2.26}$$

In the new notation the average value (or: the mean value) is the moment of order 1, namely m_1.

The generalization of the variance are *central moments* of order n, $n = 2, 3, 4, \ldots$, defined as:

$$\mu_n = \int_{-\infty}^{\infty} (x - m_1)^n \, f(x) \, dx. \tag{2.27}$$

Using definition (2.27) of the central moment we obtain the following relation between central moments and ordinary moments:

$$\mu_n = \int_{-\infty}^{\infty} (x - m_1)^n \, f(x) \, dx$$

$$= \int_{-\infty}^{\infty} \left(\sum_{j=0}^{n} (-1)^j \binom{n}{j} x^{n-j} m_1^j \right) f(x) \, dx = \sum_{j=0}^{n} (-1)^j \binom{n}{j} m_{n-j} m_1^j. \tag{2.28}$$

In particular, the variance σ^2 can be represented as:

$$\sigma^2 = \mu_2 = m_2 - m_1^2. \tag{2.29}$$

Except for the ordinary and central moments defined above, the *absolute moments* (*ordinary* and *central*), that is, average values of powers of the absolute value of x, can be defined by the following formulas:

$$m_n^{abs} = \int_{-\infty}^{\infty} |x|^n \, f(x) \, dx, \tag{2.30}$$

$$\mu_n^{abs} = \int_{-\infty}^{\infty} |x - m_1|^n \, f(x) \, dx. \tag{2.31}$$

The most often used absolute moment is the *average deviation d*, defined by (2.25).

Let us remark that for even values of n, the absolute moments and moments (ordinary and central) are identical.

Existence of moments is strongly connected with integrability of the probability density function $f(x)$ multiplied by some power of x. The condition of existence of the moment of a given order n is the convergence of the integral $\int_{-\infty}^{\infty} x^n f(x)dx$; from the existence of the moment for a certain given range $n = n_0$ we obtain the moments of lower orders. Therefore, the greatest n for which the moments exist is called the *range of the random variable*. In applications, the most often required assumption is that random variables have finite variances, that is, they are random variables of the second order.

Example 2.7 (The Cauchy distribution) The probability distribution with the probability density function

$$f(x) = \frac{1}{\pi b \{[(x-a)/b]^2 + 1\}} \tag{2.32}$$

and the cumulative distribution function

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x-a}{b}\right), \tag{2.33}$$

is called the Cauchy distribution. It is an example of distribution which has no moments (for each $n = 1, 2, \ldots$ the integral

$$m_n = \int_{-\infty}^{\infty} \frac{x^n dx}{\pi b \{[(x-a)/b]^2 + 1\}}$$

is divergent).

Example 2.8 (The normal distribution) The probability distribution with the probability density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(x-m)^2}{2\sigma^2}\right] \tag{2.34}$$

is called the normal distribution. It is an example of distribution which has moments of any order (for each $n = 1, 2, \ldots$ the integral

$$m_n = \int_{-\infty}^{\infty} \frac{x^n}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(x-m)^2}{2\sigma^2}\right] dx$$

is finite).

Remark 2.1 Assume that a certain random variable X has a finite mean value m_X and variance σ_X^2. Then we can consider the new random variable $\tilde{X}$, defined as

$$\tilde{X} = X - m_X$$

and called the centered random variable. This new random variable $\widetilde{X}$ (sometimes called the fluctuation of X) has zero average (mean) value and a variance equal to the variance of the original random variable X,

$$m_{\widetilde{X}} = 0, \qquad \sigma_{\widetilde{X}}^2 = \sigma_X^2.$$

Such decompositions of random variables are often applied in error analysis. In the above procedure we interpret the random variable X as the result of a measurement with some random error, the mean value m_X as the nominal value of the measured quantity, and the fluctuation $\widetilde{X}$ as the random measurement error itself.

2.4 The Normal Probability Distribution

The *normal distribution*, called also the *Gaussian distribution*, plays a basic role in error calculus. In most engineering applications random variables, such as small errors of measurements, small errors of positioning accuracy of certain mechanisms, e.g., robot manipulators or small deviations of magnitudes of certain parameters of objects in mass production, may be treated as those having normal probability distribution. They are called normal (Gaussian) random variables.

In the normal distribution, the probability density function takes the form [2]:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\overline{x})^2}{2\sigma^2}\right], \tag{2.35}$$

where $\overline{x}$ is the average value, comp. (2.22), and σ stands for the standard deviation (comp. (2.24)).

Introducing a new random variable

$$T = \frac{X-\overline{x}}{\sigma}, \tag{2.36}$$

which is called the *normalized random variable corresponding to X* (comp. [19]), we get another form of the probability density function,

$$\phi(t) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{t^2}{2}\right]. \tag{2.37}$$

Between the two forms of the probability density function, there exists the relation

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\overline{x})^2}{2\sigma^2}\right] = \frac{1}{\sigma}\phi(t), \quad t = \frac{x-\overline{x}}{\sigma}. \tag{2.38}$$

The numerical values of the normalized Gaussian distribution $\phi(t)$ may be calculated with the use of a computer or even a pocket calculator. Moreover, they are

Table 2.1 The probability density function $\phi(t)$ of the normalized Gaussian distribution

t	0	2	4	6	8
0.0	0.3989	0.3989	0.3986	0.3982	0.3977
0.1	0.3970	0.3961	0.3951	0.3939	0.3925
0.2	0.3910	0.3894	0.3876	0.3857	0.3836
0.3	0.3814	0.3790	0.3765	0.3739	0.3712
0.4	0.3683	0.3653	0.3621	0.3589	0.3555
0.5	0.3521	0.3485	0.3443	0.3410	0.3372
0.6	0.3332	0.3292	0.3251	0.3209	0.3166
0.7	0.3123	0.3079	0.3034	0.2989	0.2943
0.8	0.2897	0.2850	0.2803	0.2756	0.2709
0.9	0.2661	0.2613	0.2565	0.2516	0.2468
1.0	0.2420	0.2371	0.2323	0.2275	0.2227
1.1	0.2179	0.2131	0.2033	0.2036	0.1989
1.2	0.1942	0.1895	0.1849	0.1804	0.1758
1.3	0.1714	0.1669	0.1626	0.1582	0.1539
1.4	0.1497	0.1456	0.1415	0.1374	0.1334
1.5	0.1295	0.1257	0.1219	0.1182	0.1145
1.6	0.1109	0.1074	0.1040	0.1006	0.0973
1.7	0.0940	0.0909	0.0878	0.0848	0.0818
1.8	0.0790	0.0761	0.0734	0.0707	0.0681
1.9	0.0656	0.0632	0.0608	0.0584	0.0562
2.0	0.0540	0.0519	0.0498	0.0478	0.0459
2.1	0.0440	0.0422	0.0404	0.0387	0.0371
2.2	0.0355	0.0339	0.0325	0.0310	0.0297
2.3	0.0283	0.0270	0.0258	0.0246	0.0235
2.4	0.0224	0.0213	0.0203	0.0194	0.0184
2.5	0.0175	0.0167	0.0158	0.0151	0.0143
2.6	0.0136	0.0129	0.0122	0.0116	0.0110
2.7	0.0104	0.0099	0.0093	0.0089	0.0084
2.8	0.0079	0.0075	0.0071	0.0063	0.0063
2.9	0.0060	0.0056	0.0053	0.0050	0.0047
3.0	0.0044	0.0042	0.0039	0.0037	0.0035

tabulated in numerous books (comp. [9, 10]). To make this book sufficiently self-contained, the values are given in Table 2.1.[4] Knowing the function $\phi(t)$ and the standard deviation σ of a particular non-normalized normal distribution, we may

---

[4]The numbers 0, 2, 4, 6 and 8 in the heading of the table are values of the second fractional digit of the number t.
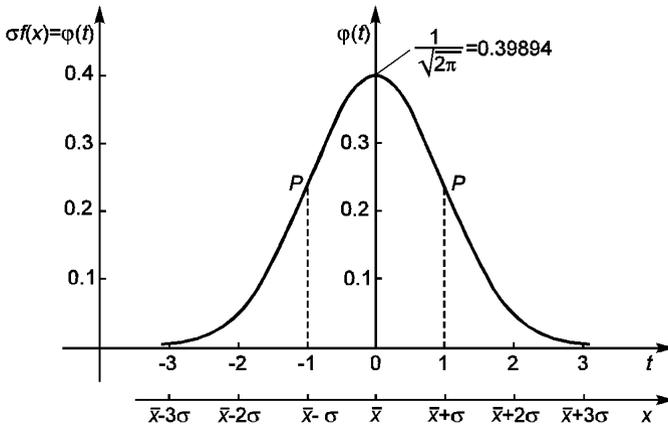
Fig. 2.4 The normalized probability density function of the normal (Gaussian) distribution

calculate by means of formula (2.38) the values of $f(x)$ for any value of the independent variable x. In practical calculations one can use the graph of the function $\phi(t)$ shown in Fig. 2.4. The graph has two inflexion points P, for $t = +1$ and for $t = -1$.

In Fig. 2.4 is also shown a simple graphical procedure allowing us to find the graph of the function $f(x)$ if the graph of the normalized density function $\phi(t)$ is given. The smaller is the standard deviation σ of the normal distribution, the smaller will be the dispersion of the random variable X around the average value $\overline{x}$. This property of normal distribution is illustrated in Fig. 2.5, in which three various normal distributions are presented. Their average value is of the same magnitude $\overline{x} = 0$, while standard deviations are different having the values $\sigma = 0.5$, $\sigma = 1.0$, and $\sigma = 2.0$, respectively.

The diagrams of normal probability densities are symmetrical with respect to the average value $\overline{x}$, at which they have a maximum. This maximum value of the density is given by the formula

$$f(\overline{x}) = \frac{1}{\sigma\sqrt{2\pi}}. \tag{2.39}$$

The relation between the half-width t_α of any arbitrarily chosen range $(-t_\alpha, t_\alpha)$ and the probability α that the random variable T takes the value located inside this range is of great practical significance. Some selected values of the pairs t_α, $(1 - \alpha)$ are collated in Table 2.2 (comp. Fig. 2.6). The quantity $(1 - \alpha)$ is called the *residual probability*.

In practice, certain specific ranges are often used, bounded by the multiplicities of the standard deviation σ, namely:

$$(-\sigma, \sigma), \quad \text{the probability is } \alpha = 0.6826,$$

$$(-2\sigma, 2\sigma), \quad \text{the probability is } \alpha = 0.9544,$$
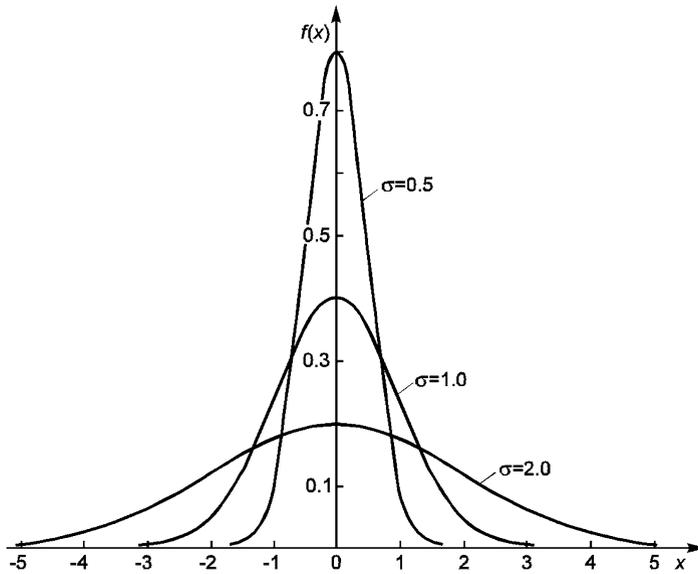
Fig. 2.5 The probability density function of the normal distribution for several values of the standard deviation

Table 2.2 The residual probabilities of the normal distribution

t_α	$1 - \alpha$
0.0	1.0000
0.5	0.6170
1.0	0.3174
1.5	0.1336
2.0	0.0456
2.5	0.0124
3.0	0.0027

$$(-3\sigma, 3\sigma), \quad \text{the probability is } \alpha = 0.9973.$$

These numbers indicate that the normal distribution of a random variable is concentrated in the vicinity of the average value $\overline{x}$. The probability that the value of a random variable X with the normal distribution differs from its average value by more than 3σ equals 0.0027. Such a significant property justifies to a certain degree the so-called *three-sigma rule*, that is, often used also in cases when other distributions are involved, not only when the normal distribution is considered. This rule should not, however, be used uncritically for any arbitrary probability distribution (comp. [6]).

Fig. 2.6 The residual
probability $(1 - \alpha)$
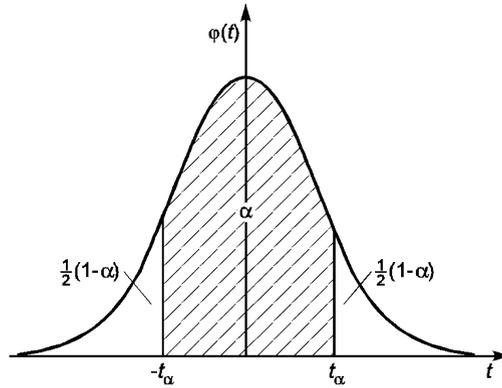
Fig. 2.7 The cumulative
distribution function of the
standard normal distribution
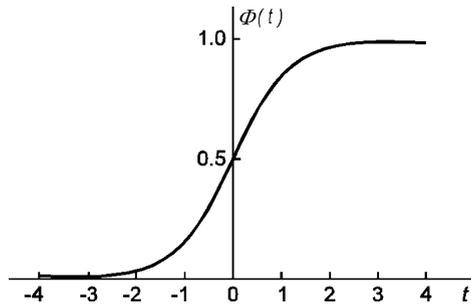
Between the standard deviation σ of the normal probability distribution and its average deviation d we have the following, sometimes useful, relation:

$$d = \sqrt{\frac{2}{\pi}}\sigma \approx 0.798\sigma. \tag{2.40}$$

The cumulative distribution function of the normal random variable is given by the formula

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{(\xi - \bar{x})^2}{2\sigma^2}\right] d\xi \equiv \Phi(t), \quad t = \frac{x - \bar{x}}{\sigma}, \tag{2.41}$$

where $\Phi(t)$ is the cumulative distribution function of the normalized Gaussian random variable:

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left[-\frac{1}{2}\tau^2\right] d\tau. \tag{2.42}$$

The function $\Phi(t)$ is tabulated, see Table 2.3; its graph is presented in Fig. 2.7. This distribution is sometimes called the *standard normal distribution* $N(0, 1)$.

Table 2.3 The cumulative distribution function of the normalized Gaussian distribution

t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$
-3.00	0.0013	-1.75	0.0401	-0.50	0.3085	0.75	0.7734	2.00	0.9773
-2.95	0.0016	-1.70	0.0446	-0.45	0.3264	0.80	0.7881	2.05	0.9798
-2.90	0.0019	-1.65	0.0495	-0.40	0.3446	0.85	0.8023	2.10	0.9821
-2.85	0.0022	-1.60	0.0548	-0.35	0.3632	0.90	0.8159	2.15	0.9842
-2.80	0.0026	-1.55	0.0606	-0.30	0.3821	0.95	0.8289	2.20	0.9861
-2.75	0.0030	-1.50	0.0668	-0.25	0.4013	1.00	0.8413	2.25	0.9878
-2.70	0.0035	-1.45	0.0745	-0.20	0.4207	1.05	0.8531	2.30	0.9893
-2.65	0.0040	-1.40	0.0808	-0.15	0.4404	1.10	0.8643	2.35	0.9906
-2.60	0.0047	-1.35	0.0885	-0.10	0.4602	1.15	0.8749	2.40	0.9918
-2.55	0.0056	-1.30	0.0968	-0.05	0.4801	1.20	0.8849	2.45	0.9929
-2.50	0.0062	-1.25	0.1056	0.00	0.5000	1.25	0.8944	2.50	0.9938
-2.45	0.0071	-1.20	0.1151	0.05	0.5199	1.30	0.9032	2.55	0.9946
-2.40	0.0082	-1.15	0.1251	0.10	0.5398	1.35	0.9115	2.60	0.9953
-2.35	0.0094	-1.10	0.1357	0.15	0.5596	1.40	0.9192	2.65	0.9960
-2.30	0.0107	-1.05	0.1469	0.20	0.5793	1.45	0.9265	2.70	0.9965
-2.25	0.0122	-1.00	0.1587	0.25	0.5987	1.50	0.9332	2.75	0.9979
-2.20	0.0139	-0.95	0.1711	0.30	0.6179	1.55	0.9394	2.80	0.9974
-2.15	0.0158	-0.90	0.1841	0.35	0.6368	1.60	0.9452	2.85	0.9978
-2.10	0.0179	-0.85	0.1977	0.40	0.6554	1.65	0.9505	2.90	0.9981
-2.05	0.0202	-0.80	0.2119	0.45	0.6736	1.70	0.9554	2.95	0.9984
-2.00	0.0227	-0.75	0.2266	0.50	0.6915	1.75	0.9599	3.00	0.9987
-1.95	0.0256	-0.70	0.2420	0.55	0.7088	1.80	0.9641	–	–
-1.90	0.0287	-0.65	0.2578	0.60	0.7257	1.85	0.9678	–	–
-1.85	0.0322	-0.60	0.2743	0.65	0.7422	1.90	0.9713	–	–
-1.80	0.0359	-0.55	0.2912	0.70	0.7580	1.95	0.9744	–	–

In practical calculations often the so-called *error function*,

$$\mathrm{erf}(t) = \frac{1}{\sqrt{2\pi}} \int_0^t \exp\left[-\frac{1}{2}\tau^2\right]d\tau, \tag{2.43}$$

is used instead of the cumulative distribution function of the normalized Gaussian random variable. The error function is tabulated and given in various books (comp., e.g., [1, 10]). Its values are also given in Table 2.4.

The error function is directly connected with the cumulative distribution function $\Phi(t)$ by the simple formulas:

$$\Phi(t) = \frac{1}{2} - \mathrm{erf}(-t) \quad \text{for } t \le 0,$$

$$\Phi(t) = \frac{1}{2} + \mathrm{erf}(t) \quad \text{for } t > 0. \tag{2.44}$$

Table 2.4 The error function

t	erf(f)	t	erf(t)	t	erf(t)
0.00	0.0000	1.00	0.3413	2.00	0.4773
0.05	0.0199	1.05	0.3531	2.05	0.4798
0.10	0.0398	1.10	0.3643	2.10	0.4821
0.15	0.0596	1.15	0.3749	2.15	0.4842
0.20	0.0793	1.20	0.3849	2.20	0.4861
0.25	0.0987	1.25	0.3944	2.25	0.4878
0.30	0.1179	1.30	0.4032	2.30	0.4893
0.35	0.1368	1.35	0.4115	2.35	0.4906
0.40	0.1554	1.40	0.4192	2.40	0.4918
0.45	0.1736	1.45	0.4265	2.45	0.4929
0.50	0.1915	1.50	0.4332	2.50	0.4938
0.55	0.2088	1.55	0.4394	2.55	0.4946
0.60	0.2257	1.60	0.4452	2.60	0.4953
0.65	0.2422	1.65	0.4505	2.65	0.4960
0.70	0.2580	1.70	0.4554	2.70	0.4965
0.75	0.2734	1.75	0.4599	2.75	0.4979
0.80	0.2881	1.80	0.4641	2.80	0.4974
0.85	0.3023	1.85	0.4678	2.85	0.4978
0.90	0.3159	1.90	0.4713	2.90	0.4981
0.95	0.3289	1.95	0.4744	2.95	0.4984
–	–	–	–	3.00	0.4987

The error function erf(t) is a special function and has no representation in the form of a combination of elementary functions. However, in certain books one can find approximate expressions allowing one to calculate the values of that function by means of elementary functions. Two such practical methods are presented below. They are based on the asymptotic expansions (comp. [1, 7]).

Method 2.1

$$\text{erf}(t) = 1 - \left(a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4 + a_5 z^5\right) \exp\left[-t^2\right] + \varepsilon(t), \qquad (2.45)$$

where

$$z = \frac{1}{1 + pt}, \qquad |\varepsilon(t)| \leq 1.5 \times 10^{-7}$$

and

$$p = 0.3275911, \qquad a_1 = 0.254829592,$$
$$a_2 = -0.284496736, \qquad a_3 = 1.421413741,$$
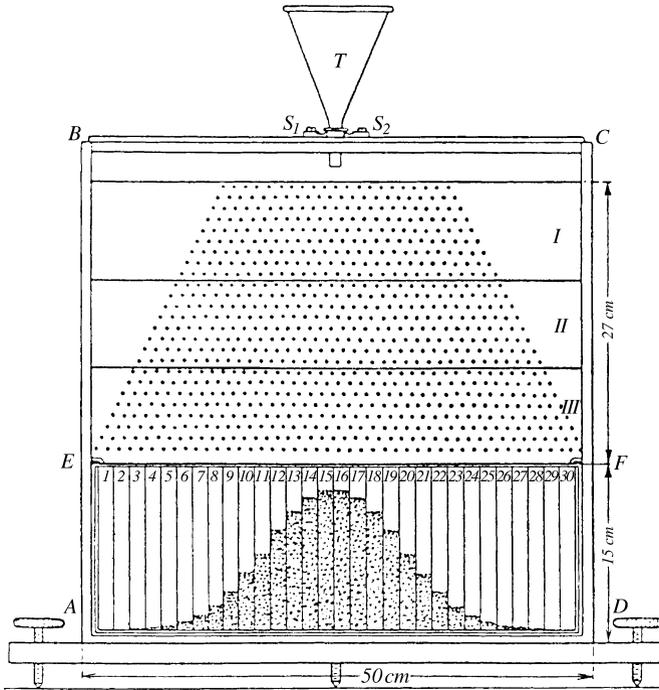
Fig. 2.8 Experimental generation of the normal distribution, the Galton box

$$a_4 = -1.453152027, \qquad a_5 = 1.061405429.$$

Method 2.2

$$\text{erf}(t) = 1 - \frac{1}{(a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5 + a_6 t^6)^{16}} + \varepsilon(t), \qquad (2.46)$$

where

$$|\varepsilon(t)| \leq 3 \times 10^{-7}$$

and

$$a_1 = 0.0705230784, \qquad a_2 = 0.0422820123,$$
$$a_3 = 0.0092705272, \qquad a_4 = 0.0001520143,$$
$$a_5 = 0.0002765672, \qquad a_6 = 0.0000430638.$$

The numerical values of the error function calculated according to each of these approximate formulas are often more accurate (the accuracy order 10^{-7} for all $x \in [0, \infty)$) than those given in the popular textbooks.

For clarity, the generation of the normal distribution can be demonstrated by using simple devices, such as that shown in Fig. 2.8 (cf. [5]). Small metal balls
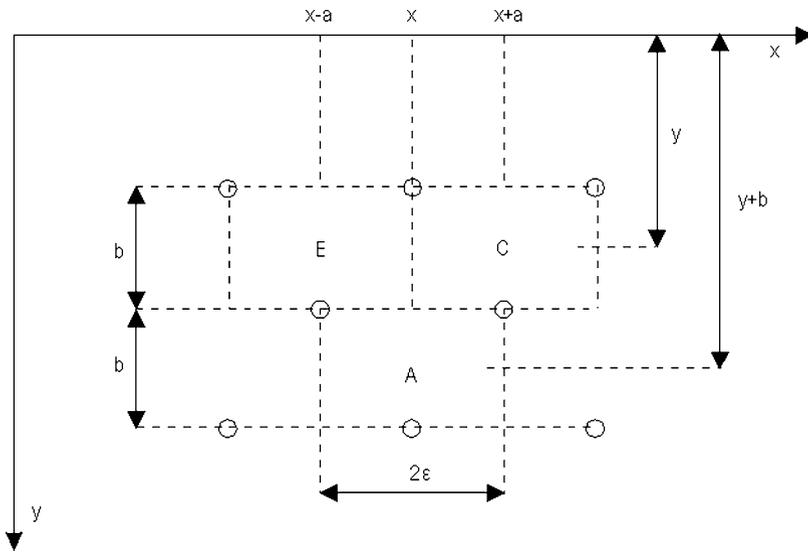
Fig. 2.9 The scheme of cells in the Galton box

falling down from a container T and striking numerous metal pins are randomly directed to the right or to the left. Finally, they fall at random into one of the separate small containers at the bottom of the device. The distribution of the number of balls in consecutive containers is close to the normal distribution. Similar examples may be found in [21].

Such a result of this educational experiment may be interpreted in two ways. From the mathematical point of view we can say that the normal distribution is formed as the consequence of the so-called central limit theorem, comp. [6]. Each ball falling down is randomly directed to the right or to the left, suffering a unit displacement with the same probability. Its final location in a specific container at the bottom is the sum of such displacements, which is the sum of independent random variables.

The fact that the distribution obtained in such an experimental device tends to the normal distribution may be proved by simple calculus, see [17]. Let us consider an arbitrary set of three cells A, B, C separated from the device in Fig. 2.8, and, moreover, let us assume that the probability distribution in the model may be described by a continuous function $P(x, y)$, if the distances between the pins are tending to zero ($a \to 0$ and $b \to 0$). The configuration of these separated cells is shown in Fig. 2.9.

Let the probabilities that a moving downwards ball falls to the cell B or C are:

$$P(x - a, y) \quad \text{and} \quad P(x + a, y),$$

respectively. Hence, we may express the momentary probabilities of migration of balls to the cell A located in the lower layer (marked as $y + b$), using the formula

for the complete probability, see (2.8). We obtain:

$$P(x, y + b) = \frac{1}{2} P(x - a, y) + \frac{1}{2} P(x + a, y). \tag{2.47}$$

Then, using Taylor's expansion of all terms of (2.47) around the point (x, y), we can write:

$$P(x, y + b) - P(x, y) = b \frac{\partial P(x, y)}{\partial y} + \frac{1}{2} b^2 \frac{\partial^2 P(x, y)}{\partial y^2} + \cdots,$$

$$P(x - a, y) - P(x, y) = -a \frac{\partial P(x, y)}{\partial x} + \frac{1}{2} a^2 \frac{\partial^2 P(x, y)}{\partial x^2} + \cdots,$$

$$P(x + a, y) - P(x, y) = a \frac{\partial P(x, y)}{\partial x} + \frac{1}{2} a^2 \frac{\partial^2 P(x, y)}{\partial x^2} + \cdots.$$

Substituting the above equations in (2.47) and decreasing the dimensions of the cells to a zero limit in such a way that simultaneously two conditions are satisfied:

$$a \to 0, \qquad b \to 0, \quad \text{and} \quad \frac{a^2}{2b} = D = \text{const.}, \tag{2.48}$$

we obtain the following partial differential equation for the probability density function $P(x, y)$:

$$\frac{\partial P(x, y)}{\partial y} - D \frac{\partial^2 P(x, y)}{\partial y^2} = 0. \tag{2.49}$$

Equation (2.49), obtained in [17], is of the same type as the equation of conduction of heat in solids, cf., e.g., [4]. Its solution can be written in the form

$$P(x, y) = \frac{\beta}{\sqrt{y}} \exp \left[-\frac{x^2}{4Dy} \right]. \tag{2.50}$$

To make the solution $P(x, y)$ of (2.50) to be a probability density function we take the parameter β such that the integral with respect to x of the right-hand of (2.50), for each $y = \text{const.}$ is equal to 1,

$$\int_{-\infty}^{\infty} P(x, y) \, dx = 2\beta \sqrt{\pi D} = 1,$$

from which it follows that

$$\beta = \frac{1}{2\sqrt{\pi D}}. \tag{2.51}$$

Thus, the solution to (2.49) can be written as

$$P(x, y) = \frac{1}{2\sqrt{\pi Dy}} \exp \left[-\frac{x^2}{4Dy} \right], \tag{2.52}$$

Fig. 2.10 Approximation of
a histogram presented in
Fig. 1.2 by the normal
probability density function
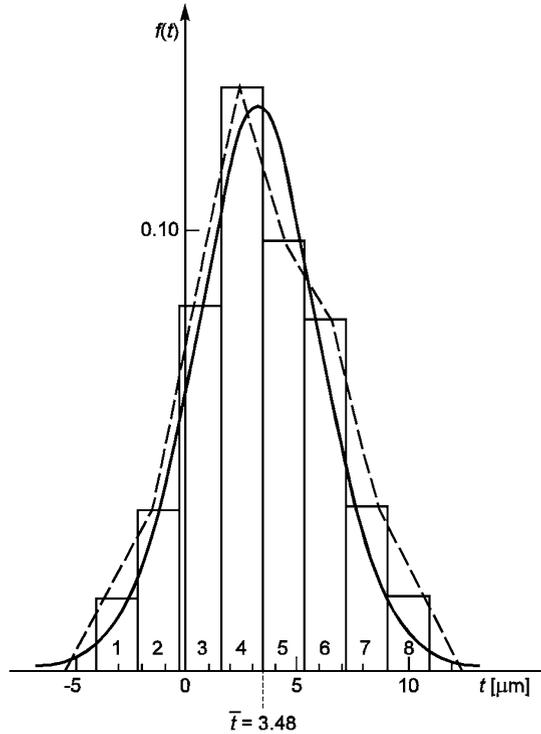
and after substitution

$$\sigma = \sqrt{2Dy} \tag{2.53}$$

as

$$P(x, y) \equiv f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{x^2}{2\sigma^2}\right]. \tag{2.54}$$

Comparing the obtained expression (2.54) with the known probability density
function of the normal distribution (2.35) of a zero mean value ($\bar{x} = 0$) we obtain
an argument that the probability distribution, which is a result of random symmetric
(that is with probability $\frac{1}{2}$ in every side) reflections of balls on pins of the Galton
box presented in Fig. 2.8, is really the normal distribution.

This result can be also interpreted more generally: we deal with the normal prob-
ability distribution of a random variable when this variable is influenced by numer-
ous independent factors. Such an interpretation explains why the normal distribution
corresponds so well to the distribution of errors of measurements, which usually
arise as a result of numerous unknown external factors.

Let us now consider an example of application of the continuous normal dis-
tribution to the description of the quasi-stepwise distribution shown in the form of
the histogram presented in Fig. 1.2. For the quasi-stepwise distribution, the average
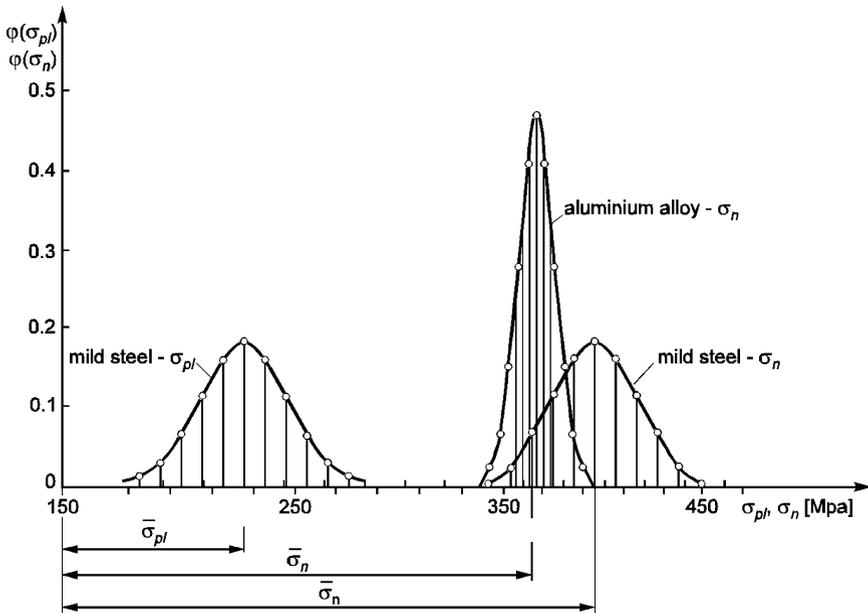
Fig. 2.11 Examples of distributions of certain mechanical properties of metal alloys, see [8]

value and the standard deviation are equal, respectively, to

$$\bar{t} = 3.48 \ \mu m, \qquad \sigma = 3.11 \ \mu m.$$

The diagram of the normal distribution calculated for such values of $\bar{t}$ and σ is presented in Fig. 2.10 along with the transformed original histogram. The area below the upper stepwise boundary of the transformed histogram equals unity. It is seen that this stepwise boundary corresponds fairly well to the graph of the normal probability distribution.

As another example, Fig. 2.11 shows distributions of the yield locus σ_{pl} and limit stress σ_n under uniaxial tension of a steel sheet 2 mm thick, measured during a tension test on 330 specimens cut out at various places of the same sheet, comp. [8]. For an aluminum alloy sheet, results of similar tests also shown in the figure display much smaller dispersion of the limit stress.

Random distributions of mechanical properties observed even in one large piece of a material contribute to the so-called scale effect: large specimens display smaller limit stress and yield locus than small specimens made of the same material.

Another example of a practical application of the normal distribution to the description of the cohesion c of soils is presented in Fig. 2.12. The figure was prepared on the basis of the experimental results given in [18].
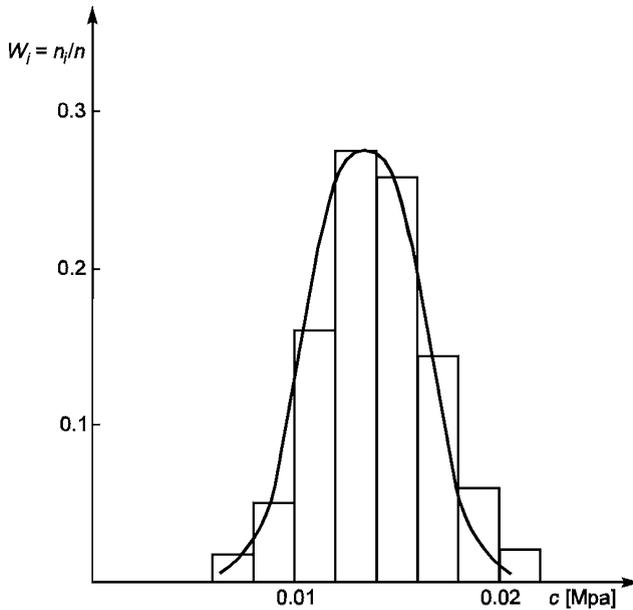
Fig. 2.12 Application of the normal distribution to the description of the cohesion c of soils, see [18]

2.5 Two-dimensional Gravity Flow of Granular Material

Before giving more information concerning probability distributions, let us analyze an example showing that using even the elementary theory of probability one can solve numerous problems of real practical significance. In the papers [15–17] J. Litwiniszyn ingeniously analyzed the inverse problem in which the cavities in a bulk of a loose material move randomly upwards from the bottom. To illustrate his idea, let us consider a two-dimensional problem of a relatively wide container with an outlet at the middle of the bottom. Fig. 2.13 shows the assumed initial system of finite cells analogous to that shown previously in Fig. 2.9. The width to height ratio of cells connected with the parameter (2.48) should be determined experimentally for the granular medium in question; for details see [23].

  A portion of the loose medium has just now left cell A leaving an empty space in it. The cavity in A formed in such a manner migrates upwards. We assume, as in the inverse problem shown in Fig. 2.8 that each time a portion of that cavity moves upwards, the probability of migrating into the right-hand or into the left-hand cell lying just above is equal to $\frac{1}{2}$. It means that at the beginning of the migration process, one half of the initial cavity A moves to the cell B and the other half is shifted to the cell C. If the volume of each cell is assumed to be a unit volume, the numbers in consecutive cells indicate how large a portion of the initial unit volume A passed through the cell during the migration process. Since after migration each portion of empty space must be filled by the granular medium falling downwards,
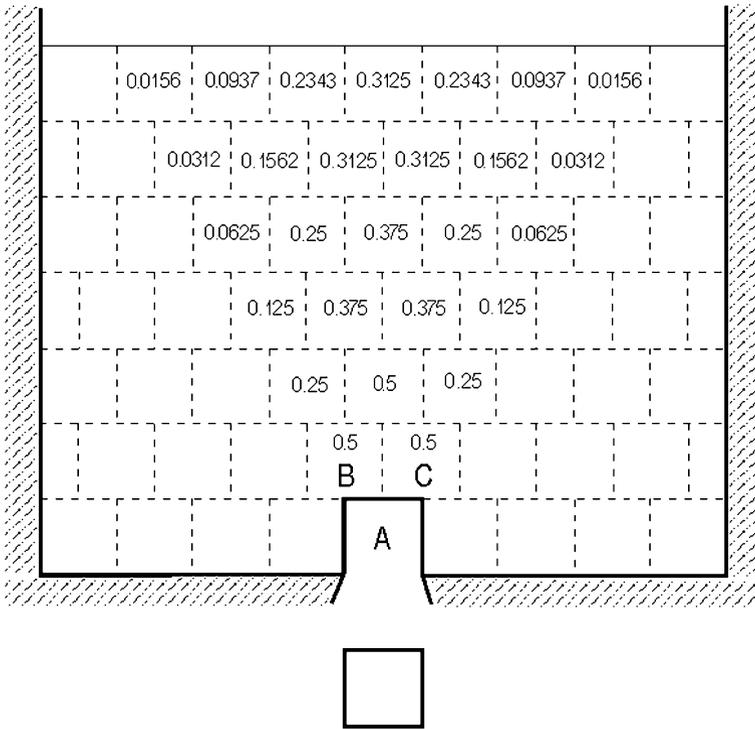
Fig. 2.13 Assumed system of cells for the problem of gravity flow from a bin

these numbers correspond to the average vertical displacement of the medium in particular cells. These vertical displacements are represented in Fig. 2.14. However, each particle of the medium is displaced also horizontally.

Below is presented a simple approximate method of determining total displacements [22]. Let us analyze an arbitrary set of three adjacent cells taken from the system of cells shown in Fig. 2.13. They are represented in Fig. 2.15a. The numbers in them correspond to the fraction of the initial volume of the cavity A, which passed through the cell during the migration towards the free surface of the bulk of the medium. According to the finite cells methodology, only one half of these fractions migrates from each cell A and B to the cell C. It is assumed that this migration takes place along the respective lines $A - C$ or $B - C$ joining central points of the cells. Directions and magnitudes of these migrating portions of the cavity may be represented by vectors W_{BC} and W_{AC} as shown in Fig. 2.15b. They may be treated as components of the resulting vector W_{cav} representing the direction and the magnitude of the averaged momentary flux of the cavity into cell C during the migration process. The opposite vector W_{mat} may be treated as a representation of the flux of the mass of granular medium filling the space left by cavities moving upwards. In order to calculate the magnitude of the averaged displacement vector u of the particles of the medium, it is assumed that its direction coincides with the direction
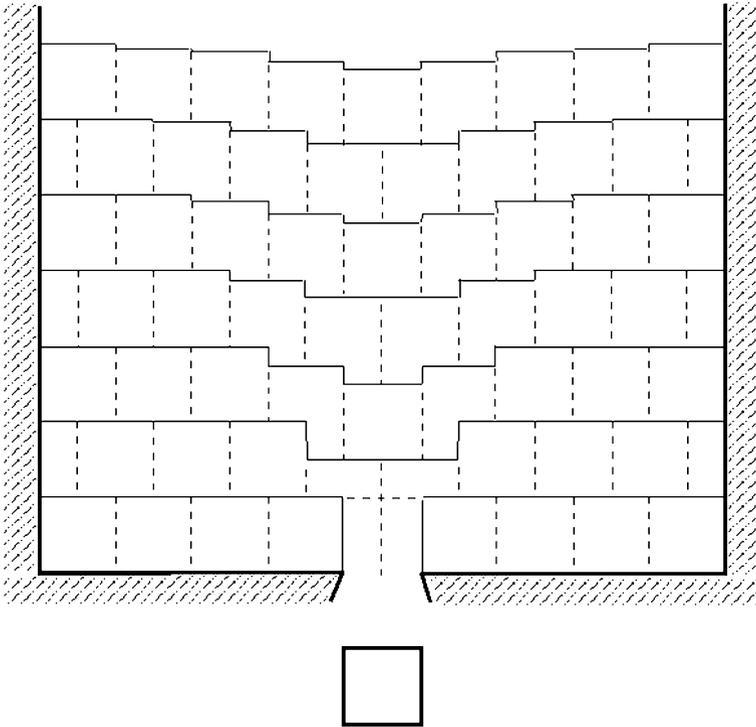
Fig. 2.14 Vertical displacement of granular material in cells

of the vector W_{mat}. To make this procedure consistent with that described before, it is assumed that the vertical component of the displacement vector u is equal to the vertical displacement of the respective sector of the stepwise deformed boundary between the rows of cells (cf. Fig. 2.14). Using this approximate procedure, the vectors of displacements have been calculated for the problem shown in Fig. 2.13 and Fig. 2.14. Results are shown in Fig. 2.16.

In Fig. 2.17 is presented an analogous solution for prediction of the movements of a crowd in a relatively narrow exit [14].

Figure 2.18 shows the theoretical field of displacements vectors calculated in the manner described above.

In order to verify experimentally such a theoretical motion pattern, a preliminary simple experimental simulation model composed of an assembly of coins of three different diameters has been used. The initial configuration of the assembly corresponding to the theoretical problem shown in Fig. 2.17 is presented in Fig. 2.19. The coins are located on a glass plate in the initial horizontal position. Then the plate is inclined with respect to the horizontal plane and the coins begin to slide downwards due to the gravity forces. This movement is disturbed by random mutual contacts between neighbors. The final configuration of displaced coins is shown in Fig. 2.20.

Fig. 2.15 Calculation of displacements of the granular material in cells, after [22]

a)

b)

W_{CAV}

$W_{BC} = 0.1875$

0.25

C

$W_{AC} = 0.0625$

0.125   0.375

A   B

W_{MAI}
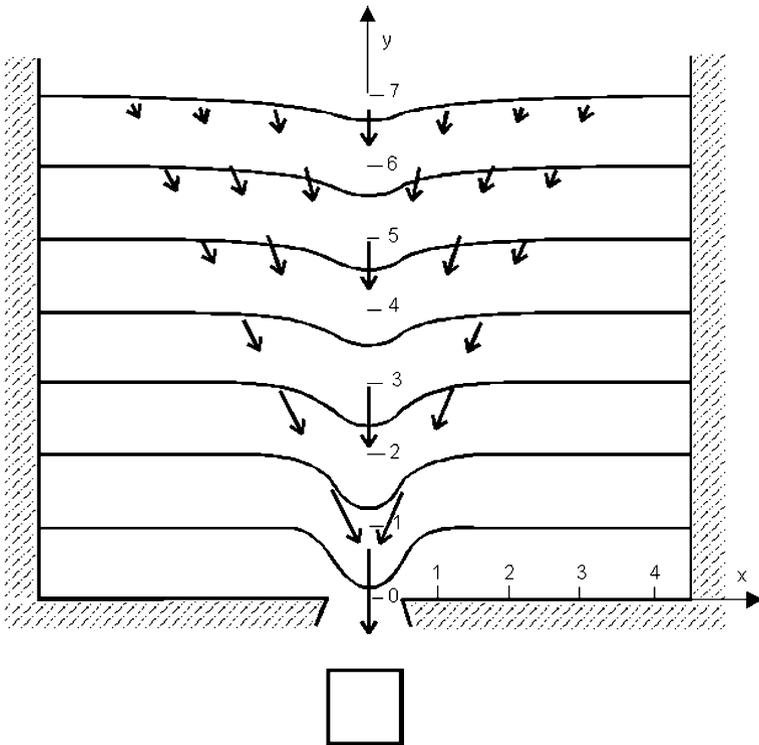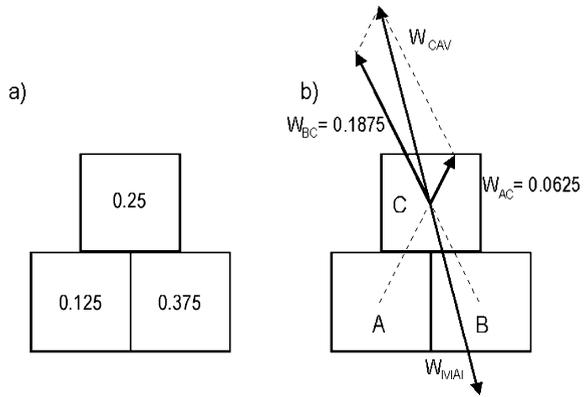
Fig. 2.16 Calculated displacements of granular medium in a bin, after [22]

The experiment was performed in three stages. In each stage one of the blocking strips at the bottom was removed. For each stage displacements of particular coins were measured. They are shown in Fig. 2.21. The stochastic nature of the movements of coins is visible. Let us notice, however, that their general layout is close to that shown previously in Fig. 2.18.
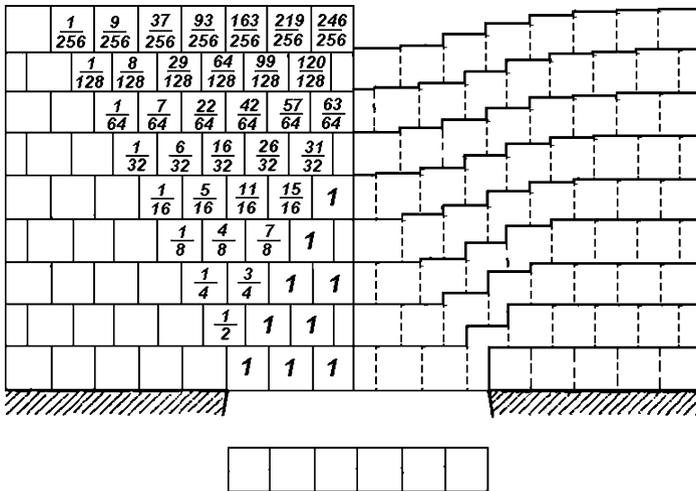
Fig. 2.17 Assumed system of finite cells and vertical displacements in a crowd in narrow exits, see [14]
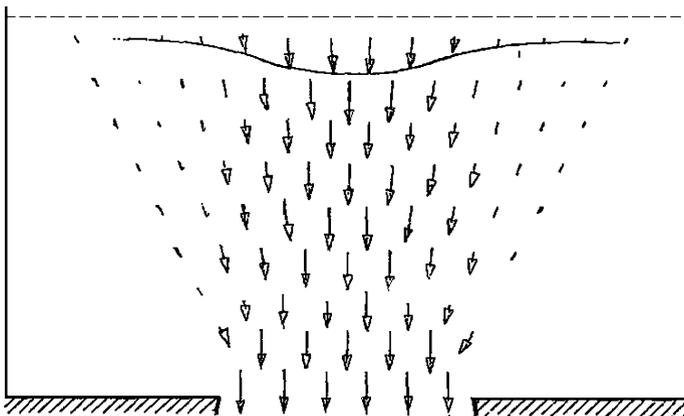
Fig. 2.18 Calculated displacements of a crowd in a narrow exit, see [14]

The next example concerns the problem of terrain subsidence caused by subterranean exploitation. The solution is shown in Fig. 2.22 (cf. [24]).

In the lower part of the soil resting on a bedrock, the empty space $A - B - C - D$ has been left by underground exploitation. In the following process of subsidence this empty space will be filled by the soil migrating downwards. Let us divide this empty space into a number of cells, each of them being of unit volume. These unit cavities migrate upwards through the system of cells shown in the figure. It is assumed that each time a cavity in the particular cell migrates upwards, the probability that it moves to the left or to the right cell, just above it, is equal to $1/2$. Numbers shown in particular cells indicate how large was the portion of a unit cavity which
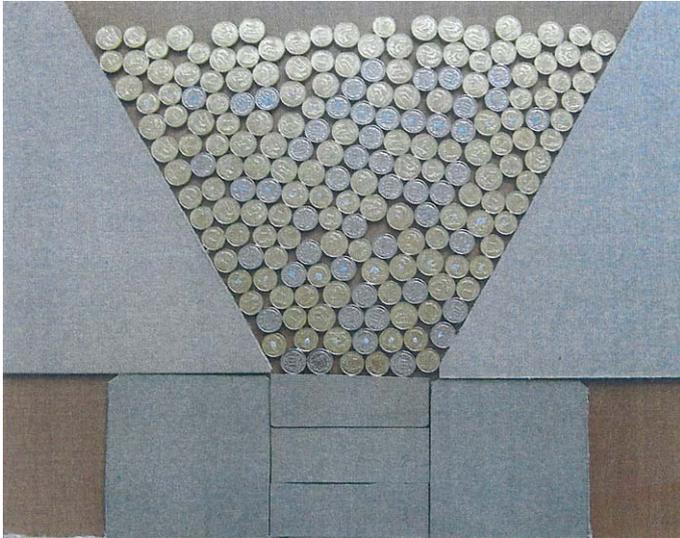
Fig. 2.19 Initial configuration of coins located on a glass plate, see [14]
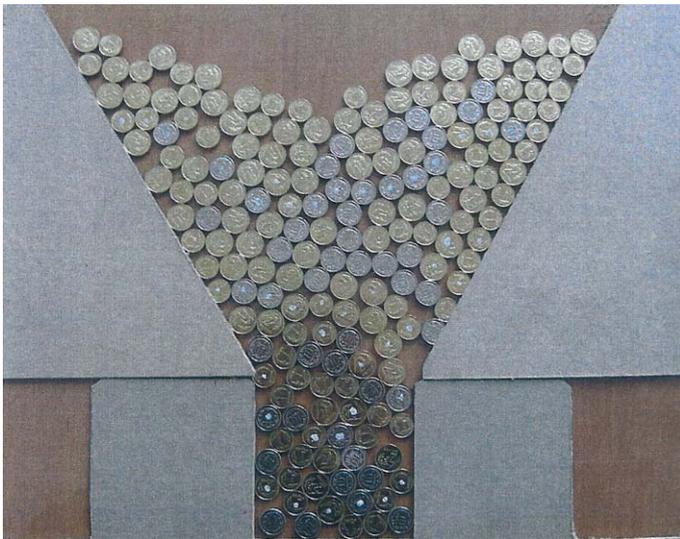
Fig. 2.20 Final configuration of coins in an experimental simulation of movements of a crowd, see [14]

has passed through the cell during the migration process. On the basis of these numbers, the diagram representing a stepwise approximation of the final subsidence shown in Fig. 2.23 has been prepared. The procedure described above allows us to calculate the vectors of displacements in the entire deformation zone Fig. 2.24.

Fig. 2.21 Experimentally determined displacements of coins in the test shown in Figs. 2.19 and 2.20
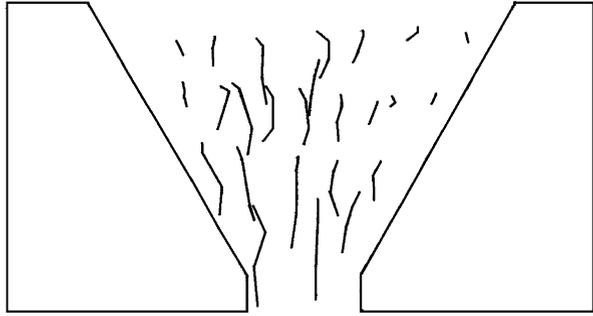
Fig. 2.22 Assumed system of finite cells for the analysis of terrain subsidence, see [24]
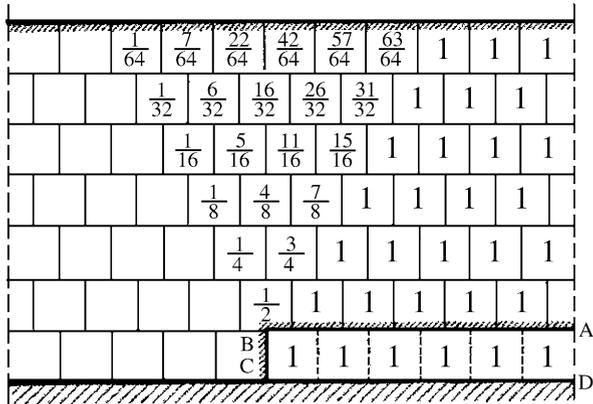
Fig. 2.23 Vertical displacement of granular medium in cells of the assumed system, see [24]
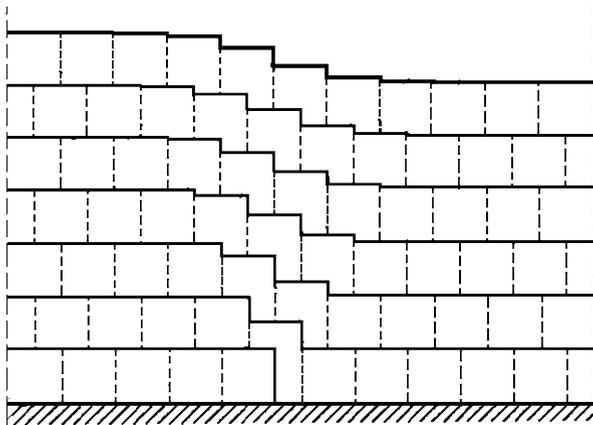
In Fig. 2.25 is presented a simple experimental simulation of such a subsidence process. The coins of different diameters are located on a glass plate as shown in the photograph. To simulate the initial configuration corresponding to that shown in Fig. 2.22, two bottom rows on the right side have been left without coins. Then

Fig. 2.24 Calculated displacements of a granular medium in the process of terrain subsidence shown in Figs. 2.22 and 2.23, after [24]
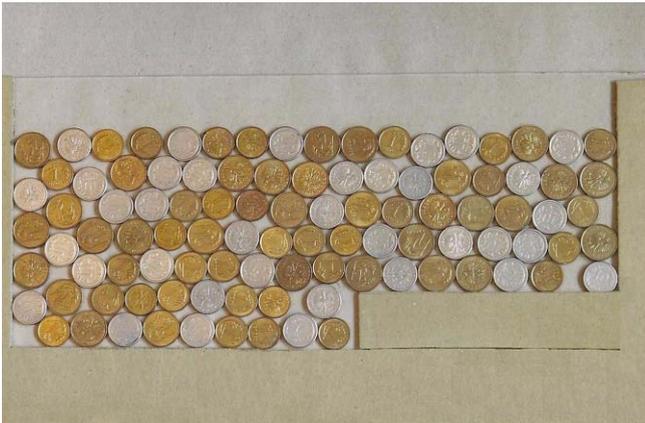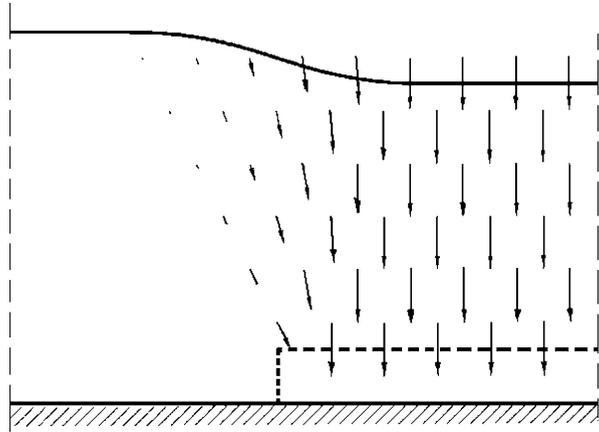
Fig. 2.25 Initial configuration of coins located on a glass plate, see [24]

the blocking strip at the bottom has been removed and the plate was inclined with respect to its initial horizontal position. The coins slid downwards due to gravity force. The final configuration of coins is shown in Fig. 2.26.

The displacements of central points of several coins resulting from this experimental simulation are shown in Fig. 2.27.

Let us note that this experimental result is similar to that resulting from theoretical solution shown in Fig. 2.24.

Summarizing the considerations of this subsection we see that the calculation methods proposed by J. Litwiniszyn were, both, very effective in solving quite involved geotechnics problems and very illustrative. They were also an inspiration for mathematically more advanced models, e.g., description of the random walk of voids by means of diffusive Markov processes, cf. [3].
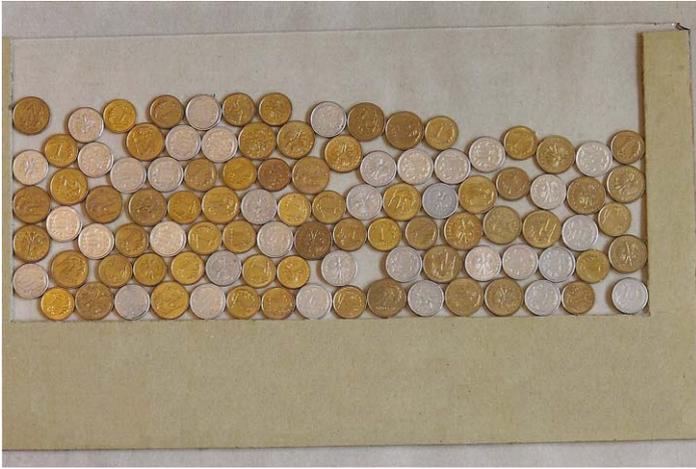
Fig. 2.26 Final configuration of coins in an experimental simulation of terrain subsidence
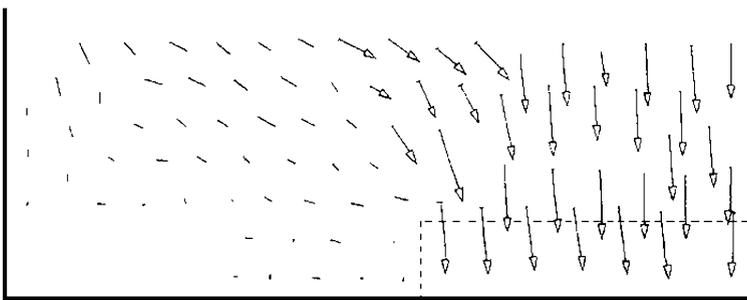
Fig. 2.27 Experimentally determined displacements of coins in the test shown in Figs. 2.25 and 2.26, after [23]

Problem 2.1 Consider a system of four electric elements connected in series. The probability of defective operation of these elements after one year of work is, respectively, 0.6, 0.5, 0.4, and 0.3, and is independent one from the other. Calculate the probability of defective operation of the system of elements. Calculate the probability that the system works correctly.

Problem 2.2 A sample of 200 mass-produced elements is tested by random choice of 10 elements. It is rejected if at least one of the elements is defective. Calculate the probability of the rejection of the sample of elements if 5% of the elements in the sample are defective.

Problem 2.3 Calculate the probability that a sample of 100 mass-produced elements will be accepted if it contains 5 defective elements and we test 50 elements allowing at the most two defective elements among them.

References

1. Abramowitz, M., Stegun, I.: Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. Dover, New York (1965)
2. Bryc, W.: The Normal Distribution. Characterization with Applications. Lecture Notes in Statistics, vol. 100. Springer, New York (1995)
3. Brząkała, W.: Diffusion of voids in the stochastic loose medium. Bull. Acad. Pol. Sci., Sér. Sci. Tech. **30**, 487–491 (1982)
4. Carslaw, H.S., Jaeger, J.C.: Conduction of Heat in Solids, 2nd edn. Oxford University Press, London (1986)
5. Cranz, H.: Aussere Ballistik. Springer, Berlin (1926)
6. Fisz, M.: Probability Theory and Mathematical Statistics, 3rd edn. Krieger, Melbourne (1980)
7. Hastings, N.A.J., Peacock, J.B.: Statistical Distributions, 2nd edn. Wiley–Interscience, New York (1993)
8. Jastrzębski, P.: Strength and carrying capacity of steel and aluminum strips. Scientific Reports of Warsaw University of Technology, No. 1 (1968) (in Polish)
9. Knuth, D.E.: Seminumerical Algorithms, 3rd edn. The Art of Computer Programming, vol. 2. Addison–Wesley, Reading (1998)
10. Korn, G.A., Korn, Th.M.: Mathematical Handbook for Scientists and Engineers, 2nd edn. Dover, New York (2000)
11. Kotulski, Z.: Random walk with finite speed as a model of pollution transport in turbulent atmosphere. Arch. Mech. **45**(5), 537–562 (1993)
12. Kotulski, Z.: On efficiency of identification of a stochastic crack propagation model based on Virkler experimental data. Arch. Mech. **50**(5), 829–847 (1998)
13. Kotulski, Z., Sobczyk, K.: Effects of parameter uncertainty on the response of vibratory systems to random excitation. J. Sound Vib. **119**(1), 159–171 (1987)
14. Kotulski, Z., Szczepiński, W.: On a model for prediction of the movements of a crowd in narrow exits. Eng. Trans. **53**(4), 347–361 (2005)
15. Litwiniszyn, J.: Application of the equation of stochastic processes to mechanics of loose bodies. Arch. Mech. **8**, 393–411 (1956)
16. Litwiniszyn, J.: An application of the random walk argument to the mechanics of granular media. In: Proc. IUTAM Symp. on Rheology and Soil Mechanics, Grenoble, April 1964. Springer, Berlin (1966)
17. Litwiniszyn, J.: Stochastic methods in mechanics of granular bodies. In: CISM Course and Lectures No. 93. Springer, Udine (1974)
18. Matsuo, M., Kuroda, K.: Probabilistic approach to design of embankments. Soil Found. **14**(2), 1–17 (1974)
19. Papoulis, A.: Probability, Random Variables, and Stochastic Processes with Errata Sheet, 4th edn. McGraw–Hill, New York (2002)
20. Sobczyk, K.: Stochastic Differential Equations with Applications to Physics and Engineering. Kluwer Academic, Dordrecht (1991)
21. Steinhaus, H.: Mathematical Kaleidoscope, 8th edn. Polish Educational Editors, Warsaw (1956) (in Polish)
22. Szczepiński, W.: On the movement of granular materials in bins and hoppers. Part I—Two-dimensional problems. Eng. Trans. **51**(4), 419–431 (2003)
23. Szczepiński, W.: On the stochastic approach to the three-dimensional problems of strata mechanics. Bull. Acad. Pol. Sci., Sér. Sci. Tech. **51**(4), 335–345 (2003)
24. Szczepiński, W., Zowczak, W.: The method of finite cells for the analysis of terrain subsidence caused by tectonic movements or by subterranean exploitation. Arch. Mech. **59**(6), 541–557 (2007)

Chapter 3
Probability Distributions and Their Characterizations

3.1 The Characteristic Function of a Distribution

A distribution of the random variable X may be described (equivalently to its description by the probability density function) by the *characteristic function of the distribution*. It is defined as the Fourier transform of the probability density function $f(x)$ of the random variable X:

$$\varphi(\lambda) = \int_{-\infty}^{\infty} \exp[i\lambda x] f(x) dx, \tag{3.1}$$

where $i = \sqrt{-1}$.

The equivalence of two descriptions, by the probability density function and by the characteristic function, is assured by the following Bochner theorem.

Theorem 3.1 *If a function $\varphi(\lambda)$, defined for $-\infty < \lambda < \infty$, satisfies the equality $\varphi(0) = 1$, then it is a characteristic function if and only if*:

(1) *it is continuous, and*
(2) *for $n = 1, 2, 3, \ldots$, for any real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ and for any complex numbers $a_1, a_2, \ldots, a_n$, the relation*

$$\sum_{j,k=1}^{n} \varphi(\lambda_j - \lambda_k) a_j a_k \geq 0 \tag{3.2}$$

is satisfied.

In other words, any continuous function equal to 1 at zero is the characteristic function of a distribution if and only if it is positive definite (that is, it satisfies condition (3.2)).

The verification whether some function is positive definite by checking the inequality (3.2) is, in a general case, very difficult. For this reason the condition is

usually tested by calculation of the inverse Fourier transform of $\varphi(\lambda)$ and examining whether the obtained function is positive. Another possibility is given by the following theorem presented by Cramer (cf. [4]):

Theorem 3.2 *A bounded and continuous function $\varphi(\lambda)$ is the characteristic function of a distribution if and only if $\varphi(0) = 1$ and the function*

$$\psi(x, a) = \int_0^a \int_0^a \varphi(\lambda - \tau) \exp[ix(\lambda - \tau)] \, d\lambda \, d\tau \tag{3.3}$$

is real-valued and non-negative for all x and for all $a > 0$.

The knowledge of the characteristic function of a random variable enables us to calculate its moments. If the characteristic function $\varphi(\lambda)$ of the distribution of random variable X is n times differentiable, then this random variable has a finite moment of n-th order m_n, and the following relation is satisfied:

$$m_n = \frac{1}{i^n} \frac{d^n}{d\lambda^n} \varphi(\lambda) \big|_{\lambda=0}. \tag{3.4}$$

Example 3.1 (Characteristic function of the normal distribution) Consider, as an example, the normal (Gaussian) distribution. The characteristic function $\varphi(\lambda)$ of this distribution has the form

$$\varphi(\lambda) = \exp\left[im_1\lambda - \frac{1}{2}\sigma^2\lambda^2 \right], \tag{3.5}$$

where m_1 is the mean value and $\sigma^2 = \mu_2$ is the variance of the distribution.

If the characteristic function of a distribution is known, we are able to calculate all the moments of this distribution using formula (3.4). Concerning the Gaussian distribution we find that its central moments have simple forms. Calculating the moments and using formula (2.28) we obtain:

$$\mu_n = \begin{cases} 0 & \text{for odd } n, \\ \dfrac{\sigma^n n!}{2^{\frac{n}{2}} [(\frac{n}{2})!]} & \text{for even } n. \end{cases} \tag{3.6}$$

As it is seen, the Gaussian distribution is completely described by its two lowest-order moments: the mean value and the variance. All the higher-order moments can be represented by polynomial functions of the mean value and the variance.

Equation (3.4) gives us the relation between the characteristic function of a random variable in question and its moments of any order. It is seen that for small values of λ the characteristic function can be expanded into the Maclaurin series, where the moments play the role of coefficients of the expansion:

$$\varphi(\lambda) = 1 + \sum_{k=1}^{n} \frac{m_k}{k!} (i\lambda)^k + o(\lambda^n). \tag{3.7}$$

If we consider Maclaurin's expansion of the natural logarithm of the characteristic function $\varphi(\lambda)$,

$$\log \varphi(\lambda) = \sum_{k=1}^{n} \frac{\kappa_k}{k!} (i\lambda)^k + o\left(\lambda^n\right), \tag{3.8}$$

then the coefficients κ_k of the obtained series are called the *cumulants* of random variable X or the *cumulants of the distribution.*

The cumulants of a certain distribution can be represented as polynomials of moments of this distribution. (Analogously: moments can be expressed as polynomials of cumulants.) The exact formulas can be obtained by formal comparison of the logarithm of the expansion (3.7) and the expansion (3.8) (see, e.g., [4]) and equivalence of the coefficients at equal powers of λ. Thus, we obtain the (usually infinite) system of equations relating the moments and cumulants.

Example 3.2 (Cumulants of the normal distribution) For the normal distribution we have:

$$\log \varphi(\lambda) = i m_1 \lambda - \frac{1}{2}\sigma^2 \lambda^2$$

and, as the consequence of the definition of cumulants, we obtain:

$$\kappa_1 = m_1,$$
$$\kappa_2 = \sigma^2,$$
$$\kappa_j = 0 \quad \text{for } j > 2.$$

The normal distribution is the only continuous one which possesses a finite number of non-zero cumulants. For other distributions, infinitely many cumulants possess non-zero values.

3.2 Constants Characterizing the Random Variables

The information on random variables and the probability distributions presented above shows that they can be described in various ways. However, the simplest characterization of them is as numbers (e.g., moments of one-dimensional distributions). Now we give some additional parameters describing the properties of random variables.

The most fundamental information about a random variable is connected with determining the point around which its distribution is concentrated. Introducing moments we defined the *average value* of the random variable (sometimes called the *mean value*, the *expected value* or, simply, the *expectation*), calculated according to the formula

$$\overline{x} = \int_{-\infty}^{\infty} x f(x)\, dx. \tag{3.9}$$

It describes the "center of gravity" of the random variable, calculated with respect to the measure described by the probability density function $f(x)$.

Another constant characterizing the center of distribution of a random variable can be defined by a choice of such a value of the argument x of a probability density function $f(x)$, where this function has a local maximum. We call it the *modal value* (*mode*). If the probability density function has only one maximum (the distribution is *unimodal*) then we observe that possible values of the random variable concentrate around this point (the point where the probability density function has the maximum value). If the probability density function has two maximums, we call it *bimodal* (examples of bimodal distributions are shown in Figs. 1.6 and 1.7b), etc. In the case of several maximums, the modal values do not give us direct information about the points of concentration of the distribution.

The *median* (*median value*), denoted by m can be some other number describing the "middle of the distribution". It is defined as the value of the variable x for which the probability density function $f(x)$ satisfies simultaneously the two following conditions:

$$
\begin{aligned}
F(x) \leq \frac{1}{2} \quad \text{for } x \leq m, \\
F(x) \geq \frac{1}{2} \quad \text{for } x \geq m.
\end{aligned}
\tag{3.10}
$$

The number m satisfying the conditions (3.10) always exists but sometimes it cannot be determined uniquely.

In a general case, for a given random variable, the three defined numbers: the mean value, the modal value, and the median value differ from each other. In the special case of unimodal symmetric distributions these numbers are equal to each other. Then, the value searched for is the point x, where the axis of symmetry of the probability density function $f(x)$ crosses the x-axis.

To describe the distribution of a random variable more precisely, we must have some other quantity, describing its dispersion around the point of concentration (the mean value). Such a number has been defined in Sect. 2.3, concerning the moments of random variables. We mean here the *standard deviation* σ being the square root of the variance σ^2, see (2.24). The standard deviation, however, does not describe dispersion of the random variable completely. We know that if the standard deviation equals zero, then the random variable is a constant number:

$$
\sigma = 0 \quad \Leftrightarrow \quad P(X = c) = 1.
\tag{3.11}
$$

For non-zero standard deviations we may compare random variables having the same type of distribution. Then we say that the one of the compared random variables is more dispersed around its mean value which has a greater standard deviation. For different distributions, the standard deviation gives only some intuitive information about dispersion of random variables.

Another measure of dispersion, closely related to the standard deviation, is the following *variation ratio* v of a random variable, defined as

$$v = \frac{\sigma}{m_1}. \tag{3.12}$$

In other words, the variation ratio is equivalent to the standard deviation measured in units of the mean value of the random variable under consideration.

The next measure of dispersion, applicable to random variables with discrete distributions or finite samples of random variables with continuous distributions, is the *range* δ defined as the distance between the maximum and the minimum values of the random variable. The concept of the range can be also applied to any random variable with continuous distribution provided it takes values from a bounded interval. Then the range δ is the length of that interval.

One can find a relation between the variation ratio v and the range δ. Let $m_1 > 0$ be the mean value of the random variable X and assume that there exist constants a and A such that the condition $P(a \leq X \leq A) = 1$ is satisfied. Then

$$v \leq \sqrt{\left(\frac{A}{m_1} - 1\right)\left(1 - \frac{a}{m_1}\right)}. \tag{3.13}$$

In this case the range δ satisfies the condition

$$\delta \leq A - a. \tag{3.14}$$

The average deviation d defined in (2.25) can be also considered as a certain measure of dispersion of a random variable.

