Andrzej Kubiaczyk
Laboratory of Physics I
Faculty of Physics
Warsaw University of Technology

EVALUATION OF UNCERTAINITY IN MEASUREMENTS

(Student's guide to Physics Laboratories)

Chapter 1 - Introduction

International Standard Organization (ISO) prepared *„Guide to the Expression of Uncertainty in Measurement"*, which is definitive document describing norms and procedures in the measurements uncertainty evaluation. Based on the international ISO standard, Polish norm *„Wyrażanie niepewności pomiaru. Przewodnik"* [1] was accepted in the 1999.

This guidebook is addressed to Physics Laboratories students and its main goal is to provide information necessary to understand international regulations concerning measurements of physical quantities as well as evaluation of measurement uncertainties in the everyday laboratory experiments.

Chapter 2 - Base definitions

- ***Measurement*** – set of activities determining the measured value.
- ***Measurement uncertainty*** - parameter associated with the result of measurement characterizing dispersion of the values attributed to the measured quantity (measurand).
- ***Standard uncertainty u(x)*** – the uncertainty of measurement expressed as a standard deviation. Uncertainty can be reported in three different ways: *u, u(x)* or *u*(*acceleration*), where quantity *x* can be expressed also in words (in the example *x* is *acceleration*). Please note, that ***u* is a number**, not a function.
- ***Type A evaluation of uncertainty*** – the evaluation of uncertainty by the statistical analysis of series of observations.
- ***Type B evaluation of uncertainty*** – the evaluation of uncertainty by means other than the statistical analysis of series of observations, thus using method other than in type A.
- ***Combined standard uncertainty $u_c(x)$*** – represents the estimated standard deviation of the result and is obtained by combining the individual standard uncertainties (both type A and type B), using the usual method combining standard deviations.
- ***Expanded uncertainty U(x)*** or ***$U_c(x)$*** – the measure of uncertainty that defines interval about the measurement that may be expected to encompass a large fraction of the distribution of values that

Warsaw, October 2017

could reasonably be attributed to the measurand. The expanded uncertainty is especially applicable for some commercial, industrial and regulatory applications.

- ***Coverage factor*** *k* – numeric multiplier of the standard uncertainty used for expanded uncertainty determination. Typically, *k* is in the range 2 to 3. For the most laboratory cases it is recommended to use k = 2, which defines an interval having a level of confidence of approximately 95%.

Chapter 3 - Uncertainty in measurements

The aim of the measurement is to determine the measured value. Thus, the measurement begins with specifying the quantity to be measured, the method used for measurement (e.g. comparative, differential, etc.) and the measurement procedure (set of steps described in detail and applied while measuring with the selected measuring method). In general, the result of a measurement is only an approximation or estimate of the value of the specific quantity subject to measurement, that is, the **measurand**. Thus, the result of measurement is complete only when accompanied by a quantitative statement of its uncertainty.

It must be underlined and always remembered, that report of measurement result consists of determined value of the measured quantity and measurement uncertainty.

There are many possible sources of uncertainty in a measurement, including:

1. incomplete definition of the measurand;
2. imperfect realization of the definition of the measurand;
3. non-representative sampling — the sample measured may not represent the defined measurand;
4. inadequate knowledge of the effects of environmental conditions on the measurement or imperfect measurement of environmental conditions;
5. personal bias in reading analogue instruments;
6. finite instrument resolution or discrimination threshold;
7. inexact values of measurement standards and reference materials;
8. inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm;
9. approximations and assumptions incorporated in the measurement method and procedure;
10. variations in repeated observations of the measurand under apparently identical conditions.

These sources are not necessarily independent, and some of sources (sources from 1 to 9) may contribute to source 10.

Two categories of uncertainty can be distinguished based on their method of evaluation, "Type A" and "Type B" [2].

Type A evaluation of standard uncertainty may be based on every statistical data analysis methods. For example, the standard deviation of a series of independent observations can be calculated, or at least squares method can be applied to fit the data with a curve and determine its parameters and their standard uncertainties.

Type B evaluation of standard uncertainty is usually based on scientific judgment taking into account all available information including:

- previous measurement data;
- experience with or general knowledge of the behavior and properties of relevant materials and instruments;
- manufacturer's specifications;
- data provided in calibration and other certificates;
- uncertainties assigned to reference data taken from handbooks.

The proper use of the pool of available information for a Type B evaluation of standard uncertainty calls for insight based on experience and general knowledge, and is a skill that can be learned with practice.

Chapter 4 - Evaluation of uncertainties

If the measured quantity can be directly compared with the external standard, or if the measurement is made using a single instrument giving result straightaway, is called *direct measurement*. This type of measurement include: length measurement with a ruler, diameter measurement of the rod using a micrometer, time measurement with a timer, measurement of electric current with an ampermeter or a voltage measurement with a voltmeter.

In many cases, it is necessary to measure one or more physical quantities to determine quantity dependent on them. This type of measurement is called *indirect measurement*. These include for example: the measurement (determination) of resistance by measuring current and voltage, determination of the cylinder volume by measuring its diameter and height, the measurement of gravitational acceleration based on the length and the period of oscillation pendulum.

Methods used for calculating the measurement uncertainty depend on whether the measurements were made directly or indirectly.

4.1 Direct measurements

Consider input quantity X determined in a direct way, which value is estimated from n independent measurements x_1, x_2,, ..., x_n. If one of the measured values differ significantly from the other (*gross error*), it should be disregarded and must not be taken into account in further calculations. In the most cases gross errors are caused by the investigator (e.g. reading 121 V instead of 12.1 V) or by a momentary disruption of the measurement conditions. The decision to recognize gross error depends on the investigator and is usually taken at the stage of interpretation of the results.

Type A evaluation of standard uncertainty

The set of independent n measurements x_1, x_2,, ..., x_n can considered as a n-element random sample of the infinite set of measurements. If the probability distribution for x_i is described by Gaussian function (see Appendix B), the following way of data analysis can be applied. The arithmetic mean value should be considered as a result of the measurement:

$$x \equiv \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \,. \qquad \textbf{(1)}$$

It should be noted, that the bigger number of measurements results in a better mean value. Standard deviation of the mean value of the measurement of quantity X is equal:

$$u(x) = \sqrt{s_{\bar{x}}^2} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2}\,, \qquad \textbf{(2)}$$

where $s_{\bar{x}}$ ***is called the standard deviation of the mean value*.**

Type B evaluation of standard uncertainty

Quite often in the laboratory work only one measurement is performed (or a single measurement of each measured quantities) or measured values show no spreading. This can happen especially, when the measuring device has low accuracy. For example, when measuring the thickness of the plate with micrometer screw, set of different results will be obtained, but if the millimeter ruler will be used to measure the same plate we will get always the same result. The accuracy of used meter device determines the ***calibration uncertainty Δx*** (also called the ***uncertainty level***). This is the number specified by the manufacturer of the measuring device or estimated on the basis of the scale interval used in the device. The probability of any result within the range defined by measured value and calibration accuracy is the same. Such a probability distribution is called a *uniform distribution*, and the standard deviation in this case is defined by formula $\Delta x/\sqrt{3}$ (Appendix B). ***The Type B standard uncertainty*** is equal to this value:

$$u(x) = \frac{\Delta x}{\sqrt{3}} = \sqrt{\frac{(\Delta x)^2}{3}} \tag{3}$$

The other source of Type B measurement uncertainty may be ***investigator uncertainty*** Δx_e and its value is estimated on the basis of investigators experimental skills and type of performed measurement. The standard uncertainty in this case is also calculated using the formula (3), where Δx should be replaced with Δx_e. In the case, where two sources of Type B uncertainty are observed both standard deviations should be added:

$$u(x) = \sqrt{\frac{(\Delta x)^2}{3} + \frac{(\Delta x_e)^2}{3}}\,. \tag{3a}$$

Combination of uncertainties

If there are two types of uncertainties in the same experiment (Type A - dispersion of results and Type B - the uncertainty of calibration and investigator), they should be added using following formula for the standard total uncertainty:

$$u(x) = \sqrt{s_{\bar{x}}^2 + \frac{(\Delta x)^2}{3} + \frac{(\Delta x_e)^2}{3}}\,. \tag{4}$$

It should be noted, that if one of the calculated uncertainties is more than one order of magnitude smaller it could be neglected. Equation (4) should be applied only in the case, when all of calculated standard deviations are the same order of magnitude.

