

Error Theory and Regression Analysis

Keywords:

Measurand, measured value, result of a measurement, uncertainty, Systematic and random error, absolute and relative error, frequency distribution, density function, Gaussian error curve, mean, standard deviation, variance, root mean square error, maximum error, law of error propagation, linear regression.

References:

- /1/ BIPM¹: “Evaluation of measurement data — Guide to the expression of uncertainty in measurement” (GUM), 2008
http://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf
- /2/ DIN 1319-3²: “Grundlagen der Messtechnik - Teil 3: Auswertung von Messungen einer einzelnen Messgröße, Messunsicherheit“, 1996
- /3/ DIN 1319-4: “Grundlagen der Messtechnik - Teil 4: Auswertung von Messungen, Messunsicherheit“, 1999
- /4/ NIST³ Technical Note 1297: “Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results”, 1994
<http://physics.nist.gov/Pubs/guidelines/TN1297/tn1297s.pdf>
- /5/ TAYLOR, J. R.: “Fehleranalyse“, VCH Verlagsgesellschaft mbH, Weinheim
- /6/ YOUNG, H. D.: “Statistical Treatment of Experimental Data“, McGraw-Hill, New York among others

1 Introduction

An experiment to determine the gravitational acceleration of the earth g is performed in a lecture hall. A ball is suspended in the air by a magnetic holder. After activation of the magnet, the ball drops onto a platform. It requires the time t for the fall through the distance s . By measuring the *measurands* s and t it is possible to determine the *measurand* g :

$$(1) \quad g = \frac{2s}{t^2}$$

The apparatus is constructed in such a manner that a flash of light is emitted upon the release of the ball and its impact on the platform. The time t between the two light flashes is measured by the students in the lecture hall using a stop watch. Nobody would expect that everyone will measure the same time. The individual *measured values* will deviate from each other. This is, for one, caused by the different reaction times of the students, and another thing to consider is the variation in drift and calibration among the individual stop watches. Likewise, the measured values for the distance s will differ when measured by several students since the positioning and reading of the measuring rod will vary from person to person. In addition, the measuring rod's calibration accuracy is limited as well.

This leads to the following questions:

- (1) Which values of s and t should be used for the determining g by Eq. (1) ?
- (2) How can the facts that the individual measured values for s and t deviate from each other, and the limited accuracy of the measuring instruments be accounted for?
- (3) How good is the value of g obtained from the measured values?

The answers to these questions are:

To (1): A single *result of the measurement* must be derived from the individual *measured values* by *fixed rules*. The *results of the measurement* for s and t are inserted into Eq. (1), thus providing a *result of the measurement* for g .

To (2): *Uncertainties* of the results of measurement for s and t have to be calculated according to *fixed rules*. The *uncertainties* provide a statistical measure for the deviations among the individual measured values. They (the measuring inaccuracies) are defined in a manner that an additional measurement of s or t will yield a result within the interval *result of the measurement* \pm *uncertainty* with a well-defined *probability*.

¹ BIPM: Bureau International des Poids et Mesures

² DIN: Deutsches Institut für Normung e.V.

³ NIST: National Institute of Standards and Technology des United States Department of Commerce - Technology Administration.

To (3): From the results of measurement and uncertainties for s and t an *uncertainty* for g must be calculated according to *fixed rules*. A further measurement of g using the same measuring procedure will then yield a result within the interval *result of the measurement* \pm *uncertainty* with a well-defined *probability*. Hence, both values together compose the *full result of the measurement*.

The *rules* mentioned above are internationally valid. They are described in great detail for all conceivable applications in norms and manuals. (e.g. /1/ - /4/). In addition, a number of books dedicated to this subject are available (e.g. /5/ and /6/). It is out of scope of this script to reiterate these rules in full detail. We will hence restrict ourselves to present some basics and provide the tools needed in the laboratory course for calculating *results of measurements* and *uncertainties*.

2 Direct and Indirect Measurement

In the example considered above, the measurands s and t may be measured directly, that is, by using a stop watch and measuring rod. One speaks of a *direct measurement* in such a case. The value g in the example is measured indirectly, that is by obtaining the result of measurement for g from the results of measurement of s and t . In such a case one speaks of an *indirect measurement*.

3 Notation

Following /2/, in the context of the measurement of a *measurand* the terms *result of a measurement* and *uncertainty* should be used. This terminology has, however, not become commonly used as of today in physics laboratories. More often the term “*error*” is used instead of *uncertainty*. For this reason, it is common to speak of “*error analysis*” instead of *calculation of uncertainties*, or about “*error bars*” instead of *bars of uncertainty*. In this text we will use both terms “*error*” and “*uncertainty*” (of a measurement).

4 Possible Kinds of Errors

4.1 Systematic Errors

Systematic errors may result from imperfect measuring instruments, from unavoidable influences of the surroundings on the measurement, or also from an unsuitable measurement method. This is illustrated by some examples from the lab course:

- (1) ***Imperfect measuring instruments:*** These are, e.g., an oscilloscope with a poorly adjusted time base unit, a multimeter with residual deviation, poorly adjusted electronic scales etc. The unpleasant thing about these faults is that one does not always notice them during the measurement. On the contrary: the value obtained (e.g., 27.5 μs , 147 Ω , 5.389 g) feigns an accuracy which one expects from these types of instrument, so there is no reason to doubt the results.
- (2) ***Influence of the surroundings:*** An example for this is the temperature dependence of measuring instruments. Normally, these dependencies are known quantitatively. They can be taken from manuals and considered for the analysis.
- (3) ***Unsuitable measurement methods:*** When one tries to determine the mass of a magnet using an electronic scale, one soon notices that the measured result is obviously absurd. The measurement method is unsuitable, because the magnetic field affects the mechanism of the scale; one needs to use another type of scale. It is considerably more difficult to judge whether the current measured in an electric circuit is negligibly affected by the circuit wiring and the internal resistance of the device. In this case it is not so obvious at first sight, whether the measured value is “right” or “wrong”. Hence, one can, in general, not depend on simply hoping to notice when a wrong method of measurement was used. Rather, it is necessary to carefully consider the appropriate measurement method to be applied when planning the experiment.

Systematic errors cannot be excluded entirely. They influence the results in a very certain way - there is a “system” in the errors. This means that one cannot reduce their influence on the result even by frequently repeating the measurement. However, if the extent of the systematic error (e.g. residual deflection of an ohmmeter, the temperature course of an amplifier or the calibration error of a pressure sensor) is known, one can account for it in the result.

4.2 Random Errors

Random errors affect the result of a measurement in an unpredictable and uncontrollable (which means a purely random) way. Causes of random errors, which may occur during a lab course may, for example, be the following:

- (1) The randomness of a natural process such as, e.g., radioactive decay or emission of photons from a light source. For example, it may result in an accidental fluctuation of the number of detected events during a measuring time t .
- (2) The stop watch is pressed too early or too late according to the individual reaction time.
- (3) The measuring rod or the caliper gauge from which sometimes a value too small or too large is read off.
- (4) The electronic noise of a measuring amplifier, which causes fluctuations in the output voltage.

Random errors always lead to the fact that the measured result deviates from the "true" value in one direction or the other (cf. Chap. 5 for the term "true" value). If the measurement is repeated several times, the deviations in both directions balance each other. If this was not true, the observed errors would not be purely random.

