Matthias Hengsberger
Physik-Institut der Universität Zürich
HS 2020/FS 2021

Data and error calculus

PHY112/122 – Praktikum zur Physik I/II

1 Measurements and uncertainties

A measurement can never be perfectly precise. Whenever a measurement is performed, measuring errors are being made. Error calculus aims to estimate and calculate measuring errors and so-called uncertainties as well as estimating their consequences. This document provides a short outline of the field of error calculus as you will need it in the lab courses, and in a shortened version.

1.1 Measurements

To measure means to match a physical quantity against a normalized unit, like e.g. the length of a rod with the help of a measuring tape. Doing so, one reads off the number which is closest to the end of the rod. The result will be represented in form of a number and a unit, e.g.

$$x = 213 \text{ mm}.$$

In reality, this measurement is only approximate, since the result gives only an estimation of the actual length. The exact, "true" value is unknown. The deviation from the true length of the measured result is denoted as the **measurement error of a single measurement**. In this example, it will probably be of the order of a millimetre. Since the read-off is the only thing we know, neither the true value, nor the error in measurement is known.
Error calculus deals with **estimating** both, the true value of a quantity and the presumable error in the measurement. One could, for example, repeat a measurement multiple times and obtain the mean $\overline{x}$ of the measured values, which is a better approximation for the true value than the single measurement.
The uncertainty in this measurement can be estimated from the distribution of the measured values or the quality and the accuracy of the measurement device.
A correct result *always* has to including the estimated measurement error:

$$x = (213 \pm 1) \text{ mm}.$$

1.2 Systematic and statistical (random) errors

Measurement errors can be categorized into **systematic** and **statistical (or random)** errors. Often, both sources of errors produce both types of uncertainties (like e.g. the background radiation in experiment Ab), and thus, the frontier in between them is not always to be drawn unambiguously.

Systematic errors are caused by the measurement process itself, an inaccurate **calibration** of the measuring device, defective functionality or even by a faulty **read out** of the pointer

instruments. Systematic errors exist, when the measurement (the measuring process or the measuring device) modifies the measured quantity. A practical example is the fact that during the measurement of a temperature of an object, this object is being cooled down due to the contact with the thermometer. These systematic errors are characterized by the fact that they deviate the measured value always in one (predictable) direction. Once determined, such errors can in principle be eliminated or reduced by adapting the experimental procedure and/or appratus. Systematic errors can *not* be minimized by repeated measurements under the same conditions.

Statistical (random) errors Statistical uncertainty results in fluctuations of the measured quantity. After repeated measurements of the same quantity and calculation of the mean value, we say that the measured values scatter around the mean value to higher and lower values with the same probability and with a given spread which we are to quantify. Random erros cannot be avoided but minimized with repeated measurements. A statistical analysis includes a calculation of the scattering of the measured values around the mean (best) value, given the measurements has been undertaken equal circumstances. One distinguishes the following cases:

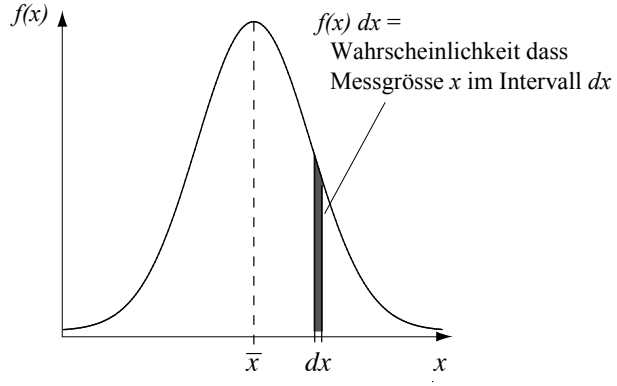
Measurement uncertainty If the measured quantity has a fixed value, the results of repeated mesurements will scatter around a mean value. The mean will generally get closer to the true value, the more measurements are being made. The distribution of the measured values around the mean allows for an estimation of the error.

Distribution of the measured quantity In this case the main interest lies in the distribution of the measured values, which reflects the distribution of the measured quantity for many different samples/objects, like e.g. the body height of many people. The most important attributes of this distribution are its mean value and a criterion for the statistical scatter of the results.