Determination of calibration uncertainties and investigator uncertainties for basic instruments used in the laboratory

- **Mechanical devices** (rulers, micrometers, calipers) - Δx is equal to half of the interval. Calibration uncertainty for mercury barometers, thermometers, stoppers etc. is defined the same way.

- **Analogue devices** - calibration uncertainty Δx can be determined using the class of the instrument and chosen measuring range:

$$\Delta x = \frac{class \cdot range}{100}\,. \tag{5}$$

Only the investigator observing the pointer during the measurement process can estimate the investigator uncertainty.

- **Digital devices** (electronic) - measurement uncertainty is defined in the technical data of the device specified in the manual. It depends mostly on the measured value of x and in the smaller extent on the used range:

$$\Delta x = c_1 x + c_2 z\,, \tag{6}$$

where c_1 and c_2 are device constants specified in user manual - for the most of digital voltmeters used in the laboratory $c_1 = 0.05\%$ and $c_2 = 0.01\%$, however in many cases one can apply $c_2 = 0$.

Please note that in order to determine Type B standard uncertainty, calculated above Δx values should be divided by $\sqrt{3}$ and possibly use the law of propagation of uncertainty.

4.2 Indirect measurements

Indirect measurement of the quantity Z (called output) is performed by measuring k values measured directly (called input), marked as $x_1, x_2, ..., x_k$. Desired value z depends on x_k:

$$z = f(x_1, x_2, ..., x_k) \text{ or } z = f(x_k).$$

For every directly measured quantity, its mean value $\bar{x}_1, \bar{x}_2, ..., \bar{x}_k$ and standard uncertainties $u(x_1)$, $u(x_2)$, ..., $u(x_k)$ should be determined. Standard uncertainties can be calculated by both, Type A and Type B methods. Obviously, in the case of method B, there is no average value but only the measurement result. The measured value of Z is calculated using the formula:

$$\bar{z} = f(\bar{x}_1, \bar{x}_2, ..., \bar{x}_k).$$

The uncertainty of Z is called the ***combined uncertainty*** $\boldsymbol{u_c}$ and is calculated using the following formula (the ***law of propagation of uncertainty***):

$$u_c(z) = \sqrt{\sum_{j=1}^{k} \left(\frac{\partial f(x_j)}{\partial x_j} \right)^2 u^2(x_j)}. \quad \textbf{(7)}$$

When the derivative functions are calculated, x_j values should be replaced with the mean values $\bar{x}_j$.
In the case of two directly measured quantities (x and y) combined uncertainty is equal:

$$u_c(z) = \sqrt{\left(\frac{\partial f(x,y)}{\partial x} \right)^2 u^2(x) + \left(\frac{\partial f(x,y)}{\partial y} \right)^2 u^2(y)}. \quad \textbf{(8)}$$

In the case of direct measurements one can consider measurements of ***correlated*** and ***uncorrelated quantities*** (correlated and uncorrelated measurements).

In the ***uncorrelated measurements*** every quantity is measured directly but in a different, independent experiment (e.g. measured and calculated at different time). As an example of uncorrelated measurements we could consider experiments for determination of the gravitational acceleration using a pendulum. There are two independent measurements of the length of the pendulum and the oscillation period and based on this the acceleration is calculated using following formula: $g = \frac{4\pi^2 l}{T^2}$.

In the ***correlated measurements*** measured quantities are somehow dependent on each other. In this case it is important to perform all measurements of input quantities under exactly the same experimental conditions, without introducing any changes in the measurement system. However, virtually in all experiments in the Physics Laboratory only uncorrelated measurements are performed and therefore combined uncertainty should be calculated using formula (7) or (8).

4.3 Extended uncertainty

The standard uncertainty $u(x)$ defines the interval from $\bar{x} - u(x)$ to $\bar{x} + u(x)$, within which the value of X can be asserted to lie with the probability of 68% for the Type A uncertainty, and with the probability of 58% for Type B uncertainty (these values are the result of Gaussian and uniform probability distributions, respectively). The standard uncertainty is a measure of the accuracy of measurements and allows for the comparison of different methods of measurement.

To make possible comparison of measurement results obtained in different laboratories and under different conditions, the concept of the extended uncertainty $\boldsymbol{U}$ was introduced. Extended uncertainty is commonly used to allow comparison of measured results with the reference data, but also for commercial purposes and to establish standards of industrial health, safety, etc. ***Extended uncertainty*** $\boldsymbol{U(x)}$ is equal to standard uncertainty value $u(x)$ multiplied by the *coverage factor k*, so that in the in-

terval $\bar{x} \pm U(x)$ the ***large fraction*** of the distribution of measured values could reasonably be attributed to *X*.

$$U(x) = k \cdot u(x), \tag{9}$$

In the most cases in the laboratory practice the coverage factor $k = 2$ should be used, since for this value the probability of finding the true value of *X* in the interval $\bar{x} \pm U(x)$ is equal to 95% to Type A uncertainty and 100% for Type B uncertainty (probability equal to 100% for Type B uncertainty is achieved already for k = 1.73!).

For the laboratory reference devices usually extended uncertainty with the coverage factor to $k = 3$ is presented as a measure of uncertainty. In this case the standard uncertainty, corresponding to a probability of 68%, is equal to 1/3 of the specified value of uncertainty.

4.4 Reporting uncertainty

It should be noted, that ***the report of measurement result must consists of measured value and measurement uncertainty*** and both of quantities should be expressed using SI units (see Appendix A). To report results of measurements properly one should start from the correct recording of uncertainty. Uncertainty is presented with accuracy (rounded) to two significant digits. The measurement result (the most probable value) is presented with an accuracy specified by the uncertainty, which means that the last digit of the measurement result and the measurements uncertainty must be at the same decimal place. Rounding of uncertainties and measurement follows the mathematical rules of rounding: digits 0-4 are rounded down and the 5-9 digits are rounded up.

Standard uncertainty can be reported in a number of ways:
(1) $t = 21.364$ s, $u(t) = 0.023$ s
(2) $\boldsymbol{t = 21.364(23)}$ **s**, – this is the most common and recommended way to report results for scientific publications and reference data
(3) $t = 21.364(0.023)$ s

In notation (2) in the brackets two significant digits of standard uncertainty are presented whereas in the notation (3) standard uncertainty is presented in full form but with the same accuracy of 2 significant digits.