After determination of the concentration point and the level of dispersion of the distribution of a random variable, one can try to describe the properties (the shape) of the probability density function more precisely. As a curve of reference we can use the normal distribution. Observing its probability density function (Fig. 2.4) one can see its characteristic properties: it has one maximum (it is a unimodal distribution), it is symmetric with respect to the central point, and it is concentrated around the mean value. Therefore, when a unimodal distribution is considered, we may compare its symmetry and flattening to an appropriate (with the same mean value and variance as those for the distribution under consideration) normal distribution.

The most frequently used measure of asymmetry of the probability density function is the *asymmetry (skewness) coefficient* γ_1 defined as (see [4, 18])

$$\gamma_1 = \frac{\mu_3}{\sigma^3}, \tag{3.15}$$

where μ_3 is the central moment of the third order, and σ is the standard deviation. For a normally distributed random variable or any other symmetrical one, the asymmetry coefficient γ_1 equals zero. If γ_1 is negative, we deduce that the probability density function has a long tail on the left-hand side of the mean value. The positive value of γ_1 indicates the existence of the long tail at the right-hand side (see Fig. 3.1).

Fig. 3.1 The asymmetry coefficient, see (3.15)
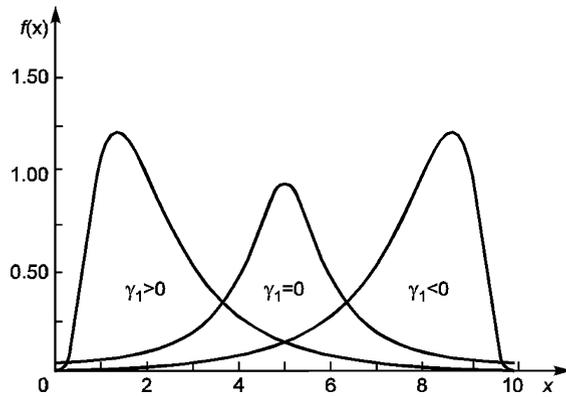
Fig. 3.2 The excess coefficient, see (3.16)
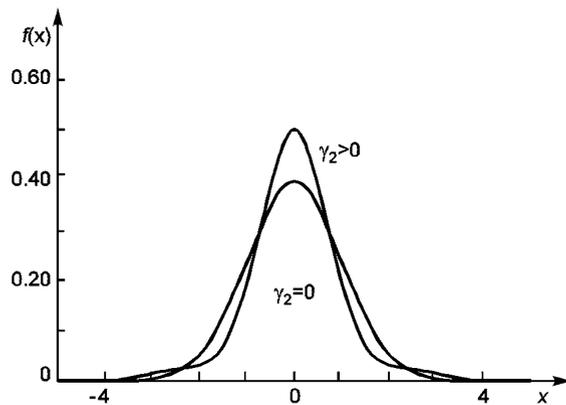
The *excess (kurtosis) coefficient* γ_2, defined as

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3, \tag{3.16}$$

is the measure of flattening of the probability density functions. As in the previous case, for the normal distribution the coefficient γ_2 is equal to zero. If the excess coefficient γ_2 is positive, then the given probability density function is taller and its shape is slimmer around the modal value than the density of the normal distribution. The negative value of γ_2 leads to the opposite conclusion (see Fig. 3.2).

Another possibility of description of the shape of the probability density function is the application of the location parameters, called *quantiles (fractiles, percentiles)*. For a random variable X of any type of distribution and a given p, $0 < p < 1$, the quantile of order p is such a number α_p which satisfies the condition

$$P\left(X < \alpha_p\right) \le p \le P\left(X \le \alpha_p\right). \tag{3.17}$$

If the random variable X has a continuous distribution, then condition (3.17) may be written in the following form:

$$P\left(X < \alpha_p\right) = p, \tag{3.18}$$

or

$$F\left(\alpha_p\right) = \int_{-\infty}^{\alpha_p} f\left(x\right) dx = p. \tag{3.19}$$

Let us note that the *median* (median value) is the quantile of order $1/2$.

The appropriate set of quantiles allows us to describe approximately the probability density function (or cumulative distribution function) of any random variable with a continuous distribution. They determine "gates" guiding the curve of the density function to obtain the approximate plot. In the limiting case the knowledge of all the quantiles (that is, the quantile function) is equivalent to the knowledge of the cumulative distribution function (see Sect. 2.2).

The quantiles can be applied for constructing the measures of dispersion, skewness and kurtosis. For example, the *quarter distance* τ_X, defined as

$$\tau_X = \frac{1}{2} \left(\alpha_{3/4} - \alpha_{1/4}\right), \tag{3.20}$$

may be considered as a measure of dispersion.

To complete this section we give one more characteristic of the probability distribution. We define the *entropy* which for a continuous distribution with the probability density function $f(x)$ has the form[1]

$$H = - \int f\left(x\right) \log f\left(x\right) dx = -E\left\{\log f(X)\right\}. \tag{3.21}$$

For discrete distributions the entropy is defined as

$$H = - \sum_{k} p_k \log p_k = -E\left\{\log P(X)\right\}. \tag{3.22}$$

The entropy describes the amount of information contained in the distribution of a random variable. If the entropy is greater, the random variable is more dispersed and one needs more information to describe the distribution of the random variable. For example, the normal distribution has the greatest entropy among all the continuous distributions defined along the entire straight line and possessing the same variance. The exponential distribution (see Sect. 3.4.2) has the analogous property among the continuous distributions defined along the positive axis.

---

[1] Symbol $E\{.\}$ denotes the expectation of a random variable; it will be defined in Sect. 3.3.

3.3 Deterministic Functions of Random Variables

In many problems of applied probability theory we use random variables which are deterministic functions of other random variables. Such a situation occurred during calculation of the moments of random variables: starting from the random variable X with a given distribution, we constructed a new random variable being a power of X and then we calculated its mean value.

Before we start considerations on functions of random variables, let us introduce some notation. Let E be the symbol of *expectation* of a random variable, that is, the integral with respect to the probability P (defined in the σ-algebra of all events). For the random variable $X(\omega)$ possessing the probability density function $f_X(x)$, the integration is equivalent to the integration with respect to the probability density function:

$$E\{X(\omega)\} = \int_{\Omega} X(\omega)\, dP(\omega) = \int_{-\infty}^{\infty} x f_X(x)\, dx. \qquad (3.23)$$

Thus, the integration of the P-measurable function $X(\omega)$ over the set of elementary events Ω is replaced by the integration of the real function $x f_X(x)$ over the set of real numbers $R = (-\infty, \infty)$.

Let $g(x)$ be a deterministic measurable function, $g : R \to R$. Since $g(x)$ is measurable, the function of the random variable $X(\omega)$, namely $Y(\omega) = g(X(\omega))$, is also measurable and, as a consequence, is a random variable. To calculate the mean value of the random variable $Y(\omega)$ we can perform one of the two equivalent integrations:

$$E\{Y(\omega)\} = E\{g(X(\omega))\} = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx = \int_{-\infty}^{\infty} y f_Y(y)\, dy. \qquad (3.24)$$

Integration with respect to the probability density function of random variable $Y(\omega)$ requires the knowledge of this function. We can find it by changing variables under the integration sign in (3.24) and by comparing the corresponding functions. Finally, we obtain

$$f_Y(y) = f_X(\psi(y)) |\psi(y)|, \qquad (3.25)$$

provided a unique inverse transformation $X = \psi(Y)$ exists. If the function inverse to $g(x)$ is non-unique, that is, some values of y are assigned to several values of x: $\psi_1(y), \psi_2(y), \ldots, \psi_k(y)$, then the probability density function of a random variable Y can be represented by

$$f_Y(y) = \sum_{j=1}^{k} f_X(\psi_j(y)) |\psi_j(y)|. \qquad (3.26)$$

Performing integration with respect to the probability density function $f_Y(y)$ we must remember the correct values of the limits of integration resulting from the form of function $g(x)$.

Example 3.3 (The arcsine distribution) If we assume that the random variable X has a uniform distribution in the interval $(-\frac{1}{2}T, \frac{1}{2}T)$, then the random variable $Y = a \sin(\frac{2\pi X}{T})$ has the following arcsine distribution:

$$f_Y(y) = \begin{cases} \frac{1}{\pi \sqrt{a^2 - y^2}} & \text{for } |y| < a, \\ 0 & \text{for } |y| \geq a. \end{cases} \tag{3.27}$$

One can give a number of useful inequalities for functions of random variables. One of them is the Jensen inequality. If an arbitrary real-valued function $g(x)$ is concave and for the random variable $X(\omega)$ both $E\{X(\omega)\}$ and $E\{g(X(\omega))\}$ exist, then

$$E\{g(X(\omega))\} \leq g(E\{X(\omega)\}); \tag{3.28}$$

the equality applies only to the linear function $g(x) = ax + b$.

Assume that $g(x) = x^{p/q}$, $X(\omega) = |Z(\omega)|^q$, $0 < p < q$. Substituting $g(x)$ and $X(\omega)$ in Jensen inequality (3.28) we get

$$E\{|Z(\omega)|^p\} \leq \left(E\{|Z(\omega)|^q\}\right)^{p/q}. \tag{3.29}$$

The inequality (3.29) confirms the fact that for any given random variable $Z(\omega)$ we can deduce the existence of lower-order moments from the existence of higher-order moments.

Now a number of basic probability distributions used in engineering applications will be discussed.

3.4 Some Other One-dimensional Distributions

In addition to the normal probability distribution, which plays a fundamental role, some other distributions are often used in the error calculus. In this section we present examples of distributions which are most often encountered in engineering applications. It is not our purpose to classify all the known probability distributions. However, we shall present some distributions which are not used directly in practical applications discussed in this book, but which can be useful for the readers in their further studies.

3.4.1 Discrete Probability Distributions

We say that a random variable X has a *discrete distribution* (is discrete) if it takes at most a countable number of values $x_j \in R$, $j = 1, 2, \ldots$, with a non-zero probability. This means that its probability distribution is completely described by the

sequence of numbers p_j, $j = 1, 2, \ldots$, (finite or countable) satisfying the following conditions:

$$p_j = P(\omega \in \Omega : X(\omega) = x_j) > 0 \quad \text{for } x_j \in R, \ j = 1, 2, \ldots, \tag{3.30a}$$

$$\sum_j p_j = 1. \tag{3.30b}$$

In (3.30b) the subscript j takes a finite or countable number of values, depending on the particular distribution.

For a discrete random variable the cumulative distribution function is a piecewise constant function and the probability density function does not exist.

3.4.1.1 The Binomial Distribution

We say that the discrete random variable X, taking values from the finite subset of natural numbers $A = \{1, 2, \ldots, n\} \subset N$, has the *binomial distribution* $b(n, p)$, if the probabilities p_j are defined by

$$p_j = P(X = j) = \binom{n}{j} p^j (1 - p)^{n-j}, \tag{3.31}$$

where the parameter p takes values from the interval $(0, 1)$.

Two lowest-order moments of the binomial random variable X are

$$E\{X\} = np, \qquad \sigma_X^2 = np(1 - p). \tag{3.32}$$

The binomial distribution can be applied to describe the number of successes produced in a sequence (succession) of n Bernoulli trials. In this process we describe a sequence of n independent experiments which give a result called a "success", with probability p, and a "failure", with probability $1 - p$. Then the random variable X describing the number of successes in n Bernoulli trials has the binomial $b(n, p)$ distribution. The number p_j defined in (3.31) represents the probability of j successes and $n - j$ failures in the sequence.

Example 3.4 (An unbiased coin flipping, continuation) An unbiased single coin flip is a Bernoulli trial where the probability of success (e.g., of heads) is $p = \frac{1}{2}$. Then the number of heads in n coin flips has the binomial distribution $b(n, \frac{1}{2})$.

The binomial distribution is often applied in the reliability theory of compound systems (describing failures of large systems), telecommunications (transmission of bits through noisy or non-perfect channels), cryptography (encryption of binary sequences), etc.

3.4.1.2 The Multinomial Distribution

The multinomial distribution is a generalization of the binomial distribution. Assume that we make an experiment which can give one of several (say r) results B_j, $j = 1, 2, \ldots, r$, each with probability p_j, $j = 1, 2, \ldots, r$, respectively. Consider a succession of n such experiments. The probability that in the succession the result B_j occurs k_j times, $j = 1, 2, \ldots, r$, has the *multinomial distribution*

$$P\{N(B_1) = k_1, N(B_2) = k_2, \ldots, N(B_r) = k_r\} = \frac{n!}{k_1! k_2! \cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r},$$
(3.33)

provided the following conditions are satisfied:

$$p_1 + p_2 + \cdots + p_r = 1,$$
$$k_1 + k_2 + \cdots + k_r = n.$$
(3.34)

Formula (3.33) is called multinomial because its right-hand side expression is the general term of the multinomial expansion of the n-th power of the sum of probabilities p_j: $(p_1 + p_2 + \cdots + p_r)^n$.

For $r = 2$ the multinomial distribution reduces to the binomial one.

Example 3.5 (Dice casting, continuation) Results of multiple (honest) die casting can be described by the multinomial distribution with $r = 6$ and

$$p_1 = p_2 = \cdots = p_6 = \frac{1}{6}.$$

3.4.1.3 The Poisson Distribution

Let X be a discrete random variable taking values from the set of natural numbers. If for some real number $\lambda > 0$ the probabilities are of the form

$$p_j = P(X = j) = e^{-\lambda} \frac{\lambda^j}{j!},$$
(3.35)

for $j = 1, 2, \ldots$, then we say that a random variable X has the *Poisson distribution*. Two lowest-order moments of the Poisson random variable are

$$E\{X\} = \lambda, \qquad \sigma_X^2 = \lambda.$$
(3.36)

Random variables with the Poisson distributions are often used for description of physical phenomena. A typical example of application is a model of the natural radioactivity of matter (e.g., counting the number of pulses by the Geiger-Mueller detector, see [27]). Another application is modeling of stochastic excitations acting on engineering structures, generated by e.g. road traffic, wind shocks, hail, etc. (see [11]).

3.4.2 Continuous Probability Distributions

We say that a random variable X has a *continuous distribution* if it can take all values from some real line interval (also: the whole line or the half-line).

3.4.2.1 The Rectangular (Uniform) Distribution

The *rectangular distribution* is also called the *uniform distribution* or the *equal probability distribution* (comp. [21]). The probability density function for such a distribution (Fig. 3.3a) takes the form

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b, \\ 0 & \text{for } x < a \text{ and } x > b. \end{cases} \tag{3.37}$$

The cumulative distribution function is shown in Fig. 3.3b.

It is not expected that such a distribution can really be observed in practice. It can, however, be useful in various engineering applications, especially, if the so-called tolerance limits, when certain dimensions of the object may change within a range between a strictly defined limits from a to b, are dealt with. Examples of such applications will be given in Sect. 6.2.

The rectangular distribution (3.37) is the simplest one for the analysis. Since the mean value and the variance have the following simple forms:

$$m = \frac{a+b}{2}, \qquad \sigma^2 = \frac{(b-a)^2}{12}, \tag{3.38}$$

respectively, it can be shown that for a given mean value m and a given variance σ^2, limiting values a and b must be equal to

$$a = m - \sqrt{3\sigma^2}, \qquad b = m + \sqrt{3\sigma^2}. \tag{3.39}$$
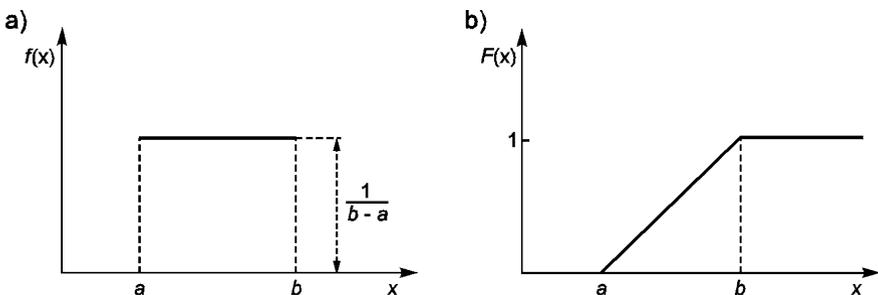
Fig. 3.3 The uniform distribution: (**a**) the probability density function, (**b**) the cumulative distribution function

Returning to the three-sigma rule presented in Sect. 2.4, one should note that the random variable of the uniform distribution takes, by definition, all its values within the range $(m - \sqrt{3}\sigma, m + \sqrt{3}\sigma)$. Thus, one can say that the uniform distribution is more concentrated than the normal distribution. (The normally distributed random variable assumes the values within the three-sigma range with probability 0.9973, see Table 2.2). The three-sigma range for the rectangular distribution equals: $3\sigma = \frac{\sqrt{3}}{2}(b - a)$, and it is larger than the half-width of the distribution range.

The characteristic function of the rectangular distribution takes the form

$$\varphi(\lambda) = \frac{\exp[ib\lambda] - \exp[ia\lambda]}{i(b - a)\lambda}, \tag{3.40}$$

where $i = \sqrt{-1}$. The central moments of an arbitrary order of that distribution are

$$\mu_n = \begin{cases} 0 & \text{for } n \text{ odd,} \\ \frac{1}{n+1}(\frac{b-a}{2})^n & \text{for } n \text{ even.} \end{cases} \tag{3.41}$$

In particular, for $n = 3$ and $n = 4$ these moments are

$$\mu_3 = 0, \qquad \mu_4 = \frac{(b - a)^4}{80}. \tag{3.42}$$

The rectangular (uniform) distribution plays a basic role in computer simulations when we want to find the random variables of an arbitrary distribution for problems in which such variables are dealt with. In such simulations the so-called random number generators producing a sequence of numbers with the uniform distribution are used. They usually operate using numerical algorithms allowing one to form sequences of non-periodic (that means practically: having a sufficiently long period) numbers within the range [0, 1]. Having generated such a random variable one can find a random variable of an arbitrary probability density distribution by using the respective transformation (e.g., inversion of the cumulative distribution function).

Example 3.6 (Random numbers generator: the uniform distribution) As an example, we obtain the sequence $\{x_n, n = 1, 2, \ldots\}$, using the algorithm:

$$y_{n+1} = ay_n + b \pmod{M}, \quad b \neq 0,$$

$$x_n = \frac{y_n}{M},$$

where a, b, and y_0 are natural numbers from the range $[0, M)$. These numbers must be carefully selected to obtain a sequence satisfying the conditions of randomness, uniformity of the distribution, and independence of elements, usually verified by statistical tests (see Chap. 10).

Examples of other random numbers generators may be found in [3, 12, 22].

3.4.2.2 The Lognormal Distribution

In numerous cases of random distribution of errors we know that the measured deviations from the desired value of the parameter in question must be positive. Typical examples are those in which the desired magnitude of a measured parameter equals zero, e.g., the distance between centers of parts which should be coaxial, the contamination of soil, water or air. In such cases the measured deviation cannot be treated as a random variable with the normal distribution. It should be treated as the one that assumes only positive values. The most frequently used distribution corresponding to such a situation is the lognormal distribution.

We say that a random variable X has the *lognormal distribution* if the natural logarithm of the variable displays the normal distribution. The probability density function of the lognormal distribution is

$$f(x) = \begin{cases} 0 & \text{for } x < 0, \\ \frac{1}{x\sqrt{2\pi\sigma^2}} \exp[\frac{-(\log x/m)^2}{2\sigma^2}] & \text{for } x \geq 0. \end{cases} \tag{3.43}$$

The lognormal distribution is two-parametric. The parameter m is called the median (comp. Sect. 3.2), and the other parameter σ stands for the standard deviation of the natural logarithm of the random variable X. Both these parameters are associated with the first-order moments.

Namely, the first-order moments (the mean value and the variance) are expressed by the formulas

$$m_X = m_1 = m \exp\left[\frac{1}{2}\sigma^2\right],$$

$$\sigma_X^2 = \mu_2 = m^2\left\{\exp[\sigma^2] - \exp\left[\frac{1}{2}\sigma^2\right]\right\}. \tag{3.44}$$

In Figs. 3.4 and 3.5 are presented two examples of histograms summarizing the results of large scale measurements, the random nature of which is well described by the lognormal distribution. These results represent the distribution of contamination of water sediments by vanadium (Fig. 3.4) and of the soil by cadmium (Fig. 3.5) over large regions in Poland. For the two histograms in which the independent variable x represents the amount of contamination in $[g/t]$, two first-order moments and then the corresponding parameters of the lognormal distribution have been calculated.

In both cases, in spite of different forms of the histograms, fairly good coincidence of empirical distributions with the respective approximation by the lognormal distribution is clearly seen. The values of the calculated parameters are given in Table 3.1.

For the lognormal distribution one can calculate the moments of an arbitrary order. They are expressed by the simple formula

$$m_n = m^n \exp\left[\frac{1}{2}n^2\sigma^2\right]. \tag{3.45}$$
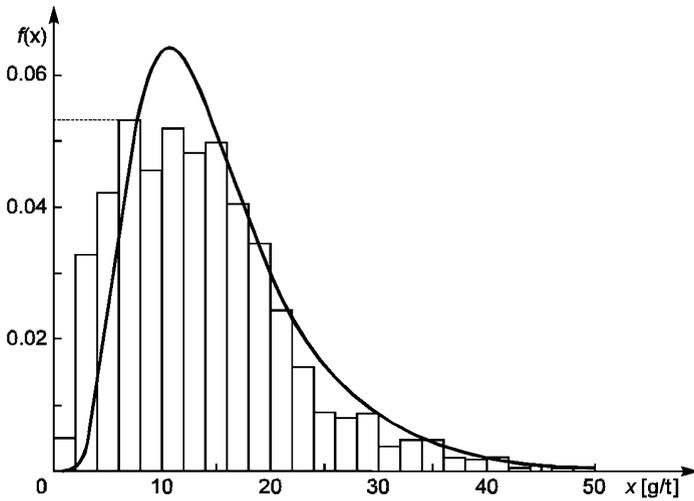
Fig. 3.4 The distribution of contamination of water sediments by vanadium, see [8]
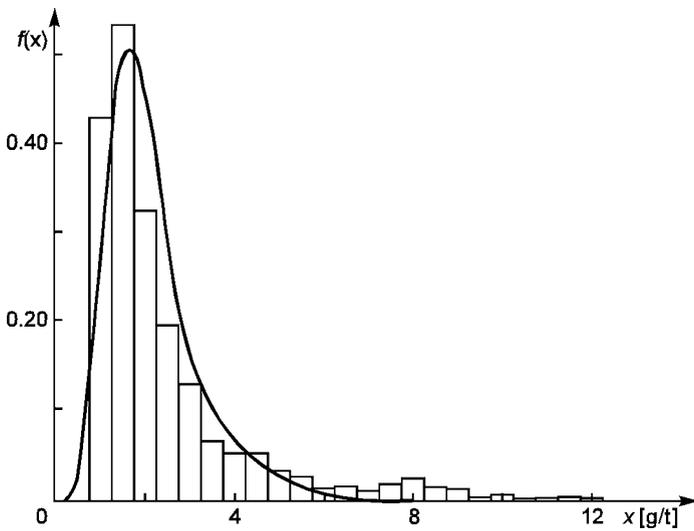
Fig. 3.5 The distribution of contamination of the soil by cadmium, see [8]

Having found the expressions for ordinary moments, one can then calculate the central moments by means of formula (2.28). For example, the central moments of

Table 3.1 The parameters of the lognormal distribution used for approximation of the histograms from Figs. 3.4 and 3.5

	Vanadium	Cadmium
The mean value m_X	13.77 $[g/t]$	2.41 $[g/t]$
The second-order moment $E\{X^2\}$	254.6 $[g^2/t^2]$	9.12 $[g^2/t^2]$
The parameter m	11.92	1.92
The parameter σ^2	0.28	0.455
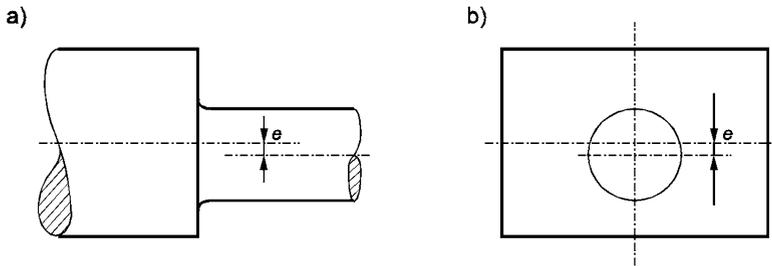
Fig. 3.6 Applications of the lognormal distribution: (**a**) the distance of centers, error e, of the two adjoining parts of a shaft, (**b**) the deviation e of the hole axis from its desired central position

the third-order and fourth-order are:

$$\mu_3 = m_3 - 3m_2 m_1 + 2m_1^3$$
$$= m^3 \left\{ \exp\left[\frac{9}{2}\sigma^2\right] - 3\exp\left[\frac{5}{2}\sigma^2\right] + 2\exp\left[\frac{3}{2}\sigma^2\right] \right\}, \tag{3.46}$$

$$\mu_4 = m_4 - 4m_3 m_1 + 6m_2 m_1^2 - 3m_1^4$$
$$= m^4 \left\{ \exp\left[\frac{16}{2}\sigma^2\right] - 4\exp\left[\frac{10}{2}\sigma^2\right] + 6\exp\left[\frac{6}{2}\sigma^2\right] - 3\exp\left[\frac{4}{2}\sigma^2\right] \right\}. \tag{3.47}$$

The lognormal distribution may also be used in other practically important cases as a good approximation of real distributions. Let us present two examples: the distribution of distance between the centers, error e, of the two adjoining parts of a shaft (Fig. 3.6a), or the deviation e of the hole axis from its desired central position (Fig. 3.6b).

The lognormal distribution may also be adequate for a description of the results of the so-called fragmentation tests of fiber-reinforced composites. In such tests a number of identical cylindrical specimens, each with a single central fiber surrounded by the matrix material, is set in tension as shown in Fig. 3.7. As the deformation process proceeds, the number of fractured sections of the central fiber increases. However, at a certain specific elongation of the whole specimen the fracturing process in the fiber comes to an end. This is referred to as the saturation state (comp. [2]). In this state the fiber is divided into a number of separate sectors, each of length l_i. In [26] the results of numerous tests were presented in the form of a histogram representing the distribution of length of these separate sectors of the fiber

Fig. 3.7 A sample obtained in the fragmentation test of fiber-reinforced composites
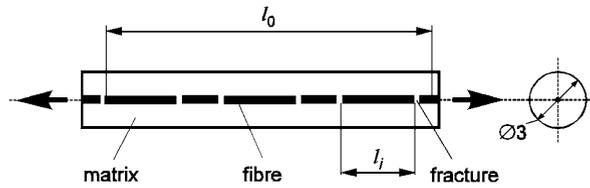
Fig. 3.8 The distribution of length of separate sectors of the fiber in the saturation state (the fragmentation test), see [26]
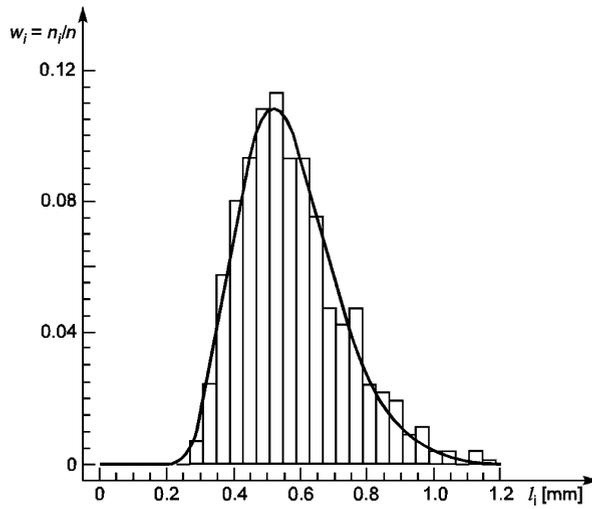
in the saturation state (Fig. 3.8). The corresponding approximation of that experimental histogram by the continuous lognormal distribution is also presented in the figure.

In papers concerning engineering problems one can find many applications of the lognormal distribution. An interesting example of using this distribution for description of a parameter in a model of fatigue crack growth may be found in [6]. Other applications deal with some problems of the reliability theory (see [5, 20, 25]).

3.4.2.3 The Exponential Distribution

As another example of distribution of a random variable X which may take positive values only, the *exponential distribution* may be mentioned. It is a one-parameter distribution with the probability density given by

$$f(x) = \begin{cases} 0 & \text{for } x < 0, \\ \frac{1}{b} \exp\left[-\frac{x}{b}\right] & \text{for } x \geq 0, \end{cases} \tag{3.48}$$

where the parameter b equals the mean value of the distribution. The variance of that distribution equals b^2. Thus, the standard deviation is equal to the mean value.

Fig. 3.9 The probability density function of the exponential distribution
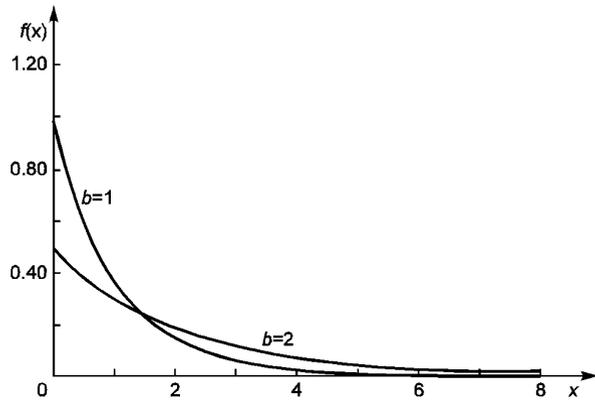
The characteristic function of the exponential distribution has the form

$$\varphi(\lambda) = \frac{1}{1 - ib\lambda}, \tag{3.49}$$

where $i = \sqrt{-1}$. The n-th order ordinary moment is given by

$$m_n = n!b^n. \tag{3.50}$$

Once knowing the ordinary moments, we may calculate (using (2.28)) the central moments of an arbitrary order. For example, when $n = 3$ or $n = 4$ we receive

$$\mu_3 = 2b^3, \qquad \mu_4 = 9b^4. \tag{3.51}$$

According to the exponential distribution, the probability that the random variable in question takes large values decreases exponentially (see Fig. 3.9). Therefore, this distribution finds numerous applications in the theory of reliability, where the probability that a system under consideration works without breakdown exhibits such a behavior.

The exponential probability distribution finds also applications in the analysis of layered geological strata, where the sedimentation process of successive layers happens to be randomly interrupted. For such layered strata one can assume, with a fair approximation, that the thickness of the layers h_i is a random independent variable with the exponential distribution (one can assume that the sedimentation time of specific layers, that is, the time between cataclysms interrupting sedimentation, has the exponential distribution). An example of such a layered medium is shown in Fig. 3.10.

Numerous problems concerning propagation of elastic waves in layered media with exponential distribution of the thickness of layers are analyzed in [13–17].

Example 3.7 (Generation of random numbers: the exponential distribution) Generation of random variables with the exponential distribution may be performed with the use of generators of random numbers. It can be done by means of the so-called

Fig. 3.10 An example of the geological stratified medium: investigation of material properties by acoustic waves
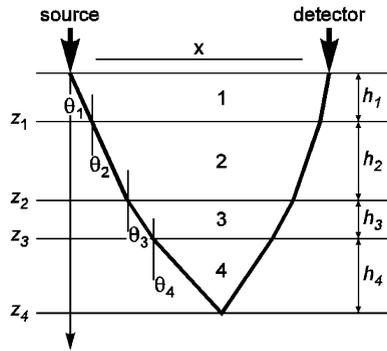
inversion of the cumulative distribution function. If $\{y_n,\ n = 1, 2, \ldots\}$ is a sequence of independent realizations of random variable Y with a distribution uniform along the interval $[0, 1]$ (generated by an appropriate random numbers generator), then the sequence $\{x_n = F^{-1}(y_n),\ n = 1, 2, \ldots\}$ becomes the sequence of independent realizations of a random variable with the exponential distribution (comp. [3, 13]). Function $y = F(x)$ represents here the cumulative distribution function of the exponential distribution

$$y = F(x) = \begin{cases} 0 & \text{for } x < 0, \\ 1 - \exp[-\frac{x}{b}] & \text{for } x \geq 0, \end{cases}$$

while function $x = F^{-1}(y)$ is the function inverse to it:

$$x = F^{-1}(y) = -b \log y \quad \text{for } y \in [0, 1].$$

3.4.2.4 The Gamma Distribution

The *gamma distribution* is defined on a half-line of positive numbers, $x \geq 0$. It has two parameters: $p > 0$ describing the scale and $b > 0$ describing the shape of the distribution. If p is an integer, then the distribution represents the sum of p independent exponentially distributed random variables with the same mean value b. The gamma distribution is often used as a model of waiting times in mass-service systems and reliability theory. The probability density function of the gamma distribution has the following form:

$$f(x) = \begin{cases} 0 & \text{for } x < 0, \\ x^{p-1} \frac{\exp[-\frac{x}{b}]}{\Gamma(p)b^p} & \text{for } x \geq 0, \end{cases} \tag{3.52}$$

where the gamma function $\Gamma(p)$ for $p > 0$ is defined as

$$\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx. \tag{3.53}$$
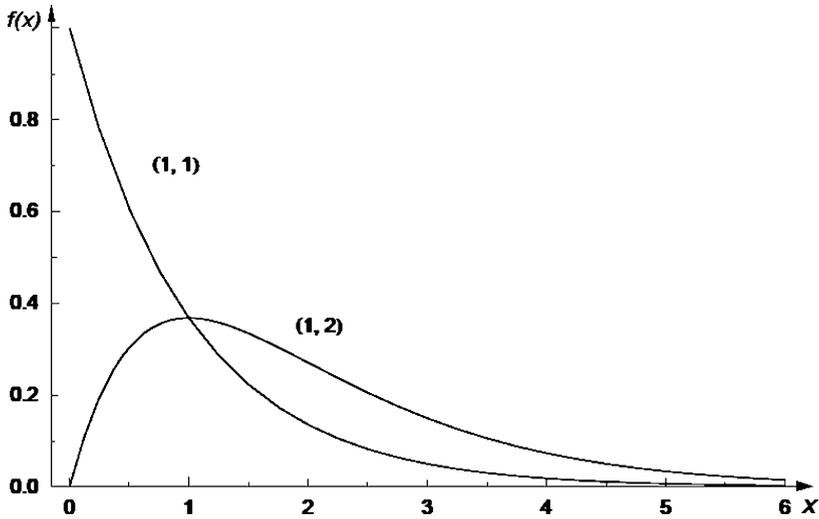
Fig. 3.11 Probability density function of the gamma distribution for $b = 1$, $p = 1$ and $b = 1$, $p = 2$

The plot of the probability density function of the gamma distribution for $b = 1$, $p = 1$ and $b = 1$, $p = 2$ is presented in Fig. 3.11.

The expressions for basic quantities characterizing the gamma distribution can be easily calculated. Thus, the mean value and the variance of the random variable X are:

$$m_X = pb, \qquad \sigma_X^2 = pb^2. \tag{3.54}$$

Also the skewness and the kurtosis are elementary functions of the distribution's parameters:

$$\gamma_1 = \frac{2}{\sqrt{p}}, \qquad \gamma_2 = \frac{6}{p}. \tag{3.55}$$

The characteristic function of the gamma distribution has the form

$$\varphi(\lambda) = \frac{1}{(1 - ib\lambda)^p}, \tag{3.56}$$

where $i = \sqrt{-1}$.

The special cases of the gamma distribution are known as distributions with their own names. If p is an integer, it is the Erlang distribution. For $p = 1$ we obtain the exponential distribution. In a special case of $b = 2$ and an integer $p = \frac{n}{2}$ the gamma distribution becomes the chi-squared distribution with n degrees of freedom.

Fig. 3.12 The chi-squared distribution with n degrees of freedom
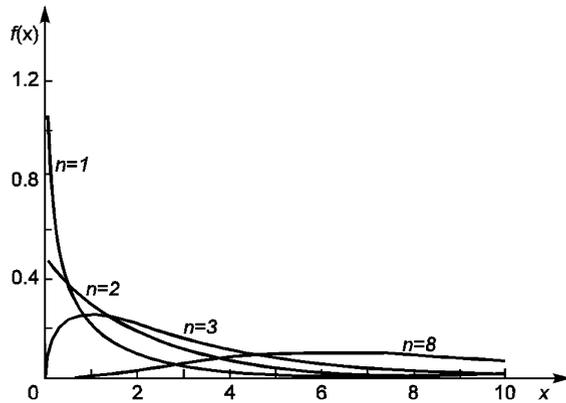
3.4.2.5 The Chi-Squared Distribution

The probability density of the chi-squared distribution has been determined by F.R. Helmert in 1876, [10]. We say that the random variable X has the *chi-squared distribution with n degrees of freedom* (denoted by $\chi^2(n)$) for $n = 1, 2, \ldots$, if its probability function is

$$f(x) = \begin{cases} 0 & \text{for } x < 0, \\ \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{n/2-1} e^{-x/2} & \text{for } x \geq 0, \end{cases} \qquad (3.57)$$

where the gamma function $\Gamma(p)$ for $p > 0$ is defined in (3.53).

In Fig. 3.12 is shown the probability density function of the chi-squared distribution with $n = 1, 2, 3$, and 8 degrees of freedom.

The mean value and the variance of the chi-squared random variable X with n degrees of freedom are

$$m_X = n, \qquad \sigma_X^2 = 2n. \qquad (3.58)$$

The convergence of the distribution of the chi-squared random variable X for the number of degrees of freedom tending to infinity is its important property. The probability distribution function of the random variable $(X - n)/\sqrt{2n}$ tends to the cumulative distribution function of the normal distribution $N(0, 1)$ for $n \to \infty$.

The chi-squared probability distribution, except for the ellipses and ellipsoids of the probability concentration presented in this book, is widely applied in mathematical statistics, especially in the chi-squared distribution agreement tests. In Chap. 5 we shall obtain in a natural way a random variable with the chi-squared distribution during summation of squares of normal random variables. In this section we have introduced its formal definition given by the probability density function (3.57).

3.4.2.6 The Weibull Distribution

The Weibull distribution is an example of a two-parametric probability distribution.
We say that a random variable X has the *Weibull distribution* with parameters $b > 0$
and $c > 0$, if its probability density function has the form

$$f(x) = \begin{cases} 0 & \text{for } x < 0, \\ cb\,(cx)^{b-1}\exp[-(cx)^b] & \text{for } x \geq 0. \end{cases} \tag{3.59}$$

In Fig. 3.13 the probability density function of the Weibull distribution is shown
for the fixed value of parameter $c = 1$ and several values $b = 0.5$, 1, and 2.

The mean value and the variance of the random variable with the Weibull distri-
bution are

$$m_X = \frac{1}{c}\Gamma\left(\frac{1}{b} + 1\right),$$

$$\sigma_X^2 = \frac{1}{c}^2\left\{\Gamma\left(\frac{2}{b} + 1\right) - \left[\Gamma\left(\frac{1}{b} + 1\right)\right]^2\right\}, \tag{3.60}$$

where the function $\Gamma(p)$ was defined in (3.53).

Example 3.8 (Generation of random numbers: the Weibull distribution) In a way
analogous to Example 3.7 we can generate a sequence of random numbers of the
Weibull distribution. Having generated $\{y_n,\ n = 1, 2, \ldots\}$, the sequence of indepen-
dent realizations of random variable Y with a distribution uniform along the interval
$[0, 1]$ we obtain the numbers of the Weibull distribution by inverting the cumulative
distribution function, $\{x_n = F^{-1}(y_n),\ n = 1, 2, \ldots\}$. Now the function $y = F(x)$ is

$$y = F(x) = \begin{cases} 0 & \text{for } x < 0, \\ 1 - \exp[-(cx)^b] & \text{for } x \geq 0, \end{cases}$$

and its inverse function $x = F^{-1}(y)$ is:

$$x = F^{-1}(y) = \frac{1}{c}(\log y)^{\frac{1}{b}} \quad \text{for } y \in [0, 1].$$

The Weibull distribution is widely applied in the probabilistic descriptions of stochastic reliability problems of engineering structures. More information on this subject can be found in the monographs concerned with this field of research, e.g., [5, 20, 25], and also [19]. The books contain some definitions, and examples of applications of the probability distributions characteristic for the reliability theory, such as the lognormal, gamma, and Gumbel distributions.

3.4.2.7 The Gumbel Distribution

The *Gumbel distribution* is a case of extreme value distribution. It is used to model the distribution of the maximum value of observations of variables with different distributions. The Gumbel distribution is useful for description of the maximum level of a river at a period on a basis of historical observations, for predicting extreme earthquakes, and in mechanics, for modeling survival time of a structure in reliability theory.

The probability density function $f(x)$ of the Gumbel distribution is defined on the whole real line, $x \in R$, and, in a general case, it depends on two parameters, $a \in R$ and $b > 0$. It has the following form:

$$f(x) = \frac{1}{b}\exp\left[-\frac{x-a}{b}\right]\exp\left[-\exp\left[-\frac{x-a}{b}\right]\right]. \tag{3.61}$$

The plot of the probability density function of the Gumbel distribution for $a = 0$, $b = 1$ and $a = 1$, $b = 2$ is shown in Fig. 3.14.

Its cumulative distribution function $F(x)$ is given in (3.62);

$$F(x) = \exp\left[-\exp\left[-\frac{x-a}{b}\right]\right]. \tag{3.62}$$

The lowest-order moments of the Gumbel distribution are:

$$m_X = a + b\gamma, \qquad \sigma_X^2 = \frac{\pi^2}{6}b^2, \tag{3.63}$$

where γ is the Euler-Mascheroni constant, $\gamma \approx 0.57721$, see [1].

In a special case $a = 0$ and $b = 1$ we have the standard Gumbel distribution with the probability density and the cumulative distribution functions, respectively,

$$f(x) = e^{-x}e^{-e^{-x}}, \tag{3.64}$$

$$F(x) = e^{-e^{-x}}. \tag{3.65}$$
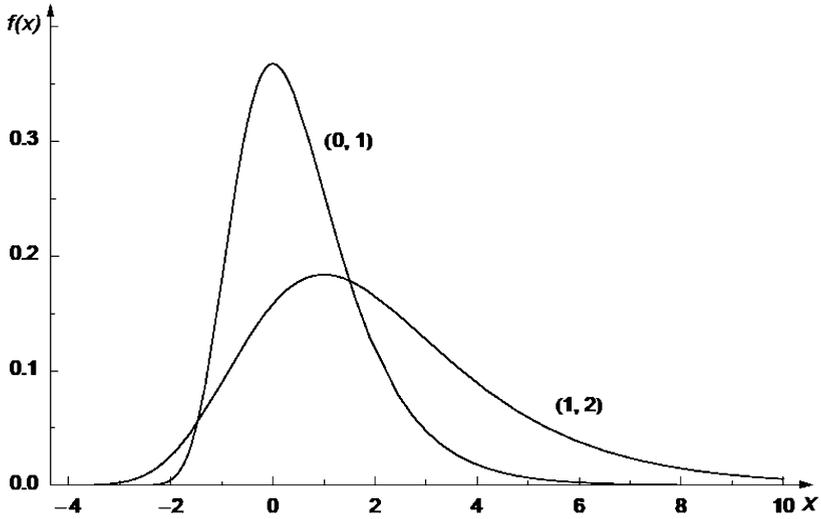
Fig. 3.14 Probability density function of the Gumbel distribution for $a = 0$, $b = 1$ and $a = 1$, $b = 2$

In this case the moments (3.63) are

$$m_X = \gamma, \qquad \sigma_X^2 = \frac{\pi^2}{6},$$
(3.66)

which gives one of the equivalent definitions of the Euler-Mascheroni constant, γ.

3.4.2.8 The Student t-Distribution

One of distributions playing an important role in testing statistical hypotheses is the Student t-distribution. We say that the random variable X has the *Student t-distribution with n degrees of freedom* if its probability density function is of the form

$$f(x) = \frac{\Gamma([n+1]/2)}{\sqrt{n\pi}\,\Gamma(n/2)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}},$$
(3.67)

for $x \in R$.

The probability density function (3.67) has a shape similar to the normal probability density function, and for $n \to \infty$ it tends to $N(0, 1)$, so in practice, for large n ($n > 30$) this distribution can be replaced by the normal distribution.

The basic parameters of the Student t-distribution: the mean value, variance, skewness and kurtosis exist only for a sufficiently large value of the parameter n. They are, respectively,

$$m_X = 0, \quad \text{for } n > 1,$$
(3.68)
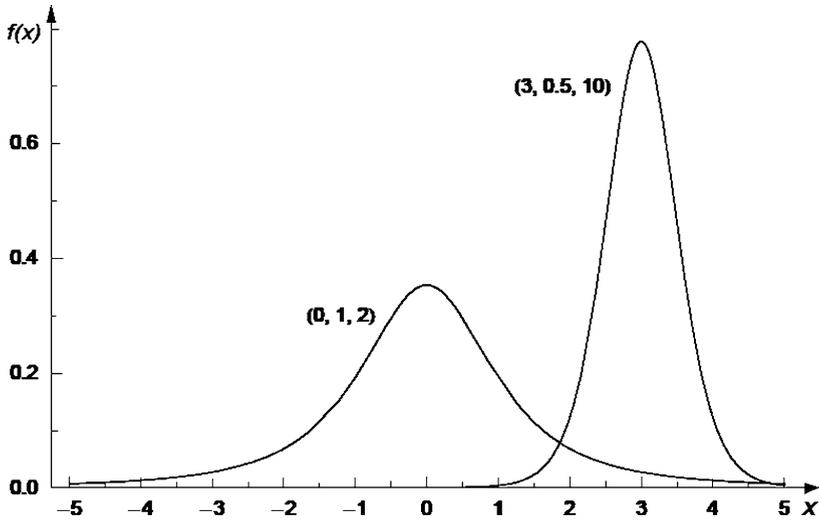
Fig. 3.15 Probability density function of the Student t-distribution

$$\sigma_X^2 = \frac{n}{n-2}, \quad \text{for } n > 2, \tag{3.69}$$

$$\gamma_1 = 0, \quad \text{for } n > 3, \tag{3.70}$$

$$\gamma_2 = \frac{6}{n-4}, \quad \text{for } n > 4. \tag{3.71}$$

For the lowest values of the parameter n the Student t-distribution takes an especially simple form. For $n = 1$ we obtain the *arcus tangens distribution* with the probability density function

$$f(x) = \frac{1}{\pi(1+x^2)}, \tag{3.72}$$

and the cumulative distribution function:

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan(x), \tag{3.73}$$

for $x \in R$.

For $n = 2$ these functions are, respectively:

$$f(x) = \frac{1}{(2+x^2)^{3/2}}, \tag{3.74}$$

$$F(x) = \frac{1}{2}\left[1 + \frac{x}{\sqrt{2+x^2}}\right], \tag{3.75}$$

also defined for $x \in R$.
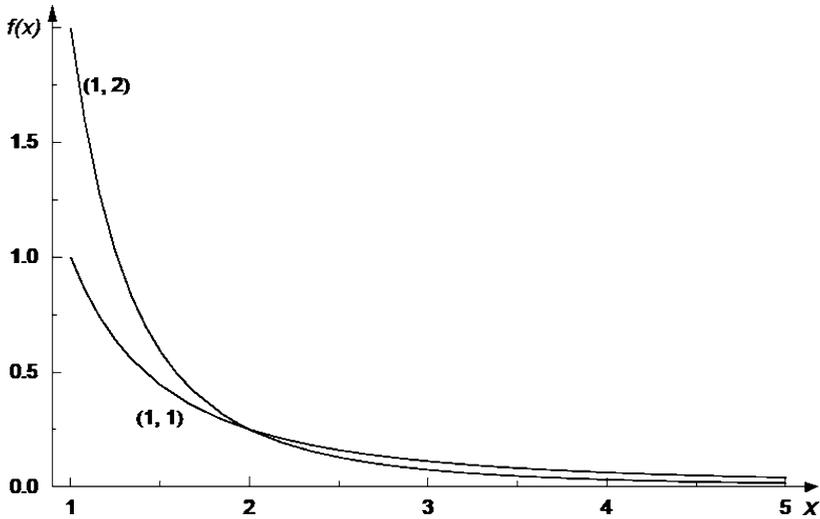
Fig. 3.16 Probability density function of the Pareto distribution for $a = 1$, $p = 1$ and $a = 1$, $p = 2$

Examples of application of the Student t-distribution can be found in any handbook of mathematical statistics, e.g., [7, 23].

3.4.2.9 The Pareto Distribution

The *Pareto distribution* (also called the *Bradford distribution*) describes random variables appearing in scientific, geophysical and also economical phenomena. It is often used for description of allocation of resources, dimension of grains of a material (e.g., sizes of sand particles) and in computer sciences, for modeling lengths of packages in TCP protocol or execution time of jobs in computer processing.

The probability density function of the Pareto distribution is defined on the positive half-line and has two parameters $a > 0$, referring to the scale of the distribution, and $p > 0$, reflecting its shape. The probability density function has the following form:

$$f(x) = \begin{cases} 0 & \text{for } x < a, \\ \frac{pa^p}{x^{(p-1)}} & \text{for } x \geq a. \end{cases} \tag{3.76}$$

An example of such a probability density function is presented in Fig. 3.16 for two pairs of the parameters: $a = 1$, $p = 1$ and $a = 1$, $p = 2$.

Equation (3.76) represents an elementary function and therefore the Pareto distribution is easy in calculations, which made it popular in many practical applications. It is also easy to obtain the exact equations for the cumulative distribution function of the Pareto distribution and its moments and to calculate its entropy. The mean

value and the variance are:

$$m_X = \frac{pa}{p-1}, \quad \text{for } p > 1, \tag{3.77}$$

$$\sigma_X^2 = \frac{pa^2}{(p-1)^2(p-2)}, \quad \text{for } p > 2. \tag{3.78}$$

From the third-order and fourth-order moments we can calculate the skewness and the kurtosis of the Pareto distribution. They exist for sufficiently large values of p:

$$\gamma_1 = \frac{2(1+p)2}{\sqrt{p-3}}\sqrt{\frac{p-1}{p}}, \quad \text{for } p > 3, \tag{3.79}$$

$$\gamma_2 = \frac{6(p^3 + p^2 - 6p - 2)}{p(p-3)(p-4)}, \quad \text{for } p > 4. \tag{3.80}$$

3.4.3 Remarks on Other Probability Distributions

In practical applications, especially in statistics, numerous probability distributions, also other than those presented in the previous section, are being used. Some of them, characterized by particular features of frequent occurrence, have been given names, e.g., "gamma", "Fisher's", "Rayleigh" distributions. Such distributions are dealt with in books on general problems of statistics, or in special specifications of distributions. A comprehensive list of various distributions along with the index of applications in mathematical statistics can be found in [9].

Any continuous probability distribution can be defined by choosing a proper non-negative function $g(x)$ of the variable x, integrable along the entire line

$$\int_{-\infty}^{\infty} g(x)\,dx = C < \infty \tag{3.81}$$

and then by normalizing it

$$f(x) = \frac{g(x)}{C}, \tag{3.82}$$

where C is a constant calculated in (3.81). Such a function $f(x)$ may be taken as the probability density function of a specific random variable.

Example 3.9 (The piecewise linear distribution) Let us assume that the function $g(x)$ is piecewise linear along a number N of consecutive intervals, and is equal to zero for other values of x:

$$g(x) = \begin{cases} a_j x + b_j & \text{for } x \in [l_j, r_j), \ j = 1, 2, \dots N, \\ 0 & \text{for other values of } x, \end{cases} \tag{3.83}$$

Fig. 3.17 An example of the
piecewise linear distribution
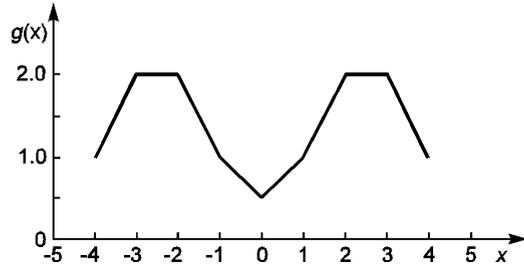
where $l_1 < r_1 \le l_2 < r_2 \le l_3 < \cdots < r_{N-1} \le l_N < r_N$. An example of such a function ($N = 8$) is given in Fig. 3.17.

The end points of adjacent sectors can overlap, that is, for some j we may have $r_j = l_{j+1}$. The constant C which normalizes the function $g(x)$ for obtaining the probability density function is

$$C = \int_{-\infty}^{\infty} g(x)\,dx = \sum_{j=1}^{N} \frac{1}{2} a_j \left(r_j^2 - l_j^2 \right) + \sum_{j=1}^{N} b_j \left(r_j - l_j \right). \tag{3.84}$$

Now lower-order moments of the distribution defined in this manner can be calculated: the expected value is

$$E\{X\} = m = \frac{1}{3C} \sum_{j=1}^{N} a_j \left(r_j^3 - l_j^3 \right) + \frac{1}{2C} \sum_{j=1}^{N} b_j \left(r_j^2 - l_j^2 \right), \tag{3.85}$$

the second-order ordinary moment is

$$E\{X^2\} = \frac{1}{4C} \sum_{j=1}^{N} a_j \left(r_j^4 - l_j^4 \right) + \frac{1}{3C} \sum_{j=1}^{N} b_j \left(r_j^3 - l_j^3 \right), \tag{3.86}$$

and the variance (the second-order central moment) is

$$\sigma^2 = E\{X^2\} - m^2. \tag{3.87}$$

The piecewise linear distribution may be used as an approximation of an arbitrary distribution by assuming a sufficiently dense division into a number of sectors. Such a procedure may be applied for an approximation of the experimental data (histograms) of a distribution defined by a complex function which may be difficult for practical calculations.

Example 3.10 (The triangular distribution) Let a piecewise linear function $g(x)$, discussed in the previous example, be limited to two symmetrically located sectors only. These sectors have a common point at $x = 0$, and their outer end points lie on the x-axis. The probability density depends on one parameter, namely, the length of

Fig. 3.18 The triangular
distribution
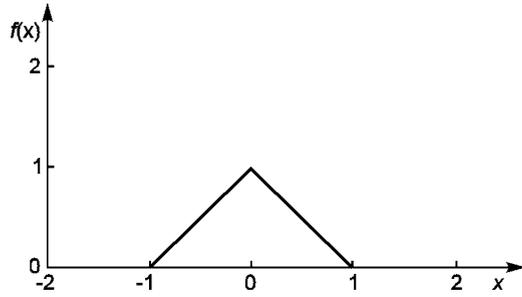
the single interval l, and it has the form

$$f(x) = \begin{cases} \frac{1}{l^2}x + \frac{1}{l} & \text{for } x \in [-l, 0), \\ -\frac{1}{l^2}x + \frac{1}{l} & \text{for } x \in [0, l), \\ 0 & \text{otherwise.} \end{cases} \qquad (3.88)$$

An example of such a probability density function (for $l = 1$) is presented in
Fig. 3.18.

Moments of odd orders (including the mean value) of the triangular distribution
are equal to zero owing to its symmetry. The variance of that distribution is given
by the formula

$$\sigma^2 = \frac{1}{6}l^2, \qquad (3.89)$$

while the fourth-order moment becomes

$$E\{x^4\} = \frac{1}{15}l^4. \qquad (3.90)$$

3.4.4 Measures of Deviation from the Normal Distribution

Results of measurements or various accuracy tests form usually a sequence of num-
bers. Once knowing them one can construct a corresponding histogram, which con-
stitutes a certain approximation of the probability density function. One can also
calculate the so-called estimators of these sequences of numbers, such as the ex-
pected value, variance, and moments of various orders, by using formulas analogous
to (1.2) and (1.4a). It is often useful to know which of the theoretical distributions of
random variables corresponds well to that resulting from the specific measurements.
This problem concerns mainly the question whether the result of measurements may
be treated as the normal distribution.

Comparing our empirical distributions with the normal distribution we should
analyze certain similarities. The diagram of the probability density of the normal
distribution (Fig. 2.4) displays certain characteristic features: a single maximum
(it means that it belongs to the class of one-modal distributions), symmetry with

Table 3.2 The asymmetry γ_1 and the excess γ_2 coefficients for certain distributions

Distribution	γ_1	γ_2
Gaussian	0	0
Rectangular	0	$\frac{9}{5}$
Lognormal	$(w+2)\sqrt{w-1}$	$w^4 + 2w^3 + 3w^2 - 6$ for $w = \exp[\sigma^2]$
Exponential	2	6
Triangular	0	$-\frac{3}{5}$

respect to its central point, and the fact that it is concentrated in the vicinity of the average value. Thus, if an experimental distribution is to be treated as the normal distribution, it has to display the features mentioned above. Moreover, the degree of symmetry and flattening of the probability density graph of the experimental distribution may be calculated.

As the measure of the asymmetry of the probability density function, the coefficient γ_1 defined in (3.15) is used. For a random variable with the normal distribution and for each other random variable with a distribution symmetrical with respect to its central point, the asymmetry (inclination) coefficient γ_1 is equal to zero. The more γ_1 differs from zero, the more the distribution in question differs from the normal distribution, see Fig. 3.1.

As the measure of flattening of the probability density function, the coefficient γ_2 defined by the relation (3.16) is used. Like the inclination coefficient γ_1, the excess (flattening) coefficient γ_2 equals zero in the case of the normal distribution. The more γ_2 differs from zero the more the distribution under consideration differs from the normal distribution, see Fig. 3.2.

After finding the third-order and the fourth-order moments of certain distributions, we can find the values of their asymmetry and excess coefficients. For some distributions these formulas are given in Table 3.2.

It is seen that the coefficients γ_1 and γ_2 do not depend upon the variance for the rectangular, the exponential, and the triangular distributions, while they do depend upon it in the case of the lognormal distribution. Comparing the coefficients γ_1 and γ_2 for the uniform and the lognormal distributions we can say that the uniform distribution is "more normal" than the lognormal one, since its asymmetry and excess coefficients are closer to zero than the analogous quantities of the lognormal distribution. We can also say that the triangular distribution is more flat around the modal value ($x = 0$) than the normal distribution, which is confirmed by the negative value of the excess coefficient γ_2. The positive value of γ_2 for the uniform distribution indicates that it is more concentrated around zero than the normal distribution with the same variance.

To give a definitive answer whether the results of measurements can be treated as a normally distributed set of data needs application of statistical tests concerning the problem of normality.

3.5 Approximate Methods for Constructing a Probability Density Function

Many types of probability distributions are used in engineering applications. After gathering experimental data one usually tries to combine the set of numbers and some a priori assumed distribution fitting the density curve to the data by an appropriate selection of the distribution parameters. The statistical regression gives us a possibility of such an approximate construction of the probability density functions. Other possibilities of generation of the densities on the basis of experimental data are asymptotic expansions of the functions.

One of the possible expansions is a representation of the probability density function in the form of a series of derivatives of the normal distribution probability density function. Such a series is called the Gram-Charlier series (see [4]). The given probability density function $f(x)$ can be represented in the form of the series

$$f(x) = c_0 \phi(x) + \frac{c_1}{1!} \phi'(x) + \frac{c_2}{2!} \phi''(x) + \cdots, \tag{3.91}$$

where $\phi(x)$ is the probability density function of the standard Gaussian distribution,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right], \tag{3.92}$$

and the coefficients c_i can be expressed by the Hermitian polynomials

$$c_i = (-1)^i \int_{-\infty}^{\infty} H_i(x) f(x) \, dx. \tag{3.93}$$

The Hermitian polynomials are defined by the following equation:

$$\left(\frac{d}{dx}\right)^n \exp\left[-\frac{x^2}{2}\right] = (-1)^n H_n(x) \exp\left[-\frac{x^2}{2}\right]. \tag{3.94}$$

The coefficients c_i in the Gram-Charlier expansion are polynomials in terms of moments of the corresponding probability density function $f(x)$. For the normalized distribution (that is, with a zero mean and unit variance) they have the following form:

$$
\begin{aligned}
&c_0 = 1, \qquad c_1 = c_2 = 0, \\
&c_3 = -\frac{\mu_3}{\sigma^3} = \gamma_1, \qquad c_4 = \frac{\mu_4}{\sigma^4} - 3 = \gamma_2, \\
&c_5 = -\frac{\mu_5}{\sigma^5} + 10\frac{\mu_3}{\sigma^3}, \qquad c_6 = \frac{\mu_6}{\sigma^6} - 15\frac{\mu_4}{\sigma^4} + 30.
\end{aligned}
\tag{3.95}
$$

As it is seen, knowing several lower-order moments of a certain distribution, we can formulate its approximate probability density function using the expansion (3.91).

Let us remark that in the Gram-Charlier expansion, two initial non-zero coefficients next to the derivatives of the density of the normal distribution are the known inclination γ_1 and the excess γ_2 coefficients. It is seen that a distribution having the same inclination (that is, symmetric around the mean value) and excess like the normal distribution, differ from the normal one by terms starting from order 5.

The Gram-Charlier asymptotic expansion is not the only possible one. For example, the probability density function can be represented in the form of the following Edgeworth series (see [4]):

$$
\begin{aligned}
f(x) = \phi(x) &- \frac{1}{3!}\frac{\mu_3}{\sigma^3}\phi^{(3)} + \frac{1}{4!}\left(\frac{\mu_4}{\sigma^4} - 3\right)\phi^{(4)} + \frac{10}{6!}\left(\frac{\mu_3}{\sigma^3}\right)^2\phi^{(6)} \\
&- \frac{1}{5!}\left(\frac{\mu_5}{\sigma^5} - 10\frac{\mu_3}{\sigma^3}\right)\phi^{(5)} - \frac{35}{7!}\frac{\mu_3}{\sigma^3}\left(\frac{\mu_4}{\sigma^4} - 3\right)\phi^{(7)} \\
&- \frac{280}{9!}\left(\frac{\mu_3}{\sigma^3}\right)^3\phi^{(9)} + \cdots .
\end{aligned}
\tag{3.96}
$$

Another possibility of construction of the approximate expression for the probability density function, provided several lower-order moments m_k, $k = 1, 2, \ldots, N$ are given, is based on application of the maximum entropy principle (see [24]). In this method we choose this probability density function from densities having moments equal to the given ones, which has the maximum value of the entropy H, defined in (3.21). Any probability density function satisfying the maximum entropy principle can be represented in the form

$$
p(x) = C \exp\left[-\sum_{k=1}^{N}\lambda_k x^k\right].
\tag{3.97}
$$

To find the probability density function for a given set of moments m_k, we must choose the coefficients λ_k in such a way that

$$
\int x^n C \exp\left[-\sum_{k=1}^{N}\lambda_k x^k\right]dx = m_n, \quad n = 1, 2, \ldots, N.
\tag{3.98}
$$

Moreover, the constant C should normalize the density $p(x)$ in such a way that its integral is equal to 1:

$$
\frac{1}{C} = \int \exp\left[-\sum_{k=1}^{N}\lambda_k x^k\right]dx.
\tag{3.99}
$$

In the above formulas the limits of integrals are chosen according to the areas where the probability density function searched for differs from zero.

As it is seen from the above considerations, in the maximum entropy method the construction of the probability density function leads to solving the system of $N + 1$ equations for the constants C and λ_k, $k = 1, 2, \ldots, N$.

To conclude, let us remark that the problem of convergence of asymptotic expansions of probability densities is very difficult, and some known conditions of the convergence are hard to verify or are non-effective. For this reason in many practical tasks, taking into account only a few terms of the expansion of the probability density function can give a better result than summation of too many terms.

3.6 Multi-dimensional Probability Distributions

In this section we present some introductory, theoretical material which will be developed in Chaps. 5 and 7, where we shall analyze problems of the error analysis in engineering applications.

Consider an n-dimensional vector, where each of its coordinates is a random variable:

$$\mathbf{X} = [X_1(\omega), X_2(\omega), \ldots, X_n(\omega)]^T. \tag{3.100}$$

The vector $\mathbf{X}$ will be called the n-dimensional random vector or the n-dimensional random variable. The distribution of multi-dimensional random variables, similarly to the one-dimensional case, can be described by the probability distribution function and, in the continuous case, by the n-dimensional probability density function. The (*joint*) *probability distribution function* of the n-dimensional random variable $\mathbf{X}$ is defined as:

$$F(x_1, x_2, \ldots, x_n) = P(\{\omega : X_1(\omega) < x_1, X_2(\omega) < x_2, \ldots, X_n(\omega) < x_n\}). \tag{3.101}$$

Similarly to the one-dimensional case, the (*joint*) *probability density function* $f(x_1, x_2, \ldots, x_n)$ is such a function, which allows us to represent the probability distribution function in the form of the following integral:

$$F(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(\xi_1, \xi_2, \ldots, \xi_n) \, d\xi_1 d\xi_2 \cdots d\xi_n. \tag{3.102}$$

As it results from the definition of the vector-valued random variable, each of its coordinates is a (one-dimensional) random variable. We can also define m-dimensional random variables, $m < n$, consisting of m coordinates of n-dimensional vector $\mathbf{X}$. It is possible to obtain the probability distribution function (probability density function) of such an m-dimensional vector from the probability distribution (density) function of the initial n-dimensional random vector. It can be done in a way analogous to that presented in the following example.

Example 3.11 (The marginal probability distributions) Consider two-dimensional vectorial random variable $[X(\omega), Y(\omega)]^T$ with the cumulative distribution function $F_{XY}(x, y)$ and the probability density function $f_{XY}(x, y)$. Each of the random variables $X(\omega)$ and $Y(\omega)$ has its own cumulative distribution function and its own probability density function. These functions can be obtained by integration of the cumulative distribution function (respectively: the probability density function) of the

random vector with respect to the corresponding independent variable y or x:

$$F_X(x) = F_{XY}(x, \infty) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{XY}(\xi, y) \, d\xi dy, \quad f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, dy,$$

(3.103)

$$F_Y(y) = F_{XY}(\infty, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{y} f_{XY}(x, \zeta y) \, dx d\zeta, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, dx.$$

(3.104)

The obtained cumulative distribution functions $F_X(x)$ and $F_Y(y)$ are called the marginal (or projective) distribution functions and the probability densities $f_X(x)$ and $f_Y(y)$ are the marginal probability densities.

We say that two continuous random variables $X(\omega)$ and $Y(\omega)$ are independent (statistically independent[2]) if their joint bivariate probability density function is the product of their marginal (one-dimensional) probability density functions

$$f_{XY}(x, y) = f_X(x) f_Y(y).$$

(3.105)

Similarly, we say that n continuous random variables are independent if their joint n-dimensional probability density function is a product of n marginal, univariate probability density functions.

A weaker condition on a pair of random variables (resulting from their independence) is non-correlation. We will say that the random variables X and Y are uncorrelated, if their mutual correlation is equal to zero:

$$E\left\{(X - \overline{X})(Y - \overline{Y})\right\} = 0.$$

(3.106)

In a general case, uncorrelated random variables need not be independent. An exception here is the normal distribution, where the conditions of independence and uncorrelation are equivalent.

Similarly to the conditional probabilities, we can define the *conditional probability density functions*. For the two-dimensional random variable $[X(\omega), Y(\omega)]^T$ the conditional probability density function of $X(\omega)$ under the condition $Y(\omega) = y$, that is, $f(x|Y(\omega) = y)$, has the following form:

$$f(x|Y(\omega) = y) = \frac{f(x, y)}{f_Y(y)} = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dx}.$$

(3.107)

The conditional probability density function allows to calculate the joint probability of two random variables X and Y in two steps:

$$P(X \in A, Y \in B) = \int_B \int_A f(x, y) dx dy = \int_B \int_A f(x|Y = y) dx f_Y(y) dy,$$

(3.108)

---

[2]We say: "statistically independent" or "statistically dependent" in opposition to "functional dependence". Two quantities are functionally dependent if one is a certain function of another.

that is, first calculating the conditional probability of one random variable and then integrating the obtained result with respect to the marginal probability density of another random variable.

Similarly to the calculation of the expected value (as an integral with respect to the probability density function), we can calculate the conditional expectation (the integral with respect to the conditional probability density function):

$$E\{X|Y=x\} = \int_{-\infty}^{\infty} xf(x|Y=y)\,dx. \tag{3.109}$$

The conditional expected value is often called *regression*.

It results from the definition of the conditional probability density function that if the random variables X and Y are statistically independent, then the conditional probability density function of X under a condition on Y equals the unconditioned probability density function of X:

$$f(x|Y(\omega)=y) = \frac{f(x,y)}{f_Y(y)} = \frac{f_X(x)f_Y(y)}{f_Y(y)} = f_X(x). \tag{3.110}$$

Remark 3.1 In this book, which is devoted to problems of measurement errors, we concentrate on continuous probability distributions and related problems. However, most of such problems have their countermeasures in discrete probability distributions, the same as in a case of multi-dimensional probability distributions. If a multi-dimensional random variable of the form (3.100) is discrete, then its probability distribution can be represented in the following form:

$$p_{i_1 i_2 \cdots i_n} = P\left(\{\omega : X_1(\omega) = x_{i_1}, X_2(\omega) = x_{i_2}, \ldots, X_n(\omega) = x_{i_n}\}\right), \tag{3.111}$$

where $x_{i_1}, x_{i_2}, \ldots, x_{i_n}$, $i_1, i_2, \ldots, i_n = 1, 2, \ldots$, are discrete values that are taken by corresponding coordinates of the random variable $\mathbf{X}$. Consider an example of two-dimensional random variable (X, Y) with discrete probability distribution. This distribution is completely described by the matrix of probabilities $\mathbf{P} = [p_{ij}]$ (we assume that the random variable can take only a finite number of values):

$$p_{ij} = P\left(\{\omega : X(\omega) = x_i, Y(\omega) = y_j\}\right), \ i = 1, 2, \ldots, n, \ j = 1, 2, \ldots, m. \tag{3.112}$$

For such a random variable, likewise as for random variables with a continuous distribution, we can calculate the marginal probability distributions

$$p_i = P(\{\omega : X(\omega) = x_i\}) = \sum_{j=1}^{m} p_{ij}, \quad i = 1, 2, \ldots, n,$$

$$p_j = P\left(\{\omega : Y(\omega) = y_j\}\right) = \sum_{i=1}^{n} p_{ij}, \quad j = 1, 2, \ldots, m, \tag{3.113}$$

conditional probability distributions

$$p\left(x_i|y_j\right) = P\left(\{\omega : X\left(\omega\right) = x_i|Y\left(\omega\right) = y_j\}\right) = \frac{p_{ij}}{p_j}, \quad i = 1, 2, \ldots, n \tag{3.114a}$$

calculated for every y_j, $\quad j = 1, 2, \ldots, m$,

and

$$p\left(y_j|x_i\right) = P\left(\{\omega : Y\left(\omega\right) = y_j|X\left(\omega\right) = x_i\}\right) = \frac{p_{ij}}{p_i}, \quad j = 1, 2, \ldots, m \tag{3.114b}$$

calculated for every x_i, $\quad i = 1, 2, \ldots, n$,

and also calculate the conditional expected values:

$$E\left\{X|y_j\right\} = E\left\{X|Y = y_j\right\} = \sum_{i=1}^{n} x_i\, p\left(x_i|y_j\right). \tag{3.115}$$

As in the one-dimensional case, we can consider vectorial random variables being functions of other vectorial random variables. Assume now that the n-dimensional random variable Y is a vector-valued, single-valued function of the n-dimensional random variable X. Then the probability density function of the vector Y has the following form:

$$f_Y\left(y_1, y_2, \ldots, y_n\right) = |D|\; f_X\left(x_1, x_2, \ldots, x_n\right), \tag{3.116}$$

where the variables $x_1, x_2, \ldots, x_n$ are expressed by $y_1, y_2, \ldots, y_n$ and D is the Jacobian determinant of the transformation of variables

$$D = \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(y_1, y_2, \ldots, y_n)}. \tag{3.117}$$

Example 3.12 (Difference of two normal random variables) If the vectorial random variable $[X\left(\omega\right), Y\left(\omega\right)]^T$ has the joint probability density function of the following Gaussian form:

$$f_{XY}\left(x, y\right) = \frac{1}{2\pi\,\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right], \tag{3.118}$$

then the random variable Z defined by $Z = X - Y$ has the density of the form:

$$f_Z\left(z\right) = \frac{1}{\sqrt{2\pi}\,\sigma_z}\exp\left(-\frac{z^2}{2\sigma_z^2}\right), \tag{3.119}$$

where

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2. \tag{3.120}$$

Fig. 3.19 The domain of the
two-dimensional distribution
for which the marginal
distributions are uncorrelated
but dependent
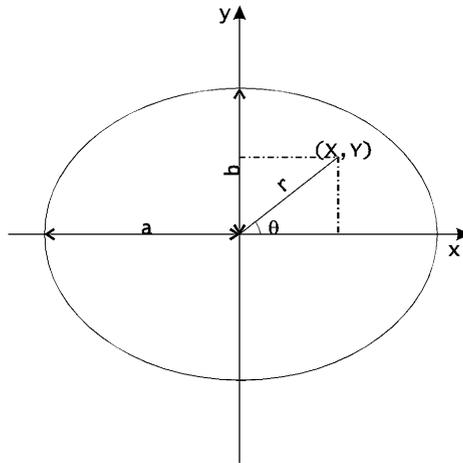
Example 3.12 shows that the difference (sum) of two independent normal random
variables is also a normal random variable.

