CHAPTER 1
Evaluation of measurement data*

Andrzej Zięba

Measurements of physical quantities can only be made with finite accuracy. The reason for this is the imperfection of measuring instruments and the imprecision of our senses participating in the observations. Providing only the result of the measurement itself is insufficient; the evaluation of measurements should also include a measure of their reliability, i.e., measurement uncertainty. The theory of measurement uncertainty (interchangeably called the calculus of measurement uncertainty) arose from the need to solve the above problems. It constitutes a professional skill necessary for everyone performing measurements. This document presents its most important results, illustrated with examples.

The theory of measurement uncertainty does not belong to the field of exact sciences; it is rather an approximate mathematical description of the imperfections of an experiment. Its methods and results are not limited to physics, but are the same or very similar for all experimental sciences. The international scientific community has long strived to harmonize the terminology and methods of uncertainty estimation. The result is the document *Guide to the Expression of Uncertainty in Measurement*, adopted in 1995 by the International Organization for Standardization (ISO) in agreement with a number of global scientific and technical organizations[1]. In this text, the document will be cited as the *Guide*. It is the most important, but not the only document of the international convention on the evaluation of measurement uncertainty, referred to as the GUM convention[2]. The terminology, symbolism, and uncertainty calculation methods used are consistent with the recommendations of the *Guide*. They constitute a professional skill necessary for everyone performing measurements.

Understanding the content related to mathematical statistics requires knowledge of a few elementary concepts. The necessary results of this branch of mathematics are briefly presented in Statistical Appendices A, B, and C (separate file).

1.1. Measurement error and its types

Until recently, the words "error" and "uncertainty" were used interchangeably. The *Guide* introduces a clear distinction between these concepts. In quantitative terms, by measurement error we mean the difference between the measured value x_i and the true value x_0,

$$\text{measurement error} = x_i - x_0. \tag{1}$$

In doing so, we make the tacit assumption that a true value exists. In practice, the true value can be identified with the result of a measurement performed using another, much more accurate method.

*The text is a modified chapter 1 of the script: ed. A. Zięba, PHYSICS LABORATORY of the Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, Part 1, Third revised edition. AGH Publishing House, Kraków 2002. Made available with the Author's permission. Translated by AI.

[1]Official Polish translation: *Wyrażanie Niepewności Pomiaru. Przewodnik.* Warsaw, Central Office of Measures 1999.

[2]Concise information on the history of the GUM convention and its current status can be found on the internet: http://physics.nist.gov/Uncertainty

The fundamental meaning of the word "error" is qualitative, as a name for the fact that the measured value differs from the true value. Figure 1.1 shows on a number line the mutual relationship between the true value x_0 and a series of values x_i obtained in an experiment, illustrating three types of measurement error.

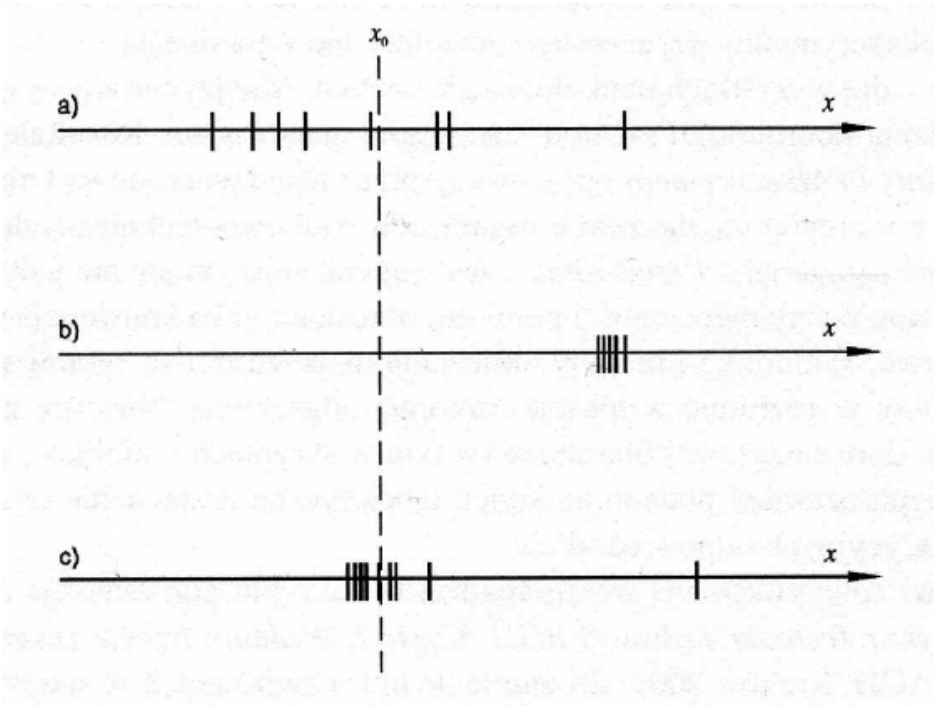


Figure 1.1: Mutual relationship of the true value x_0 and the set of measurement results (marked with dashes) on a number line for: a) random error, b) systematic error, c) a combination of random error and gross error. Figures c) and a) show, on different scales, the numerical results used in Examples 1.1 and 1.2

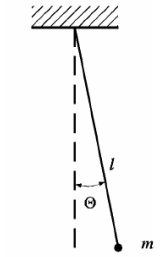
With a **random error**, we observe a scatter of measurement results around the true value (Fig. 1.1a). The result of each subsequent measurement is different, and there is an approximately equal chance of obtaining results larger as well as smaller than x_0.

What are the causes of the statistical scatter of measurement results in classical physics, where most phenomena are described by deterministic laws? Most often, the source of random error is the inaccuracy and randomness of human senses. When performing another measurement, a human will do it slightly differently, hence a statistical scatter of results will arise. For example, the results of measuring the fall time of a ball from a two-meter height using a stopwatch are characterized by a certain scatter, despite the fact that the stopwatch itself runs evenly. Noise generated within the measuring system itself and external disturbances can also be a source of statistical scatter of measurement results.

We deal with a **systematic error** when, upon repeating a measurement, the same difference occurs between the measured values and the true value, while the scatter of the results of individual measurements is small or non-existent. If, for example, we measure the value of a standard resistor (which is a realization of the true value) using an ohmmeter, we will find a systematic difference, which remains the same when the measurement is repeated.

We speak of a **gross error** (mistake) when the difference between the measurement result and the true value is large or drastically large. A gross error occurs as a result of inability to use a given instrument, mistakes in reading and recording results, etc. We deal with the case of a gross error occurring in a series of measurements when one of the results deviates significantly from

the others. Example 1.1 illustrates two of the most varied possibilities of making a gross error.



Rys. 1.2. Wahadło proste

Example 1.1. Pendulum - gross errors in period measurement
Examples are an integral part of the measurement uncertainty calculus lecture. Most of them (Examples 1.1, 1.2, 1.3, 1.5, 1.6, and 1.7) concern one simple experiment: investigating the motion of a simple pendulum.
A simple (or mathematical) pendulum is a point mass m suspended on a massless and inextensible string of length l (Fig. 1.2). A practical realization of this idealized object can be, for example, a metal ball suspended on ordinary sewing thread. When the swing angle Θ is small, the pendulum's period T_0 depends only on its length l and the acceleration due to gravity g,

$$T_0 = 2\pi\sqrt{\frac{l}{g}}. \tag{2}$$

A digital stopwatch was used to measure the pendulum's period. The duration of 50 periods was measured 9 times. The results were copied from the instrument's display as numbers:
103.88 104.16 105.26 104.03 103.90 103.97 103.85 104.02 103.92
The gravitational acceleration calculated based on these data turned out to be three times too small. The experimenter looked at the stopwatch and realized he had misread it: the first digit in the window (the one) indicates the number of minutes. The time for 50 periods expressed in seconds is actually:
63.88 64.16 65.26 64.03 63.90 63.97 63.85 64.02 63.92
Thus, the first gross error was detected and corrected.
Looking closely at the results shows that 8 numbers are clustered around 64 seconds, but the third result, 65.26 s, is more than a second larger. We begin to suspect that we measured 51 periods instead of 50. We are confirmed in this belief by the fact that the result 65.26 s differs from the others by a value close to one period. We reject the doubtful result.

1.2. What is measurement uncertainty?

The quantitative description of any phenomenon must begin with defining the measures that characterize it.

The measurement error defined by formula (1.1) does not constitute a measure of the accuracy of the measurement method, because a similar measurement, but performed with a different instrument, at a different time and place, will give a different value. Therefore, Δx_i is a random number whose value cannot be predicted, just as one cannot predict the result of a dice roll.

But regarding dice roll results, one can know that they are contained within a set of integers from 1 to 6. Similarly, the purpose of the uncertainty calculus is to at least approximately

estimate the scatter of measurement results, and the measure of this scatter is the measurement uncertainty.