Extended uncertainty is only noted with the ***symbol*** ±.
For the above mentioned example $U(t) = k \cdot u(t)$, $t = 21.364$ s, $U(t) = 0.046$ s ($k = 2$), $n = 11$

$t =$ **(21.364±0.046)** s.

Chapter 5 - Measurements of functional type relations

In the typical laboratory cases of indirect measurements are ***multiple, almost simultaneous measurements*** of two quantities *x* and *y* dependent on each other. For different values of x_i different values of y_i will be obtained and *n* pairs of numbers (x_i, y_i) will be the result of the measurement.

If there is a known function relating measured quantities *x* and *y* (e.g. $y = ax^2$) it is possible to determine function parameters (in this example, the parameter *a* by fitting collected data with the known functional relationship. The fitting procedure provides also standard uncertainty of determined parameters (Type A uncertainty). However this uncertainty doesn't depend on the uncertainty of the measured values x_i, y_i and to take it into account complex uncertainty $u_c(z)$ has to be calculated for one of the measured pairs x_i, y_i (Type B uncertainty). Two types of uncertainties should be combined using uncertainty propagation rule.

It is also possible to verify theoretical dependence between two measured quantities by fitting collected data with the known function. For example one can check the applicability of Ohm's law by measurement of current-voltage characteristic for unknown resistance. For this purposes different methods of fitting can be applied with the most known and widely used least squares method, which will be described later in this guidebook.

5.1 Linear type functions

To verify theoretical model, dependence between measured quantities collected data points have to be fitted to the known function. In some cases this relation can be quite complicated including also implicit type function. Usually, the physical model also provides a range of values, for which model equation is valid. The task of the investigator is to perform as many measurements as possible in the range of applicability of the model to obtain the best fit. Modern computer software allows fitting any functional relationship to the set of measurements. However, the most of relationships in physics can be reduced to a linear form (linearized). Such a linearization is based on the transformation of function $y = f(x)$ to another function $Y = F(X)$, which will take the form of a first degree polynomial $Y = BX + A$.

Some examples of transformation to linear dependence:

$$y\lambda = 2d\sin x \qquad \Rightarrow \qquad X = 2\sin x,\ \ Y = y\lambda,\ \ B = d$$

$$y = y_0 e^{-\frac{x}{\tau}} \qquad \Rightarrow \qquad X = x,\ \ Y = \ln\left(\frac{y}{y_0}\right),\ \ B = -\frac{1}{\tau}$$

$$y = y_0 e^{-\frac{x}{\tau}} \qquad \Rightarrow \qquad X = \ln\left(\frac{y_0}{y}\right),\ \ Y = x,\ \ B = \tau$$

Mentioned above transformations are unique for function $y = f(x)$. Two sets of experimental data $(x_1, x_2, x_3, \ldots x_n)$ and $(y_1, y_2, y_3, \ldots y_n)$, should be transformed to $(X_1, X_2, X_3, \ldots X_n)$ and $(Y_1, Y_2, Y_3, \ldots Y_n)$ and plotted $Y=F(X)$. In the next step fitting procedure using linear type function has to be applied.

5.2 Least squares method

The most commonly used method of linear regression analysis is the least squares method described in details in Appendix C. The goal of this method is to determine parameters of modelling function to lie as close as possible to all experimental points. For this purpose sum of squared residuals is calculated, where residual is the difference between an observed value y_i and the value provided by the model $(Bx_i + A)$. Parameters A and B are modified iteratively to minimize sum of squared residuals.

$$\sum_{i=1}^{n}[y_i - (Bx_i + A)]^2 = \min.$$

Students can prepare their own procedures to apply least squares method, following the information in the Appendix C, but it is much more convenient to use some software with *linear fit* or *linear regression* options. As a result slope of the line (parameter B) and its intercept (parameter A) as well as their standard uncertainties $u(B)$ and $u(A)$ can be obtained. It should be noted, that there are different types of least squares methods. In the simplest case uncertainties of measured values x_i and y_i are unknown, but assumed to be equal. In this case standard uncertainties $u(B)$ and $u(A)$ doesn't depend on the uncertainties of measured values. *The **least squares method** is one of the statistical methods and therefore it gives a **Type A uncertainty**.*

Appendix D shows an example of using least squares method with MicroCal Origin 8 software.

5.3 Verification of linear dependence hypothesis

Linear regression models, including least squares method, can be used to model experimental data. However, not always linear function offers proper fit to measured values. Most of the fitting software packages calculates also value of *correlation factor*, which is a number from the range [-1, 1] describing correlation between variables. Unfortunately in the most laboratory cases this parameter is close to 1 or -1 and doesn't provide enough information on the deviation from linearity. Therefore other test should be performed to verify linear dependence for measured values.

One of the possibilities is the graphical analysis. At first data points should be plotted, including uncertainty line segments of measured quantities. If the theoretical model line crosses uncertainty line segments for less than 2/3 of experimental data points, hypothesis of linear dependence should be rejected, even if the correlation factor is high.

The most commonly method used for hypothesis verification is χ^2 test. The variable χ^2 is defined:

$$\chi^2 = \sum_{i=1}^{n} w_i (y_i - y(x_i))^2 \,, \tag{10}$$

where $y(x_i)$ value of the tested hypothesis function at point x_i, w_i is a statistical weight value for point number i, which is calculated according to following formula:

$$w_i = [u(y_i)]^{-2} \,. \tag{11}$$

Let us consider verification of linear dependence using χ^2 test. In this case value of $(y_i - y(x_i))^2$ is equal to squared difference between measured value y_i and value of tested function at point $x = x_i$, $(y_i - B(x_i) - A)^2$. For every measurement point this value divided by the squared standard uncertainty shouldn't be bigger than 1. In fact value bigger than 1 means, that fitted line doesn't cross uncertainty line segment for this point. If the fitted line crosses all uncertainty line segments, test function ***χ^2 should not be bigger than number of measured points n*.**

Calculated value of χ^2 function provides information on correlation between experimental data and fitted line. The second very important parameter of χ^2 test is statistical ***significance value α***. Significance level describes level of confidence about linear dependence of measured data. The value of significance varies from 0 to 1 and is an input parameter set by investigator. For the high value of α also many good points (lying close to model line) will be rejected and for the very small values also points not fitting to model line will be considered. In the Laboratory experiments typical value of ***α = 0.05*** should be applied. It is also strongly suggested to analyse series of data not smaller than n = 6 points.

In the typical case χ^2 value determined for experimental data points is compared with critical value $\chi^2_{critical}$ for given *significance value* and given *number of degree of freedom*. Degree of freedom is equal to n number of data points reduced by number of parameters used in fitting procedure (for least squares procedure there are two parameters B and A). Critical values $\chi^2_{critical}$ for different significance value and degree of freedom values are presented in Appendix E.

1. $\chi^2 \leq \chi^2_{critical}$ - hypothesis on linear dependence between measured values can be accepted
2. $\chi^2 > \chi^2_{critical}$ - hypothesis on linear dependence between measured values should be rejected

If the value of χ^2 function for experimental data is *significantly* higher than critical value, one should consider if the measurement uncertainty isn't too big or measurements shouldn't be repeated with higher accuracy measurement devices.

If the linear dependence hypothesis is rejected, the other model should be tested.