From the statements above the following conclusions may be drawn: if there are no experiences with a certain measuring method, a single measured value does not, in principle, provide any information. The measured value may randomly deviate more or less from the "true" value. Only with frequent repetition of the measurement or previous experience with the measuring method one can get an idea of the value about which the individual measured values vary and it becomes possible to evaluate the validity of such a measured value. In the following chapters, these relationships are described quantitatively by means of formulas.

5 The Frequency Distribution of Measured Values

Let us assume that a measurand, e.g. the time t which a body takes to get from A to B, was measured N times ⁴. Thus, N measurement values were obtained which deviate from each other governed by the laws of randomness. The question is: Which of these values is closest to the "true" value?

In order to solve that question, one must examine whether certain values occur more frequently than others, and if so, which ones. Because one may rightly expect that the *most frequent*, i.e. *most probable*, values are closest to the "true" value. For this purpose, the N measurement values, ranging between t_{\min} and t_{\max} , are grouped in j classes with a class range Δt . The range of measured values

$$(2) \quad t_{\min} + (i - 1) \Delta t \leq t < t_{\min} + i \Delta t$$

is assigned to the class i ($i = 1, 2, \dots, j$)⁵. Each class i is assigned a time t_i , which corresponds to the centre of the respective time interval. Now, for each class i , the number of values per Δt , $n(t_i)$ belonging to this class is plotted over the corresponding time as a *bar*. By this method, one obtains a *bar diagram* from which may be gathered how often the measured values have occurred (Fig. 1).

The envelope curve of this diagram, $n(t_i)$, is called *distribution graph* of measured values. According to its definition the area below the distribution graph always equals the total number of measured values, N .

Question 1:

- What is the (very simple!) relationship between the rates of a measured value Δt , $n(t_i)$, and the number of measured values within a class, $m(t_i)$? What is the unit of the quantity $n(t_i)$ in the chosen example?
- What is the equation for calculating N from the distribution graph $n(t_i)$?

⁴ The following considerations apply to every physical measurand. The quantity t (time) only serves as an example here.

⁵ An example for this: Let the measured values range between $t_{\min} = 20.4$ s and $t_{\max} = 22.3$ s, the precision of the watch reading be 0.1 s. The measured values are hence divided into $j = (22.3 - 20.4)/0.1 = 19$ classes, each of which represents a time interval of $\Delta t = 0.1$ s width

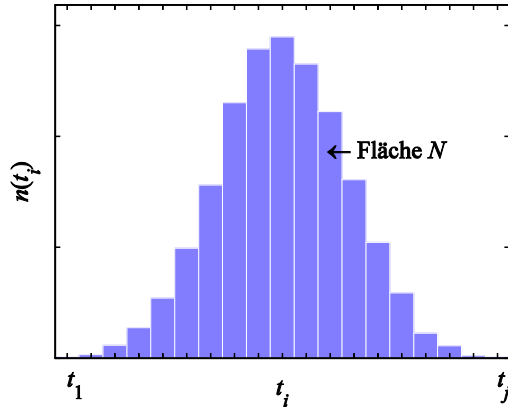


Fig. 1: Distribution graph of measured values. The width of a time interval (bar) is Δt . (German “Fläche” = area).

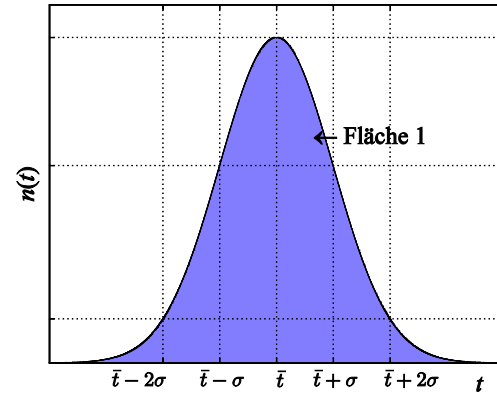


Fig. 2: Distribution graph (*density function*) of the GAUSS distribution or normal distribution (GAUSS curve).

Past experience has shown, and the theory established by CARL FRIEDRICH GAUß (Fig. 3) substantiates that for $N \rightarrow \infty$ and $\Delta t \rightarrow 0$ (and thus $t_i \rightarrow t$) the distribution graph $n(t)$ for values that have been *obtained independently of each other* and are affected by random errors has a quite characteristic form: The form of a *Gaussian bell-shaped curve* or, briefly, *Gaussian* (Fig. 2). In those cases it is common practice to say that the measured values are *Gaussian distributed* or *normally distributed*.



Fig. 3: CARL FRIEDRICH GAUß (1777 - 1855) ⁶

The area under the Gaussian again equals the total number of measured values, N . In general it is normalized to 1. As explained below, this means that the probability of finding a value in the entire range of values between $-\infty$ and $+\infty$ equals 1 ⁷.

The course of the Gaussian normalized to the area 1 is given by:

$$(3) \quad n(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\bar{t})^2}{2\sigma^2}} \quad \text{with} \quad \int_{-\infty}^{\infty} n(t)dt = 1$$

where \bar{t} is the *mean* and σ the *standard deviation* of the Gaussian, the square of the standard deviation, σ^2 , is called *variance*. At $t = \bar{t} \pm \sigma$ the Gaussian has its inflection point. The quantities \bar{t} and σ are of great practical importance:

⁶ Image source: GELLERT, W. et al. [Eds.]: „Kleine Enzyklopädie Mathematik“, VEB Bibliographisches Institut, Leipzig, 1969.

⁷ In the observed example, the time t is the physical quantity, its real value range can only be within the interval $0 \leq t < \infty$. Hence, it appears to be formally wrong or at least nonsensical to extend the range of values to $-\infty$. However, in practice the part of the integral of Eq. (3) is so small within the range $-\infty \leq t < 0$ (≈ 0) that it can be neglected. This is why the limits of the range of values are determined to be $\pm \infty$ for reasons of mathematical simplification.

- The *mean* \bar{t} is the value where $n(t)$ has a maximum. This is the most frequently occurring value in a series of infinitely many measurements. Thus, it represents the *most probable result* of the measurement. In other words: a series of measurements never yields a *true* value, but always a *most probable* value.
- The *standard deviation* σ is a measure of the *scattering* of measured values around the mean \bar{t} . The greater the scattering, the larger σ is, the larger the width of the distribution graph becomes (while the area remains unchanged), and the less single values differ from the other values.

Question 2:

- Calculate and sketch $n(t)$ with the aid of Matlab according to Eq. (3) in the time interval $121.5 \text{ s} \leq t \leq 123.5 \text{ s}$ for $\bar{t} = 122.5 \text{ s}$ as well as a) $\sigma = 0.1 \text{ s}$ and b) $\sigma = 0.2 \text{ s}$. Show both curves in one diagram (Matlab -Command `hold on`). Eq. (3) is written in the following form in Matlab:

$$n = (1/(\sigma*\sqrt{2*\pi}))*\exp(-((t - t_quer).^2)/(2*\sigma^2))$$

Let us pretend that we have performed our experiment such that the conditions $N \rightarrow \infty$ and $\Delta t \rightarrow 0$ are almost fulfilled - so that the distribution graph of our measured values is approximately given by a Gaussian according to Eq. (3). Then through integration of $n(t)$ one can *calculate* (hence, one does not need to count!), how many values are found in the time interval $[\bar{t} - \sigma, \bar{t} + \sigma]$, i.e. within the range $\bar{t} \pm \sigma$.