The *Guide* adopts the following definition:

> **Measurement uncertainty is a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.**

The definition suggests that different measures of uncertainty are possible. To determine the uncertainty of a direct measurement, we use two measures: the primary one is the **standard uncertainty** $u(x)$, and the second measure useful in specific situations is the **limiting uncertainty** Δx[3].

In the case of limiting uncertainty Δx, we try to define an interval

$$x_0 - \Delta x < x_i < x_0 + \Delta x \tag{3}$$

within which all measurement results x_i, both currently performed and future ones, fall (Fig. 1.3).

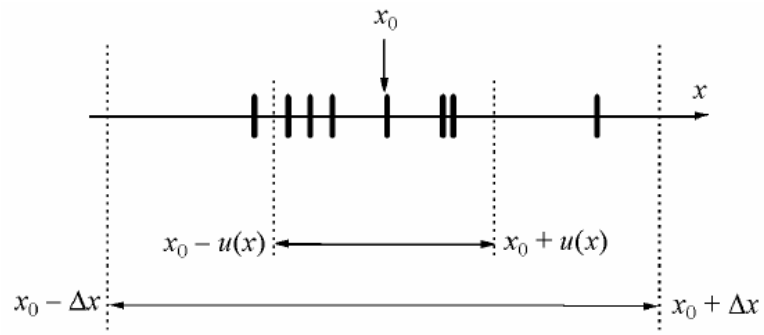


Figure 1.3: Illustration of standard uncertainty $u(x)$ and limiting uncertainty Δx

Limiting uncertainty is a deterministic measure because we claim that the true value is certainly contained within the interval $x_0 \pm \Delta x$. Limiting uncertainty is used in specific situations, e.g., as a measure of the accuracy of electrical measuring instruments.

The measure of measurement accuracy most widely used and recognized as the primary one by the *Guide* is the **standard uncertainty**. Its shortest definition is the sentence:

> **Standard uncertainty is the estimate of the standard deviation.**

Let's comment on the key words of this definition:

(i) The presented formulation conceals the assumption that the measurement result is a random variable, whose scatter is characterized by a parameter called the standard deviation. The standard deviation can be defined as the square root of the mean value of the squared difference between the measured and true values. (Formula (A6a) in Appendix A; basic information on this statistical parameter is provided there.)

[3]In the past, the word "error" was used in various meanings, also instead of the current "uncertainty". This is still done by those who do not know or do not accept the findings of the GUM convention. The traditional name is "limiting error" or "maximum error". The GUM convention is inconsistent regarding this term, as it introduces no name for this quantity, which is indispensable in measurement technique. In this study, the term "limiting uncertainty" is used to remain consistent with the postulated distinction between the meanings of the words "error" and "uncertainty".

(ii) We do not know the exact value of the standard deviation; standard uncertainty is its not very precise estimate (estimator, evaluation). The two basic methods of estimating the uncertainty of a direct measurement, Type A and Type B, will be discussed in sections 1.3 and 1.4.

Figure 1.3 graphically compares both measures of uncertainty. The standard uncertainty u is a measure of the average deviation of measurement results from the true value, thus a portion of the results (about 1/3) will be found outside the interval $(x_0 - u(x), x_0 + u(x))$.

In the remainder of the text, the word "uncertainty" without an adjective always means standard uncertainty. We denote standard uncertainty as $u(x)$. The symbol u comes from the English word *uncertainty*, and the symbol inside the parenthesis determines what the measured quantity is[4]. The advantage of the notation introduced by the *Guide* is clarity and the avoidance of indices. The ability to write the measured quantity in words, such as u(NaCl concentration), makes it easier to create measurement documentation.

Uncertainty u has a dimension, the same as the dimension of the measured quantity.

Relative uncertainty is the ratio of the (absolute) uncertainty to the measured quantity,

$$\frac{u(x)}{x}. \tag{4}$$

Relative uncertainty is a dimensionless quantity, often expressed in %. It gives a better idea of the measurement's accuracy than the absolute uncertainty u. It also allows comparing the uncertainties of physical quantities having different dimensions.

A qualitative concept associated with the word uncertainty is accuracy (of measurement). A more accurate measurement is a measurement with a smaller uncertainty.

1.3. Type A evaluation of uncertainty

It can be applied to measurements where a random error occurs. The simplest case is the analysis of a series of n observations $x_1, \ldots, x_i, \ldots, x_n$. We treat them as n realizations of a random variable with an expected value (which we identify with the true value x_0) and standard deviation σ (Appendix A). To calculate the approximate values of these parameters, we use the results of estimation theory (Appendix B). In most cases, we accept the value of the arithmetic mean as the measurement result x (closest to the unknown true value x_0)

$$x \equiv \overline{x} = \frac{1}{n} \sum x_i. \tag{5}$$

In formula 1.5, as in all formulas in chapter 1, the sum sign without indices means summation from $i = 1$ to n.

The measure of the scatter of measurement results is a quantity called the standard deviation estimator,

$$s_x = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}}. \tag{6}$$

The quantity s_x could be identified with the measurement uncertainty if we accepted any of the x_i values as its result. However, when calculating the mean, there is a partial compensation of deviations $x_i - \overline{x}$ of different signs, making it closer to the true value x_0 than the result of a single observation. Quantitatively, the standard deviation estimator of the mean $s_{\overline{x}}$ is $\sqrt{n}$ times smaller than the estimator s_x,

$$s_{\overline{x}} = \frac{s_x}{\sqrt{n}}. \tag{7}$$

[4]The adopted notation "illegally" uses the symbol of a mathematical function. Let us remember that $u(x)$ is a number, not a function. It is not possible, for example, to calculate the derivative du/dx!

Since we accept the mean as the measurement result, we identify the measurement uncertainty $u(x)$ with the standard deviation estimator of the mean, $u(x) \equiv s_{\overline{x}}$. Combining formulas (1.6) and (1.7a), we obtain

$$u(x) \equiv s_{\overline{x}} = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n(n-1)}}. \tag{8}$$

The quantities s_x and $s_{\overline{x}}$ are called estimators because, although calculated from unambiguous formulas, they are equal to the true values of the standard deviation only in the limit $n \to \infty$. When the number of measurements n is finite, the standard deviation of the mean—i.e., the measurement uncertainty—is known with finite, not very high precision (Table 1.1).

Table 1: Relative uncertainty of standard deviation evaluation s_x, $s_{\overline{x}}$ for a series of n measurements

Number of measurements	2	3	4	5	6	8	10	100
Uncertainty of evaluation	43%	38%	34%	31%	28%	25%	22%	7%

Repeating a measurement therefore brings two benefits: it reduces the uncertainty caused by random error and allows estimating the uncertainty. The question of how many measurements are worth performing cannot be answered unequivocally. It is considered that to determine the standard deviation, at least $5 \div 10$ measurements must be performed. This allows for estimating the uncertainty with an accuracy of the order of $30 \div 20\%$ (cf. Table 1.1). Moreover, for a series of, say, 9 measurements, the uncertainty of the mean is 3 times smaller than the uncertainty of a single measurement. It generally does not pay off to perform too large a number of measurements, because the increase in accuracy with increasing n is slow.

Performing a very small number of measurements, for example 2 or 3, makes sense as a repeatability check. We take the arithmetic mean as the measurement result, but to obtain the uncertainty, it is better to use a Type B evaluation (section 1.4).

Example 1.2. Calculation of the uncertainty in measuring the pendulum's period of oscillation (continuation of Example 1.1)
After rejecting the result $50T = 65.26$ s affected by a gross error and dividing the remaining values by 50, we obtain eight values of the pendulum's period (in seconds):
1.2776 1.2832 1.2806 1.2780 1.2794 1.2770 1.2804 1.2784
These values are shown on an appropriate scale in Figure 1.3.
The scheme for calculating the mean and standard uncertainties of a single measurement and the mean is as follows:

$$\begin{aligned} T_0 &= (1.2776 + 1.2832 + \cdots + 1.2784)/8 = 1.27933 \text{ s}, \\ s(T_0) &= \sqrt{\frac{(1.2776 - 1.27933)^2 + \cdots + (1.2784 - 1.27933)^2}{8-1}} = 0.0020 \text{ s}, \\ u(T_0) &= \frac{0.0020 \text{ s}}{\sqrt{8}} = 0.00071 \text{ s}. \end{aligned}$$

The calculation of T_0 and $u(T_0)$ using a calculator is discussed in subsection 1.11.

1.4. Type B evaluation of uncertainty

It is used when statistical analysis of a series of observations is not possible. For example, for systematic error or when a random error occurs but we only have one measurement result. Type B evaluation of uncertainty is based on the scientific judgment of the experimenter using all information about the measurement and the sources of its uncertainty.

Type B evaluation can utilize, among others:

- data from previous measurements,
- experience and knowledge about the instruments and measured objects,
- instrument manufacturer's information,
- uncertainties assigned to data taken from literature.

When this information is good, the accuracy of Type B evaluation is comparable to the accuracy of Type A evaluation. (Statistical evaluation is also not very accurate, cf. Table 1.1). In more difficult situations, Type B evaluation only allows estimating the order of magnitude of the uncertainty.