An example of χ^2 test is presented in the Example 3 in the next Chapter.

SUMMARY: least squares method applied for correlated quantities allows determining value of parameter correlating these quantities. However, at first fitted line should be plotted and investigator should check if it crosses uncertainty line segments. The second stronger condition is a χ^2 test, which allows accepting or rejecting hypothesis regarding set significance level.

Examples

Reporting measurement results

Results of experiment and calculations	**Reporting results**
a = 321.735 m/s; $u(a)$ = 0.24678 m/s	a = 321.74 m/s; $u(a)$ = 0.25 m/s a = 321.74(25) m/s a = 321.74(0.25) m/s
b = 321785 m; $u(b)$ = 1330 m	b = 321800 m; $u(b)$ = 1300 m b = 321800(1300) m $b = 321.8(1.3)\cdot 10^3$ m b = 321.8(1.3) km
C = 0.0002210045 F; $u_c(C)$ = 0.00000056 F	C=0.00022100 F; $u_c(C)$=0.00000056F $C = 221.00(56)\cdot 10^{-6}$ F $C = 221.00(0.56)\cdot 10^{-6}$ F C = 221.00(56) μF
T = 373.4213 K; $u(T)$ = 2.3456 K	T = 373.4 K; $u(T)$ = 2.3 K T = 373.4(23) K **$U(T)$ = 4.7 K (k=2)** **T = (373.4 ± 4.7) K**
R = 7885.666 Ω; $u_c(R)$ = 66.6667 Ω	R = 7886 Ω; $u_c(R)$ = 67 Ω R = 7886(67) Ω R = 7.886(0.067) kΩ **$U_c(R)$ = 130 Ω (k=2)** **R = (7890 ± 130) Ω** **R = (7.89 ± 0.13) kΩ**
x = 1.12345 A; $u(x)$ = 0.00011111 A	x = 1.12345 A; $u(x)$ = 0.00011 A x = 1.12345(11) A x = 1.12345(0.00011) A
y = 1.12 A; $u(y)$ = 0.00011111 A	y = 1.12000 A; $u(y)$ = 0.00011 A y = 1.12000(11) A y = 1.12000(0.00011) A

NOTES:
(1) Notation underlined is RECOMMENDED and should be used
(2) Notation bolded refers only to **extended uncertainty**
(3) It is allowed also to report results by text e.g. „The speed of sound in air is equal to 321.74 m/s with standard uncertainty 0.25 m/s”

Example 1

The dimension of squared cross-section rod was measured with calliper with accuracy of 0.1 mm. Obtained results are: 12.5; 12.3; 12.6; 12.5; 12.3; 12.5; 12.7; 12.3; 12.7; 12.4; 12.3. Determine size of the rod. Report result properly.

Measured quantity (the length of rod side - *d*) was measured directly by a series of measurements. The result of experiment is a mean value according to formula (1):

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i = \frac{1}{11}\sum_{i=1}^{11} d_i = 12.46364 \text{ mm}$$

Since the calliper's interval is equal to 0.1 mm calibration uncertainty Δd is twice smaller – 0.05 mm. Type B uncertainty can be calculated:

$$u(d) = \frac{\Delta d}{\sqrt{3}} = \frac{0.05}{\sqrt{3}} = 0.02887 \text{ mm}$$

Since distribution of results is observed Type A standard uncertainty should be calculated using formula (2):

$$u(d) = \sqrt{s_{\bar{d}}^2} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(d_i - \bar{d})^2} = \sqrt{\frac{1}{11(11-1)}\sum_{i=1}^{11}(d_i - 12.46364)^2} = 0.047238 \text{ mm}$$

Both uncertainties Type A and Type B are the same order of magnitude, therefore uncertainty propagation rule should be used to determine overall standard uncertainty:

$$u(d) = \sqrt{s_{\bar{d}}^2 + \frac{(\Delta d)^2}{3}} = \sqrt{0.047238^2 + 0.02887^2} = 0.055361 \text{ mm.}$$

Standard uncertainty can be presented in different ways:
u = 0.055 mm
u(d) = 0.055 mm
u(side of the rod) = 0.055 mm

Proper reporting of the result:
d = 12.464 mm. u(d) = 0.055 mm
d = 12.464(55) mm
d = 12.464(0.055) mm

If the extended uncertainty is required (e.g. to be compared with literature reference data), then result should be reported:
U(d) = k·u(d). d = 12.46 mm. U(d) = 0.11 mm (k = 2). n = 11
d = (12.46±0.11) mm.

Example 2 [4]

To determine gravitational acceleration, time of flight of a body dropped from height h was measured. The height h was measured 3 times with a line and every time the same value of 1270 mm was obtained. Time of flight was measured 5 times using a stopper with an accuracy 0.001 s and recorded values are: $t_1 = 0.509$, $t_2 = 0.512$, $t_3 = 0.510$, $t_4 = 0.504$, $t_5 = 0.501$ s. The uncertainty related with the investigator and start and stop moments was estimated to be 0.01 s. Calculate gravitational acceleration and its uncertainty.

Gravitational acceleration g can be calculated using formula $g = \frac{2h}{t^2}$. At first the average height $\bar{h}$ and average time $\bar{t}$ is calculated using formula (1):

$$\bar{h} = 1270mm = 1.27m,\ \bar{t} = 0.5072s\ .$$

Now, based on $\bar{h}$ and $\bar{t}$ values the $\bar{g}$ value can be calculated:

$$\bar{g} = \frac{2 \cdot 1.27}{0.5072^2} = 9.87359\ \mathrm{m/s^2}.$$

To determine uncertainty of indirectly measured g standard uncertainties of time and height should be calculated at first.

Standard uncertainty *u(t)*:

Type A uncertainty:

Formula (2) will be used:

$$u(t) = \sqrt{s_{\bar{t}}^2} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(t_i - \bar{t})^2} = \sqrt{\frac{1}{5 \cdot 4}\sum_{i=1}^{5}(t_i - 0.5072)^2} = 2.035 \cdot 10^{-3}\ \text{s} = 2.035\ \text{ms}.$$

Type B uncertainty:

The uncertainty related with the investigator and his decision when to start and stop measuring time of flight was estimated to be $\Delta t_e = 0{,}01\ s = 10\ ms$ (accuracy of the stopper can be neglected since is one order of magnitude smaller). Therefore Type B standard uncertainty is equal:

$$u(t) = \frac{\Delta t_e}{\sqrt{3}} = \frac{10}{\sqrt{3}} = 5.7735\ \text{ms}.$$

It should be noted, that both types of uncertainties are the same order of magnitude and therefore uncertainty propagation rule should be used to determine overall standard uncertainty:

$$u(t) = \sqrt{2.035^2 + 5.7735^2} = 6.122\ \text{ms}.$$

The result of time measurement: **t = 0.5072(61) s, n=5**.

Standard uncertainty *u(h)*:

In this case there is no distribution of results and standard uncertainty of height measurement should be determined as the Type B uncertainty. The smallest interval in the line is equal to 1 mm but taking into account other factors (non-vertical alignment of line, investigators error) uncertainty of this measurement should be estimated as Δh = 2 mm.

Therefore standard uncertainty Type B is equal:

$$u(h) = \frac{\Delta h}{\sqrt{3}} = \frac{2}{\sqrt{3}} = 1.15\ \text{mm}.$$

Measurement of height can be reported: **h = 1270.0(12) mm** or **h = 1.2700(12) m, n = 1**.