The following examples show that two random variables can be uncorrelated but
statistically dependent.

Example 3.13 (Random variables uncorrelated but dependent) Consider the two-di-
mensional uniform distribution over the surface of the ellipse with its semi-axes a
and b lying on the system of coordinate axes x and y (Fig. 3.19). The probability
density function of this distribution has the following form:

$$f_{XY}(x, y) = \begin{cases} \frac{1}{\pi ab} & \text{for } \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1, \\ 0 & \text{for } \frac{x^2}{a^2} + \frac{y^2}{b^2} > 1. \end{cases}$$

Now we can calculate the marginal probability distributions of the random vari-
ables X and Y. From Fig. 3.19 we see that $f_X(x) = 0$ for $|x| > a$ and it can be
calculated for $|x| \leq a$ by appropriate integration of the two-dimensional probability
density function $f_{X,Y}(x, y)$ with respect to the second variable:

$$f_X(x) = \int_{-b\sqrt{1-x^2/a^2}}^{b\sqrt{1-x^2/a^2}} \frac{dy}{\pi ab} = \frac{2\sqrt{1-x^2/a^2}}{\pi a}.$$

By an analogous integration we can obtain the marginal probability density function
for the random variable Y. Finally, the marginal probability density functions can
be written as:

$$f_X(x) = \begin{cases} \frac{2\sqrt{1-x^2/a^2}}{\pi a} & \text{for } |x| \leq a, \\ 0 & \text{for } |x| > a, \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{2\sqrt{1-y^2/b^2}}{\pi b} & \text{for } |y| \le b, \\ 0 & \text{for } |y| > b. \end{cases}$$

Now we can see that the random variables X and Y are dependent, because

$$f_{X,Y}(x, y) \ne f_X(x) f_Y(y).$$

To calculate the covariance of the random variables X and Y,

$$\text{cov}(X, Y) = \iint_{\frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1} xy f_{XY}(x, y) dx dy,$$

we change the variables:

$$x = ar\cos\theta,$$

$$y = br\sin\theta,$$

and obtain the integral

$$\text{cov}(X, Y) = \int_0^{2\pi} \int_0^1 ar br \cos\theta \sin\theta \frac{1}{\pi ab} ab r dr d\theta = 0.$$

It is seen that X and Y are uncorrelated.

Example 3.14 (Random variables uncorrelated but dependent) Consider two independent, uniformly distributed random variables $R(\omega)$ and $\theta(\omega)$, the first one over the interval $0 < r_1 \le R(\omega) \le r_2$ and the second one over the interval $0 < \theta(\omega) \le 2\pi$. From these random variables we construct two new random variables $X(\omega)$ and $Y(\omega)$ being functions of the random variables $R(\omega)$ and $\theta(\omega)$, defined as:

$$X(\omega) = R(\omega)\cos\theta(\omega),$$

$$Y(\omega) = R(\omega)\sin\theta(\omega).$$

The joint distribution of the pair of random variables $(R(\omega), \theta(\omega))$ is uniform, with the probability density function defined as:

$$f_{R\theta}(r, \theta) = \begin{cases} \frac{1}{2\pi(r_2 - r_1)} & \text{for } (r, \theta) \in [r_1, r_2] \times (0, 2\pi], \\ 0 & \text{otherwise,} \end{cases}$$

The domain of this probability density function is presented in Fig. 3.20.

The probability density function of the pair of random variables $(X(\omega), Y(\omega))$ can be obtained by a formula analogous to (3.26) and (3.116). It has the form:

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{\pi(r_2^2 - r_1^2)} & \text{for } r_1^2 \le x^2 + y^2 \le r_2^2, \\ 0 & \text{otherwise.} \end{cases}$$

Fig. 3.20 The domain of the two-dimensional distribution for which the marginal distributions are uncorrelated but dependent
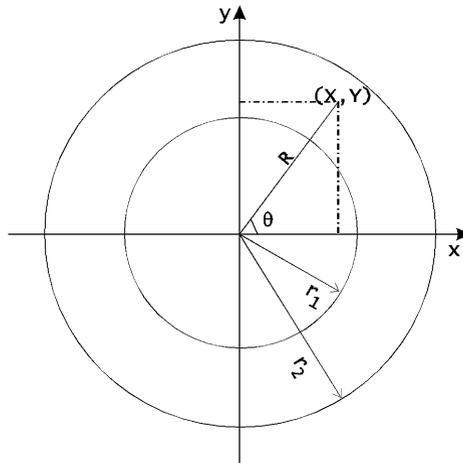
We can calculate the marginal densities for each of the variables $X(\omega)$ and $Y(\omega)$. They are:

$$f_X(x) = \begin{cases} \dfrac{2\sqrt{r_2^2-x^2}}{\pi(r_2^2-r_1^2)} & \text{for } x \in [-r_2, -r_1], \\[3mm] \dfrac{2(\sqrt{r_2^2-x^2}-\sqrt{r_1^2-x^2})}{\pi(r_2^2-r_1^2)} & \text{for } x \in (-r_1, r_1), \\[3mm] \dfrac{2\sqrt{r_2^2-x^2}}{\pi(r_2^2-r_1^2)} & \text{for } x \in [r_1, r_2], \\[3mm] 0 & \text{otherwise,} \end{cases}$$

$$f_Y(y) = \begin{cases} \dfrac{2\sqrt{r_2^2-y^2}}{\pi(r_2^2-r_1^2)} & \text{for } y \in [-r_2, -r_1], \\[3mm] \dfrac{2(\sqrt{r_2^2-y^2}-\sqrt{r_1^2-y^2})}{\pi(r_2^2-r_1^2)} & \text{for } y \in (-r_1, r_1), \\[3mm] \dfrac{2\sqrt{r_2^2-y^2}}{\pi(r_2^2-r_1^2)} & \text{for } y \in [r_1, r_2], \\[3mm] 0 & \text{otherwise.} \end{cases}$$

We note that the random variables $X(\omega)$ and $Y(\omega)$ are dependent because the joint probability density function is not the product of the marginal distributions:

$$f_{XY}(x, y) \neq f_X(x) f_Y(y).$$

Now we calculate the mutual correlation of the random variables $X(\omega)$ and $Y(\omega)$. Changing variables from the rectangular to a polar system of coordinates, we calcu-

late the covariance function:

$$\mathrm{cov}(X, Y) = \iint_{r_1^2 \leq x^2 + y^2 \leq r_2^2} xy f_{XY}(x, y) dx dy$$

$$= \int_0^{2\pi} \int_{r_1}^{r_2} r^2 \cos\theta \sin\theta \frac{1}{2\pi(r_2 - r_1)} r dr d\theta = 0.$$

In this way we have obtained another example of a pair of random variables which are dependent but uncorrelated.

Problem 3.1 The probability density function of the random variable X has the following form:

$$f(x) = cx^2 e^{-kx}, \quad k > 0, \ 0 \leq x < \infty.$$

Calculate the normalizing coefficient c. Find the cumulative distribution function of the random variable X. Find the probability $P(X \in (0, 1/2k))$.

Problem 3.2 The Weibull distribution has the following cumulative distribution function:

$$F(x) = 1 - \exp(-cx^b), \quad x \geq 0.$$

Find: (a) the probability density function, (b) the quantile of p-th order, (c) the modal value of the Weibull distribution.

Problem 3.3 The random variable X has the beta distribution with the following probability density function:

$$f(x) = \begin{cases} Ax^{a-1}(1-x)^{b-1} & \text{for } x \in [0, 1], \\ 0 & \text{otherwise.} \end{cases}$$

$a > 0$, $b > 0$. Determine the parameter A and calculate the mean value and the variance of the random variable X.

Problem 3.4 For the random variable X with the probability density function (the Laplace distribution)

$$f(x) = \frac{1}{2} e^{-|x|},$$

find the mean deviation defined as the number E satisfying the following condition:

$$P(|X - \overline{x}| < E) = \frac{1}{2}.$$

Problem 3.5 The coordinates X and Y of a random point are independent and have uniform distributions over the intervals $[a, b]$ and $[c, d]$, respectively. Find the joint probability density function and the two-dimensional cumulative distribution function of the vectorial random variable (Y, Y).

Problem 3.6 The probability density function of the vectorial random variable (X, Y) has the form

$$f_{XY}(x, y) = C \exp\left[-\alpha x^2 + \xi xy - \beta y^2\right].$$

Find the marginal probability density functions $f_X(x)$ and $f_Y(y)$. When are the random variables X and Y independent?

Problem 3.7 The random variables X and Y are independent and identically distributed with the probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{x^2}{2\sigma^2}\right].$$

Find the joint probability density function of the random vector (R, Φ) satisfying the relations

$$X = R\cos\Phi, \qquad Y = R\sin\Phi.$$

Problem 3.8 Calculate the entropy of the Pareto distribution.

Problem 3.9 Propose a method of generation of sequences of random numbers with the Pareto distribution.

References

1. Abramowitz, M., Stegun, I.: Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. Dover, New York (1965)
2. Box, G.E.P., Jenkins, G.M.: Time Series Analysis. Forecasting and Control, 4th edn. Wiley, New York (2008)
3. Bratley, P., Fox, B.L., Schrage, L.E.: A Guide to Simulation, 2nd edn. Springer, New York (1987)
4. Cramer, H.: Mathematical Methods of Statistics. Princeton University Press, Princeton (1946)
5. Ditlevsen, O., Madsen, H.O.: Structural Reliability Methods. Wiley, Chichester (1996)
6. Doliński, K.: Comparison of a stochastic model of fatigue crack growth with experiments. Fatigue Fract. Eng. Mater. Struct. **16**(10), 1021–1034 (1993)
7. Fisz, M.: Probability Theory and Mathematical Statistics, 3rd edn. Krieger, Melbourne (1980)
8. Geochemical Atlas of Poland. Polish Geological Institute, Warsaw (1996) (in Polish)
9. Haight, F.A.: Index to the distribution of mathematical statistics. J. Res. Natl. Bur. Stand. B, Math. Math. Phys. **65B**(1) (1961)
10. Helmert, F.R.: Über die Wahrscheinlichkeit der Potenzsummen und bereinige damit in Zusammenhange stehende Fragen. Z. Math. Phys. **21** (1876)
11. Iwankiewicz, R.: Dynamical Mechanical Systems under Random Impulses. World Scientific, Singapore (1995)
12. Knuth, D.E.: Seminumerical Algorithms, 3rd edn. The Art of Computer Programming, vol. 2. Addison–Wesley, Reading (1998)
13. Kotulski, Z.: Wave propagation in randomly stratified medium. J. Sound Vib. **128**(2), 195–208 (1989)

14. Kotulski, Z.: Elastic waves in randomly stratified medium. Analytical results. Acta Mech. **83**, 61–75 (1990)
15. Kotulski, Z.: II Numerical calculations. Acta Mech. **92**, 143–163 (1992)
16. Kotulski, Z.: On the effective reflection properties of the randomly segmented elastic bar. Eur. J. Mech. A, Solids **13**(5), 677–696 (1994)
17. Kotulski, Z.: Wave pulses in two-dimensional randomly stratified elastic media. Arch. Mech. **47**(1), 125–139 (1995)
18. Kotulski, Z., Sobczyk, K.: Linear systems and normality. J. Stat. Phys. **24**(2), 359–373 (1981)
19. Lipson, Ch., Sheth, N.J.: Statistical Design and Analysis of Engineering Experiments. McGraw–Hill, Kogakusha (1973)
20. Murzewski, J.: Reliability of Engineering Structures. Arkady, Warsaw (1989)
21. Papoulis, A.: Probability, Random Variables, and Stochastic Processes with Errata Sheet, 4th edn. McGraw–Hill, New York (2002)
22. Press, W.H., Flannery, B.F., Teukolsky, S.A., Vetterling, W.T.: Numerical Recipes. The Art of Scientific Computing, 2nd edn. Cambridge University Press, Cambridge (1992)
23. Sachs, L.: Applied Statistics. A Handbook of Techniques, 2nd edn. Springer, New York (1984)
24. Sobczyk, K.: Stochastic Differential Equations with Applications to Physics and Engineering. Kluwer Academic, Dordrecht (1991)
25. Spaethe, G.: Die Sicherheit tragender Baukonstruktionen. VEB Verlag für Bauwesen, Berlin (1987)
26. Suchodolski, S.: Probabilistic estimation of strength of glass fibers and the epoxy matrix-fiber cohesion by means of the fragmentation test. In: Proceedings of XVII Polish Conference on Experimental Mechanics of Solid Materials, pp. 550–557, Jachranka, Poland, 1996 (in Polish)
27. Thewlis, J. (ed.): Encyclopedic Dictionary of Physics. Pergamon, Elmsford (1967)

Chapter 4
Functions of Independent Random Variables

4.1 Basic Relations

In engineering applications, for example, when accuracy of mechanism movements is dealt with, or in various non-direct measurements, we have to deal with important cases when a random variable U being under consideration is a function of certain other independent variables $X_1, X_2, \ldots, X_n$, that is:

$$U = g(X_1, X_2, \ldots, X_n). \tag{4.1}$$

Before treating such a problem in general, let us consider at first some important particular cases.

As the simplest case, let us consider the relation

$$U = aX, \tag{4.2}$$

where a is a certain constant. As an example, a simple feeding mechanism shown schematically in Fig. 4.1 may be taken. Arm AB turns around point A during repeated working cycles. Point D in the figure stands for the desired (nominal) position of the grip located at the point B of the arm at the end of each working cycle. This nominal position is determined by the zero value of the deviation angle ($\theta = 0$). In practical implementation, the deviation angle θ is of certain small non-zero magnitude ($\theta \neq 0$), and therefore, the final position of the grip is slightly shifted ,with respect to the desired position, by a small distance

$$U = a\theta.$$

The probabilistic nature of the error [5] of positioning angle θ results in the same nature of the positioning accuracy of the grip. If the standard deviation of the angle θ positioning error from its desired position $\theta = 0$ is σ_θ, then the standard deviation of the grip positioning error will be $\sigma_U = a\sigma_\theta$. Consequently, the variances will satisfy the relation $\sigma_U^2 = a^2\sigma_\theta^2$. In general, returning to relation (4.2) we can write

$$\sigma_U^2 = a^2\sigma_X^2. \tag{4.3}$$

Fig. 4.1 An example of a simple feeding mechanism
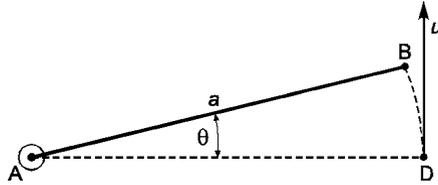
If a random variable U is determined as a sum of two independent (uncorrelated) random variables X and Y, then between the variances the relation

$$\sigma_U^2 = \sigma_{X \pm Y}^2 = \sigma_X^2 + \sigma_Y^2 \tag{4.4}$$

resulting from (3.105) will hold valid, see also Example 3.12. Such situations are illustrated in Fig. 4.2. Variable U (resulting length) presented in Fig. 4.2a results from the subtraction of two independent variables X and Y, while that shown in Fig. 4.2b is obtained as the sum of the two. Note that if such elements are mass-produced, the deviations of the two independent dimensions X and Y will have a distribution close to the normal distribution.

In general, when a dependent random variable U arises as a consequence of summation of several independent random variables X_j,

$$U = \sum_{j=1}^{n} X_j, \tag{4.5}$$

its variance σ_U^2 is defined as

$$\sigma_U^2 = \sum_{j=1}^{n} \sigma_{X_j}^2. \tag{4.6}$$

Note that the expression (4.6) holds valid also in the cases when some X_j in (4.5) are negative, that is, if a dimension is subtracted as it was shown in Fig. 4.2a:

$$U = X - Y.$$

A typical example of summing independent variables is shown in Fig. 4.3. The length U of the set of gears results from adding the lengths X_1, X_2, X_3, X_4 of the elements of the set.

From relations (4.3) and (4.6) may be deduced the formula for the variance of a dependent variable U being a linear combination of a number of independent variables X_j, that is, for

$$U = \sum_{j=1}^{n} a_j X_j. \tag{4.7}$$

Fig. 4.2 Examples, where
the resulting dimension U is:
(**a**) the difference of the
random variables X and Y,
(**b**) the sum of the random
variables X and Y
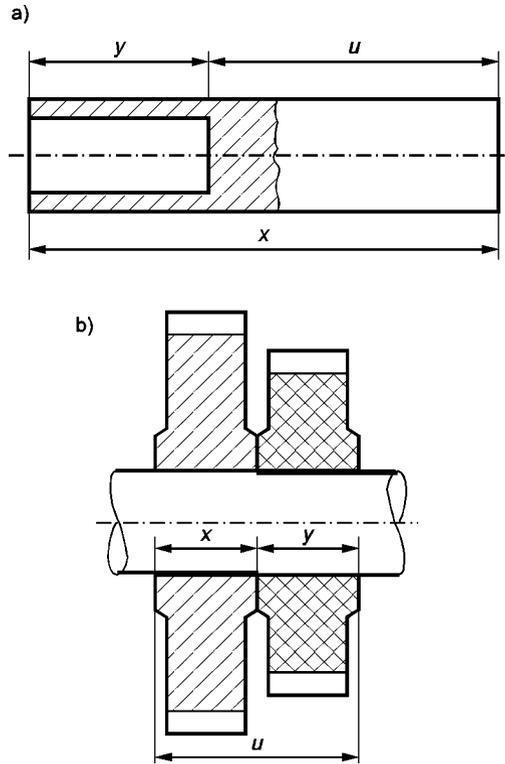
The variance σ_U^2 in this case can be calculated from the formula

$$\sigma_U^2 = \sum_{j=1}^n a_j^2 \sigma_{X_j}^2. \qquad (4.8)$$

Let us return now to the general case (4.1), when a dependent variable U is an arbitrary differentiable function of a number of independent random variables X_i. The average values of the variable U and of the independent variables X_i will be denoted by $\overline{u}$ and $\overline{x}_i$ respectively. If the distributions of these independent random variables are concentrated in the vicinity of their average values, as it happens in most cases of practical significance, then we may write

$$U - \overline{u} \approx \frac{\partial g}{\partial x_1}(X_1 - \overline{x}_1) + \frac{\partial g}{\partial x_2}(X_2 - \overline{x}_2) + \cdots + \frac{\partial g}{\partial x_n}(X_n - \overline{x}_n), \qquad (4.9a)$$

or in the compact form

$$U - \overline{u} \approx \sum_{j=1}^n \frac{\partial g}{\partial x_j}(X_j - \overline{x}_j), \qquad (4.9b)$$

Fig. 4.3 The example, where the resulting dimension U is the sum of several random variables X_1, X_2, X_3, X_4
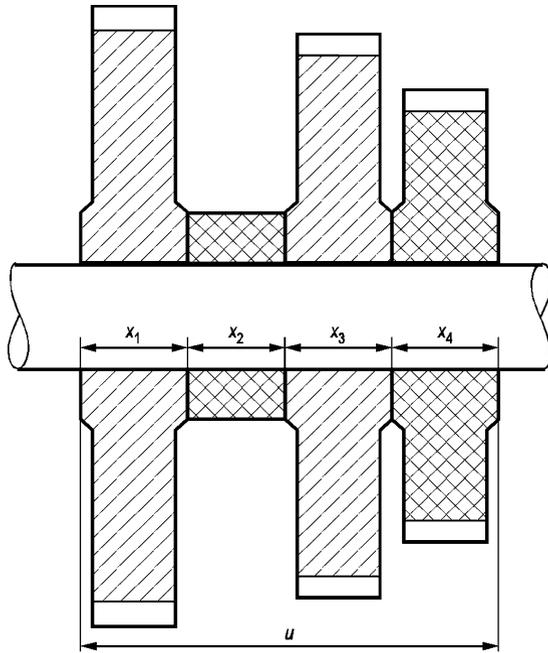
where in (4.9a), (4.9b) the values of partial derivatives $\frac{\partial g}{\partial x_j}$, $j = 1, 2, \ldots, n$ are taken at points $x_1 = \overline{x}_1$, $x_2 = \overline{x}_2$, $\ldots$, $x_n = \overline{x}_n$.

This relation may be treated as an approximation analogous to the expression for the total differential

$$dg = \sum_{j=1}^{n} \frac{\partial g}{\partial x_j} dx_j$$

at small finite deviations of variables from their average values.

According to (4.8) we can write an approximate expression for the variance of a function of a number of independent random variables

$$\sigma_U^2 \approx \left(\frac{\partial g}{\partial x_1}\right)^2 \sigma_{X_1}^2 + \left(\frac{\partial g}{\partial x_2}\right)^2 \sigma_{X_2}^2 + \cdots + \left(\frac{\partial g}{\partial x_n}\right)^2 \sigma_{X_n}^2, \qquad (4.10a)$$

or in compact form

$$\sigma_U^2 \approx \sum_{j=1}^{n} \left(\frac{\partial g}{\partial x_j}\right)^2 \sigma_{X_j}^2. \qquad (4.10b)$$

As in (4.9a), (4.9b) the derivatives are calculated at points equal to the mean values of the independent errors of measurements X_j, $j = 1, 2, \ldots, n$. The formula (4.10a), (4.10b) is called *the rule of propagation of errors*.

4.2 Simple Examples of Applications

Example 4.1 (Resistance of an ohmic element) The ohmic resistance of a certain element of an electric device is given by the formula

$$R = \rho \frac{4a}{\pi d^2}, \tag{4.11}$$

where

ρ is the specific ohmic resistance of the material,
a is the length of the wire,
d is the wire diameter.

Let us calculate the standard deviation σ_R of that element in mass-production if standard deviations of independent variables ρ, a, d are:

σ_ρ of the specific ohmic resistance,
σ_a of the wire length,
σ_d of the wire diameter.

Calculating the derivatives of resistance R treated as the function of independent random variables ρ, a, d and then substituting them into (4.10a), we arrive at the expression for the variance σ_R of the resistance R:

$$\sigma_R^2 = \left(\frac{4a}{\pi d^2} \right)^2 \sigma_\rho^2 + \left(\frac{4\rho}{\pi d^2} \right)^2 \sigma_a^2 + \left(\frac{8\rho a}{\pi d^3} \right)^2 \sigma_d^2. \tag{4.12}$$

The average values of ρ, a, and d should be taken in (4.12).

Example 4.2 (A system of resistors) From mass-produced electric elements A and B, ohmic resistances of which are random variables with average values R_A and R_B and standard deviations σ_A and σ_B, respectively, are manufactured electric devices for various purposes. In the first type of devices the elements are connected in series, while in the other type they are connected in parallel. Let us calculate the resultant standard deviations of ohmic resistance in the two cases.

In products where the elements are connected in series, the resultant ohmic resistance is

$$R = R_A + R_B$$

and the variance of that resistance is

$$\sigma_R^2 = \sigma_A^2 + \sigma_B^2.$$

In the other case when elements A and B are connected in parallel, the resultant resistance is

$$\frac{1}{R} = \frac{1}{R_A} + \frac{1}{R_B},$$

which gives:

$$R = \frac{R_A R_B}{R_A + R_B}.$$

Calculating partial derivatives of R with respect to R_A and R_B, we obtain from (4.10a) the expression for the variance of the resultant ohmic resistance R:

$$\sigma_R^2 = \left(\frac{R_B}{R_A + R_B} \right)^4 \sigma_A^2 + \left(\frac{R_A}{R_A + R_B} \right)^4 \sigma_B^2.$$

The square root of this magnitude stands for the standard deviation σ_R of the resistance R.

4.3 Examples of Applications in Non-direct Measurements

In metrology we often deal with numerous cases when it is difficult, or even impossible, to measure directly a specific parameter U of an object under investigation. In such cases methods of non-direct measuring can be helpful if other parameters $X_1, X_2, \ldots, X_n$, connected with the parameter U in question by a mathematical relation generally written as (4.1), can be directly measured. Random errors, unavoidable when parameters X_i are measured, are transferred on the non-directly measured parameter U according to the error propagation rule (4.10a). It is assumed that random measuring errors of parameters X_i have the normal distribution.

Example 4.3 (Non-direct measurements) In Fig. 4.4 is schematically presented an arrangement for non-direct measuring of the diameter D of a hole in a configuration
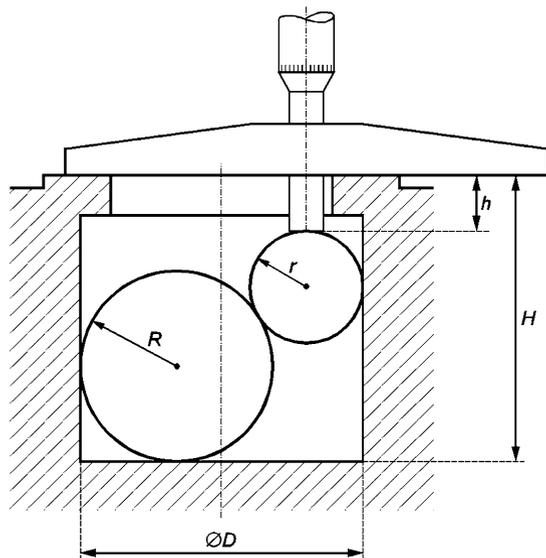

Fig. 4.4 Non-direct measurements of the diameter with the measuring steel balls

Fig. 4.5 Non-direct
measurements of the angle
with the measuring steel balls
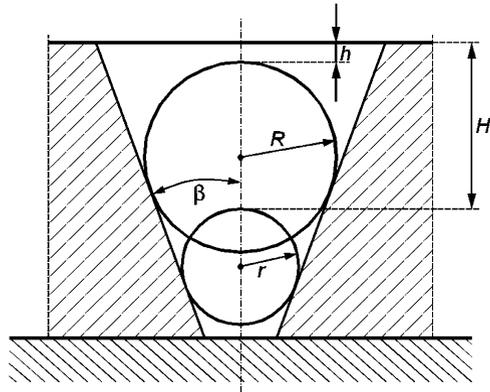
of difficult direct access. Two measuring steel spheres of the radii R and r, respectively, are used. Non-direct procedure consists in measuring first the depth H of the hole and then, after inserting both spheres, in measuring the distance h of the highest point of the upper sphere from the reference plane. The diameter D is related with the four distances R, r, H, h by the formula

$$D = (R + r)\left(1 + \sqrt{1 - \left(\frac{H - h - R - r}{R + r}\right)^2}\right).$$

Referring to the formula (4.10a), we can write the expression for the variance of the non-direct measuring accuracy of diameter D,

$$\sigma_D^2 = \left(\frac{\partial D}{\partial R}\right)^2 \sigma_R^2 + \left(\frac{\partial D}{\partial r}\right)^2 \sigma_r^2 + \left(\frac{\partial D}{\partial H}\right)^2 \sigma_H^2 + \left(\frac{\partial D}{\partial h}\right)^2 \sigma_h^2,$$

if the estimators of variances of the measuring accuracy of dimensions H and h, and σ_R^2, σ_r^2, which are the variances of the diameters R and r of the two spheres, are known.

Example 4.4 (Non-direct measurements) In Fig. 4.5 is shown a scheme of a procedure allowing to measure the angle β of the conical hole with the use of two measuring spheres. Knowing the radii R and r of the spheres and measuring consecutively the depths H and h, we can calculate the angle β using the formula

$$\beta = \arcsin \frac{R - r}{H + r - (h + R)}.$$

The expression for the variance takes the form

$$\sigma_\beta^2 = \left(\frac{\partial \beta}{\partial R}\right)^2 \sigma_R^2 + \left(\frac{\partial \beta}{\partial r}\right)^2 \sigma_r^2 + \left(\frac{\partial \beta}{\partial H}\right)^2 \sigma_H^2 + \left(\frac{\partial \beta}{\partial h}\right)^2 \sigma_h^2.$$

As in the previous example, the variance of the measuring accuracy of both dimensions H and h and of the radii R and r are assumed to be known.

4.4 Remarks on Applications in the Calculus of Tolerance Limits

In engineering applications, especially in machine design, it is generally assumed
that dimensions of mass-produced machine parts may vary only within strictly lim-
ited allowable small deviations $+\Delta q_i$ and $-\Delta q_i$, referred to as the tolerance limits,
see e.g. [3, 4] . If a resultant dimension is a function of other independent dimen-
sions limited by their tolerance limits, then the question may arise: how large will be
the resultant tolerance limits of that resultant dimension? This problem may be an-
alyzed in terms of the error calculus by assuming that deviations of the dimensions
vary within the limits bounded by a respective pair of tolerance limits according to
the rectangular distribution (comp. Fig. 3.3a).

Let the resultant dimension U be a certain function of a set of other dimensions
$q_j, j = 1, 2, \ldots, n$ (comp. (4.1)),

$$U = g(q_1, q_2, \ldots, q_n). \tag{4.13}$$

Each dimension q_j is bounded by its upper Δq_j^+ and lower Δq_j^- tolerance limits.

The resultant tolerance limit ΔU of the resultant dimension U can be calculated
from the general formula

$$\Delta U = \sum_{j=1}^{n} \frac{\partial U}{\partial q_j} \Delta q_j. \tag{4.14}$$

We get the upper tolerance limit ΔU^+ and the lower one ΔU^- by substituting
appropriately the tolerance limits Δq_i^+ and Δq_i^- in order, at first, to obtain the largest
possible value of the sum in (4.14), that is, the upper tolerance limit ΔU^+, and then,
its smallest possible value ΔU^- corresponding to the lower tolerance limit.

Note that the tolerance limits may be symmetrical, in which case we have
$|\Delta q_j^+| = |\Delta q_j^-|$; or non-symmetrical, when $|\Delta q_j^+| \neq |\Delta q_j^-|$, with respect to the
desired (nominal) value of a specific dimension $q_j, j = 1, 2, \ldots, n$.

Example 4.5 (Tolerances) Let the centers of two holes A and B in a rectangular ob-
ject be determined by four dimensions: x_A, y_A, x_B, y_B with non-symmetrical toler-
ance limits as shown in Fig. 4.6. Find the resultant tolerance limits ΔR^+ and ΔR^-
of the resultant distance R between the axes of the holes. The desired (nominal)
value of that distance is

$$R = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2},$$

where on the right-hand side the nominal values of the coordinates should be sub-
stituted. In this particular case we obtain $R = 58.31$ mm.

The partial derivatives are

$$\frac{\partial R}{\partial x_A} = -\frac{1}{R}(x_B - x_A) = -0.857, \qquad \frac{\partial R}{\partial x_B} = \frac{1}{R}(x_B - x_A) = 0.857,$$

Fig. 4.6 System of dimensions for calculation of the tolerance limits of the distance between the centers of two holes
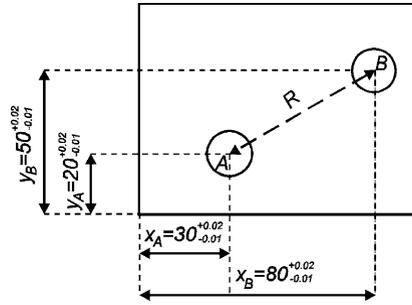
Fig. 4.7 The crankshaft mechanism
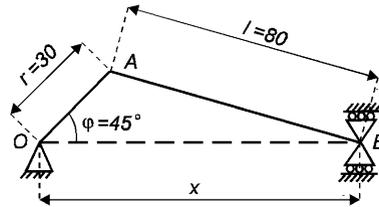
$$\frac{\partial R}{\partial y_A} = -\frac{1}{R}(y_B - y_A) = -0.514, \qquad \frac{\partial R}{\partial y_B} = \frac{1}{R}(y_B - y_A) = 0.514.$$

These derivatives may be treated as the coefficients of sensitivity of the resultant dimension to deviations of the basic dimensions.

Now, making use of general formula (4.14) we find the resulting tolerance limits of the distance R,

$$\Delta R^+ = 0.041 \text{ mm}, \qquad \Delta R^- = -0.041 \text{ mm}.$$

Calculating the upper tolerance limit ΔR^+ we should substitute the upper tolerance limit of a basic dimension if the numerical value of the corresponding derivative has a positive sign. If the numerical value of the derivative is negative, then the lower tolerance limit of the corresponding basic dimension should be taken. When calculating the lower tolerance limit ΔR^- we should substitute the tolerance limits of the basic dimensions in the opposite manner to that for ΔR^+ in order to get the smallest possible value of ΔR^-.

Example 4.6 (Sensitivity of positioning accuracy) As another simple example we shall analyze the sensitivity of the positioning accuracy of the end effector B of a crankshaft mechanism (shown in Fig. 4.7) to the variations of dimensions (e.g., caused by temperature changes) of the arms *OA* and *AB*, or to the positioning errors of the angle φ. The position of point B is determined by the length $OB = x$ given by

$$x = r\cos\varphi + \sqrt{l^2 - r^2\sin^2\varphi}.$$

The sensitivity coefficients are:

$$\frac{\partial x}{\partial l} = \frac{1}{\sqrt{l^2 - r^2 \sin^2 \varphi}} = 1.037 \text{ mm/mm},$$

$$\frac{\partial x}{\partial r} = \cos \varphi - \frac{1}{\sqrt{l^2 - r^2 \sin^2 \varphi}} = 0.5126 \text{ mm/mm},$$

$$\frac{\partial x}{\partial \varphi} = -r \sin \varphi - \frac{r^2 \sin \varphi \cos \varphi}{\sqrt{l^2 - r^2 \sin^2 \varphi}} = -27.097 \text{ mm/rad}.$$

The numerical values of the coefficients have been calculated for $\varphi = 45°$.

4.5 Statical Analogy in the Analysis of Complex Dimension Nets

As an introductory example of application of the statical analogy method let us consider a system of bars connected by bolt joints A, B, C, D, E, F (Fig. 4.8). If all the bars are made with such an ideal accuracy that they are of nominal (desired) length, then the resultant distance R between points A and B will have a certain nominal value R_0. However, if the real lengths of the bars are slightly different within the tolerance limits, then also the resultant distance R will differ from its nominal value. Using the analytical method described in the previous section we should at first determine the function:

$$R = R(l_1, l_2, \ldots, l_9), \tag{4.15}$$

where $l_1, l_2, \ldots, l_9$ stand for the lengths of particular bars.

The deviation of distance R from its nominal value R_0 is, comp. (4.14),

$$\Delta R = \frac{\partial R}{\partial l_1} \Delta l_1 + \frac{\partial R}{\partial l_2} \Delta l_2 + \cdots + \frac{\partial R}{\partial l_9} \Delta l_9. \tag{4.16}$$

If the tolerance limits of all lengths l_j, $j = 1, 2, \ldots, 9$ are given, then the resulting tolerance limits of R can be calculated in the manner described in Sect. 4.4.
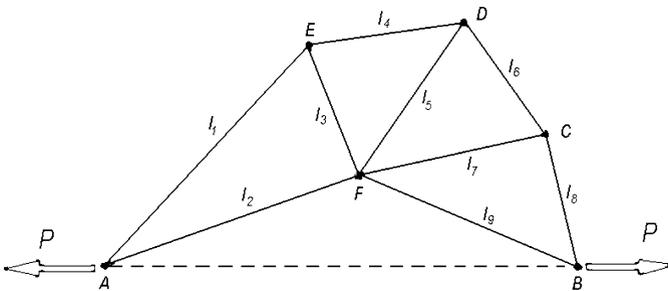
Fig. 4.8 The system of bars connected by bolt joints
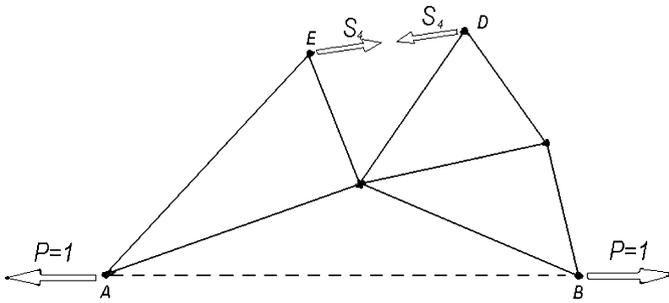
Fig. 4.9 The equivalent system of bars and forces for the system of bars connected by bolt joints, presented in Fig. 4.8

Note, however, that even in the relatively simple case as that shown in Fig. 4.8 it is not easy to write the specific form of relation (4.15). In such cases the statical analogy method based on the Maxwell-Mohr method used in structural mechanics for determining deformation of structures may be useful, see e.g. [1].

Let us treat the system of bars shown in Fig. 4.8 as a statically determinate truss. To find the change of distance R between points A and B we should at first, using the Maxwell-Mohr procedure, apply at these points two unit forces acting along AB in opposite directions. Suppose now that one of the bars, say bar ED, has been removed and instead of it, two forces s_4, as shown in Fig. 4.9, have been applied. To keep the system in equilibrium, the relation

$$P \Delta R = s_4 \Delta l_4 \qquad (4.17)$$

must be satisfied, where ΔR and Δl_4 are virtual small changes of the lengths of sectors AB, and ED respectively. It is assumed that tensile internal forces s_i in the bars resulting from the unit loading of the system have positive sign, while the sign of compressive forces is negative. Elongations of bars are treated as having positive sign, while the sign of shortening is negative. For example, positive ΔR causes shortening of bar number 4 and a compressive force in it. Thus, Δl_4 and s_4 should be taken with negative signs.

Generally, we can write relation (4.17) for each bar of the system in the form

$$\Delta R = s_j \Delta l_j. \qquad (4.18)$$

Force P does not appear in this equation because we have assumed earlier that it is equal to unity.

Thus, if the length l_j of each bar changes by a small value Δl_j, then the change of the distance R between points A and B will be

$$\Delta R = \sum_{j=1}^{n} s_j \Delta l_j. \qquad (4.19)$$

Note that no restrictions concerning the nature of length changes Δl_j have been made. In structural mechanics, when a truss is loaded by external forces, elastic deformations of particular bars stand for Δl_j in (4.19). However, formula (4.19) holds valid also in such cases when the change Δl_j in length l_j results from temperature changes or simply when Δl_j is a manufacturing error, that is, the bar has been made slightly longer or shorter than the bar of nominal length.

The structure of relation (4.19) is the same as that of the previous relation (4.16), except that instead of derivatives $\frac{\partial R}{\partial l_j}$, internal forces s_j appear in it. The method of statical analogy in the analysis of errors in complex dimension nets is based upon this similarity [2]. For every dimension network we construct a representative system of bars and hinges, and then find the numerical values of derivatives appearing in (4.14) as internal forces in this system loaded by appropriate unit forces. Note that for dimension networks without unnecessary "closing" dimensions, the representative bar systems are statically determinate.

In Fig. 4.10 is explained the principle of constructing the representative bar systems. A simple example of dimension nets determining the location of four points was chosen. Generalized unit loads for the analysis of a resultant linear R or angular Ω dimension are shown in the figure. Every linear dimension l_j is represented by a bar, while every angular dimension θ_j is represented by a special node (hinge) able to transmit a moment. Such nodes are shown schematically as having a "spring" between the bars. In column A, various combinations of basic dimensions (linear or angular) determining mutual positions of four points are shown. The resultant linear dimension under analysis is denoted by R, and the resultant angular dimension is shown as angle Ω. In column B, the systems of bars for the analysis of deviations of resultant distance R are presented along with the respective unit load. In column C, analogous schemes for the analysis of deviations of angular dimension Ω are given. Note that in such cases, the unit load reduces to two unit moments applied to the bars forming an angle. These moments are implemented as pairs of forces applied at the end points of respective bars.

Note that for elastic hinges corresponding to basic angular dimensions, we should substitute in formula (4.19) the moment acting in the hinge as a generalized force s_j, and the angular deviation as a generalized deformation Δl_j.

As an example of application of the statical analogy method let us analyze a system of four holes shown in Fig. 4.11. The configuration of the holes is determined by five dimensions: four linear a, b, c, d, and one angular α, each with tolerance limits given in the figure. Let us find the resultant tolerance limits of the resultant distance R between points B and D.

The corresponding statical system of bars is shown in Fig. 4.12. Note that the elastic hinge with a "spring" at point B represents the angular basic dimension α. The system is loaded by two unit forces acting in opposite directions along BD. The internal forces in the bars are:

$$\text{bar } AB \quad s_a \equiv \frac{\partial R}{\partial a} = -0.15,$$

$$\text{bar } BC \quad s_b \equiv \frac{\partial R}{\partial b} = 0.63,$$
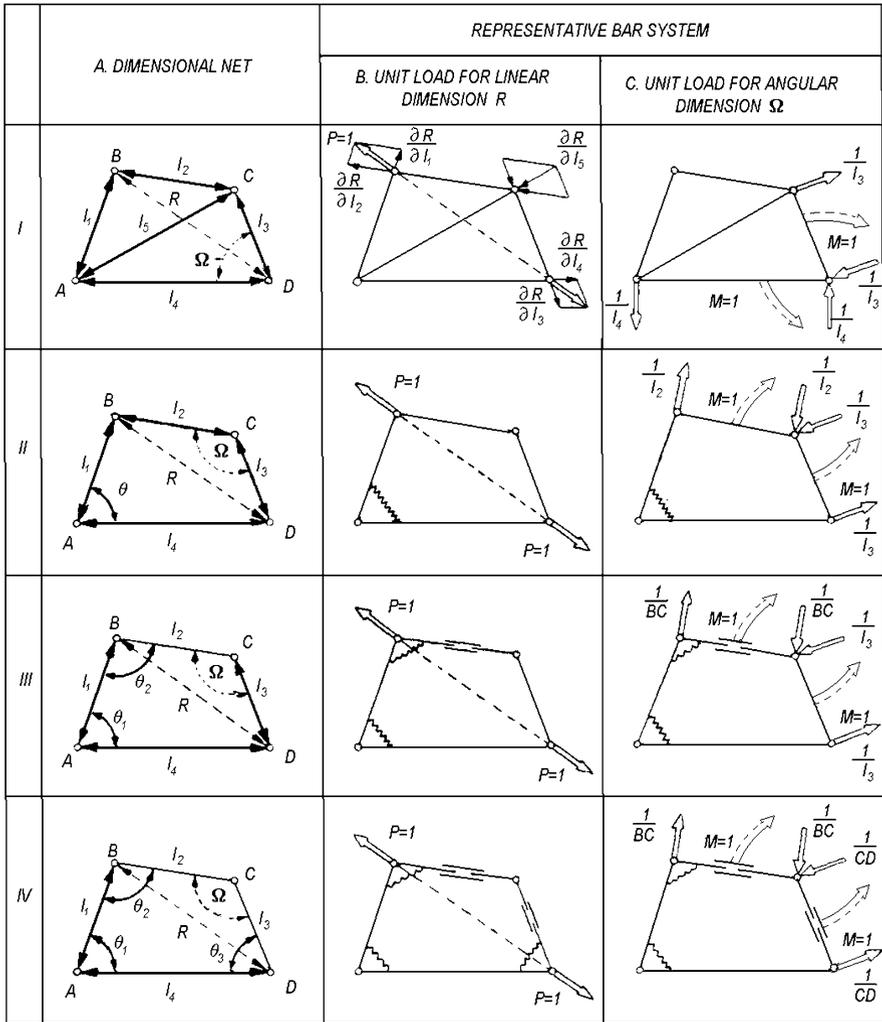
Fig. 4.10 The principle of constructing representative bar systems

$$\text{bar } CD \quad s_c \equiv \frac{\partial R}{\partial c} = 0.95,$$

$$\text{bar } DA \quad s_d \equiv \frac{\partial R}{\partial d} = 0.45.$$

The moment in the hinge B is

$$m_B = a s_1 = b s_2 \equiv \frac{\partial R}{\partial a} = -45 \text{ mm.}$$

Fig. 4.11 An example of
four holes
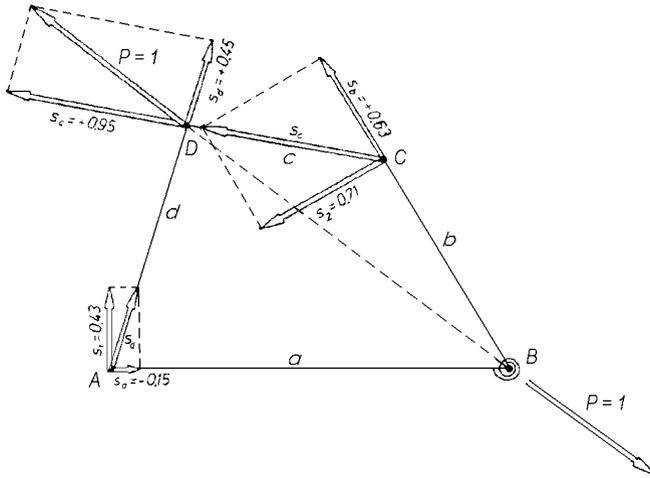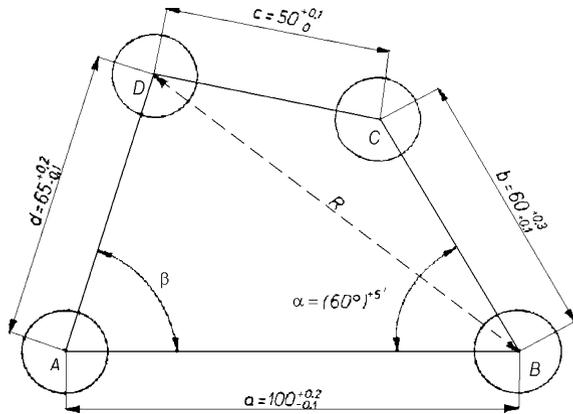
Fig. 4.12 The statical system of bars corresponding to the system of four holes presented in
Fig. 4.11

Thus, the sensitivity coefficients have been determined without determining the analytical relation

$$R = R\,(a, b, c, d, \alpha)\,.$$

They have been found on the basis of graphical representation of this relation given
in Fig. 4.12.

Now, having known numerical values of the sensitivity coefficients, the resulting
tolerance limits of the distance R were calculated:

$$\Delta R^{+} = +0.40 \text{ mm,} \qquad \Delta R^{-} = -0.07 \text{ mm.}$$

Fig. 4.13 Graphical
illustration of Problem 4.3
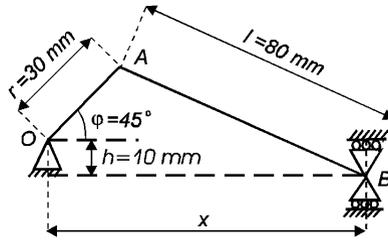
Problem 4.1 Find the expression for the variance of non-direct measuring accuracy of distance R between holes A and B (Fig. 4.6) if the distances x_A, y_A, x_B, y_B have been measured with the accuracy defined by the variance σ^2.

Problem 4.2 Find the expression for the variance of the desired location of point B of a simple mass-produced mechanism shown in Fig. 4.7 in the position defined by the angle $\varphi = 45°$. Dimensions are $r = 30$ mm, $l = 80$ mm. Their variances are σ_r^2 and σ_l^2, respectively. The variance of adjusting the angle is σ_φ^2.

Problem 4.3 Find the sensitivity coefficients of positioning accuracy of the end effector B of the mechanism shown in the figure for the particular value of the angle $\varphi = 45°$.

Problem 4.4 Using the statical analogy method find the representative unit load for determining the coefficients of sensitivity of angle EAF in the network shown in Fig. 4.8, to the deviations of all basic dimensions.

Problem 4.5 Using the statical analogy method find for the mechanism shown in Fig. 4.7 coefficients of sensitivity of the distance X to the deviations of dimensions r and l and the adjusting error of angle $\varphi = 45°$. Compare the results with those calculated analytically in Sect. 4.4, Example 4.6.

References

1. Fertis, D.G.: Advanced Mechanics of Structures. CRC Press, Boca Raton (1996)
2. Szczepiński, W.: Statical analogy in the study of deviations in complex dimension nets. Bull. Acad. Pol. Sci., Sér. Sci. Tech. **7**(6), 365–370 (1966)
3. Szczepiński, W.: Theory of polyhedron of positioning accuracy of manipulators. Mech. Mach. Theory **26**(7), 697–709 (1991)
4. Szczepiński, W.: Introduction to the Calculus of Tolerance Limits in Mechanical Engineering. Polish Scientific Editors, Warsaw (1993) (in Polish)
5. Taylor, J.R.: An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, 2nd edn. University Science Books (1997)

Chapter 5
Two-dimensional Distributions

5.1 Introductory remarks

As a typical example of two-dimensional distribution, the errors in accuracy of artillery fire shown in Fig. 1.4 may be taken. The total error is composed of two independent deviations: error in the shooting distance X and lateral deviation Y from the shooting direction. Such a relatively small number of points classifies this distribution among the discrete (quasi-step) distributions.

Using this simple example we shall show how to determine the average point M of all the points hit by shells and how to check whether the two deviations in x and y directions are really independent.

Coordinates of central (average) point M are, comp. (1.1),

$$\overline{x}_M = \frac{1}{n} \sum_{j=1}^{n} x_j, \qquad \overline{y}_M = \frac{1}{n} \sum_{j=1}^{n} y_j, \tag{5.1}$$

where n is the number of all the hit points. Point M determined in such a manner is referred to as the estimator of the average values of random variables X and Y. An obvious analogy can be seen with the procedure of calculating the position of gravity center of the set of points, each with unit mass.

In the particular problem at hand we get in the appropriately shifted reference system

$$\overline{x}_M = 28.7 \text{ m}, \qquad \overline{y}_M = 15.3 \text{ m}.$$

Estimators of variances of the random variables X and Y are given by the formulas, comp. (1.4a),

$$\sigma_X^2 = \frac{1}{n} \sum_{j=1}^{n} \left(x_j - \overline{x}_M\right)^2, \qquad \sigma_Y^2 = \frac{1}{n} \sum_{j=1}^{n} \left(y_j - \overline{y}_M\right)^2. \tag{5.2}$$

Fig. 5.1 Hit points of
artillery fire
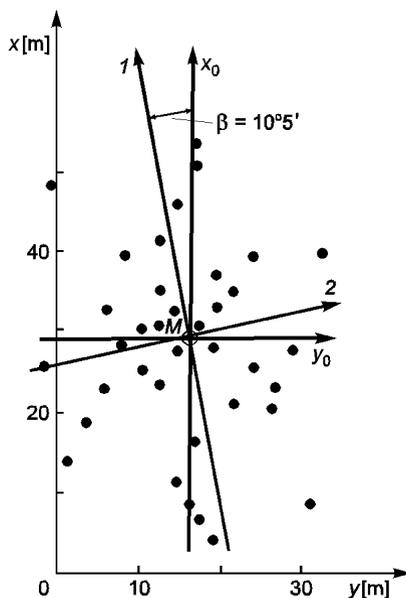
Note that the variances calculated in such a manner may be treated as the inertia
moments of that set of hit points with respect to central axes x_0, y_0 passing through
the central point of the set (Fig. 5.1), divided by the number of all points.

In the particular problem in question we get

$$\sigma_X^2 = 137.6 \text{ m}^2, \qquad \sigma_Y^2 = 62.13 \text{ m}^2.$$

Thus, the estimators of standard deviations in this case are

$$\sigma_X = 11.71 \text{ m}, \qquad \sigma_Y = 7.88 \text{ m}.$$

To check whether random variables (with the normal distribution) X and Y are
independent, we shall calculate their covariance.[1] The formula for the estimator of
the covariance takes the form:

$$\text{cov}(X, Y) = \frac{1}{n} \sum_{j=1}^{n} (x_j - \overline{x}_M)(y_j - \overline{y}_M). \tag{5.3}$$

The covariance may be interpreted as the deviation moment (mixed moment) of the
set of points with respect to central axes x_0, y_0, divided by the number of points.

---

[1]Random variables (as functions of an elementary event) are denoted here by capital letters, while
their actual values (values for a specific event), such as results of measurements, are denoted by
small letters.

For the particular set of points shown in Fig. 5.1 the estimator of the covariance was found to be

$$\text{cov}(X, Y) = -13.81 \text{ m}^2.$$

The three magnitudes: variances (5.2) and covariance (5.3) stand for the components of a certain symmetrical tensor, called the covariance tensor

$$T_{\sigma^2} = \begin{bmatrix} \sigma_X^2 & \text{cov}(X, Y) \\ \text{cov}(X, Y) & \sigma_Y^2 \end{bmatrix}. \tag{5.4}$$

Principal directions of this tensor and its principal components may be determined, for example, by using the Mohr circles representation, see Remark below.

Remark 5.1 (Mohr circles representation of two-dimensional tensors) In two-dimensional problems the tensors often appear determined, for example, by their principal components along with the principal directions. Such a tensor can be written as

$$T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix}, \tag{5.5}$$

where T_1 and T_2 are the principal components. As examples of such tensors in applied mechanics may be mentioned: stress tensor in plane stress problems, strain tensor in plane strain problem, tensor of inertia moments of the section of beams subjected to bending. The covariance tensor in two-dimensional distributions of random variables belongs also to this class of tensors.

If the reference system is turned through angle β with respect to the principal axes 1 and 2, then the tensor will take the form

$$T = \begin{bmatrix} T_{xx} & T_{xy} \\ T_{yx} & T_{yy} \end{bmatrix}, \tag{5.6}$$

where the transformed components are

$$T_{xx} = T_1 \cos^2 \beta + T_2 \sin^2 \beta = \frac{1}{2}(T_1 + T_2) + \frac{1}{2}(T_1 - T_2)\cos 2\beta,$$

$$T_{yy} = T_1 \sin^2 \beta + T_2 \cos^2 \beta = \frac{1}{2}(T_1 + T_2) - \frac{1}{2}(T_1 - T_2)\cos 2\beta, \tag{5.7}$$

$$T_{xy} = T_{yx} = \frac{1}{2}(T_1 - T_2)\sin 2\beta.$$

Let us assume now a certain coordinate system with the horizontal axis representing the components of the tensor having repeated indexes, and with the vertical axis along which the components with mixed indexes are laid off. It is seen that in such a reference system, relations (5.7) represent a circle with a radius equal to $\frac{1}{2}(T_1 - T_2)$ and with the center lying on the horizontal axis at a distance equal to

$\frac{1}{2}(T_1 + T_2)$ from the origin, Fig. 5.2. Components T_{xx}, T_{yy}, and T_{xy} of the tensor are
represented for a given angle by a pair of points A and B on that circle. Position of
the two points is determined by the angle 2β as shown in the figure. Such a graphi-
cal method was proposed by Otto Mohr in 1882 for the analysis of stress tensors. It
finds now applications in other problems.

Let us note that the inverse formulas result directly from the Mohr circle, when
the components T_{xx}, T_{yy}, and T_{xy} of the tensor are known, while its principal com-
ponents T_1, T_2 and their directions are to be found. These formulas have the form:

$$T_1 = \frac{1}{2}\left(T_{xx} + T_{yy}\right) + \sqrt{\frac{1}{4}(T_{xx} - T_{yy})^2 + T_{xy}^2},$$

$$T_2 = \frac{1}{2}\left(T_{xx} + T_{yy}\right) - \sqrt{\frac{1}{4}(T_{xx} - T_{yy})^2 + T_{xy}^2}, \tag{5.8}$$

$$\tan 2\beta = \frac{2T_{xy}}{T_{xx} - T_{yy}}.$$

In general three-dimensional problems, when in a chosen coordinate system all
six components of a tensor have non-zero values, the interpretation of the Mohr
circles representation is much more complex. In such three-dimensional cases rather
analytical methods of tensor components transformation are used.

However, in numerous special three-dimensional problems, such as those ana-
lyzed in Chap. 8, when one of the principal components of the covariance tensor
is known, the Mohr circles representation is very useful. The analysis can be done
in the same manner as for two-dimensional problems. The covariance tensor has in
such cases a simplified form

$$T = \begin{bmatrix} T_{xx} & T_{xy} & 0 \\ T_{yx} & T_{yy} & 0 \\ 0 & 0 & T_{zz} \end{bmatrix}. \tag{5.9}$$

Its four-element part may be analyzed as a two-dimensional tensor of the type (5.6).
Following the procedure described above, one can find the values of the two re-

Fig. 5.3 Determination of
principal directions of the
covariance tensor by means
of the Mohr circles
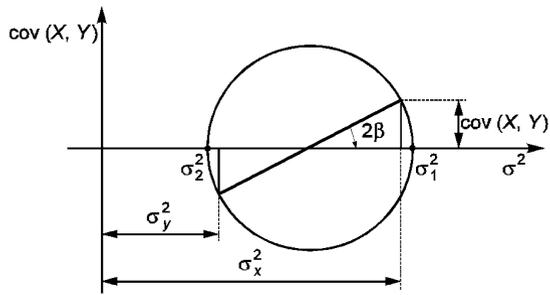
maining principal components and the orientation of the two remaining principal
directions.

Returning to our problem from Fig. 5.1, we find (see Fig. 5.3) the formula for the
angle β of deviation of principal axes of the covariance tensor from the directions
of coordinate axes x_0, y_0:

$$\tan 2\beta = \frac{2\,\text{cov}(X, Y)}{\sigma_X^2 - \sigma_Y^2}. \tag{5.10}$$

Principal components of the covariance tensor are

$$\sigma_1^2 = \frac{1}{2}\left(\sigma_X^2 + \sigma_Y^2\right) + \sqrt{\frac{1}{4}\left(\sigma_X^2 + \sigma_Y^2\right)^2 + [\text{cov}\,(X, Y)]^2},$$

$$\sigma_2^2 = \frac{1}{2}\left(\sigma_X^2 + \sigma_Y^2\right) - \sqrt{\frac{1}{4}\left(\sigma_X^2 + \sigma_Y^2\right)^2 + [\text{cov}\,(X, Y)]^2}. \tag{5.11}$$

For the particular set of points shown in Fig. 5.1 we get the following numerical
value of the estimator of deviation angle:

$$\tan 2\beta = -\frac{2 * 13.81}{137.16 - 62.13} = -0.3676, \quad \beta = 10°5',$$

and values of estimators of principal components of the covariance tensor

$$\sigma_1^2 = 140.05 \text{ m}^2, \qquad \sigma_2^2 = 59.68 \text{ m}^2.$$

In Fig. 5.1 is shown the calculated deviation of principal axes 1 and 2 of the
covariance tensor from central axes x_0, y_0. Such a slight deviation may be in this
case connected with a faulty choice of the reference coordinate system x, y, which
may not coincide with the shooting direction. Such a possibility should always be
considered, because it may happen that the independence of random variables in
question is illusory.

Results presented in Fig. 5.1 illustrate how such and similar practical problems
may be analyzed. Note, however, that in this specific case we should take into ac-
count that the number of registered points (only 38) is too small to ensure that the
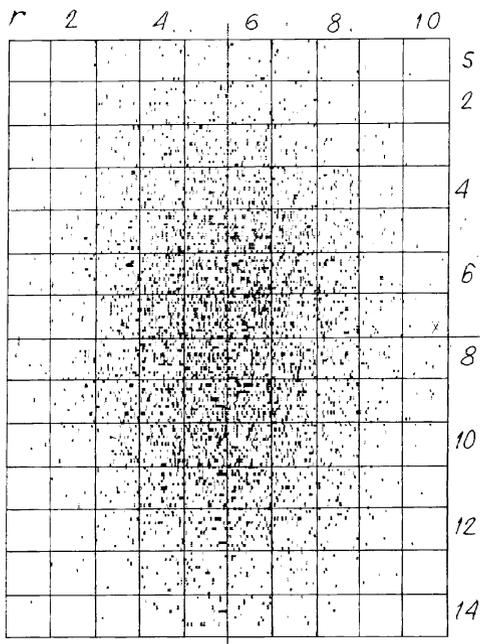
estimator of angle β has been calculated with sufficient accuracy. In more accurate analysis, statistical tests should be used (see Chap. 10) in order to verify the hypothesis concerning the estimated value of the deviation angle.

As another example in Fig. 5.4 is shown a set of 3100 points generated by a personal computer. A simple program generating random numbers with the two-dimensional normal distribution was used (comp. Chap. 6, Sect. 6.3—[14, 15]). The entire field has been divided into 140 small squares. The numbers of points in each square have been counted and the two-dimensional histogram shown in Fig. 5.5 was built. The height of each column corresponds to the frequency defined as the quotient of number n_{rs} of points in a specific square $r - s$ divided by the total number of all generated points $n = 3100$. If the area of each square in Fig. 5.4 is assumed to be equal to unity, then the sum of the volumes of all columns will be equal to unity.

The irregularities in the histogram and its deviation from the regular two-dimensional distribution (comp. Sect. 5.5) are caused by the approximate nature of the numerical simulation performed.

5.2 Linear Regression of Experimental Observations

In experimental investigations the results observed are usually presented in the form of a set of points in a certain coordinate system, say, a rectangular x, y system. Each coordinate represents one of the two magnitudes under investigation, which often
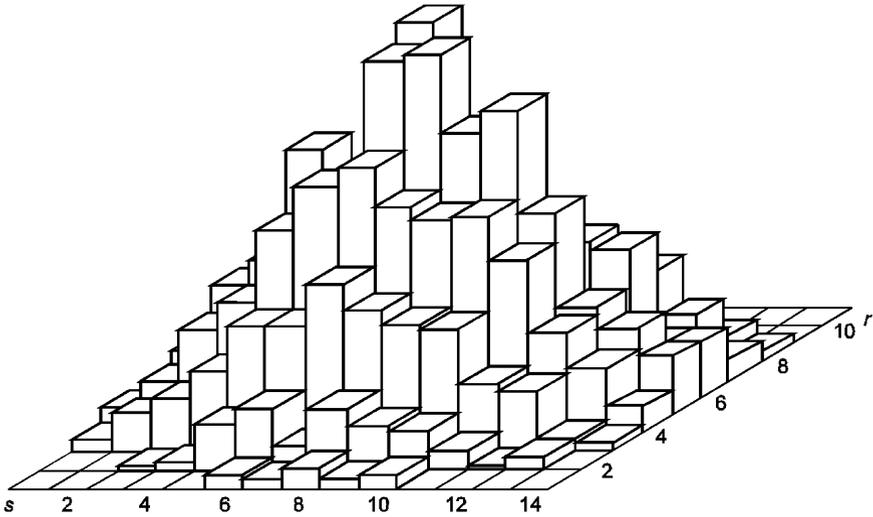
Fig. 5.5 Histogram built from the data presented in Fig. 5.4

are of different physical dimensions. If we want such a set of experimental points to be treated as a two-dimensional random variable, we should lay off on the coordinate axes the non-dimensional magnitudes resulting from dividing the measured quantities by appropriately chosen constant values of the magnitudes measured. In Fig. 5.6, as an example, is presented a set of experimental points showing how the austenite share in an alloy steel's internal structure changes with the amount of introduced plastic deformation (comp. [1]). Such sets of experimental points may be treated as a realization of the distribution of a certain two-dimensional random variable. In statistics much attention is paid to such a choice of a certain mathematical curve defined by the generally written equation

$$y = g_1(x), \tag{5.12a}$$

or that it might possibly accurately represent the results of experiment or observation. Such an equation in mathematical statistics is referred to as the regression of variable Y with respect to variable X (comp. [10]). Changing the names of variables, one can also write generally the following equation of regression of variable X with respect to variable Y:

$$x = g_2(y). \tag{5.12b}$$

Functions (5.12a), (5.12b) represent certain curves in the x, y plane. These curves are called regression lines. Usually it is not easy to find an appropriate form of functions $g_1(x)$ or $g_2(y)$ which assure sufficiently good representation of experimental points. Selection of the mathematical form of these functions is often arbitrary and

Fig. 5.6 Content of austenite in the alloy steel versus the range of plastic deformation [1]
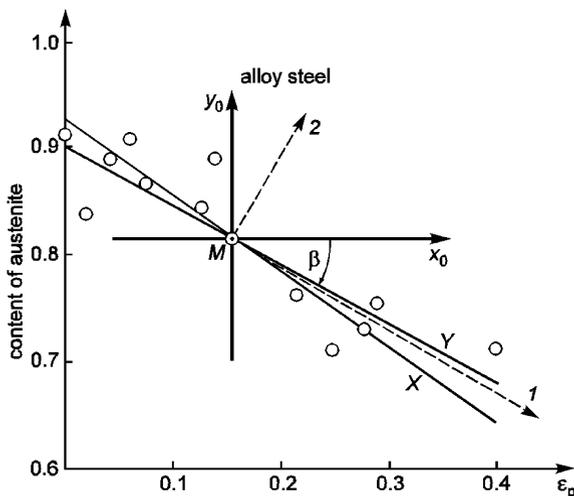
depends on the researcher's experience. This problem is still being discussed. More information may be found in specialized books (see, e.g., [16]).

5.2.1 Nonparametric Regression

Among methods applied for identifying laws resulting from experimental data and expressed by the curves of the form (5.12a) or (5.12b) we can indicate those that are more theoretical and those of practical importance, useful for calculations. The nonparametric regression, which we present now, belongs to more theoretical methods.

Assume that the random variables X and Y are measured simultaneously and the measurement procedure is repeated n times. As a result we obtain the sequence of measurement points of the form (x_k, y_k), $k = 1, 2, \ldots, n$. For such a set of measurement data we wish to find a curve of the form (5.12a) or (5.12b), which is the best approximation of the relation between the two random variables and will enable us to predict[2] from a known value of x the corresponding value of y.

If we know the joint two-dimensional probability distribution of the random variables X and Y, then we can calculate the conditional expected value of one of them with respect to the other. We obtain, according to (3.109), for a continuous probability distribution:

$$E\{X|y\} = \int_{-\infty}^{\infty} x f(x|y)\, dx, \tag{5.13a}$$

---

[2]The relation we defined here makes it possible to complement missing experimental data in a case where one measurement of the pair is unavailable, or to reduce costs of an experiment omitting one of two measurements and calculating the missing number from the regression curve.

$$E\{Y|x\} = \int_{-\infty}^{\infty} y f(y|x) \, dy, \tag{5.13b}$$

and (3.115) for a discrete probability distribution

$$E\{X|y\} = \sum_i x_i p(x_i|y), \tag{5.14a}$$

$$E\{Y|x\} = \sum_j y_j p(y_j|x). \tag{5.14b}$$

Definition 5.1 The set of points of the plane OXY with co-ordinates $(x, E\{Y|x\})$ is called the regression curve of the random variable Y with respect to the random variable X and the set of points with co-ordinates $(E\{X|y\}, y)$ is the regression curve of the random variable X with respect to the random variable Y.

The regression curve can be determined in such a way if the random variable has the finite expected value. For instance, if $E\{Y\}$ exists then also $(x, E\{Y|x\})$ exists and one can draw the regression curve of the random variable Y with respect to the random variable X.

The concept of nonparametric regression is a tool identifying optimal curves of the form (5.12a), (5.12b). Assume that the two random variables (X, Y) are given and we are looking for the formula of the form (5.12a) that minimizes the mean-square distance of the exact value of the random variable Y and its approximate value, e.g.,

$$E\{(Y - g_1(X))^2\} = \min. \tag{5.15}$$

It proves to be that such a function minimizes (5.15) if and only if it is of the form

$$g_1(x) = E\{Y|x\}. \tag{5.16}$$

Analogously, in a case of (5.12b) we have:

$$g_2(y) = E\{X|y\} \tag{5.17}$$

and

$$E\{(X - g_2(Y))^2\} = \min. \tag{5.18}$$

Example 5.1 Let us consider a continuous two-dimensional random variable (Y, Y) with a joint probability distribution function

$$f_{XY}(x, y) = \begin{cases} ((1 + \alpha x)(1 + \alpha y) - \alpha)e^{-(x+y+\alpha xy)} & \text{for } x > 0, y > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{5.19}$$

We will find the regression of the random variable Y with respect to X. We calculate in sequence:

$$f_X(x) = \begin{cases} e^{-x} & x > 0, \\ 0 & x \le 0, \end{cases} \tag{5.20}$$

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} ((1+\alpha x)(1+\alpha y) - \alpha) e^{-(y+\alpha xy)} & \text{for } x > 0, y > 0, \\ 0 & \text{otherwise,} \end{cases}$$

$$(5.21)$$

and

$$E\{Y|x\} = \int_0^{+\infty} yf(y|x) \, dy = \frac{1+\alpha+\alpha x}{(1+\alpha x)^2},$$

that is

$$g_1(x) = \frac{1+\alpha+\alpha x}{(1+\alpha x)^2}. \tag{5.22a}$$

From the symmetry of the probability distribution function (5.19) it results that in the case where the regression of the random variable X with respect to the random variable Y, we obtain the approximation function

$$g_2(y) = \frac{1+\alpha+\alpha y}{(1+\alpha y)^2}. \tag{5.22b}$$

The method of nonparametric regression presented in this subsection is very general but it needs knowledge of two-dimensional joint probability distributions of the random variables for which we seek mutual dependencies. An attempt to find an approximate expression for the probability distribution (which can be easier in practice than finding the exact expression) and then calculating the expression (5.12a) can lead to results less accurate than an alternative method: postulating a priori some functional dependence (with unknown parameters) between the random variables and next identifying the values of the parameters. The method of such a kind is called parametric regression. In a special case of linear dependence of the variables we have linear regression.