Most often, Type B evaluation concerns determining uncertainties resulting from the finite accuracy of instruments. As a result of the revolution in metrology stemming from advances in electronics, almost all instruments used today are either simple mechanical instruments or digital electronic meters. Nevertheless, a method will also be given for determining the uncertainty for analog (pointer) instruments still in use.

Simple mechanical instruments

Manufacturers of instruments such as a millimeter ruler, a caliper, or a liquid thermometer generally do not specify their accuracy. It is widely believed that "unspecified accuracy" is equal to the value of the smallest scale division, hereinafter referred to as the elementary division. Its value is 1 mm for a ruler, 0.05 mm for a caliper, 0.01 mm for a micrometer screw, and 0.1°C for a medical thermometer. As a first approximation for standard uncertainty, we take:

$$u(x) \approx \text{elementary division.} \tag{9}$$

This assessment can be adjusted up or down according to possessed knowledge and experience. For example, if we measure the diameter of a one-grosz coin with a ruler and evaluate "by eye" also tenths of a millimeter, the standard uncertainty can decrease to 0.2 mm. On the other hand, when measuring the dimensions of a room with a measuring tape, the uncertainty should be assumed to be greater than 1 mm, even though we have a millimeter scale on the entire five-meter tape.

Digital and analog electrical meters

In instruments with a digital readout, the value corresponding to a change in the last digit, conventionally also called the elementary division, determines the instrument's resolution. The measurement uncertainty is larger and is provided by the manufacturer in the instrument's manual. Under the name "limiting error", "accuracy", etc., lies the limiting uncertainty, defined most often as a specific fraction of the measured value plus a fraction of the range,

$$\Delta x = C_1 \cdot x + C_2 \cdot \text{range}. \tag{10}$$

For example, for type 1321 ohmmeters used in the Laboratory, we have $C_1 = 0.2\%$ and $C_2 = 0.1\%$. When measuring a $10\,\text{k}\Omega$ resistor on the $20\,\text{k}\Omega$ range, we get $\Delta x = 0.04\,\text{k}\Omega$, the equivalent of 4 elementary divisions.

In the case of analog (pointer) meters, the value of the limiting uncertainty is determined by the formula

$$\Delta x = \frac{\text{instrument class}}{100} \cdot \text{range} \tag{11}$$

where the parameter called the instrument class, equal to 0.5; 1; 1.5 or 2.5, can be found as one of the symbols determining the instrument's properties (under the glass). Note that also in the case of pointer meters, the value of Δx should not be identified with the elementary scale division of the instrument.

The *Guide* recommends converting the limiting uncertainty obtained from formulas (1.9a) or (1.9b) into standard uncertainty using the formula[5]

$$u(x) = \frac{\Delta x}{\sqrt{3}}. \tag{12}$$

Formula (1.10) follows from the simplifying assumption that if x is the measured value, then the unknown true value can occur with equal probability in the interval $(x - \Delta x, x + \Delta x)$. In other words, we assume we are dealing with a uniform distribution, for which the standard deviation is equal to half the width of the probability density function divided by $\sqrt{3}$. (Appendix A, formula (A7)).

Example 1.3. Type B evaluation of uncertainty for pendulum length measurement
We measure the length of the pendulum with a millimeter ruler, obtaining the value $l = 410$ mm. We assume an uncertainty equal to the scale division: $u(l) = 1$ mm. This assessment takes into account the difficulty of properly applying the ruler to the segment: center of the ball - suspension point of the pendulum.

1.5. Law of propagation of uncertainty

Many physical quantities cannot be measured with a single instrument, but are determined by an indirect measurement method. For example, gravitational acceleration can be determined based on the measurement of the length and period of oscillation of a pendulum. Suppose that the quantity y we are interested in is calculated from the functional formula

$$y(x_1, \dots, x_k, \dots),$$

where the successive x_k can be measured directly. The uncertainties $u(x_1), \dots, u(x_k)$ of directly measured quantities "propagate" to the calculated quantity y, causing it to be burdened with finite uncertainty. Therefore, the methods of calculating the uncertainty of quantity y are called the law of propagation of uncertainty.

Function of one variable

We will begin the analysis of the problem with a function of one variable $y = f(x)$. The uncertainty $u(x)$ is small compared to the measured value x, so the uncertainty y can be calculated as the absolute value of the product of the function's derivative and the uncertainty $u(x)$,

$$u(y) = \left| \frac{dy}{dx} \cdot u(x) \right|. \tag{13}$$

(The product of a derivative and a small increment of the independent variable is called a differential in mathematics. We take its absolute value so that the uncertainty $u(y)$ is a positive number.) The law of propagation of uncertainty for a function of one variable is illustrated by Fig. 1.4 and Example 1.4.

[5] We perform this conversion only when it is needed, in particular to apply the law of propagation of uncertainty (section 1.5).

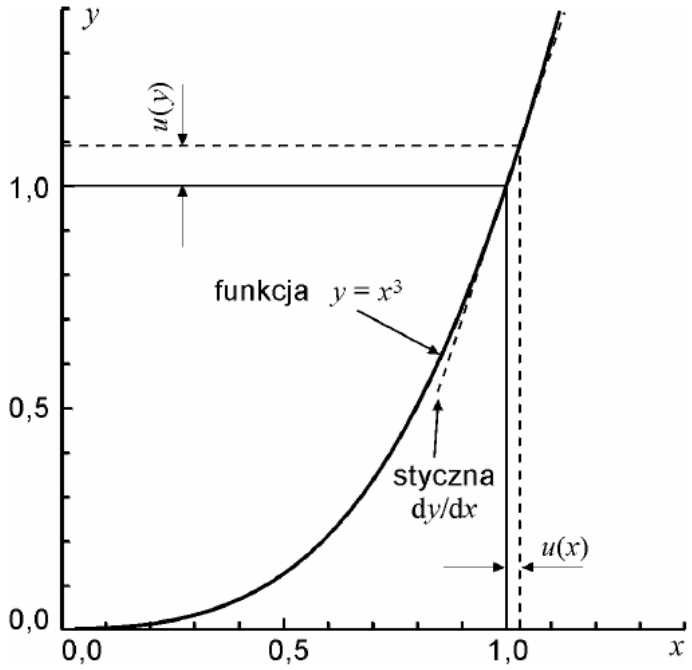


Figure 1.4: Illustration of the law of propagation of uncertainty

Example 1.4. Uncertainty of the volume of a sphere of known diameter
We measured the diameter D of a steel ball with a caliper, obtaining the value $D = 2.45$ mm with an uncertainty $u(D) = 0.05$ mm. The volume of the sphere is calculated from the formula $(4/3)\pi r^3 = (\pi/6)D^3 = 7.70$ mm^3. The uncertainty of the sphere's volume is

$$u(V) = \frac{d}{dD}\left(\frac{\pi}{6}D^3\right)u(D) = \frac{\pi}{2}D^2u(D) = \frac{3.1416}{2}(2.45\text{ mm})^2 \cdot 0.05\text{ mm} = 0.47\text{ mm}^3.$$

Function of many variables

In the case of a function of many variables, we calculate the partial differentials for successive variables $x_1 \ldots x_k \ldots$ using formula (1.11) and form their geometric sum[6]

$$u_c(y) = \sqrt{\sum_k \left[\frac{\partial y}{\partial x_k}u(x_k)\right]^2}. \tag{14}$$

The calculated value of the uncertainty of function y is called the combined uncertainty and is denoted by the symbol u_c or $u_c(y)$[7]. The geometric summation is a consequence of the theorem on the standard deviation of a sum of random variables, assuming that the random variables are uncorrelated. The condition of lack of correlation is fulfilled if each of the quantities x_k is measured with a different instrument.

The simplest case of the law of propagation of (absolute) uncertainty occurs when the function y is a sum or difference of any number of components. The partial derivatives $\partial y/\partial x_k$ are equal to unity and as a result, the combined uncertainty is the geometric sum of the uncertainties of the individual components:

$$y = x_1 + x_2 - x_3 + \cdots \Rightarrow u_c(y) = \sqrt{u^2(x_1) + u^2(x_2) + u^2(x_3) + \ldots}. \tag{15}$$

[6]The geometric sum is the square root of the sum of the squares of the components.
[7]The index c comes from the English word *combined*.