Combined uncertainty of gravitational acceleration $u_c(g)$:

Since gravitational acceleration is measured based on two direct and uncorrelated measurements uncertainty propagation rule should be used (7):

$$u_c(g) = \sqrt{\left(\frac{\partial g}{\partial h}\right)^2 u^2(h) + \left(\frac{\partial g}{\partial t}\right)^2 u^2(t)} = \sqrt{\left(\frac{2}{\bar{t}^2}\right)^2 u^2(h) + \left(\frac{4\bar{h}}{\bar{t}^3}\right)^2 u^2(t)} =$$

$$= \sqrt{\left(\frac{2}{0.5072^2}\right)^2 \cdot 0.0012^2 + \left(\frac{4 \cdot 1.27}{0.5072^3}\right)^2 \cdot 0.006122^2} = \sqrt{8.7 \cdot 10^{-5} + 0.0564} = 0.237 \text{ m/s}^2.$$

If we compare two components of combined uncertainty under square root, one can clearly find that the first one related with the height uncertainty is negligibly small compared to time uncertainty component.

Notation reporting gravitational acceleration measurement results:
$g = 9.87$ m/s^2, $u_c(g) = 0.24$ m/s^2
$g = 9.87(24)$ m/s^2
$g = 9.87(0.24)$ m/s^2

Extended uncertainty $U_c(g)$:
According to formula (9):
$U_c(g) = 2 \cdot u_c(g) = 2 \cdot 0.237$ m/s^2 = 0.474 m/s^2.
The final result of gravitational acceleration:

$$\mathbf{g = (9.87 \pm 0.47) \text{ m/s}^2.}$$

This value can be compared with reference data according to which gravitational acceleration in Warsaw is equal to 9,80665 m/s^2. This value lay within range defined by extended uncertainty, and therefore it can be considered as a correct measurement.

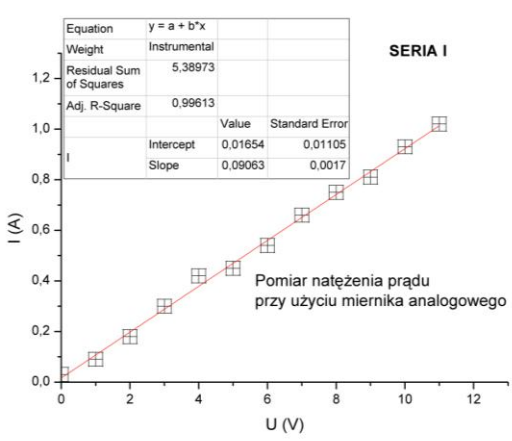
Example 3

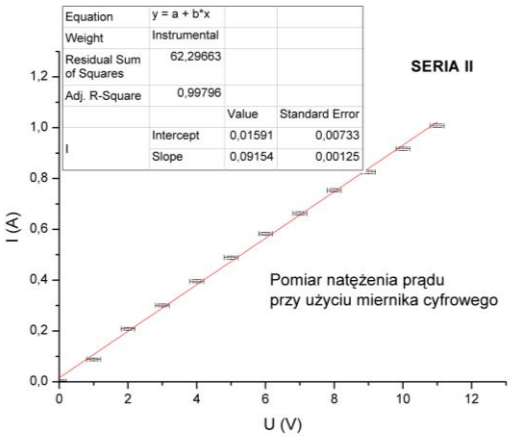
To verify linear type dependence between electric voltage applied to the resistor and measured current through this resistor two series of I and U measurements were performed for the same resistor of a nominal resistance equal to 10 Ω and a class 5% (extended uncertainty is equal to 0.5 Ω). In two series two different ammeters were used. In the first experiment analogue ammeter of class 2.5 measurement range 1.5 A, and with 50 intervals was used. In the second experiment digital ammeter was used and in this case formula (6) can be used to determine device uncertainty with c_1 = 0.2% and c_2 = 0.1% for the measurement range 10 A. The voltage was measured with analogue voltmeter of a range 15 V and class 2, in both cases. Measured values are presented in table. Is it possible to confirm existence of linear dependence between I and U? Does the measured resistance correspond to nominal value?

Measurement results:

SERIES I				***SERIES II***			
U (V)	***I (A)***	$u_i(y_i)$	$u_i(x_i)$	***U (V)***	***I (A)***	$u_i(y_i)$	$u_i(x_i)$
0.0	0.03	0.028	0.17	0.0	0.0030	0.0058	0.17
1.0	0.09	0.028	0.17	1.0	0.0877	0.0059	0.17
2.0	0.18	0.028	0.17	2.0	0.2080	0.0060	0.17
3.0	0.30	0.028	0.17	3.0	0.3005	0.0061	0.17
4.0	0.42	0.028	0.17	4.0	0.3960	0.0062	0.17
5.0	0.45	0.028	0.17	5.0	0.4886	0.0063	0.17
6.0	0.54	0.028	0.17	6.0	0.5823	0.0065	0.17
7.0	0.66	0.028	0.17	7.0	0.6626	0.0065	0.17
8.0	0.75	0.028	0.17	8.0	0.7536	0.0066	0.17
9.0	0.81	0.028	0.17	9.0	0.8255	0.0067	0.17
10.0	0.93	0.028	0.17	10.0	0.9172	0.0068	0.17
11.0	1.02	0.028	0.17	11.0	1.0073	0.0069	0.17

Validity of Ohm's law was assumed U = R·I and I(U) plots were prepared for both data series.





The linear fitting procedure using Microcal Origin software was applied for experimental data (see insert in the figures above) and obtained parameters of the fit are listed below:

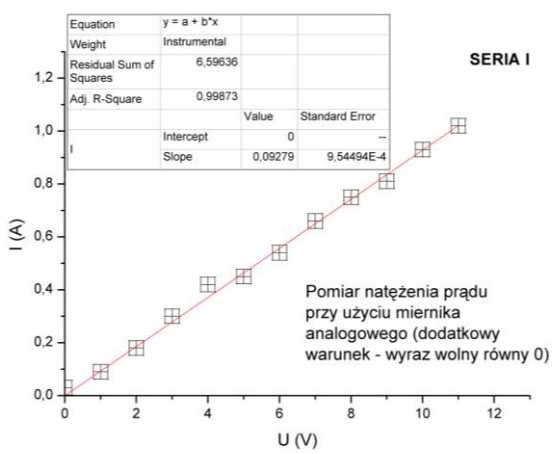
Series I
B = 0.09063
A = 0.01654
χ^2 = 5.39

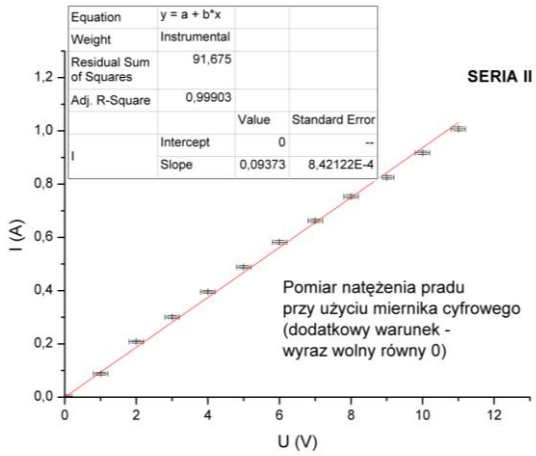
Series II
B = 0.09154
A = 0.01526
χ^2 = 62.3

In both cases, fitting lines cross all of the uncertainty boxes and the correlation factor for both series is similar and close to 1 (see Appendix D). In addition, slope values (parameter B) for both series are very similar, thus linear type of dependence seems to be confirmed. However, there is a significant difference in the χ^2 parameter values. For the studied case number of degrees of freedom is equal to 10 (12 measurement points minus 2 fitted parameters) and for the significance value 0.05 critical

value of χ^2 is equal to 18.3. Therefore, based on the first experimental series linear dependence hypothesis cannot be rejected – 5.39 < 18.3. However, based on the second experimental series, with a higher accuracy devices, **linear dependence hypothesis have to be rejected.** It is worth to note, that small vale of χ^2 for the first data series is suggesting, that the uncertainty of measurements are too high to use it in verification of linearity hypothesis. Indeed, only second series of measurements was good enough to reject mentioned hypothesis.