In this lab course we are mostly concerned with the first variant, the measurements uncertainty due to the fluctuations of the measurement results.

1.3 Statistical measurement uncertainty

Consider the measurement of a physical quantity. $\hat{x}$ shall be its true value. Let x_i be a single measurement. We assume that the probability to measure some pre-defined value x increases the more x is close to $\hat{x}$, like in this gaussian distribution:



$w(x)$ be the probability to measure the value x. Often, it is more convenient to use the so-called probability density:

$$f(x) = \frac{dw(x)}{dx} \tag{1}$$

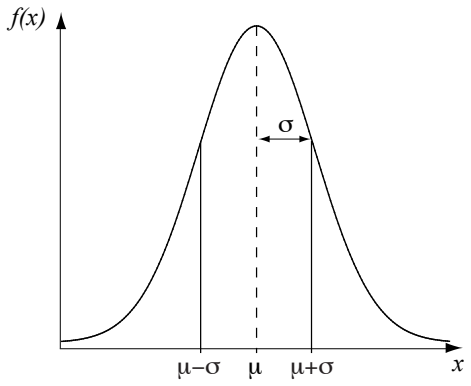
which is defined as probability dw for the measured value x to lie within the interval $[x, x + dx]$: $dw(x) = f(x) \cdot dx$. Since any measurement must yield some result the function $f(x)$ must be normalized:

$$\int_{-\infty}^{+\infty} dw = \int_{-\infty}^{+\infty} f(x) \cdot dx = 1 \tag{2}$$

Key parameters of a probability distribution		
Quantity	Meaning	Definition
1st Statistical moment	**Mean value**	$\mu = \int_{-\infty}^{\infty} f(x) \cdot x \, dx$
2nd Statistical moment	Variance	$V := \sigma^2 = \int f(x) \cdot (x - \hat{x})^2 \cdot dx$
Standard deviation	**Scatter of the data** **Uncertainty of a single measurement**	$\sigma = \sqrt{V} = \sqrt{\sigma^2}$

In many cases the distribution of the measured values follows a gaussian distribution. Other important distributions include the binominal distribution or Poisson statistics like e.g. for radioactive decays (see experiment Ab).
The Gaussian distribution reads:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{3}$$



The parameter σ is the standard deviation which defines the width of the distribution. It is related to the full width at half maximum (FWHM) by FWHM $\approx 2.35\,\sigma$. The area underneath the graph of $f(x)dx$ represents the probability to find a measured value within the interval dx. Statistically, 68% of all measured values lie inside $[\mu - \sigma, \mu + \sigma]$, and 95% or 99.7% of all measured values lie within the 2σ- or the 3σ-intervals, respectively.

1.4 Estimation of the true values

In practice, neither μ nor σ are known and have to be estimated based on a number of n sample measurements. Considering that, one has to make a distinction between two cases, which you are going to encounter both in these lab courses:

1.4.1 ... in the case of repeated measurements under identical conditions

This means, that the very same measurement is being repeated multiple times, such as the measuring of the period of a pendulum with constant length and for always the same initial amplitude. As a consequence we may assume that the statistical uncertainties, i.e. the standard deviation be always the same for all measurements.
Let $x_1, x_2, ..., x_n$ be the n sample measurements. One then may use the following estimations:

- Estimation for the mean value μ using the arithmetic mean value of the single measurements:

$$\mu \approx \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{4}$$

- Estimation for the standard deviation σ:[1]

$$\sigma \approx s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{5}$$

The denominator $(n-1)$ for s takes into account that we use the same n data for calculating the mean value $\bar{x}$ which in turn enters the calculation of s. One speaks about n degrees of freedom of our sampling with one being used for the mean value.
It is obvious that the more measurements are taken, the more accurate the estimations will be:

$$\lim_{n\to\infty} \bar{x} = \mu \quad \text{and} \quad \lim_{n\to\infty} s = \sigma.$$