5.2.2 The Method of Least Squares for Determining the Linear Regression Line

One can distinguish an important group of experiments when experimental points are laid out in such a manner that one can expect a linear relation between the two magnitudes that are under investigation. In such cases we can assume that the generally written relation (5.12a), (5.12b) takes the form of the equation of a straight line. Two various procedures of calculation are usually recommended.

Procedure 5.1 (Y) *Instead of the general relation* (5.12a) *we assume a linear relation*

$$y = Ax + B, \tag{5.23}$$

where A and B are certain constants to be suitably determined in order to ensure that the calculated straight line well corresponds to the configuration of experimental points.

As the criterion for determining both constants A and B it is assumed that the mean value of squares of all deviations of experimental points from the line (5.23) should have the minimum value. Such a procedure is called the least squares criterion. In the Y-procedure the sum of squared deviations Δy_j^2 of experimental points from the corresponding ordinates of straight line (5.23) being sought should take the minimum value.

Let x_j^* be the abscissa of the j-th experimental point. Then the corresponding ordinate on the straight line (5.23) is

$$y_j^0 = Ax_j^* + B.$$

Thus, the deviation of real ordinate y_j^* of the j-th experimental point from the line (5.23) is

$$\Delta y_j = y_j^* - y_j^0 = y_j^* - Ax_j^* - B.$$

Now the problem consists in finding the values of constants A and B which make the sum of squared deviations possibly small:

$$H = \sum_{j=1}^{n} \left(y_j^* - Ax_j^* - B\right)^2 \rightarrow \min. \tag{5.24}$$

To find these values of the two constants we should differentiate (5.24) first with respect to A, and then with respect to B. The two derivatives should then be equal to zero:

$$\frac{\partial H}{\partial A} = 0, \qquad \frac{\partial H}{\partial B} = 0.$$

Solving the obtained system of equations we arrive at the expressions for A and B

$$A = \frac{\sum_{j=1}^{n} x_j y_j - \frac{1}{n}(\sum_{j=1}^{n} x_j)(\sum_{j=1}^{n} y_j)}{\sum_{j=1}^{n} x_j^2 - \frac{1}{n}(\sum_{j=1}^{n} x_j)^2}, \qquad B = \overline{y} - A\overline{x},$$

where

$$\overline{x} = \frac{1}{n} \sum_{j=1}^{n} x_j, \qquad \overline{y} = \frac{1}{n} \sum_{j=1}^{n} y_j.$$

As an example, in Fig. 5.6 the linear regression line calculated for this particular case is denoted by Y. The angle of its inclination is

$$\beta = 28°57'.$$

Procedure 5.2 (X) *In this procedure, instead of generally written relation (5.12b), we assume the linear form*

$$x = Cy + D, \tag{5.25}$$

where C and D are certain constants to be found. Their values will be calculated using the least squares method, as in the previous Procedure Y.

Let the actual ordinate of a j-th experimental point be y_j^*. To this value of y_j^* corresponds on the straight line (5.25) *a point with the abscissa*

$$x_j^0 = C y_j^* + D.$$

Thus, the deviation of the actual abscissa x_j^* of the j-th experimental point from that on the linear regression line is

$$\Delta x_j = x_j^* - x_j^0 = x_j^* - C y_j^* - D.$$

The sum of squares of all such deviations should have the minimum value. Thus, we can write

$$S = \sum_{j=1}^{n} \left(x_j^* - C y_j^* - D \right)^2 \to \min. \tag{5.26}$$

Now the two constants C and D may be calculated from the system of two equations

$$\frac{\partial S}{\partial C} = 0, \qquad \frac{\partial S}{\partial D} = 0.$$

Finally, we get

$$C = \frac{\sum_{j=1}^{n} x_j y_j - \frac{1}{n} (\sum_{j=1}^{n} x_j)(\sum_{j=1}^{n} y_j)}{\sum_{j=1}^{n} y_j^2 - \frac{1}{n} (\sum_{j=1}^{n} y_j)^2}, \qquad D = \overline{x} - C \overline{y},$$

where

$$\overline{x} = \frac{1}{n} \sum_{j=1}^{n} x_j, \qquad \overline{y} = \frac{1}{n} \sum_{j=1}^{n} y_j.$$

Figure 5.6 *shows the linear regression line calculated in this manner. The angle of its inclination is*

$$\beta = 35°6'.$$

The two linear regression lines *X* and *Y* intersect each other at the "gravity" center of the set of all the experimental points.

5.2.3 The Method of Moments for Determining the Linear Regression Line

Having assumed the coordinate system x, y (comp. Fig. 5.6) we calculate first the position of central point *M* of the set (gravity center) according to formulas (5.1).

Fig. 5.7 Calculation of the
principal components of the
tensor of the second-order
moments by means of Mohr
circles
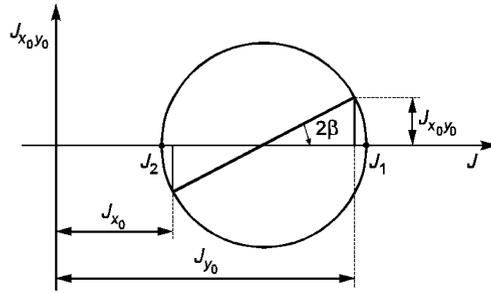
Then as the coordinate reference system, central axes x_0 and y_0 passing through M
are taken.

In the method of moments, instead of variance (5.2), central inertia moments J_{x_0}
and J_{y_0} of the set of all experimental points with respect to the axes x_0 and y_0,

$$J_{x_0} = \sum_{j=1}^{n} (y_j - y_0)^2, \qquad J_{y_0} = \sum_{j=1}^{n} (x_j - x_0)^2, \qquad (5.27)$$

should be calculated. Moreover, instead of covariance (5.3), the mixed moment of
the second order with respect to these axes,

$$J_{x_0 y_0} = \sum_{j=1}^{n} (x_j - x_0)(y_j - y_0), \qquad (5.28)$$

is calculated.

Moments (5.27) and (5.28) represent the components of a symmetrical tensor

$$T_J = \begin{bmatrix} J_{x_0} & J_{x_0 y_0} \\ J_{x_0 y_0} & J_{y_0} \end{bmatrix}. \qquad (5.29)$$

The principal directions and principal components of this tensor may be found, for
example, using the Mohr circles representation.

The angle β of the principal direction of that tensor with respect to the assumed
central coordinate axes x_0, y_0 is determined by the formula

$$\tan 2\beta = \frac{2 J_{x_0 y_0}}{J_{y_0} - J_{x_0}}. \qquad (5.30)$$

Principal components of the tensor T_J of inertia moments are

$$J_1 = \frac{1}{2}(J_{y_0} + J_{x_0}) + \sqrt{\frac{1}{4}(J_{y_0} - J_{x_0})^2 + J_{x_0 y_0}^2},$$

$$J_2 = \frac{1}{2}(J_{y_0} + J_{x_0}) - \sqrt{\frac{1}{4}(J_{y_0} - J_{x_0})^2 + J_{x_0 y_0}^2}. \qquad (5.31)$$

Fig. 5.8 Anisotropy coefficient versus sheet deformation
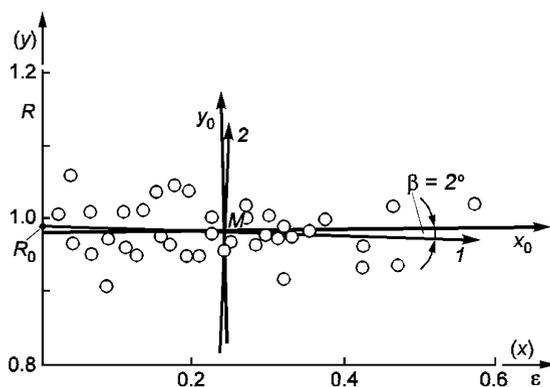
These principal components and their directions have been also determined by using the Mohr circle shown in Fig. 5.7. In the particular case of the set of experimental points shown previously in Fig. 5.6, using formulas (5.27) and (5.28) we get the values of inertia moments

$$J_{y_0} = 0.176, \qquad J_{x_0} = 0.065, \qquad J_{x_0 y_0} = -0.097.$$

The angle β given by (5.30) is

$$\tan 2\beta = -2 \frac{0.097}{0.176 - 0.065} = -1.748,$$

and

$$\beta = 30°7'.$$

This value of the angle is close to that which resulted from procedure Y. Such a linear regression line is shown in Fig. 5.6 as line 1. It represents also the principal axis of the tensor of inertia moments. The principal inertia moment calculated with respect to this line is the smallest of all central inertia moments. Thus, the straight line 1 has such a property that the sum of squares of all orthogonal deviations of experimental points from it assumes the smallest possible value. In the particular case shown in Fig. 5.6, this smallest sum of squared deviations, calculated from the second of formulas (5.31), is

$$J_2 = 0.008.$$

This value may be treated as a measure of the scatter of experimental points.

Another example of a set of experimental points is shown in Fig. 5.8 (comp. [17]). In this case the scatter of the points is considerable, which frequently happens in various investigations. Here a certain coefficient characterizing the initial anisotropy of a sheet metal was measured. Such a coefficient is often determined experimentally with the use of a specimen cut out from the sheet pulled in tension. The deformation, changing in the lateral direction and simultaneously in the thickness direction,

is measured during the process of tensile loading. The most often used anisotropy coefficient R, known as *the Lankford coefficient*, is defined as

$$R = \frac{\ln(\frac{b}{b_0})}{\ln(\frac{h}{h_0})},$$

where h_0 and b_0 stand for the initial thickness and for the initial width of the specimen, respectively, while h is the actual thickness and b is width measured during the test.

Coefficient R changes its value with progressing deformation of the specimen. Note that the initial value R_0 of the sheet cannot be measured directly. It can be estimated, however, in an indirect manner by calculating the linear regression line for experimental points shown in Fig. 5.8, and then by extrapolating it to the zero-deformation point of the diagram.

Following the procedure described above we find inertia moments for that set of points:

$$J_{y0} = 0.679, \qquad J_{x0} = 0.048, \qquad J_{x0y0} = -0.022.$$

The angle β of inclination of the linear regression line determined by (5.30) is

$$\tan 2\beta = -0.0702,$$

and

$$\beta = -2°.$$

The regression line is shown in Fig. 5.8.

5.3 Linear Correlation Between Experimentally Determined Quantities

As mentioned above, in numerous problems of interpreting experimental results it is reasonable to assume a hypothesis of a linear relation between the quantities in question in the form (5.23) or (5.25). In most cases a certain scattering of experimental points takes place as, for example, in Fig. 5.8. Thus, it is important to assess to what degree the assumption of linearity of that relation is justified. After determining the linear regression line we should analyze the so-called correlation between the two quantities which are under investigation. Having the results of measurements in the form of a set of experimental points in the x, y-plane we may, following the procedure described in Sect. 5.1, calculate the estimators of the components of covariance tensor (5.4). First let us assume that this tensor has a particular form represented by a Mohr circle tangent to the vertical covariance axis (Fig. 5.9a). In such a particular case, the smaller principal component of the tensor (comp. Fig. 5.3) is zero

$$\sigma_2^2 = 0.$$
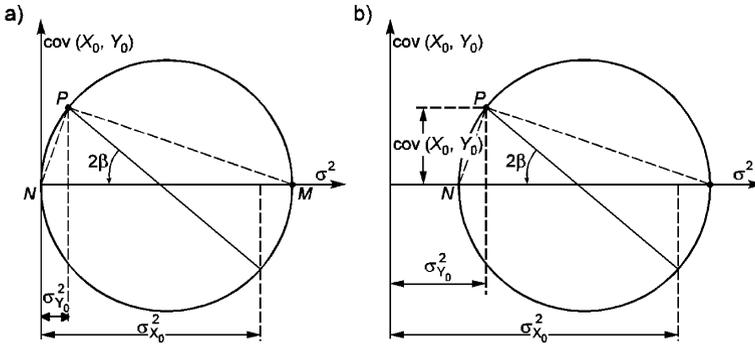
Fig. 5.9 Illustration of the concept of linear correlation by means of Mohr circles
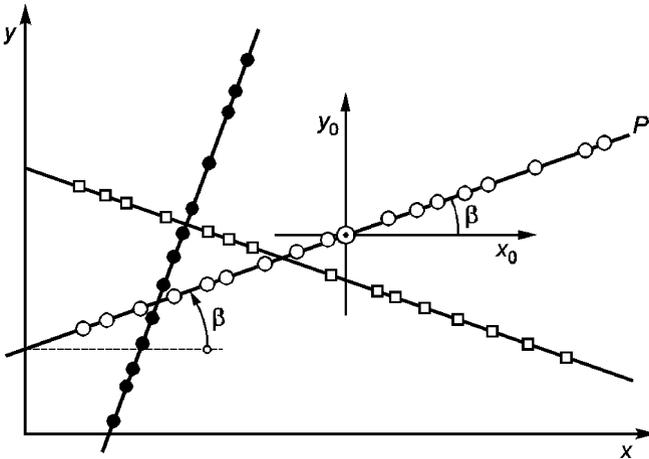
Fig. 5.10 Illustration of a case of the collinear location of sample points

It means that in this particular case all experimental points lie without any deviation on a single straight line inclined to the x-axis by an angle β. This straight line is shown in Fig. 5.10 as line P, if the components of covariance tensor, determined in the manner described in Sect. 5.1, are represented by point P on the Mohr circle, Fig. 5.9a. For other positions of point P on the circle, the corresponding straight lines in Fig. 5.10 will have other positions. Considering a right-angled triangle NPM (Fig. 5.9a) we get the relation

$$\mathrm{cov}\,(X, Y) = \sigma_X \sigma_Y, \tag{5.32}$$

which holds valid for all points on the Mohr circle, when it takes the specific position shown in Fig. 5.9a. For other covariance tensors, when none of the two principal

components vanishes (Fig. 5.9b), we can write instead of (5.32) the inequality

$$|\text{cov}(X, Y)| \leq \sigma_X \sigma_Y, \tag{5.33}$$

which is known as *the Schwartz inequality* (comp., e.g., [9, 16]).

On the basis of relation (5.32), an important coefficient r may be deduced. It allows us to assess to what extent the layout of experimental points may be approximated by a linear relation between the measured quantities X and Y. This coefficient, defined as

$$r = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}, \tag{5.34}$$

is called the *coefficient of linear correlation*, or simply the *correlation coefficient*. It may take values within the range

$$-1 \leq r \leq 1. \tag{5.35}$$

The two limiting values correspond to particular cases when the covariance tensor is represented by a Mohr circle in the position shown in Fig. 5.9a. This means that all experimental points are located exactly on a single straight line.

When the correlation coefficient r calculated for a specific case is close to one of the limiting values (5.35), then the correlation between the measured quantities is high. Small values of this coefficient indicate that the correlation is weak. Small values of the correlation coefficient may be caused by:

— Considerable scatter of experimental points which leads to large values of the two variances σ_X^2 and σ_Y^2; the Mohr circle is then at a considerable distance from the coordinate origin, Fig. 5.9b.
— Small (or large) inclination of the linear regression line, in other words, the angle β being close to zero or to $\pi/2$. Such a situation occurs when the absolute value of the covariance is small. When $\beta = 0$ or $\beta = \pi/2$, there will be no correlation at all.

Returning to the sets of experimental points shown in Fig. 5.1 and in Fig. 5.8 we find that they are characterized by small values of the correlation coefficient.

For the set shown in Fig. 5.1 we get (substituting into (5.34) the values of variances and of the covariance calculated in Sect. 5.1) the correlation coefficient $r = -0.149$. Such a weak correlation of deviations in x and y directions is caused by significant scattering of hit points and additionally by the small value of the angle β. One can say, with a rather high degree of certainty, that during artillery fire from a fixed gun the lateral and longitudinal deviations of hit points from the target point are independent. For the set of experimental points shown in Fig. 5.8 we obtain $r = -0.123$. Such a weak correlation results first of all from the almost horizontal orientation of the correlation line and also from the significant scattering of points.

Another result we obtain by analyzing the set of experimental points shown in Fig. 5.6. The correlation is strong. The estimator of the correlation coefficient is $r = -0.907$.

Remark 5.2 Let us consider again the problem of calculating the regression line. According to the Procedure Y the regression line (5.23) can be written as

$$y - \bar{y} = \hat{r}\frac{\hat{\sigma}_Y}{\hat{\sigma}_X}(x - \bar{x}),$$

where $\bar{x}$ and $\hat{\sigma}_X$ are the estimators of the mean value and the standard deviation of the random variable X, calculated from the experimental data using the known formulas:

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{n}x_j, \qquad \hat{\sigma}_X = \sqrt{\sum_{j=1}^{n}x_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n}x_j\right)^2},$$

and $\bar{y}$ and $\hat{\sigma}_Y$ are the analogous quantities calculated for the random variable Y,

$$\bar{y} = \frac{1}{n}\sum_{j=1}^{n}y_j, \qquad \hat{\sigma}_Y = \sqrt{\sum_{j=1}^{n}y_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n}y_j\right)^2},$$

while $\hat{r}$ is the estimator of the linear correlation coefficient of the two random variables X and Y and is calculated according to the formula

$$\hat{r} = \frac{\sum_{j=1}^{n}x_jy_j - \frac{1}{n}(\sum_{j=1}^{n}x_j)(\sum_{j=1}^{n}y_j)}{\sqrt{\sum_{j=1}^{n}x_j^2 - \frac{1}{n}(\sum_{j=1}^{n}x_j)^2}\sqrt{\sum_{j=1}^{n}y_j^2 - \frac{1}{n}(\sum_{j=1}^{n}y_j)^2}}.$$

The analogous regression line calculated in the Procedure X has the form

$$x - \bar{x} = \hat{r}\frac{\hat{\sigma}_X}{\hat{\sigma}_Y}(y - \bar{y}).$$

Now we can write the regression lines in the slope-intercept form $y = ax + b$. The slope coefficient obtained in the Procedure Y is

$$a = a_{YX} = \hat{r}\frac{\hat{\sigma}_Y}{\hat{\sigma}_X},$$

while the same coefficient obtained in Procedure X is

$$a = a_{XY} = \frac{1}{\hat{r}}\frac{\hat{\sigma}_Y}{\hat{\sigma}_X}.$$

We see that the two coefficients are equal, ($a_{YX} = a_{XY}$) if and only if $\hat{r} = \pm 1$, that is, if the random variables X and Y are linearly correlated.

It is possible to assess the probability that the real value of the correlation coefficient is larger than its value r^* calculated on the basis of formula (5.34) and results

Table 5.1 Probabilities $P_N(|r| \geq r^*)$, that for N measurements of random variables, X and Y will give the correlation coefficient $|r| \geq r^*$ (the empty cells are for probabilities less than 0.01)

N	r^*								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
4	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10
6	0.85	0.70	0.56	0.43	0.31	0.21	0.12	0.056	0.014
8	0.81	0.63	0.47	0.33	0.21	0.12	0.053	0.017	
10	0.78	0.58	0.40	0.25	0.14	0.067	0.024		
15	0.72	0.47	0.28	0.14	0.058	0.018			
20	0.67	0.40	0.20	0.081	0.025				
25	0.63	0.34	0.15	0.048	0.011				
30	0.60	0.29	0.11	0.029					
35	0.57	0.25	0.08	0.017					
40	0.54	0.22	0.06	0.011					
45	0.51	0.19	0.045						
50	0.49	0.16	0.034						
60	0.45	0.13	0.020						
70	0.41	0.097	0.012						
80	0.38	0.075							
90	0.35	0.059							
100	0.32	0.046							

of N measurements. That probability may be written as

$$P_N\left(|r| \geq r^*\right). \tag{5.36}$$

Such an assessment is specially important when the number N is small and, therefore, doubts concerning the degree of correlation arise. Table 5.1 contains the values of the probability (5.36) for various numbers N. Details concerning the method used for calculating the numbers given in the table may be found in specialized monographs [11, 16].

A more precise way of establishing whether the random variables are correlated would be the application of statistical tests (see Chap. 10).

5.4 Two-dimensional Continuous Random Variables

When in such problems as that shown in Fig. 5.4 the number of points tends to infinity, and the number of squares, as those in Fig. 5.5, also tends to infinity, we get instead of a stepwise upper surface of a two-dimensional histogram, a continuous surface $f(x, y)$ of a probability density function. Introducing the concept of

probability distribution function $F(x, y) = P(\{\omega : X(\omega) < x, Y(\omega) < y\})$ of a two-dimensional distribution, comp. (3.101), and assuming its differentiability, we can express the probability density function as the mixed derivative of $F(x, y)$ at a given point x, y (see [2, 10])

$$f(x, y) = \frac{\partial^2 F}{\partial x \partial y}. \tag{5.37}$$

For continuous random variables X, Y the probability distribution function $F(x, y)$ is a continuous function. By analogy to (2.14) we can write

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(\xi, \zeta) \, d\xi d\zeta. \tag{5.38}$$

A probability density function must satisfy the obvious condition that the volume V between the surface representing it and the plane of variables x, y should be equal to unity (comp. (2.16)). Thus, we have

$$V = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx dy = 1. \tag{5.39}$$

Knowing the specific form of the two-dimensional probability density function $f(x, y)$ of a pair of random variables X, Y, we can find the one-dimensional probability density for each of the two variables. They are the so-called marginal distributions, defined as

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy, \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx. \tag{5.40}$$

They describe the distribution of each of the one-dimensional variables, which does not depend upon the distribution of the other variable. They represent projections of the two-dimensional distribution $f(x, y)$ on one of the planes passing through the x or the y-axis. An example of an old-time application of such marginal distributions was presented in Fig. 1.4. For a long time the marginal distributions were used in order to assess the accuracy of artillery fire.

Having found the marginal probability densities we can calculate the first order moments of random variables X and Y by using definition (2.21). They are, respectively, moment of variable X with respect to the y-axis

$$m_X = \int_{-\infty}^{\infty} x f_X(x) \, dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) \, dx dy,$$

moment of variable Y with respect to the x-axis

$$m_Y = \int_{-\infty}^{\infty} y f_Y(y) \, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) \, dx dy.$$

Assuming such specific values $\overline{x}$ and $\overline{y}$ that the equalities

$$V\overline{x} = m_X,$$

$$V\overline{y} = m_Y,$$

hold valid, we can write

$$\overline{x} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y)\, dx dy, \tag{5.41a}$$

$$\overline{y} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y)\, dx dy. \tag{5.41b}$$

Equality (5.39), according to which volume V is equal to unity, has been taken here into account.

The two quantities $\overline{x}$ and $\overline{y}$ are the average values of the distribution (comp. (2.22)). In other words, $\overline{x}$ and $\overline{y}$ are the coordinates of the gravity center of volume V, while moments m_X and m_Y are the static moments of that volume with respect to the axes x and y, respectively. Axes x_0 and y_0, parallel to the initially chosen axes x and y and passing through central point $(\overline{x}, \overline{y})$, are called the central axes of the distribution.

Second-order central moments of volume V with respect to central axes x_0 and y_0 are

$$J_X = \int_{-\infty}^{\infty} (x - \overline{x})^2 f_X(x)\, dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \overline{x})^2 f(x, y)\, dx dy, \tag{5.42a}$$

$$J_Y = \int_{-\infty}^{\infty} (y - \overline{y})^2 f_Y(y)\, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \overline{y})^2 f(x, y)\, dx dy. \tag{5.42b}$$

Assume now two such quantities σ_X^2 and σ_Y^2 that

$$V\sigma_X^2 = J_X, \qquad V\sigma_Y^2 = J_Y.$$

Taking into account that $V = 1$, comp. (5.39), we get the relations:

$$\sigma_X^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \overline{x})^2 f(x, y)\, dx dy, \tag{5.43a}$$

$$\sigma_Y^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \overline{y})^2 f(x, y)\, dx dy, \tag{5.43b}$$

where σ_X^2 and σ_Y^2 are the variances of random variables X and Y, respectively. Square roots of the variances, that is, the quantities σ_X and σ_Y, are the standard deviations of the two random variables.

Note that moments J_X and J_Y, given by formulas (5.42a), (5.42b), are in engineering terms the inertia moments of volume V with respect to central axes y_0 and x_0, respectively. In such an interpretation standard deviations σ_X and σ_Y stand for the so-called inertia radii of volume V.

Continuous random variables X and Y are independent when their joint probability density function $f(x, y)$ can be expressed as a product of univariate probability

densities $f_X(x)$ and $f_Y(y)$,

$$f(x, y) = f_X(x) f_Y(y). \tag{5.44}$$

Similarly, we say that random variables X and Y are uncorrelated when the covariance of their two-dimensional distribution equals zero:

$$\text{cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \overline{x})(y - \overline{y}) f(x, y) \, dx dy = 0. \tag{5.45}$$

Note that covariance (5.45) may be treated as a mixed inertia moment of volume V with respect to central axes y_0 and x_0. Let us assume now that random variables X and Y are independent, which means that their joint probability density function satisfies condition (5.44). Introducing (5.44) into (5.45) and taking into account definitions of average values $\overline{x}$ and $\overline{y}$, we arrive at the conclusion that independent random variables are uncorrelated. As in the previously discussed (in Sect. 5.1) case, the two variances (5.43a), (5.43b) and covariance (5.45) stand for the components of a certain symmetrical tensor

$$T_{\sigma^2} = \begin{bmatrix} \sigma_X^2 & \text{cov}(X, Y) \\ \\ \text{cov}(X, Y) & \sigma_Y^2 \end{bmatrix}. \tag{5.46}$$

The form of this tensor is identical with the form (5.4), but now the interpretation of elements of the matrix on the right-hand side is different.

Principal directions of that tensor and its principal components can be calculated with the use of formulas (5.10) and (5.11).

5.5 The Two-dimensional Normal Distribution

5.5.1 The Case of Independent Random Variables

Let us assume that the one-dimensional random variable X has the normal distribution determined by the probability density function, comp. (2.35),

$$f_X(x) = \frac{1}{\sqrt{2\pi \sigma_X^2}} \exp \left[-\frac{(x - \overline{x})^2}{2\sigma_X^2} \right], \tag{5.47}$$

while the other one-dimensional random variable Y has also the normal distribution with the probability density

$$f_Y(y) = \frac{1}{\sqrt{2\pi \sigma_Y^2}} \exp \left[-\frac{(y - \overline{y})^2}{2\sigma_Y^2} \right]. \tag{5.48}$$
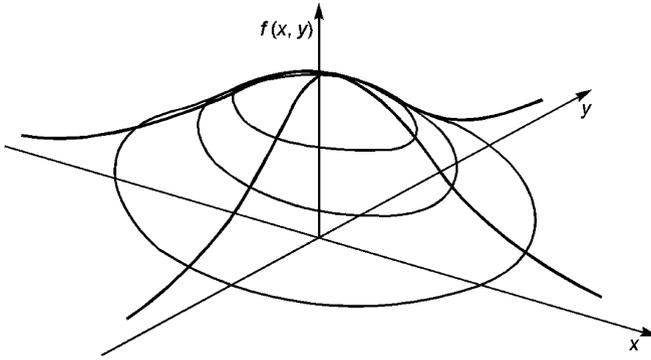
Fig. 5.11 The probability density of the two-dimensional normal distribution of independent random variables

The two-dimensional probability density function $f(x, y)$ of the pair (X, Y) is called their joint probability density function. Since the two random variables are independent, the condition (5.44) must be satisfied. Thus, we can write

$$f(x, y) = \frac{1}{2\pi \sigma_X \sigma_Y} \exp\left[-\frac{(x - \overline{x})^2}{2\sigma_X^2} - \frac{(y - \overline{y})^2}{2\sigma_Y^2} \right]. \tag{5.49}$$

This probability density function is represented by a certain hat-like surface having a single maximum at the central point $(\overline{x}, \overline{y})$. The probability density takes at this point the value $\frac{1}{2\pi \sigma_X \sigma_Y}$. The axes of symmetry of that surface are parallel to the coordinate axes x and y and they pass through central point $(\overline{x}, \overline{y})$. In Fig. 5.11 such a surface is schematically shown for the case when coordinate axes x and y are chosen as principal axes, which means that the mean values of both one-dimensional distributions are assumed to have zero values. In such an important particular case the probability density function (5.49) takes a simpler form

$$f(x, y) = \frac{1}{2\pi \sigma_X \sigma_Y} \exp\left[-\frac{1}{2} \left(\frac{x}{\sigma_X} \right)^2 - \frac{1}{2} \left(\frac{y}{\sigma_Y} \right)^2 \right]. \tag{5.50}$$

The ellipses

$$\left(\frac{x - \overline{x}}{\sigma_X} \right)^2 + \left(\frac{y - \overline{y}}{\sigma_Y} \right)^2 = \text{const.}, \tag{5.51a}$$

or, in the particular case of probability density function (5.50), the ellipses

$$\left(\frac{x}{\sigma_X} \right)^2 + \left(\frac{y}{\sigma_Y} \right)^2 = \text{const.}, \tag{5.51b}$$

on the plane x, y are the lines along which the probability density has a constant value. They are often referred to as the lines of probability concentration. In the

Fig. 5.12 The ellipse of
accuracy of artillery fire [3]
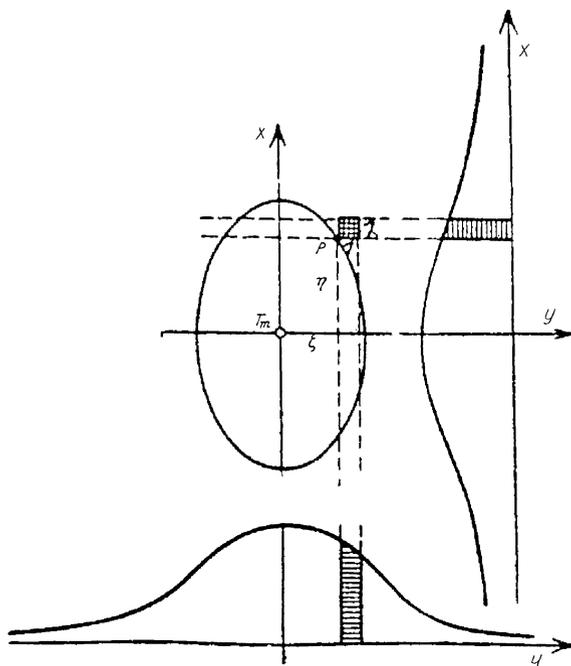
past, practical use of such ellipses was limited to problems of ballistics connected
with the assessment of the accuracy of artillery fire (comp., e.g., Fig. 5.12, [3]). At
present they find also other applications, particularly for assessing the positioning
accuracy of robot manipulators and other similar mechanisms. Such problems will
be discussed in Chap. 6.

5.5.2 The Circular Normal Distribution

The circular normal distribution may be treated as a particular case of the two-di-
mensional normal distribution when

$$\sigma_X = \sigma_Y = \sigma. \tag{5.52}$$

Assume that the coordinate axes x, y of independent random variables X and Y
are central axes. The probability density of such a distribution can be represented
by a certain function of the radius r of a circle on which the point x^*, y^* under
consideration is located. This function results from general formulation (5.50) after
substituting equality (5.52) and after replacing the previous coordinate system x, y
by a polar coordinate system r, θ. Finally, the circular normal probability density
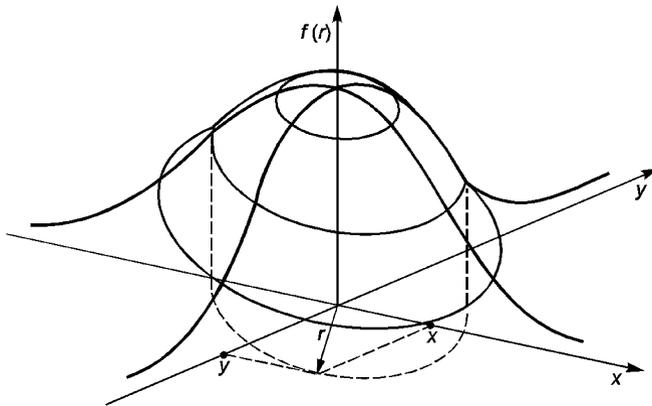
Fig. 5.13 The probability density function of the circular normal distribution

function takes the form

$$f(r, \theta) = \frac{r}{2\pi\sigma^2} \exp\left[-\frac{r^2}{2\sigma^2}\right]. \tag{5.53a}$$

It represents the probability density at each point of a circle with the radius r. However, circular normal probability density may also be presented as a one-dimensional distribution with respect to the radius r (comp., e.g., [4]). In such a case, the probability density for a given radius r will be 2π times larger, see expression (5.53b). The probability density function of the two-dimensional circular normal distribution of a random variable is represented by a surface of revolution shown in Fig. 5.13. The circles $r = $ const are the lines of constant probability. They are also referred to as circles of probability concentration. Such a concentration increasing towards the central point is seen in Fig. 5.14 showing results of approximate computer simulation.[3]

Probability that a two-dimensional random variable X, Y is located inside or on a circle of a radius r^* is equal to the probability that random variable R with the density function $f_R(r)$,
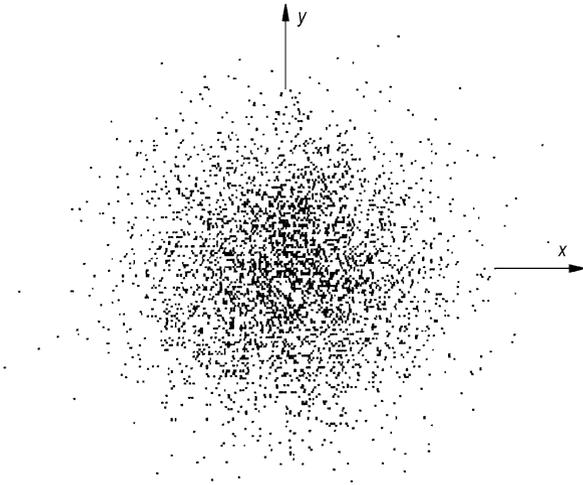
$$f_R(r) = \int_0^{2\pi} f(r, \theta)\, d\theta = \frac{r}{\sigma^2} \exp\left[-\frac{r^2}{2\sigma^2}\right], \tag{5.53b}$$

will be located within the range $(0, r^*)$. Hence, this probability $P(0 \le R \le r^*)$ will be

$$P(0 \le R \le r^*) = \int_0^{r^*} f_R(r)\, dr. \tag{5.54}$$

---

[3]This simulation was made by Professor Z. Wesolowski at the authors' request.

Fig. 5.14 Computer simulation of the circular normal distribution

These probabilities are tabulated in the form of the so-called $\chi^2(2)$ distributions (the chi-squared distribution with two degrees of freedom).

Remark 5.3 Usually, the chi-squared distributions are interpreted in the following way (comp., e.g., [4]). If m independent normalized random variables

$$U_k = \frac{X_k - \overline{x}_k}{\sigma_k} \tag{5.55}$$

have the normal distribution, then the sum of their squares

$$\chi^2 = \sum_{k=1}^{m} U_k^2, \tag{5.56}$$

has the chi-squared distribution with m degrees of freedom.

The quantiles[4] of an order p of such a distribution are denoted by $\chi_p^2(m)$. In numerical tables of the chi-squared distribution usually the values of $\chi_{1-\alpha}^2(m)$ depending on α are given.

The chi-squared distributions for various m find many applications in statistics, and monographs are devoted to them (comp., e.g., [5]). Special attention to them is given also in various books dealing with statistics problems. In the present book we shall limit our attention to the two cases: when $m = 2$ and when $m = 3$, which find direct applications in engineering problems. Examples of practical applications will be given in Chaps. 6 and 8.

---

[4]For a continuous distribution with the cumulative distribution function $F(x)$, the quantile of an order p is such a value x (not necessarily unique) for which equality $F(x) = p$ is satisfied, comp. Sect. 3.4.2

Table 5.2 The quantiles of the $\chi^2(2)$ distribution

$1 - \alpha$	$\chi_\alpha^2(2)$	$1 - \alpha$	$\chi_\alpha^2(2)$
0.99	0.0201	0.30	2.408
0.98	0.0404	0.20	3.219
0.95	0.103	0.10	4.605
0.90	0.211	0.05	5.991
0.80	0.446	0.02	7.824
0.70	0.713	0.01	9.210
0.50	1.386	0.001	13.855

In the case of the two-dimensional circular distribution discussed here, the chi-squared distribution has a direct interpretation. Consider a cylinder whose radius r_α (Fig. 5.13) is determined by the assumed value of probability α that random variable R will lie inside this cylinder (comp. Fig. 2.6 and the analogous problem for the one-dimensional normal distribution). Note that the expression for the square of that radius

$$(r_\alpha)^2 = x^2 + y^2 \equiv \chi_\alpha^2(2) \qquad (5.57)$$

is, according to definition (5.56), a quantile $\chi_\alpha^2(2)$ of an order α of the chi-squared distribution with two degrees of freedom, marked by the number 2 in parentheses. Values of $\chi_\alpha^2(2)$ are given in Table 5.2.

In practical calculations the corresponding diagram presented in Fig. 5.15 may be used.

In statistical tables, usually the values of probability $P(\chi^2 \geq u_0)$ are given. They are compiled under the assumption that the summed up distributions have unit variance, in other words, random variables in question have been determined according to (5.55). Using such tables in the cases when random variables have identical but not unit variances σ^2, we can replace expression $P(\chi^2 \geq u_0)$ by $P(\chi^2 \geq z_0\sigma^2)$, where $u_0 = z_0\sigma^2$. For example, for an arbitrary σ, the quantile of order $\alpha = 0.5$ of the chi-squared distribution with two degrees of freedom is

$$r_{0.5} = \sigma\sqrt{\chi_{0.5}^2(2)} = 1.177\sigma.$$

5.5.3 Three-dimensional Gravity Flow of Granular Media

Before discussing in the following section the normal distribution of dependent random variables, let us present some examples of practical applications of the normal distribution of independent random variables. They concern three-dimensional problems of gravity flow of granular media such as the terrain subsidence caused by underground exploitation or the movements of such media in bins. In Sect. 2.5 were presented two-dimensional solutions of such problems based on the original idea of
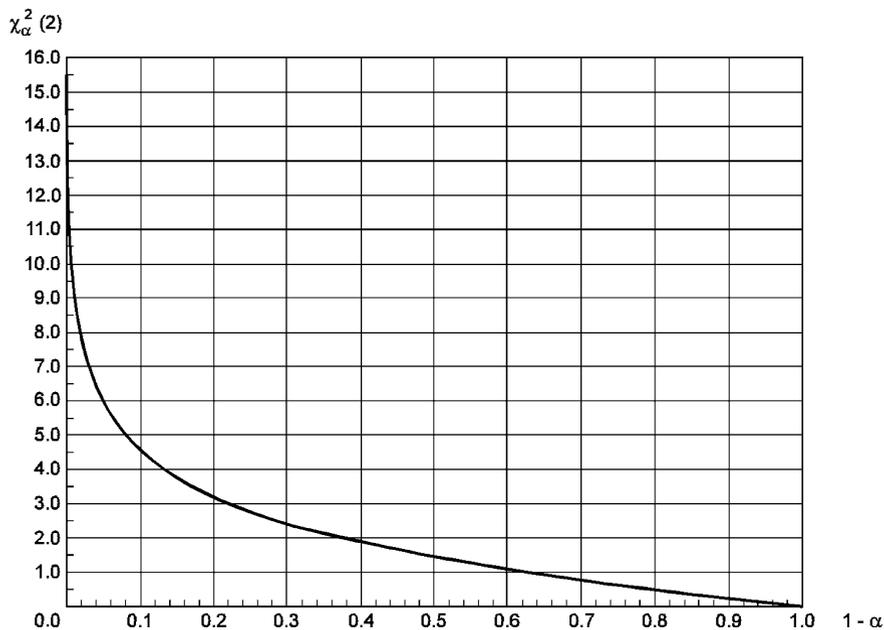
Fig. 5.15 The quantiles of the $\chi^2(2)$ distribution

J. Litwiniszyn [6]. According to Litwiniszyn's approach the displacements in granular medium caused by gravity forces are of the mass character of random changes of mutual contacts between the particles. Consequently, displacements of particles are random. As the starting point for the three-dimensional analysis of such random movements of granular media, let us imagine a demonstrating device composed of a number of plates (layers) resting one on the other. Each plate is formed by a regular array of cuboid cells with h square holes. Four upper plates in the expanded form are shown in Fig. 5.16. The cells in subsequent plates are arranged with respect to each other in such a manner that central axes of holes in a plate coincide with the common line of the four corners of cells in the plate located just below or above, see [12].

Let us assume, similarly as in the so-called Galton's box (Fig. 2.8) for two-dimensional cases, that small balls falling down from a particular cell in plate I and striking the common vertical edges of four cells in plate II below are randomly directed into one of these cells with the probability equal to $\frac{1}{4}$. The random path of the consecutive falling balls is repeated for each plate below. Finally they fall at random into one of the separate containers at the bottom. The probabilities that a ball migrating downwards will pass through a particular cell in a subsequent plate below are shown as an example in Fig. 5.17 for a few upper plates.

If containers at the bottom of the device are located just below the plate IX, then the probabilities that a particular ball will fall into one of them will be as presented in Table 5.3.

Fig. 5.16 Three-dimensional analogue of Galton box, after [12]
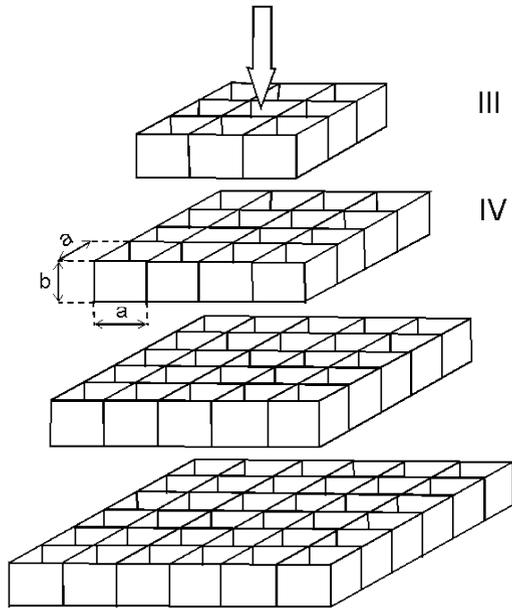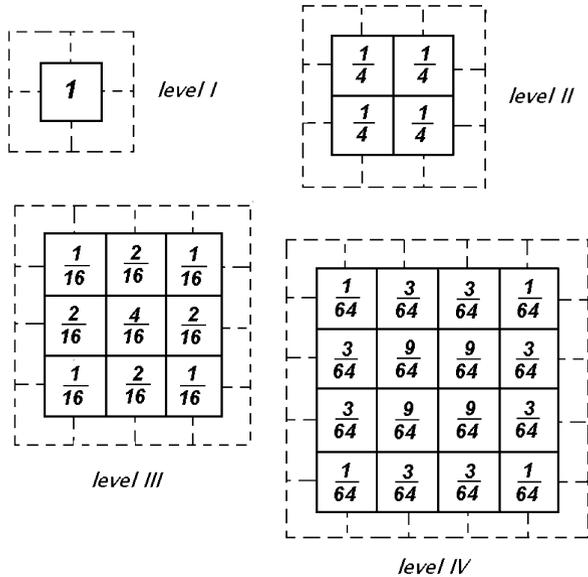

Fig. 5.17 Probabilities that a falling ball passes through cells in four lower plates of the device shown in Fig. 5.16

These calculated values of probabilities are shown in the form of a histogram (Fig. 5.18). It can be seen that the larger is the number of plates in our experimental device, the more the calculated probability distribution approaches the circular normal distribution.

Table 5.3 Probability distribution of a ball falling into containers below level IX

| $\frac{1}{65536}$	$\frac{8}{65536}$	$\frac{28}{65536}$	$\frac{56}{65536}$	$\frac{70}{65536}$	$\frac{56}{65536}$	$\frac{28}{65536}$	$\frac{8}{65536}$	$\frac{1}{65536}$
$\frac{8}{65536}$	$\frac{64}{65536}$	$\frac{224}{65536}$	$\frac{448}{65536}$	$\frac{560}{65536}$	$\frac{448}{65536}$	$\frac{224}{65536}$	$\frac{64}{65536}$	$\frac{8}{65536}$
$\frac{28}{65536}$	$\frac{224}{65536}$	$\frac{784}{65536}$	$\frac{1568}{65536}$	$\frac{1960}{65536}$	$\frac{1568}{65536}$	$\frac{784}{65536}$	$\frac{224}{65536}$	$\frac{28}{65536}$
$\frac{56}{65536}$	$\frac{448}{65536}$	$\frac{1568}{65536}$	$\frac{3136}{65536}$	$\frac{3920}{65536}$	$\frac{3136}{65536}$	$\frac{1568}{65536}$	$\frac{448}{65536}$	$\frac{56}{65536}$
$\frac{70}{65536}$	$\frac{560}{65536}$	$\frac{1960}{65536}$	$\frac{3920}{65536}$	$\frac{4900}{65536}$	$\frac{3920}{65536}$	$\frac{1960}{65536}$	$\frac{560}{65536}$	$\frac{70}{65536}$
$\frac{56}{65536}$	$\frac{448}{65536}$	$\frac{1568}{65536}$	$\frac{3136}{65536}$	$\frac{3920}{65536}$	$\frac{3136}{65536}$	$\frac{1568}{65536}$	$\frac{448}{65536}$	$\frac{56}{65536}$
$\frac{28}{65536}$	$\frac{224}{65536}$	$\frac{784}{65536}$	$\frac{1568}{65536}$	$\frac{1960}{65536}$	$\frac{1568}{65536}$	$\frac{784}{65536}$	$\frac{224}{65536}$	$\frac{28}{65536}$
$\frac{8}{65536}$	$\frac{64}{65536}$	$\frac{224}{65536}$	$\frac{448}{65536}$	$\frac{560}{65536}$	$\frac{448}{65536}$	$\frac{224}{65536}$	$\frac{64}{65536}$	$\frac{8}{65536}$
$\frac{1}{65536}$	$\frac{8}{65536}$	$\frac{28}{65536}$	$\frac{56}{65536}$	$\frac{70}{65536}$	$\frac{56}{65536}$	$\frac{28}{65536}$	$\frac{8}{65536}$	$\frac{1}{65536}$
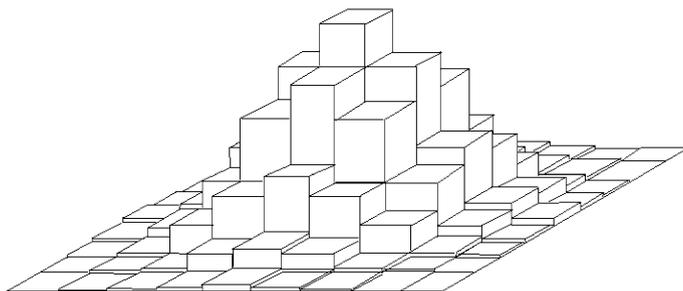
Fig. 5.18 Two-dimensional histogram of the distribution of location of a falling ball in the device presented in Fig. 5.16

As stated in Sect. 2.5, Litwiniszyn analyzed an inverse problem in which the cavities existing in the bulk of a loose medium migrate randomly upwards from the bottom [6–8]. His approach may be used as the basis for the finite cells procedure for the analysis of three-dimensional movements of granular material in bins. As an example let us analyze displacements of a granular medium in a bin of rectangular shape. The outlet has the form of a relatively narrow slit as shown in Fig. 5.19. Let us assume a system of seven plates with square holes (Fig. 5.20), cf. [12].

When a portion of six elementary unit volumes falls down from the bin through the outlet at the bottom, the elementary cavities begin to migrate upwards. They migrate according to the stochastic distribution described above. The probabilities that they pass during the migration through a particular cell in three lower plates are shown in Fig. 5.21.

The corresponding probabilities calculated for the upper plate VII corresponding to the free surface of the bulk are shown in Table 5.4.

Using the calculated data for the upper layer VII we can construct, in a manner analogous to that described in Sect. 2.5 for a two-dimensional flow, the step-wise deformation pattern of the upper surface of the bulk contained in the bin. One half of
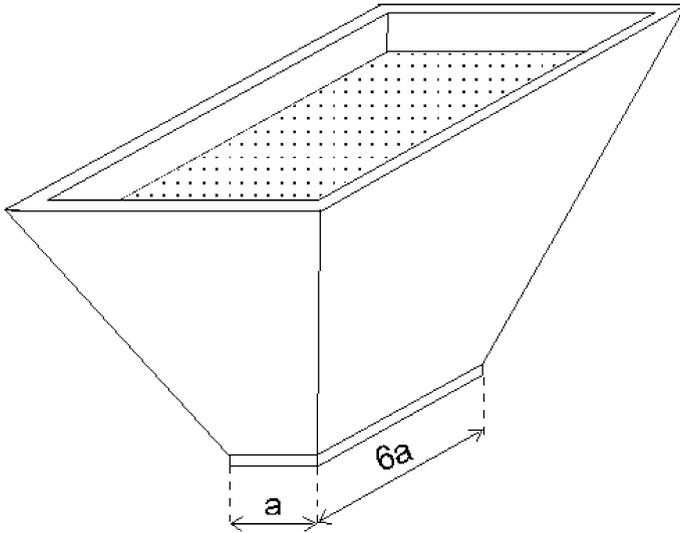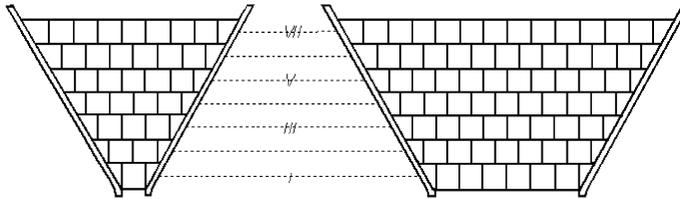
Fig. 5.19 Bin with a rectangular cross-section

Fig. 5.20 Assumed system of unit cells for the problem shown in Fig. 5.19

Table 5.4 Measures that a unit cavity reaches cells of the upper layer VII

$\frac{1}{4096}$	$\frac{7}{4096}$	$\frac{22}{4096}$	$\frac{42}{4096}$	$\frac{57}{4096}$	$\frac{63}{4096}$	$\frac{63}{4096}$	$\frac{57}{4096}$	$\frac{42}{4096}$	$\frac{22}{4096}$	$\frac{7}{4096}$	$\frac{1}{4096}$
$\frac{6}{4096}$	$\frac{42}{4096}$	$\frac{132}{4096}$	$\frac{252}{4096}$	$\frac{342}{4096}$	$\frac{378}{4096}$	$\frac{378}{4096}$	$\frac{342}{4096}$	$\frac{252}{4096}$	$\frac{132}{4096}$	$\frac{42}{4096}$	$\frac{6}{4096}$
$\frac{15}{4096}$	$\frac{105}{4096}$	$\frac{330}{4096}$	$\frac{630}{4096}$	$\frac{855}{4096}$	$\frac{945}{4096}$	$\frac{945}{4096}$	$\frac{855}{4096}$	$\frac{630}{4096}$	$\frac{330}{4096}$	$\frac{105}{4096}$	$\frac{15}{4096}$
$\frac{20}{4096}$	$\frac{140}{4096}$	$\frac{440}{4096}$	$\frac{840}{4096}$	$\frac{1140}{4096}$	$\frac{1260}{4096}$	$\frac{1260}{4096}$	$\frac{1140}{4096}$	$\frac{840}{4096}$	$\frac{440}{4096}$	$\frac{140}{4096}$	$\frac{20}{4096}$
$\frac{15}{4096}$	$\frac{105}{4096}$	$\frac{330}{4096}$	$\frac{630}{4096}$	$\frac{855}{4096}$	$\frac{945}{4096}$	$\frac{945}{4096}$	$\frac{855}{4096}$	$\frac{630}{4096}$	$\frac{330}{4096}$	$\frac{105}{4096}$	$\frac{15}{4096}$
$\frac{6}{4096}$	$\frac{42}{4096}$	$\frac{132}{4096}$	$\frac{252}{4096}$	$\frac{342}{4096}$	$\frac{378}{4096}$	$\frac{378}{4096}$	$\frac{342}{4096}$	$\frac{252}{4096}$	$\frac{132}{4096}$	$\frac{42}{4096}$	$\frac{6}{4096}$
$\frac{1}{4096}$	$\frac{7}{4096}$	$\frac{22}{4096}$	$\frac{42}{4096}$	$\frac{57}{4096}$	$\frac{63}{4096}$	$\frac{63}{4096}$	$\frac{57}{4096}$	$\frac{42}{4096}$	$\frac{22}{4096}$	$\frac{7}{4096}$	$\frac{1}{4096}$

the deformed surface for the particular case when three portions of the unit volumes left the bin through the outlet at the bottom is presented in Fig. 5.22. The step-wise image of the surface may be treated as the first approximation resulting from the
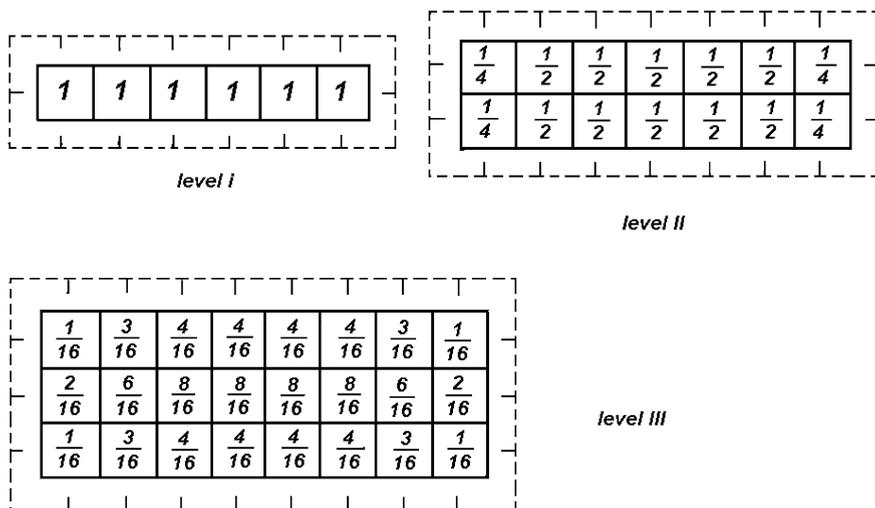
Fig. 5.21 Measures that a unit cavity will pass through cells in three lower layers shown in Fig. 5.20

distribution of the probabilities shown in Table 5.4. Displacement vectors may be found in a manner analogous to that described in Sect. 3.1, see also [12].

As an example of terrain subsidence (cf. [13]) let us analyze displacements in a particular configuration shown in Fig. 5.23. In a large bulk of granular medium, part of which is shown in the figure, we assume the initial cavity in the form of a letter L. Then in the process of gravity flow the cavity is completely filled up by the medium migrating downwards. As the result of this migration process the upper surface of the bulk suffers local deformation and a depression is formed in it.

The assumed system of five plates, each with a regular array of square holes, is presented in Fig. 5.24. The plates are arranged, one with respect to the adjacent one, in the manner shown in Fig. 5.16. Each cell in the system is assumed to represent a unit volume. During the process of deformation all seven elementary empty unit volumes, forming the initial cavity, begin to migrate upwards. They migrate according to the stochastic distribution algorithm described above. The fractions of a unit cavity that finally will appear in particular cells of the upper plate V are shown in Table 5.5. Note that in our problem the sum of these fractions must be equal to seven.

These numbers indicate also how much the surface of the granular medium has been lowered in particular cells as the result of migration of voids. The deformation of the upper surface calculated in this manner is shown in Fig. 5.25. To make the deformation pattern more clearly visible the portion of the bulk bounded in Fig. 5.23 by the planes $ABGF$ and $BCDG$ has been removed.

The method presented in this subsection as well as its two-dimensional variant described in Chap. 2 use elementary facts of probability theory and the calculations can be made even on a pocket calculator. However, the analogous methods can
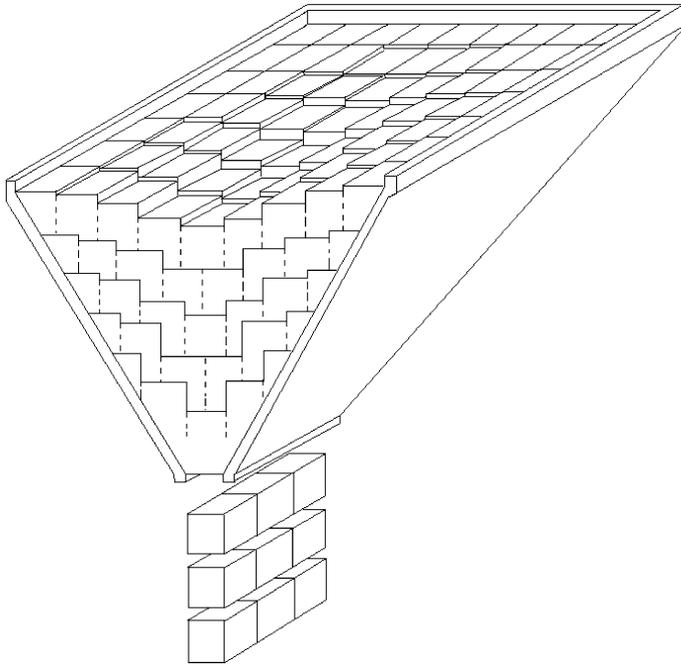
Fig. 5.22 Vertical displacements of granular material in the bin shown in Fig. 5.19

Fig. 5.23 Example of a cavity in the bulk of granular medium
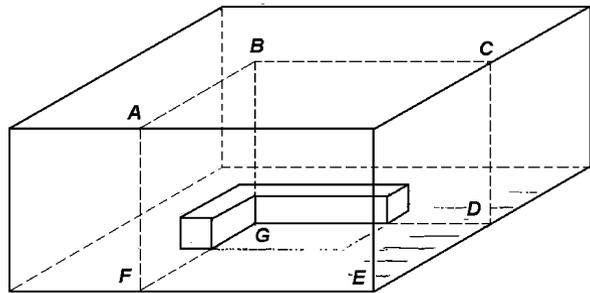
be applied to realistic problems of geotechnics, e.g., the terrain surface subsidence over complex shape drifts in mines or several drifts located at different levels. In such cases one must use computer calculations. Analogously as in widely-used finite elements methods, also in these problems one can apply automatic design of partition of cavities and the ground around it into finite cells.

Fig. 5.24 Assumed system of cells for the terrain subsidence problem shown in Fig. 5.23

Table 5.5 Measures that a unit cavity reaches cells in the upper layer V

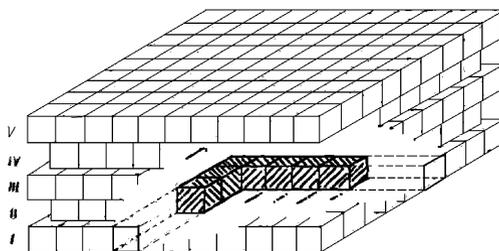|										
0	0	0	0	0	0	0	0	0	0	0
0	$\frac{1}{256}$	$\frac{5}{256}$	$\frac{11}{256}$	$\frac{15}{256}$	$\frac{16}{256}$	$\frac{15}{256}$	$\frac{11}{256}$	$\frac{5}{256}$	$\frac{1}{256}$	0
0	$\frac{5}{256}$	$\frac{24}{256}$	$\frac{50}{256}$	$\frac{64}{256}$	$\frac{65}{256}$	$\frac{60}{256}$	$\frac{44}{256}$	$\frac{20}{256}$	$\frac{4}{256}$	0
0	$\frac{11}{256}$	$\frac{50}{256}$	$\frac{96}{256}$	$\frac{110}{256}$	$\frac{101}{256}$	$\frac{90}{256}$	$\frac{66}{256}$	$\frac{30}{256}$	$\frac{6}{256}$	0
0	$\frac{14}{256}$	$\frac{60}{256}$	$\frac{104}{256}$	$\frac{100}{256}$	$\frac{74}{256}$	$\frac{60}{256}$	$\frac{44}{256}$	$\frac{20}{256}$	$\frac{4}{256}$	0
0	$\frac{11}{256}$	$\frac{45}{256}$	$\frac{71}{256}$	$\frac{55}{256}$	$\frac{26}{256}$	$\frac{15}{256}$	$\frac{11}{256}$	$\frac{5}{256}$	$\frac{1}{256}$	0
0	$\frac{5}{256}$	$\frac{20}{256}$	$\frac{30}{256}$	$\frac{20}{256}$	$\frac{5}{256}$	0	0	0	0	0
0	$\frac{1}{256}$	$\frac{4}{256}$	$\frac{6}{256}$	$\frac{4}{256}$	$\frac{1}{256}$	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Fig. 5.25 Calculated step-wise approximation of terrain subsidence for the problem shown in Fig. 5.23
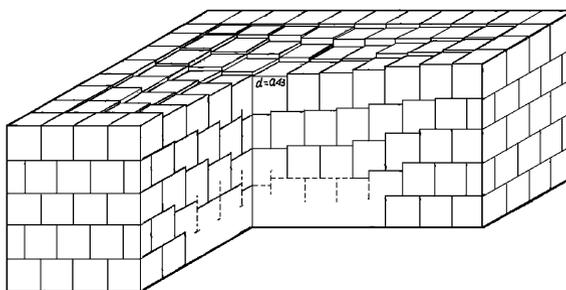
5.5.4 The Case of Dependent Random Variables

If random variables X_1, X_2, each with the normal distribution, are dependent, then their two-dimensional probability density is also referred to as the normal distribution. In the exponent of its probability density function, not only the squares of the two variables x_1, x_2 but also their product appear. Generally, if the axes x_1 and x_2 are chosen to be central axes, the probability density function takes the form (see

Remark 5.4 below)

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\det[k_{ij}]}} \exp\left[-\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2} K_{ij}x_i x_j\right], \qquad (5.58a)$$

or simply

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\det[k_{ij}]}} \exp\left[-\frac{1}{2}\left(K_{11}x_1^2 + 2K_{12}x_1 x_2 + K_{22}x_2^2\right)\right]. \qquad (5.58b)$$

Remark 5.4 The case discussed above is the simplest of n-dimensional distributions of mutually dependent random variables. For such n-dimensional centered normal distributions, the probability density function takes the general form

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det[k_{ij}]}} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} K_{ij}x_i x_j\right]. \qquad (5.59)$$

The normal distribution is fully determined by its central point (the mean value) and by the matrix of second-order moments (the covariance matrix)

$$[k_{ij}] = [K_{ij}]^{-1}, \qquad k_{ij} = k_{ji}, \quad i, j = 1, 2, \ldots, n, \qquad (5.60)$$

where:

for $j = i$, $k_{jj} = \sigma_j^2$ is the variance of X_j,

for $j \neq i$, $k_{ij} = \text{cov}\{X_i, X_j\}$ is the covariance of X_i and X_j.

Returning to the two-dimensional distribution we can write

$$x_1 = x, \qquad x_2 = y,$$

$$K_{11} = K_{XX}, \qquad K_{12} = K_{XY}, \qquad K_{22} = K_{YY}.$$

Coefficients K_{XX}, K_{XY}, and K_{YY} stand for the elements of the inverse matrix to the matrix of second-order moments $[k_{ij}]$. Thus, we have

$$[K_{ij}] = \begin{bmatrix} K_{XX} & K_{XY} \\ K_{YX} & K_{YY} \end{bmatrix} = [k_{ij}]^{-1} = \frac{1}{k_{XX}k_{YY} - k_{XY}^2}\begin{bmatrix} k_{YY} & -k_{XY} \\ -k_{YX} & k_{XX} \end{bmatrix} \qquad (5.61)$$

where

$k_{XX} = \sigma_X^2$ is the variance of the random variable X;
$k_{YY} = \sigma_Y^2$ is the variance of the random variable Y;
$k_{XY} = k_{YX} = \text{cov}\{X, Y\}$ is the covariance of variables X and Y, called also the correlation moment.
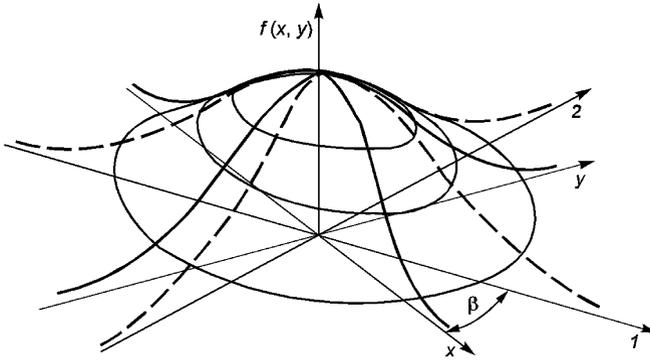
Fig. 5.26 The probability density function of the normal distribution of statistically dependent random variables

Now we can write the probability density function (5.58a) of the two-dimensional normal distribution in a form more convenient for practical applications

$$f\left(x,y\right)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}}$$

$$\times\exp\left[-\frac{1}{2(1-\rho_{XY}^2)}\left(\frac{x^2}{\sigma_X^2}-2\rho_{XY}\frac{x}{\sigma_X}\frac{y}{\sigma_Y}+\frac{y^2}{\sigma_Y^2}\right)\right],\qquad(5.62)$$

where

$$\rho_{XY}=\frac{k_{XY}}{\sqrt{k_{XX}k_{YY}}}=\frac{\mathrm{cov}\{X,Y\}}{\sqrt{\sigma_X^2\sigma_Y^2}}\qquad(5.63)$$

is the coefficient of correlation between variables X and Y. For the zero value of this coefficient, in other words, when $\mathrm{cov}\{X,Y\}=0$, expression (5.62) takes the form (5.50) for independent variables. Thus, the lack of correlation between normal random variables indicates that the variables are mutually independent. The distribution (5.62) is represented by a hat-like surface like that shown in Fig. 5.11. However, now the surface is turned by an angle β with respect to the coordinate axes x and y (Fig. 5.26). The two axes 1 and 2, lying in the symmetry planes of the surface, are the principal axes of the distribution (5.62). Knowing the values of variances σ_X^2 and σ_Y^2 and of the covariance $\mathrm{cov}\{X,Y\}$, we may find the angle β as the angle of each of the principal directions of the covariance tensor (5.46). The Mohr circles representation may be used as it was shown before in Fig. 5.3.

Sections of the surface shown in Fig. 5.26 by planes $f(x,y)=\mathrm{const}$ have an elliptic form (Fig. 5.27). Their projections on the x,y-plane are determined by the equation

$$\frac{1}{1-\rho_{XY}^2}\left[\left(\frac{x}{\sigma_X}\right)^2-2\rho_{XY}\frac{x}{\sigma_X}\frac{y}{\sigma_Y}+\left(\frac{y}{\sigma_Y}\right)^2\right]=\lambda^2,\qquad(5.64)$$

Fig. 5.27 The ellipse of a
constant probability of the
normal distribution of
statistically dependent
random variables
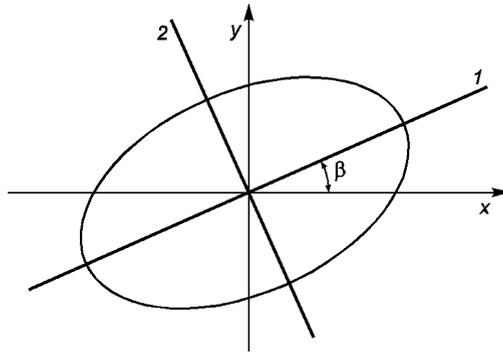
where λ is a certain constant, whose numerical value depends on the assumed probability α that particular values x and y of random variables X and Y will lie inside the ellipse.

If the distribution (5.62) is presented according to (5.55) as a distribution of normalized random variables

$$U_x = \frac{X}{\sigma_X}, \qquad U_y = \frac{Y}{\sigma_Y},$$

it will become a circular distribution with density (5.53a). Ellipse (5.64) will then be transformed into a circle. The procedure showing how to find radius r_α of that circle for an assumed probability α that a point determined by particular values of variables will be located inside it, has been discussed in the previous section. This radius is defined by formula (5.57). Returning to our original distribution (5.62) and to the ellipse (5.64), we do not change this probability. Then, we can write

$$\lambda^2 = \chi_\alpha^2 (2). \tag{5.65}$$

Therefore, the size of the ellipse (5.64) for any assumed probability concentration level may be found.

Example 5.2 Consider a pair of random variables (X, Y) that have jointly a Gaussian distribution with the probability density function of the form:

$$f(x, y) = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho_{XY}^2}}$$

$$\times \exp \left\{ -\frac{1}{2(1 - \rho_{XY}^2)} \left(\frac{(x - m_X)^2}{\sigma_X^2} - 2\rho_{XY} \frac{(x - m_X)}{\sigma_X} \cdot \frac{(y - m_Y)}{\sigma_Y} \right. \right.$$

$$\left. \left. + \frac{(y - m_Y)^2}{\sigma_Y^2} \right) \right\}. \tag{5.66}$$

Now we find the curve of regression of the random variable Y with respect to random variable X. We calculate first the marginal distribution of the random variable which is the condition, that is X:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy = \int_{-\infty}^{\infty} \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho_{XY}^2}}$$

$$\times \exp\left\{ -\frac{1}{2(1 - \rho_{XY}^2)} \left(\frac{(x - m_X)^2}{\sigma_X^2} - 2\rho_{XY} \frac{(x - m_X)}{\sigma_X} \frac{(y - m_Y)}{\sigma_Y} \right. \right.$$

$$\left. \left. + \frac{(y - m_Y)^2}{\sigma_Y^2} \right) \right\} dy$$

$$= \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left\{ -\frac{(x - m_X)^2}{\sigma_X^2} \right\}. \tag{5.67}$$

In the next step we calculate the conditional probability distribution of random variable Y with respect to random variable X:

$$f(y|x) = \frac{f(x, y)}{f_X(x)}$$

$$= \frac{1}{\sigma_Y \sqrt{2\pi(1 - \rho_{XY}^2)}}$$

$$\times \exp\left\{ -\frac{1}{2(1 - \rho_{XY}^2)} \left(\frac{(2\rho_{XY}^2 - 1)(x - m_X)^2}{\sigma_X^2} \right. \right.$$

$$\left. \left. - 2\rho_{XY} \frac{(x - m_X)}{\sigma_X} \frac{(y - m_Y)}{\sigma_Y} + \frac{(y - m_Y)^2}{\sigma_Y^2} \right) \right\}$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{(1 - \rho_{XY}^2)\sigma_Y^2}} \exp\left\{ -\frac{(y - m_Y - \rho_{XY}\frac{\sigma_Y}{\sigma_X}(x - m_X))^2}{2(1 - \rho_{XY}^2)\sigma_Y^2} \right\}. \tag{5.68}$$

The line of regression is the expected value of the random variable conditioned with the second of the pair random variable X, that is the integral of the conditional probability density function (5.68) multiplied by y:

$$E\{Y|x\} = \int_{-\infty}^{\infty} yf(y|x)dy$$

$$= \int_{-\infty}^{\infty} \frac{y}{\sqrt{2\pi}\sqrt{(1 - \rho_{XY}^2)\sigma_Y^2}} \exp\left\{ -\frac{(y - m_Y - \rho_{XY}\frac{\sigma_Y}{\sigma_X}(x - m_X))^2}{2(1 - \rho_{XY}^2)\sigma_Y^2} \right\} dy$$

$$= m_Y + \rho_{XY} \frac{\sigma_Y}{\sigma_X}(x - m_X). \tag{5.69}$$

As it is seen from the above equation, for the two-dimensional Gaussian distribution the regression curve is the straight line, which means that the general regression is always the linear regression. In other words, the regression line is the best approximation of the dependence of the two random variables Y and X among all the curves of the general form (5.12a).