This somehow surprising result can be also confirmed in other calculations. If we consider, that for an applied voltage equal to 0 no current should be measured, then fitting linear function has to cross additionally (0,0) point and consequently intercept parameter A should be zero. In this case, number of degrees of freedom is actually 11 (12 measurement points minus 1 for one fitted parameter), and for significance value of 0.05 critical value for χ^2 is equal to 19.7. The new fitting parameters are listed below:





Series I
B = 0.09279
χ^2 = 6.59

Series II
B = 0.09373
χ^2 = 91.7

The χ^2 value for the first data series almost didn't change for the new fit, but for the second series there is almost 50% change of χ^2 value confirming, that hypothesis on linear dependence between *U* and *I* should be rejected. Obtained results suggest, that the measured resistance is changing with the applied voltage. Indeed, additional experiment showed that the temperature of the resistor was increasing due to the Joule-Lenz heat and therefore its resistance was ca. 10% higher at the end of experiment. Experimental data points were fitted with a new model function including temperature changes $I = U/R - k \cdot U^2$ (*k* is a parameter) (Fig.1) and a resistance was determined to be *9.72 Ω* with a standard uncertainty of *0.12 Ω*. Obtained resistance result can be reported:

$$\mathbf{R = (9.72 \pm 0.24)\ \Omega.}$$

This agree with the technical specification for this resistor: R = (10.0 ± 0.5) Ω.

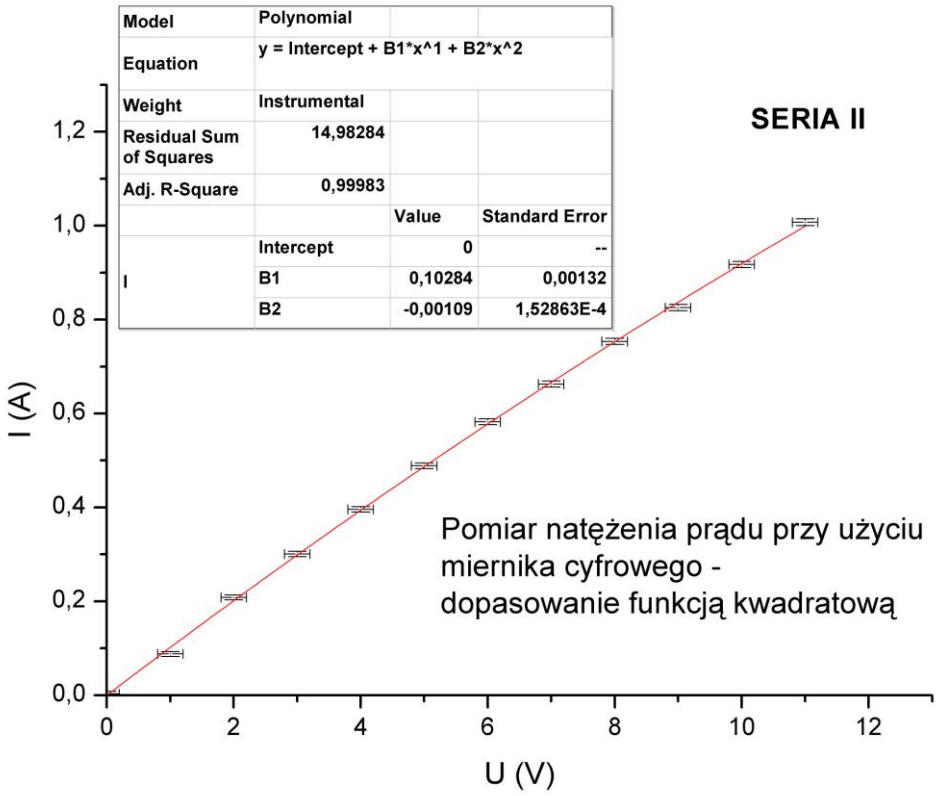


Fig. 1. Polynomial function fitting of data from Example 3.

I would like to thank Prof. Irmina Śledzińska and Piotr Panecki PhD for their help, labor and commitment during creating this guide. Also I would like to thank Prof. Andrzej Zięba for his detailed and critical revision of the text.

References

[1] *Wyrażanie niepewności pomiaru. Przewodnik*, Główny Urząd Miar, Warszawa 1999.
[2] *Guideline for Evaluating and Expressing the Uncertainty of NIST Measurements Results*, NIST Technical Note 1297
[2] H. SZYDŁOWSKI, *Niepewności w pomiarach*, Wydawnictwo Naukowe UAM, Poznań 2001.
[3] A. ZIĘBA, *Pracownia fizyczna*, Wydawnictwo AGH, Kraków 2002.
[4] M. LEWANDOWSKA, *Analiza niepewności pomiarowych*, in the internet.
[5] Wikipedia

APPENDIX A
SI Units

A.1 Base units

METRE (m) - The distance travelled by light in vacuum in 1/299792458 s (1983)

KILOGRAM (kg) - The mass of the international prototype kilogram placed in Sevres (1901).

SECOND (s) - The duration of 9192631770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium 133 atom ^{133}Cs (1964).

KELVIN (K) - 1/273.16 of the thermodynamic temperature of the triple point of water (1967/68).

MOL (mol) - The amount of substance of a system which contains as many elementary entities as there are atoms in 0,012 kilogram of carbon 12, ^{12}C (1971).

AMPER (A) - The constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 m apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newtons per meter of length (1948).

CANDELA (cd) - The luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of 1/683 watt per steradian (1979).