We know that all N values must be in the time interval $[-\infty, +\infty]$. Due to normalization of the area under the Gaussian to 1 (cf. Eq. (3)) this means:

$$(4) \quad \int_{-\infty}^{+\infty} n(t) dt = 1 \triangleq N \triangleq 100\% \text{ of all measured values.}$$

For the interval $[\bar{t} - \sigma, \bar{t} + \sigma]$ we obtain:

$$(5) \quad \int_{\bar{t}-\sigma}^{\bar{t}+\sigma} n(t) dt = \int_{\bar{t}-\sigma}^{\bar{t}+\sigma} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\bar{t})^2}{2\sigma^2}} dt \approx 0,683 \triangleq 0,683 N \triangleq 68,3\% \text{ of all measured values.}$$

➤ Those who want to repeat that calculation should be warned: The integral over the Gaussian according to Eq. (3) can only be solved numerically and not analytically! It is given in Table 2 (Chap. 11.5).

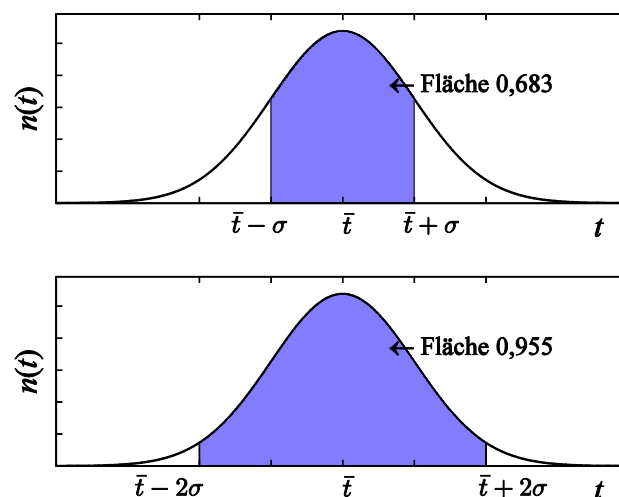


Fig. 4: Partial areas below a Gaussian with a total area normalized to 1. *Top:* Partial area in the range $\bar{t} \pm \sigma$. *Bottom:* Partial area in the range $\bar{t} \pm 2\sigma$.

If the distribution graph of the measured values is represented by a Gaussian (a fact from which we mostly start in practice), *always* about 68.3% of all values are found within the interval $\bar{t} \pm \sigma$ (Fig. 4). For the

range $\bar{t} \pm 2\sigma$, we *always* obtain a percentage of about 95.5 % (cf. Fig. 4) and for the range $\bar{t} \pm 3\sigma$ *always* a percentage of about 99.7 %. In the laboratory jargon we usually say: 68 % of all measured values are found within the 1σ range around the mean, 95 % in the 2σ range, and 99 % in the 3σ range around the mean.

Of course, the conditions $N \rightarrow \infty$ and $\Delta t \rightarrow 0$ cannot be fulfilled in practice. Thereby the interval is increased, in which we find, e.g., 68.3 % of all measured values. In this case, $\bar{t} \pm \sigma$ has to be replaced by $\bar{t} \pm p\sigma$ in which the value of the quantity $p \geq 1$ depends on N and can be calculated by statistical methods (e.g. $p = 1.32$ for $N = 3$, $p = 1.15$ for $N = 5$, $p = 1.06$ for $N = 10$ and $p \rightarrow 1$ for $N \rightarrow \infty$). However, for the evaluation of measurements within the laboratory course we won't take this into account.

6 Mean and Standard Deviation

In the preceding chapter it was detailed that, respecting the assumptions one can make the following statement about a single measurement (one measured value) from a series of measurements:

➤ *The result of a single measurement is within the range $\bar{t} \pm \sigma$ with a probability of 68 %.*

For practical work the question arises of how to determine \bar{t} and σ . Since neither the condition $N \rightarrow \infty$ nor the condition $\Delta t \rightarrow 0$ can be met in a real experiment, one needs to find out how one can calculate the best assessed value, briefly: calculate the *optimum* for \bar{t} and σ from a finite number of values (a so-called *sample*). We will dispense of the theoretical deduction for calculating the optimum values and state the results.

6.1 Mean Value

If a measurand, e.g. the time t , is measured N times, the *optimum for the mean* \bar{t} , resulting from the case $N \rightarrow \infty$, is the arithmetic average of the measured values t_i :

$$(6) \quad \bar{t} = \frac{1}{N} \sum_{i=1}^N t_i$$

6.2 Standard Deviation for a Single Measurement

The *optimum for the standard deviation* σ of the single measurement is:

$$(7) \quad \sigma = \sqrt{\left(\frac{1}{N-1} \sum_{i=1}^N (t_i - \bar{t})^2 \right)}$$

This can be explained: The standard deviation for a single measurement is a measure for the deviation of measured values t_i around the mean \bar{t} . The deviation⁸ of a single measured value t_i from \bar{t} is given by the difference $(t_i - \bar{t})$ (cf. Fig. 5). If the arithmetic average of these differences was taken as a measure of deviation, we would obtain zero as a direct consequence of the definition of the average, because positive and negative differences would cancel each other out. The information about the existing deviation of measured values would thus get lost. In order to prevent such a loss of information, the differences are squared:

$$(t_i - \bar{t})^2$$

This turns all quantities into positive numbers. Then, the arithmetic average of these squares and finally the root of this value is calculated.

⁸ According to /2/ this value is called *uncertainty of measurement*.

The fact that the values are not divided by N but by $N - 1$ for calculating the arithmetic average can be justified by a detailed statistical analysis, which particularly goes into differences between random sample and parent population. However, we will not go into further detail here, since the deviation between $1/N$ and $1/(N-1)$ is infinitesimal for high N .

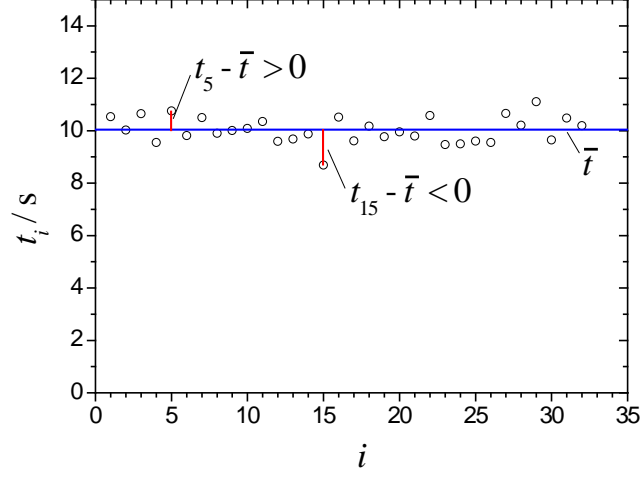


Fig. 5: Illustration of the standard deviation: 32 measured values t_i of time t are plotted over the number of measurement i . \bar{t} is the average of t_i . For $i = 5$ and $i = 15$, the deviations have been drawn exemplarily into the diagram between t_i and \bar{t} .

The standard deviation σ of the single measurement is often termed *error (uncertainty) of a single measurement*, or, in accordance with Eq. (7), *root-mean-square error (rms error)*.