Problem 5.1 In the table are given the independently observed values of some two-dimensional random variable (X, Y). Check if the one-dimensional random vari-

X	12.5	4.7	4.8	3.0	4.0	3.6	6.9	6.9	16.4	5.0
Y	333	123	183	112	118	133	281	281	765	198
X	4.5	6.9	5.0	4.6	4.6	4.5	5.8	6.2	5.6	6.1
Y	135	238	202	213	149	187	341	170	177	310

ables X and Y are correlated.

Problem 5.2 For the random data given in the table calculate the variances and

X	1.91	1.77	1.93	1.67	1.32	1.61	1.52	1.41	1.55	1.96
Y	1.15	1.27	1.22	1.21	1.27	1.29	1.3	1.25	1.18	1.34
X	1.09	1.72	1.93	1.81	1.37	1.25	1.49	1.73	1.11	1.83
Y	1.21	1.25	1.35	1.32	1.32	1.12	1.26	1.27	1.23	1.31

standard deviations of the random variables X and Y and their covariance. Write down their covariance matrix. What is the correlation coefficient of X and Y? Treat the covariance matrix as a tensor and find its principal axes.

Problem 5.3 The results of an experiment are given in the table (X is the strength of fiber in grams, and Y is the strain in %)

X	130	185	165	150	185	165	220	190	170	180
Y	3.8	4.6	4.6	3.8	5.6	4.4	5.6	5.6	4.4	5.2
X	165	185	190	210	150	160	170	175	165	145
Y	4.8	5.2	5.2	5.6	3.8	4.6	3.2	3.6	4.6	4.2

Do the following: calculate the correlation coefficient for the two random variables; calculate the regression coefficients by means of the three methods presented in Sect. 5.2. Compare the obtained results.

Problem 5.4 The two-dimensional random variable (X, Y) has the normal distribution with the mean values vector $(m_X, m_Y) = (1, 2)$ and the covariance matrix $\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$. Find the probability density function of this random variable.

Problem 5.5 Find the ellipses of constant probability for the two-dimensional normal distribution with zero mean value and the covariance matrix:

$$(a) \quad \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}, \qquad (b) \quad \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}, \qquad (c) \quad \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}.$$

Problem 5.6 The random variables X and Y are independent and they have the normal distributions with the probability densities

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right] \quad \text{and} \quad f_Y(y) = \frac{1}{\sqrt{8\pi}} \exp\left[-\frac{y^2}{8}\right],$$

respectively. Find the joint probability density function of the vectorial random variable (X, Y). Find the joint probability density function of the vector-valued random variable (U, V) obtained from the vector (X, Y) by the clockwise rotation of the coordinate system (x, y) through the angle $\alpha = 30^0$.

Problem 5.7 The hit points of artillery fire are concentrated around the point $(x_0, y_0) = (1000, 0)$ and have the normal distribution in x and y directions independent one from another. The variance of the distribution in the x-direction is 100 and in the y-direction it is 50. Find a set on the plane outside of which the probability of the hit point is less than 0.001.

References

1. Breczko, T.: Residual stresses in the shape-memory materials. In: Proceedings of XVII Conference on Experimental Mechanics of Solid Materials, pp. 112–117, Jachranka, Poland, 1996 (in Polish)
2. Fisz, M.: Probability Theory and Mathematical Statistics, 3rd edn. Krieger, Melbourne (1980)
3. Hänert, L.: Geschütz und Schuß. Eine Einführung in die Geschützmechanik und Ballistik. Springer, Berlin (1928)
4. Korn, G.A., Korn, Th.M.: Mathematical Handbook for Scientists and Engineers, 2nd edn. Dover, New York (2000)
5. Lancaster, F.: The Chi-Squared Distribution. Wiley, Chichester (1969)
6. Litwiniszyn, J.: Application of the equation of stochastic processes to mechanics of loose bodies. Arch. Mech. **8**, 393–411 (1956)
7. Litwiniszyn, J.: An application of the random walk argument to the mechanics of granular media. In: Proc. IUTAM Symp. on Rheology and Soil Mechanics, Grenoble, April 1964. Springer, Berlin (1966)
8. Litwiniszyn, J.: Stochastic methods in mechanics of granular bodies. In: CISM Course and Lectures, No. 93. Springer, Udine (1974).
9. Maurin, K.: Methods of Hilbert Spaces. Polish Scientific Editors, Warsaw (1972)
10. Papoulis, A.: Probability, Random Variables, and Stochastic Processes with Errata Sheet, 4th edn. McGraw–Hill, New York (2002)

11. Pugh, E.M., Winslow, G.H.: The Analysis of Physical Measurements. Addison–Wesley, Reading (1966)
12. Szczepiński, W.: On the movement of granular materials in bins. Part II: Three-dimensional problems. Eng. Trans. **51**(4), 433–444 (2003)
13. Szczepiński, W.: On the stochastic approach to the three-dimensional problems of strata mechanics. Bull. Acad. Pol. Sci., Sér. Sci. Tech. **51**(4), 335–345 (2003)
14. Szczepiński, W., Wesolowski, Z.: On certain method of determining the ellipses and ellipsoids of the positioning accuracy of robot manipulators. Appl. Theor. Mech. **31**(3), 465–482 (1993)
15. Szczepiński, W., Wesołowski, Z.: On determining the positioning accuracy of robot manipulators. Bull. Acad. Pol. Sci., Sér. Sci. Tech. **42**(2), 177–186 (1994)
16. Taylor, J.R.: An Introduction to Error Analysis. Oxford University Press, Oxford (1982)
17. Truszkowski, W.: Stress-strain relation of non-homogeneous and anisotropic polycrystalline metals. Arch. Metall. **28**(4), 429–440 (1983) (in Polish)

Chapter 6
Two-dimensional Functions of Independent Random Variables

6.1 Basic Relations

Let us analyze at the beginning a simple example of a manipulator shown schematically in Fig. 6.1. It consists of two arms AB and BO of the length l_1 and l_2, respectively, and of two revolving joints A and B. In each of the repeated working cycles the hand located at the end point O of arm BO should reach the desired nominal position which is determined by nominal coordinates X^0 and Y^0. These coordinates are certain functions of positioning angles θ_1 and θ_2:

$$X = l_1 \cos\theta_1 + l_2 \cos(\theta_1 + \theta_2),$$
$$Y = l_1 \sin\theta_1 + l_2 \sin(\theta_1 + \theta_2). \qquad (6.1)$$

Let θ_1^0 and θ_2^0 denote the desired values of positioning angles. However, in each working cycle they are positioned with a certain small random error $\Delta\theta_1$ and $\Delta\theta_2$ respectively.

In the analysis we shall use a local coordinate system x, y with the origin 0 at the point of nominal position of the hand. We shall analyze the positioning error of the hand caused by random small errors $\Delta\theta_1^0$ and $\Delta\theta_2^0$ in positioning the revolving joints. The positioning error of the hand may be represented as a vector with components

$$U = X - X^0, \qquad V = Y - Y^0, \qquad (6.2)$$

where X and Y are the actual coordinates of the hand's position, and X^0 and Y^0 stand for the coordinates of the nominal position. Generally, in such problems we have to deal with a random variable in the form of a vector of deviation from the average position. Components U and V of this vector are certain functions of independent random variables $Q_1, Q_2, \ldots, Q_n$,

$$U = g(Q_1, Q_2, \ldots, Q_n),$$
$$V = h(Q_1, Q_2, \ldots, Q_n). \qquad (6.3)$$

Fig. 6.1 The scheme of a simple manipulator
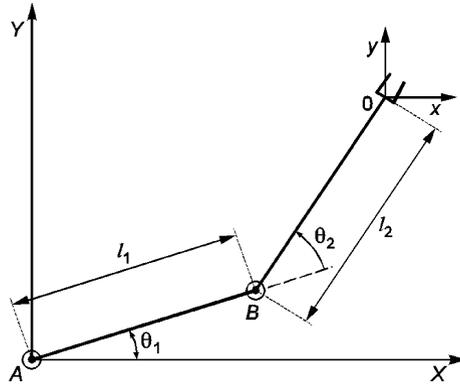
In the particular case discussed here, the angles θ_1 and θ_2 stand for the independent random variables, while the errors in the hand's position (6.2) are the components of this vector.

By analogy with expression (4.10a) for a single function of independent random variables, we can write expressions for the variances of the two random variables U and V

$$\sigma_U^2 = \left(\frac{\partial g}{\partial q_1}\right)^2 \sigma_{Q_1}^2 + \left(\frac{\partial g}{\partial q_2}\right)^2 \sigma_{Q_2}^2 + \cdots + \left(\frac{\partial g}{\partial q_n}\right)^2 \sigma_{Q_n}^2,$$

$$\sigma_V^2 = \left(\frac{\partial h}{\partial q_1}\right)^2 \sigma_{Q_1}^2 + \left(\frac{\partial h}{\partial q_2}\right)^2 \sigma_{Q_2}^2 + \cdots + \left(\frac{\partial h}{\partial q_n}\right)^2 \sigma_{Q_n}^2, \tag{6.4}$$

or in compact form

$$\sigma_U^2 = \sum_{j=1}^n \left(\frac{\partial g}{\partial q_j}\right)^2 \sigma_{Q_j}^2, \qquad \sigma_V^2 = \sum_{j=1}^n \left(\frac{\partial h}{\partial q_j}\right)^2 \sigma_{Q_j}^2. \tag{6.5}$$

The total error of both functions (6.3), caused by deviations Δq_j of independent variables Q_j from their nominal (expected) values, is represented by a vector $\mathbf{p}$ with components

$$\Delta u = U - u^0, \qquad \Delta v = V - v^0, \tag{6.6}$$

where u^0 and v^0 are the nominal values of the two functions, while U and V are random variables representing their actual values.

If the values Δq_i of the small deviations of independent variables are known, then the components of the error vector (6.6) may be calculated from the linearized

relations

$$\Delta u = \frac{\partial g}{\partial q_1} \Delta q_1 + \frac{\partial g}{\partial q_2} \Delta q_2 + \cdots + \frac{\partial g}{\partial q_n} \Delta q_n,$$

$$\Delta v = \frac{\partial h}{\partial q_1} \Delta q_1 + \frac{\partial h}{\partial q_2} \Delta q_2 + \cdots + \frac{\partial h}{\partial q_n} \Delta q_n.$$

(6.7a)

They are written as an approximate generalization of the expressions for total differential

$$du = \sum_{j=1}^{n} \frac{\partial g}{\partial q_j} dq_j, \qquad dv = \sum_{j=1}^{n} \frac{\partial h}{\partial q_j} dq_j.$$

Expressions (6.7a) may be written in a matrix form

$$\begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} = \begin{bmatrix} \frac{\partial g}{\partial q_1} & \frac{\partial g}{\partial q_2} & \cdots & \frac{\partial g}{\partial q_n} \\ \frac{\partial h}{\partial q_1} & \frac{\partial h}{\partial q_2} & \cdots & \frac{\partial h}{\partial q_n} \end{bmatrix} \begin{bmatrix} \Delta q_1 \\ \Delta q_2 \\ \vdots \\ \Delta q_n \end{bmatrix},$$

(6.7b)

or in a short form

$$\mathbf{p} = \mathbf{A} \Delta \mathbf{q},$$

(6.7c)

where the matrix

$$\mathbf{A} = \begin{bmatrix} \frac{\partial g}{\partial q_1} & \frac{\partial g}{\partial q_2} & \cdots & \frac{\partial g}{\partial q_n} \\ \frac{\partial h}{\partial q_1} & \frac{\partial h}{\partial q_2} & \cdots & \frac{\partial h}{\partial q_n} \end{bmatrix}$$

(6.8)

may be treated as the matrix of coefficients of sensitivity of the resultant error $\mathbf{p}$ to the errors Δq_j of independent variables. Sometimes they are called the coefficients of influence. In expression (6.7c) symbols $\mathbf{p}$ and $\Delta \mathbf{q}$ represent the column matrices

$$\mathbf{p} = \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix}, \qquad \Delta \mathbf{q} = \begin{bmatrix} \Delta q_1 \\ \Delta q_2 \\ \vdots \\ \Delta q_n \end{bmatrix}.$$

(6.9)

Let us assume now that errors (fluctuations) Δq_j of independent variables Q_j have a certain given distribution and that they are statistically independent. Moments of the second order of the resultant error $\mathbf{p}$ may be determined in the following manner:

$$[k_{ij}] \overset{\text{def}}{=} \mathbf{A} \sigma^2 \mathbf{A}^T = \begin{bmatrix} k_{UU} & k_{UV} \\ k_{VU} & k_{VV} \end{bmatrix},$$

(6.10)

where

$$\boldsymbol{\sigma}^2 = \begin{bmatrix} \sigma_{Q_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{Q_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{Q_n}^2 \end{bmatrix} \qquad (6.11)$$

is the matrix of variances of independent random variables Q_j, while the matrix

$$\mathbf{A}^T = \begin{bmatrix} \frac{\partial g}{\partial q_1} & \frac{\partial h}{\partial q_1} \\ \frac{\partial g}{\partial q_2} & \frac{\partial h}{\partial q_2} \\ \vdots & \vdots \\ \frac{\partial g}{\partial q_n} & \frac{\partial h}{\partial q_n} \end{bmatrix} \qquad (6.12)$$

is the transposed matrix $\mathbf{A}$ in which respective rows have been presented as columns and vice versa, [1].

Expression (6.10) for the matrix of the second-order moments we shall present in the expanded form of a product of three matrices,

$$[k_{ij}] = \begin{bmatrix} k_{UU} & k_{UV} \\ k_{VU} & k_{VV} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial g}{\partial q_1} & \frac{\partial g}{\partial q_2} & \cdots & \frac{\partial g}{\partial q_n} \\ \frac{\partial h}{\partial q_1} & \frac{\partial h}{\partial q_2} & \cdots & \frac{\partial h}{\partial q_n} \end{bmatrix} \begin{bmatrix} \sigma_{Q_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{Q_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{Q_n}^2 \end{bmatrix} \begin{bmatrix} \frac{\partial g}{\partial q_1} & \frac{\partial h}{\partial q_1} \\ \frac{\partial g}{\partial q_2} & \frac{\partial h}{\partial q_2} \\ \vdots & \vdots \\ \frac{\partial g}{\partial q_n} & \frac{\partial h}{\partial q_n} \end{bmatrix}.$$

The elements of the matrix (6.10) resulting from the product are

$$k_{UU} = \sigma_U^2 = \left(\frac{\partial g}{\partial q_1} \right)^2 \sigma_{Q_1}^2 + \left(\frac{\partial g}{\partial q_2} \right)^2 \sigma_{Q_2}^2 + \cdots + \left(\frac{\partial g}{\partial q_n} \right)^2 \sigma_{Q_n}^2,$$

$$k_{VV} = \sigma_V^2 = \left(\frac{\partial h}{\partial q_1} \right)^2 \sigma_{Q_1}^2 + \left(\frac{\partial h}{\partial q_2} \right)^2 \sigma_{Q_2}^2 + \cdots + \left(\frac{\partial h}{\partial q_n} \right)^2 \sigma_{Q_n}^2, \qquad (6.13)$$

$$k_{UV} = \mathrm{cov}\,(U, V) = \frac{\partial g}{\partial q_1} \frac{\partial h}{\partial q_1} \sigma_{Q_1}^2 + \frac{\partial g}{\partial q_2} \frac{\partial h}{\partial q_2} \sigma_{Q_2}^2 + \cdots + \frac{\partial g}{\partial q_n} \frac{\partial h}{\partial q_n} \sigma_{Q_n}^2.$$

These elements are:

$k_{UU} = \sigma_U^2$ is the variance of the component U of the error vector $\mathbf{p}$;
$k_{VV} = \sigma_V^2$ is the variance of the component V of the error vector $\mathbf{p}$;
$k_{UV} = k_{VU} = \mathrm{cov}(U, V)$ is the covariance of the two components U and V of the error vector $\mathbf{p}$, or in other words, the correlation moment.

Matrix (6.10) is the matrix of a certain tensor. It has the same meaning as matrix (5.46). However, owing to functional relations (6.3), its components may be calculated by using formulas (6.13). Distributions of errors of the two variables U and V are now associated with distributions of errors of independent variables Q_i.

6.2 The Rectangular Distribution of Independent Random Variables

Let us assume that each independent random variable Q_i appearing in two-dimensional functions (6.3) has the uniform (rectangular) distribution, comp. Sect. 3.4.2.

In such a distribution the values of random errors are bounded by two marginal values a and b (Fig. 6.2a) or, alternatively, by the values Δq^+ and Δq^- (Fig. 6.2b), when the nominal value q^0 is taken as a reference point. In practical situations it is difficult to expect that the uniform distribution strictly corresponds to reality, however, it may be useful when a certain class of engineering problems is analyzed. Using such a distribution as an auxiliary one, we can solve in terms of the error calculus various complex problems of the so-called tolerance limits analysis, comp. Sect. 4.4.

6.2.1 Analytical Method for Determining Two-dimensional Tolerance Limits Polygons

Let us analyze the movement of the end-point of the vector **p** of a resultant error with components Δu and Δv, when the error Δq_r of one of the independent variables, namely q_r, is changing, while errors of the remaining independent variables do not change, comp. [5]. Assume that Δq_r is a changing parameter in the equations,
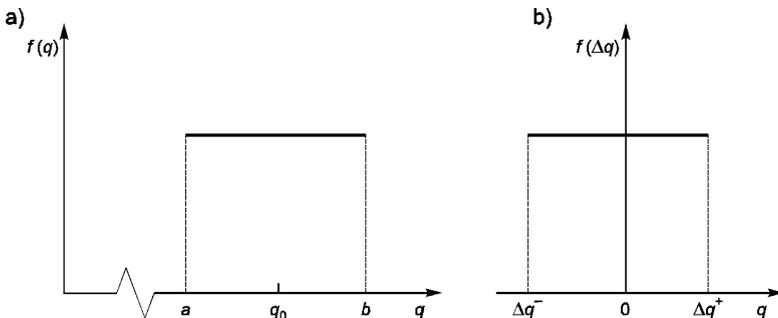

Fig. 6.2 The probability density function of the uniform distribution

comp. (6.7a),

$$\Delta u = \frac{\partial g}{\partial q_1} \Delta q_1 + \cdots + \frac{\partial g}{\partial q_r} \Delta q_r + \cdots + \frac{\partial g}{\partial q_n} \Delta q_n,$$

$$\Delta v = \frac{\partial h}{\partial q_1} \Delta q_1 + \cdots + \frac{\partial h}{\partial q_r} \Delta q_r + \cdots + \frac{\partial h}{\partial q_n} \Delta q_n. \tag{6.14}$$

From the first of these equations we get the expression for Δq_r,

$$\Delta q_r = \frac{1}{\frac{\partial g}{\partial q_r}} \Delta u - \frac{1}{\frac{\partial g}{\partial q_r}} \left(\sum_{j=1}^{r-1} \frac{\partial g}{\partial q_j} \Delta q_j + \sum_{j=r+1}^{n} \frac{\partial g}{\partial q_j} \Delta q_j \right).$$

Then introducing this expression into the second of (6.14) we obtain the equations of a family of straight lines

$$\frac{\partial g}{\partial q_r} \Delta v = \frac{\partial h}{\partial q_r} \Delta u + \sum_{j=1}^{n} \left\{ \det \begin{bmatrix} \frac{\partial g}{\partial q_r} & \frac{\partial g}{\partial q_j} \\ \frac{\partial h}{\partial q_r} & \frac{\partial h}{\partial q_j} \end{bmatrix} \Delta q_j \right\}. \tag{6.15}$$

On the right-hand side the free term is represented in a form of Jacobian.[1] Note that for $j = r$ all the Jacobians are equal to zero because two of their columns are identical.

The distance of each of straight lines (6.15) from the origin (located at the nominal position point) depends on the assumed values of independent variable errors Δq_j. The two extreme positions of each of these straight lines are obtained by substituting appropriately the extreme values (tolerance limits) Δq^- or Δq^+ in order to obtain, first, the largest possible value of the sum on the right-hand side of (6.15), and then its smallest possible value.

This procedure should be repeated for each number r from 1 to n. This means that we should take consecutively each error Δq_j as the changing parameter Δq_r. In this way we obtain n equations of various families of parallel straight lines and then their extreme positions. Note, however, that in practical applications of the theory some families of lines may have identical equations. This means that these families overlap and that the number of independent families is smaller than n. By determining extreme positions of the lines of each family we find a polygon of tolerance limits. All possible positions and magnitudes of the resultant error vector are limited by boundaries of the polygon. Inside the polygon, the probability distribution is uniform according to the presented theory.

Example 6.1 Let us analyze the error of positioning the hand O of the manipulator shown in Fig. 6.1 if its nominal position is determined by the nominal values of the

---

[1] det in (6.15) denotes the determinant of the matrix.

positioning angles of revolving joints A and B,

$$\theta_1^0 = 0, \qquad \theta_2^0 = \frac{3}{4}\pi.$$

The nominal position of the hand is then determined by coordinates X_0, Y_0 given by expressions (6.1) for these values of positioning angles.

The polygon of errors of positioning the hand with respect to its nominal position will be bounded by two pairs of parallel extreme straight lines, each belonging to one of the two families of lines determined by the following equations:

for the first family

$$\frac{\partial Y}{\partial \theta_1} x - \frac{\partial X}{\partial \theta_1} y = \det \begin{bmatrix} \frac{\partial X}{\partial \theta_2} & \frac{\partial X}{\partial \theta_1} \\ \frac{\partial Y}{\partial \theta_2} & \frac{\partial Y}{\partial \theta_1} \end{bmatrix} \Delta \theta_2, \tag{6.16a}$$

and for the second family

$$\frac{\partial Y}{\partial \theta_2} x - \frac{\partial X}{\partial \theta_2} y = \det \begin{bmatrix} \frac{\partial X}{\partial \theta_1} & \frac{\partial X}{\partial \theta_2} \\ \frac{\partial Y}{\partial \theta_1} & \frac{\partial Y}{\partial \theta_2} \end{bmatrix} \Delta \theta_1. \tag{6.16b}$$

Calculations will be performed for

$$l_1 = l_2 = l = 1000 \text{ mm},$$

and for the tolerance limits of the positioning angles θ_1 and θ_2:

$$\Delta \theta_1 = \Delta \theta_2 = \Delta \theta = \pm 0.001\sqrt{3} \text{ rad}.$$

Calculating the respective derivatives of (6.1) and substituting them into (6.16a), (6.16b) we obtain the equations of two pairs of straight lines in extreme positions which determine the polygon of positioning accuracy of the hand in the given position

$$\left(\sqrt{2} - 1\right)x + y = \pm l\Delta\theta = \pm 1.73 \text{ mm},$$

$$x - y = \pm l\Delta\theta = \pm 1.73 \text{ mm}.$$

The polygon is shown in Fig. 6.3.

As stated, the probability distribution is uniform as shown in Fig. 6.4. The volume of this parallelepiped must be equal to unity. Thus, we have

$$f(x, y) = \frac{1}{td}. \tag{6.17}$$

In Fig. 6.3 are also shown the marginal distributions $f(x)$ and $f(y)$, the first being triangular, while the second is trapezoidal.

The ordinate of the marginal distribution $f(x)$ for an arbitrary value of variable x was calculated as the product of density (6.17) and the length of the distribution's
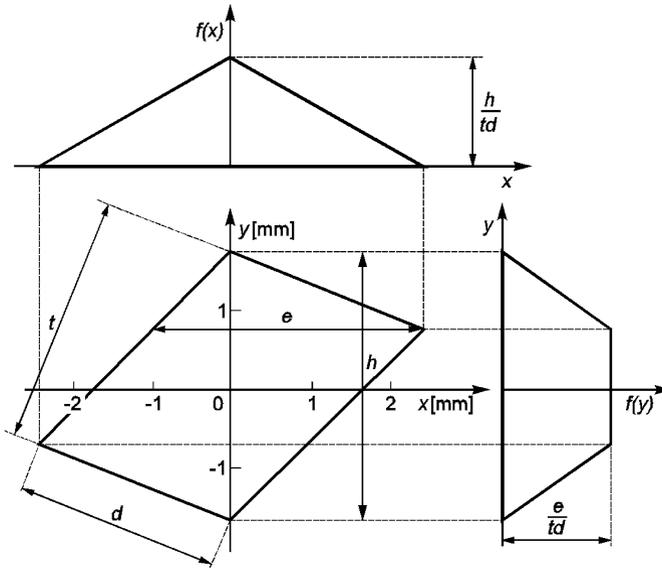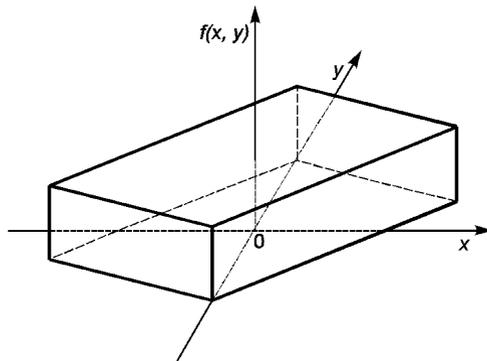
Fig. 6.3 The range of the positioning errors of the hand of the manipulator presented in Fig. 6.1 for the position analyzed in Example 6.1


Fig. 6.4 Two-dimensional probability density function of the positioning error of the hand of the manipulator presented in Fig. 6.1 for the position analyzed in Example 6.1

extent in the y-direction corresponding to that value of x. In a similar manner the ordinates of the marginal distribution $f(y)$ for arbitrary values of variable y may be calculated. Note that the area between the graph of each marginal distribution and the corresponding axis of variable x or y is equal to unity, comp. Sect. 3.6.

Example 6.2 In problems such as that analyzed in the previous example, the distribution of the resultant error in the hand's position usually depends on the nominal position itself. Analyzing the positioning errors of the hand of the mechanism shown

Fig. 6.5 The range of the uniform distribution of the positioning errors of the hand of the manipulator presented in Fig. 6.1 for the position analyzed in Example 6.2
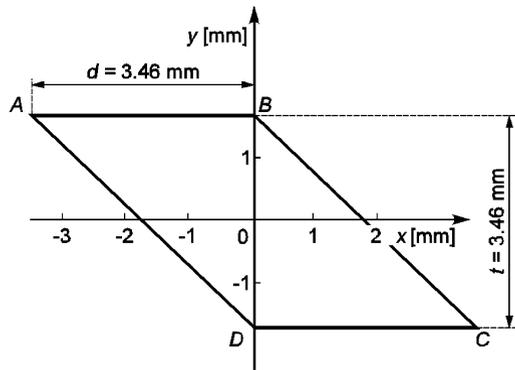
in Fig. 6.1 for another nominal positioning angles of revolving joints, namely for

$$\theta_1^0 = 0, \qquad \theta_2^0 = \frac{\pi}{2},$$

we obtain the polygon of positioning accuracy completely different than that obtained in the previous example.

Repeating the calculations of Example 6.1 we obtain equations of two pairs of straight lines forming a simple polygon of positioning accuracy. For the given positioning angles of revolving joints, and for the same remaining data as in Example 6.1, the equations are

$$
\begin{aligned}
x + y &= \pm l\Delta\theta = \pm 1.73 \text{ mm}, \\
y &= \pm l\Delta\theta = \pm 1.73 \text{ mm}.
\end{aligned}
\tag{6.18}
$$

The polygon is shown in Fig. 6.5. The comparison of the positioning accuracy polygons shown in Figs. 6.3 and 6.5 demonstrates how strongly the resultant positioning error may depend on the nominal position of the hand.

6.2.2 Statical Analogy Method for Determining Two-dimensional Tolerance Limit Polygons

The statical analogy method described in Sect. 4.5 may be useful in determining the tolerance limit polygons in more complex cases, when analytical formulation of functions (6.3) is difficult.

We shall explain the concept of the statical analogy method by considering a particular practical example of a mass-produced mechanism shown in Fig. 6.6. Let us determine the tolerance limits polygon of the position of point D when dimensions $l_1, l_2, \ldots, l_5$ determining that position have tolerance limits indicated in the figure. These dimensions form a certain network as shown in Fig. 6.7. Let us assume the
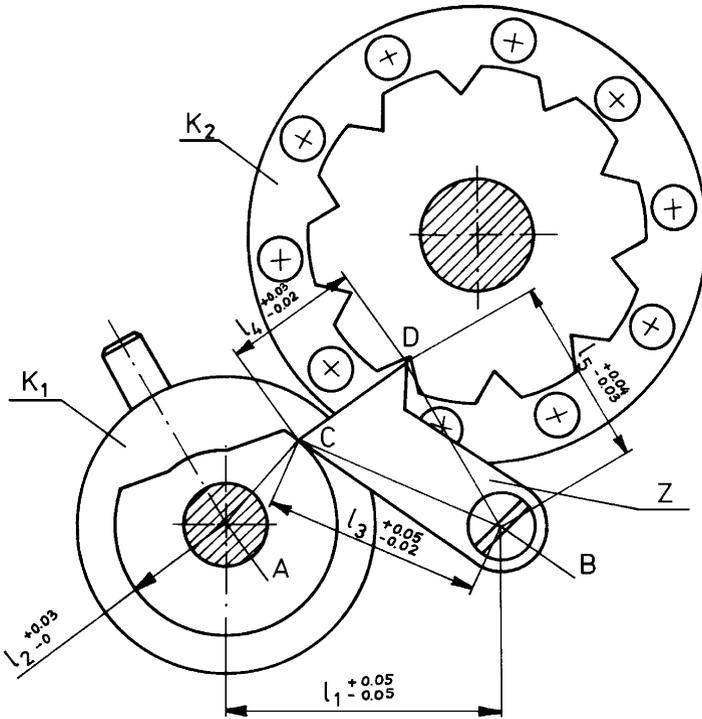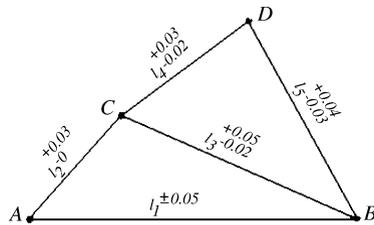
Fig. 6.6 The example of a mass produced mechanism

Fig. 6.7 A network of dimensions for the mechanism presented in Fig. 6.6

coordinate system X, Y with the X-axis passing through reference points A and B. The position of point D is determined by certain functions of dimensions

$$X = X(l_1, l_2, \ldots, l_5),$$

$$Y = Y(l_1, l_2, \ldots, l_5).$$

For nominal values of dimensions these functions determine the nominal position of point D, which will be taken as the origin of the local coordinate system x, y with axes parallel to the axes X and Y, respectively.

Let $\Delta l_1, \Delta l_2, \ldots, \Delta l_5$ be the deviations of dimensions $l_1, l_2, \ldots, l_5$ from their nominal values. Then coordinates x, y of the corresponding position of point D in

Fig. 6.8 A system of rods
corresponding to the network
of dimensions presented in
Fig. 6.7
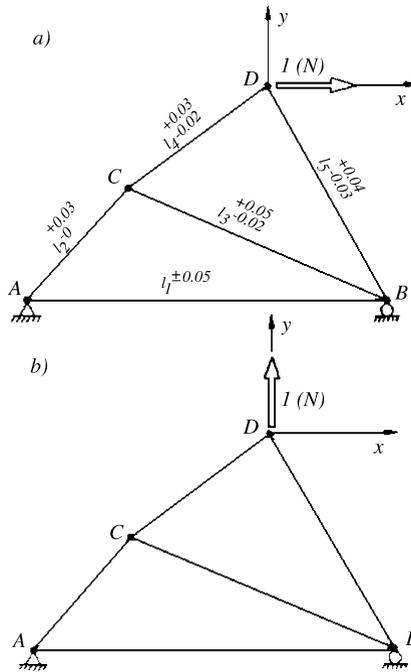
the local coordinate system will be

$$
x = \frac{\partial X}{\partial l_1}\Delta l_1 + \frac{\partial X}{\partial l_2}\Delta l_2 + \cdots + \frac{\partial X}{\partial l_5}\Delta l_5,
$$

$$
y = \frac{\partial Y}{\partial l_1}\Delta l_1 + \frac{\partial Y}{\partial l_2}\Delta l_2 + \cdots + \frac{\partial Y}{\partial l_5}\Delta l_5.
$$

(6.19)

Now let us assume a representative system of rods corresponding to the dimension network shown in Fig. 6.7. It consists of five rods connected by four hinge-joints, see Fig. 6.8. In order to satisfy our assumption that sector AB is taken as the reference X-axis, the system of rods has to be supported at points A and B as shown in the figure.

Let us now apply at point D a unit force $F_x = 1$ acting in the x-direction (Fig. 6.8a). Under such a loading, that point suffers a small displacement whose horizontal component in the local coordinate system is determined by the abscissa of the shifted position of point D from its nominal position. According to the procedure given in Sect. 4.5 we can write for each rod the equality

$$
F_x x_j = s_j^x \Delta l_j,
$$

(6.20)

where s_j^x denotes the internal force in the j-th rod caused by unit force F_x, and x_j is that part of the x-component of the complete displacement of point D which is

Table 6.1 The numerical
values of partial derivatives
for the network of dimensions
presented in Fig. 6.7

j	$s_j^x = \frac{\partial X}{\partial l_j}$	$s_j^y = \frac{\partial Y}{\partial l_j}$
1	$+0.48$	-0.28
2	$+0.79$	$+0.45$
3	-0.19	-0.11
4	$+0.87$	$+0.50$
5	-0.61	$+0.80$

caused by deformation Δl_j of the j-th rod. Noting that $F_x = 1$, we can write

$$x_j = s_j^x \Delta l_j.$$

Summing up all the x_j we finally obtain expressions for complete component x
of the displacement of point D,

$$x = s_1^x \Delta l_1 + s_2^x \Delta l_2 + \cdots + s_5^x \Delta l_5. \tag{6.21a}$$

Applying next a unit force $F_y = 1$ at point D (Fig. 6.8b) we arrive at the analo-
gous expression for the y-component of the displacement of that point,

$$y = s_1^y \Delta l_1 + s_2^y \Delta l_2 + \cdots + s_5^y \Delta l_5. \tag{6.21b}$$

Comparing expressions (6.21a), (6.21b) with (6.19) it is seen that, for $j =
1, 2, \ldots, 5$,

$$s_j^x = \frac{\partial X}{\partial l_j}, \qquad s_j^y = \frac{\partial Y}{\partial l_j}. \tag{6.22}$$

Thus, the numerical values of partial derivatives $\frac{\partial X}{\partial l_j}$ and $\frac{\partial Y}{\partial l_j}$ may be found as the
values of internal forces in a representative statical system loaded by respective unit
forces applied at the point under analysis.

In the example being under consideration these internal forces have been found
to be those in Table 6.1.

Now, introducing these values into the general equation (6.15) and proceeding
in the manner discussed in Sect. 6.2.1 we obtain equations of three families of par-
allel straight lines. Theoretically there should be in this particular case five various
families of lines, each of them corresponding to the changing error in the length
of consecutive five dimensions. However, three families are overlapping and have
the same equation. Thus, the number of independent families reduces to three. Such
situations are often met in practice.

The extreme positions of the lines of the three families form a tolerance limits
polygon shown in Fig. 6.9. All allowable vectors of the displacement of point D lie
inside that polygon.

Inside the polygon are shown two variants of the Williot's diagram for ready
comparison of the results obtained analytically with those obtained with the use
of a graphical method. The latter shall be described in the next section. Note that
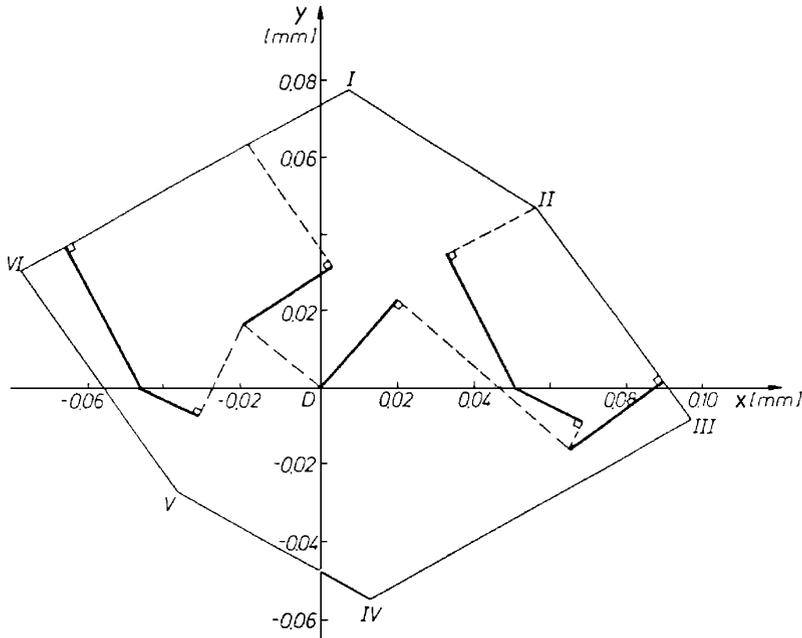
Fig. 6.9 The tolerance limits polygon for the network of dimensions presented in Fig. 6.7

the method of statical analogy may be useful also in the cases when the network contains not only linear but also angular dimensions. In Sect. 4.5 it was shown how to proceed in such cases.

6.2.3 Graphical Method for Determining Two-dimensional Tolerance Limits Polygon. Williot's Diagram

The graphical method of constructing Williot's diagram is mostly used in the mechanics of structures for the analysis of deformations in truss structures. It may, however, be used also in the analysis of tolerance limits in complex nets of dimensions or in the analysis of positioning accuracy of certain classes of mechanisms. We shall present the principles of the graphical method analyzing the problem of positioning accuracy of a mechanism shown schematically in Fig. 6.10. The mechanism has three kinematic joints.

Two of them are prismatic joints allowing the length of a respective arm to be changed, and a single revolving joint at point B. The procedure of constructing the Williot's diagram is shown in Fig. 6.11. Assume that dimension $a = 700$ mm has been given with the upper tolerance limit $\Delta a^+ = +0.1$ mm. It means that the end point of the arm AB has been shifted from the nominal position B to the position B' (in the figure, distance BB' is exaggerated for clarity).
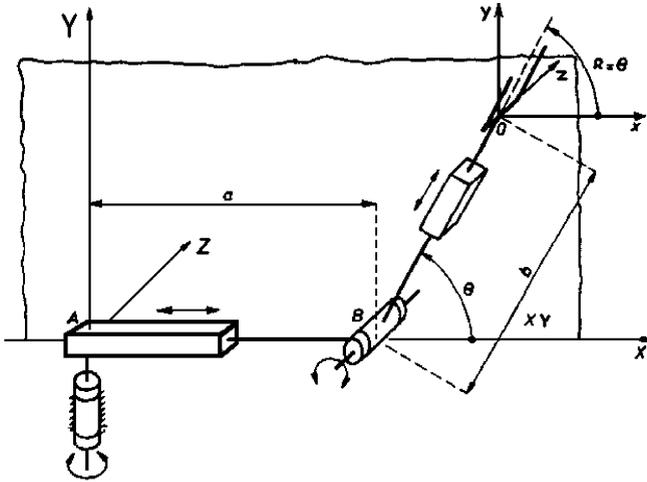
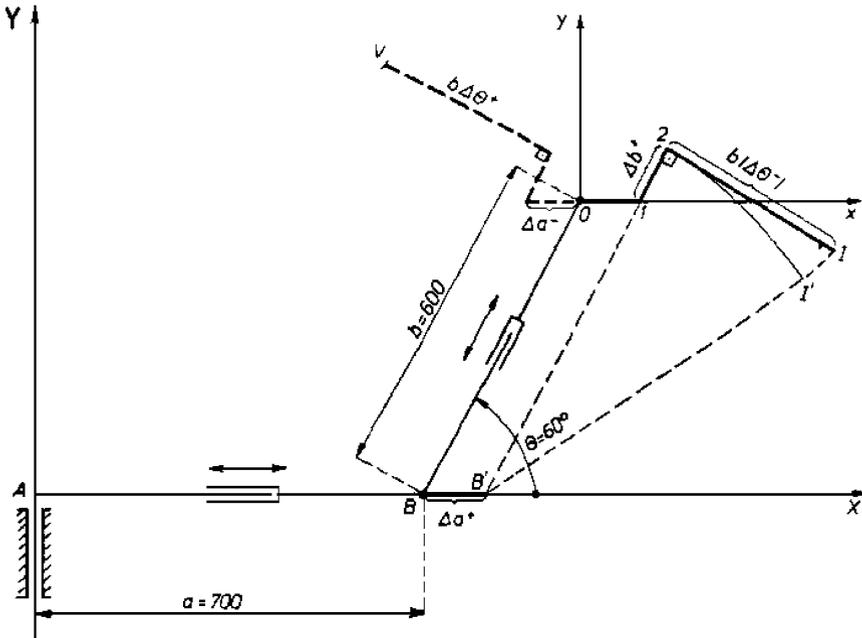**Fig. 6.10** The scheme of a mechanism with three kinematic joints

Fig. 6.11 The procedure of constructing Williot's diagram for the mechanism presented in Fig. 6.10

Thus, the other arm of the mechanism has been also shifted from its nominal position BO to the new position $B'1$. Now let us assume that adjustment of the
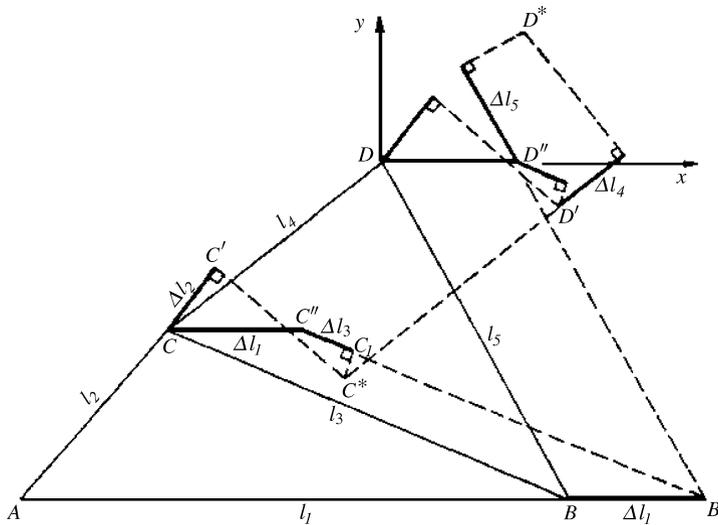
Fig. 6.12 An example of Williot's diagram for the network presented in Fig. 6.7

length of arm BD was done with the deviation from the required nominal length $b = 600$ mm equal to the upper tolerance limit $\Delta b^+ = +0.1$ mm. This deviation is represented in the Williot's diagram by segment 12. Next, assuming that the angle θ, whose nominal value is $\theta = 60°$, has been adjusted with the error $\Delta\theta = -2' = -0.00058$ rad, we should draw an arc $1I'$ of a circle with the center at point B'. The length of that arc is $l = |\Delta\theta^-|b = 600 \times 0.00058 = 0.349$ mm. However, noting that all the deviations are small, we can linearize the problem by plotting a straight segment $2I$ of that length perpendicularly to the preceding sector 12. Point I of the diagram represents in the local enlarged scale of errors the final deviation of the point O of the mechanism from its required nominal position for the chosen combination of the tolerance limits

$$\Delta a^+ = \Delta b^+ = +0.1 \text{ mm}, \qquad \Delta\theta^- = -2'.$$

The graph of deviations in the local coordinate system in the vicinity of point O under consideration is referred to as Williot's diagram. In Fig. 6.11 the broken lines represent such a diagram for a different combination of tolerance limits, namely for

$$\Delta a^- = -0.1 \text{ mm}, \qquad \Delta b^+ = +0.1 \text{ mm}, \qquad \Delta\theta^+ = +2'.$$

Generally, if a net of dimensions is composed of n dimensions (linear or angular), we can draw 2^n different Williot's diagrams.

In Fig. 6.12 is presented another example of the application of Williot's diagram. The network of dimensions shown in Fig. 6.7 of Sect. 6.2.2 shall now be analyzed by the graphical method. Each of five dimensions forming the network has its tolerance limits shown in Fig. 6.7 and collected in Table 6.2.

Table 6.2 Tolerance limits for the network presented in Fig. 6.7 (in mm)

Dimension	l_1	l_2	l_3	l_4	l_5
Upper tolerance limit	+0.05	+0.03	+0.05	+0.03	+0.04
Lower tolerance limit	−0.05	0	−0.02	−0.02	−0.03

Table 6.3 Tolerance limits for Variant I and Variant II (in mm)

Dimension	l_1	l_2	l_3	l_4	l_5
Tolerance limits, Variant I	+0.05	+0.03	−0.02	+0.03	+0.04
Tolerance limits, Variant II	−0.05	0	−0.02	+0.03	+0.04

The number of possible combinations of these tolerance limits, and therefore, the number of variants of Williot's diagrams, is $2^5 = 32$. As in the previous section, we shall take segment AB as the reference axis and we shall analyze the deviation of the position of point D from the nominal position. The analysis will be presented for a particular combination of tolerance limits given in Table 6.3 as Variant I.

Point B is horizontally shifted to the position B' (Fig. 6.12), while point C is shifted at first to the position C' along the AC direction. To determine the final position of that point, it is convenient to first find the auxiliary position of segment $B'C''$ obtained by shifting in parallel segment BC. Then this shifted segment must be shortened by the tolerance limit $\Delta l_3 = -0.02$ mm, as shown in enlarged scale in the figure. Now the final position C^* of this point is found at the intersection of two segments $C'C^*$ and C_1C^* perpendicular to AC and BC, respectively. Next, the whole configuration of displacement at point C has to be translated without rotation to the nominal position of point D. One can say that dimension l_4 has been shifted to the position $C^* - D'$. Similarly, sector $B'D''$ represents the shifted position of dimension l_5. The two dimensions are assumed to be reached with their upper tolerance limits as indicated in Table 6.2 and these tolerance limits have been laid off as shown in the figure. The final position D^* of point D is found at the intersection of the segments perpendicular to these tolerance limits plotted at their end points.

In practice, only the final configuration in the vicinity of the point under consideration is constructed, as shown in Fig. 6.13. In the figure there is also shown another Williot's diagram for a combination of upper and lower tolerance limits given in Table 6.3 (variant II). The two Williot's diagrams are also shown in Fig. 6.9 in order to demonstrate that using the graphical method, one can determine polygons of tolerance limits or polygons of positioning accuracy of a certain class of two-dimensional mechanisms.

Note, however, that in the particular case discussed here, it would be necessary to construct all $2^5 = 32$ Williot's diagrams. Thus, in such cases analytical methods seem to be more useful.
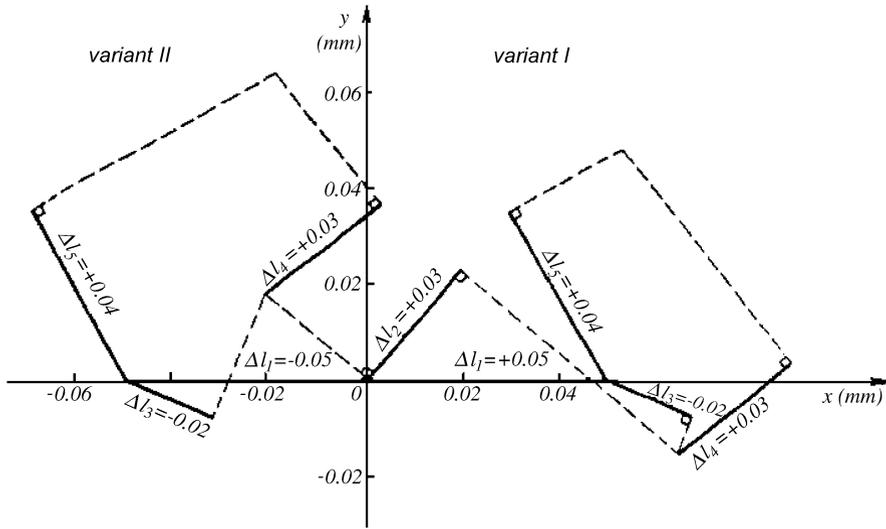
Fig. 6.13 Williot's diagram for Variants I and II

6.3 The Normal Distribution of Independent Random Variables

Assume that each of independent random variables Q_j appearing in the two-dimensional functions (6.3) has the normal distribution determined by the relation, see [3],

$$f\left(\Delta q_j\right) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\Delta q_j}{\sigma_j}\right)^2\right], \qquad (6.23)$$

which defines the probability density that the deviation from the nominal value of variable Q_j equals Δq_j, comp. Sect. 2.4, relation (2.35). If distributions of independent random variables are normal, then the distribution of two-dimensional errors Δu and Δv, in problems of practical interest, can be also considered to be normal.[2]

The probability density function of the two-dimensional normal distribution is usually written with respect to its central axes in the following form:

$$f\left(\Delta u, \Delta v\right) = \frac{1}{2\pi \sigma_U \sigma_V \sqrt{1 - \rho_{UV}^2}}$$

$$\times \exp\left[-\frac{1}{2(1 - \rho_{UV}^2)}\left(\frac{\Delta u^2}{\sigma_U^2} - 2\rho_{UV}\frac{\Delta u}{\sigma_U}\frac{\Delta v}{\sigma_V} + \frac{\Delta v^2}{\sigma_V^2}\right)\right], \quad (6.24)$$

---

[2]Let us recall that the errors of positions U and V are calculated with linear approximation (6.7a), possible to justify for a small range of errors. In general, the distribution of two-dimensional deviations, as a nonlinear function of the normal variables, may not be normal.

where

$$\rho_{UV} = \frac{k_{UV}}{\sqrt{k_{UU}k_{VV}}} = \frac{\text{cov}\{U, V\}}{\sqrt{\sigma_U^2 \sigma_V^2}} \tag{6.25}$$

stands for the coefficient of correlation between the two variables U and V.

Expressions (6.24) and (6.25) are identical with (5.62) and (5.63). However, now they are not obtained as results of statistical analysis of experimental data or recorded observations, but they can be calculated analytically with the use of formulas (6.13) if functions (6.3) are known.

The equation of ellipses of uniform probability, also called the ellipses of probability concentration,

$$\frac{1}{(1 - \rho_{UV}^2)} \left(\frac{\Delta u^2}{\sigma_U^2} - 2\rho_{UV} \frac{\Delta u}{\sigma_U} \frac{\Delta v}{\sigma_V} + \frac{\Delta v^2}{\sigma_V^2} \right) = \lambda^2 = \text{const.} \tag{6.26}$$

is identical with (5.64). But now all coefficients in it may be determined analytically. Below are given some examples of practical applications of such ellipses of probability concentration.

Example 6.3 Let us analyze the positioning error of the hand O of a simple manipulator shown in Fig. 6.1. Assume that its nominal position under analysis is the same as in Example 6.1. Thus, we have

$$\theta_1 = 0, \qquad \theta_2 = \frac{3\pi}{4}. \tag{6.27}$$

The standard deviation of the normal distribution of adjusting accuracy of the two angles is assumed to be

$$\sigma_{\theta_1} = \sigma_{\theta_2} = \sigma_\theta = 0.001 \text{ rad.}$$

Geometrical data are the same as in Example 6.1,

$$l_1 = l_2 = l = 1000 \text{ mm.}$$

The elements of the matrix (6.10) of second-order moments calculated with the use of formulas (6.13) are

$$k_{XX} = \sigma_X^2 = \left[\left(\frac{\partial X}{\partial \theta_1} \right)^2 + \left(\frac{\partial X}{\partial \theta_2} \right)^2 \right] \sigma_\theta^2 = l^2 \sigma_\theta^2 = 1 \text{ mm}^2,$$

$$k_{YY} = \sigma_Y^2 = \left[\left(\frac{\partial Y}{\partial \theta_1} \right)^2 + \left(\frac{\partial Y}{\partial \theta_2} \right)^2 \right] \sigma_\theta^2 = \left(2 - \sqrt{2} \right) l^2 \sigma_\theta^2 = 0.586 \text{ mm}^2,$$

$$k_{XY} = \text{cov}\{X, Y\} = \left(\frac{\partial X}{\partial \theta_1} \frac{\partial Y}{\partial \theta_1} + \frac{\partial X}{\partial \theta_2} \frac{\partial Y}{\partial \theta_2} \right) \sigma_\theta^2$$

$$= \left(1 - \frac{1}{\sqrt{2}}\right) l^2 \sigma_\theta^2 = 0.293 \text{ mm}^2.$$

The coefficient of correlation of the two components of the error vector (formula (6.25)) is

$$\rho_{XY} = \frac{\text{cov}\{X, Y\}}{\sqrt{\sigma_X^2 \sigma_Y^2}} = \frac{1}{2}\sqrt{2 - \sqrt{2}} = 0.383.$$

The equation of the ellipses of uniform probability density (ellipses of probability concentration) is obtained by substituting the magnitudes calculated above into (6.26). After some rearrangements this equation takes the form

$$x^2 - xy + 1,707y^2 = 0.853\lambda^2 l^2 \sigma_\theta^2. \tag{6.28}$$

As stated in Sects. 5.5.2 and 5.5.4, the value of constant λ^2 depends on how large is the required probability that error vector **p** will lay inside the ellipse. Assuming that probability has the value $\alpha = 0.683$, which corresponds to the so-called one-sigma range in the one-dimensional normal distribution (comp. Sect. 2.4), by using the table of the $\chi^2(2)$ distribution (Table 5.2), or the diagram in Fig. 5.15, we obtain the value

$$\lambda^2 = \chi_{0.683}^2 (2) = 2.26.$$

Finally, the equation of the ellipse of probability concentration for $\alpha = 0.683$ is

$$x^2 - xy + 1.707y^2 = 1.928. \tag{6.29}$$

The longer principal axis makes with the x-axis an angle whose magnitude is $\gamma = 27°22'$, Fig. 6.14. The lengths of semi-axes are

$$a = 1.75 \text{ mm}, \qquad b = 1.07 \text{ mm},$$

and their ratio is

$$\frac{a}{b} = 1.63.$$

The theoretical ellipse shown in Fig. 6.14 has been compared with the results of a numerical experiment (comp. [6]). Random small Gaussian deviations from the desired revolving joint positions (6.27) have been numerically generated by a program for a personal computer calculating the displacement of the hand from its nominal position. The random displaced positions of the hand calculated by the program are shown in Fig. 6.15 as the corresponding points. In total, five thousand repeated cycles of the movement have been simulated with randomly generated joint positioning errors. The theoretical ellipse from Fig. 6.14 is also shown in Fig. 6.15 for ready comparison. Note that the theoretical ellipse fits well with the set of numerically simulated points.

Fig. 6.14 The ellipse of the probability concentration for the position (6.27) of the hand of the manipulator presented in Fig. 6.1
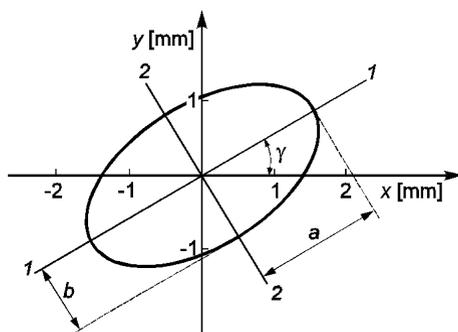

Fig. 6.15 The results of numerical simulation for the position (6.27) of the hand of the manipulator presented in Fig. 6.1
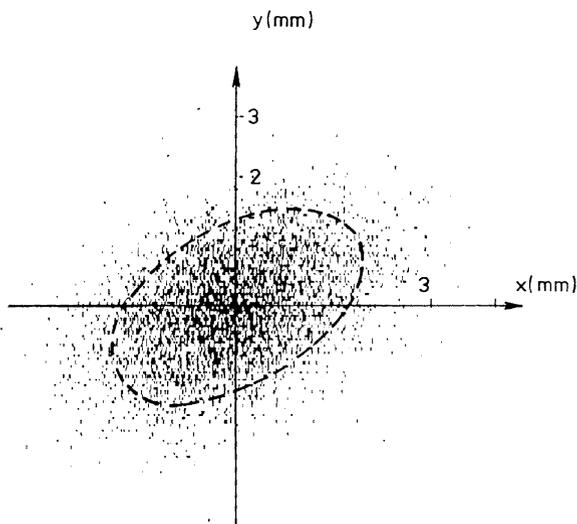
Example 6.4 We shall determine the ellipse of probability concentration for a different position of the same manipulator, see Fig. 6.1. Now the nominal position of the hand is given by the angles

$$\theta_1 = 0, \qquad \theta_2 = \frac{\pi}{2}.$$

Proceeding in the same manner as in Example 6.3, we find the elements of the matrix of second-order moments (6.10).

They are

$$k_{XX} = \sigma_X^2 = 2l^2\sigma_\theta^2 = 2 \text{ mm}^2,$$

$$k_{YY} = \sigma_Y^2 = l^2\sigma_\theta^2 = 1 \text{ mm}^2,$$

$$k_{XY} = \text{cov}(X, Y) = -l^2\sigma_\theta^2 = -1 \text{ mm}^2.$$

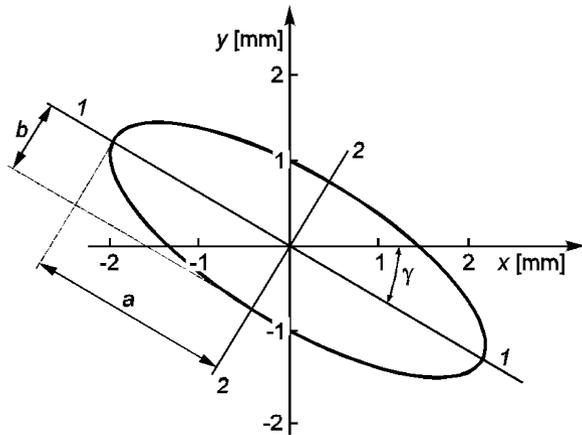The same lengths of manipulator's arms as those used in the previous example and
the standard deviations of positioning accuracy of revolving joints have been as-
sumed.

The correlation coefficient was found to be

$$\rho_{XY} = -\frac{1}{\sqrt{2}}.$$

The ellipse of probability concentration for $\alpha = 0.683$ is determined by the equa-
tion

$$x^2 + 2xy - 2y^2 = \lambda^2 l^2 \sigma_\theta^2 = 2.26. \tag{6.30}$$

The ellipse is shown in Fig. 6.16. The longer axis of the ellipse makes the angle
$\gamma = 31°43'$ with the x-axis. The semi-axes are

$$a = 2.43 \text{ mm}, \qquad b = 0.93 \text{ mm}.$$

Their ratio is

$$\frac{a}{b} = 2.62.$$

Note how different are the ellipses of probability concentration for the same mech-
anism for different nominal positions of the hand.

6.4 Indirect Determination of the Ellipses of Probability Concentration

The ellipses of probability concentration may also be found indirectly, provided that
all independent random variables Q_j in functions (6.3) have the normal distribu-
tion, comp. [2, 4]. First, we assume that independent variables Q_j have a substitute

rectangular distribution instead of the actual normal distribution. Next, the polygon which limits the region of uniform distribution of errors Δu and Δv of the two functions (6.3) should be determined in the manner discussed in Sect. 6.2.1. The limiting values of the substitute rectangular distributions of independent variables errors

$$\Delta q_j^+ = +\kappa \sigma_j \quad \text{and} \quad \Delta q_j^- = -\kappa \sigma_j,$$

depend on the assumed probability that end points of the error vector $\mathbf{p}$, whose components are Δu and Δv, will lie inside the polygon. Here, σ_j stands for the standard deviation of the actual normal distribution of variable Q_j. The polygon calculated under the assumption that $\kappa = \sqrt{3}$ corresponds to the probability $\alpha = 1$. For each $\kappa < \sqrt{3}$ the corresponding probability may be calculated as the ratio of the area of the polygon obtained for $\Delta q_j = \pm \kappa \sigma_j$ and the area of the polygon for $\Delta q_j = \pm \sqrt{3} \sigma_j$. Thus, the multiplier κ should have the value

$$\kappa = \sqrt{3}\sqrt{\alpha}, \tag{6.31}$$

where α is the assumed probability level.

Having found the polygon we calculate its second-order moments (inertia moments). Then its principal axes 1 and 2 and its principal second-order moments I_1, I_2 should be determined. The ellipse of probability concentration may be found as the ellipse with the same principal axes as those of the polygon. Moreover, its principal second-order moments should have the same values as the corresponding moments of the polygon. Thus, the lengths a and b of this ellipse may be found by solving the system of two equations

$$\frac{\pi a b^3}{4} = I_1, \qquad \frac{\pi a^3 b}{4} = I_2, \tag{6.32}$$

where I_1 and I_2 are the principal second-order moments of the polygon.

Example 6.5 Let us find by means of the indirect method the ellipse of probability concentration for the problem of positioning accuracy of the mechanism shown in Fig. 6.1. Assume the probability level to be $\alpha = 0.683$ and the nominal position to be determined by the positioning angles

$$\theta_1 = 0, \qquad \theta_2 = \frac{\pi}{2}.$$

These data correspond to those assumed in Examples 6.2 and 6.4 analyzed in Sects. 6.2.1 and 6.3.

The polygon determined in Example 6.4 is shown in Fig. 6.5. It was found under the assumption that the multiplier κ has the value $\kappa = \sqrt{3}$, that is for the probability $\alpha = 1$. In the present case the probability level is assumed to be $\alpha = 0.683$. Thus, using formula (6.31) we get

$$\kappa = \sqrt{3}\sqrt{\alpha} = 1.43.$$

Fig. 6.17 An application of the indirect method of determination of the ellipse of probability concentration, Example 6.5
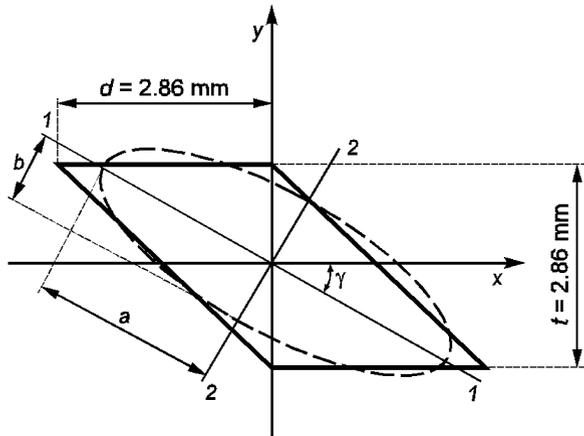
Now the polygon, being geometrically similar to that shown in Fig. 6.5, has smaller dimensions, namely

$$t = 2.86 \text{ mm}, \qquad d = 2.86 \text{ mm}.$$

The second-order moments (inertia moments) with respect to coordinate axes are

$$I_x = \frac{1}{12}dt^3 = 5.576 \text{ mm}^4,$$

$$I_y = \frac{1}{12}d^3t = 5.576 \text{ mm}^4,$$

$$I_{xy} = \frac{1}{12}d^2t^2 = 5.576 \text{ mm}^4.$$

The principal second-order moments are

$$I_1 = \frac{1}{2}\left(I_x + I_y\right) - \sqrt{\frac{1}{4}\left(I_x - I_y\right)^2 + I_{xy}^2} = 2.128 \text{ mm}^4,$$

$$I_2 = \frac{1}{2}\left(I_x + I_y\right) + \sqrt{\frac{1}{4}\left(I_x - I_y\right)^2 + I_{xy}^2} = 14.599 \text{ mm}^4.$$

Note that these formulas are analogous to (5.11).

By solving the system of (6.32) we find the lengths a and b of semi-axes of the ellipse of probability concentration

$$a = 2.64 \text{ mm}, \qquad b = 1.01 \text{ mm}.$$

The ellipse is shown in Fig. 6.17 along with the polygon. The ratio of semi-axes lengths is

$$\frac{a}{b} = 2.61.$$

The angle γ, which determines the position of the principal axis of the polygon, which is the longer axis of the ellipse, is given by the formula

$$\tan 2\gamma = \frac{2I_{XY}}{I_Y - I_X} = -2.$$

Hence,

$$\gamma = -31°43'.$$

This value is identical with that obtained by the direct method, comp. Example 6.4 and Fig. 6.16. The two ellipses are geometrically similar, their dimensions being only slightly different.

Problem 6.1 Determine analytically the polygon of positioning tolerance limits for the manipulator shown in Fig. 6.10. Assume the data:

$$a = 700 \text{ mm}, \qquad b = 600 \text{ mm}, \qquad \theta = 60°,$$

and the position tolerance limits:

$$\Delta a = \pm 0.1 \text{ mm}, \qquad \Delta b = \pm 0.1 \text{ mm}, \qquad \Delta \theta = \pm 2' = \pm 0.00058 \text{ rad}.$$

Problem 6.2 Find the Williot's diagrams for all eight combinations of the joint positioning tolerance limits of the manipulator shown in Fig. 6.10.

Problem 6.3 Assuming for the manipulator shown in Fig. 6.10, dimensions a and b and positioning tolerance limits Δa, Δb, $\Delta \theta$ of the joints are the same as in Problem 6.1, determine analytically the polygons of the positioning accuracy for two special positions of the hand determined by the positioning angles

$$\theta = 0°, \qquad \theta = 90°.$$

Problem 6.4 Determine analytically the ellipse of probability concentration for the manipulator shown in Fig. 6.1, assuming the same data as in Examples 6.3 and 6.4. The position of the hand is now determined by the joint positioning angles

$$\theta_1 = 0, \qquad \theta_2 = \frac{1}{4}\pi.$$

Problem 6.5 Using the indirect method find the lengths of the semi-axes and the orientation of the ellipse of probability concentration for the problem of positioning accuracy of the manipulator shown in Fig. 6.1. Assume the data as in Example 6.5 and the hand's position determined by the positioning angles

$$\theta_1 = 0, \qquad \theta_2 = \frac{3}{4}\pi.$$

References

1. Bellman, R.: Introduction to Matrix Analysis. McGraw–Hill, New York (1970)
2. Cramer, H.: Mathematical Methods of Statistics. Princeton University Press, Princeton (1946)
3. Korn, G.A., Korn, Th.M.: Mathematical Handbook for Scientists and Engineers, 2nd edn. Dover, New York (2000)
4. Kotulski, Z., Szczepiński, W.: On two methods of determining the ellipses and ellipsoids of the positioning accuracy of robot manipulators. Ing. Arch. (Arch. Appl. Mech.) **63**(1), 25–41 (1993)
5. Szczepiński, W.: Theory of polyhedron of positioning accuracy of manipulators. Mech. Mach. Theory **26**(7), 697–709 (1991)
6. Szczepiński, W., Wesołowski, Z.: On a certain method of determining the ellipses and ellipsoids of the positioning accuracy of robot manipulators. Appl. Theor. Mech. **31**(3), 465–482 (1993)

Chapter 7
Three-dimensional Distributions

7.1 General Remarks

Three-dimensional random distributions of points do not have such wide applications in engineering practice as two-dimensional distributions. Of practical significance, however, are three-dimensional (vectorial) functions of random variables which will be discussed in Chap. 8, see [6, 8]. In this chapter we shall discuss the basic principles of the theory of three-dimensional distributions of errors.

Let us consider a certain random distribution of points in a Cartesian three-dimensional coordinate system x, y, z. By analogy with the two-dimensional distributions analyzed in Sect. 5.1 we can write expressions for the coordinates of central point M of the set of points being analyzed. These expressions are

$$\overline{x}_M = \frac{1}{n} \sum_{j=1}^{n} x_j, \qquad \overline{y}_M = \frac{1}{n} \sum_{j=1}^{n} y_j, \qquad \overline{z}_M = \frac{1}{n} \sum_{j=1}^{n} z_j, \qquad (7.1)$$

where n is the number of all the points in the set. If coordinates x, y, z of particular points are treated as random variables, then the estimators of their variances are defined by the formulas

$$\sigma_X^2 = \frac{1}{n} \sum_{j=1}^{n} \left(x_j - \overline{x}_M\right)^2, \qquad \sigma_Y^2 = \frac{1}{n} \sum_{j=1}^{n} \left(y_j - \overline{y}_M\right)^2,$$

$$\sigma_z^2 = \frac{1}{n} \sum_{j=1}^{n} \left(z_j - \overline{z}_M\right)^2. \qquad (7.2)$$

The quantities $\overline{x}_M, \overline{y}_M, \overline{z}_M$ defined by (7.1) are treated here as estimators of expected values of the random variables.

The variances (7.2) may be also treated as inertia moments of the set of points in question divided by the number of all points. These inertia moments should be calculated with respect to the central planes $x_0 = 0$, $y_0 = 0$, $z_0 = 0$ of the coordinate

system x_0, y_0, z_0 with origin at the central point M and the axes parallel to original coordinate axes x, y, z.

To verify whether random variables X, Y, Z are independent (under assumption that their distributions are normal), see [4, 7], one should calculate the covariances

$$\mathrm{cov}\,(X, Y) = \frac{1}{n} \sum_{j=1}^{n} (x_j - \overline{x}_M)(y_j - \overline{y}_M),$$

$$\mathrm{cov}\,(Y, Z) = \frac{1}{n} \sum_{j=1}^{n} (y_j - \overline{y}_M)(z_j - \overline{z}_M), \tag{7.3}$$

$$\mathrm{cov}\,(Z, X) = \frac{1}{n} \sum_{j=1}^{n} (z_j - \overline{z}_M)(x_j - \overline{x}_M).$$

The covariances may be treated as the so-called mixed inertia moments of the set of points under consideration divided by the number of all the points. These moments should be calculated with respect to the central coordinate system x_0, y_0, z_0.

If the random variables (being coordinates of particular points) are independent, the corresponding covariances (7.3) are equal to zero. For random variables with the normal distribution, the inverse theorem stating that their independence results from the lack of correlation is also valid [3]. In Chap. 8 this will be demonstrated on practical applications.

Variances (7.2) and covariances (7.3) are, like the mentioned above inertia moments, the components of a certain symmetrical tensor

$$T_{\sigma^2} = \begin{bmatrix} \sigma_X^2 & \mathrm{cov}(X, Y) & \mathrm{cov}(X, Z) \\ \mathrm{cov}(Y, X) & \sigma_Y^2 & \mathrm{cov}(Y, Z) \\ \mathrm{cov}(Z, X) & \mathrm{cov}(Z, Y) & \sigma_Z^2 \end{bmatrix}, \tag{7.4}$$

which is called the covariance tensor.