Propagation of relative uncertainty

The law of propagation of uncertainty takes a particularly clear and convenient form for practical calculations when, instead of absolute uncertainties, we calculate the combined relative uncertainty $u_c(y)/y$. For this purpose, equation (1.12) is divided on both sides by y, and then, inside the square bracket, multiplied and divided by x_k,

$$\frac{u_c(y)}{y} = \frac{1}{y}\sqrt{\sum_k \left[\frac{\partial y}{\partial x_k}u(x_k)\right]^2} = \sqrt{\sum_k \left[\left(\frac{\partial y}{\partial x_k}\cdot\frac{x_k}{y}\right)\cdot\frac{u(x_k)}{x_k}\right]^2}.$$

The obtained expression is written in a compact form

$$\frac{u_c(y)}{y} = \sqrt{\sum_k \left[p_k\frac{u(x_k)}{x_k}\right]^2} \tag{16}$$

expressing the law of propagation of relative uncertainties: The combined relative uncertainty $u_{c,r}(y) = u_c(y)/y$ is the geometric sum of the relative uncertainties $u(x_k)/x_k$ of directly measured quantities multiplied by dimensionless sensitivity coefficients p_k equal to

$$p_k = \frac{x_k}{y}\frac{\partial y}{\partial x_k}. \tag{17}$$

Formula (1.14a) seems more complicated than formula (1.13) expressing the "ordinary" law of propagation of uncertainty. The point is that when calculating the weights, most symbols cancel out (see Example 1.5) and the formulas for p_k turn out to be amazingly simple. They are collected in Table 1.2.

Table 2: Sensitivity coefficients p_k for the most important functions

Function form	$p_k = \frac{x_k}{y}\frac{\partial y}{\partial x_k}$
$y = \text{const}\cdot x_k$	1
$y = \frac{\text{const}}{x_k}$	-1
$y = \text{const}\cdot x_k^n$	n
$y = \text{const}\cdot\exp(ax_k)$	ax_k
$y = \text{const}\cdot\ln(ax_k)$	$\frac{\text{const}}{y}$

In the "function form" column, the symbol const means not only a constant, but also the remaining part of the functional formula not containing the variable x_k, which is treated as a constant factor when calculating the appropriate partial derivative.

The simplest and practically important case of the law of propagation of relative uncertainty occurs when quantity y is a product or quotient of directly measured quantities. The coefficients p_k are then equal to $+1$ or -1 (Table 1.2). As a result, the combined relative uncertainty is the geometric sum of the relative uncertainties of the factors x_k:

$$y = \frac{x_1\cdot x_2\cdot\ldots}{x_3\cdot\ldots} \Rightarrow \frac{u_c(y)}{y} = \sqrt{\left[\frac{u(x_1)}{x_1}\right]^2 + \left[\frac{u(x_2)}{x_2}\right]^2 + \left[\frac{u(x_3)}{x_3}\right]^2 + \ldots}. \tag{18}$$

A qualitative conclusion from the law of propagation of uncertainty is to determine which quantity x_k gives the largest contribution to the combined uncertainty. It is usually, but not always, the variable whose relative uncertainty is the largest.

Example 1.5. Uncertainty of the value of gravitational acceleration determined from the measurement of the period of oscillation and length of a simple pendulum.
For a pendulum, we determined the values and uncertainties of the oscillation period $T = 1279.33$ ms, $u(T) = 0.72$ ms and length $l = 410$ mm, $u(l) = 1$ mm (Examples 1.2 and 1.3). Gravitational acceleration is calculated as

$$g = \frac{4\pi^2 l}{T^2} = \frac{4 \cdot 3.1416^2 \cdot 410 \text{ mm}}{(1.27933 \text{ s})^2} = 9890 \frac{\text{mm}}{\text{s}^2} = 9.890 \frac{\text{m}}{\text{s}^2}.$$

Note: In the calculation, we write both numerical values and units. We record the result with a number of digits "matched" to the predicted measurement uncertainty. In most cases, writing 4 significant digits is sufficient.
Calculating the combined uncertainty using formula (1.12) requires computing the expression

$$u_c(g) = \sqrt{\left[\frac{4\pi^2}{T^2} u(l)\right]^2 + \left[-\frac{8\pi^2 l}{T^3} u(T)\right]^2}.$$

Applying formula (1.14a) for relative uncertainty we obtain:

$$\frac{u_c(g)}{g} = \sqrt{\left[\left(\frac{4\pi^2}{T^2} / \frac{4\pi^2 l}{T^2}\right) \frac{u(l)}{l}\right]^2 + \left[\left(-\frac{8\pi^2 l}{T^3} / \frac{4\pi^2 l}{T^2}\right) \frac{u(T)}{T}\right]^2} = \sqrt{\left[\frac{u(l)}{l}\right]^2 + \left[-2\frac{u(T)}{T}\right]^2}.$$

The obtained coefficients p_k, equal to 1 and -2 for l and T respectively, can be written down immediately using Table 1.2. Numerical calculations and recording of uncertainties are performed to an accuracy of 2 significant digits (see section 1.7). It is convenient to list them in a table:

	x_k	$u(x_k)$	$\frac{u(x_k)}{x_k}$	p_k	$p_k \frac{u(x_k)}{x_k}$
length l	409 mm	1 mm	0.24%	1	0.24%
period T	1279 ms	0.72 ms	0.056%	-2	-0.11%
				Geometric sum:	0.26%

From the obtained relative uncertainty, we calculate the absolute uncertainty

$$u_c(g) = 9.890 \cdot \frac{0.26\%}{100\%} = 0.026 \frac{\text{m}}{\text{s}^2}.$$

(Note: the original text says $0.028 \frac{\text{m}}{\text{s}^2}$ likely due to intermediate rounding differences or typos in the original source, but I preserve the mathematical procedure). Comparing the contributions $p_k \frac{u(x_k)}{x_k}$ originating from $u(l)$ and $u(T)$ shows that the greater source of uncertainty in gravitational acceleration is the uncertainty of the pendulum length measurement.

1.6. Expanded uncertainty

A property of standard uncertainty is that the true value lies within the interval from $x - u(x)$ to $x + u(x)$ with a probability of about 2/3 (exactly: 68% for a Gaussian distribution, 58% for a uniform distribution). Standard uncertainty is a measure of the accuracy of measurements, enables comparing the accuracy of different measurement methods; this measure of uncertainty is shown on graphs (about which in section 1.8).

To infer the agreement of a measurement result with other results, the *Guide* introduces the concept of **expanded uncertainty**. As the name suggests, this is an "enlarged" standard

uncertainty, chosen so that the prevailing majority of measurement results needed for specific applications in industry, medicine, and environmental protection fall within the interval $(y - U(y), y + U(y))$. The value U is calculated by multiplying the combined uncertainty by a dimensionless expansion factor (coverage factor) k

$$U(y) = k \cdot u_c(y). \tag{19}$$

According to international practice, a conventional value of $k = 2$ is most often adopted for calculating U. Values of k other than 2 can be used only for specific applications and should be dictated by established and documented requirements[8]. The value $k = 2$ corresponds to a probability of realization of a random variable in the interval $(y - U(y), y + U(y))$ equal to 95% for a Gaussian distribution and 100% for a uniform distribution.

Typical applications of expanded uncertainty are inferring the agreement of the obtained result with the exact value, or with another measured value of known uncertainty.

Comparison with an exact value (theoretical or tabular)

A theoretical value is a quantity, usually dimensionless, which can be determined flawlessly or with negligibly small uncertainty using theory. For example, using a flexible measuring tape and a round basin, one can experimentally determine the ratio of the circumference to the diameter of a circle. The measured value can be compared with the theoretical value $\pi = 3.1415927\ldots$.

Exact tabular values include, among others, physical constants, whose values come from measurement but are known with very high accuracy.

Checking agreement consists of verifying whether the exact value y_0 falls within the interval $(y-U(y), y+U(y))$. A mathematically equivalent way is to check if the inequality $|y-y_0| < U(y)$ is satisfied. A negative result of the comparison (assuming calculations were flawless) usually indicates the presence of an undetected systematic or gross error.

Comparison of the results of two measurements

The results of two independent measurements of the same quantity (e.g., the refractive index of glass) generally have different values. The question arises: do these quantities really differ (because different types of glass were measured), or are they equal within the limits of measurement uncertainty. Uncertainty theory helps answer this quantitatively.

The calculation proceeds as follows. We have at our disposal two measured values, y_1 and y_2, and their standard uncertainties, $u(y_1)$ and $u(y_2)$. According to the law of propagation of uncertainty (formula (1.13)), the difference $y_1 - y_2$ has an uncertainty equal to the geometric sum of $u(y_1)$ and $u(y_2)$. Therefore, the expanded uncertainty is

$$U(y_1 - y_2) = k\sqrt{[u(y_1)]^2 + [u(y_2)]^2}. \tag{20}$$

Measurement results are considered consistent with each other if $|y_1 - y_2| < U(y_1 - y_2)$.

[8]The last two sentences quote the official position of the National Institute of Standards and Technology USA, the most important metrological institute in the world.