6.3 Standard Deviation of the Mean

In practice the standard deviation σ of the *single measurement* is oftentimes not the fundamental quantity. It is not very interesting to know with which probability a single measured value lies within the range $\bar{t} \pm \sigma$. Much more important is the question of how reliable and reproducible is the *mean* \bar{t} , which was found with a *series of measurements* and which represents the result of a measurement. In other words: with which probability would the *measured result* of another *measurement series*, i.e., a second mean, be found in a given interval around the first one? In order to answer that question, a statement about the *standard deviation of the mean* $\sigma_{\bar{t}}$ is required analogous to the standard deviation of the single measurement.

Assuming we repeated a *series of measurements* consisting of N single measured values M times, so that M means \bar{t}_i ($i = 1, 2, \dots, M$) are obtained. It can be shown that for $M \rightarrow \infty$ the distribution graph of these means of the measurement series would again yield a Gaussian curve with the standard deviation $\sigma_{\bar{t}}$.

In practice one does not wish to repeat the measurement series M times in order to determine the standard deviation of the mean, $\sigma_{\bar{t}}$. Instead, the aim is to determine the optimum for $\sigma_{\bar{t}}$ from *one series of measurements* with N measured values). That optimum results in:

$$(8) \quad \sigma_{\bar{t}} = \frac{\sigma}{\sqrt{N}} = \sqrt{\left(\frac{1}{N(N-1)} \sum_{i=1}^N (t_i - \bar{t})^2 \right)}$$

Now the following probabilistic statement about the mean \bar{t} of the measurement series representing the *result of the measurement* can be made.

- *The result of measurement of another measurement series is in the range $\bar{t} \pm \sigma_{\bar{t}}$ with a probability of about 68 %.*

Further, it holds that:

- The standard deviation of the mean $\sigma_{\bar{t}}$ is the *uncertainty of the measurement* as mentioned in Chap. 1, which, is stated together with the *result of the measurement* (the mean) as the *full result of a measurement* of a *series of measurements* for determining a measurand. It is often termed “*error of the mean*“.

From Eq. (7) it can be seen that the standard deviation σ of a *single measurement* remains almost constant with an increasing number N of measured values in a measurement series, as illustrated in Fig. 5. Addition of further measured values to the diagram does not affect the deviation of the measured values from the mean.

Eq. (8), however, demonstrates that the standard deviation $\sigma_{\bar{t}}$ of the *mean*, i.e., the *uncertainty of the measurement* decreases with increasing N : The uncertainty decreases by $1/\sqrt{N}$. In principle, one can make it arbitrarily small if only enough measurements are performed. In practice, however, one will repeat the measurements only until the uncertainty of the measurement meets a given precision requirements. At this the condition $N \geq 4$ must always be met because otherwise no standard deviation (according to Eq.(7)) can be indicated. This follows from statistical considerations which cannot be discussed in detail here.

In summary, it can be stated:

- The result of a measurement series must always be given in the form $\bar{t} \pm \sigma_{\bar{t}}$.
- The uncertainty of a measurement $\sigma_{\bar{t}}$ (the error of the mean) decreases by a factor of $1/\sqrt{N}$ for an increasing number of measured values N .
- Unless stated otherwise, the result of a measurement having the form $t = (100.6 \pm 1.2)$ s is always interpreted as follows: Result of the measurement 100.6 s, uncertainty of the measurement (standard deviation of the mean) 1.2 s.

6.4 Absolute and Relative Errors

The quantity $\sigma_{\bar{t}}$ represents the *absolute error* of the measured result, the quantity $\sigma_{\bar{t}} / \bar{t}$ the *relative error*, which generally is given as a relative value. During the lab course we will mainly confine ourselves to stating the absolute errors.

7 Maximum Error of a Single Measurement

It often happens, in the lab course very frequently, that the value of a measurand a is not determined by a series of measurements but instead only by a single measurement, as e.g. for measurements of lengths. In that case the *result of a single measurement* is stated in the laboratory course instead of the mean and the *maximum error* Δa instead of the standard deviation of the mean. This is the biggest possible error occurring in the single measurement of the quantity. It has to be assessed according to reasonable projections. For example, when the length of a distance is measured with a ruler, then the reading precision of the ruler will be taken as the maximum error. E.g. using millimetre measuring tape, the value 0.5 mm is used, whereas for a vernier caliper a value of 0.1 mm or 0.05 mm is correct, and a value of 0.01 mm is used for a micrometer gauge.

8 Accuracy

The *accuracy* of a measured result for the measurand a , i.e. the number of *significant digits*, is limited by the uncertainty of the measurement, i.e. the standard deviation $\sigma_{\bar{a}}$ of the mean or the maximum error Δa of a single measurement, respectively.

$\sigma_{\bar{a}}$ and Δa are to be rounded to maximally *two significant digits* in this laboratory course!⁹

⁹ This means that the uncertainty can be stated with an accuracy of approx. 1 %. A better accuracy cannot be achieved with the equipment in the lab!

The mean, or, the single measured value resp. is then to be rounded so that its last significant digit has the same order of magnitude as the last significant digit of $\sigma_{\bar{a}}$ or, respectively, of Δa .

Some examples: An accuracy feigned by calculations, or the number of digits of an electronic clock of the form: $t = 90.4671$ s is simply incorrect, if the maximum error of the time measurement is 1.1 s for example. In this case it should read (rounded): $t = (90.5 \pm 1.1)$ s. It is also incorrect to state $R = (83.62 \pm 2.624) \Omega$; it should read $R = (83.6 \pm 2.6) \Omega$ because of the limitation to two significant digits for the uncertainty of the measurement.

The *significance* of a digit is independent of its *order of magnitude* (position with respect to the decimal point). Thus, all of the following numbers contain two significant digits:

18 1.8 0.18 0.018 0.0018 etc.

This becomes quite obvious when, as recommended, using powers of ten, i.e., writing the above numbers as follows:

1.8×10^1 1.8×10^0 1.8×10^{-1} 1.8×10^{-2} 1.8×10^{-3} .

In rounding numbers, the question arises in which direction the digit 5 should be rounded. For example, take the number 4.135, to be rounded to two digits after the decimal point. It would be possible to round it to 3.14 to 3.13. The rule to apply is to round in such a way, that the last digit of the resulting number is *even*. In the example, 4.135 would be *rounded upwards* to give 4.14. The number 4.125 would in contrast be *rounded downwards* to 4.12. The reasoning behind this rule is the fact, that a division by 2 results in both numbers (rounded and non-rounded) having the same rounded result. For the examples above, we have:

$$4.135 : 2 = 2.0675 \approx 2.07 \quad \text{and likewise} \quad 4.14 : 2 = 2.07$$

$$4.125 : 2 = 2.0625 \approx 2.06 \quad \text{and likewise} \quad 4.12 : 2 = 2.06$$

9 Error Propagation, Combined Measurands

It happens frequently that the measured quantity itself (direct measured result) is of no interest, but a quantity calculated from it (indirect measured result, cf. Chap. 2). Again, let us take as an example the acceleration of gravity g (Eq. (1) which depends on the measured quantities s and t :

$$g = \frac{2s}{t^2}$$

Further examples are the density ρ of a physical body which is a function of the measurands mass m and volume V :

$$(9) \quad \rho = \frac{m}{V}$$

or the capacity C of a plate capacitor in vacuum which depends on the measurands surface A of the plates and the distance d between them. With the electrical field constant ϵ_0 we have:

$$(10) \quad C = \epsilon_0 \frac{A}{d}$$

All these examples clearly show that the error of the quantity of interest has to be calculated from different measured quantities. The following chapters describe how this is done.