A.2 Derived units

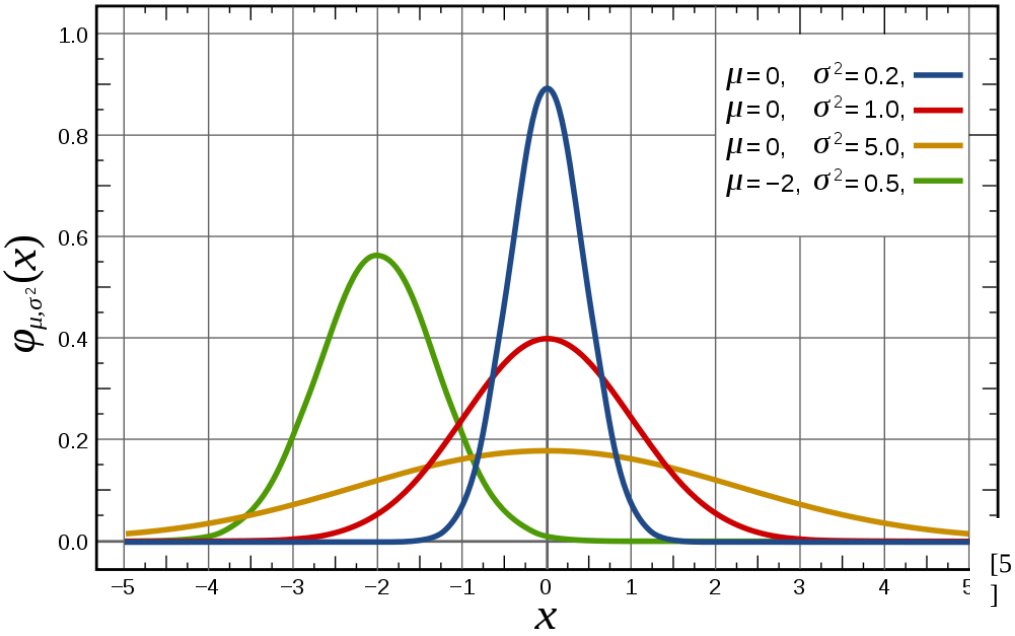
RADIAN (rd) - One radian is the angle subtended at the center of a circle by an arc that is equal in length to the radius of the circle.

STERADIAN (sr) - The solid angle subtended at the center of a unit sphere by a unit area on its surface.

A.3 Prefixes

Prefix name	Symbol	Factor
Exa	E	10^{18} = 1 000 000 000 000 000 000
Peta	P	10^{15} = 1 000 000 000 000 000
Tera	T	10^{12} = 1 000 000 000 000
Giga	G	10^{9} = 1 000 000 000
Mega	M	10^{6} = 1 000 000
Kilo	k	10^{3} = 1 000
Hecto	*h*	*10^{2} = 100*
Deca	*da*	*10^{1} = 10*
-	-	1
Deci	*d*	*10^{-1} = 0.1*
Centi	*c*	*10^{-2} = 0.01*
Milli	m	10^{-3} = 0.001
Micro	μ	10^{-6} = 0.000 001
Nano	n	10^{-9} = 0.000 000 001
Pico	p	10^{-12} = 0.000 000 000 001
Femto	f	10^{-15} = 0.000 000 000 000 001
Atto	a	10^{-18} = 0.000 000 000 000 000 001

APPENDIX B
B.1 Gaussian distribution (normal distribution)



[5]

If we assume that during our measurements, obtaining a result bigger or smaller compared to the mean value has the same probability, and the results far away from the mean value are less probable then ones close to the mean value, the distribution of results for **big number of measurements** may by estimated by the Gaussian curve:

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

Function $\varphi(x)$ is called the Gaussian or normal distribution. It depends on two parameters μ and σ (mean and variance). The integral of this function fulfill the condition:

$$\int_{-\infty}^{+\infty} \varphi(x)dx = 1.$$

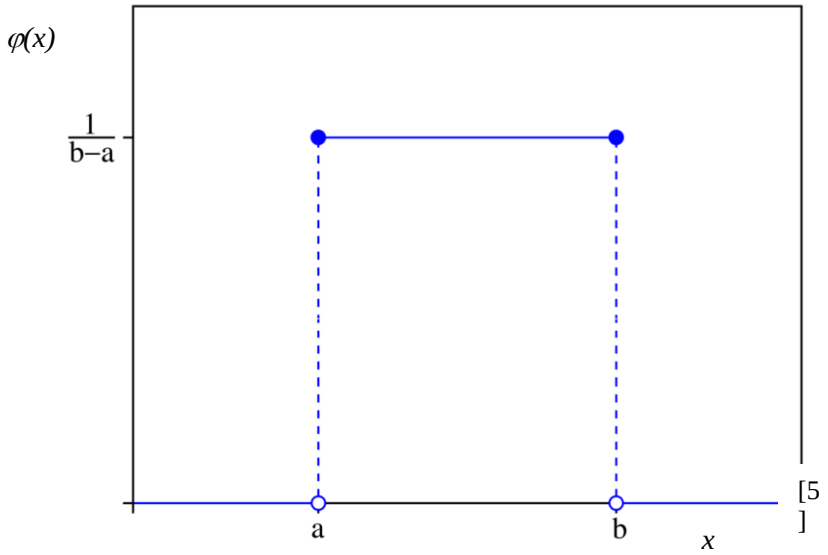
This condition means, that finding a result of the measurement in the interval from x to $x+dx$ is equal $\varphi(x)dx$, and the probability of finding any result in the interval from -∞ to ∞ must be equal 1. Parameters μ and σ can be easily interpreted analytically and statistically. For value $x = \mu$ function $\varphi(x)$ has its maximum. Parameter σ defines two points $\mu - \sigma$ and $\mu + \sigma$, where the Gaussian graph has inflection points. Hence, the value σ can be regarded as the measure of the distribution width. From the statistical point of view, μ is the expectation value $E(X)$, and the parameter σ is the square root of the variance $D^2(X)$, called the standard deviation.

The integrals of the function $\varphi(x)$ presented below define probability of finding the specific number of the measurements (68.3%. 95.4% and 99.7%) in the intervals, which width is the standard deviation multiple:

$$\int_{-\sigma}^{+\sigma} \varphi(x)dx = 0.683\,, \quad \int_{-2\sigma}^{+2\sigma} \varphi(x)dx = 0.954\,, \quad \int_{-3\sigma}^{+3\sigma} \varphi(x)dx = 0.997\,.$$

Gaussian distribution is continuous probability distribution, which can properly approximate the experimental dispersion of the measurements resulting from the causes described in Chapter 3. This distribution can be adapted to finite number of measurements to evaluate Type A measurement uncertainty. In such case, the expected value of the distribution is mean value (1), and the standard deviation is the standard deviation of the mean value (2).

B.2 Uniform distribution



The uniform distribution (also known as rectangular) is a symmetric probability distribution, in which probability density function is constant (not zero) in the interval from *a* to *b*, and outside this interval is equal to zero.

The probability density function of the uniform distribution is:

$$\varphi(x) = \frac{1}{2\sigma\sqrt{3}} \qquad \text{for } -\sigma\sqrt{3} \le x-\mu \le \sigma\sqrt{3},$$

$$\varphi(x) = 0 \qquad \text{for other } x,$$

where μ and σ (mean and variance) are as follows:

$$\mu = \frac{a+b}{2}, \text{ and } \sigma^2 = \frac{(b-a)^2}{12}.$$

If we assume that for the Type B measurement uncertainty the calibration uncertainty determines the interval with $2\boldsymbol{\Delta x}$ width about the μ value, then the formula (3) results directly from the variance definition (substitute $a = -\boldsymbol{\Delta x}$ i $b = \boldsymbol{\Delta x}$).

Note for the inquisitive persons:

For the uniform distribution, the coverage factors which guarantee probability for finding specific number of measurements (95.4% and 99.7%) in the intervals being the multiple of the standard deviation are different than for the normal distribution. 95% gets for k = 1.65, 99% for k = 1.71, and for k = 1.73 the probability is equal 100%.