Example 1.6. Comparison of the obtained gravitational acceleration value with the tabular value
We obtained, using a simple pendulum, the value $g = 9.866$ m/s^2 with an uncertainty $u(g) = 0.028$ m/s^2. The tabular value for Kraków is $g_0 = 9.811$ m/s^2. We calculate the difference

$$g - g_0 = 9.890 \text{ m/s}^2 - 9.811 \text{ m/s}^2 = 0.079 \text{ m/s}^2.$$

We calculate the expanded uncertainty, assuming the value $k = 2$,

$$U(g) = k \cdot u(g) = 2 \cdot 0.028 \text{ m/s}^2 = 0.056 \text{ m/s}^2.$$

Since $|g - g_0| > U(g)$, we must conclude that the measured gravitational acceleration is inconsistent with the tabular value. We intentionally give such a result to show that such cases can also appear in measurement practice. Stating an inconsistency should be a stimulus to analyze possible reasons for its occurrence and look for ways to improve the measurement.
The difference $g - g_0$ is small, which indicates that no gross error was made. Rather, unavoidable measurement uncertainties were evaluated too low. For example, an additional systematic error might have occurred during period measurement, which cannot be detected by statistical analysis of the result of 8 measurements. A radical reduction of $u(T)$ is possible by applying electronic time measurement.
Also, the length measurement uncertainty might have been evaluated too optimistically, considering the difficulty of determining "by eye" where the center of the ball is. A way to increase the accuracy of measuring l could be, for example, measuring with a ruler the distance from the suspension point to the top point of the ball and adding half the ball's diameter, measured using a caliper.

1.7. Recording of measurement uncertainty

We present recommended ways of recording uncertainty using an example. Our example distinguishes word recording (i), using symbols (ii), and abbreviated (iii), but any combination of the presented recording elements can be used.

Standard uncertainty

(i) gravitational acceleration is equal to 9.866 m/s^2 with an uncertainty of 0.028 m/s^2;

(ii) $g = 9.866$ m/s^2; $u(g) = 0.028$ m/s^2;

(iii) $g = 9.866(28)$ m/s^2.

Expanded uncertainty

(i) gravitational acceleration is 9.866 m/s^2 with an expanded uncertainty of 0.056 m/s^2;

(ii) $g = 9.866$ m/s^2; $U(g) = 0.056$ m/s^2;

(iii) $g = (9.866 \pm 0.056)$ m/s^2.

The example illustrates the rules for recording uncertainty recommended by the *Guide*.

- We record the uncertainty to the accuracy of two significant digits[9]. When rounding to two significant digits, the limiting uncertainty caused by rounding ranges from 5% to 0.5%

[9]The uncertainty value $u(g) = 0.028$ m/s^2 used in the example has two significant digits. Leading zeros are not significant digits - they may disappear with a different choice of units, e.g., $u(g) = 28$ mm/s^2. A long-standing custom (confirmed by another document of the GUM Convention from 2009) also allows recording one significant digit. After all, we know the measurement uncertainty so inaccurately! On the other hand, three or more digits should not be given, because there is no situation in which additional digits would have real significance.

(respectively, for digits 10 and 99). Such accuracy is sufficient because the uncertainty estimate is more inaccurate (see Table 1.1).

- We round the measured value to the same place as the uncertainty. If the last digit of the result is a zero, it should be kept as a significant digit.
- In abbreviated records (iii), the symbol ± should be used for expanded uncertainty, while a notation using parentheses should be used for standard uncertainty.

Additional remarks on recording numbers and units

Measurement and calculation results are best given in units for which the numerical value is contained within the range from approximately 0.1 to 1000. Numbers from this range can be called "friendly" — they are easy to pronounce and remember, and recording them requires the fewest printing characters. To keep the numerical result in this range, prefixes have been introduced to the SI system, such as: $p \equiv 10^{-12}$, $n \equiv 10^{-9}$, $\mu \equiv 10^{-6}$, $m \equiv 10^{-3}$, $k \equiv 10^{3}$, $M \equiv 10^{6}$, $G \equiv 10^{9}$, ... (we haven't listed them all). They can be attached to any unit having its own symbol (m, s, A, W, F, Hz, etc.).

When an SI unit is a combination of symbols (e.g., kg/m^3, V/m, W/(K·m) - units of density, electric field strength, and thermal conductivity), prefixes can be attached to any symbol. For example, writing the density of mercury as 13.6 g/cm^3 is friendlier than $13.6 \cdot 10^3$ kg/m^3.

1.8. Graphs of functional dependencies

The essence of physics methodology is the experiment and its theoretical interpretation. In graphs illustrating the functional dependencies of two or more variables, this is reflected in the clear marking of both experimental points and the theoretical curve interpreting the course of the phenomenon.

Creating graphs is an important skill, useful in other experimental sciences. The principles presented apply to both manually drawn graphs and those prepared using a computer.

Standard graph elements are discussed below.

Coordinate system

The coordinate system must have a scale and designations of quantities and units. The graph scale should be chosen so that it is clear and uses the paper surface well. To fulfill this, the scale does not have to start from zero (e.g., the vertical axes of Figures 1.5 and 1.9).

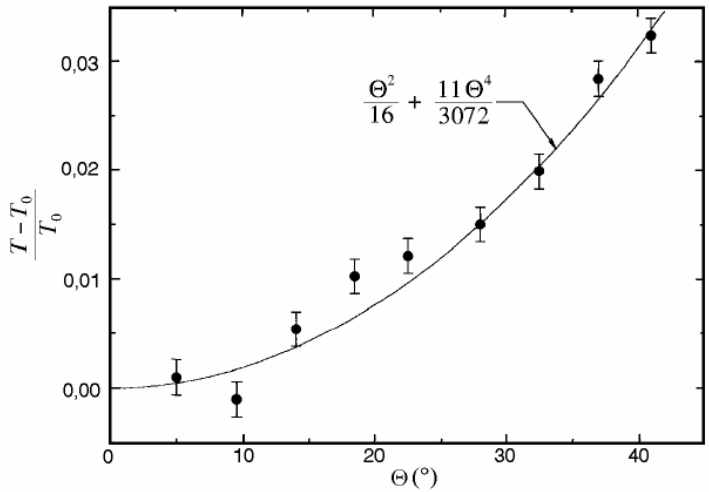


Figure 1.5: Dependence of the relative change in the pendulum's oscillation period on the amplitude of oscillation. An example of comparing experimental points with a theoretical curve

From mathematics, we are accustomed to drawing a coordinate system as two perpendicular line segments (Fig. 1.9). A second method is to enclose the graph area in a rectangular frame (Figs. 1.5, 1.6, 1.8). The symbol or word description is then placed in the middle of the frame side (units in square brackets). The description is made with technical lettering. Scale tick marks, pointing inward the frame, are repeated on the other two sides. If necessary, parallel frame sides can be used to show different scales (Fig. 1.8b). We encourage making graphs in a "frame" form, as it is currently the standard way of presenting experimental results in technical documentation and scientific publications.

Manually prepared graphs are made with a pencil on graph paper. Ordinary graph paper has a linear grid. You can also purchase graph paper with a logarithmic grid on one (Fig. 1.6, 1.8b) or both axes. Nonlinear scales of various types can also be realized using a computer.

Graphs with a logarithmic scale are used for two different reasons. First, for sensible presentation of a quantity that varies by many orders of magnitude (Fig. 1.6). Second, to linearize exponential and power functions (which will be discussed in subsection 1.9).

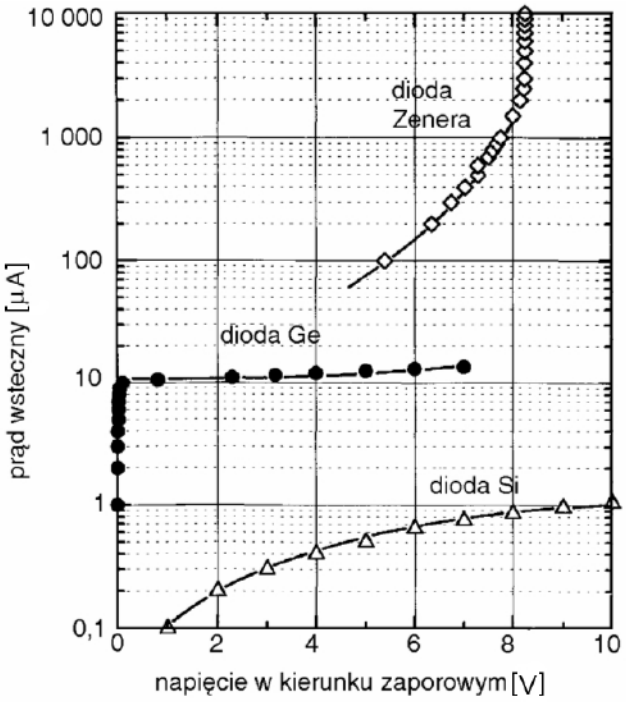


Figure 1.6: Current-voltage characteristics of semiconductor diodes of various types, reverse-biased. An example of dependencies that cannot be described by a simple mathematical function

Experimental points

The basis for making a graph is a table (Example 1.7). Experimental points are images of corresponding pairs of numbers from the table. Points drawn with a pencil on a graph are poorly visible. In many sciences (e.g., economics), these points are connected by a broken line.

In physics and other exact sciences, we generally proceed differently: we make the points visible by surrounding them with a symbol in the shape of a circle, square, etc. (Fig. 1.7). For aesthetic manual drawing of symbols, it is worth using plastic templates equipped with appropriate holes. Different symbol shapes can be used to convey additional information, e.g., to distinguish points belonging to different curves.