9.1 The Most Probable Error of a Combined Measured Quantity

In case the result of a measurement for a measurand y is calculated from the results of measurement for several Gaussian distributed measurands, for which means and standard deviations have been gained from *measurement series*, the *most probable error* for y will be determined by the *Gaussian error propagation law*, which is defined in the following.

Let us assume that the quantity of interest y depends on the measurands a, b, c , etc.:

$$(11) \quad y = f(a, b, c, \dots)$$

We assume that the measured values for the measurands a, b, c, \dots are Gaussian distributed and do not influence each other, i.e., they are independent from each other in a statistical sense. We further assume that the mean values $\bar{a}, \bar{b}, \bar{c}, \dots$ and the standard deviations of the means $\sigma_{\bar{a}}, \sigma_{\bar{b}}, \sigma_{\bar{c}}, \dots$ are known. Then the *optimum* y_B ¹⁰ for the measurand of interest y is the value obtained by calculating y from the means $\bar{a}, \bar{b}, \bar{c}, \dots$:

$$(12) \quad y_B = f(\bar{a}, \bar{b}, \bar{c}, \dots)$$

This is plausible. The *standard deviation* σ_{y_B} of y_B is given by the *Gaussian error propagation law* (illustrated in Chap. 9.2) which reads:

$$(13) \quad \sigma_{y_B} = \sqrt{\left[\left(\frac{\partial y}{\partial a} \right)_B \sigma_{\bar{a}} \right]^2 + \left[\left(\frac{\partial y}{\partial b} \right)_B \sigma_{\bar{b}} \right]^2 + \left[\left(\frac{\partial y}{\partial c} \right)_B \sigma_{\bar{c}} \right]^2 + \dots} := \sqrt{\Delta y_a^2 + \Delta y_b^2 + \Delta y_c^2 + \dots}$$

The expressions $\partial y / \partial a$, $\partial y / \partial b$, etc. in Eq. (13) are the "partial derivatives" of y with respect to the quantities a, b, c, \dots . They state how y would change, if one changes *only a*, *only b*, or *only c* etc. keeping the other quantities constant. (Mathematically: One forms the derivative of y with respect to *each* of the quantities a, b, c, \dots and considers the other quantities to be constant). The index B in the partial derivatives states that the numerical values of the partial derivatives are to be calculated at the best values (means) $\bar{a}, \bar{b}, \bar{c}, \dots$ of the measurands a, b, c, \dots .

As an example for the calculation of partial derivatives we take again Eq. (1) for the acceleration of gravity g , which depends on the quantities s and t . The partial derivative of g by s is:

$$\frac{\partial g}{\partial s} = \frac{2}{t^2}$$

and the partial derivative of g by t is:

$$\frac{\partial g}{\partial t} = -\frac{4s}{t^3}$$

9.2 Illustration of the Error Propagation Law

To illustrate the error propagation law we examine once more the acceleration of gravity g given by Eq.(1). We therefore have a function, which depends on the two variables s and t . Eq. (11) then reads with $y := g$, $a := s$ and $b := t$:

$$(14) \quad g = f(s, t) = \frac{2s}{t^2}$$

¹⁰ The index B stands for best value.

In Fig. 6, g is shown as a function of s and t in a 3D-plot, in which the linear dependence of the acceleration of gravity g of s and the reciprocal-square dependence of t is made clear.

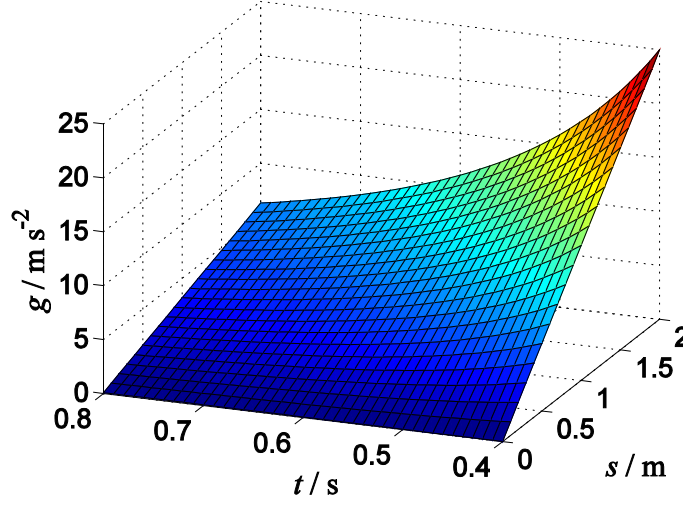


Fig. 6: Illustration of the error propagation law.

Looking at Fig. 6 we examine the different terms in Eq. (13) more closely, and - as an example - pick the second one: The quantity to be determined, y (g in this case), depends on the measurand b (t in this case) among others. When b changes, y will change as well. The partial derivative $\partial y / \partial b$ states how big this change is, i.e., how steep the *slope* of the function $y = f(a, b, c, \dots)$ as a function of b is, if one assumes the remaining quantities a, c, \dots to be constant. In the considered example we have:

$$(15) \quad \frac{\partial y}{\partial b} := \frac{\partial g}{\partial t} = -\frac{4s}{t^3}$$

Since this slope is not equally steep everywhere (in the example it changes with t^{-3} , see Eq.(15)), it is useful to calculate it at the position $\bar{a}, \bar{b}, \bar{c}, \dots$ (here \bar{s}, \bar{t}) which is also of relevance for calculating the best value y_B , (here g_B). Therefore the index B stands in Eq. (13): $(\partial y / \partial b)_B$.

Now one needs to know for the term picked out how big the change of y_B is, which is caused by the error $\sigma_{\bar{b}}$. From basic differential calculus it is known that this change is given by the *differential*:

$$(16) \quad \Delta y_b := \left(\frac{\partial y}{\partial b} \right)_B \sigma_{\bar{b}} \quad \text{here:} \quad \Delta g_t = \left(\frac{\partial g}{\partial t} \right)_B \sigma_{\bar{t}} = -\frac{4\bar{s}}{\bar{t}^3} \sigma_{\bar{t}}$$

In the same way we can determine the errors

$$(17) \quad \Delta y_a := \left(\frac{\partial y}{\partial a} \right)_B \sigma_{\bar{a}} \quad \text{here:} \quad \Delta g_s = \left(\frac{\partial g}{\partial s} \right)_B \sigma_{\bar{s}} = \frac{2}{\bar{t}^2} \sigma_{\bar{s}}$$

$$(18) \quad \Delta y_c := \left(\frac{\partial y}{\partial c} \right)_B \sigma_{\bar{c}} \quad \text{(here without importance)}$$

etc., which all add to the total error, i.e. to the standard deviation σ_{y_B} of y_B . Thus it is obvious that they have to be added in order to calculate σ_{y_B} . Since the single errors according to Eqs. (16) and (18) may be positive or negative, they may partly or entirely neutralize each other and may therefore suggest overall too small of an error. To avoid this it is useful to square the single errors (which makes them all positive) and to take the square root of the sum of squares afterwards. Due to this *geometrical (quadratic)* addition of

single errors, the total error will be smaller than the sum of the single errors. This takes into consideration that the single errors of the independent quantities a, b, c etc. may not be equally represented in the final result but may at least partly compensate one another. Hence one speaks of the *most probable error*.