In practical applications, some of which are discussed in Chap. 8, often two of the covariances (7.3) take zero values. In the cases when a covariance tensor takes such a simple form, its principal directions and principal components may be determined in the elementary manner described in Sect. 5.1 for a two-dimensional distribution. Note that if there appears $\mathrm{cov}(X, Y)$ only, the two others being equal to zero, the principal variances may be calculated with the use of formulas (5.11), because the third principal variance is known in such a particular case. In general cases, when all the covariances have non-zero values, determining the principal directions and components of the covariance tensor is more difficult. Since such a general case seldom happens in practical applications of the error calculus, we shall not give here the details of how to transform the covariance tensor in a general case. The corresponding procedures are described in detail in monographs dealing with the tensor calculus or with the theory of elasticity.

7.2 Continuous Three-dimensional Random Variables

The probability density $f(x, y, z)$ for the continuous three-dimensional random variables X, Y, Z may be interpreted as a distribution of a "specific measure" of the probability in the whole region of all three variables. The equality

$$W = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z)\, dxdydz = 1 \tag{7.5}$$

must be satisfied. It means that the whole "mass" of the probability must be equal to unity. The first-order moments of random variables X, Y, Z with respect to the planes of the rectilinear coordinate system x, y, z, are defined as

$$m_{yz} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x, y, z)dxdydz, \quad \text{with respect to } x = 0 \text{ plane,}$$

$$m_{zx} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x, y, z)dxdydz, \quad \text{with respect to } y = 0 \text{ plane,} \tag{7.6}$$

$$m_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} zf(x, y, z)dxdydz, \quad \text{with respect to } z = 0 \text{ plane.}$$

Assume now such quantities $\bar{x}$, $\bar{y}$, and $\bar{z}$ which ensure that the equalities

$$W\bar{x} = m_{yz}, \qquad W\bar{y} = m_{zx}, \qquad W\bar{z} = m_{xy},$$

hold valid. Note that W is determined by relation (7.5).

Since $W = 1$ we can write

$$\bar{x} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x, y, z)\, dxdydz,$$

$$\bar{y} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x, y, z)\, dxdydz, \tag{7.7}$$

$$\bar{z} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} zf(x, y, z)\, dxdydz.$$

These quantities represent mean values of the distribution, comp. (5.41a), (5.41b). In other words, $\bar{x}$, $\bar{y}$, and $\bar{z}$ are the coordinates of the central point of the "mass" W of the probability distribution. The coordinate axes x_0, y_0, z_0 with their origin at that point are called the central coordinate axes.

The second-order moments with respect to the planes of the central coordinate system x_0, y_0, z_0 are

$$J_{yz} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \overline{x})^2 \, f(x, y, z) \, dx \, dy \, dz,$$

$$J_{zx} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \overline{y})^2 \, f(x, y, z) \, dx \, dy \, dz, \tag{7.8}$$

$$J_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (z - \overline{z})^2 \, f(x, y, z) \, dx \, dy \, dz,$$

for $x_0 = 0$, $y_0 = 0$, $z_0 = 0$ planes, respectively. They are called the central second-order moments.

Let us assume three quantities σ_X^2, σ_Y^2, and σ_Z^2, such that the following equalities are valid:

$$W\sigma_X^2 = J_{yz}, \qquad W\sigma_Y^2 = J_{zx}, \qquad W\sigma_Z^2 = J_{xy}.$$

Remembering that $W = 1$ (comp. (7.5)) we can write

$$\sigma_X^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \overline{x})^2 \, f(x, y, z) \, dx \, dy \, dz, \tag{7.9a}$$

$$\sigma_Y^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \overline{y})^2 \, f(x, y, z) \, dx \, dy \, dz, \tag{7.9b}$$

$$\sigma_Z^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (z - \overline{z})^2 \, f(x, y, z) \, dx \, dy \, dz, \tag{7.9c}$$

where σ_X^2, σ_Y^2, and σ_Z^2 are the variances of the corresponding variables. Similarly to the two-dimensional cases, square roots of the variances represent the standard deviations of the respective random variables X, Y, Z.

Let us note that moments J_{yz}, J_{zx}, and J_{xy} determined by relations (7.8) may be treated, to a certain degree, as the "inertia moments" of the "probability mass" W with respect to the planes of central coordinate system x_0, y_0, z_0. According to such an interpretation, the standard deviations σ_X, σ_Y, and σ_Z may be treated as the so-called "inertia radii" of the whole mass W.

In the formulas given above for the mean values and for the variances, the integration has been carried out using the probability density of three-dimensional distribution of random variables. However, while calculating these quantities we can, similarly to two-dimensional problems, make use of marginal probability densities, comp. (5.40). Knowing the analytical form of the probability density function we can perform the integration in an arbitrary order. However, when the probability density function has been determined as a result of measurements (empirical probability density), it may be easier to find the marginal probability density than the complete three-dimensional probability density. Such a possibility of calculating the moments of three-dimensional (or more generally, of a multi-dimensional) distribution should be always taken into account.

A significant role in the analysis of continuous three-dimensional distributions of errors is played by the covariances

$$\text{cov}\,(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \overline{x})\,(y - \overline{y})\, f\,(x, y, z)\, dx dy dz, \qquad (7.10a)$$

$$\text{cov}\,(Y, Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \overline{y})\,(z - \overline{z})\, f\,(x, y, z)\, dx dy dz, \qquad (7.10b)$$

$$\text{cov}\,(Z, X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (z - \overline{z})\,(x - \overline{x})\, f\,(x, y, z)\, dx dy dz. \qquad (7.10c)$$

Note that covariances (7.10a)–(7.10c) may be treated as mixed inertia moments of the "probability mass" with respect to the planes of the central coordinate system x_0, y_0, z_0.

Similarly to the case discussed in Sect. 5.1, variances (7.9a)–(7.9c) and covariances (7.10a)–(7.10c) are the components of a certain symmetrical tensor

$$T_{\sigma^2} = \begin{bmatrix} \sigma_X^2 & \text{cov}\,(X, Y) & \text{cov}\,(X, Z) \\ \text{cov}\,(Y, X) & \sigma_Y^2 & \text{cov}\,(Y, Z) \\ \text{cov}\,(Z, X) & \text{cov}\,(Z, Y) & \sigma_Z^2 \end{bmatrix}. \qquad (7.11)$$

In practical applications often one of the three random variables is independent of the two others. In this case let it be variable Z. It means that in this case the three-dimensional probability density may be presented as a product of a two-dimensional probability density of variables X and Y, and of a one-dimensional density of variable Z. Thus, we can write

$$f(x, y, x) = f_{XY}(x, y) f_Z(z).$$

Hence, two of the covariances appearing in (7.10a)–(7.10c) have zero values,

$$\text{cov}(X, Z) = \text{cov}(Y, Z) = 0.$$

If all three random variables X, Y, Z are independent, their joint probability density function shall be expressed as a product of one-dimensional densities

$$f(x, y, x) = f_X(x) f_Y(y) f_Z(z). \qquad (7.12)$$

All three covariances have then zero values

$$\text{cov}(X, Z) = \text{cov}(Y, Z) = \text{cov}(X, Y) = 0.$$

The covariance tensor (7.11) becomes a diagonal tensor with variances of one-dimensional distributions located along its diagonal.

7.3 The Three-dimensional Normal Distribution

7.3.1 Independent Random Variables

Let us assume that each of the three independent random variables X, Y, Z has the normal distribution determined by the probability density function, comp. (2.35),

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left[-\frac{(x-\bar{x})^2}{2\sigma_X^2}\right], \tag{7.13a}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_Y^2}} \exp\left[-\frac{(y-\bar{y})^2}{2\sigma_Y^2}\right], \tag{7.13b}$$

$$f_Z(z) = \frac{1}{\sqrt{2\pi\sigma_Z^2}} \exp\left[-\frac{(z-\bar{z})^2}{2\sigma_Z^2}\right]. \tag{7.13c}$$

According to (7.12) we obtain for such independent variables their three-dimensional joint probability density function $f(x, y, z)$ in the following form:

$$f(x, y, z) = \frac{1}{\sqrt{(2\pi)^3}\sigma_X\sigma_Y\sigma_Z}$$
$$\times \exp\left[-\frac{1}{2}\left[\left(\frac{x-\bar{x}}{\sigma_X}\right)^2 + \left(\frac{y-\bar{y}}{\sigma_Y}\right)^2 + \left(\frac{z-\bar{z}}{\sigma_Z}\right)^2\right]\right]. \tag{7.14}$$

Assuming a system of central coordinates, that is, a system in which mean (expected) values of all three variables X, Y, Z are equal to zero, we can write the probability density (7.14) in a simpler form:

$$f(x, y, z) = \frac{1}{\sqrt{(2\pi)^3}\sigma_X\sigma_Y\sigma_Z} \exp\left[-\frac{1}{2}\left[\left(\frac{x}{\sigma_X}\right)^2 + \left(\frac{y}{\sigma_Y}\right)^2 + \left(\frac{z}{\sigma_Z}\right)^2\right]\right]. \tag{7.15}$$

The ellipsoids determined in an arbitrary coordinate system by the equation

$$\left(\frac{x-\bar{x}}{\sigma_X}\right)^2 + \left(\frac{y-\bar{y}}{\sigma_Y}\right)^2 + \left(\frac{z-\bar{z}}{\sigma_Z}\right)^2 = \text{const.}, \tag{7.16a}$$

or in the central coordinate system by a simpler equation

$$\left(\frac{x}{\sigma_X}\right)^2 + \left(\frac{y}{\sigma_Y}\right)^2 + \left(\frac{z}{\sigma_Z}\right)^2 = \text{const.}, \tag{7.16b}$$

are the surfaces on which the probability density has a constant value. They are also known as the ellipsoids of probability concentration. Examples of practical applications of such ellipsoids will be given in Chap. 8.

7.3.2 The Spherical Normal Distribution

The spherical normal distribution may be treated as a special case of the three-dimensional normal distribution when

$$\sigma_X = \sigma_Y = \sigma_Z = \sigma. \tag{7.17}$$

Let us assume an arbitrarily oriented system of central coordinate axes x, y, z. The probability density of the spherical normal distribution may be expressed as a certain function of the radius r of a sphere on which the point x^*, y^*, z^* in question is located, and of the angles θ and φ of the spherical coordinate system. Accounting for equalities (7.17) we obtain this function by replacing Cartesian coordinate system x, y, z with a spherical coordinate system r, θ, φ. Finally, the probability density function for the spherical normal distribution takes the form [5]

$$f(r, \theta, \varphi) = \frac{r^2}{\sqrt{(2\pi)^3 \sigma^3}} \exp\left[-\frac{r^2}{2\sigma^2}\right]. \tag{7.18a}$$

This function represents the probability density at each point on the surface of a sphere of radius r. Similarly to the case of the two-dimensional circular normal distribution (see Sect. 5.5.2), the three-dimensional spherical normal distribution may be also interpreted as a one-dimensional distribution depending on the radius r only (the value of the probability density function $f(r, \theta, \varphi)$ remains constant if r is unchanged but the angles θ, φ vary over their domains). Such a one-dimensional distribution determines the probability of an event that a point is located on the sphere of radius r. Hence, the probability density will now be for a given radius 4π times larger than in the previous definition (7.18a). Thus, we can write finally

$$f_R(r) = \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(r, \theta, \varphi)\, d\varphi d\theta = \frac{2r^2}{\sqrt{2\pi}\,\sigma^3} \exp\left[-\frac{r^2}{2\sigma^2}\right]. \tag{7.18b}$$

Surfaces of spheres of a radius $r = $ const., over which the probability density has a constant value (it does not depend on the two angular coordinates θ, φ), are referred to as the spheres of probability concentration. The level of that concentration increases toward the central point, as in the case of the two-dimensional circular distribution, comp. computer simulation shown in Fig. 5.14. The probability that a three-dimensional random variable X, Y, Z will lie inside the sphere of a specific radius r^*, in other words that variable R will have a value within the range $(0, r^*)$, is determined by the integral

$$P\left(0 \leq R \leq r^*\right) = \int_0^{r^*} f(r)\, dr. \tag{7.19}$$

These probabilities have been tabulated in the $\chi^2(3)$ distribution (the chi-squared distribution with three degrees of freedom), see [1].

Table 7.1 The quantiles of
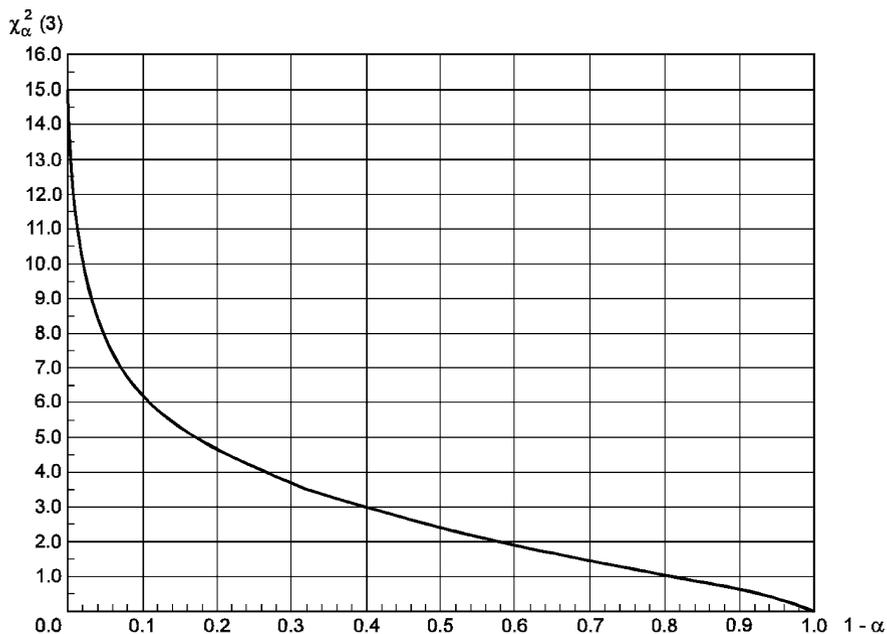the $\chi^2(3)$ distribution

$1 - \alpha$	$\chi_\alpha^2(3)$	$1 - \alpha$	$\chi_\alpha^2(3)$
0.99	0.115	0.30	3.665
0.98	0.185	0.20	4.642
0.95	0.352	0.10	6.251
0.90	0.584	0.05	7.815
0.80	1.005	0.02	9.837
0.70	1.424	0.01	11.341
0.50	2.366	0.001	16.268

Similarly to the case of two-dimensional circular distribution, there is a direct
interpretation of the $\chi^2(3)$ distribution. Given a radius r_α of such a sphere, let us
determine that the probability of the variable R to be located inside this sphere has
a prescribed value α. Note that the expression for the square of that radius,

$$(r_\alpha)^2 = x^2 + y^2 + z^2 \equiv \chi_\alpha^2(3) \tag{7.20}$$

represents, according to definition (5.56), the quantile $\chi_\alpha^2(3)$ of the order α of the
chi-squared distribution with three degrees of freedom, as indicated by the number
3 in parentheses. The values of $\chi_\alpha^2(3)$ are given in Table 7.1.

In practical applications the diagram shown in Fig. 7.1 may be useful.

Fig. 7.1 The quantiles of the $\chi^2(3)$ distribution

7.3.3 The Case of Dependent Random Variables

If random variables X_1, X_2, X_3 with the normal distributions are dependent, their joint probability density function is also normal. It contains in its exponent not only the squares of variables x_1, x_2, x_3 but also their products. Generally, if axes x_1, x_2, x_3 are central axes, the probability density function takes the form

$$f(x_1, x_2, x_3) = \frac{1}{\sqrt{(2\pi)^3 \det[k_{ij}]}} \exp\left[-\frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3} K_{ij}x_i x_j\right], \qquad (7.21a)$$

or, after expansion of sums,

$$f(x_1, x_2, x_3) = \frac{1}{\sqrt{(2\pi)^3 \det[k_{ij}]}}$$

$$\times \exp\left[-\frac{1}{2}(K_{11}x_1^2 + K_{22}x_2^2 + K_{33}x_3^2 + 2K_{12}x_1x_2 + 2K_{23}x_2x_3\right.$$

$$\left. + 2K_{13}x_1x_3)\right]. \qquad (7.21b)$$

Introducing new notation X, Y, Z for the three random variables, we can write

$$x_1 = x, \qquad x_2 = y, \qquad x_3 = z,$$

$$K_{11} = K_{XX}, \qquad K_{22} = K_{YY}, \qquad K_{33} = K_{ZZ},$$

$$K_{12} = K_{21} = K_{XY} = K_{YX}, \qquad (7.22)$$

$$K_{23} = K_{32} = K_{YZ} = K_{ZY},$$

$$K_{13} = K_{31} = K_{XZ} = K_{ZX}.$$

Coefficients K_{ij} are, according to (5.60), the elements of the matrix inverse to the matrix of second-order moments

$$[k_{ij}] = \begin{bmatrix} k_{XX} & k_{XY} & k_{XZ} \\ k_{YX} & k_{YY} & k_{YZ} \\ k_{ZX} & k_{ZY} & k_{ZZ} \end{bmatrix}, \qquad (7.23)$$

where

$k_{XX} = \sigma_X^2$ is the variance of random variable X;

$k_{YY} = \sigma_Y^2$ is the variance of random variable Y;

$k_{ZZ} = \sigma_Z^2$ is the variance of random variable Z;

$k_{XY} = k_{YX} = \text{cov}\{X, Y\}$ is the covariance of random variables X and Y;

$k_{YZ} = k_{ZY} = \text{cov}\{Y, Z\}$ is the covariance of random variables Y and Z;

$k_{ZX} = k_{XZ} = \text{cov}\{X, Z\}$ is the covariance of random variables X and Z.

Knowing the elements of matrix $[k_{ij}]$ one can calculate coefficients $[K_{ij}]$ appearing in the probability density function (7.21b) by solving, for example, the system of linear equations resulting from the definition of the inverse matrix: the product of an original matrix and an inverse matrix is equal to the unit matrix. Hence, such a system of linear equations may be written in the matrix form

$$\begin{bmatrix} k_{XX} & k_{XY} & k_{XZ} \\ k_{YX} & k_{YY} & k_{YZ} \\ k_{ZX} & k_{ZY} & k_{ZZ} \end{bmatrix} \begin{bmatrix} K_{XX} & K_{XY} & K_{XZ} \\ K_{YX} & K_{YY} & K_{YZ} \\ K_{ZX} & K_{ZY} & K_{ZZ} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad (7.24)$$

In general, the expressions for coefficients K_{ij} are rather long and complex. Their usefulness in practical calculations is limited. In most practical problems some of the elements of matrix $[k_{ij}]$ are equal to zero and the solution of the system of equations resulting from (7.24) becomes significantly simplified. Examples of such calculations will be given in Chap. 8.

On the surfaces of ellipsoids determined by the equation

$$K_{XX}x^2 + K_{YY}y^2 + K_{ZZ}z^2 + 2K_{XY}xy + 2K_{YZ}yz + 2K_{ZX}zx = \lambda^2 = \text{const.}, \tag{7.25}$$

the probability density (7.21b) has a constant value, comp. (7.16b). Such ellipsoids are called ellipsoids of probability concentration. Examples of practical applications will be given in Chap. 8. If distribution (7.21b) were presented according to (5.55) in the form of a distribution of normalized random variables

$$U_x = \frac{X}{\sigma_X}, \qquad U_y = \frac{Y}{\sigma_Y}, \qquad U_z = \frac{Z}{\sigma_Z},$$

it would become a spherical distribution with the probability density function (7.18a), and the ellipsoid (7.25) would be transformed into a sphere. The procedure of determining the radius r_α of this sphere for an assumed probability level α such that a point determined by particular values of the variables will lie inside it, has been discussed in Sect. 7.3.2. Such a radius is determined by formula (7.20). Returning to our original distribution (7.21b) and the ellipsoid (7.25), we do not change this probability. Thus, we can write

$$\lambda^2 = \chi_\alpha^2(3). \tag{7.26}$$

Remark 7.1 Calculation of the inverse matrix $[K_{ij}]$ required to construct the probability density function of normal distribution (two-dimensional, three-dimensional and also for more dimensions, see Chap. 9) can be done according to the following general method of calculating inverse matrices [2].

Let $\mathbf{A} = [A_{ij}]$ be a square matrix of dimension $n \times n$, non-singular, that is such that its determinant satisfies $\det \mathbf{A} \neq 0$. The algebraic complement (or signed minor) of the element A_{ij} of the matrix $\mathbf{A}$ (for $i, j = 1, 2, \ldots, n$) is the number a_{ij} calculated as

$$a_{ij} = (-1)^{i+j} D_{ij},$$

where D_{ij} is the minor of the element A_{ij}, that is the determinant of the square matrix $(n - 1) \times (n - 1)$ created from the matrix $\mathbf{A} = [A_{ij}]$ by removing its i-th row and j-th column (they are the row and the column of the matrix $\mathbf{A}$ that intersect at the element A_{ij}). Now we can generate a new matrix of dimension $n \times n$, which we denote as $\mathbf{A}^D$. This matrix is called the adjacent matrix of the matrix $\mathbf{A}$ and its elements are the algebraic complements of the corresponding elements of the matrix $\mathbf{A}$ written down in a transposed order, that is:

$$\mathbf{A}^D = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}.$$

The inverse matrix $\mathbf{A}^{-1}$ is composed of the elements of the adjacent matrix $\mathbf{A}^D$ after dividing each element of this matrix by the determinant of the matrix $\mathbf{A}$, that is:

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \mathbf{A}^D.$$

Problem 7.1 We are given the covariance matrix of the three-dimensional normal random variable (X, Y, Z) and the vector of its mean values:

$$[k_{ij}] = \begin{bmatrix} 6 & 3 & 0 \\ 3 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}, \qquad \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}.$$

Find the probability density function of the distribution of (X, Y, Z).

Problem 7.2 We are given the covariance matrix of the three-dimensional normal random variable (X, Y, Z) and the vector of its mean values:

$$[k_{ij}] = \begin{bmatrix} 11 & 4 & 2 \\ 4 & 10 & 6 \\ 2 & 6 & 9 \end{bmatrix}, \qquad \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} -3 \\ -2 \\ 1 \end{bmatrix}.$$

Find the probability density function of the distribution of (X, Y, Z).

Problem 7.3 We are given the covariance matrix of the four-dimensional normal random variable (X, Y, Z, U) and the vector of its mean values:

$$[k_{ij}] = \begin{bmatrix} 15 & 3 & 1 & 0 \\ 3 & 16 & 6 & -2 \\ 1 & 6 & 4 & 1 \\ 0 & -2 & 1 & 3 \end{bmatrix}, \qquad \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \\ \bar{u} \end{bmatrix} = \begin{bmatrix} -5 \\ 2 \\ 0 \\ 4 \end{bmatrix}.$$

Find the probability density function of the distribution of (X, Y, Z, U).

Problem 7.4 A point in the three-dimensional space (X, Y, Z) is located in such a way, that its coordinates are independent and they have the normal distributions with the mean values and the variances, respectively: $m_X = 0$, $\sigma_X^2 = 1$, $m_Y = 1$, $\sigma_X^2 = 4$, $m_Z = 0$, $\sigma_Z^2 = 9$. Find the probability that the point will be located within the walls of two rectangular prisms having their edges parallel to the axes of the coordinate system and bounded by the planes: $x = 0$, $x = 2$, $y = 0$, $y = 2$, $z = 0$, $z = 2$ and $x = 0.5$, $x = 1$, $y = 0.5$, $y = 1$, $z = 0.5$, $z = 1$.

Problem 7.5 The random point (X, Y, Z) has the normal distribution with the probability density function

$$f_{XYZ}(x, y, z) = \frac{\sqrt{3}}{16\pi\sqrt{\pi}} \exp\left\{-\frac{1}{8}\left[2x^2 + 4(y-1)^2\right.\right.$$

$$\left.\left. - 2(y-1)(z+9) + (z+9)^2\right]\right\}.$$

Find the set of points in which the probability density function takes the value 0.02.

Problem 7.6 The hit points of artillery fire are concentrated around the point $(x_0, y_0, z_0) = (0, 0, 3000)$ and have the normal distribution in x, y and z directions, independent one from the others. The variance of the distribution in the x-direction is 100, in the y-direction it is 100 and in the z-direction it is 500. Find a set in the space, outside of which the probability of the hit point is less than 0.001.

References

1. Abramowitz, M., Stegun, I.: Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. Dover, New York (1965)
2. Bellman, R.: Introduction to Matrix Analysis. McGraw–Hill, New York (1970)
3. Bryc, W.: The Normal Distribution. Characterization with Applications. Lecture Notes in Statistics, vol. 100. Springer, New York (1995)
4. Fisz, M.: Probability Theory and Mathematical Statistics, 3rd edn. Krieger, Melbourne (1980)
5. Korn, G.A., Korn, Th.M.: Mathematical Handbook for Scientists and Engineers, 2nd edn. Dover, New York (2000)

6. Kotulski, Z., Szczepiński, W.: On two methods of determining the ellipses and ellipsoids of the positioning accuracy of robot manipulators. Ing. Arch. (Arch. Appl. Mech.) **63**(1), 25–41 (1993)
7. Papoulis, A.: Probability, Random Variables, and Stochastic Processes with Errata Sheet, 4th edn. McGraw–Hill, New York (2002)
8. Szczepiński, W., Kotulski, Z.: Error Analysis with Applications in Engineering. Lastran, Rochester (2000)

Chapter 8
Three-dimensional Functions of Independent Random Variables

8.1 Basic Relations

As a typical example of practical application of three-dimensional functions of independent random variables, the analysis of the positioning accuracy of a certain class of mechanisms, especially robot manipulators, may be mentioned. An example of such a manipulator with four revolving joints is shown in Fig. 8.1. In this mechanical system, axis 22 is perpendicular to axis 11. Axes 22 and 33 are parallel one to the other. Nominal positions of joints are determined by the nominal values of positioning angles θ_1, θ_2, and θ_3. The position of joint 44 has no influence on the position of the manipulator's hand. It is used for turning the hand around the axis of the manipulator's end arm.

In the basic coordinate system X, Y, Z, the position of the hand is determined by three coordinates:

$$X = [l_1 \cos \theta_2 + l_2 \cos (\theta_2 + \theta_3)] \cos \theta_1,$$
$$Y = l_1 \sin \theta_2 + l_2 \sin (\theta_2 + \theta_3), \tag{8.1}$$
$$Z = [l_1 \cos \theta_2 + l_2 \cos (\theta_2 + \theta_3)] \sin \theta_1.$$

In each working cycle, the manipulator's hand should be positioned at a certain desired point O determined by its nominal coordinates X^0, Y^0, Z^0. In practice, however, real values of these coordinates are certain functions of the positioning angles $\theta_1 = \theta_1^0 + \Delta\theta_1$, $\theta_2 = \theta_2^0 + \Delta\theta_2$, and $\theta_3 = \theta_3^0 + \Delta\theta_3$, where θ_1^0, θ_2^0, and θ_3^0 are the desired nominal values of the positioning angles. In each repeated working cycle they are realized with a certain small random error $\Delta\theta_1$, $\Delta\theta_2$, and $\Delta\theta_3$, respectively.

In the analysis of influence of these errors on the resultant positioning error of the hand we shall use a local coordinate system x, y, z with the origin at the point of the nominal hand's position. The positioning error of the hand may be represented as a vector with three components:

$$U = X - X^0, \qquad V = Y - Y^0, \qquad W = Z - Z^0, \tag{8.2}$$

Fig. 8.1 An example of a manipulator with four revolving joints
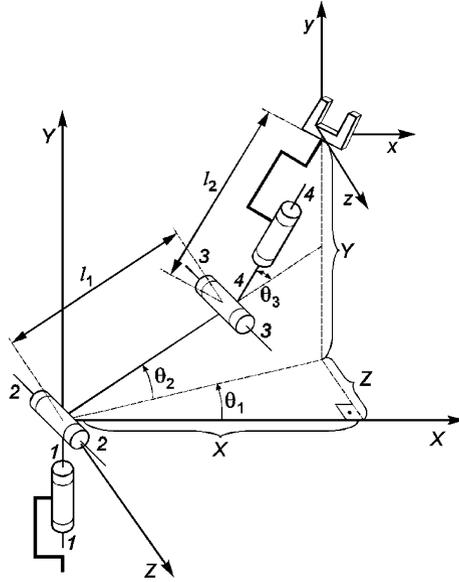
where X, Y, Z are the actual coordinates of the hand's position, and X_0, Y_0, Z_0 stand for the coordinates of the nominal position.

Generally, in such problems we have to deal with a random variable in the form of a vector of deviation from the average position. Components U, V, W of this vector are certain functions of independent random variables $Q_1, Q_2, \ldots, Q_n$,

$$
\begin{aligned}
U &= g\,(Q_1, Q_2, \ldots, Q_n)\,, \\
V &= h\,(Q_1, Q_2, \ldots, Q_n)\,, \\
W &= k\,(Q_1, Q_2, \ldots, Q_n)\,.
\end{aligned}
\tag{8.3}
$$

In the particular case discussed above, the angles θ_1, θ_2, and θ_3 stand for these independent random variables, while the errors in the hand's position (8.2) are the components of this vector.

By analogy with the two-dimensional functions (6.4), we can write the expressions for the variances of the three resultant random variables U, V, and W:

$$
\begin{aligned}
\sigma_U^2 &= \left(\frac{\partial g}{\partial q_1}\right)^2 \sigma_{Q_1}^2 + \left(\frac{\partial g}{\partial q_2}\right)^2 \sigma_{Q_2}^2 + \cdots + \left(\frac{\partial g}{\partial q_n}\right)^2 \sigma_{Q_n}^2\,, \\
\sigma_V^2 &= \left(\frac{\partial h}{\partial q_1}\right)^2 \sigma_{Q_1}^2 + \left(\frac{\partial h}{\partial q_2}\right)^2 \sigma_{Q_2}^2 + \cdots + \left(\frac{\partial h}{\partial q_n}\right)^2 \sigma_{Q_n}^2\,, \\
\sigma_W^2 &= \left(\frac{\partial k}{\partial q_1}\right)^2 \sigma_{Q_1}^2 + \left(\frac{\partial k}{\partial q_2}\right)^2 \sigma_{Q_2}^2 + \cdots + \left(\frac{\partial k}{\partial q_n}\right)^2 \sigma_{Q_n}^2\,,
\end{aligned}
\tag{8.4}
$$

or in the compact form

$$\sigma_U^2 = \sum_{j=1}^{n} \left(\frac{\partial g}{\partial q_j} \right)^2 \sigma_{Q_j}^2,$$

$$\sigma_V^2 = \sum_{j=1}^{n} \left(\frac{\partial h}{\partial q_j} \right)^2 \sigma_{Q_j}^2, \tag{8.5}$$

$$\sigma_W^2 = \sum_{j=1}^{n} \left(\frac{\partial k}{\partial q_j} \right)^2 \sigma_{Q_j}^2.$$

Similarly to the case of two-dimensional functions of independent random variables (Chap. 6), these expressions may be obtained as a result of more general considerations. The total error of the three functions (8.3), caused by deviations Δq_j of independent variables Q_j from their nominal values, is represented by a vector **p** with components

$$\Delta u = U - u^0, \qquad \Delta v = V - v^0, \qquad \Delta w = W - w^0, \tag{8.6}$$

where u^0, v^0, and w^0 are the nominal values of the three functions, while U, V, and W are their actual values.

If the values Δq_j of small deviations of independent variables are known, then the components of the error vector (8.6) may be calculated from the linearized relations

$$\Delta u = \frac{\partial g}{\partial q_1} \Delta q_1 + \frac{\partial g}{\partial q_2} \Delta q_2 + \cdots + \frac{\partial g}{\partial q_n} \Delta q_n,$$

$$\Delta v = \frac{\partial h}{\partial q_1} \Delta q_1 + \frac{\partial h}{\partial q_2} \Delta q_2 + \cdots + \frac{\partial h}{\partial q_n} \Delta q_n, \tag{8.7a}$$

$$\Delta w = \frac{\partial k}{\partial q_1} \Delta q_1 + \frac{\partial k}{\partial q_2} \Delta q_2 + \cdots + \frac{\partial k}{\partial q_n} \Delta q_n.$$

Expressions (8.7a) may be written in a matrix form

$$\begin{bmatrix} \Delta u \\ \Delta v \\ \Delta w \end{bmatrix} = \begin{bmatrix} \frac{\partial g}{\partial q_1} & \frac{\partial g}{\partial q_2} & \cdots & \frac{\partial g}{\partial q_n} \\ \frac{\partial h}{\partial q_1} & \frac{\partial h}{\partial q_2} & \cdots & \frac{\partial h}{\partial q_n} \\ \frac{\partial k}{\partial q_1} & \frac{\partial k}{\partial q_2} & \cdots & \frac{\partial k}{\partial q_n} \end{bmatrix} \begin{bmatrix} \Delta q_1 \\ \Delta q_2 \\ \vdots \\ \Delta q_n \end{bmatrix}, \tag{8.7b}$$

or

$$\mathbf{p} = \mathbf{A} \Delta \mathbf{q}, \tag{8.7c}$$

where matrix

$$\mathbf{A} = \begin{bmatrix} \frac{\partial g}{\partial q_1} & \frac{\partial g}{\partial q_2} & \cdots & \frac{\partial g}{\partial q_n} \\ \frac{\partial h}{\partial q_1} & \frac{\partial h}{\partial q_2} & \cdots & \frac{\partial h}{\partial q_n} \\ \frac{\partial k}{\partial q_1} & \frac{\partial k}{\partial q_2} & \cdots & \frac{\partial k}{\partial q_n} \end{bmatrix} \tag{8.8}$$

may be treated as the matrix of coefficients of sensitivity of the resultant error $\mathbf{p}$ to the errors Δq_j of independent variables. In expression (8.7c) symbols $\mathbf{p}$ and $\Delta\mathbf{q}$ represent the column matrices

$$\mathbf{p} = \begin{bmatrix} \Delta u \\ \Delta v \\ \Delta w \end{bmatrix}, \qquad \Delta\mathbf{q} = \begin{bmatrix} \Delta q_1 \\ \Delta q_2 \\ \vdots \\ \Delta q_n \end{bmatrix}. \tag{8.9}$$

Let us assume now that errors Δq_j of independent variables Q_j with respect to their nominal values have a certain given distribution and that they are statistically independent. Moments of the second order of the resultant error $\mathbf{p}$ may be represented in the form of a matrix (comp. (7.23), (6.10)) and its expansion in the form of a product of three matrices

$$\left[k_{ij}\right] = \begin{bmatrix} k_{UU} & k_{UV} & k_{UW} \\ k_{VU} & k_{VV} & k_{VW} \\ k_{WU} & k_{WV} & k_{WW} \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} \frac{\partial g}{\partial q_1} & \frac{\partial g}{\partial q_2} & \cdots & \frac{\partial g}{\partial q_n} \\ \frac{\partial h}{\partial q_1} & \frac{\partial h}{\partial q_2} & \cdots & \frac{\partial h}{\partial q_n} \\ \frac{\partial k}{\partial q_1} & \frac{\partial k}{\partial q_2} & \cdots & \frac{\partial k}{\partial q_n} \end{bmatrix}$$

$$\times \begin{bmatrix} \sigma_{Q_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{Q_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{Q_n}^2 \end{bmatrix} \begin{bmatrix} \frac{\partial g}{\partial q_1} & \frac{\partial h}{\partial q_1} & \frac{\partial k}{\partial q_1} \\ \frac{\partial g}{\partial q_2} & \frac{\partial h}{\partial q_2} & \frac{\partial k}{\partial q_2} \\ \vdots & \vdots & \vdots \\ \frac{\partial g}{\partial q_n} & \frac{\partial h}{\partial q_n} & \frac{\partial k}{\partial q_n} \end{bmatrix}. \tag{8.10}$$

The elements of matrix $[k_{ij}]$ resulting from the operation given just above are

$$k_{UU} = \left(\frac{\partial g}{\partial q_1}\right)^2 \sigma_{Q_1}^2 + \left(\frac{\partial g}{\partial q_2}\right)^2 \sigma_{Q_2}^2 + \cdots + \left(\frac{\partial g}{\partial q_n}\right)^2 \sigma_{Q_n}^2,$$

$$k_{VV} = \left(\frac{\partial h}{\partial q_1}\right)^2 \sigma_{Q_1}^2 + \left(\frac{\partial h}{\partial q_2}\right)^2 \sigma_{Q_2}^2 + \cdots + \left(\frac{\partial h}{\partial q_n}\right)^2 \sigma_{Q_n}^2,$$

$$k_{WW} = \left(\frac{\partial k}{\partial q_1}\right)^2 \sigma_{Q_1}^2 + \left(\frac{\partial k}{\partial q_2}\right)^2 \sigma_{Q_2}^2 + \cdots + \left(\frac{\partial k}{\partial q_n}\right)^2 \sigma_{Q_n}^2,$$

$$k_{UV} = \frac{\partial g}{\partial q_1}\frac{\partial h}{\partial q_1}\sigma_{Q_1}^2 + \frac{\partial g}{\partial q_2}\frac{\partial h}{\partial q_2}\sigma_{Q_2}^2 + \cdots + \frac{\partial g}{\partial q_n}\frac{\partial h}{\partial q_n}\sigma_{Q_n}^2,$$

$$\tag{8.11}$$

$$k_{VW} = \frac{\partial h}{\partial q_1}\frac{\partial k}{\partial q_1}\sigma_{Q_1}^2 + \frac{\partial h}{\partial q_2}\frac{\partial k}{\partial q_2}\sigma_{Q_2}^2 + \cdots + \frac{\partial h}{\partial q_n}\frac{\partial k}{\partial q_n}\sigma_{Q_n}^2,$$

$$k_{UW} = \frac{\partial g}{\partial q_1}\frac{\partial k}{\partial q_1}\sigma_{Q_1}^2 + \frac{\partial g}{\partial q_2}\frac{\partial k}{\partial q_2}\sigma_{Q_2}^2 + \cdots + \frac{\partial g}{\partial q_n}\frac{\partial k}{\partial q_n}\sigma_{Q_n}^2.$$

These elements stand for:

$k_{UU} = \sigma_U^2$ the variance of the component U of the error vector $\mathbf{p}$,

$k_{VV} = \sigma_V^2$ the variance of the component V of the error vector $\mathbf{p}$,

$k_{WW} = \sigma_W^2$ the variance of the component W of the error vector $\mathbf{p}$,

$k_{UV} = k_{VU} = \text{cov}(U, V)$ the covariance of the components U and V of the error vector $\mathbf{p}$ or, in other words, the correlation moment of U and V,

$k_{VW} = k_{WV} = \text{cov}(V, W)$ the covariance of the components V and W of the error vector $\mathbf{p}$ or, in other words, the correlation moment of V and W,

$k_{UW} = k_{WU} = \text{cov}(U, W)$ the covariance of the components U and W of the error vector $\mathbf{p}$ or, in other words, the correlation moment of U and W.

The matrix (8.10) is the matrix of a certain tensor. It has the same meaning as the matrix (7.11). Owing, however, to functional relations (8.3), its components may be calculated with the use of formulas (8.11). The distributions of errors of the three variables U, V, W are now associated with the distributions of the errors of independent variables Q_j.

8.2 The Rectangular Distribution of Independent Random Variables

Let us assume that each independent random variable Q_i appearing in three-dimensional functions (8.3), has the uniform (rectangular) distribution, comp. Sect. 3.4.2. In such a distribution the values of random errors are bounded by two marginal values a and b (Fig. 6.2a) or, alternatively, by the values Δq^+ and Δq^- (Fig. 6.2b), when the nominal value q^0 is taken as a reference point.

As in the analogous two-dimensional problems (Sect. 6.2), the rectangular distribution will be used here as an auxiliary distribution in solving in terms of the error calculus the so-called tolerance limits problems, comp. Sect. 4.4.

Let us analyze the movements of the end-point of the vector of the resultant error $\mathbf{p}$ with the components Δu, Δv, and Δw, when the errors Δq_r and Δq_s of two arbitrarily chosen independent variables, namely Q_r and Q_s, are changing, while errors of the remaining independent variables do not change (comp. [2, 3]). Assume that the errors q_r and q_s will be treated as changing parameters in equations (8.7a),

which can be written in a more convenient form

$$\Delta u = \frac{\partial g}{\partial q_1}\Delta q_1 + \cdots + \frac{\partial g}{\partial q_r}\Delta q_r + \cdots + \frac{\partial g}{\partial q_s}\Delta q_s + \cdots + \frac{\partial g}{\partial q_n}\Delta q_n,$$

$$\Delta v = \frac{\partial h}{\partial q_1}\Delta q_1 + \cdots + \frac{\partial h}{\partial q_r}\Delta q_r + \cdots + \frac{\partial h}{\partial q_s}\Delta q_s + \cdots + \frac{\partial h}{\partial q_n}\Delta q_n, \qquad (8.12)$$

$$\Delta w = \frac{\partial k}{\partial q_1}\Delta q_1 + \cdots + \frac{\partial k}{\partial q_r}\Delta q_r + \cdots + \frac{\partial k}{\partial q_s}\Delta q_s + \cdots + \frac{\partial k}{\partial q_n}\Delta q_n.$$

Let us now solve the first two of (8.12) with respect to Δq_r and Δq_s. Then we shall introduce expressions for Δq_r and Δq_s into the third equation (8.12). Finally, we obtain the equation of a family of parallel planes in the local rectangular coordinate system $\Delta u, \Delta v, \Delta w$. This equation may be written in the following form:

$$\det\begin{bmatrix} \frac{\partial h}{\partial q_r} & \frac{\partial h}{\partial q_s} \\ \frac{\partial k}{\partial q_r} & \frac{\partial k}{\partial q_s} \end{bmatrix}\Delta u + \det\begin{bmatrix} \frac{\partial k}{\partial q_r} & \frac{\partial k}{\partial q_s} \\ \frac{\partial g}{\partial q_r} & \frac{\partial g}{\partial q_s} \end{bmatrix}\Delta v + \det\begin{bmatrix} \frac{\partial g}{\partial q_r} & \frac{\partial g}{\partial q_s} \\ \frac{\partial h}{\partial q_r} & \frac{\partial h}{\partial q_s} \end{bmatrix}\Delta w$$

$$= \sum_{j=1}^{n}\left\{\det\begin{bmatrix} \frac{\partial g}{\partial q_r} & \frac{\partial g}{\partial q_s} & \frac{\partial g}{\partial q_j} \\ \frac{\partial h}{\partial q_r} & \frac{\partial h}{\partial q_s} & \frac{\partial h}{\partial q_j} \\ \frac{\partial k}{\partial q_r} & \frac{\partial k}{\partial q_s} & \frac{\partial k}{\partial q_j} \end{bmatrix}\Delta q_j\right\}. \qquad (8.13a)$$

The coefficients on the left-hand side of this equation and free terms on its right-hand side are represented in the form of a Jacobian.[1] Note that for $j = r$ and for $j = s$, the corresponding Jacobians on the right-hand side are equal to zero because two of their columns are identical.

Equation (8.13a) can be written down in a more compact form:

$$\det\begin{bmatrix} \frac{\partial g}{\partial q_r} & \frac{\partial g}{\partial q_s} & \Delta u \\ \frac{\partial h}{\partial q_r} & \frac{\partial h}{\partial q_s} & \Delta v \\ \frac{\partial k}{\partial q_r} & \frac{\partial k}{\partial q_s} & \Delta w \end{bmatrix} = \sum_{j=1}^{n}\left\{\det\begin{bmatrix} \frac{\partial g}{\partial q_r} & \frac{\partial g}{\partial q_s} & \frac{\partial g}{\partial q_j} \\ \frac{\partial h}{\partial q_r} & \frac{\partial h}{\partial q_s} & \frac{\partial h}{\partial q_j} \\ \frac{\partial k}{\partial q_r} & \frac{\partial k}{\partial q_s} & \frac{\partial k}{\partial q_j} \end{bmatrix}\Delta q_j\right\}. \qquad (8.13b)$$

Expanding the left-hand side of (8.13b) into sub-determinants, one obtains (8.13a).

Taking consecutively all the possible combinations of pairs of errors Δq_r and Δq_s of independent variables as changing parameters, we obtain equations of various families of parallel planes. The number of these families m depends on the number of independent variables n. It may be calculated with the use of the formula

$$m = \binom{n}{2} = \frac{n!}{2(n-1)!}, \qquad (8.14)$$

or taken directly from Table 8.1:

---

[1] det in (8.13a) denotes the determinant of the matrix.

Table 8.1 Number of the plane families versus number of independent variables

n	3	4	5	6	7	8
m	3	6	10	15	21	28

Note, however, that in practical applications of the theory, some families of the planes may have identical equations. This means that these families are overlapping and that the number of independent families is smaller than that resulting from formula (8.14).

The distance of any plane belonging to a particular family from the origin depends on the assumed values of errors Δq_j of independent variables q_j on the right-hand side of (8.13a). The two extreme positions of these planes are obtained by substituting appropriately the extreme values of errors Δq_j^+ and Δq_j^- in order to obtain at first the largest possible value of the sum on the right-hand side of (8.13a), and then its smallest possible value.

By determining the extreme positions of planes of each family we find a polyhedron bounded by these extreme pairs of planes. We shall call it the polyhedron of tolerance limits. All possible vectors of resultant error will lie inside the polyhedron. Note that inside the polyhedron, the distribution is uniform according to the theory presented above.

Example 8.1 Let us determine the polyhedron of the positioning accuracy of the hand of the manipulator shown schematically in Fig. 8.1.

Let the nominal position of the hand be determined by the positioning angles

$$\theta_1^0 = 0, \qquad \theta_2^0 = 0, \qquad \theta_3^0 = \frac{\pi}{2}. \tag{8.15}$$

In the basic coordinate system X, Y, Z, the position of the hand is given by expressions (8.1) for these values of the positioning angles $\theta_1, \theta_2, \theta_3$. In the following calculations we shall assume

$$l_1 = l_2 = l = 1000 \text{ mm}. \tag{8.16}$$

The particular values of partial derivatives of functions (8.1) with respect to independent variables $\theta_1, \theta_2, \theta_3$ for the assumed values (8.15) of positioning angles are given in Table 8.2.

Table 8.2 Values of the partial derivatives in Example 8.1

j	1	2	3
$\frac{\partial X}{\partial \theta_j}$	0	$-l$	$-l$
$\frac{\partial Y}{\partial \theta_j}$	0	l	0
$\frac{\partial Z}{\partial \theta_j}$	l	0	0

The polyhedron of the positioning accuracy will be bounded by extreme positions of three pairs of planes determined by (8.13a). To find equations of some particular families of planes we shall substitute into (8.13a) the appropriate values of partial derivatives for all possible combinations of pairs (r, s). For example, when $\Delta\theta_1$ and $\Delta\theta_2$ are changing, (8.13a) takes the particular form

$$
\det \begin{bmatrix} \frac{\partial Y}{\partial \theta_1} & \frac{\partial Y}{\partial \theta_2} \\ \frac{\partial Z}{\partial \theta_1} & \frac{\partial Z}{\partial \theta_2} \end{bmatrix} x + \det \begin{bmatrix} \frac{\partial Z}{\partial \theta_1} & \frac{\partial Z}{\partial \theta_2} \\ \frac{\partial X}{\partial \theta_1} & \frac{\partial X}{\partial \theta_2} \end{bmatrix} y + \det \begin{bmatrix} \frac{\partial X}{\partial \theta_1} & \frac{\partial X}{\partial \theta_2} \\ \frac{\partial Y}{\partial \theta_1} & \frac{\partial Y}{\partial \theta_2} \end{bmatrix} z
$$

$$
= \det \begin{bmatrix} \frac{\partial X}{\partial \theta_1} & \frac{\partial X}{\partial \theta_2} & \frac{\partial X}{\partial \theta_3} \\ \frac{\partial Y}{\partial \theta_1} & \frac{\partial Y}{\partial \theta_2} & \frac{\partial Y}{\partial \theta_3} \\ \frac{\partial Z}{\partial \theta_1} & \frac{\partial Z}{\partial \theta_2} & \frac{\partial Z}{\partial \theta_3} \end{bmatrix} \Delta\theta_3.
\tag{8.17}
$$

Substituting the values of derivatives given in Table 8.2, we obtain the equation of the first family of planes

$$
\det \begin{bmatrix} 0 & l \\ l & 0 \end{bmatrix} x + \det \begin{bmatrix} l & 0 \\ 0 & -l \end{bmatrix} y + \det \begin{bmatrix} 0 & -l \\ 0 & l \end{bmatrix} z = \det \begin{bmatrix} 0 & -l & -l \\ 0 & l & 0 \\ l & 0 & 0 \end{bmatrix} \Delta\theta_3.
\tag{8.18}
$$

Determining the dimensions of the polyhedron we shall assume the following tolerance limits of the positioning angles θ_1, θ_2, and θ_3:

$$
\Delta\theta_1 = \Delta\theta_2 = \Delta\theta_3 = \pm 0.001525 \text{ rad.}
\tag{8.19}
$$

Repeating the procedure described above for all three pairs of changing positioning errors $\Delta\theta_r$ and $\Delta\theta_s$, we arrive at the equations of extreme positions of planes forming the faces of the polyhedron. The equations of these planes are

$$
\begin{aligned}
x + y &= \pm 1.525 \text{ mm,} \\
y &= \pm 1.525 \text{ mm,} \\
z &= \pm 1.525 \text{ mm.}
\end{aligned}
\tag{8.20}
$$

The polyhedron is shown in Fig. 8.2.

Example 8.2 In Fig. 8.3 there is presented a scheme of a more complex manipulator with five revolving joints. The position of the manipulator's hand in the basic coordinate system X, Y, Z is determined by three coordinates:

$$
\begin{aligned}
X &= [a \cos\theta_2 + b \cos(\theta_2 + \theta_3) + c \cos(\theta_2 + \theta_3 + \theta_4)] \cos\theta_1, \\
Y &= [a \cos\theta_2 + b \cos(\theta_2 + \theta_3) + c \cos(\theta_2 + \theta_3 + \theta_4)] \sin\theta_1, \\
Z &= a \sin\theta_2 + b \sin(\theta_2 + \theta_3) + c \sin(\theta_2 + \theta_3 + \theta_4).
\end{aligned}
\tag{8.21}
$$

These coordinates are functions of four independent variables: positioning angles θ_1, θ_2, θ_3, and θ_4. General expressions for the partial derivatives of these functions with respect to the angles are rather lengthy. Thus, the solution in a general form

Fig. 8.2 The polyhedron of the positioning accuracy for the manipulator presented in Fig. 8.1—Example 8.1
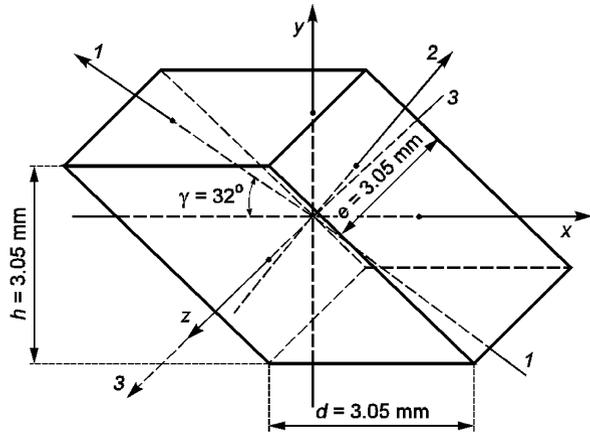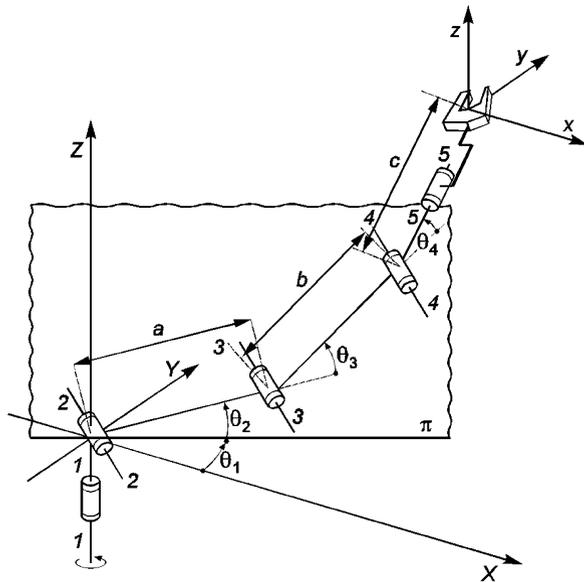

Fig. 8.3 The scheme of a manipulator with five revolving joints

is not useful for practical applications. It is more convenient to perform numerical calculations for the manipulator's position in question.

Assume the following data for the dimensions of the manipulator and for its position

$$a = 700 \text{ mm}, \qquad b = 600 \text{ mm}, \qquad c = 250 \text{ mm},$$
$$\theta_1 = 30°, \qquad \theta_2 = 30°, \qquad \theta_3 = 30°, \qquad \theta_4 = -60°. \tag{8.22}$$

We shall derive all equations for an arbitrary value of the angle θ_1.

Table 8.3 Values of the
partial derivatives in
Example 8.2

i	1	2	3	4
$\frac{\partial X}{\partial \theta_i}$	-578.1	$-869.6 \cos \theta_1$	$-519.6 \cos \theta_1$	0
$\frac{\partial Y}{\partial \theta_i}$	1001.3	$-869.6 \sin \theta_1$	$-519.6 \sin \theta_1$	0
$\frac{\partial Z}{\partial \theta_i}$	0	1156.2	550	-250

The bounding values of the errors (tolerance limits) in positioning the angles are
assumed to be

$$\Delta\theta_1 = \Delta\theta_2 = \Delta\theta_3 = \Delta\theta_4 = \pm 2' = \pm 0.00058 \text{ rad.} \tag{8.23}$$

For the data (8.22) the partial derivatives constituting the elements of the matrix
(8.8) take the values collated in Table 8.3, where dimensions are given in millime-
ters.

Below are derived the equations of particular families of planes by substituting
respective values of derivatives (sensitivity coefficients) into (8.13a) for all possible
combinations of pairs (r, s). We shall denote these planes by the symbol $P(r; s)$.
For example, $P(1; 2)$ denotes the family of planes obtained when $\Delta\theta_1$ and $\Delta\theta_2$ are
changing.

For the planes $P(2; 3)$, when positioning errors $\Delta\theta_2$ and $\Delta\theta_3$ are changing while
$\Delta\theta_1$ and $\Delta\theta_4$ have constant values, (8.13a) takes the following form:

$$\det \begin{bmatrix} \frac{\partial Y}{\partial \theta_2} & \frac{\partial Y}{\partial \theta_3} \\ \frac{\partial Z}{\partial \theta_2} & \frac{\partial Z}{\partial \theta_3} \end{bmatrix} x + \det \begin{bmatrix} \frac{\partial Z}{\partial \theta_2} & \frac{\partial Z}{\partial \theta_3} \\ \frac{\partial X}{\partial \theta_2} & \frac{\partial X}{\partial \theta_3} \end{bmatrix} y + \det \begin{bmatrix} \frac{\partial X}{\partial \theta_2} & \frac{\partial X}{\partial \theta_3} \\ \frac{\partial Y}{\partial \theta_2} & \frac{\partial Y}{\partial \theta_3} \end{bmatrix} z$$

$$= \det \begin{bmatrix} \frac{\partial X}{\partial \theta_2} & \frac{\partial X}{\partial \theta_3} & \frac{\partial X}{\partial \theta_1} \\ \frac{\partial Y}{\partial \theta_2} & \frac{\partial Y}{\partial \theta_3} & \frac{\partial Y}{\partial \theta_1} \\ \frac{\partial Z}{\partial \theta_2} & \frac{\partial Z}{\partial \theta_3} & \frac{\partial Z}{\partial \theta_1} \end{bmatrix} \Delta\theta_1 + \det \begin{bmatrix} \frac{\partial X}{\partial \theta_2} & \frac{\partial X}{\partial \theta_3} & \frac{\partial X}{\partial \theta_4} \\ \frac{\partial Y}{\partial \theta_2} & \frac{\partial Y}{\partial \theta_3} & \frac{\partial Y}{\partial \theta_4} \\ \frac{\partial Z}{\partial \theta_2} & \frac{\partial Z}{\partial \theta_3} & \frac{\partial Z}{\partial \theta_4} \end{bmatrix} \Delta\theta_4. \tag{8.24}$$

Substituting numerical values of the partial derivatives given in Table 8.3 we observe
that the third determinant on the left-hand side and the second determinant on the
right-hand side of (8.24) are equal to zero for any value of angle θ_1. Finally, the
equation of planes $P(2; 3)$ takes the form

$$x \sin \theta_1 - y \cos \theta_1 = -1156.2 \Delta\theta_1. \tag{8.25}$$

These planes are perpendicular to the reference plane $x-y$.

Repeating this procedure for all possible pairs of changing joint positioning er-
rors $\Delta\theta_r$ and $\Delta\theta_s$ and solving all determinants, we arrive at the equations of all
families of planes.

For example, for $P(2; 4)$ planes, when positioning errors $\Delta\theta_2$ and $\Delta\theta_4$ are chang-
ing, and for $P(3; 4)$ planes, when $\Delta\theta_3$ and $\Delta\theta_4$ are changing, we find that the cor-
responding equations are identical with the equation (8.25). Thus, the families of
planes $P(2; 3)$, $P(2; 4)$, and $P(3; 4)$ are overlapping.

In the case of $P(1;2)$ planes, when errors $\Delta\theta_1$ and $\Delta\theta_2$ are changing while errors $\Delta\theta_2$ and $\Delta\theta_4$ preserve constant values, their equation takes the form

$$x\cos\theta_1 + y\sin\theta_1 + 0.7521z = -105.93\Delta\theta_3 - 188.03\Delta\theta_4. \tag{8.26}$$

These planes are inclined with respect to all three axes of the local coordinate system x, y, z.

For the next $P(1;3)$ family of planes, when errors $\Delta\theta_1$ and $\Delta\theta_3$ are changing with the remaining joint positioning errors being kept constant, the equation has the form

$$x\cos\theta_1 + y\sin\theta_1 + 0.945z = 222.69\Delta\theta_2 - 236.18\Delta\theta_4. \tag{8.27}$$

These planes are also inclined with respect to all three coordinate axes x, y, z.

The last family of $P(1;4)$ planes, when errors $\Delta\theta_1$ and $\Delta\theta_4$ are changing, is determined by the equation

$$x\cos\theta_1 + y\sin\theta_1 = -869.6\Delta\theta_2 - 519.6\Delta\theta_3. \tag{8.28}$$

These planes are perpendicular to the reference plane $x-y$.

Thus, we have four independent families of planes determined by (8.25)–(8.28). The polyhedron of the positioning accuracy will be, therefore, bounded by four pairs of parallel planes only, instead of six pairs as predicted by formula (8.14). In practical applications such particular cases may often happen. Extreme positions of planes of these families will be obtained by substituting the tolerance limits (8.23) of joint positioning errors into the corresponding equations. The sign of these tolerance limits should be suitably taken in order to obtain first the largest possible value of the free term on the right-hand side of each equation, and then its smallest possible value. In this manner, we arrive at the equations of planes in their extreme positions. They constitute the faces of the polyhedron of the positioning accuracy when the value $\theta_1 = 30°$:

Faces $P(2;3)$—symbolic notation A

$$A_1\ y = 0.577x + 0.774,$$
$$A_2\ y = 0.577x - 0.774,$$

Faces $P(1;2)$—symbolic notation B

$$B_1\ 0.866x + 0.500y + 0.7521z = 0.1705,$$
$$B_2\ 0.866x + 0.500y + 0.7521z = -0.1705,$$

Faces $P(1;3)$—symbolic notation C

$$C_1\ 0.866x + 0.500y + 0.945z = 0.266,$$
$$C_2\ 0.866x + 0.500y + 0.945z = -0.266,$$
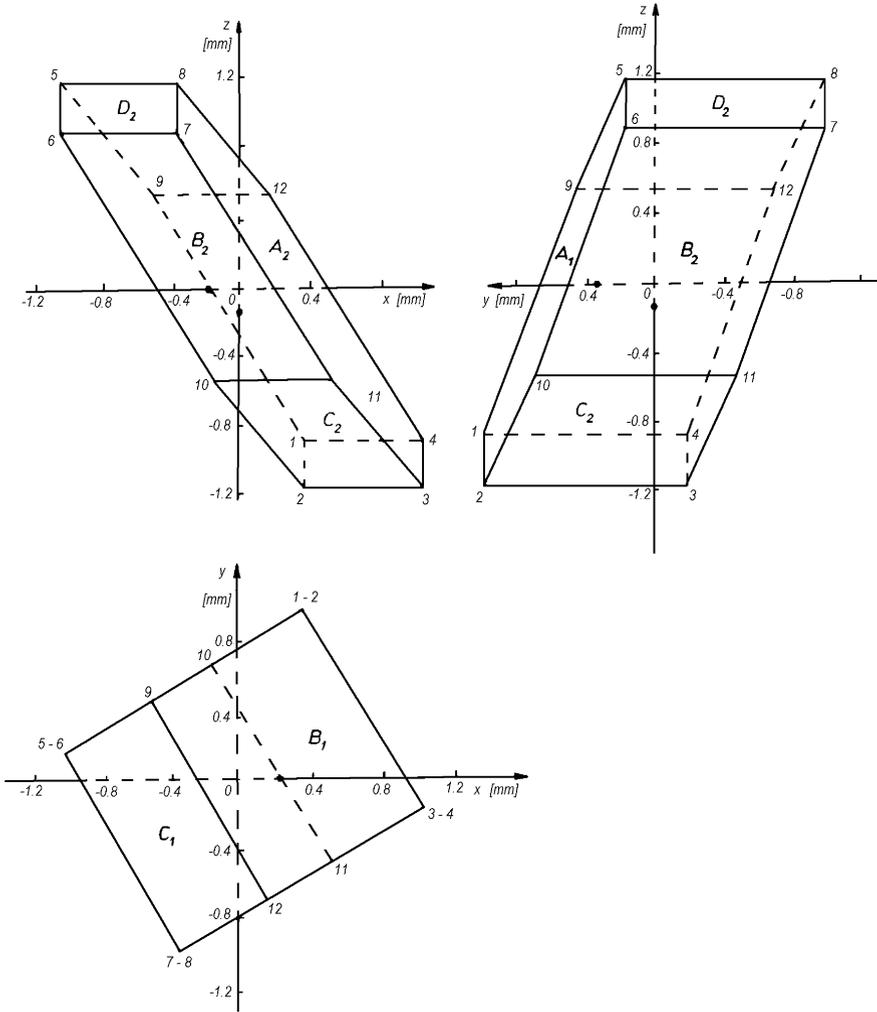
Fig. 8.4 Three projections of the polyhedron of the positioning accuracy for the manipulator presented in Fig. 8.3—Example 8.2

Faces $P(1; 4)$—symbolic notation D

$$D_1 \ y = -1.732x + 1.612,$$
$$D_2 \ y = -1.732x - 1.612.$$

In Fig. 8.4 are presented three projections of the polyhedron. Note that for each position of the manipulator the polyhedron has a different form, comp. [2].

8.3 The Normal Distribution of Independent Random Variables

Let us assume that each of the independent random variables Q_j appearing in three-dimensional functions (8.3) has the normal distribution. Then the fluctuations Δq_j of these random variables, describing the deviation from the nominal value of the variable q_j, have the probability density functions determined by the relations, comp. (2.35) in Sect. 2.4,

$$f\left(\Delta q_j\right) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\Delta q_j}{\sigma_j}\right)^2\right]. \tag{8.29}$$

If distributions of independent random variables Q_j are normal, then the three-dimensional distribution of resultant errors Δu, Δv, and Δw, in problems of practical interest, can be also considered to be normal.[2]

The probability density function of the three-dimensional normal distribution takes the following form, comp. (7.21b),

$$f(\Delta u, \Delta v, \Delta w) = \frac{1}{\sqrt{(2\pi)^3 \det[k_{ij}]}}$$

$$\times \exp\left[-\frac{1}{2}(K_{uu}\Delta u^2 + K_{vv}\Delta v^2 + K_{ww}\Delta w^2 + 2K_{uv}\Delta u\Delta v\right.$$

$$\left. + 2K_{vw}\Delta v\Delta w + 2K_{uw}\Delta u\Delta w)\right]. \tag{8.30}$$

The elements of matrix $[k_{jl}]$ are determined by relations (8.11). Their relations with coefficients K_{jl} result from (7.24).

The ellipsoids of probability concentration are determined by (7.25). We shall demonstrate in the following examples how to determine them analytically.

Example 8.3 Determine the ellipsoid of probability concentration for the problem of positioning accuracy of the hand of a simple manipulator shown previously in Fig. 8.1. The position in question is determined by the nominal positioning angles

$$\theta_1 = 0, \qquad \theta_2 = 0, \qquad \theta_3 = \frac{\pi}{2}, \tag{8.31}$$

thus, the position is identical with that assumed in Example 8.1. However, the distributions of positioning errors of joints ($j = 1, 2, 3$) are now assumed to be normal:

$$f\left(\Delta\theta_j\right) = \frac{1}{\sigma_j\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\Delta\theta_j}{\sigma_{\theta_j}}\right)^2\right]. \tag{8.32}$$

---

[2]Note that errors of positions U, V, W are calculated with linear approximation (8.7a). In general, if errors Δq_i are large, the distribution of resultant errors Δu, Δv, and Δw, being non-linear functions of normal variables, may not be normal.

The manipulator's dimensions are the same as in Example 8.1. Thus,

$$l_1 = l_2 = l = 1000 \text{ mm}.$$

Let us assume now that

$$\sigma_{\theta_1} = \sigma_{\theta_2} = \sigma_{\theta_3} = \sigma_0 = 0.001 \text{ rad.} \tag{8.33}$$

The values of partial derivatives of functions (8.1) with respect to the positioning angles θ_j have been collated previously in Table 8.2. Substituting these values into formulas (8.11) we get

$$k_{XX} = \sigma_X^2 = \left(\frac{\partial X}{\partial \theta_1}\right)^2 \sigma_{\theta_1}^2 + \left(\frac{\partial X}{\partial \theta_2}\right)^2 \sigma_{\theta_2}^2 + \left(\frac{\partial X}{\partial \theta_3}\right)^2 \sigma_{\theta_3}^2 = 2l^2\sigma_\theta^2,$$

$$k_{YY} = \sigma_Y^2 = \left(\frac{\partial Y}{\partial \theta_1}\right)^2 \sigma_{\theta_1}^2 + \left(\frac{\partial Y}{\partial \theta_2}\right)^2 \sigma_{\theta_2}^2 + \left(\frac{\partial Y}{\partial \theta_3}\right)^2 \sigma_{\theta_3}^2 = l^2\sigma_\theta^2,$$

$$k_{ZZ} = \sigma_Z^2 = \left(\frac{\partial Z}{\partial \theta_1}\right)^2 \sigma_{\theta_1}^2 + \left(\frac{\partial Z}{\partial \theta_2}\right)^2 \sigma_{\theta_2}^2 + \left(\frac{\partial Z}{\partial \theta_3}\right)^2 \sigma_{\theta_3}^2 = l^2\sigma_\theta^2,$$

$$k_{XY} = k_{YX} = \frac{\partial X}{\partial \theta_1}\frac{\partial Y}{\partial \theta_1}\sigma_{\theta_1}^2 + \frac{\partial X}{\partial \theta_2}\frac{\partial Y}{\partial \theta_2}\sigma_{\theta_2}^2 + \frac{\partial X}{\partial \theta_3}\frac{\partial Y}{\partial \theta_3}\sigma_{\theta_3}^2 = -l^2\sigma_\theta^2,$$

$$k_{XZ} = k_{ZX} = \frac{\partial X}{\partial \theta_1}\frac{\partial Z}{\partial \theta_1}\sigma_{\theta_1}^2 + \frac{\partial X}{\partial \theta_2}\frac{\partial Z}{\partial \theta_2}\sigma_{\theta_2}^2 + \frac{\partial X}{\partial \theta_3}\frac{\partial Z}{\partial \theta_3}\sigma_{\theta_3}^2 = 0,$$

$$k_{ZY} = k_{YZ} = \frac{\partial Y}{\partial \theta_1}\frac{\partial Z}{\partial \theta_1}\sigma_{\theta_1}^2 + \frac{\partial Y}{\partial \theta_2}\frac{\partial Z}{\partial \theta_2}\sigma_{\theta_2}^2 + \frac{\partial Y}{\partial \theta_3}\frac{\partial Z}{\partial \theta_3}\sigma_{\theta_3}^2 = 0.$$

The matrix $[k_{jl}]$ of second-order moments, comp. (8.10), takes, therefore, the form

$$[k_{jl}] = \begin{bmatrix} 2l^2\sigma_\theta^2 & -l^2\sigma_\theta^2 & 0 \\ -l^2\sigma_\theta^2 & l^2\sigma_\theta^2 & 0 \\ 0 & 0 & l^2\sigma_\theta^2 \end{bmatrix}.$$

The inverse matrix $[K_{jl}] = [k_{jl}]^{-1}$, calculated with the use of (7.24), is

$$[K_{jl}] = \begin{bmatrix} K_{XX} & K_{XY} & K_{XZ} \\ K_{YX} & K_{YY} & K_{YZ} \\ K_{ZX} & K_{ZY} & K_{ZZ} \end{bmatrix} = \begin{bmatrix} \frac{1}{l^2\sigma_\theta^2} & \frac{1}{l^2\sigma_\theta^2} & 0 \\ \frac{1}{l^2\sigma_\theta^2} & \frac{2}{l^2\sigma_\theta^2} & 0 \\ 0 & 0 & \frac{1}{l^2\sigma_\theta^2} \end{bmatrix}.$$

Let us assume now the probability $\alpha = 0.683$ that the error vector **p** of the accuracy of positioning the hand in its nominal position will lie inside the ellipsoid of probability concentration. Using Table 7.1 or the diagram in Fig. 7.1 we obtain for
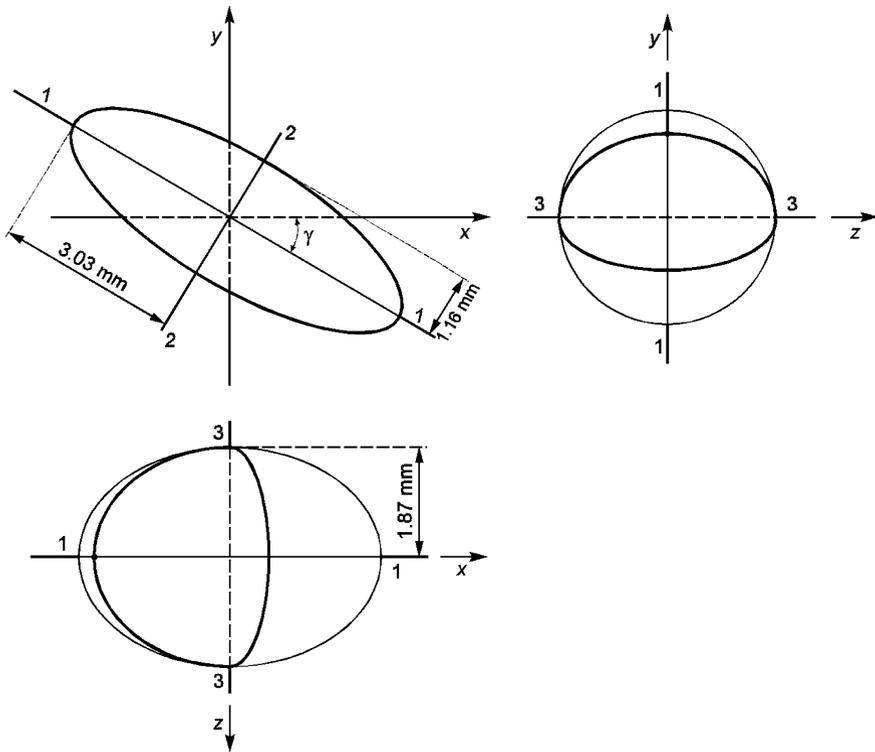
Fig. 8.5 Three projections of the ellipsoids of probability concentration of the positioning accuracy for the manipulator presented in Fig. 8.1, Example 8.3

the assumed probability level that $\chi_p^2(3) = \lambda^2 = 3.5$. By substituting all the data into formula (7.25) we arrive at the equation of the ellipsoid of probability concentration

$$x^2 + 2xy + 2y^2 + z^2 = \lambda^2 l^2 \sigma_\theta^2 = 3.5.$$

The ellipsoid is shown in Fig. 8.5. Its longer axis makes an angle $\gamma = 31°43'$ with the x-axis. The principal semi-axes are

$$a = 3.03 \text{ mm}, \qquad b = 1.16 \text{ mm}, \qquad c = 1.87 \text{ mm}.$$

The oblateness of the ellipsoid is given by the parameters

$$\frac{a}{b} = 2.61, \qquad \frac{a}{c} = 1.62, \qquad \frac{b}{c} = 0.62.$$

The theoretical ellipsoid shown in Fig. 8.5 has been compared with the results of a numerical experiment, comp. [4]. Random small Gaussian deviations from the desired revolving joint positions (8.31) have been numerically generated by a computer program followed by displacements of the manipulator's hand from its
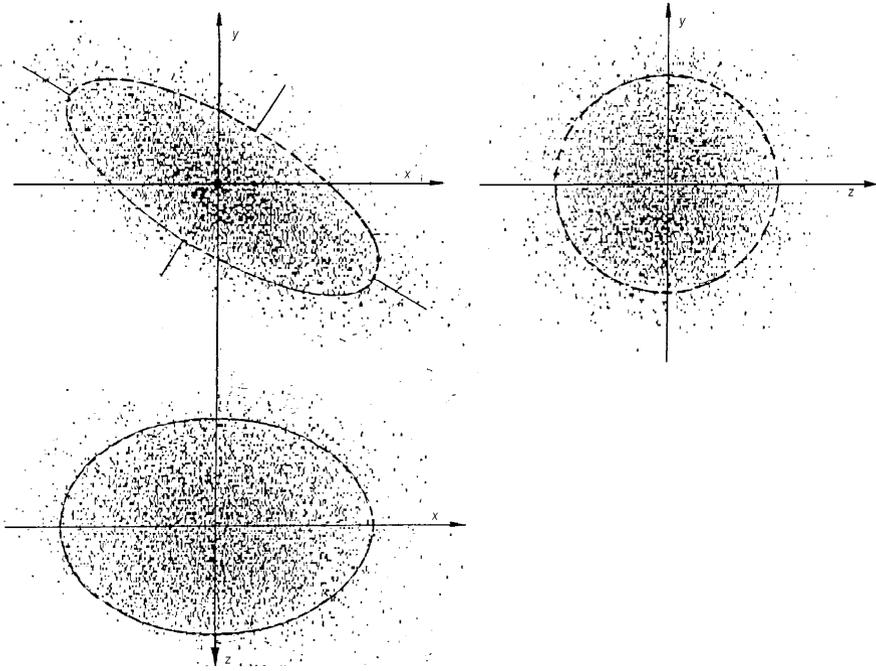
Fig. 8.6 The comparison of the ellipsoids presented in Fig. 8.5 with the results of computer simulation

nominal position. Calculated random displaced positions of the hand are shown in Fig. 8.6 as the corresponding points. The theoretical ellipsoid from Fig. 8.5 is also shown in the figure. The theoretical ellipsoid fits well with the assembly of numerically simulated points.