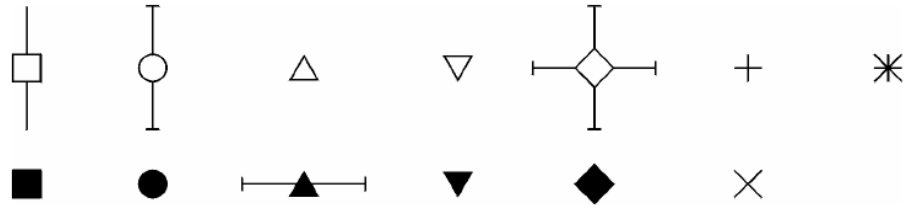


Figure 1.7:
Sample symbols of experimental points and ways of drawing uncertainty bars

We can also show measurement uncertainties on the graph. The generally accepted way is to draw an uncertainty bar of length $\pm u(y)$ or $\pm u(x)$, as shown in Figures 1.5 and 1.8b. We plot them when they are large on the scale of the drawing, i.e., the size of the uncertainty bar exceeds the dimensions of the experimental point symbol. Good computer programs also allow plotting specified uncertainty bars. Marking uncertainties serves, among other things, to infer the agreement of experiment with theory. If the values of uncertainty bars have been correctly estimated, on average 2/3 of them should intersect the theoretical curve.

Curve interpreting the experiment results

The rules for drawing the curve depend on the "quality" of the theoretical description we have at our disposal.

- We have an algorithm allowing us to calculate the theoretical curve independently of the position of the experimental points. The graph consists of these points and the calculated curve (Fig. 1.5). An "experimental" curve is not needed!

- We know the type of function from theory (e.g., we know it is an exponential function $y = Ae^{-ax}$), but we do not know its parameters A and a. Then, a function of a given kind must be fitted as best as possible to the position of the experimental points; the parameters of the fitted function are the measurement results (Fig. 1.8). Methods of fitting the straight line $y = ax + b$ are discussed in section 1.10.

- We do not have a specific functional formula (e.g., for the dependence of thermocouple voltage on temperature). Then, through the experimental points, we manually (or with the help of a French curve) draw a smooth "experimental" curve (Fig. 1.6). The procedure of "smoothing" measurement results is based on the assumption that an unknown smooth function $y(x)$ exists, thus it can be approximated by a power series. Therefore, in the case of using a computer (which cannot do anything "by eye"), one way to generate a smooth curve is to fit a power series, i.e., a polynomial, the degree of which we choose by trial and error.

Beside the curve, in the graph area, one can and should place additional inscriptions, lines, arrows, etc., facilitating its understanding. The above, somewhat schematic remarks do not, of course, exhaust all graph possibilities and forms.

Example 1.7. Graph of pendulum period versus amplitude
The pendulum period measurement elaborated in Example 1.2 was performed at a small oscillation amplitude. Let's recall the result: $T_0 = 1.2793$ s, $u(T_0) = 0.0072$ s (note: previously calculated standard uncertainty of the mean was 0.00071 s, possibly a typo in source text, keeping numbers as in source text).
Next, single measurements of 50 periods were taken for the swinging pendulum as a function of the increasing oscillation amplitude Θ. The table below presents the measured period values T and the calculated values of the relative change in period $(T-T_0)/T_0$. The quantity $(T-T_0)/T_0$ is introduced because it does not depend on the pendulum length and gravitational acceleration; moreover, the dependence of $(T-T_0)/T_0$ on the swing angle Θ is the same for a swinging body of any shape.

Θ [°]	5	9.5	14	18.5	22.5	28	32.5	37	41
T [s]	1.2808	1.2780	1.2862	1.2926	1.2950	1.2986	1.3090	1.3158	1.3208
$\frac{T-T_0}{T_0}$	0.0012	-0.0010	0.0054	0.0104	0.0123	0.0151	0.0232	0.0285	0.0324

Since each measurement of period T was performed only once, the standard deviation estimator of a single measurement can be taken as the measurement uncertainty: $u(T) = 0.0020$ s (see Example 1.2). Therefore, we applied a Type B uncertainty evaluation, based on the results of the previous measurement.
The combined uncertainty of the quantity $(T-T_0)/T_0$ determined from the law of propagation of uncertainty is

$$u_c\left(\frac{T-T_0}{T_0}\right) = \sqrt{\left[\frac{1}{T_0}u(T)\right]^2 + \left[-\frac{T}{T_0^2}u(T_0)\right]^2} = \sqrt{\left(\frac{0.020\text{ s}}{1.2793\text{ s}}\right)^2 + \left(-\frac{1.32\text{ s}\cdot 0.0072\text{ s}}{(1.2793\text{ s})^2}\right)^2} \approx 0.017.$$

Calculating u_c for different values of T, we find that the uncertainty is practically the same for all points on the graph. It has been marked on the graph (Fig. 1.5). On the other hand, the relatively small uncertainty in angle measurement was not shown.
The theoretical value of the relative change in the pendulum's period can be calculated, independently of the experiment results, using the formula

$$(T-T_0)/T_0 = \theta^2/16 + (11\theta^4)/3072 + \ldots$$

(see exercise 2, the swing angle θ must be substituted in radians). The data needed to draw the curve, listed in the table below, are calculated at equal angle intervals.

Θ [°]	5	10	15	20	25	30	35	40	45
$\frac{T-T_0}{T_0}$	0.0005	0.0019	0.0043	0.0076	0.0119	0.0172	0.0234	0.0305	0.0387

We do not visualize the plotted theoretical points, but we draw a smooth curve through them. In the case of using a computer, it is simplest to calculate more theoretical points (e.g., every 0.5 degrees) and order the machine to draw a broken line through them; the impression of a smooth curve is guaranteed!
Since we mark standard uncertainty on the graph, only a part of the uncertainty bars (circa 2/3) should intersect the theoretical curve. Figure 1.5 thus demonstrates the agreement of theory and experiment.

1.9. Linearization of nonlinear functional dependencies

A straight line is the easiest curve to draw, and our eye easily distinguishes a curve from a straight line. Since both drawing and mathematical analysis of nonlinear dependencies are more difficult than linear ones, a common practice is reducing nonlinear functional dependencies to a linear form.

For example, if we are dealing with an exponential dependency

$$y = A\exp(-ax), \tag{21}$$

then to linearize it we calculate the logarithm

$$\ln y = \ln A - ax. \tag{22}$$

By placing $\ln y$ on the vertical axis and x on the horizontal axis, we obtain a line with a slope coefficient equal to $-a$, intersecting the vertical axis at the point $\ln A$ (Fig. 1.8).

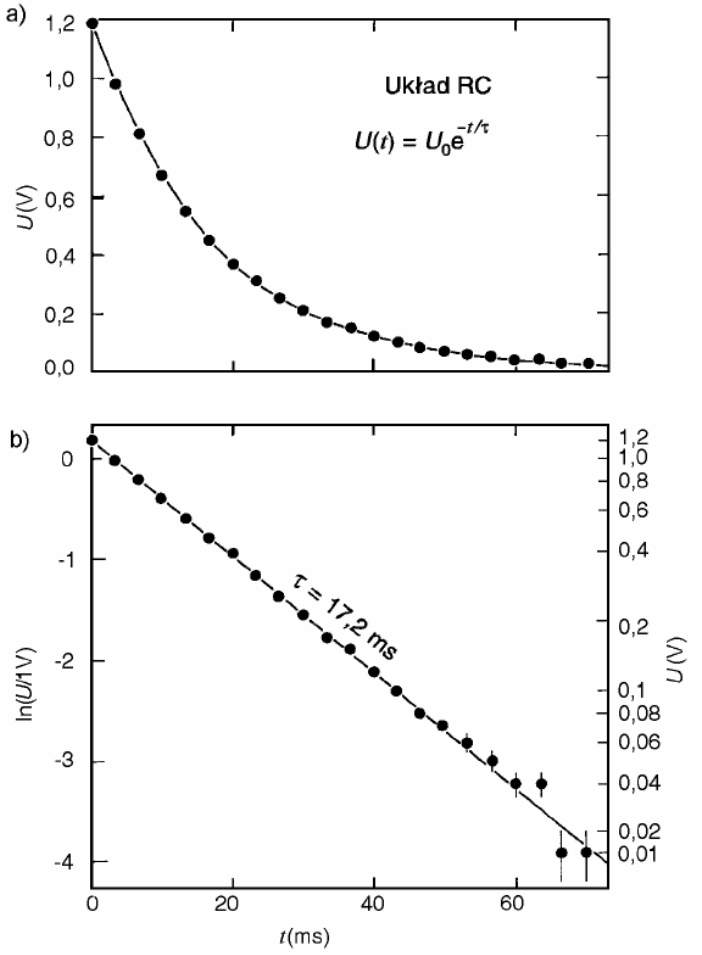


Figure 1.8: Dependence of voltage U on time t during the discharge of capacitor C through resistor R, shown on a normal (a) and linearized (b) graph. The measurement uncertainty of U is 0.005 V.