9.3 Maximum Error of a Combined Measured Quantity

Let us now consider the case (as an example), that the quantities a, b, c etc. either contain no random errors or their errors have partially not been obtained from measurement series. The latter case occurs frequently in practice (even in the lab course), namely when at least some of the results of measurement for the measurands a, b, c etc. have been obtained from single measurements for which only the respective maximum errors $\Delta a, \Delta b, \Delta c$ are known.

In such a case the *maximum error* Δy is stated instead of the *standard deviation* according to Eq. (13) for the combined measurand y . It results from the most unfavourable, i.e. *arithmetic (linear)* addition of all single errors and is given by:

$$(19) \quad \Delta y_B = \left| \frac{\partial y}{\partial a} \right|_B \Delta a + \left| \frac{\partial y}{\partial b} \right|_B \Delta b + \left| \frac{\partial y}{\partial c} \right|_B \Delta c + \dots := |\Delta y_a| + |\Delta y_b| + |\Delta y_c| + \dots$$

where the maximum errors, or respectively, standard deviations are to be inserted for the quantities $\Delta a, \Delta b, \Delta c, \dots$

Except for taking the absolute values of the differentials, Eq. (19) is the *total differential* of y_B .

➤ *If not explicitly required otherwise, the maximum error is always to be stated for combined measured quantities in the laboratory course.*

10 A Concrete Example

The gravitational acceleration g can be determined using a “mathematical pendulum“. The mathematical pendulum which can only be realized approximately in practice, consists of a point-like mass in the ideal case, which is suspended on a massless thread such that it can swing without externally disturbing influences (especially no friction). For displacements of the pendulum by a small angle α below about 5° it holds that $\alpha \approx \sin \alpha \approx \tan \alpha$ (α in radian) and the following relationship between the period of oscillation T of the pendulum, the length of the thread l , and the gravitational acceleration g is in good approximation given by:

$$(20) \quad T = 2\pi \sqrt{\frac{l}{g}} \quad \text{and} \quad g = \frac{4\pi^2 l}{T^2}, \text{ respectively.}$$

By measuring l and T it is thus possible to determine g . Already prior to the measurement we can state the possible systematic errors:

- Contrary to theory the mass is not pointlike and the thread is not massless. How this affects the measurement is difficult to state. One tries to approach the mathematical pendulum as far as possible by using a very thin thread and a mass with little spatial expansion and expects that the remaining errors will be so small compared to the measurement uncertainties of the measurands l and T , that they can be neglected.
- The pendulum cannot be suspended so that it is completely frictionless. Hence, sufficient effort needs to be made towards achieving an optimal suspension so that the error caused by friction is small compared to the measurement uncertainties.

When preparing the experiment care must be taken, so that the watch for measuring T as well as the ruler for measuring l are calibrated in order to exclude systematic errors caused by inadequate instruments. Additionally, the length of the thread l must be chosen such that the measurement can be carried out using pendulum swings below approx. 5° , since Eq. (20) is only valid in good approximation under that condition.

After completing these preparations, the measurement can be performed. It is known that the influence of random errors on the uncertainty of measurement can be minimized by repeating the measurement as often as possible. At the same time one recognizes that repeated measurements of the length l are of no use at all. If the ruler is placed and read carefully, the measured value will not change when the measurement is repeated several times. Nevertheless, the measured value for l is, of course, also subject to an error: On the one hand, the ruler only has a certain accuracy even when following calibration, on the other hand, it can only be placed and read with finite accuracy. The result of the measurement can then be stated as follows:

$$(21) \quad l = L \pm \Delta L ; \quad \text{e.g.} \quad l = (2,5580 \pm 0,0020) \text{ m}$$

L being the value read and ΔL its maximum error.

The oscillation period T is determined using a stop watch, the internal inaccuracy of which is assumed to be negligible. Starting and stopping the watch depends on the reaction time of the user and are thus subject to random fluctuations. Its influence on the uncertainty of the measured result can be minimized by repeated measurements. After a total of N measurements yielding the values T_i , the result of the time measurement is:

$$(22) \quad T = \bar{T} \pm \sigma_{\bar{T}} ; \quad \text{e.g.} \quad T = (3.210 \pm 0.010) \text{ s}$$

with \bar{T} being the arithmetic mean of measured values T_i according to Eq. (6) and hence the best value for T , and $\sigma_{\bar{T}}$ being the standard deviation of the mean \bar{T} according to Eq. (8).

Hence, the best value g_B for g is according to eq. (20):

$$(23) \quad g_B = \frac{4\pi^2 L}{\bar{T}^2} ; \quad \text{in the example} \quad g_B = \frac{4\pi^2 \times 2,5580 \text{ m}}{(3,210 \text{ s})^2} = 9,801 \frac{\text{m}}{\text{s}^2}$$

Because L has not been determined from a measurement series, not the standard deviation is calculated using Eq. (13), but the maximum error Δg_B according to Eq.(19). This results in:

$$(24) \quad \Delta g_B = \left| \frac{\partial g}{\partial l} \right|_B \Delta L + \left| \frac{\partial g}{\partial T} \right|_B \sigma_{\bar{T}}$$

First, the absolute values of the partial derivatives at the positions of the best values B (in this case for the values L and \bar{T}) are calculated:

$$(25) \quad \left| \frac{\partial g}{\partial l} \right|_{L,\bar{T}} = \left| \frac{4\pi^2}{T^2} \right|_{L,\bar{T}} = \frac{4\pi^2}{\bar{T}^2} = \frac{4\pi^2}{(3,210 \text{ s})^2} = \dots$$

$$\left| \frac{\partial g}{\partial T} \right|_{L,\bar{T}} = \left| -\frac{8\pi^2 l}{T^3} \right|_{L,\bar{T}} = \frac{8\pi^2 L}{\bar{T}^3} = \frac{8\pi^2 \times 2,5580 \text{ m}}{(3,210 \text{ s})^3} = \dots$$

Substituting L and \bar{T} , equation (25) yields two numbers which are multiplied with two other numbers according to Eq. (24) (namely ΔL and $\sigma_{\bar{T}}$), and must be added to find the value of interest, Δg_B :

$$\begin{aligned}
 \Delta g_B &= \frac{4\pi^2}{\bar{T}^2} \Delta L + \frac{8\pi^2 L}{\bar{T}^3} \sigma_{\bar{T}} \\
 (26) \quad &= \frac{4\pi^2}{(3,210 \text{ s})^2} 0,0020 \text{ m} + \frac{8\pi^2 \times 2,5580 \text{ m}}{(3,210 \text{ s})^3} 0,010 \text{ s} = 0,069 \frac{\text{m}}{\text{s}^2}
 \end{aligned}$$

In the statement of the numerical value the rounding to two significant digits must be adhered to. Summarizing, the *full result of the measurement* reads:

$$(27) \quad g = g_B \pm \Delta g_B = (9.801 \pm 0.069) \frac{\text{m}}{\text{s}^2}$$

Since a value for g is available from the literature in this example that value must be compared with the result of the measurement. If the value for g in the literature is within the range $g_B \pm \Delta g_B$ the experiment is completed and concluded by the statement that a “good agreement within the frame of the measuring accuracy“ was achieved. If the value from the literature is not within the range $g_B \pm \Delta g_B$ however, there is a relatively high probability that the measurement has been falsified by a systematic error.