Example 8.4 Determine the ellipsoid of probability concentration for the problem of positioning accuracy of a more complex manipulator shown schematically in Fig. 8.3. We assume a certain particular position of the hand which is determined by the nominal values of the positioning angles of revolving joints

$$\theta_1 = 0°, \qquad \theta_2 = 90°, \qquad \theta_3 = -90°, \qquad \theta_4 = 30°.$$

We shall perform calculations for the following data:

$$a = 700 \text{ mm}, \qquad b = 600 \text{ mm}, \qquad c = 250 \text{ mm},$$

$$\sigma_{\theta_1} = \sigma_{\theta_2} = \sigma_{\theta_3} = \sigma_{\theta_4} = 0.001 \text{ rad}.$$

In the coordinate system X, Y, Z the position of the hand is determined by functions (8.21) given in Example 8.2, Sect. 8.2. The numerical values of partial derivatives of these functions with respect to the independent variables for the manipulator's position in question are given in Table 8.4.

Table 8.4 Values of the partial derivatives in Example 8.4

j	1	2	3	4
$\frac{\partial X}{\partial \theta_j}$	0	−825	−125	−125
$\frac{\partial Y}{\partial \theta_j}$	816.5	0	0	0
$\frac{\partial Z}{\partial \theta_j}$	0	816.5	816.5	216.5

The matrix $[k_{jl}]$ (comp. (8.10)) of the second-order moments takes the following form (numerical values in mm^2):

$$\left[k_{jl}\right] = \begin{bmatrix} 0.7119 & 0 & -0.8028 \\ 0 & 0.6667 & 0 \\ -0.8028 & 0 & 1.3803 \end{bmatrix},$$

and the inverse matrix is

$$\left[K_{jl}\right] = \left[k_{jl}\right]^{-1} = \begin{bmatrix} 4.028 & 0 & 2.374 \\ 0 & 1.500 & 0 \\ 2.374 & 0 & 2.105 \end{bmatrix}.$$

By substituting the respective values into equation (7.25) we obtain for the probability $\alpha = 0.683$ the equation of the ellipsoid of probability concentration

$$4.082x^2 + 1.500y^2 + 2.105z^2 + 4.748xz = 3.50.$$

The ellipsoid is shown in Fig. 8.7 along with the result of a computer simulation.

8.4 Indirect Determination of the Ellipsoids of Probability Concentration

The ellipsoids of probability concentration may be also found indirectly, provided that all independent variables Q_j in functions (8.3) have the normal distribution, comp. [1]. First we assume, as in the analogous case of two-dimensional distributions (comp. Sect. 6.4), that independent random variables Q_j have a substitute rectangular distribution instead of the actual normal distribution. Then the polyhedron bounding the region of uniform distribution of errors Δu, Δv, and Δw of all three functions (8.3) should be determined by using the procedure described in Sect. 8.2. The bounding values of the substitute rectangular distribution of independent variables errors

$$\Delta q_j^+ = +\kappa\sigma_j \quad \text{and} \quad \Delta q_j^- = -\kappa\sigma_j,$$

depend on the assumed probability α that end points of the resultant error vector **p**, whose components are Δu, Δv, and Δw, will lie inside the polyhedron. Here
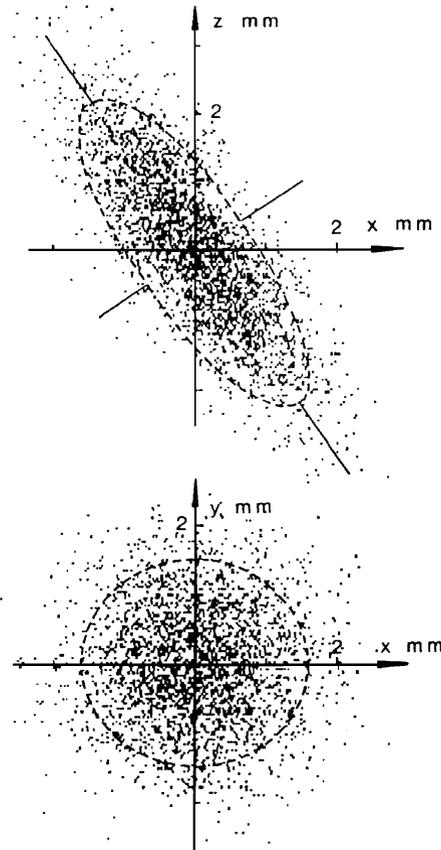
σ_j stands for the standard deviation of the actual normal distributions of any variable Q_j.

The polyhedron has been calculated under the assumption that $\kappa = \sqrt{3}$ corresponds to the probability $\alpha = 1$. For each $\kappa < \sqrt{3}$ the corresponding probability may be calculated as the ratio of the volume of the polyhedron obtained for $\Delta q_j = \pm \kappa \sigma_j$ to the volume of the polyhedron obtained for $\Delta q_j = \pm \sqrt{3}\sigma_j$. Thus, for a prescribed probability level α one should assume, comp. (6.31),

$$\kappa = \sqrt{3}\sqrt[3]{\alpha}. \tag{8.34}$$

Having found the polyhedron for a chosen probability level we calculate at first its second-order moments (inertia moments) with respect to the local central coordinate system. Then its principal central axes 1, 2, 3, and principal second-order moments I_1, I_2, and I_3 should be determined.

The ellipsoid of probability concentration may now be found as the ellipsoid having the same principal axes as the polyhedron. Moreover, its principal second-order moments should have the same values as the corresponding principal second-

order moments of the polyhedron. Thus, the lengths a, b, c of the semi-axes of the ellipsoid may be found by solving the system of three equations

$$\frac{4}{15}\pi ab^3c = I_1, \qquad \frac{4}{15}\pi a^3bc = I_2, \qquad \frac{4}{15}\pi abc^3 = I_3, \qquad (8.35)$$

where I_1, I_2, and I_3 are the principal second-order moments of the polyhedron.

Example 8.5 Find by means of the indirect method the ellipsoid of probability concentration for the problem of the positioning accuracy of the manipulator shown previously in Fig. 8.1. We shall assume the same data as in Examples 8.1 and 8.3. The nominal position of the manipulator's hand is assumed to be determined by the positioning angles

$$\theta_1 = 0, \qquad \theta_2 = 0, \qquad \theta_3 = \frac{\pi}{2}.$$

The lengths of the arms are

$$l_1 = l_2 = l = 1000 \text{ mm},$$

and standard deviations of the normal distribution of errors in positioning the revolving joints are

$$\sigma_{\theta_1} = \sigma_{\theta_2} = \sigma_{\theta_3} = \sigma_\theta = 0.001 \text{ rad}.$$

Assume the probability level $\alpha = 0.683$ that the error vector, in positioning the hand, will lie inside the ellipsoid of probability concentration. Thus, using formula (8.34) we get $\kappa = 1.525$ and

$$\Delta\theta_1 = \Delta\theta_2 = \Delta\theta_3 = \pm\kappa\sigma_\theta = \pm0.001525 \text{ rad}.$$

The polyhedron for these data was determined in Example 8.1 and shown in Fig. 8.2. Its second-order moments with respect to the reference planes of a local coordinate system shown in that figure are:

$$I_{xx} = \frac{1}{6}d^3he = 43.99 \text{ mm}^5, \quad \text{moment with respect to } x = 0 \text{ plane};$$

$$I_{yy} = \frac{1}{12}dh^3e = 21.98 \text{ mm}^5, \quad \text{moment with respect to } y = 0 \text{ plane};$$

$$I_{zz} = \frac{1}{12}dhe^3 = 21.98 \text{ mm}^5, \quad \text{moment with respect to } z = 0 \text{ plane}.$$

The mixed second-order moment with respect to the reference planes $x = 0$ and $y = 0$ is

$$I_{xy} = -\frac{1}{12}d^2h^2e = -21.98 \text{ mm}^5.$$

Principal axis 1 of the polyhedron coincides with the plane $z = 0$ and makes with the x-axis an angle γ determined by the relation, comp. formula (5.10),

$$\tan 2\gamma = \frac{2I_{xy}}{I_{xx} - I_{yy}} = -2.$$

Hence, we get $\gamma = -31°43'$. Thus, the angle is the same as that obtained for the ellipsoid by means of a direct method, see Fig. 8.5.

The principal second-order moments of the polyhedron are

$$I_1 = \frac{1}{2}(I_{xx} + I_{yy}) - \sqrt{\frac{1}{4}(I_{xx} - I_{yy})^2 + I_{xy}^2} = 8.42 \text{ mm}^5,$$

$$I_2 = \frac{1}{2}(I_{xx} + I_{yy}) + \sqrt{\frac{1}{4}(I_{xx} - I_{yy})^2 + I_{xy}^2} = 57.56 \text{ mm}^5,$$

$$I_3 = I_{zz} = 21.98 \text{ mm}^5.$$

By solving the system of (8.35) we shall find the lengths a, b, c of semi-axes of the ellipsoid of probability concentration

$$a = 3.11 \text{ mm}, \qquad b = 1.19 \text{ mm}, \qquad c = 1.92 \text{ mm}.$$

The ratios of semi-axes are

$$\frac{a}{b} = 2.61, \qquad \frac{a}{c} = 1.62, \qquad \frac{b}{c} = 0.62.$$

Thus, the ellipsoid calculated by means of the indirect method differs only slightly from that obtained analytically in Example 8.3.

Note, however, that in more complex problems (comp. Fig. 8.4) calculations of second-order moments for the polyhedron may not be easy. Generally speaking, the direct method discussed in Sect. 8.3 is more universal in practical applications.

Problem 8.1 Find the polyhedron of the positioning accuracy for the manipulator shown in Fig. 8.3. Assume the position determined by the positioning angles

$$\theta_1 = 0°, \qquad \theta_2 = \theta_3 = 30°, \qquad \theta_4 = -60°,$$

and the data (8.22) and (8.23) given in Example 8.2.

Problem 8.2 Find the ellipsoid of probability concentration for the problem of positioning accuracy of the hand of a manipulator shown in Fig. 8.3. Assume the position determined by the positioning angles

$$\theta_1 = 0°, \qquad \theta_2 = \theta_3 = 30°, \qquad \theta_4 = -60°.$$

Dimensions of the manipulator and the standard deviations of joint positioning errors are:

$$a = 700 \text{ mm}, \qquad b = 600 \text{ mm}, \qquad c = 250 \text{ mm}.$$

$$\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = 0.001 \text{ rad}.$$

Assume the probability that the end point of the error displacement vector lying inside the ellipsoid is $p = 0.683$.

References

1. Kotulski, Z., Szczepiński, W.: On two methods of determining the ellipses and ellipsoids of the positioning accuracy of robot manipulators. Ing. Arch. (Arch. Appl. Mech.) **63**(1), 25–41 (1993)
2. Szczepiński, W.: Theory of polyhedron of positioning accuracy of manipulators. Mech. Mach. Theory **26**(7), 697–709 (1991)
3. Szczepiński, W.: Introduction to the Calculus of Tolerance Limits in Mechanical Engineering. Polish Scientific Editors, Warsaw (1993) (in Polish)
4. Szczepiński, W., Wesołowski, Z.: On determining positioning accuracy of robot manipulators. Bull. Acad. Pol. Sci., Sér. Sci. Tech. **42**, 177–186 (1994)

Chapter 9
Problems Described by Implicit Equations

9.1 Introduction

In previous chapters we considered the basic problems of the error calculus that, in their simplest form, can be formulated as calculation (or estimation) of the probability of an event of the form

$$A(\omega) < b, \tag{9.1}$$

where $A(\omega)$ is the value of the measurement's error, which is a measurable function of the parameter $\omega \in \Omega$ (an elementary event) from the probability space (Ω, F, P), see Chap. 2 and b is a deterministic constant. The parameter b can be interpreted as an acceptable value of the measurement error. The basic inequality in a form formulated in (9.1) refers to a one-dimensional random variable. However, such an inequality can be generalized to more complicated cases. In Chap. 4 we presented the problem where the random variable $A(\omega)$ is a function of n statistically independent random variables $X_1, X_2, \ldots, X_n$. In such a case one should find the probability distribution of the random variable

$$A(\omega) = A(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)), \tag{9.2}$$

which is a resultant distribution of the distributions of independent random variables, and from it calculate appropriate probabilities. However, even in the simplest practical cases of engineering applications we must apply approximate probability distributions. In the problems of error calculus presented in previous sections, we have used as an approximation the uniform distribution or the normal distribution, depending on the assumed distributions of arguments, with the parameters of the distribution obtained by linearization of the function $A(x_1, x_2, \ldots, x_n)$.

Further generalizations of (9.1) lead to increasing the dimension of the state space (the dimension of the space of values of the random variable $A(\omega)$). Problems in engineering practice have been presented in Chaps. 5 and 6 (two-dimensional problems) and in Chaps. 7 and 8 (three-dimensional problems). In such cases instead

of one-dimensional inequalities estimating measurement errors one must consider ellipses and ellipsoids of probability concentration (for normal statistically independent random variables) and polygons or polyhedrons of tolerances (for uniformly distributed statistically independent random variables) in two and three-dimensional problems, respectively. In previous chapters we already presented computational problems resulting from such engineering tasks. A wide area of applications can be described by another generalization of the formula (9.1). Let us assume that both quantities in (9.1) are the random variables and that

$$A(\omega) < B(\omega). \tag{9.3}$$

The inequality of the form (9.3) leads to more complicated calculations than the calculations arising from (9.1), since they require application of the joint two-dimensional probability distributions of the random variables $A(\omega)$ and $B(\omega)$ to approximate the probability of the event that $\{\omega : A(\omega) < B(\omega)\}$. In practical applications when the two random variables are certain functions of some other random variables, this condition leads to calculation of the probabilities of events of the form

$$A(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)) < B(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)), \tag{9.4}$$

where the probability distributions of the random variables $X_1, X_2, \ldots, X_n$ are known, which is an involved computation problem. Examples of such tasks can be found in the following paragraphs of this chapter.

Before starting presentation of concrete computational tasks, we indicate areas of possible applications of the generally formulated (9.4). In technological sciences it is most frequently used in reliability theory, see e.g., [3, 5, 6, 9]. This is because the inequality (9.4) can be interpreted as the condition that the set of structural parameters represented by the random variable $A(\omega)$ is within the area of safe values bounded by their critical values (above which a crash is possible) represented by the random variable $B(\omega)$.

Example 9.1 Consider the problem of reliability of the stretched steel rod. In such a case the random variable $A(\omega)$ is the value of the stretching force, the random variable $B(\omega)$ is the rod's strength resistance and the condition (9.4) defines the state of the rod's failure-free work. Exceeding the condition defined by the inequality (9.4) results in the rod's destruction (exceeding plasticity limit, rupture, etc.).

Analogously, the inequality (9.4) can be used in congestion problems to compare nominal and actual time of functioning of a service and also in optimization problems with constraints, e.g., in reliability-based optimization, see [10]. However, such practical applications exceed the scope of this book so we omit them and return to computational problems of estimating the probability of satisfying the inequality of the form (9.4).

Fig. 9.1 The domain of integration for calculation of the value of probability (9.5)
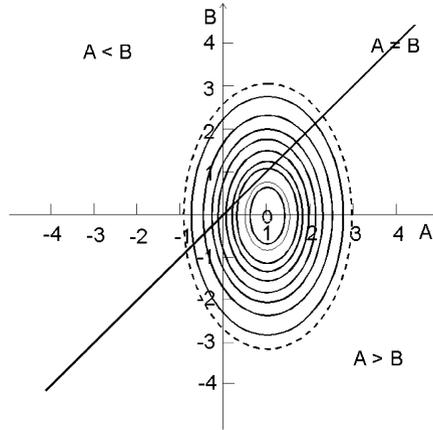
9.2 Statistically Independent Random Variables

9.2.1 Two Independent Random Variables

Consider two independent random variables $A(\omega)$ and $B(\omega)$ and the inequality of the form (9.3) written for the two variables. To simplify our reasoning, let us assume that their probability distribution is continuous. To obtain the probability of the event $\{\omega : A(\omega) < B(\omega)\}$ we must calculate the following integral of the joint probability density function of the two random variables $A(\omega)$ and $B(\omega)$, which in this particular case is a product of one-dimensional probability density functions $f_A(x)$, $f_B(y)$ (see (3.105)), over an area of the form $\{(x, y) : x < y\}$. Thus, we have

$$P\left(A\left(\omega\right) < B\left(\omega\right)\right) = \iint_{x<y} f_{AB}\left(x, y\right) dx dy = \iint_{x<y} f_A\left(x\right) f_B\left(y\right) dx dy. \quad (9.5)$$

The domain of integration is shown in Fig. 9.1.

It is seen in Fig. 9.1 that the integral of the form (9.5) can be transformed to the following simpler form:

$$P\left(A\left(\omega\right) < B\left(\omega\right)\right) = \int_{-\infty}^{\infty} \int_{-\infty}^{y} f_A\left(x\right) dx f_B\left(y\right) dy = \int_{-\infty}^{\infty} F_A\left(y\right) f_B\left(y\right) dy,$$
$$(9.6)$$

where $F_A(x)$ is the cumulative distribution function of the random variable $A(\omega)$. Alternatively, we can transform it to

$$P\left(A\left(\omega\right) < B\left(\omega\right)\right) = \int_{-\infty}^{\infty} \left[1 - F_B\left(x\right)\right] f_A\left(x\right) dx = 1 - \int_{-\infty}^{\infty} F_B\left(x\right) f_A\left(x\right) dx,$$
$$(9.7)$$

where now $F_B(y)$ is the cumulative distribution function of the random variable $B(\omega)$.

As we can see, in a case of independent random variables with an arbitrary known distribution, calculation of the probability (9.5) is a rather simple task, especially if we apply numerical computations. However, to do this one should know explicit expressions of the functions in (9.6) or (9.7).

Consider now the particular case of independent random variables that have the normal distribution, see (2.35). Thus, we assume that the random variable $A(\omega)$ has the distribution $N(m_A, \sigma_A^2)$, and the random variable $B(\omega)$ has the distribution $N(m_B, \sigma_B^2)$. In such a case, instead of the inequality (9.3), we can consider the other equivalent inequality, that is to calculate the following inequality:

$$P\left(A\left(\omega\right) < B\left(\omega\right)\right) = P\left(A\left(\omega\right) - B\left(\omega\right) < 0\right) = P\left(Z\left(\omega\right) < 0\right). \qquad (9.8)$$

The random variable occurring in (9.8),

$$Z\left(\omega\right) = A\left(\omega\right) - B\left(\omega\right), \qquad (9.9)$$

which is a difference of the two normally distributed statistically independent random variables, is also normally distributed, see (3.118)–(3.120); the parameters of its distribution are

$$\begin{aligned} m_Z &= m_A - m_B, \\ \sigma_Z^2 &= \sigma_A^2 + \sigma_B^2. \end{aligned} \qquad (9.10)$$

Thus, it suffices to perform a single integration of the probability density function of the normal distribution with the parameters given in (9.10) or read out the appropriate value from the standardized normal distribution tables according to (9.11)

$$P\left(A\left(\omega\right) < B\left(\omega\right)\right) = P\left(Z\left(\omega\right) < 0\right) = \Phi\left(\frac{-m_Z}{\sigma_Z}\right) = \Phi\left(\frac{-m_A + m_B}{\sqrt{\sigma_A^2 + \sigma_B^2}}\right). \quad (9.11)$$

Example 9.2 In Fig. 9.2 is presented the problem of calculation of the probability defined in (9.8) for the particular case of two independent normal distributions such that the random variable $A(\omega)$ has $N(3, 1)$ distribution and the random variable $B(\omega)$ has $N(1, 4)$ distribution. In such a case, the random variable $Z(\omega)$ has $N(2, 5)$ distribution. The value of the probability that we seek is equal to the measure of the marked area under the probability density plot $f_Z(x)$, or, according to (9.11), equal to $\Phi(-0.894) \approx 0.19$.

As we can see, the calculation of the probability of a random event of the form (9.3) is rather complicated even if we consider a single random variable instead of a function of random variables, as in (9.4). It is also seen that the analogous calculation becomes much easier if the random variables under consideration are normally distributed. This observation suggests an idea to make calculations on Gaussian random variables also in more general cases than (9.3). In the following part of this chapter we will show that, after applying a specific transformation, we can bring calculations on a wide class of distributions to calculations on the Gaussian distribution.
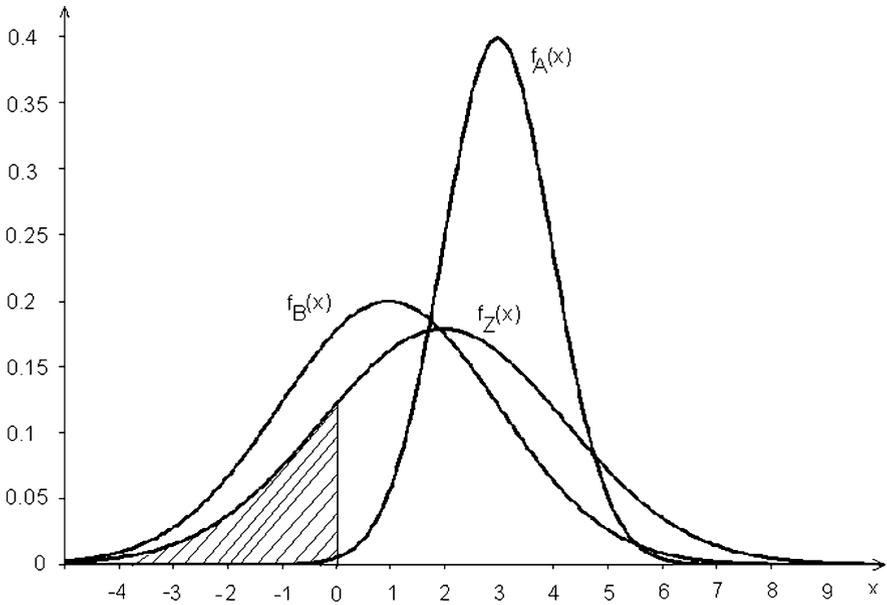
Fig. 9.2 Probability density functions of two independent random variables and of their difference

9.2.2 A Function of Independent Random Variables

Let us consider an inequality of the form (9.4) in a particular case where the random variables $X_1(\omega), X_2(\omega), \ldots, X_n(\omega)$ are statistically independent. In such a case the statistically independent random variables are described by the joint probability density function which is a product of one-dimensional probability density functions describing distributions of each of the random variables $X_i(\omega)$, $i = 1, 2, \ldots, n$. Thus, the probability density function $f_{\mathbf{X}}(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ has the following form:

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n). \qquad (9.12)$$

The probability density function (9.12) can be used for calculation of the probability of the event (9.4), that is

$$p = P(\omega : A(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)) < B(X_1(\omega), X_2(\omega), \ldots, X_n(\omega))). \qquad (9.13)$$

Applying (9.12) we can write

$$p = \int_{A(x_1, x_2, \ldots, x_n) < B(x_1, x_2, \ldots, x_n)} f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n) dx_1 dx_2 \cdots dx_n$$

$$= \int_{g(x_1, x_2, \ldots, x_n) < 0} f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n) dx_1 dx_2 \cdots dx_n, \qquad (9.14)$$

where in (9.14) we introduced the function

$$g(x_1, x_2, \ldots, x_n) = A(x_1, x_2, \ldots, x_n) - B(x_1, x_2, \ldots, x_n) \qquad (9.15)$$

to define the boundary of integration area. As we can see, to calculate the probability (9.14) one should know the adequate area (in n-dimensional space) defined by a hypersurface of the form

$$g(x_1, x_2, \ldots, x_n) = 0, \qquad (9.16a)$$

or after introducing the new notation $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $x_i \in R$, $i = 1, 2, \ldots, n$,

$$g(\mathbf{x}) = 0. \qquad (9.16b)$$

Thus, the area of integration in (9.14) is such an area of the n-dimensional Euclidean space R^n, where the following condition is satisfied :

$$g(\mathbf{x}) < 0. \qquad (9.17)$$

Calculation of the integral (9.14), for arbitrary probability distributions $f_i(x)$, $i = 1, 2, \ldots, n$, although they are statistically independent, is not an easy task. In such a case the transformation leading to normally distributed random variables can be helpful, as we have observed in the case of the inequality (9.3) for two random variables, when the result for the normal distribution was easy to obtain.

Thus, consider the n-dimensional standardized normally distributed random variable $\mathbf{u} = (u_1, u_2, \ldots, u_n)$. Let us assume that we are able to make such a transformation of an n-dimensional random variable $\mathbf{X}$ with statistically independent coordinates of arbitrary continuous distributions into n-dimensional standardized random variable $\mathbf{U}$ with normal distributions,

$$\mathbf{U} = \mathbf{T}(\mathbf{X}), \qquad (9.18)$$

that will make it possible to calculate the probability p. It proves to be that in a case of statistically independent random variables such a transformation has a very simple form; it can be written for each coordinate as

$$u_i = \Phi^{-1}\left(F_{X_i}(x_i)\right), \qquad (9.19)$$

for $i = 1, 2, \ldots, n$, where $\Phi^{-1}(.)$ is a map inverse to the cumulative distribution function of the Gaussian distribution and $F_{X_i}(x_i)$ for $i = 1, 2, \ldots, n$, are the cumulative distribution functions of the transformed random variables,

$$F_{X_i}(x) = \int_{-\infty}^{x} f_{X_i}(s)\, ds. \qquad (9.20)$$

This procedure also changes the equation of the hypersurface limiting the area of integration and writes it down in the new coordinates, corresponding to the domain

of the n-dimensional standardized normal distribution. The transformation of the integration boundary has the following form:

$$g(\mathbf{x}) = 0 \quad \Rightarrow \quad g\left(\mathbf{T}^{-1}(\mathbf{u})\right) = G(\mathbf{u}) = 0. \tag{9.21}$$

Before we start more precise explanations of the relationship (9.21), we will give several useful formulas. Thus, the equation (9.19) can be written formally as

$$u_i = T_i(x_i), \tag{9.22}$$

where $T_i(x_i)$ is the i-th coordinate of the vector-valued function $\mathbf{T}(\mathbf{x})$ and its inverse function $T_i^{-1}(u_i)$ (the i-th coordinate of the vector-valued function $\mathbf{T}^{-1}(\mathbf{u})$) can be represented as

$$x_i = F_{X_i}^{-1}(\Phi(u_i)), \tag{9.23}$$

where $F_{X_i}^{-1}(.)$ are the inverse functions to the cumulative distribution functions of the random variables $X_i, i = 1, 2, \ldots, n$. Using (9.23) we can write the equation of hypersurface (9.21) in an expanded form:

$$G(u_1, u_2, \ldots, u_n) = g\left(F_{X_1}^{-1}(\Phi(u_1)), F_{X_2}^{-1}(\Phi(u_2)), \ldots, F_{X_n}^{-1}(\Phi(u_n))\right) = 0. \tag{9.24}$$

Since the new probability distribution, which is the standardized n-dimensional normal distribution, is well known, the key problem in calculation of the probability (9.13) is constructing an adequate transformation of the hypersurface (9.24). How does such a construction look practically? To show this we will start from the particular case when both the initial and the transformed distributions are the normal distributions.

Example 9.3 Consider an example of the Gaussian random variables X_i, $i = 1$, $2, \ldots, n$ with arbitrary finite expected values m_i and arbitrary finite variances σ_i^2. In such a case we can easily find the relationships (9.19) and (9.23). In turn, we make the following transformations (using the explicit form of the known cumulative distribution functions):

$$F_{X_i}(x_i) = \Phi(u_i),$$

$$\int_{-\infty}^{x_i} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}\left(\frac{v - m_i}{\sigma_i}\right)^2\right) dv = \int_{-\infty}^{u_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}s^2\right) ds,$$

$$\int_{-\infty}^{\frac{x_i - m_i}{\sigma_i}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}s^2\right) ds = \int_{-\infty}^{u_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}s^2\right) ds,$$

so,

$$\frac{x_i - m_i}{\sigma_i} = u_i$$
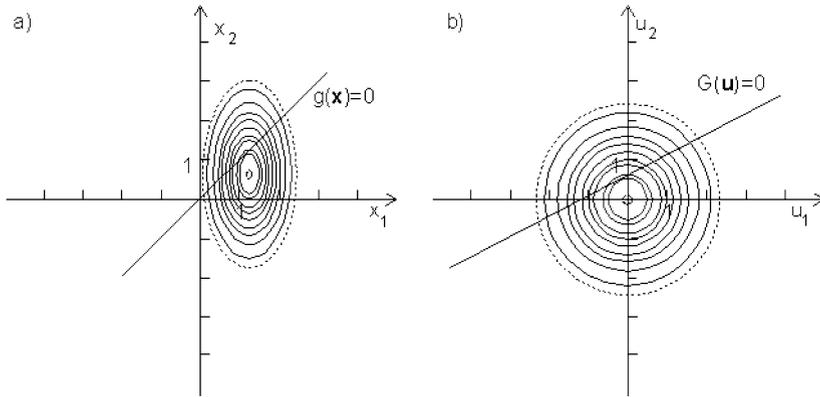
Fig. 9.3 Transformation of a system of independent normal random variables (**a**) to the system of standardized normal random variables (**b**)

and

$$x_i = \sigma_i u_i + m_i.$$

Finally we obtain that for any function $g(\mathbf{x})$, the function obtained after transformation $G(\mathbf{u})$ has the following form:

$$G(\mathbf{u}) = G(u_1, u_2, \ldots, u_n)$$
$$= g(\sigma_1 u_1 + m_1, \sigma_2 u_2 + m_2, \ldots, \sigma_n u_n + m_n). \tag{9.25}$$

In the particular case when we are considering the inequality (9.3), that is we deal with two random variables only, the resulting equations of the hypersurfaces (which in this particular case are the straight lines) have the following form:

$$g(x_1, x_2) = x_1 - x_2 = 0,$$
$$G(u_1, u_2) = \sigma_1 u_1 + m_1 - \sigma_2 u_2 - m_2 = 0.$$

Thus, the transformation of the straight line is its shifting and proportional lengthening along the axis of the coordinates system as it is shown in Fig. 9.3.

As we see from the presented example, for independent Gaussian distributions the map transforming the hypersurface $g(\mathbf{x}) = 0$ into the hypersurface $G(\mathbf{u}) = 0$ is the composition of parallel shift along the axes of the coordinate system and proportional extension of the hypersurfaces in the directions parallel to these axes. Let us remark that even if the above transformation is not complicated, obtaining practical results needs numerical calculations. It is not in fact an essential constraint because usually the calculation of the probability (9.13) is numerical, due to a complicated shape of the hypersurface $G(\mathbf{u}) = 0$ defining the integration area.

For arbitrary non-Gaussian statistically independent probability distributions, the maps leading to obtaining the functions $G(\mathbf{u})$ from the functions $g(\mathbf{x})$ have the
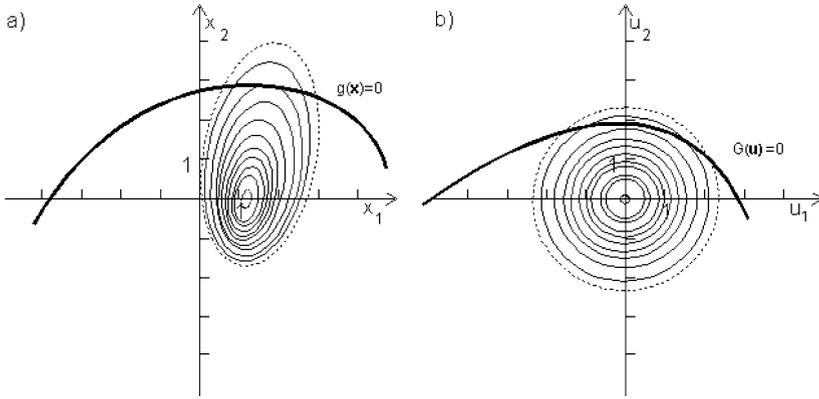
Fig. 9.4 Transformation of a system of arbitrary random variables to the system of standardized normal random variables

form (9.24). They cannot be represented in a linear form (9.25), as it is in the case of the Gaussian distributions. The result of this fact is that now the straight line from Fig. 9.3a is transformed into some arbitrary curve depending on the distribution of the random variables X_1, X_2, and in the case of more-dimensional spaces: the plane is transformed into some non-plane surface in three dimensions and hyperplanes are transformed into general hypersurfaces. In a general case, when the constraints put on the random variables $X_1, X_2, \ldots, X_n$ are defined by an arbitrary function $g(x_1, x_2, \ldots, x_n)$, the hypersurface $G(\mathbf{u}) = 0$ obtained after the transformation (9.18) can be of quite general shape. An example of such a transformation in two dimensions is shown in Fig. 9.4.

9.3 Statistically Dependent Random Variables

9.3.1 Two Dependent Random Variables

Calculating the probability of an event of the form (9.3) or (9.4) for dependent random variables is much more difficult than it is in a case of independent random variables presented in Sect. 9.2. Also in this case the advantageous operation is such a change of variables that leads the dependent random variables $X_1, X_2, \ldots, X_n$ to the form of a system of statistically independent random variables $U_1, U_2, \ldots, U_n$ with a known, optimally normal probability distribution. The appropriate transformation will be of the general form (9.18). However, it cannot be represented as a vector of one-dimensional transformations, as it was in a case of statistically independent random variables. To illustrate the problem we will start from the particular example of two statistically dependent random variables and the inequality (9.3).
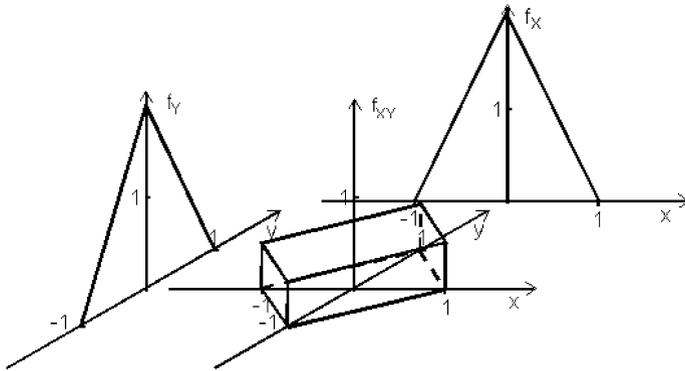
Fig. 9.5 The probability density function (9.26)

Example 9.4 Consider the two-dimensional random variable (X, Y) with probability density function of the form

$$f_{XY}(x, y) = \begin{cases} \frac{1}{2} & \text{for } |x| + |y| \leq 1, \\ 0 & \text{for } |x| + |y| > 1. \end{cases} \tag{9.26}$$

The function defined in (9.26) is plotted in Fig. 9.5.

We can see that the two-dimensional joint probability density function of the random variables X and Y is not a product of the two one-dimensional marginal probability density functions of X and Y,

$$f_{XY}(x, y) \neq f_X(x) f_Y(y),$$

from which follows that the random variables X and Y are statistically dependent. Now we calculate the probability of the random event $X(\omega) < Y(\omega)$. To do this we find the probability density function of the random variable $Z(\omega) = X(\omega) - Y(\omega)$ and next we calculate the probability of the random event that $Z(\omega) < 0$. We find the probability density function of the random variable Z in two steps. In the first step we find the joint probability density function of the vector-valued random variable (V, Z), where $V(\omega) = X(\omega) + Y(\omega)$ and $Z(\omega) = X(\omega) - Y(\omega)$. Using (3.116) we obtain

$$f_{VZ}(v, z) = \begin{cases} \frac{1}{4} & \text{for } -1 \leq v \leq 1, -1 \leq z \leq 1, \\ 0 & \text{for the other } v, z. \end{cases} \tag{9.27}$$

This new probability density function is shown in Fig. 9.6.

The second step is calculating the marginal probability density function $f_Z(z)$, that is calculating the integral of the joint two-dimensional probability density function (9.27) with respect to the argument v. As a result we obtain

$$f_Z(z) = \begin{cases} \frac{1}{2} & \text{for } -1 \leq z \leq 1, \\ 0 & \text{for the other } z. \end{cases}$$
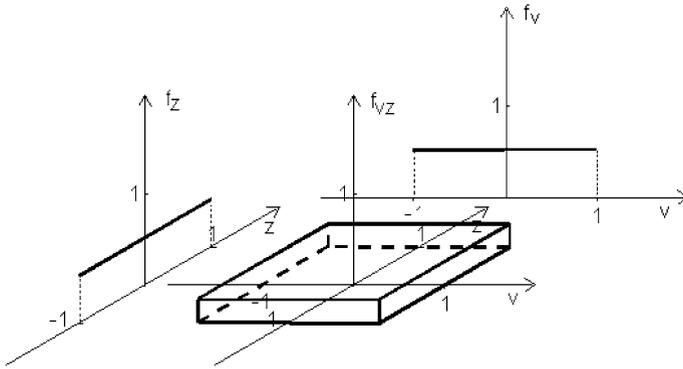
Fig. 9.6 The probability density function (9.27)

Finally we have: $P(Z(\omega) < 0) = \frac{1}{2}$.

The result of calculations in Example 9.4 is rather natural and it results from the symmetry of the probability distribution of the two-dimensional random variable (X, Y). However, the calculations performed show that for dependent random variables obtaining the probability of the form (9.13), especially in the multi-dimensional case, is not easy. As in the case of independent random variables, a transformation of the arbitrarily distributed dependent random variables into independent standardized normal random variables can be helpful. Before discussing the general example of multi-dimensional dependent random variables of arbitrary distribution, let us have a look at dependent Gaussian random variables.

9.3.2 The Case of Gaussian Random Variables

The multi-dimensional normal distribution discussed in this section is a specific distribution from the point of view of the transformation (9.18). As we observed in Example 9.3, considered in Sect. 9.2.2, in a case of independent random variables of the normal distribution, the appropriate transformation is a linear function (or, more precisely, affine transformation) of coordinates. The analogous property can be observed also for dependent normal random variables with the claim that now the transformation does not refer to particular coordinates but to vectors of coordinates.

In previous chapters we defined the examples of the probability density functions of the one-, two-, and three-dimensional normal distributions. It is known that such a distribution is completely defined by two parameters: the expected value (in the general n-dimensional case, by the vector of expected values) and by the variance (the covariance matrix, in the general case). Using these quantities we can write the probability density function of the n-dimensional normal distribution in a general

form:

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det[k_{ij}]}} \exp\left[-\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} K_{ij}(x_j - m_j)(x_k - m_k) \right],$$
(9.28)

where $\mathbf{k} = [k_{ij}]$ is the covariance matrix of the random variables $X_1, X_2, \ldots, X_n$, the matrix $\mathbf{K} = [K_{ij}]$ is the inverse matrix to $\mathbf{k}$, and $\mathbf{m} = (m_1, m_2, \ldots, m_n)^T$ is the vector of the average values of the random variables $X_1, X_2, \ldots, X_n$. Such a form of the probability density function is valid if it defines a non-degenerate normal distribution, that is all the random variables $X_1, X_2, \ldots, X_n$ are linearly independent: none of them can be represented as a linear combination of some of the others.

Now we rewrite the probability density function (9.28) using matrix and vector notation. The obtained formula for the probability density function is

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det \mathbf{k}}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{m})^T \mathbf{K}(\mathbf{x} - \mathbf{m}) \right],$$
(9.29)

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ is the vector of arguments of the probability density function. Now we can use the fact that the covariance matrix $\mathbf{k}$ is non-degenerate and positive definite. In such a case we can represent it as a product of two triangular positive definite matrices in the form

$$\mathbf{k} = \mathbf{P}\mathbf{P}^T,$$
(9.30)

that is, as a product of the matrix $\mathbf{P}$ and its transpose matrix. Also the matrix $\mathbf{K}$ can be represented in the form of two triangular matrices:

$$\mathbf{K} = \left(\mathbf{P}^T \right)^{-1} \mathbf{P}^{-1}.$$
(9.31)

The matrix decomposition of the form (9.31) makes it possible to construct an orthogonalizing transformation of the system of random variables, which means a transformation that builds an equivalent system of statistically independent random variables.

Now let us try to calculate the probability (9.14) for the probability density function of the form (9.29). As a result we obtain (this notation is written symbolically since in fact the integration is with respect to n arguments):

$$p = \int_{g(\mathbf{x})<0} \frac{1}{\sqrt{(2\pi)^n \det \mathbf{k}}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{m})^T \mathbf{K}(\mathbf{x} - \mathbf{m}) \right] d\mathbf{x}$$

$$= \int_{g(\mathbf{x})<0} \frac{1}{\sqrt{(2\pi)^n \det \mathbf{k}}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{m})^T \left(\mathbf{P}^T \right)^{-1} \mathbf{P}^{-1}(\mathbf{x} - \mathbf{m}) \right] d\mathbf{x}$$

$$= \int_{g(\mathbf{x})<0} \frac{1}{\sqrt{(2\pi)^n \det \mathbf{k}}} \exp\left[-\frac{1}{2} \left(\mathbf{P}^{-1}(\mathbf{x} - \mathbf{m}), \mathbf{P}^{-1}(\mathbf{x} - \mathbf{m}) \right)_{R^n} \right] d\mathbf{x}, \quad (9.32)$$

where $(.,.)_{R^n}$ is the scalar product of two vectors in R^n. Now we can introduce the new argument $\mathbf{u}$,

$$\mathbf{u} = \mathbf{P}^{-1}(\mathbf{x} - \mathbf{m}), \tag{9.33}$$

or equivalently

$$\mathbf{x} = \mathbf{P}\mathbf{u} + \mathbf{m}, \tag{9.34}$$

and write down the last integral in (9.32) in the following form:

$$
\begin{aligned}
p &= \int_{g(\mathbf{P}\mathbf{U}+\mathbf{m})<0} \frac{1}{\sqrt{(2\pi)^n}} \exp\left[-\frac{1}{2}(\mathbf{u},\mathbf{u})_{R^n}\right] d\mathbf{u} \\
&= \int_{G(\mathbf{U})<0} \frac{1}{\sqrt{(2\pi)^n}} \exp\left[-\frac{1}{2}(\mathbf{u},\mathbf{u})_{R^n}\right] d\mathbf{u},
\end{aligned} \tag{9.35}
$$

where during a change of variables in (9.35) we used the fact that $\sqrt{\det \mathbf{k}} = \det \mathbf{P} = (\det \mathbf{P}^{-1})^{-1}$. As we can see from (9.35), the transform of variables of the form (9.33) makes it possible to find the equation of a hypersurface defining the area of integration after transforming the set of statistically dependent Gaussian random variables into the appropriate system of independent Gaussian random variables. Thus, in the new system of variables we have

$$G(\mathbf{U}) = g(\mathbf{P}\mathbf{u} + \mathbf{m}). \tag{9.36}$$

A simple example of the two-dimensional Gaussian distribution illustrates the method presented above.

Example 9.5 Consider the two-dimensional normal distribution with a known vector of expected values and a known covariance matrix. The vector and the matrix are

$$\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}, \qquad \mathbf{k} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

The matrix $\mathbf{k}$ can be decomposed into a product of matrices of the form (9.30), where the matrix $\mathbf{P}$ is

$$\mathbf{P} = \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix}.$$

The inverse matrix $\mathbf{P}$ has the form

$$\mathbf{P}^{-1} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ \frac{-\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{bmatrix}.$$

Thus, the transformation of the variables has the form

$$\mathbf{u} = \mathbf{P}^{-1}(\mathbf{x} - \mathbf{m})$$
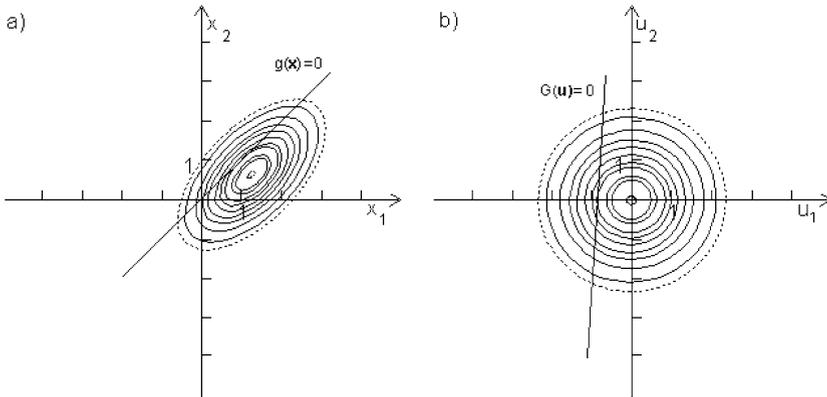
Fig. 9.7 Transformation of the two-dimensional normal random distribution to the standardized normal random distribution

$$= \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ \frac{-\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{bmatrix} \begin{bmatrix} x_1 - m_1 \\ x_2 - m_2 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{x_1-m_1}{\sigma_1} \\ -\frac{\rho(x_1-m_1)}{\sigma_1\sqrt{1-\rho^2}} + \frac{x_2-m_2}{\sigma_2\sqrt{1-\rho^2}} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix},$$

and for a given curve $g(x_1, x_2) = 0$ the transformed curve $G(u_1, u_2) = 0$ has the form

$$G(u_1, u_2) = g\left(\sigma_1 u_1 + m_1, \rho\sigma_2 u_1 + \sigma_2\sqrt{1 - \rho^2}u_2 + m_2\right) = 0.$$

In Fig. 9.7 we present as an example a transformation of the two-dimensional Gaussian random variables in a case when the boundary of the integration area takes the form of a straight line. Due to the transformation of arguments, the boundary of the integration area in the new system of coordinates has also the form of a straight line arisen from the original one by a composition of rotation, shift, lengthening and homothetic transformation of the plane.

9.3.3 More Random Variables: the Rosenblatt Transformation

After the preceding considerations of this chapter, only the most general case of the transformations of random variables remain to be analyzed: the initial random variables $X_1, X_2, \ldots, X_n$ are statistically dependent and their probability distribution is known. Before we start the analysis let us recapitulate the conclusions from the results of Sect. 9.3.2.

The first observation resulting from the previous considerations is that the transformation of the variables leading to a standardized normal distribution is of a specific structure. The result of the transformation is a vector-valued function, where the first coordinate depends on one argument, the second coordinate depends on two arguments, etc., and finally the n-th coordinate depends on n arguments being the values of the random variables under transformation.

The second observation is that the normal distribution has a specific property: each of the transforms described above is a linear map of, respectively, one, two, ..., n arguments. This is the counterpart of a fact that for the normal distribution the conditional expected value (the regression) is a linear function of the measurement data, see e.g. [2].

The above two observations suggest the following general form of the map transforming a system of n arbitrary dependent random variables $X_1, X_2, \ldots, X_n$ into the system of standardized Gaussian random variables $U_1, U_2, \ldots, U_n$. This map is called the Rosenblatt transform, see [8]. Such a transformation has two fundamental properties: the number of its arguments grows by one in each line of the transformation and if the transformed random variables are Gaussian the transformation is a linear map.

The Rosenblatt transformation is expressed by equality of the probabilities defined, respectively, by the cumulative distribution function of the Gaussian distribution and by the conditional cumulative distribution functions of the transformed, in sequence, random variables $X_i, i = 1, 2, \ldots, n$. The Rosenblatt transformation can be symbolically written by means of the conditional cumulative distribution functions as:

$$
\begin{aligned}
\Phi(u_1) &= F_{X_1}(x_1), \\
\Phi(u_2) &= F_{X_2}(x_2 | X_1 = x_1), \\
\Phi(u_3) &= F_{X_3}(x_3 | X_1 = x_1, X_2 = x_2), \\
&\cdots \\
\Phi(u_{n-1}) &= F_{X_{n-1}}(x_{n-1} | X_1 = x_1, X_2 = x_2, \ldots, X_{n-2} = x_{n-2}), \\
\Phi(u_n) &= F_{X_n}(x_n | X_1 = x_1, X_2 = x_2, \ldots, X_{n-1} = x_{n-1}),
\end{aligned}
\tag{9.37}
$$

or by means of the conditional probability density functions:

$$
\begin{aligned}
\Phi(u_1) &= \int_{-\infty}^{x_1} f_1(\tau)\, d\tau, \\
\Phi(u_2) &= \int_{-\infty}^{x_2} \frac{f_2(x_1, \tau)}{f_1(x_1)}\, d\tau, \\
\Phi(u_3) &= \int_{-\infty}^{x_2} \frac{f_3(x_1, x_2, \tau)}{f_2(x_1, x_2)}\, d\tau, \\
&\cdots
\end{aligned}
\tag{9.38}
$$

$$\Phi\left(u_{n-1}\right)=\int_{-\infty}^{x_{n-1}}\frac{f_{n-1}\left(x_1,x_2,\ldots,x_{n-2},\tau\right)}{f_{n-2}\left(x_1,x_2,\ldots,x_{n-2}\right)}d\tau,$$

$$\Phi\left(u_n\right)=\int_{-\infty}^{x_n}\frac{f\left(x_1,x_2,\ldots,x_{n-1},\tau\right)}{f_{n-1}\left(x_1,x_2,\ldots,x_{n-1}\right)}d\tau,$$

where

$$f_1\left(x_1\right)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f\left(x_1,x_2,x_3,\ldots,x_n\right)dx_2dx_3\cdots dx_n,$$

$$f_2\left(x_1,x_2\right)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f\left(x_1,x_2,x_3,\ldots,x_n\right)dx_3dx_4\cdots dx_n,$$

$$\ldots$$

$$f_{n-1}\left(x_1,x_2,\ldots,x_{n-1}\right)=\int_{-\infty}^{\infty}f\left(x_1,x_2,x_3,\ldots,x_n\right)dx_n.$$

(9.39)

Example 9.6 Consider two statistically dependent random variables X,Y having the joint probability density function of the form

$$f_{XY}\left(x,y\right)=\left(x+y+xy\right)e^{-(x+y+xy)}.\tag{9.40}$$

For the two random variables, the probability that the condition

$$X\left(\omega\right)<Y\left(\omega\right)\tag{9.41}$$

is satisfied can be calculated by integration of the two-dimensional probability density function (9.40) over such an area of the Euclidean space where $x<y$, that is

$$g\left(x,y\right)=x-y<0.\tag{9.42}$$

Thus, the boundary line separating the areas where the inequality (9.41) is satisfied or is not satisfied is the straight line. Now we apply the Rosenblatt transformation to write down this boundary line in the new coordinates system u_1,u_2, corresponding to the probability density function of the standardized Gaussian random variables U_1,U_2.

$$f_1(x)=f_X\left(x\right)=\int_0^{\infty}f_{XY}\left(x,y\right)dy=\int_0^{\infty}\left(x+y+xy\right)e^{-(x+y+xy)}dy$$

$$=e^{-x}\left[x\int_0^{\infty}e^{-(1+x)y}dy+(1+x)\int_0^{\infty}ye^{-(1+x)y}dy\right]$$

$$=e^{-x}\left[\frac{x}{1+x}+\frac{1+x}{(1+x)^2}\right]=e^{-x},\tag{9.43}$$

$$\Phi\left(u_1\right)=\int_0^x f_1\left(\tau\right)d\tau=\int_0^x e^{-\tau}d\tau=1-e^x,\tag{9.44}$$

$$\Phi(u_2) = \int_0^y \frac{f_2(x, \tau)}{f_1(x)} d\tau$$

$$= \int_0^y \frac{(x + \tau + x\tau) e^{-(x+\tau+x\tau)}}{e^{-x}} d\tau = \int_0^y [x + \tau(1+x)] e^{-(1+x)\tau} d\tau$$

$$= 1 - (1 + y) e^{-(1+x)y}. \tag{9.45}$$

The appropriate transformation of the variables in the function (9.42) allows us to obtain the arguments x, y expressed as functions of u_1, u_2 by solving the following system of equations:

$$\begin{aligned} 1 - e^x &= \Phi(u_1), \\ 1 - (1 + y) e^{-(1+x)y} &= \Phi(u_2). \end{aligned} \tag{9.46}$$

9.4 Computational Problems

In this chapter we have discussed the task of calculating the probability of an event described by the inequality (9.4), which we called for simplicity, the implicit problem. Analogously as in problems of the error calculus presented in preceding chapters, many such problems can be reduced to calculating probabilities for the multi-dimensional standardized Gaussian distribution. However, in the case of error calculus problems we considered the probability integrals over an area symmetric with respect to the center of the distribution, bounded by a spherical surface or n-dimensional ellipsoid (or in the lower dimensional spaces by, respectively, two points, ellipses and ellipsoids), but in the implicit problems the integration area is an area bounded by a general hypersurface in the n-dimensional space (analogously, in the lower dimensions: a curve of a surface). Therefore calculations of the probabilities in the implicit problems are performed with the use of approximate methods, usually numerical methods.

The mathematical methods presented in Chap. 9 are very often used in the branch of mechanics called structural reliability theory. Those readers who are interested in details of application in reliability theory may study specialized monographs, e.g., [1, 3] or [6], where one can find both theoretical aspects of reliability problems and computational methods. In particular, the extended information about applications of the Rosenblatt transformation that is an effective tool in reliability methods of engineering structures leading to unified calculations of standardized Gaussian distributions can be found in original papers, e.g. in [4, 10] and also in the monograph by Ditlevsen and Madsen [3].

Let us return to the initial problem formulated in Sect. 9.1. We presented there the two inequalities (9.3) and (9.4). The first of them described the relation of two random variables, in a general case statistically dependent ones, while the second inequality described two functions of systems of statistically dependent random variables. We presented a detailed analysis of the inequality (9.4), obtaining as a result an explicit expression for the probability that the inequality is satisfied, which is in

a form of integral of the probability density function in n-dimensional Euclidean space. The problem presented in such a form makes possible an extended analysis taking into account the influence of each of the constituent random variables $X_1, X_2, \ldots, X_n$ on the resultant probabilities. It is especially important in reliability theory problems where one can analyze the influence of each of the factors in the model on the lifetime of an engineering structure.

Except for the exact analysis of the inequality (9.4) presented above, the problem of most interest is its approximate analysis which consists in reduction of the problem to the form (9.3) and calculation of the appropriate probabilities for such an inequality. It is known that random variables of the form (9.2), that is to say

$$
\begin{aligned}
A(\omega) &= A(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)), \\
B(\omega) &= B(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)),
\end{aligned}
\tag{9.47}
$$

are deterministic functions of the random variables $X_1(\omega), X_2(\omega), \ldots, X_n(\omega)$ with a known distribution, so one can find (approximately, in a general case) the joint probability distribution of the random variables (9.47). Thus, instead of integrations in the n-dimensional space one integrates in a two-dimensional space. Such calculations, although less accurate, are sufficient in some applications. If additionally we use the Rosenblatt transformation, we can calculate probabilities for functions of n random variables of arbitrary distributions using the two-dimensional probability density function of the normal distribution.

The fundamental problem to solve is calculation of an integral of the two-dimensional circular normal distribution (see Chap. 5). Difficulties in such a calculation are due to the irregular shape of the integration area resulting from the form of the boundary curve $G(u_1, u_2) = 0$. Calculation of the probability p with high accuracy, where $\phi(u_1, u_2)$ is the probability density function of the two-dimensional standardized normal distribution,

$$
p = P(A(\omega) < B(\omega)) = \int_{G(u_1,u_2)<0} \phi(u_1, u_2) \, du_1 du_2,
\tag{9.48}
$$

is a very complicated task. A high accuracy in calculations is essential in professional technological applications, see e.g. [3, 7, 10].

In some problems one need not calculate the exact value of the probability (9.48) and it suffices to estimate its value from above (in a case of a crash probability) or from below (in a case of the probability of structural reliability). Then one can approximate the exact areas of integration $\{u_1, u_2 \in R^2 : G(u_1, u_2) < 0\}$ with such subsets of the R^2 space for which the integration of the probability density function is easy. The areas of such a kind are: circles with their center coinciding with the center of the system of coordinates O or their sectors, circular rings with their middle in O or their sectors, half-planes, plane wedge-shaped sectors with their top in O and systems of such subsets.

Example 9.7 Let us consider the example presented in Fig. 9.8. In this problem one should calculate the probability of the area cut out from the plane by the curve $G(u_1, u_2) = 0$. The exact value of this probability lies in between $\chi_2^2(r)$ and $\Phi(r)$.

Fig. 9.8 An example of
approximation of the
probability (9.48)
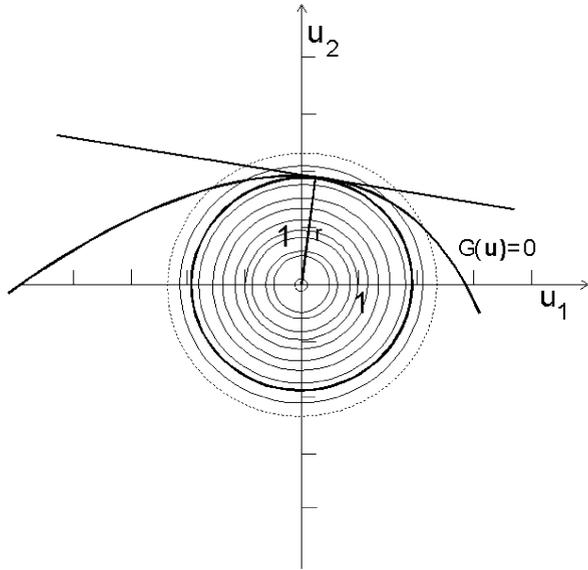
The first of the numbers is the probability of hitting the circle with its center coinciding with the center of the coordinate system and tangent to the curve $G(u_1, u_2) = 0$. As we know from Chap. 5, it is calculated according to the chi-squared distribution with two degrees of freedom. The second number is the probability that the one-dimensional standardized normally distributed random variable (that is, the random variable corresponding to the distance measured along the line normal to the curve $G(u_1, u_2) = 0$) will exceed the value r. Since the values of the normally distributed random variable are strongly concentrated around the center of the coordinate system, the presented approximation is quite precise.

Concluding this chapter we give a number of examples of the value of the probability for specific areas of integration.

The circle with a center located in the center of the system of coordinates O and the radius r:

$$p = F_{\chi^2(2)}(r),$$

where $F_{\chi^2(2)}(r)$ is the value at the point r of the cumulative distribution function of the chi-squared distribution with two degrees of freedom.

The sector of the above circle with the angle α *rad*:

$$p = \frac{\alpha}{2\pi} F_{\chi^2(2)}(r).$$

The circular ring with the center O and two radii r_1 and r_2 satisfying $r_1 < r_2$:

$$p = F_{\chi^2(2)}(r_2) - F_{\chi^2(2)}(r_1).$$

The sector of the above circular ring with the angle α *rad*:

$$p = \frac{\alpha}{2\pi}\left(F_{\chi^2(2)}(r_2) - F_{\chi^2(2)}(r_1)\right).$$

The half-plane containing the center of the coordinate system O and limited with a straight line located with a distance r from O:

$$p = \Phi(r).$$

References

1. Biegus, A.: Probabilistic Analysis of Steel Structures. Polish Scientific Editors, Warsaw (1999) (in Polish)
2. Bryc, W.: The Normal Distribution. Characterization with Applications. Lecture Notes in Statistics, vol. 100. Springer, New York (1995)
3. Ditlevsen, O., Madsen, H.O.: Structural Reliability Methods. Wiley, Chichester (1996)
4. Doliński, K.: First-order second-moment approximation in reliability of structural systems: critical review and alternative approach. Struct. Saf. **1**, 211–231 (1983)
5. Melchers, R.E.: Structural Reliability Analysis and Prediction. Wiley, Chichester (1999)
6. Murzewski, J.: Reliability of Engineering Structures. Arkady, Warsaw (1989)
7. Rackwitz, R.: Reliability analysis—a review and some perspectives. Struct. Saf. **23**(4), 365–395 (2001)
8. Rosenblatt, M.: Remarks on a multivariate transformation. Ann. Math. Stat. **23**, 47–472 (1952)
9. Spaethe, G.: Die Sicherheit tragender Baukonstruktionen, 2nd edn. Springer, Berlin (1992)
10. Stocki, R.: Reliability-based optimization of truss structures for large displacements. Theory and software. Ph.D. Thesis, Prace IPPT-IFTR Reports 13/1999, Warsaw (1999)

Chapter 10
Useful Definitions and Facts of Probability Theory for Further Reading

10.1 Statistical Linearization

In many practical problems one must calculate a moment (e.g., the mean value) of the random variable Y which is a nonlinear function of the random variables X_j, $j = 1, 2, \ldots, N$ (see Sect. 2.3). Sometimes this operation can be done accurately, however in most cases we face certain computational problems. To find a remedy, we usually try to perform some approximate calculations to express the moment of the random variable Y as a deterministic function of moments of the random variables X_j, $j = 1, 2, \ldots, N$. The most popular method of such an approximation is the so-called linearization, that is, an exchange of the nonlinear function

$$Y = g(X_1, X_2, \ldots, X_N), \qquad (10.1)$$

by the linear expression

$$Y \approx Z = A_0 + \sum_{j=1}^{N} A_j (X_j - m_j), \qquad (10.2)$$

where the coefficients A_j, $j = 0, 1, 2, \ldots, N$ and m_j, $j = 1, 2, \ldots, N$ are chosen in a specific way. The method applied for selection of the coefficients is a foundation of the quality of the approximation made. For this reason, depending on the requirements of a certain engineering problem where nonlinear functions are applied, we can choose one of several possible methods of linearization, see e.g. [51].

One of the methods of linearization, relatively easy to perform and giving good results for smooth, regular functions $g(.)$, is the method of direct linearization. In this method we expand the nonlinear function $g(x_1, x_2, \ldots, x_N)$ into the Taylor series and treat its linear terms as the approximation of the function:

$$g(x_1, x_2, \ldots, x_N) \approx g(m_1, m_2, \ldots, m_N) + \sum_{j=1}^{N} \frac{\partial g}{\partial x_j} (x_j - m_j), \qquad (10.3)$$

where the partial derivatives $\frac{\partial g}{\partial x_j}$ for $j = 1, 2, \ldots, N$ are calculated at points $x_1 = m_1, \ldots, x_N = m_n$.

In this method we proceed in two steps. First we linearize the deterministic function $g(x_1, x_2, \ldots, x_N)$; next we obtain the approximate value of the random variable Y substituting in the formula (10.3), instead of the independent variables x_j, $j = 1, 2, \ldots, N$, the random variable X_j, $j = 1, 2, \ldots, N$. The point (in R^N) around which we do the linearization is usually the vector of mean values of the random variables X_j:

$$m_j = E\{X_j\}, \quad j = 1, 2, \ldots, N. \tag{10.4}$$

Finally, the random variable Y can be represented by

$$Y = g(m_1, m_2, \ldots, m_N) + \sum_{j=1}^{N} \frac{\partial g}{\partial x_j}(X_j - m_j), \tag{10.5}$$

where the derivatives $\frac{\partial g}{\partial x_j}$ are calculated as in formula (10.3).

The method of direct linearization gives good results in the case when terms of higher order in the Taylor expansion of the function $g(.)$ are small enough, with probability equal to 1, in comparison to linear terms of the expansion. This means that the derivatives of higher order of the function $g(.)$ are bounded and the random variables X_j, $j = 1, 2, \ldots, N$ are strongly concentrated around their mean values (which gives small variances).

The method of direct linearization can be applied to vector-valued functions of random variables in a way analogous to the one presented above.

In this book, the method of direct linearization is consistently applied in the analysis of the errors of measurements. Examples of application of this method to solutions of stochastic differential equations can be found in [48].

In addition to the method presented above there are some other approaches to linearization of functions of random variables or, more generally, stochastic systems. This group of methods, called stochastic linearization, is based on the general assumption that the nonlinear function must be replaced by a linear approximation (10.2), where the coefficients $A_j = 0, 1, 2, \ldots, N$ are chosen according to some criterion of stochastic optimality. Selection of a criterion depends on its applicability in the problem under consideration. The number of variants of stochastic linearization is quite large and its usefulness has been confirmed in many practical engineering problems.

The problem of statistical linearization is extensively studied in the literature. Complete presentation of the state of the art exceeds the boundaries of this book. More information and some bibliographical references can be found in [38, 45]. Now let us present some simple examples illustrating the idea of the method.

Example 10.1 Given the random variable Y which is a nonlinear function of some random variable X with the mean value m_X and the variance σ_X^2:

$$Y = g(X). \tag{10.6}$$

We search for the coefficients A_0 and A_1, such that the random variable Z defined as

$$Z = A_0 + A_1(X - m_X) \tag{10.7}$$

and the random variable Y have equal mean values and variances, that is, they satisfy the conditions

$$E\{Y\} = E\{Z\} \quad \text{and} \quad \sigma_Y^2 = \sigma_Z^2. \tag{10.8}$$

We find that

$$A_0 = E\{g(X)\} \tag{10.9}$$

and

$$A_1 = \pm \sqrt{\frac{\sigma_{g(X)}^2}{\sigma_X^2}}. \tag{10.10}$$

The term $\sigma_{g(X)}^2$ in (10.10) is the variance of the random variable Y defined in (10.6) and is a deterministic function of the random variable X. The sign in (10.10) is the one which fits better the approximation to the function $g(.)$.

Example 10.2 Now we postulate that the random variables defined in Example 10.1 satisfy the least squares condition

$$\min E\{[Y - Z]^2\} = \min_{A_0,\, A_1} E\{[g(X) - A_0 - A_1(X - m_X)]^2\}. \tag{10.11}$$

After calculating the square of the bracket expression and the mean values we find A_0 and A_1 by minimizing the expression (10.11). We obtain A_0 as that given in (10.9) while the coefficient A_1 takes now the form

$$A_1 = \frac{E\{[g(X) - Eg(X)][X - m_X]\}}{\sigma_X^2}. \tag{10.12}$$

Both examples show how the result of the linearization depends on the optimality criterion: depending on the method, we obtained two different values of the expansion coefficients A_i in formula (10.7).

The method of stochastic linearization is sometimes difficult to apply, especially if we consider stochastic differential equations. It gives, however, a possibility to solve problems in which the nonlinear function $g(.)$ is non-differentiable, because it does not require the smoothness property. This attribute of the method is particularly important in the case of experimentally determined functions, because in such a case the existence of derivatives and possibility of estimating them is often very problematic.

10.2 Multi-dimensional Regression

In Chap. 5 we considered a problem of finding relationships between two random variables when we have a series of simultaneous measurements of them. The purpose of the analysis was to obtain an explicit expression of the form (5.12a) or (5.12b), making it possible to calculate a value of one of the variables from the value of the other.