When drawing graphs that use the logarithm operation, a common practice—instead of showing the logarithm value on a given axis—is introducing a nonlinear logarithmic scale (semi-logarithmic paper or a logarithmic scale option in a graphics program), which is equivalent to this operation. Figure 1.8b has a vertical axis described on one side by the values of the natural logarithm itself, and on the right side using a logarithmic scale.

In general, both the very possibility of linearization and the type of coordinates that must be used for this purpose depend on the form of the function. Almost all functions containing two unknown parameters can be brought to a linear form. However, functions dependent on three or more parameters (e.g. $y = ax^2 + bx + c$) cannot be linearized.

When linearizing a function, alongside the change in coordinate values of the points, the values of uncertainty bars also change. The new values of u are calculated using the law of propagation of uncertainty for a function of one variable (formula (1.12)). The comparison of Figures 1.8a and 1.8b reveals that the sizes of uncertainty bars, identical (and invisible) on the linear scale, become relatively large in the "tail" of the logarithmized dependency.

1.10. Fitting a straight line to a set of experimental points

The topic of this chapter is the issue of drawing the line $y = ax + b$ fitted as best as possible to a set of n experimental points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. The goal of fitting is not only to obtain a visual effect, but above all to obtain the values of parameters a and b describing the line, and their uncertainties $u(a)$ and $u(b)$.

The graphical method consists of making a graph, then applying a ruler (preferably transparent) and sketching the line by eye so that the distances between the line and experimental points are on average as small as possible. The graph for the graphical method should be large (A4 size), with scales chosen so that the line's slope is close to 45° (Fig. 1.9). Nowadays, the graph of points for the graphical method can be prepared using a computer.

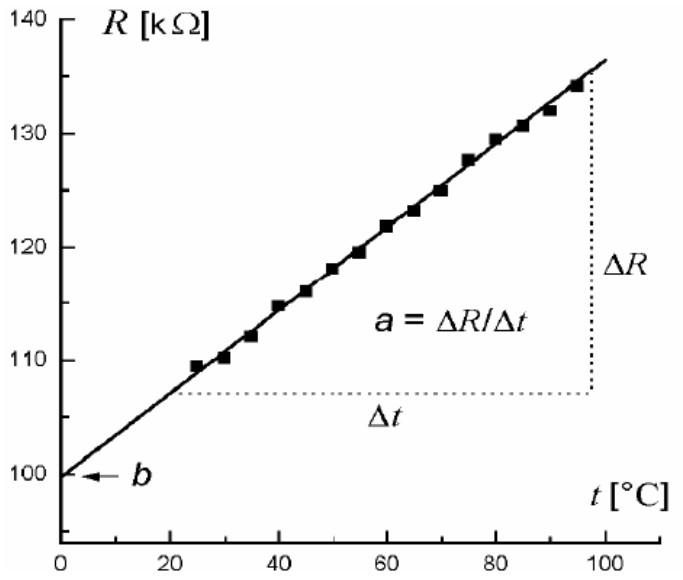


Figure 1.9: Determining line parameters by graphical method.

The slope coefficient $a = \Delta y/\Delta x$ is the ratio of the legs Δy and Δx of a large right triangle whose hypotenuse is part of the graphically drawn line (Fig. 1.9). Parameter b determines the point of intersection of the line with the y-axis.

A source of misunderstandings can be the identification, derived from mathematics courses, of the slope coefficient with the tangent of the line's angle of inclination to the x-axis. In graphs of physical quantities, the inclination angle α of a line can be different for the same measurement data depending on what scales we apply on the graph's axes. Uniquely defined remains the slope coefficient $a = \Delta y/\Delta x$ (briefly called slope). In contrast to the dimensionless tangent, slope a has a dimension, which is the ratio of the dimensions of quantities y and x.

A disadvantage of the graphical method might seem to be subjectivity - everyone will draw the line a bit differently. Tests show, however, that with a correctly prepared graph and a bit of practice, the values of the line parameters are, within uncertainty limits, the same as those obtained analytically. Furthermore, an advantage of the graphical method is the elimination of points drastically deviating from the line. The greatest disadvantage of the method is the lack of information about the uncertainty of the line's parameters.

The method of least squares is the most commonly used analytical method. It owes its name to the criterion of fit quality: such a selection of line parameters that the sum of squared differences between experimental values y_i and calculated values $ax_i + b$ is as small as possible

$$S^2 = \sum_{i=1}^{n} [y_i - (ax_i + b)]^2 = \min . \tag{23}$$

Criterion (1.19) provides the best estimate of line parameters under the assumption that all measurement points are burdened with equal random error with a normal distribution.

To find parameters a and b, we use the usual condition for the minimum of a two-variable function:

$$\frac{\partial S^2}{\partial a} = 0, \quad \frac{\partial S^2}{\partial b} = 0.$$

Calculating the above partial derivatives leads to a system of linear equations for unknowns a and b:

$$a \sum x_i^2 + b \sum x_i = \sum x_i y_i$$
$$a \sum x_i + bn = \sum y_i$$

The solution to this system of equations can be written in two equivalent ways. The formulas presented below are most convenient for manual calculations. We begin by calculating the arithmetic means for variables x and y:

$$\overline{x} = \frac{1}{n} \sum x_i, \quad \overline{y} = \frac{1}{n} \sum y_i \tag{24}$$

determining the position of the "center of gravity", $\overline{x}, \overline{y}$, of experimental points.

Line parameters are calculated from the formulas:

$$a = \frac{1}{D} \sum y_i (x_i - \overline{x}) \tag{25}$$

$$b = \overline{y} - a\overline{x}, \tag{26}$$

where

$$D = \sum (x_i - \overline{x})^2. \tag{27}$$

Notice that the formula for parameter b, i.e., the point of intersection of the line with the y-axis, results from drawing a line with slope a through the "center of gravity" $\overline{x}, \overline{y}$.

Applying the laws of mathematical statistics allows deriving formulas for the standard deviations of both line parameters. First, we calculate the quantity

$$s_y = \sqrt{\frac{S^2}{n-2}} = \sqrt{\frac{\sum [y_i - (ax_i + b)]^2}{n-2}} \tag{28}$$

which is the standard deviation estimator of the points from the fitted line. The value s_y is an intermediate result for calculating the uncertainty of line parameters, which we calculate from formulas:

$$u(a) = \frac{s_y}{\sqrt{D}} \tag{29}$$

$$u(b) = s_y \sqrt{\frac{1}{n} + \frac{\overline{x}^2}{D}} \tag{30}$$

The least squares criterion can be used to fit other functional dependencies. The necessary algorithms are discussed in mathematical statistics textbooks and implemented in computer data analysis programs.

Special cases of straight line fitting

Many laws of physics express the proportionality of one quantity to another, e.g., Ohm's law of current I to voltage U. The graphical image of such a function is a straight line passing through the origin of the coordinate system. In the graphical method, we draw the line in such a way that it passes through point (0, 0). Applying the method of least squares requires fitting a line given by equation $y = ax$ (i.e., with parameter value $b = 0$).

The derivation of the formula for the slope value a is very simple. We seek the minimum of the function

$$S^2 = \sum_{i=1}^{n} [y_i - ax_i]^2 = \min . \tag{31}$$

Squaring the expression $y_i - ax_i$ gives

$$S^2 = \sum y_i^2 - \sum 2ax_iy_i + \sum a^2x_i^2.$$

Pulling constant factors in front of the sum sign yields

$$S^2 = \sum y_i^2 - 2a \sum x_iy_i + a^2 \sum x_i^2.$$

To find the minimum, we calculate the derivative with respect to a and equate it to zero:

$$\frac{dS^2}{da} = 0 - 2 \sum x_iy_i + 2a \sum x_i^2 = 0.$$

As a result, we obtain a formula for the parameter value

$$a = \frac{\sum x_iy_i}{\sum x_i^2} \tag{32}$$

which is different from (1.21). Without derivation, we give the formula for uncertainty

$$u(a) = \sqrt{\frac{S^2}{(n-1)\sum x_i^2}} \tag{33}$$

where the value S^2 is determined by sum (1.25).

A second special case is fitting a horizontal line, described by equation $y = b$. An analogous derivation from the condition of minimum sum of squares gives

$$b = \frac{\sum y_i}{n} \tag{34}$$

which is the arithmetic mean. Using the mean as the best approximation for a sequence of n observations (formula (1.5)) is therefore a special case of the method of least squares, where the fitted function is a constant function $y = b$.

The method of least squares and the problem of systematic and gross errors

The method of least squares is a Type A evaluation—a statistical analysis of a series of n pairs of numbers x_i, y_i. It provides an evaluation of uncertainty $u(a)$ and $u(b)$ originating solely from random error.

A systematic error equal for all points causes a displacement of the entire image of experimental points and the line. Such displacement (along the x or y direction) only affects the value of line parameter b. It is therefore irrelevant in situations where the really important experiment result is the slope coefficient a.