Instead of the absolute error Δg_B of the measured result g_B one can also state the relative error ε_g for g_B :

$$(28) \quad \varepsilon_g = \frac{\Delta g_B}{g_B}$$

thus:

$$(29) \quad \varepsilon_g = \frac{\Delta L}{L} + 2 \frac{\sigma_{\bar{T}}}{\bar{T}}$$

From this equation it can be seen that the relative error of \bar{T} , $\sigma_{\bar{T}}/\bar{T}$ is embodied doubly in the result, however, the relative error of L , $\Delta L/L$, only simply. If this is to be compensated, the relative error of \bar{T} may only be half as big as the relative error of L . This can always be achieved by a sufficient number of measurements of the oscillation period (cf. Eq. (8)) and should already be considered when planning the experiment.

11 Appendix

11.1 Linear Regression, Regression Lines

11.1.1 Regression Lines of the Form $y = ax + b$

In practice it happens quite often that two quantities x and y depend linearly on each other, which means that they are linked by the equation of a line: $y = ax + b$. The aim of the measurement is then to find the quantities a and b . Let us take as an example the time behaviour of a velocity v of a constantly accelerated motion: $v(t) = at + v_0$ with a : acceleration, t : time, and v_0 initial velocity. We measure $v(t)$ (dependent variable) with certain default values of t (independent variable), to obtain a value for the acceleration a and the initial velocity v_0 .

Plotting $v(t)$ against t according to Fig. 7, we expect that the sketched points are on a line with a slope that represents the value of a and a v -interception that gives us v_0 . Trying to plot this line in the diagram of measured values, we notice that quite a few slopes and values for v at $t=0$ match our measured values more or less. Which parameters are the right ones? This question can only be answered in the sense of a probability statement. We want to give the answer in the following.

We return to our function $y = ax + b$. As frequently occurs in practice, we have N values of the quantity x , from which we determine N measured values for the quantity y . The errors of the specified values of x are

negligible; the errors of the measured values of y shall be randomly distributed. We maintain, that the *optimums* of A and B for the parameters a and b of the line equation have been found, when the sum of the squares of the vertical distances between the measured values and the „regression line” determined by A and B , are minimal:

$$(30) \quad \sum_{i=1}^N [y_i - (Ax_i + B)]^2 \rightarrow \text{Minimum}$$

Question 3:

- How can this approach be justified?

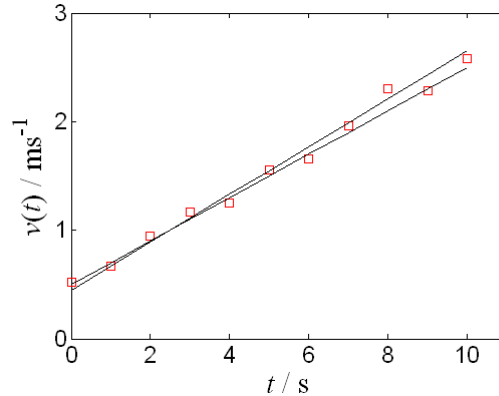


Fig. 7: Which is the best regression line through the red measured values?

By means of differential calculus one may quite easily determine a solution for the requirements described in Eq. (30). It is found for A and B (summation from 1 to N each):

$$(31) \quad A = \frac{N(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{N(\sum x_i^2) - (\sum x_i)^2}$$

$$(32) \quad B = \frac{(\sum y_i)(\sum x_i^2) - (\sum x_i y_i)(\sum x_i)}{N(\sum x_i^2) - (\sum x_i)^2}$$

Of course, these optimums are also subject to errors, which we are now looking for. The faulty quantities in Eqs. (31) and (32) are the y_i . For the variance of the y_i we obtain the optimum (cf. Eq. (7)):

$$(33) \quad \sigma_y^2 = \frac{1}{N-2} \sum (Ax_i + B - y_i)^2$$

Division by $(N - 2)$ instead of $(N - 1)$ is explained by the fact that the optimums A and B enter the calculation of the quantity σ_y^2 . Applying the error propagation to Eqs. (31) and (32) and inserting Eq. (33) for σ_y , we find as *optimums for the standard deviations* of A and B (D is an auxiliary quantity defined in Eq. (36)):

$$(34) \quad \sigma_A = \sqrt{ND}$$

$$(35) \quad \sigma_B = \sqrt{D \sum x_i^2}$$

with

$$(36) \quad D = \frac{1}{N-2} \frac{\sum (Ax_i + B - y_i)^2}{N(\sum x_i^2) - (\sum x_i)^2}$$

In this laboratory course, the software *Origin* is used for these calculations, as it can produce this data with a few clicks of the mouse (\rightarrow Analysis \rightarrow Fit \rightarrow Linear Fit). Do not calculate the parameters of regression lines „by hand“. That would be much too time-consuming!

11.1.2 Regression Lines of the Form $y = ax + b$ with Predefined b

In practice measured values are also connected with each other by a linear function $y = ax + b$ for which the intercept of the axis is predefined. Let us take OHM's law $U = RI$ as an example: if we measure the voltage U as a function of the current I , then the regression line through the measured values is a straight line through the origin ($b = 0$) with the slope R . The condition for calculating the optimum A of the slope a of the regression line in this case reads

$$(37) \quad \sum_{i=1}^N [y_i - Ax_i]^2 \rightarrow \text{Minimum}$$

analogous to Eq. (30). We obtain for A by means of differential calculus and for σ_A by means of error propagation:

$$(38) \quad A = \frac{\sum x_i y_i}{\sum x_i^2}$$

$$(39) \quad \sigma_A = \sqrt{\frac{\sum (y_i - Ax_i)^2}{(N-2) \sum x_i^2}}$$

In order to perform the corresponding calculations with *Origin*, the box labeled “Fixed intercept at” must be checked in the window *Linear Fit*, and the value of b must be entered.

11.1.3 Regression Lines of the Form $y = ax + b$ with Predefined a

The reverse case, in which the slope a of the regression line is predetermined and we are only looking for the intercept of the axis b occurs occasionally. The condition for calculating the optimum B of b again reads analogous to Eq. (30):

$$(40) \quad \sum_{i=1}^N [y_i - (ax_i + B)]^2 \rightarrow \text{Minimum}$$

only B being a free parameter for determining the minimum in this case. For B and σ_B we obtain in this case:

$$(41) \quad B = \frac{\sum y_i - a \sum x_i}{N}$$

$$(42) \quad \sigma_B = \sqrt{\frac{1}{N(N-2)} \sum (ax_i + B - y_i)^2}$$

In order to perform the corresponding calculations with *Origin*, the field *Fixed Slope* in the window *Linear Fit* must be checked and the numerical value for a must be entered.