In this section we consider a more general problem of practical importance connected with measurements of multi-dimensional random variables. Let us assume that one measures simultaneously $k + 1$ random variables that make a vector $\mathbf{U}$:

$$\mathbf{U} = (X_1, X_2, \ldots, X_k, Y), \tag{10.13}$$

written here as a row vector. The purpose of the analysis is to construct a function of k arguments of the form

$$y = g(x_1, x_2, \ldots, x_k) \tag{10.14}$$

that could express functional dependence of the random variable Y from the random variables $X_1, X_2, \ldots, X_k$ considered as arguments. The appropriate approximation should be made from a sample of n observations (measurements) of the vector $\mathbf{U}$, that is from $\mathbf{u}_i = (x_{i1}, x_{i2}, \ldots, x_{ik}, y_i)$, $i = 1, 2, \ldots, n$.

The problem defined by (10.13) and (10.14) and methods of its solution are a generalization in more dimensions of the problem of constructing the regression curve or the regression line formulated in Chap. 5 for two-dimensional spaces. In this section we will present how to find the best linear expression approximating an arbitrary function defined in (10.14).

The problem under consideration is: find a set of constants $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_k$ that satisfy the following condition:

$$E\left\{[Y - (\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_k X_k)]^2\right\} = \min. \tag{10.15}$$

After making measurements we have a set of data that can be written as an n by k matrix $\mathbf{X}$ containing values of the arguments of the function and the n-dimensional vector $\mathbf{Y}$ containing the measured values of the function:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \qquad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix}. \tag{10.16}$$

Finding a linear expression of the form

$$g(x_1, x_2, \ldots, x_k) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k, \tag{10.17}$$

which minimizes (10.15), leads to the task of calculating such values of the parameters $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_k$ that minimize the following expression:

$$H(\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_k) = \sum_{i=1}^{n} (y_i - \alpha_0 - \alpha_1 x_{i1} - \alpha_2 x_{i2} - \cdots - \alpha_k x_{ik})^2. \quad (10.18)$$

The values are solutions of a system of equations obtained by equating to zero partial derivatives with respect to α_i, $i = 0, 1, 2, \ldots, k$ of the function defined in (10.18):

$$\frac{\partial}{\partial \alpha_i} H(\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_k) = 0, \quad i = 0, 1, 2, \ldots, k. \quad (10.19)$$

The first of (10.19) makes it possible to eliminate the parameter α_0 from the rest of the equations by substitution:

$$\alpha_0 = \overline{y} - \sum_{j=1}^{k} \alpha_j \overline{x}_j, \quad (10.20)$$

where

$$\overline{y} = \frac{1}{n} \sum_{i=1}^{n} y_i, \qquad \overline{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}, \quad j = 1, 2, \ldots, k. \quad (10.21)$$

Let us observe that after the substitution of (10.20) into (10.18) we obtain a function of only k arguments and its value is calculated for centered values of the measurements:

$$H_1(\alpha_1, \alpha_2, \ldots, \alpha_k) = \sum_{i=1}^{n} \left((y_i - \overline{y}) - \alpha_1 (x_{i1} - \overline{x}_1) - \alpha_2 (x_{i2} - \overline{x}_2) - \cdots \right.$$

$$\left. - \alpha_k (x_{ik} - \overline{x}_k) \right)^2. \quad (10.22)$$

For simplicity in calculations, to obtain the parameters α_i, $i = 1, 2, \ldots, k$, of the map defined in (10.17), we can solve the following system of equations:

$$\frac{\partial}{\partial \alpha_i} H_1(\alpha_1, \alpha_2, \ldots, \alpha_k) = 0, \quad i = 1, 2, \ldots, k. \quad (10.23)$$

From (10.23) after elementary transformations we obtain a system of linear equations which in matrix-vector notation takes the form

$$\hat{\mathbf{K}}_k \hat{\mathbf{a}}_k = \hat{\mathbf{c}}_k, \quad (10.24)$$

where $\hat{\mathbf{K}}_k$ is an estimator of the covariance matrix of measured arguments $X_1, X_2, \ldots, X_k$,

$$\hat{\mathbf{K}}_k = \left[K_{ij} \right] = \left[\mathrm{cov}\left(X_i, X_j \right) \right] = \left[\frac{1}{n} \sum_{p=1}^{n} (x_{pi} - \bar{x}_i)(x_{pj} - \bar{x}_j) \right], \quad i, j = 1, 2, \ldots, k,$$

$$(10.25)$$

$\hat{\mathbf{c}}_k$ is an estimator of the covariance vector of the measured arguments $X_1, X_2,$ $\ldots, X_k$ and the resulting variable Y,

$$\hat{\mathbf{c}}_k = [c_j] = [\text{cov}\,(Y, X_j)] = \left[\frac{1}{n}\sum_{p=1}^{n}(y_p - \bar{y})(x_{pj} - \bar{x}_j)\right], \quad j = 1, 2, \ldots, k,$$
(10.26)

and $\hat{\mathbf{a}}_k$ is the vector of system parameters that are wanted,

$$\hat{\mathbf{a}}_k = [\alpha_1, \alpha_2, \ldots, \alpha_k]^T.$$
(10.27)

Finding the optimal linear function which satisfies the condition (10.15) is now reduced to solving the system of (10.24) and calculating the missing parameter α_0 from (10.20).

More results concerning multi-dimensional regression and its applications can be found in specialized monographs, e.g. [15, 40].

10.3 Limit Theorems of Probability Theory

Probability theory, being by definition a tool for description of random phenomena, has some regularities which make it possible to predict some properties of their behavior. Remember the results of the classical physical experiment shown in Fig. 2.8, where, due to multiple random reflections of a large number of balls from the pegs at the especially prepared board, we obtain a shape of the Gaussian probability density function. This and other similar processes were described in an abstract way and formalized as the limit theorems of probability theory. To present these results we must introduce the concepts of probabilistic convergence.

10.3.1 Concepts of Probabilistic Convergence

Consider the sequence of random variables X_j, $j = 1, 2, \ldots$. We want to give various definitions of convergence of this sequence to a limit. For the sake of simplicity we will assume the convergence to zero. In the case of sequences with a non-zero limit we can subtract this value from every element of the sequence, obtaining finally the required condition.

We say that the sequence X_i is *stochastically convergent* (or: *converges in probability*) to zero if for every $\epsilon > 0$ the following condition is satisfied:

$$\lim_{j \to \infty} P\left(|X_j| > \epsilon\right) = 0.$$
(10.28)

Let F_j, $j = 1, 2, \ldots$ be the sequence of the probability distribution functions corresponding to random variables X_j, $j = 1, 2, \ldots$. We say that the sequence of

random variables X_j converges to zero *according to probability distribution functions* (or *in distribution*) if the sequence of distributions satisfies the condition

$$\lim_{j \to \infty} F_j(x) = \begin{cases} 0 & \text{for } x < 0, \\ 1 & \text{for } x > 0. \end{cases} \tag{10.29}$$

It can be shown (see, e.g., [14]) that the definitions of convergence in probability and convergence in distribution (for continuous distributions) are equivalent.

The concepts of convergence defined above are weaker than the following one. We say that the sequence of random variables X_j, $j = 1, 2, \ldots$, converges to zero *with probability 1* (or: *almost everywhere*, *almost certainly*, *almost surely*) if it satisfies the condition

$$P\left(\omega : \lim_{j \to \infty} X_j(\omega) = 0\right) = 1. \tag{10.30}$$

The convergence in probability of the sequence X_j, $j = 1, 2, \ldots$, follows from its convergence with probability equal to 1.

For random variables possessing a finite second-order moment we can define the concept of *mean-square convergence*. We say that the sequence X_j, $j = 1, 2, \ldots$, converges to zero in a mean-square sense if it satisfies the condition

$$\lim_{j \to \infty} E\{X_j^2\} = 0. \tag{10.31}$$

For random variables with finite second-order moments, the mean-square convergence results from the convergence with probability equal to 1. The stochastic convergence results from the mean-square convergence. The last fact is a consequence of the following Tchebyshev inequality (see, e.g., [14]).

For an arbitrary random variable X, a non-negative measurable function λ, and for every $\epsilon \geq 0$, the following inequality is satisfied:

$$P\left(X \in A_\epsilon\right) \leq \frac{E\{\lambda(X)\}}{\epsilon}, \tag{10.32}$$

where $A_\epsilon = \{x : \lambda(x) \geq \epsilon\}$.

One can also consider a generalization of the mean-square convergence, *the convergence in the mean of order p*, reasonable for the sequences of random variables possessing finite p-th moment. The sequence X_j, $j = 1, 2, \ldots$, tends to zero in the mean of order p if

$$\lim_{j \to \infty} E\{X_j^p\} = 0. \tag{10.33}$$

10.3.2 The Law of Large Numbers

A class of limit theorems called laws of large numbers enables us to calculate fundamental numerical parameters characterizing the sums of independent, identically

distributed random variables, provided the number of elements of the sum tends to infinity.

Example 10.3 (The Bernoulli law of large numbers) Let X_j, $j = 1, 2, \ldots$ be a sequence of independent random variables which have the binomial distribution $b(n, p)$ (see (3.31)), taking values from the set $\{1, 2, \ldots, n\}$. Define a new sequence of random variables Y_n by

$$Y_n = X_n - np, \tag{10.34}$$

where p is the probability of a success in the Bernoulli trial (see Sect. 3.4.1). Then, for every $\epsilon > 0$ the following condition is satisfied:

$$\lim_{n \to \infty} P\left(|Y_n| > \epsilon\right) = 0, \tag{10.35}$$

that is, the sequence Y_n tends to zero in probability. This means that if the number n of the Bernoulli trials grows infinitely, then the ratio of the number of successes to the number of trials tends to the probability for success p in a single trial.

The law of large numbers enables us, in many situations of practical applications, to construct estimators of parameters or the state of systems (more precisely: their mathematical models) using experimental data (the measured values of appropriately chosen quantities), see Sect. 9.3.1.

10.3.3 The Central Limit Theorems

The central limit theorems present the conditions which must be satisfied to obtain a Gaussian random variable by summing up a sequence of independent random variables possessing certain properties.

Example 10.4 (The de Moivre-Laplace theorem) Let X_n, $n = 1, 2, \ldots$ be a sequence of independent random variables with the binomial distribution $b(n, p)$ (see (3.31)), taking one of the values from the set $\{1, 2, \ldots, n\}$. Consider the sequence Y_n, $n = 1, 2, \ldots$ of the normalized random variables

$$Y_n = \frac{X_n - np}{\sqrt{np(1 - p)}}. \tag{10.36}$$

Let F_n be the sequence of the cumulative distribution functions of random variables Y_n. If $0 < p < 1$, then for every real number y the following condition holds:

$$\lim_{n \to \infty} F_n(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{y^2/2} dy. \tag{10.37}$$

This means that the sequence of the normalized random variables (10.36) converges (by convergence of the cumulative distribution functions) to a certain normalized Gaussian random variable.

Let us note that the random variables X_n, $n = 1, 2, \ldots$, of the binomial distribution are in fact sums of independent Bernoulli random variables taking the value 1 with probability p and the value 0 with probability $1 - p$. This means that the de Moivre-Laplace theorem can be interpreted as the limit theorem for the sums of the Bernoulli random variables. Analogous limit theorems can be proved for sums of independent random variables with other distributions.

Example 10.5 (The Lindeberg-Levy theorem) Consider the sequence of independent, identically distributed random variables Z_j, $j = 1, 2, \ldots$, with the mean value and the variance equal, respectively, to

$$E\{Z_j\} = m, \qquad E\{(Z_j - m)^2\} = \sigma^2. \tag{10.38}$$

Define the new sequence of random variables Y_n, $n = 1, 2, \ldots$, by

$$Y_n = \frac{\sum_{j=1}^{n} Z_j - nm}{\sigma\sqrt{n}}. \tag{10.39}$$

Then, the sequence of the probability distribution functions F_n, $n = 1, 2, \ldots$, of the random variables Y_n, $n = 1, 2, \ldots$, satisfies for every value of y the following equation (cf. [14]):

$$\lim_{n \to \infty} F_n(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{y^2/2} dy. \tag{10.40}$$

The central limit theorem can be essential in many practical problems. It confirms, e.g., the fact that measurement errors, being a result of a number of independent random factors, have the normal distribution. It allows us also to apply statistical tests to verify that the model parameters estimated from the measured experimental data have acceptable values.

The limit theorems of probability theory have many generalizations. In the literature one can find theorems in which various concepts of convergence are used, the summed-up random variables are dependent ("weakly dependent" in some sense) or some of their moments are unbounded. More information on this subject can be found in [14, 18, 25, 39, 58].

10.4 Elements of Mathematical Statistics

Most of the numerical parameters describing errors of measurements, e.g., the mean value, the variance, and moments of higher order, are calculated from the experimental data obtained from measurements. Generally, it is hopeless to measure all the possible values of a given quantity. This means that the values of parameters are obtained on the basis of an arbitrary set of measurements, generated at random according to some probabilistic law. Thus, the calculated value is a random variable which depends on the event of collecting measurements. Mathematically rigorous

calculation of errors and verification of their validity requires application of statistical methods, especially from the theory of estimation and the methods of testing statistical hypotheses.

10.4.1 Estimators

First of all we should define the estimator of a quantity searched for. In the literature there are a number of abstract definitions of the estimator (comp., e.g., [27]). For our purposes it is enough to say that the estimator $\widehat{p}$ of some unknown parameter p is such a function of experimental data which makes it possible to choose the best approximation of the real value of p (choosing the value from the set of all assumed choices). Of course, one should determine what "the best" means in the above definition. The precise formulation exceeds the scope of this chapter. Therefore, we restrict ourselves to some examples of estimators used in this book.

Expression (1.1) is the estimator of the mean value, based on the data obtained from an experiment (a sample of measurements). This means that if we measure some independent random variables X_i with the same distributions of the (unknown) mean value m, then possessing a sample of n measurements x_i we can assume $m = \overline{x}$. If we continue the measurements and then calculate $\overline{x}$ using more data points (some larger n), we will obtain another value of the estimator, probably closer to the real value of m. However, for a fixed number n of the measured experimental points, the best approximation of the parameter m is given by formula (1.1).

Analogously, if the random variables have a common variance σ^2, then its estimator can be given by expression (1.4a). It is a question whether this estimator is the best one. If in formula (1.4a) we substitute the known exact value m for x, then we will obtain a "good" estimator. However, if we use in (1.4a) the value $\overline{x}$ defined in (1.1), then the obtained estimator is *biased*, that is, its expected value is not equal to the real value of the parameter searched for. Using instead of the estimator (1.4a) the following one:

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(x_j - \overline{x}\right)^2, \tag{10.41}$$

we can show by a simple calculation that $E\{\hat{\sigma}^2\} = \sigma^2$, which means the new estimator (10.40) is *unbiased*. It is seen that for large values of n both expressions for the estimator of the variance are equivalent.

In Sect. 5.2.2 we presented a method of estimating the coefficients of linear regression, called Procedure Y. The criterion of choice of the best expression for the estimators of parameters A and B is the minimum of the mean square distance (5.24). Such an estimator is called the least squares estimator. Some other criterion, e.g., the minimization of the average sum of absolute values of distances of the sample points from the regression line would give different expressions for the estimators of A and B.

The examples presented here show that constructing the estimators is usually a complicated procedure, and it is restricted by many precise assumptions. Readers interested in details should study dedicated textbooks or monographs, e.g., [27] or [41].

10.4.2 Testing Statistical Hypotheses

As it was noted in the previous section, by constructing appropriate estimators we can calculate, using experimental data, the values of parameters of distributions of the observed random variables. After such calculations, however, we would like to know the reliability of the obtained results. It can happen that we made wrong assumptions concerning a certain type of measured random variable, the set of experimental data is too small or someone preparing the data made some changes which disturbed their random character (comp. [36]). Statistical tests can help us in detecting or excluding such incorrectness. In this section we will present a general idea of testing statistical hypotheses.

The *statistical hypothesis* is any conclusion drawn from experimental data which are considered random variables. If this conclusion is associated with parameters of the measured random variables (e.g., the mean value, the correlation coefficient), the hypothesis is called a parametric one. If it is connected with other properties of the distributions of random variables (e.g., the type of distribution, the dependence or independence of two random variables), the hypothesis is nonparametric. A procedure of verification whether a hypothesis H_0 is true or false is called the *statistical test*. One of the following decisions is usually the result of such a test: we accept the initial hypothesis H_0 or reject it. The only way to make a proper decision is to perform the exhaustive test, that is, to make all possible measurements and then make the decision. In general, such a procedure is impossible and usually the investigator has a limited number of data samples. Then the following four situations are possible (with a certain probability):

The hypothesis H_0 is false and we reject it. Probability of this event is called the *power of the test*.

The hypothesis H_0 is false but we accept it. This decision is called the *error of the 2nd kind*.

The hypothesis H_0 is true and we accept it.

The hypothesis H_0 is true but we reject it. This decision is called the *error of the 1st kind*. Some arbitrarily fixed maximum probability of this error is called the *significance level of the test*.

From many possible tests with a given level of significance, we usually choose the test which has the maximum power. Such a test is called the strongest one (on a certain level of significance).

To construct a statistical test we should build a certain function of the measured statistical data, called the *statistic*, the probability distribution of which, under the assumption that the tested hypothesis is true, is known. Such a function is a measure

of difference between the theoretically assumed distribution and the really estimated probability distribution of the investigated random variable (comp. [41]). To perform the test, one should fix some area, called the critical area, depending on the postulated level of significance and the power of the test. Now, if the value of the statistic calculated for the data measured in an experiment made for verification of the hypothesis belongs to the critical area, we reject the hypothesis H_0. If it is outside of the critical area, we accept this hypothesis.

In the literature there is reported a large number of tests which make it possible to verify various postulated properties of empirical probability distributions, both one-dimensional and multi-dimensional. Here we give the most important areas where the statistical tests enable us to verify the statistical hypotheses indispensable in engineering applications. They are:

Tests of randomness, which let us decide if a sample of measurements can be considered a sequence of realizations of random variables. They also let us verify if a random sample is representative in order to make other statistical tests.

Tests of goodness of fit, which allow us to determine the distribution of some sequence of measurements or to decide if two sequences of the measured data have the same probability distribution.

Tests of normality, which are a specific kind of tests of goodness, deciding if a sample of data can be considered to be normally distributed; they are especially important in practice so they have their own name.

Tests for parameters of a certain distribution (the parametric tests), which let us verify the reliability of parameters, e.g., moments, parameters of population, estimated from a series of measurements.

Tests of independence, which make it possible to detect, after measurements, if random variables (two or more) can be considered to be independent.

Examples of tests of each of the group listed and some other special tests may be found in dedicated monographs, see, e.g., [14, 28].

10.4.3 Confidence Intervals

The problems of estimation and testing of the statistical hypotheses are strongly connected with the concept of a confidence level and confidence intervals. As we know, a certain value of a parameter estimated from a series of measurements, that is, the estimator of the real value of this parameter, is a random variable. For any series of measurements it takes some different values, according to its specific probability distribution. This probability distribution can often be calculated, provided the distributions of the measurements are known. It is interesting to find an interval where the values of the estimator will occur with some a priori assumed probability. In such a case this assumed probability is called the confidence level and the corresponding interval is called the *confidence interval* (see e.g., [14, 41]).

Let us recall the three-sigma rule for the normal distribution. If we generate a random number according to $N(0, 1)$, that is, the normal distribution with a zero

mean and a unit variance, then it will be in the interval $(-3, 3)$ with probability 0.9973.

If we now generate independently two numbers such as those in the previous case, then the sum of their squares will have the $\chi^2(2)$ probability distribution. Hence, the sum of squares, with probability of 0.5, will lie inside the circle of radius 1.386 (comp. Table 5.2). This circle can be called the *confidence set* for the level of confidence equal to 0.5.

10.5 Bibliographical Notes for Future Studies

Error calculus is a mathematical tool which enables us to consider, among others, random properties of measurements of engineering objects. It has been an element of scientific experimental analysis in engineering problems. One can find many examples in the literature dealing with both the practical aspects of measurement technique and the mathematical foundations of error calculus or, in a wider sense, data analysis (see, e.g., [7, 10, 34, 35, 43, 55]). But the error analysis enables us to take into account only the most elementary effects of randomness in the behavior of structures or technological processes. They can be, for example, clearances in mechanisms, material imperfections of mass-produced elements, or some environmental disturbances during artillery shots. Observations of practically operating engineering structures show that their behavior fluctuates randomly in time and, as a consequence, can be appropriately described by stochastic processes (or random fields with space-time arguments).

Our purpose has not been a presentation of advanced stochastic analysis of engineering structures. We give, however, some notes for the readers who, after studies of error analysis, are interested in more advanced problems of probabilistic engineering mechanics. Below we provide the titles of several monographs and textbooks which can indicate new directions of studies.

We start with some general books on probability theory, where the foundations of probability theory [13, 14, 33] and associated field of mathematics like measure theory [4, 9] or combinatorics [11] are presented. We can also mention applications of probability theory in engineering problems, such as reliability of structures [12, 30, 31, 53].

A detailed analysis of errors should involve, as we mentioned, advanced statistical methods. Information on mathematical statistics, especially testing statistical hypotheses and estimation theory, may be found in [27, 28, 36, 41, 59].

As we have mentioned, the realistic description of work of engineering mechanisms and structures requires application of stochastic processes. The theory of stochastic processes is well developed both as an abstract theory and as a tool for analysis of real phenomena. Information on stochastic processes can be found in [17, 33, 39, 57, 60, 62] and on random fields in [49].

A specific class of stochastic processes are the solutions of so-called stochastic differential equations. Stochastic differential equations are the counterpart of

deterministic differential equations (ordinary or partial) in the case when the initial conditions, parameters, and excitations are random variables or stochastic processes. There is a large literature on the theory of stochastic differential equations [20, 42, 46] and in their applications in many fields of engineering such as mechanical engineering [45, 48, 50], random vibrations [5, 38, 52, 54], wave processes [47], and stability theory [44, 56].

We should also indicate some specific methods useful in stochastic mechanics such as stochastic finite element methods [16, 22], numerical methods [23, 24], and computer simulations of physical phenomena [8, 19, 29, 37, 61].

Returning to measurements and data analysis we should mention the case when the results of individual measurements are mutually dependent. They may then be modeled as specific stochastic processes, e.g., Markov processes [3, 39], but in most cases it is convenient to treat them as time series, which are stochastic processes with discrete time argument [6, 32].

Applications of the results of measurements and data analysis are a necessary part of many practical problems of engineering. We refer here to some of them: estimation, prediction, identification, filtering of signals and stochastic control [1, 2, 21, 26], stochastic optimization of processes and structures [63].

The list of books presented in this section is far from being complete. But we think that it may show readers the directions for their own studies.

References

1. Anderson, B.D.O., Moore, J.B.: Optimal Filtering. Dover, New York (2005)
2. Bensoussan, A.: Stochastic Control of Partially Observable Systems. Cambridge University Press, Cambridge (2004)
3. Bharucha-Reid, A.T.: Elements of the Theory of Markov Processes and Their Application, 2nd edn. Dover, New York (1997)
4. Billingsley, P.: Probability and Measure, 3rd edn. Wiley–Interscience, New York (1995)
5. Bolotin, V.V.: Random Vibrations of Elastic Systems. Martinus Nijhoff, The Hague (1984)
6. Box, G.E.P., Jenkins, G.M.: Time Series Analysis. Forecasting and Control, 4th edn. Wiley, New York (2008)
7. Brandt, S.: Statistical and Computational Methods in Data Analysis, 3rd edn. Springer, Berlin (1998)
8. Bratley, P., Fox, B.L., Schrage, L.E.: A Guide to Simulation, 2nd edn. Springer, New York (1987)
9. Burrill, C.W.: Measure, Integration and Probability. McGraw–Hill, New York (1972)
10. Cowan, G.: Statistical Data Analysis. Oxford Science Publications, New York (1998)
11. Constantine, G.M.: Combinatorial Theory and Statistical Design. Wiley Series in Probability and Mathematical Statistics, Applied Probability and Statistics. Wiley, New York (1987)
12. Ditlevsen, O., Madsen, H.O.: Structural Reliability Methods. Wiley, Chichester (1996)
13. Feller, W.: An Introduction to Probability Theory and its Applications, vols. 1, 2, 2nd edn. Wiley, New York (1970)
14. Fisz, M.: Probability Theory and Mathematical Statistics, 3rd edn. Krieger, Melbourne (1980)
15. Fox, J.: Applied Regression Analysis and Generalized Linear Models, 2nd edn. Sage, Thousand Oaks (2008)
16. Ghanem, R.G., Spanos, P.D.: Stochastic Finite Elements: A Spectral Approach. Dover, New York (2003), revised edn.

17. Gihman, I.I., Skorohod, A.V.: The Theory of Stochastic Processes I. Springer, Berlin (2004)
18. Gnedenko, B.W., Kolmogorov, N.N.: Limit Distributions for Sums of Independent Random Variables. Addison–Wesley, Reading (1964)
19. Hastings, C. Jr.: Approximation for Digital Computers. Princeton Univ. Press, Princeton (1955)
20. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. North-Holland, Amsterdam (1981)
21. Kaliampur, G.: Stochastic Filtering Theory. Springer, New York (1980)
22. Kleiber, M., Hien, T.D.: The Stochastic Finite Element Method. Basic Perturbation Technique and Computer Implementation. Wiley, Chichester (1992)
23. Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations. Springer, Berlin (2000), corrected edn.
24. Knuth, D.E.: Seminumerical Algorithms. The Art of Computer Programming, vol. 2, 3rd edn. Addison–Wesley, Reading (1998)
25. Kotulski, Z., Sobczyk, K.: Linear systems and normality. J. Stat. Phys. **24**(2), 359–373 (1981)
26. Kushner, H.: Introduction to Stochastic Control. Holt, Rinehart and Winston, New York (1971)
27. Lehmann, E.L., Casella, G.: Theory of Point Estimation, 2nd edn. Springer, Berlin (2003)
28. Lehmann, E.L., Romano, J.P.: Testing Statistical Hypotheses, 3rd edn. Springer, Berlin (2008)
29. Lipson, Ch., Sheth, N.J.: Statistical Design and Analysis of Engineering Experiments. McGraw–Hill, Kogakusha (1973)
30. Melchers, R.E.: Structural Reliability Analysis and Prediction. Wiley, Chichester (1999)
31. Murzewski, J.: Reliability of Engineering Structures. Arkady, Warsaw (1989)
32. Naidu, P.S.: Modern Spectrum Analysis of Time Series. CRC Press, Boca Raton (1996)
33. Papoulis, A.: Probability, Random Variables, and Stochastic Processes with Errata Sheet, 4th edn. McGraw–Hill, New York (2002)
34. Partington, J.R.: Interpolation, Identification, and Sampling. Clarendon Press, Oxford (1997)
35. Pugh, E.M., Winslow, G.H.: The Analysis of Physical Measurements. Addison–Wesley, Reading (1966)
36. Radhakrishna Rao, C.: Statistics and Truth: Putting Chance to Work, 2nd edn. World Scientific, Singapore (1997)
37. Ripley, B.D.: Stochastic Simulation. Wiley–Interscience, New York (2006)
38. Roberts, J.B., Spanos, P.D.: Random Vibration and Statistical Linearization. Dover, New York (2003)
39. Rosenblatt, M.: Random Processes. Oxford University Press, New York (1962)
40. Seber, G.A.F., Lee, A.J.: Linear Regression Analysis, 2nd edn. Wiley–Interscience, Singapore (2003)
41. Sachs, L.: Applied Statistics. A Handbook of Techniques, 2nd edn. Springer, New York (1984)
42. Schuss, Z.: Theory and Applications of Stochastic Differential Equations. Wiley, New York (1980)
43. Sivia, D.S.: Data Analysis. A Bayesian Tutorial, 2nd edn. Oxford Scientific Publications, New York (2006)
44. Skalmierski, B., Tylikowski, A.: Stability of Dynamical Systems. Polish Scientific Editors, Warsaw (1973)
45. Skalmierski, B., Tylikowski, A.: Stochastic Processes in Dynamics. Polish Scientific Editors/Martinus Nijhoff Publishers, Warsaw/The Hague (1982)
46. Skorohod, A.V.: Stochastic Equations for Complex Systems. Reidel, Dordrecht (1988)
47. Sobczyk, K.: Stochastic Wave Propagation. Polish Scientific Editors/Elsevier, Warsaw/Amsterdam (1984)
48. Sobczyk, K.: Stochastic Differential Equations with Applications to Physics and Engineering. Kluwer Academic, Dordrecht (1991)
49. Sobczyk, K., Kirkner, D.: Stochastic Modeling of Microstructures. Birkhäuser, Basel (2001)
50. Sobczyk, K., Spencer, B.F. Jr.: Random Fatigue: From Data to Theory. Academic Press, Boston (1992)

51. Socha, L.: Linearization Methods for Stochastic Dynamic Systems. Lecture Notes in Physics, vol. 730. Springer, Berlin (2008)
52. Soong, T., Grigoriu, M.: Random Vibration of Mechanical and Structural Systems. Prentice Hall, Englewood Cliffs (1993)
53. Spaethe, G.: Die Sicherheit tragender Baukonstruktionen, 2nd edn. Springer, Berlin (1992)
54. Sveshnikov, A.A.: Problems in Probability Theory, Mathematical Statistics and Theory of Random Functions. Dover, New York (1979)
55. Taylor, J.R.: An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, 2nd edn. University Science Books (1997)
56. Tylikowski, A.: Stochastic Stability of Continuous Systems. Polish Scientific Editors, Warsaw (1991)
57. van Kampen, N.G.: Stochastic Processes in Physics and Chemistry, 2nd edn. North-Holland, Amsterdam (2001)
58. vom Scheidt, J., Fellenberg, B., Wöhrl, U.: Analyse und Simulation stochastischer Schwingungssysteme. Teubner, Leipzig (1994)
59. Wald, A.: Sequential Analysis. Dover, New York (2004)
60. Wentzel, A.D.: A Course in the Theory of Stochastic Processes. McGraw–Hill, New York (1980)
61. Wieczorkowski, R., Zieliński, R.: Computer-based Random Numbers Generators. Scientific and Technological Editors, Warsaw (1997)
62. Wong, E.: Stochastic Processes in Engineering Systems. Springer, Berlin (1984)
63. Zieliński, R., Neumann, P.: Stochastische Verfahren zur Suche nach dem Minimum einer Funktion. Akademie-Verlag, Berlin (1983)

Solutions

Problems of Chap. 1

1.1 The table of lengths for the scale to 20 cm is

Interval	1–20	21–40	41–60	61–80	81–100
No.	75	14475	10950	1100	125

Fig. S.1 Histogram 1 in Problem 1.1
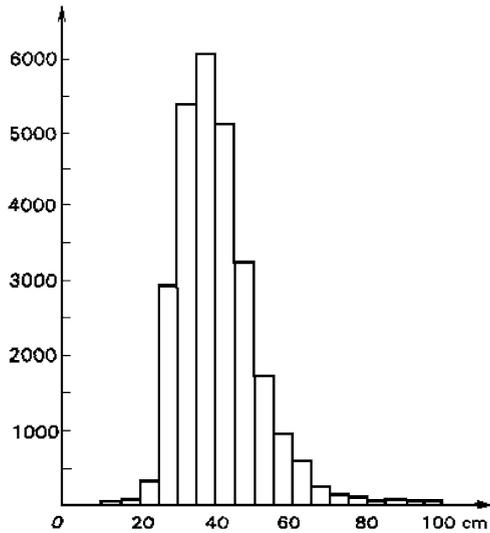
Fig. S.2 Histogram 2 in
Problem 1.1
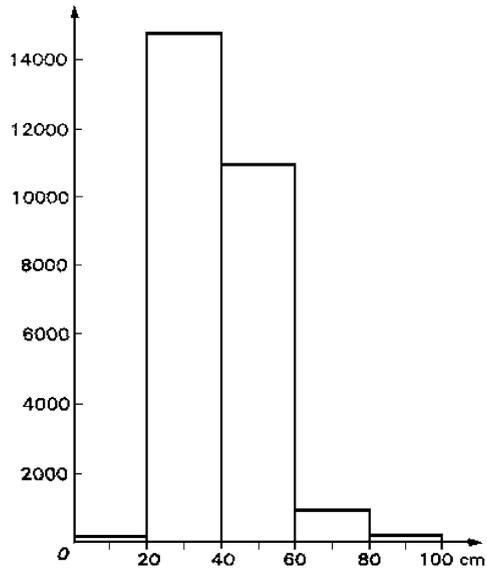
1.2

Interval	w_l	$\sum_{j=1}^{l} w_j$	Interval	w_l	$\sum_{j=1}^{l} w_j$
1−5	0	0	51−55	0.064	0.919
6−10	0	0	56−60	0.036	0.954
11−15	0.001	0.001	61−65	0.024	0.978
16−20	0.002	0.003	66−70	0.008	0.987
21−25	0.010	0.013	71−75	0.005	0.992
26−30	0.109	0.122	76−80	0.004	0.995
31−35	0.198	0.320	81−85	0.001	0.996
36−40	0.225	0.544	86−90	0.002	0.998
41−45	0.191	0.735	91−95	0.001	0.999
46−50	0.120	0.855	96−100	0.001	1.000

Interval	1−20	21−40	41−60	61−80	81−100
w_l	0.003	0.541	0.410	0.041	0.005
$\sum_{j=1}^{l} w_j$	0.003	0.544	0.954	0.995	1.000

Fig. S.3 Cumulative frequency distribution 1 in Problem 1.2
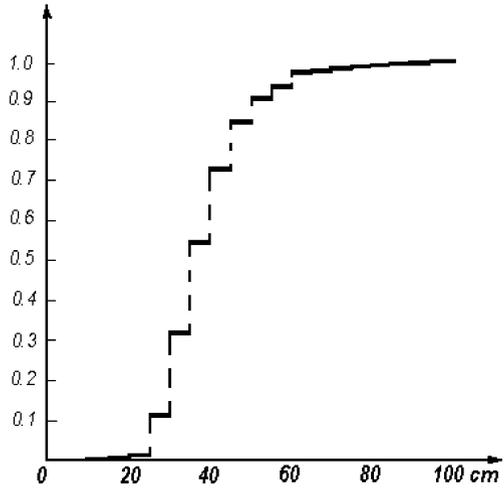
Fig. S.4 Cumulative frequency distribution 2 in Problem 1.2
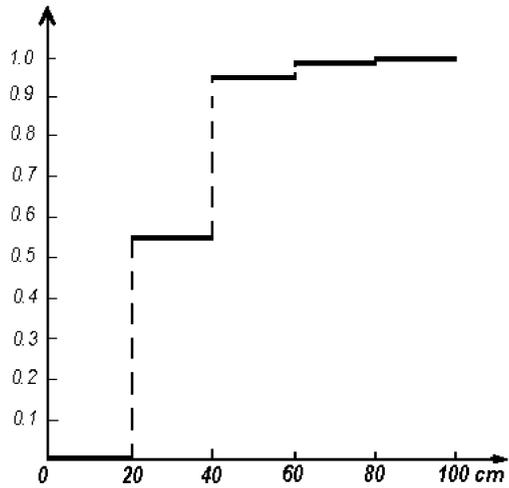
1.3 The parameters are given in the following table:

	Scale 5 cm	Scale 20 cm
The average value	42.94	50.07
The variance	103.34	146.39
The standard deviation	10.17	12.10

1.4 The parameters are given in the following table:

	Scale 5 cm	Scale 20 cm
The average deviation $\overline{d}$	7.79	11.07
The standard deviation s	10.17	12.10
$\frac{s}{\overline{d}}$	1.31	1.09

Problems of Chap. 2

2.1 The probability that the individual elements will work correctly is, respectively, 0.4, 0.5, 0.6, and 0.7. The probability that the system will work correctly is

$$P(A) = 0.4 * 0.5 * 0.6 * 0.7 = 0.084,$$

and the probability of a defect during one year of work is

$$P(A') = 1 - P(A) = 0.916.$$

2.2 First we find the probability of the complementary event that the sample will be accepted. This means that we should get a sequence of 10 correct elements, each with the probability, respectively, $\frac{190}{200}, \frac{189}{199}, \ldots, \frac{181}{191}$. The probability of the acceptance of the sample is

$$P(A') = \frac{190}{200} * \frac{189}{199} * \cdots * \frac{181}{191} = 0.5915,$$

and the probability of the rejection of the sample is

$$P(A) = 1 - P(A') = 0.4085.$$

2.3 The sample of 50 elements can be chosen in $\binom{100}{50}$ ways.[1] The sample will be accepted if it contains no defective elements $\left(\binom{95}{50}\right.$ choices$\left.\right)$, exactly one defective element $\left(\binom{95}{49}\binom{5}{1}\right.$ choices$\left.\right)$ or exactly two defective elements $\left(\binom{95}{48}\binom{5}{2}\right.$ choices$\left.\right)$. Thus, the probability of acceptance of the sample is

$$P(A) = \frac{\binom{95}{50} + \binom{95}{49}\binom{5}{1} + \binom{95}{48}\binom{5}{2}}{\binom{100}{50}}$$

$$= \frac{47 * 46 + 5 * 50 * 47 + 10 * 50 * 49}{8 * 99 * 97} = \frac{38412}{76824} = 0.5.$$

---

[1] Symbol $\binom{n}{k}$ is a number of k-elements combinations of elements of an n-element set and is defined as $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n*(n-1)*\cdots*(n-k+1)}{k*(k-1)*\cdots*2*1}$.

Problems of Chap. 3

3.1

$$c = \frac{1}{\int_0^\infty x^2 e^{-kx}dx} = \frac{k^3}{2},$$

$$F_X(x) = \int_0^x \frac{k^3}{2}x^2 e^{-kx}dx = 1 - \frac{1}{2}(k^2x^2 + 2kx + 2)e^{-kx},$$

$$P(X \in (0, 1/2k)) = F(1/2k) = 1 - \frac{1}{2}\left(\frac{1}{4} + 1 + 2\right)e^{-1/2} \approx 0.0144.$$

3.2

(a) $\quad f(x) = \begin{cases} cbx^{b-1}\exp[-cx^b] & \text{for } x > 0, \\ 0 & \text{for } x \le 0, \end{cases}$

(b) $\quad \alpha_p = \left[-\frac{1}{c}\ln(1-p)\right]^{1/b},$

(c) $\quad \left(\dfrac{b-1}{bc}\right)^{1/b}.$

3.3

$$A = \left[\int_0^1 x^{a-1}(1-x)^{b-1}dx\right]^{-1} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)},$$

$$E[X] = \frac{a}{a+b}, \qquad \text{Var}[X] = \frac{ab}{(a+b)^2(a+b+1)}.$$

3.4 Since the Laplace distribution is symmetric, $\bar{x} = 0$. Then,

$$\frac{1}{2} = P(|X - \bar{x}| < E) = \int_{-E}^{E} \frac{1}{2}e^{-|x|}dx = \int_0^E e^{-x}dx = 1 - e^{-E}$$

and $E = \ln 2 = 0.6931$.

3.5

$$f_{XY}(x, y) = \begin{cases} \frac{1}{(b-a)(d-c)} & \text{for } a \le x \le b, c \le y \le d, \\ 0 & \text{outside this rectangle} \end{cases}$$

$$F_{XY}(x, y) = F_X(x)F_Y(y),$$

where

$$F_X(x) = \begin{cases} 0 & \text{for } x < a, \\ \frac{x-a}{b-a} & \text{for } a \le x \le b, \\ 1 & \text{for } x > b, \end{cases}$$

$$F_Y(y) = \begin{cases} 0 & \text{for } y < c, \\ \frac{y-c}{d-c} & \text{for } c \le y \le d, \\ 1 & \text{for } y > d. \end{cases}$$

3.6

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,dy = C\sqrt{\frac{\pi}{\beta}}\,\exp\left[-\left(\alpha - \frac{\xi^2}{4\beta}\right)x^2\right],$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,dx = C\sqrt{\frac{\pi}{\alpha}}\,\exp\left[-\left(\beta - \frac{\xi^2}{4\alpha}\right)x^2\right].$$

The random variables are independent if

$$f_{XY}(x, y) = f_X(x) f_Y(y),$$

that is, if the condition

$$\frac{\sqrt{\alpha\beta}}{\pi C}\,\exp\left[-\frac{\xi^2}{4}\left(\frac{x^2}{\beta} - \frac{4xy}{\xi} + \frac{y^2}{\alpha}\right)\right] = 1$$

is satisfied. This is possible only if $\xi = 0$ and $C = \sqrt{\frac{\alpha\beta}{\pi}}$.

3.7 The joint probability density function of the random vector (X, Y) is

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma^2}\,\exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right].$$

The probability density function of the vector (R, Φ) can be calculated according to the formula

$$f_{R\Phi}(r, \varphi) = f_{XY}(x(r, \varphi), y(r, \varphi))\left|\frac{\partial(x, y)}{\partial(r, \varphi)}\right|,$$

where

$$\left|\frac{\partial(x, y)}{\partial(r, \varphi)}\right| = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{vmatrix} = r.$$

Finally,

$$f_{R\Phi}(r, \varphi) = \frac{r}{2\pi\sigma^2}\,\exp\left[-\frac{r^2\cos^2\varphi + r^2\sin^2\varphi}{2\sigma^2}\right]$$

$$= \frac{r}{2\pi\sigma^2}\,\exp\left[-\frac{r^2}{2\sigma^2}\right].$$

It is seen that the random variables R and Φ are independent, because

$$f_{R\Phi}(r, \varphi) = f_R(r) f_\Phi(\varphi),$$

where

$$f_R(r) = \frac{r}{\sigma^2} \exp\left[-\frac{r^2}{2\sigma^2}\right], \quad r \geq 0,$$

is the probability density function of the Rayleigh distribution, and

$$f_\Phi(\varphi) = \frac{1}{2\pi}, \quad 0 \leq \varphi < \pi,$$

is the probability density function of the uniform distribution.

3.8 For the Pareto distribution with parameters $a > 0$ and $p > 0$ the entropy (with base e) is

$$H = \log\left(\frac{p}{a}\right) - \frac{1}{p} - 1.$$

3.9 We repeat the procedure of Example 3.7 of the inversion of the cumulative distribution function, now for the Pareto distribution. Let $\{y_n, \; n = 1, 2, \ldots\}$ be a sequence of independent realizations of random variable Y with a distribution uniform along the interval $[0, 1]$. In such a case the sequence $\{x_n = F^{-1}(y_n), n = 1, 2, \ldots\}$ represents independent realizations of a random variable with the Pareto distribution provided that $y = F(x)$ is the cumulative distribution function of the Pareto distribution with the parameters $a > 0$ and $p > 0$,

$$y = F(x) = \begin{cases} 0 & \text{for } x < a, \\ 1 - [\frac{a}{x}]^p & \text{for } x \geq a, \end{cases}$$

and the function $x = F^{-1}(y)$ is the function inverse to it:

$$x = F^{-1}(y) = a\left[1 - y\right]^{-\frac{1}{p}} \quad \text{for } y \in [0, 1].$$

Problems of Chap. 4

4.1

$$R = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2},$$

$$\sigma_R^2 = \left[\left(\frac{\partial R}{\partial x_A}\right)^2 + \left(\frac{\partial R}{\partial x_B}\right)^2 + \left(\frac{\partial R}{\partial y_A}\right)^2 + \left(\frac{\partial R}{\partial y_B}\right)^2\right]\sigma^2,$$

or in numbers, comp. Example 3.5,

$$\sigma_R^2 = \left[(-0.857)^2 + 0.857^2 + (-0.514)^2 + 0.514^2\right]\sigma^2 = 1.997\sigma^2.$$

4.2

$$x = r\cos\varphi + \sqrt{l^2 - r^2\sin^2\varphi},$$

$$\sigma_x^2 = \left(\frac{\partial x}{\partial l}\right)^2\sigma_l^2 + \left(\frac{\partial x}{\partial r}\right)^2\sigma_r^2 + \left(\frac{\partial x}{\partial\varphi}\right)^2\sigma_\varphi^2,$$

or in numbers, comp. Example 3.6,

$$\sigma_x^2 = (1.037)^2\,\sigma_l^2 + (0.5126)^2\,\sigma_r^2 + (-27.047)^2\,\sigma_\varphi^2.$$

4.3

$$x = r\cos\varphi + \sqrt{l^2 - (h + r\sin\varphi)^2},$$

$$\frac{\partial x}{\partial l} = 1.08 \text{ mm/mm}, \qquad \frac{\partial x}{\partial r} = 0.41 \text{ mm/mm},$$

$$\frac{\partial x}{\partial h} = -0.42 \text{ mm/mm}, \qquad \frac{\partial x}{\partial\varphi} = -30.0 \text{ mm/rad}.$$

4.4
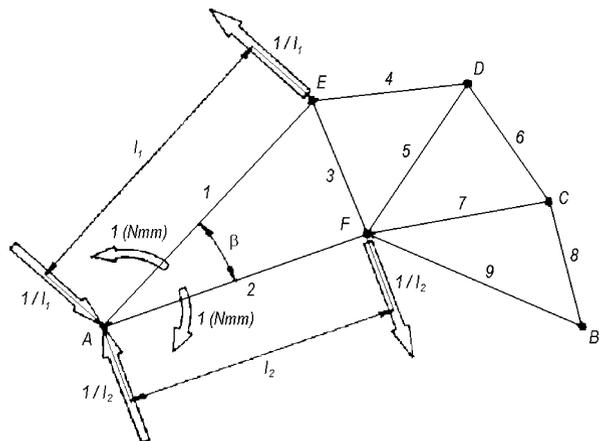
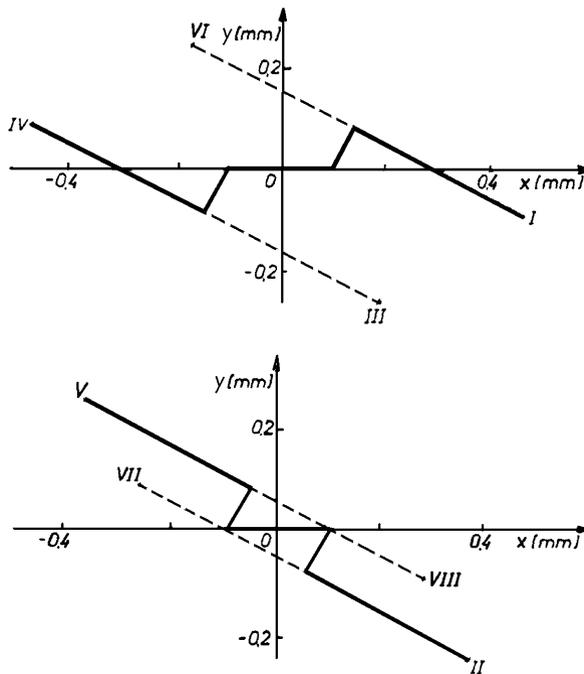**Fig. S.5** Graphical illustration of the solution of Problem 4.4
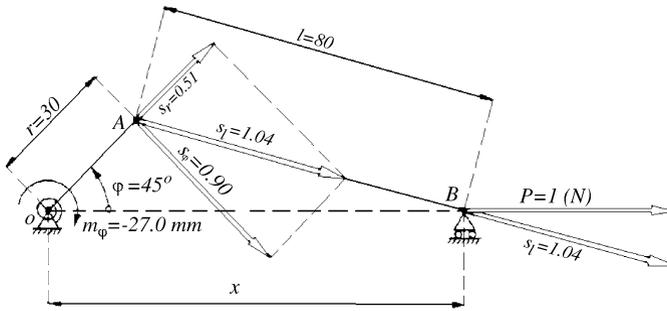
Fig. S.6 Graphical illustration of the solution of Problem 4.5

4.5

$$s_l \equiv \frac{\partial x}{\partial l} = 1.04, \qquad s_r \equiv \frac{\partial x}{\partial r} = 0.51, \qquad m_\varphi \equiv \frac{\partial x}{\partial \varphi} = -27 \text{ mm/rad.}$$

Problems of Chap. 5

5.1 We calculate:

$$\bar{x} = \frac{1}{20} \sum_{j=1}^{20} x_j = 6.08, \qquad \bar{y} = \frac{1}{20} \sum_{j=1}^{20} y_j = 232.45,$$

$$\text{cov}(X, Y) = \frac{1}{20} \sum_{j=1}^{20} (x_j - \bar{x})(y_j - \bar{y}) = 386.704,$$

so the random variables X and Y are strongly correlated.

5.2 We calculate

$$\bar{x} = \frac{1}{20} \sum_{j=1}^{20} x_j = 1.599, \qquad \bar{y} = \frac{1}{20} \sum_{j=1}^{20} y_j = 1.256,$$

$$\sigma_X^2 = \frac{1}{20} \sum_{j=1}^{20} (x_j - \bar{x})^2 = 0.0708, \qquad \sigma_Y^2 = \frac{1}{20} \sum_{j=1}^{20} (y_j - \bar{y})^2 = 0.0036,$$

$$\text{cov}(X, Y) = \frac{1}{20} \sum_{j=1}^{20} (x_j - \bar{x})(y_j - \bar{y}) = 0.005456,$$

$$\begin{bmatrix} \sigma_X^2 & \text{cov}(X, Y) \\ \text{cov}(X, Y) & \sigma_Y^2 \end{bmatrix} = \begin{bmatrix} 0.0708 & 0.005456 \\ 0.005456 & 0.0036 \end{bmatrix},$$

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = 0.034.$$

The principal axes of the covariance tensors are rotated through the angle $\beta = |\frac{1}{2}\tan^{-1}\frac{2\text{cov}(X,Y)}{\sigma_Y^2 - \sigma_X^2})| = 4.61°$ with respect to the axes of the coordinate system. We take the positive value of the rotation angle because the random variables X and Y are positively correlated.

5.3 We calculate

$$\bar{x} = \frac{1}{20}\sum_{j=1}^{20} x_j = 172.75, \qquad \bar{y} = \frac{1}{20}\sum_{j=1}^{20} y_j = 4.62,$$

$$\sigma_X^2 = \frac{1}{20}\sum_{j=1}^{20} (x_j - \bar{x})^2 = 438.6875, \qquad \sigma_Y^2 = \frac{1}{20}\sum_{j=1}^{20} (y_j - \bar{y})^2 = 0.5116,$$

$$\text{cov}(X, Y) = \frac{1}{20}\sum_{j=1}^{20} (x_j - \bar{x})(y_j - \bar{y}) = 10.995,$$

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = 0.7339.$$

Procedure Y:

$$A = 0.0251, \qquad B = 0.290,$$

and the angle of inclination is $\beta = \tan^{-1} A = 1.43°$.

Procedure X:

$$C = 24.49, \qquad D = 73.46,$$

and the angle of inclination is $\beta = \tan^{-1}\frac{1}{C} = 2.34°$.

The method of moments gives

$$T_J = \begin{bmatrix} 8773.75 & 219.9 \\ 219.9 & 10.232 \end{bmatrix}$$

and the angle of inclination is $\beta = \frac{1}{2}\tan^{-1}\frac{2J_{x_0 y_0}}{J_{y_0} - J_{x_0}} = 1.44°$. We see that the method of moments gives a result between the results of Procedure X and Procedure Y.

5.4 We find

$$\sigma_X = \sqrt{2}, \qquad \sigma_Y = 2, \qquad \rho_{XY} = \frac{1}{2\sqrt{2}}$$

and, according to formula (5.62),

$$f(x, y) = \frac{1}{2\pi\sqrt{7}}\exp\left[-\frac{1}{7}\left(2x^2 - xy + y^2\right)\right].$$

5.5 We find the standard deviations and the correlation coefficient of the random variables X and Y and then the general equations of the ellipses of constant probability:

(a) $\quad \sigma_X = \sqrt{2}, \qquad \sigma_Y = 2, \qquad \rho_{XY} = 0,$

$$\frac{x^2}{2} + \frac{y^2}{4} = \lambda^2,$$

(b) $\quad \sigma_X = \sqrt{2}, \qquad \sigma_Y = 2, \qquad \rho_{XY} = \frac{-1}{2\sqrt{2}},$

$$\frac{8}{7}\left[\frac{x^2}{2} + \frac{xy}{4} + \frac{y^2}{4}\right] = \lambda^2,$$

(c) $\quad \sigma_X = \sqrt{2}, \qquad \sigma_Y = 2, \qquad \rho_{XY} = \frac{1}{2\sqrt{2}},$

$$\frac{8}{7}\left[\frac{x^2}{2} - \frac{xy}{4} + \frac{y^2}{4}\right] = \lambda^2.$$

The principal axes of the covariance tensors are:

(a) parallel to the axes of the coordinate system; (b) rotated through the angle $-22.5°$; (c) rotated through the angle $22.5°$ with respect to the axes of the coordinate system.

5.6 Since the random variables X and Y are independent, their joint probability density function is the product of the one-dimensional probability density functions,

$$f_{XY}(x, y) = \frac{1}{4\pi} \exp\left[-\frac{x^2}{2} - \frac{y^2}{8}\right].$$

The random variables (U, V) are expressed by the random variables (X, Y) as

$$U = X \cos 30° + Y \sin 30° = \frac{\sqrt{3}}{2}X + \frac{1}{2}Y,$$

$$V = -X \sin 30° + Y \cos 30° = -\frac{1}{2}X + \frac{\sqrt{3}}{2}Y.$$

The mean values and the covariance tensor of the new random variables (U, V) are

$$\begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} \sigma_U^2 & \text{cov}(U, V) \\ \text{cov}(U, V) & \sigma_V^2 \end{bmatrix} = \begin{bmatrix} 1\frac{3}{4} & \frac{3\sqrt{3}}{4} \\ \frac{3\sqrt{3}}{4} & 3\frac{1}{4} \end{bmatrix},$$

and the standard deviations and the correlation coefficient are, respectively,

$$\sigma_U = \frac{\sqrt{7}}{2}, \qquad \sigma_V = \frac{\sqrt{13}}{2}, \qquad \rho_{UV} = \frac{\frac{3\sqrt{3}}{4}}{\frac{\sqrt{7}}{2}\frac{\sqrt{13}}{2}} = \frac{3\sqrt{3}}{\sqrt{91}} \approx 0.5447.$$

Finally, we obtain the probability density function of the random vector (U, V),

$$f_{UV}(u, v) = \frac{1}{4\pi} \exp\left\{-\frac{91}{128}\left[\frac{4u^2}{7} - \frac{3\sqrt{3}uv}{2} + \frac{4v^2}{13}\right]\right\}.$$

We observe that the random variables U and V are correlated (they are statistically dependent).

5.7 We consider the normalized random variables

$$U = \frac{X - 1000}{10} \quad \text{and} \quad V = \frac{Y}{\sqrt{50}}.$$

According to (5.56), the sum of their squares has the chi-squared distribution with two degrees of freedom, so outside of the ellipse

$$\left(\frac{X - 1000}{10}\right)^2 + \frac{Y^2}{50} = \chi^2_{0.001}(2) = 13.855$$

the probability of the hit point is less than 0.001.

Problems of Chap. 6

6.1 The polygon of positioning tolerance limits is bounded by three pairs of parallel straight lines. The equations of these lines are:

$$y = \pm 0.2606,$$
$$y - 1.732x = \pm 0.8692,$$
$$y + 0.577x = \pm 0.1732.$$

The polygon is shown in the figure.
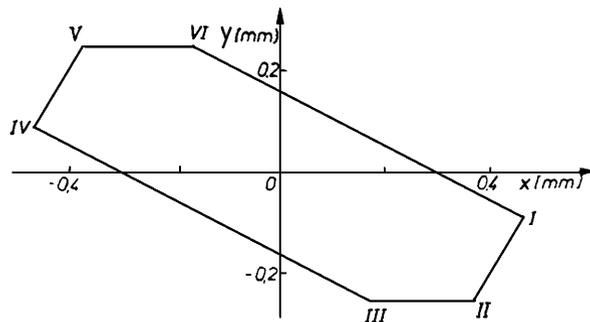
Fig. S.7 Graphical illustration of the solution of Problem 6.1

6.2 Eight combinations of joint positioning tolerance limits are collated in the table:

Variant	I	II	III	IV	V	VI	VII	VIII
Δa	+0.1	+0.1	−0.1	−0.1	−0.1	+0.1	+0.1	−0.1
Δb	+0.1	−0.1	−0.1	−0.1	+0.1	+0.1	−0.1	+0.1
$\Delta\theta$	−2	−2	−2	+2	+2	+2	+2	−2

All eight Williot's diagrams are shown in Fig. S.8. The end point of each diagram is denoted by the number of the corresponding variant of positioning tolerance limits combination.

Compare the Williot's diagrams with the tolerance limits polygon in Problem 6.1.

6.3 The polygon for $\theta = 0°$ is bounded by two pairs of parallel straight lines determined by the equations

$$y = \pm0.348, \qquad x = \pm0.200.$$

For $\theta = 90°$ the polygon is bounded also by two pairs of straight lines

$$y = \pm0.100, \qquad x = \pm0.448.$$

Fig. S.8 Graphical illustration of the solution of Problem 6.2

Fig. S.9 Graphical
illustration of the solution of
Problem 6.3
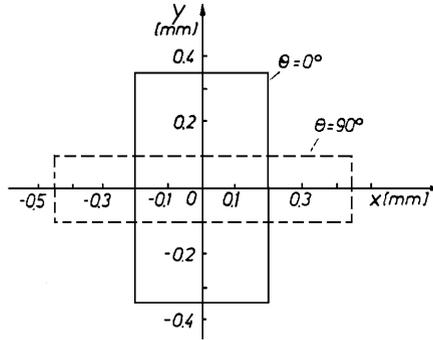
Fig. S.10 Graphical
illustration of the solution of
Problem 6.4
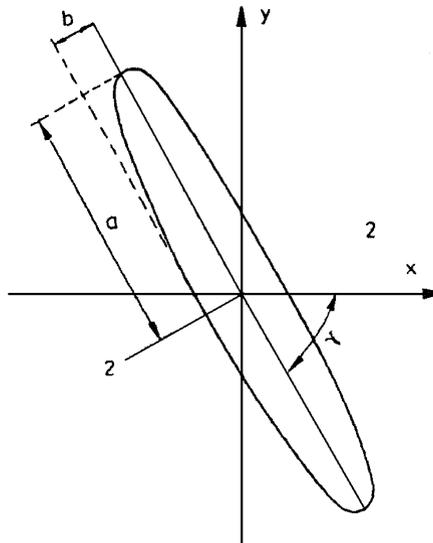
The two polygons are shown in Fig. S.9.

Note how strongly the shape and dimensions of the polygon of the positioning accuracy depend on the position of the mechanism.

6.4 The ellipse of probability concentration is determined by the equation

$$x^2 + xy + 0.293y^2 = 0.146\lambda^2 l^2 \sigma_\theta^2 = 32.0.$$

The ellipse is shown in Fig. S.10.

6.5 The lengths of the semi-axes of the ellipse are

$$a = 17.51 \text{ mm}, \qquad b = 10.75 \text{ mm}.$$

The ellipse and the polygon are shown in Fig. S.11.
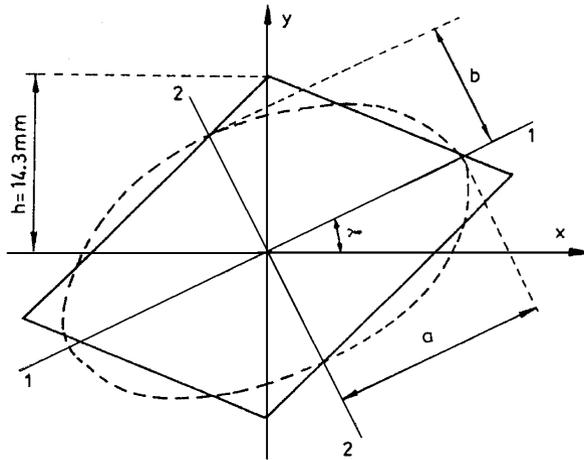
Fig. S.11 Graphical illustration of the solution of Problem 6.5

Problems of Chap. 7

7.1 We observe that random variables (X, Y) and Z are independent and

$$f_{XYZ}(x, y, z) = f_{XY}(x, y) f_Z(z),$$

where

$$f_{XY}(x, y) = \frac{1}{2\pi\sqrt{3}} \exp\left\{-2\left[\frac{(x-4)^2}{6} - \frac{(x-4)y}{2} + \frac{y^2}{2}\right]\right\}$$

and

$$f_Z(z) = \frac{1}{2\sqrt{2\pi}} \exp\left[-\frac{(z-2)^2}{8}\right].$$

7.2 We must calculate the elements of the inverse matrix $[\Lambda_{jk}] = [\lambda_{jk}]^{-1}$. First we calculate the algebraic complements of the determinant of $\Delta = |\lambda_{jk}|$,

$$\Delta_{11} = \begin{vmatrix} 10 & 6 \\ 6 & 9 \end{vmatrix} = 54, \qquad \Delta_{12} = -\begin{vmatrix} 4 & 6 \\ 2 & 9 \end{vmatrix} = -24,$$

$$\Delta_{13} = \begin{vmatrix} 4 & 10 \\ 2 & 6 \end{vmatrix} = 4, \qquad \Delta_{22} = \begin{vmatrix} 11 & 2 \\ 2 & 9 \end{vmatrix} = 95,$$

$$\Delta_{23} = -\begin{vmatrix} 11 & 4 \\ 2 & 6 \end{vmatrix} = -58, \qquad \Delta_{33} = \begin{vmatrix} 11 & 4 \\ 4 & 10 \end{vmatrix} = 94,$$

and then the determinant

$$|\lambda_{jk}| = \begin{vmatrix} 11 & 4 & 2 \\ 4 & 10 & 6 \\ 2 & 6 & 9 \end{vmatrix} = 11\Delta_{11} + 4\Delta_{12} + 2\Delta_{13} = 506.$$

Since $\Delta_{jk} = \Delta_{kj}$, the corresponding terms under the exponent in the probability density function are equal and the density can be written as

$$f(x, y, z) = \frac{1}{2\pi\sqrt{1012\pi}} \exp\left\{ -\frac{1}{1012}[54(x+3)^2 \right.$$
$$- 48(x+3)(y+2) + 8(x+3)(z-1) + 95(y+2)(y+2)$$
$$\left. - 116(y+2)(z-1) + 94(z-1)^2] \right\}.$$

7.3 We must calculate the elements of the inverse matrix $[\Lambda_{jk}] = [\lambda_{jk}]^{-1}$. First we calculate the algebraic complements of the determinant of $\Delta = |\lambda_{jk}|$,

$$\Delta_{11} = \begin{vmatrix} 16 & 6 & -2 \\ 6 & 4 & 1 \\ -2 & 1 & 3 \end{vmatrix} = 28, \qquad \Delta_{12} = -\begin{vmatrix} 3 & 6 & -2 \\ 1 & 4 & 1 \\ 0 & 1 & 3 \end{vmatrix} = -13,$$

$$\Delta_{13} = \begin{vmatrix} 3 & 16 & -2 \\ 1 & 6 & 1 \\ 0 & -2 & 3 \end{vmatrix} = 16, \qquad \Delta_{14} = -\begin{vmatrix} 3 & 16 & 6 \\ 1 & 6 & 4 \\ 0 & -2 & 1 \end{vmatrix} = -14,$$

$$\Delta_{22} = \begin{vmatrix} 15 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 3 \end{vmatrix} = 162, \qquad \Delta_{23} = -\begin{vmatrix} 15 & 3 & 0 \\ 1 & 6 & 1 \\ 0 & -2 & 3 \end{vmatrix} = -291,$$

$$\Delta_{24} = \begin{vmatrix} 15 & 3 & 1 \\ 1 & 6 & 4 \\ 0 & -2 & 1 \end{vmatrix} = 205, \qquad \Delta_{33} = -\begin{vmatrix} 15 & 3 & 0 \\ 3 & 16 & -2 \\ 0 & -2 & 3 \end{vmatrix} = 633,$$

$$\Delta_{34} = \begin{vmatrix} 15 & 3 & 1 \\ 3 & 16 & 6 \\ 0 & -2 & 1 \end{vmatrix} = -405, \qquad \Delta_{44} = -\begin{vmatrix} 15 & 3 & 1 \\ 3 & 16 & 6 \\ 1 & 6 & 4 \end{vmatrix} = 404,$$

and then the determinant itself

$$|\lambda_{jk}| = \begin{vmatrix} 15 & 3 & 1 & 0 \\ 3 & 16 & 6 & -2 \\ 1 & 6 & 4 & 1 \\ 0 & -2 & 1 & 3 \end{vmatrix} = 15\Delta_{11} + 3\Delta_{12} + \Delta_{13} = 397.$$

Since $\Delta_{jk} = \Delta_{kj}$, the corresponding terms under the exponent in the probability density function are equal and the density can be written as

$$f(x, y, z, u) = \frac{1}{4\pi^2\sqrt{397}} \exp\left\{-\frac{1}{794}\left[28(x+5)^2 - 26(x+5)(y-2)\right.\right.$$

$$+ 32(x+5)z - 28(x+5)(u-4) + 162(y-2)^2$$

$$- 582(y-2)z + 410(y-2)(u-4)$$

$$\left.\left.+ 633z^2 - 810z(u-4) + 404(u-4)^2\right]\right\}.$$

7.4 Since the random variables are independent, for numbers $a < b$, $c < d$, $e < f$ the following relation is true:

$$P(a \le X < b, c \le Y < d, e \le Z < f)$$
$$= P(a \le X < b)P(c \le Y < d)P(e \le Z < f).$$

The probabilities, for the normal distribution, may be expressed by the cumulative distribution function of the normalized distribution, $\Phi(x)$, given in Table 2.3, e.g.,

$$P(a \le X < b) = \Phi\left(\frac{b - m_X}{\sigma_X}\right) - \Phi\left(\frac{a - m_X}{\sigma_X}\right).$$

Finally, the probability P that the point is located within the walls of the rectangular prisms is the difference of the probability that it is inside the larger prism and the probability that it is inside the smaller one :

$$P = [\Phi(2) - \Phi(0)][\Phi(0.5) - \Phi(-0.5)][\Phi(0.667) - \Phi(0)]$$

$$- [\Phi(1) - \Phi(0.5)][\Phi(0) - \Phi(-0.24)][\Phi(0.333) - \Phi(0.167)]$$

$$\approx 0.0443 - 0.00114 = 0.043.$$

7.5 The set of the constant value of the probability density function contains the point coordinates which solve the equation

$$0.02 = \frac{\sqrt{3}}{16\pi\sqrt{\pi}} \exp\left\{-\frac{1}{8}\left[(2x^2 + 4(y-1)^2 - 2(y-1)(z+9) + (z+9)^2\right]\right\},$$

that is, the solutions of

$$-8\log\left(\frac{32\pi\sqrt{\pi}}{100\sqrt{3}}\right) = 2x^2 + 4(y-1)^2 - 2(y-1)(z+9) + (z+9)^2.$$

7.6 We consider the normalized random variables

$$U = \frac{X}{10}, \qquad V = \frac{Y}{10} \quad \text{and} \quad W = \frac{Z - 3000}{\sqrt{500}}.$$

The sum of their squares has the chi-squared distribution with three degrees of freedom, so outside the ellipsoid

$$\left(\frac{X}{10}\right)^2 + \left(\frac{Y}{10}\right)^2 + \left(\frac{Z-3000}{10\sqrt{5}}\right)^2 = \chi^2_{0.001}(3) = 16.268$$

the probability of the hit point is less than 0.001.

Problems of Chap. 8

8.1 The equations of the planes bounding the polyhedron are

$$1.329x + z = \pm 0.227,$$

$$1.058x + z = \pm 0.282,$$

$$x = \pm 0.806,$$

$$y = \pm 0.671.$$

The polyhedron is shown in Fig. S.12.
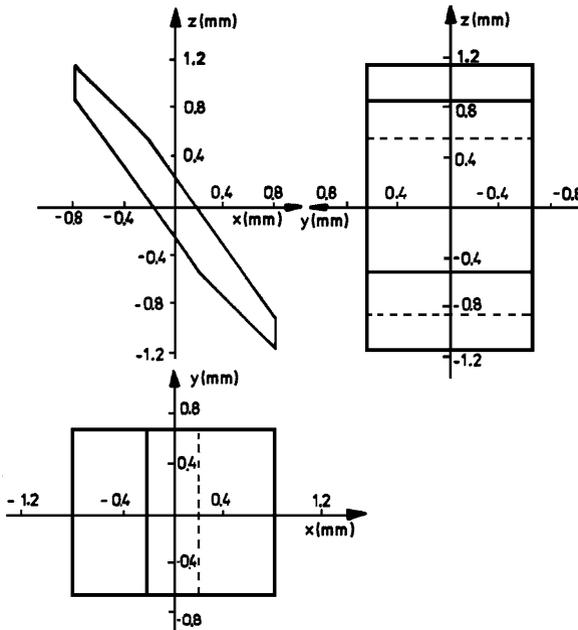
Fig. S.12 Graphical illustration of the solution of Problem 8.1

8.2 In the reference Cartesian coordinate system X, Y, Z, the position of the end-effector is determined by the following functions of the four independent variables $\theta_1, \theta_2, \theta_3, \theta_4$:

$$X = [a\cos\theta_2 + b\cos(\theta_2 + \theta_3) + c\cos(\theta_2 + \theta_3\theta_4)]\cos\theta_1,$$

$$Y = [a\cos\theta_2 + b\cos(\theta_2 + \theta_3) + c\cos(\theta_2 + \theta_3\theta_4)]\sin\theta_1,$$

$$Y = a\sin\theta_2 + b\sin(\theta_2 + \theta_3) + c\sin(\theta_2 + \theta_3\theta_4).$$

The values of partial derivatives of these functions after substituting the values of dimensions and angles are

–	$i = 1$	$i = 2$	$i = 3$	$i = 4$
$\frac{\partial X}{\partial \theta_i}$	0	−869.62	−519.62	0
$\frac{\partial Y}{\partial \theta_i}$	1156.22	0	0	0
$\frac{\partial Z}{\partial \theta_i}$	0	1156.22	550	250

Matrix (8.10) of the second-order moments takes on now the following form (in mm^2):

$$[k_{ij}] = \begin{bmatrix} 1.026 & 0 & -1.291 \\ 0 & 1.337 & 0 \\ -1.291 & 0 & 1.701 \end{bmatrix}.$$

The inverse matrix $[K_{ij}]$ reads

$$[K_{ij}] = \begin{bmatrix} 21.390 & 0 & 16.224 \\ 0 & 0.748 & 0 \\ 16.224 & 0 & 12.894 \end{bmatrix}.$$

For the assumed value of the probability $\alpha = 0.683$ we find from Fig. 7.1 that $\chi^2 = 3.50$. Substituting the derived data into (7.25) we find the ellipsoid of probability concentration

$$21.390x^2 + 0.748y^2 + 12.894z^2 + 32.448xz = 3.50.$$

Index