A second, frequently encountered type of systematic error is that points deviate from the straight line at the beginning or end of the measurement range. When fitting a line, both by least squares and graphically, points systematically deviating from linear dependence should not be taken into account (although we show all points on the graph).

A manifestation of a gross error is a graph point drastically deviating from the rest. It can be the result of a mistake during the execution of the experiment and recording its results, as well as a mistake when entering data into a calculator or computer. Even a single such point decidedly ruins the quality of fitting by the method of least squares.

In summary: a graph allowing visual evaluation of data should be made (or viewed on a computer monitor) before proceeding with calculations.

Example 1.8. Fitting a straight line by the method of least squares
The presented numerical example concerns the dependence of platinum resistor resistance on temperature. This dependence in the used temperature range is described by the line $R = at+b$, where t is the temperature measured in degrees Celsius. Through 15 experimental points (Fig. 1.9) a line must be drawn by the method of least squares.
The table below contains the coordinates of the experimental points (columns x_i, y_i) and all intermediate calculation results.

i	x_i	y_i	$(x_i-\overline{x})^2$	$(x_i-\overline{x})y_i$	$\delta y_i = y_i-(ax_i+b)$	δy_i^2
1	25	109.4	1225	-3829	+0.54	0.29
2	30	110.1	900	-3303	-0.57	0.33
3	35	112	625	-2800	-0.48	0.23
4	40	114.7	400	-2294	+0.40	0.16
5	45	116.0	225	-1740	-0.11	0.01
6	50	118.1	100	-1181	+0.18	0.03
7	55	119.5	25	-597.5	-0.23	0.05
8	60	121.8	0	0	+0.25	0.06
9	65	123.1	25	615.5	-0.26	0.07
10	70	124.9	100	1249	-0.27	0.07
11	75	127.6	225	1914	+0.62	0.38
12	80	129.4	400	2588	+0.60	0.36
13	85	130.6	625	3265	-0.01	0.00
14	90	131.9	900	3957	-0.52	0.27
15	95	134.1	1225	4693	-0.13	0.02
Sum	900	1823.2	7000	2537.5		2.33

Calculation of line parameters:

$$\begin{aligned}\overline{x} &= 900/15 = 60^\circ\mathrm{C}\\ \overline{y} &= 1823.2/15 = 121.55\,\Omega\\ D &= 7000\ \mathrm{deg}^2\\ a &= 2537.5/7000 = 0.3625\,\Omega/\mathrm{deg}\\ b &= 121.55 - 0.3625\cdot 60 = 99.80\,\Omega\end{aligned}$$

Based on calculated parameters a and b, we can plot the fitted line. To do this, we calculate coordinates of two arbitrary points of the line, e.g., $R(100^\circ\mathrm{C}) = 0.3625\cdot 100 + 99.8 = 136.05\,\Omega$ and $R(0^\circ\mathrm{C}) = b = 99.8\,\Omega$, and we connect these points with a straight line. The agreement of the line and experimental points constitutes the best check of the correctness of calculations made so far!
Based on the sum of squares of point deviations from the line (last column of the table) we calculate the value

$$s_y = \sqrt{\frac{2.33}{15-2}} = 0.42\,\Omega$$

and the standard uncertainties of the line parameters:

$$u(a) = \frac{0.42}{\sqrt{7000}} = 0.0050\,\Omega/\mathrm{deg}, \quad u(b) = 0.42\sqrt{\frac{1}{15}+\frac{60^2}{7000}} = 0.24\,\Omega.$$

In short, the regression line parameters and their uncertainties can be written as:

$$a = 0.3625(50)\,\Omega/\mathrm{deg}, \quad b = 99.80(24)\,\Omega.$$

The fact that parameter b is, within expanded uncertainty limits, equal to $100\,\Omega$ is not a coincidence. Our measurement was performed for a standard Pt resistor used for temperature measurement, made so that at a temperature of 0°C its resistance is exactly $100\,\Omega$.

1.11. Use of calculators and computers for data evaluation

Calculators

A good personal calculator is currently as indispensable for an engineer as a slide rule was years ago. It is best to acquire a "scientific" calculator having mathematical functions, performing statistical calculations, and often equipped with simple programming capability. The quality of the instrument and the convenience of work depend largely on the quality of its buttons.

Low-end calculators show all $(7 \div 9)$ digits after the decimal point. (It is reprehensible to report all digits as a calculation result - even for intermediate results, writing 4 significant digits is usually sufficient). Convenient to use are calculators that automatically round to a specified number of decimal places in standard notation, or to a specific number of digits when set to scientific (exponential) notation.

The ability to program a calculator helps streamline repetitive calculations. A typical application is calculating theoretical curve points using a formula entered into the calculator's memory.

Most calculators use standard, i.e. algebraic, notation. It is worth knowing that some calculators use "Polish notation", proposed in the interwar period by Jan Łukasiewicz (1878-1956). In such calculators, numbers are entered first, and then the mathematical operation symbol $(+, \ldots,$ etc.). The simplification of complex calculations consists in the complete elimination of the need to use parentheses and the = sign. Calculators utilizing Reverse Polish Notation (e.g., from Hewlett Packard) have the symbol RPN on the casing.

Statistical calculations using a calculator

Every scientific calculator is adapted for calculating the mean and standard deviation of a series of n numbers. The designations and way of using buttons are implemented differently, but the calculation scheme is the same:

- an introduction to calculations is zeroing the statistical memory registers,
- next, we enter any number of values x_i. They are not memorized, but used to form sums in three registers: $\sum 1$, $\sum x_i$ and $\sum x_i^2$. Summing successive ones determines the number of measurements ($\sum 1 = n$), so the number n does not have to be entered separately.
- after entering data, using appropriate buttons causes the calculation of the mean and the standard deviation estimator of a single measurement s_x,
- as a rule, there is no button for calculating the standard deviation estimator of the mean. An additional calculation must be performed: $u(x) = s_x/\sqrt{n}$ (formula (1.7a)).

Many scientific calculators allow fitting a straight line by the method of least squares. The calculation scheme is similar.

- zeroing statistical memory registers.
- entering successive pairs of numbers x_i, y_i. They will be used to form sums in 6 memory registers: $\sum 1 = n$, $\sum x_i$, $\sum x_i^2$, $\sum y_i$, $\sum y_i^2$, $\sum x_i y_i$.
- using appropriate buttons allows obtaining line parameters a, b (they are calculated based on the contents of the above memory registers).
- in most calculators there is no automatic calculation of uncertainties $u(a)$ and $u(b)$. Fortunately, there is usually a button to calculate the correlation coefficient r. This quantity is a dimensionless statistical parameter, closer to the value 1 or -1 the closer the experimental points are to a straight line. (For data from Example 1.8, $r = 0.9987$.)

Knowledge of r allows calculating the uncertainties of line parameters using formulas[10]:

$$u(a) = |a|\sqrt{\frac{r^{-2}-1}{n-2}}, \quad u(b) = u(a)\sqrt{\frac{\sum x_i^2}{n}} \tag{35}$$

The value $\sum x_i^2$ needed to calculate $u(b)$ can be extracted from the calculator's memory.

Computers

A typical application of a computer in data analysis is statistical calculations and fitting a straight line by the method of least squares. The use of a computer is absolutely necessary with a large amount of data, where its main advantage is not only the speed of calculation, but also the possibility of checking whether there was a mistake when entering data. Appropriate programs are available on computers installed in the Physics Laboratory; they also include the possibility of fitting exponential and power dependencies. Fitting a straight line and other functional dependencies is also a feature of every program for graphical data presentation.

A computer can be used to make graphs. Advanced graphical data presentation programs provide the ability to implement all graph creation rules discussed in section 1.7. Such a program makes the first version of the graph using default settings. To obtain the desired form of the graph, a number of things must be changed or added. For example, to obtain the graph shown in Fig. 1.5, the following modifications were made:

- default axis descriptions were replaced with appropriate ones,
- font size of both axis descriptions and scale ticks was enlarged,
- an additional two sides were added to form a "frame",
- a horizontal dashed line passing through zero on the vertical axis was added,
- direction of scale tick marks was changed to "inward" the frame,
- range of variables on both horizontal and vertical axes was changed,
- distance between digits describing scales was changed,
- unannotated scale tick marks were removed,
- point shape was changed from square to circular,
- the word "theory" and an arrow pointing to the theoretical curve were placed in the graph area.

If the program we use cannot do something - or we have not mastered all its capabilities - a reasonable solution is to supplement the computer drawing manually by adding uncertainty bars, additional lines and descriptions, etc. Ultimately, a human decides on the correctness of the obtained graph and the reasoning: "the graph must be good because it was obtained using a computer" is unacceptable.

[10] Formulas (1.25) appeared in the literature recently (J. Higbie, Uncertainty in the linear regression slope, Am. J. Phys. 59, 184 (1991)) and have not yet made their way into calculator textbooks and manuals.