11.2 Linearizations

By means of quite elemental mathematical rearrangement we may *linearize* non-linear relations between measured values in order to use linear regression for the calculation of the optimums of the desired values. For example, the power relationship of the form:

$$(43) \quad y = bx^a$$

becomes the linear relationship by a simple logarithm (line equation):

$$(44) \quad \underbrace{\log y}_{\tilde{y}} = \underbrace{\log b}_{\tilde{b}} + a \underbrace{\log x}_{\tilde{x}} \quad \text{hence} \quad \tilde{y} = \tilde{b} + a\tilde{x}$$

For logarithmic relationships we have to consider the following: the logarithm of a physical quantity y , which is given by the product of *numerical value* and *unit*, cannot be calculated directly, because the logarithm of a *unit* does not make sense. This is why we have to divide these values by their unit to make them dimensionless. After that we may do rewritings according to the one from Eq. (43) to Eq. (44). An example: The ohmic resistance R becomes $r = R/\Omega$, the voltage U becomes $u = U/V$, the current I becomes $i = I/A$ and so the OHM's law $R = U/I$ becomes the modified form $r = u/i$, which reads $\log r = \log u - \log i$ in the logarithmic form.

Plotting y against x (Eq. (44)) on a *double-logarithmic* scale (i.e. $\log y$ over $\log x$), we obtain a line with the intercept $\log b$ and the slope a from the best fit line through the measuring points. The optimums for a and $\log b$ are found by means of a linear regression, to which we have to apply to Eq. (44).

Using the logarithm also converts an exponential relationship of the form

$$(45) \quad y = be^{ax}$$

into a linear relationship:

$$(46) \quad \ln y = \ln b + ax \ln e = \ln b + ax$$

Plotting y against x on a *half-logarithmic* scale, we also obtain from the best regression a line with intercept $\ln b$ and slope a , the optimums of which are found by applying the linear regression to Eq. (46).

If *logarithmic papers* are used, we have to consider that this paper is always meant for the use of *decimal logarithm* (\log , basis 10). However Eq. (46) deals with *natural logarithm* (\ln , base e). Hence, when values are to be taken graphically from a diagram on logarithmic paper or calculated values are to be entered, they have to be converted appropriately. (Remember: $\log x = \ln x / \ln 10$; $\ln x = \log x / \log e$).

The software *Origin* is used in the laboratory course for the calculating parameters of regression lines in logarithmic diagrams. In order to do this, it is necessary to check the field **Apparent Fit**¹¹ in the window **Linear Fit**.

Question 4:

- The exponential attenuation law $I(x) = I_0 \exp(-\mu x)$ describes the attenuation of the intensity I of a radiation passing through a layer of matter of thickness x . I_0 is the initial intensity of the radiation at the location $x = 0$ and μ a material dependent coefficient ($[\mu] = 1/m$). Draw $I(x)$ in linear and semilog representation (abscissa x linear each). How do we obtain the attenuation coefficient μ from the semilog diagram ?

¹¹ The fit is called *apparent* since the data in the *Origin* worksheet remain in their original, linear form. The data only *appear* logarithmic in the diagram on which the fit is based, provided the scaling of the appropriate axes is set to the Type "log10", "log2" or "ln".

11.3 Correlation

Sometimes, although rather rarely in an introductory laboratory course, it has to be examined whether a presumed linear connection exists between two quantities x and y , i.e., whether the two values are *correlated*. The diagram of the measured values does not always show whether the plotted values are well fit to a line or not. In any case the question arises of how “well” is “well enough”, in order to maintain the *hypothesis* that x and y are correlated? The quantitative answer of this question is yielded by the calculation of the *correlation coefficient*. It is given by:

$$(47) \quad r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

\bar{x} and \bar{y} being the arithmetic means of the measured values x and y . The correlation coefficient can only range between -1 and +1. For judgement of the question of whether two values are correlated, the magnitude of r is decisive. For $|r| = 1$ the values are correlated perfectly, for $|r| = 0$ they are not correlated. For all values between 0 and 1 probability statements can be made which additionally depend on the amount of N , i.e. the number of measurements. For $N = 10$ and $|r| \geq 0.8$ e.g. the probability P that the values are not correlated is $P = 0.5\%$. Table 1 (Chap. 11.5) in the annex shows the probabilities for further combinations of N and $|r|$.

11.4 Errors of Weighted Means

We assume that a measurand h is measured in M measurements (No. $i = 1, \dots, M$) under *varying conditions*. The result of the single measurements yielded the results h_i and the uncertainties σ_i .

Our aim is to calculate a final result for the quantity of interest h from the M measured values h_i . In the easiest case this would be the arithmetic mean of the h_i , however, this would not consider the fact that the h_i may exhibit quite different uncertainties σ_i because, e.g., the optimum measurement accuracy was not the same in all measurement series.

In such cases a *weighted mean* h_g is calculated instead of the *arithmetic mean of the h_i* . If g_i are the *weights* to be considered for the single values h_i when calculating h_g , we obtain for summation from 1 to M :

$$(48) \quad h_g = \frac{\sum h_i g_i}{\sum g_i}$$

Generally the reciprocal values of the variances are chosen as weights:

$$(49) \quad g_i = \frac{1}{\sigma_i^2}$$

Applying the error propagation law to Eq. (48) we obtain for the uncertainty σ_g of the weighted mean when summing up from 1 to M :

$$(50) \quad \sigma_g = \sqrt{\sum \left(\frac{\partial h_g}{\partial h_i} \sigma_i \right)^2} = \left(\sum \frac{1}{\sigma_i^2} \right)^{-\frac{1}{2}}$$

Question 5:

- How do we obtain that result? What σ_g is found for the special case $g_i = \text{const.} = 1$?

11.5 Tables

Table 1:

Percentage probabilities that two quantities measured N times and having a correlation coefficient of $|r| \geq |r_b|$ are *uncorrelated* (after $1/N$).

$ r_b \rightarrow$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
N ↓											
3	100	94	87	81	74	67	59	51	41	29	0
4	100	90	80	70	60	50	40	30	20	10	0
5	100	87	75	62	50	39	28	19	10	3.7	0
6	100	85	70	56	43	31	21	12	5.6	1.4	0
7	100	83	67	51	37	25	15	8.0	3.1	0.6	0
8	100	81	63	47	33	21	12	5.3	1.7	0.2	0
9	100	80	61	43	29	17	8.8	3.6	1.0	0.1	0
10	100	78	58	40	25	14	6.7	2.4	0.5		0
11	100	77	56	37	22	12	5.1	1.6	0.3		0
12	100	76	53	34	20	9.8	3.9	1.1	0.2		0
13	100	75	51	32	18	8.2	3.0	0.8	0.1		0
14	100	73	49	30	16	6.9	2.3	0.5	0.1		0
15	100	72	47	28	14	5.8	1.8	0.4			0
16	100	71	46	26	12	4.9	1.4	0.3			0
17	100	70	44	24	11	4.1	1.1	0.2			0
18	100	69	43	23	10	3.5	0.8	0.1			0
19	100	68	41	21	9.0	2.9	0.7	0.1			0
20	100	67	40	20	8.1	2.5	0.5	0.1			0
25	100	63	34	15	4.8	1.1	0.2				0
30	100	60	29	11	2.9	0.5	0.1				0
35	100	57	25	8.0	1.7	0.2					0
40	100	54	22	6.0	1.1	0.1					0
45	100	51	19	4.5	0.6						0

$ r_b \rightarrow$	0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	
N ↓											
50	100	73	49	30	16	8.0	3.4	1.3	0.4	0.1	
60	100	70	45	25	13	5.4	2.0	0.6	0.2		
70	100	68	41	22	9.7	3.7	1.2	0.3	0.1		
80	100	66	38	18	7.5	2.5	0.7	0.1			
90	100	64	35	16	5.9	1.7	0.4	0.1			
100	100	62	32	14	4.6	1.2	0.2				