APPENDIX C
The method of least squares

The method of least squares is the most popular analytical method for fitting the linear function to the experimental points. Its name is connected with the criteria of the fitting quality – the parameters of the linear function are chosen in such manner to minimize the sum of the squares residuals, a residual being the difference between an observed value y_i and fitted value $(Bx_i + A)$ provided by the model function.

$$S^2 = \sum_{i=1}^{n} [y_i - (Bx_i + A)]^2 = \min .$$

To find out the parameters B and A the condition for the two variables function minimum is used:

$$\frac{\partial S^2}{\partial B} = 0 \text{ and } \frac{\partial S^2}{\partial A} = 0 .$$

Calculating both partial derivatives creates the linear system of equations in the two unknown variables B and A. Further calculations are presented in the form enabling manual calculations. For every measured point, the values of auxiliary functions $\tilde{X}_i$, $\tilde{Y}_i$ and $\tilde{d}_i$ must be calculated:

$$\tilde{X}_i = X_i - \frac{1}{n}\sum_{i=1}^{n} X_i , \qquad \tilde{Y}_i = Y_i - \frac{1}{n}\sum_{i=1}^{n} Y_i , \qquad \tilde{d}_i = Y_i - \overline{B}\tilde{X}_i - \frac{1}{n}\sum_{i=1}^{n} Y_i .$$

Next the value of the slope B and the intercept A (ordinate of the point in which the linear function crosses axis OY) can be obtained:

$$B = \frac{\sum_{i=1}^{n} \tilde{X}_i \tilde{Y}_i}{\sum_{i=1}^{n} \tilde{X}_i^2} \qquad A = \frac{1}{n}\sum_{i=1}^{n} Y_i - \frac{B}{n}\sum_{i=1}^{n} X_i .$$

Formulas:

$$u(B) = \sqrt{\frac{1}{n-2} \cdot \frac{\sum_{i=1}^{n} \tilde{d}_i^2}{\sum_{i=1}^{n} \tilde{X}_i^2}} \qquad u(A) = u(B)\sqrt{\frac{1}{n}\sum_{i=1}^{n} \tilde{X}_i^2 + \left(\frac{\sum_{i=1}^{n} X_i}{n}\right)^2}$$

define the standard uncertainties of B and A. Nowadays, these calculations can be done using any software with linear regression or linear fit. On fig. 1 the results of the linear fit made by MicroCal Origin software are presented.

APPENDIX D
Linear regression in the MicroCal Origin 8.0 software

On the fig.2 the least square method in the MicroCal Origin 8 software (using function *linear fit*) is presented. Basic results of the calculation are displayed in the table, which appears automatically in the window *Graph*. Detailed information about the fitting and all statistical parameter can be found in the window *Book*, tab *Data*. In the table we can find the following information:

- *Equation* – function fitted to the experimental data. In our example we have linear function $y = a + b*x$.
- *Weight* – the method of calculating the statistical weight of the specific measured point. *Instrumental* means, that weight w_i is calculating as a reciprocal of the squared uncertainty y_i (value taken from the uncertainty column for quantity *Y*).
- *Residual Sum of Squares* – value of the function χ^2 (to display this value in the table with results, the option *Residual Sum of Square* in *Quantities to Compute>Fit statistics* in the window *Fit Linear* must be checked.
- *Adj. R-Square* – number that indicates how well data fit chosen statistical model. 1 indicates that the regression line perfectly fits the data, while 0 indicates that the line does not fit the data at all.
- *Value* and *Standard Error* – uncertainties for a and b.
- *Intercept* (*a*) and *Slope* (*b*).

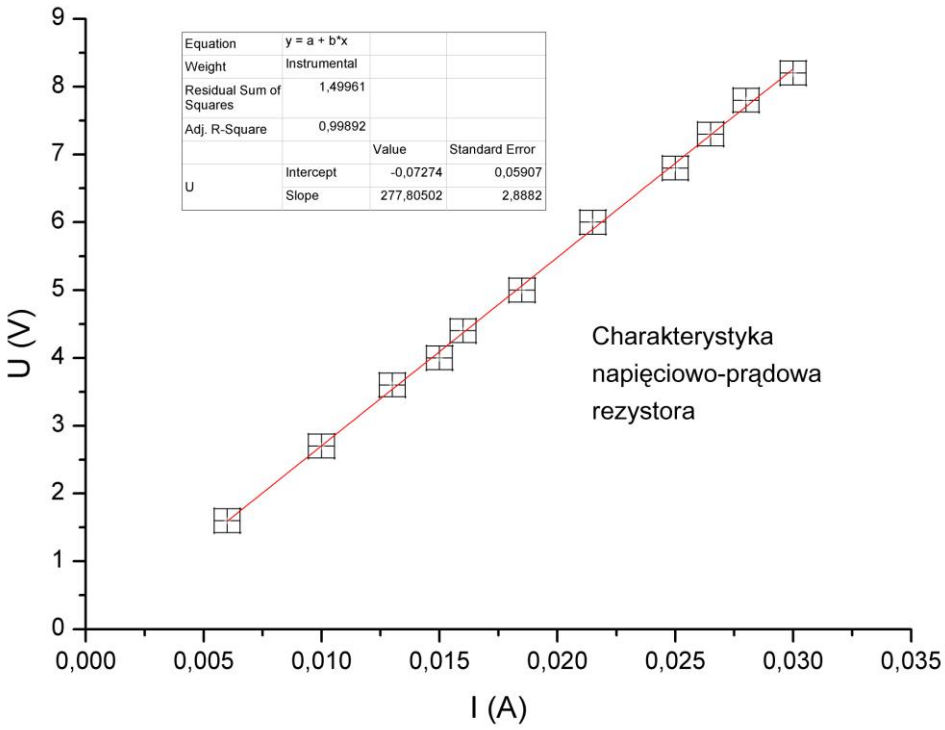


Fig. 2 Linear fit in the MicroCal Origin software

APPENDIX E
Critical values χ^2 for different significance levels α and the number of degrees of freedom

		Significance level α				
		0.20	**0.10**	**0.05**	**0.01**	**0.005**
Number of degrees of freedom	**1**	1.64	2.7	3.8	6.6	7.9
	2	3.22	4.6	6.0	9.2	11.6
	3	4.64	6.3	7.8	11.3	12.8
	4	6.0	7.8	9.5	13.3	14.9
	5	7.3	9.2	11.1	15.1	16.3
	6	8.6	10.6	12.6	16.8	18.6
	7	9.8	12.0	14.1	18.5	20.3
	8	11.0	13.4	15.5	20.1	21.9
	9	12.2	14.7	16.9	21.7	23.6
	10	13.4	16.0	18.3	23.2	25.2
	11	14.6	17.3	19.7	24.7	26.8
	12	15.8	18.5	21.0	26.2	28.3
	13	17.0	19.8	22.4	27.7	29.8
	14	18.2	21.1	23.7	29.1	31.0
	15	19.3	22.3	25.0	30.6	32.5
	16	20.5	23.5	26.3	32.0	34.0
	17	21.6	24.8	27.6	33.4	35.5
	18	22.8	26.0	28.9	34.8	37.0
	19	23.9	27.2	30.1	36.2	38.5
	20	25.0	28.4	31.4	37.6	40.0

Shaded gray column contains the critical values χ^2 most frequently used in the students laboratory.

Translation made by Wojciech Wróbel, PhD in the collaboration with Andrzej Kubiaczyk, MSc.