The **central limit theorem** states that for different samples of n measurements each, the distribution of the corresponding mean values $\bar{x}$ always converges on a Gaussian distribution with a variance σ^2/n for large n, irrespective of the initial distribution of measurement data x_i.
This theorem allows for the estimation of the statistical error, which is called the uncertainty of the estimated mean:

$$m_{\bar{x}} = \sqrt{\sigma^2} = s/\sqrt{n}.$$

In brief: For the measurement of a quantity using n samples with measured values x_i, the true mean value and its uncertainty can be estimated using the following formulas:

[1] Often, the letter σ is used for the estimated standrad deviation instead of s. Doing so one always has to keep in mind that we deal with an estimate rather than with a precise value.

Result of the measurement:

$$\bar{x} \pm m_{\bar{x}} \tag{6}$$

Where $\bar{x}$ is the estimate for the true physical quantity

$$\bar{x}_x = \frac{1}{n}\sum_{k=1}^{n} x_k, \tag{7}$$

and m the estimate for the uncertainty:

$$m = \frac{s}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)}\sum_{k=1}^{n}(x_k - \bar{x})^2} \tag{8}$$

Generally, one should specify one or maximum two decimal places for the error. As a rule of thumb, we use two digits if the first digit of the error is smaller than 4, otherwise only one. Errors are always rounded off towards higher values. The number of decimal places must be the same both for the error and the measured value.

Example: $x = 1.57 \pm 0.24$ with $m_x = 0.24$ or $r = m_x/x = 0.15$.

While the error m_x of the mean value decreases with a increasing number of samples, the average deviation of single measurements s converges on σ. The width of the distribution of the measured values, i.e. the scatter of the data, depends only on the experimental conditions but not on the number of samples taken.
For computing the standard deviation and the statistical uncertainty, one commonly uses the following conversions:

$$\sum_{k=1}^{n}(x_k - \bar{x})^2 = \sum_{k}({x_k}^2 - 2x_k\bar{x} + \bar{x}^2) = \sum_{k}{x_k}^2 - n\bar{x}^2 \tag{9}$$

$$\Rightarrow m = \sqrt{\frac{1}{n-1}(\frac{\sum_k {x_k}^2}{n} - \bar{x}^2)}. \tag{10}$$

1.4.2 ... in the case of repeated measurements with different uncertainties

If the measured quantity is measured several times with different accuracy due to different conditions like, e.g. the speed of the gyroscope in experiment K, the data taken must be weighted by their respective uncertainties *before* averaging. This is referred to as **weighted mean value**. Those values which have small uncertainties will have higher weight in the average than those with larger scatter.
We measure a quantity x n times, where we obtain the values x_i and the corresponding uncertainties m_i (we didn't use the letter σ or s;on purpose). Defining the weights as $g_i = 1/m_i^2$, one gets the mean value:

$$\bar{x} = \frac{1}{G} \sum_{i=1}^{N} g_i \, x_i = \frac{1}{G} \sum_{i=1}^{N} \frac{x_i}{m_i^2}. \tag{11}$$

The parameter G is used for the normalization:

$$G = \sum_{i=1}^{n} g_i = \sum_{i=1}^{n} \frac{1}{m_i^2}.$$

Likewise the error on the mean value has to account for the weighting. This leads to:

$$m_{\bar{x}} = \sqrt{\frac{1}{G}}. \tag{12}$$

As an exercise, you may recover the the formulae for the case of measurements with identical uncertainties in section 1.4.1 by using $m_i = \text{constant} = s$.

1.5 Error propagation

When having measured one ore various physical quantities, one often has to perform further calculations using those initial results:
Let $u = f(x_1, x_2, \ldots, x_N)$ be a function which depends on some N variables. We would like to know how large the uncertainty m_u is. In order to do this, one uses error propagation. Let m_i be the uncertainty of the value of x_i.
Equation of error propagation:

$$m_u^2 = \left(m_{x_1} \frac{\partial u}{\partial x_1} \right)^2 + \left(m_{x_2} \frac{\partial u}{\partial x_2} \right)^2 + \ldots + \left(m_{x_N} \frac{\partial u}{\partial x_N} \right)^2 \tag{13}$$

where $\partial / \partial x_i$ denotes the partial derivative with respect to variable x_i.
For relations which include only summations and subtractions or only multiplications, divisions and powers, one uses the following simplified principles:

Summations:

Let $u = x + 3y - z$, whereas one measured:

$$x = \bar{x} \pm m_x \qquad y = \bar{y} \pm m_y \qquad z = \bar{z} \pm m_z$$

Then one ahas

$$\bar{u} = \bar{x} + 3 \cdot \bar{y} - \bar{z} \qquad \text{where} \quad m_u^2 = m_x^2 + (3 \cdot m_y)^2 + m_z^2 \tag{14}$$

→ The absolute errors, multiplied with the (constant) pre-factors, sum up quadratically.

Multiplications:

Let $u = xy/z \cdot a^3$, whereas one measured:

$$\begin{array}{rrrr} x = \bar{x} \pm m_x & y = \bar{y} \pm m_y & z = \bar{z} \pm m_z & a = \bar{a} \pm m_a \\ \text{resp. } r_x = m_x/\bar{x} & r_y = m_y/\bar{y} & r_z = m_z/\bar{z} & r_a = m_a/\bar{a} \end{array}$$

Then one has:

$$\bar{u} = \bar{x}\bar{y}/\bar{z} \cdot \bar{a}^3 \qquad \text{with} \quad r_u^2 = r_x^2 + r_y^2 + r_z^2 + (3 \cdot r_a)^2 \quad \text{and} \quad m_u = r_u \bar{u} \tag{15}$$

→ The relative errors, multiplied with the respective powers, sum up quadratically.

Important! Both rules are only valid, if the measured errors are **independent** of each other.

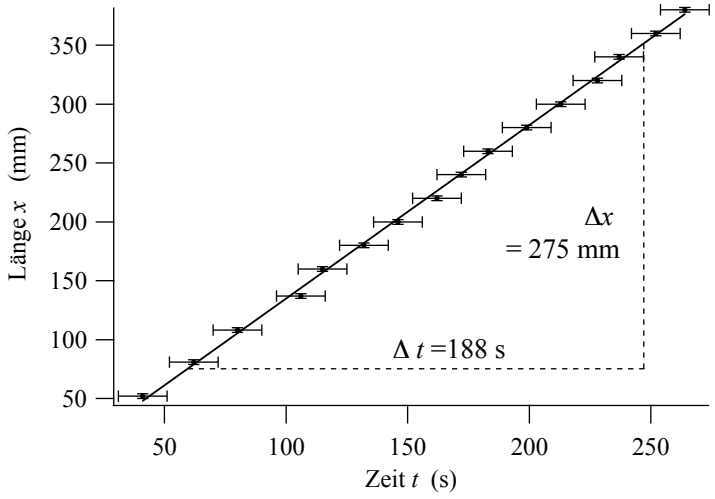
2 Graphic representation and evaluation of measured results

2.1 Graphic representation

Measured results are often displayed in form of graphs. Thereby, one plots the results on the ordinate (y-axis) against the quantity (abscissa or x-axis) which has been varied in the associated experiment.
Example: For the measurement of the velocity, one measures the walked distance as a function of time. The measuring errors have been estimated with the measurement accuracy, which is assumed to be constant in this example. They are being drawn as horizontal (error of the time measurement) or as vertical (error in the measurement of distance) bars respectively, whose lengths correspond to the particular measuring error. They are known as **error bars**.
The graph shows that it is about a straight line, i.e. a graph with constant slope. In our example, this line corresponds to the equation

$$\Delta x \;=\; v \cdot \Delta t$$

for a constant velocity v.
Remark: The data points should **not** be connected with lines, unless one has not enough measuring measuring points to be able to draw a smooth curve.



The boundaries, i.e. the minimum and maximum values, which are represented on the axis, should always correspond to the lowest and highest data values, respectively. Furthermore, graphical extrapolations are not very inaccurate in general. Suppose one would like to determine a systematic error by "drawing" (extrapolating) a line to zero, which lies outside of the actual measuring interval, one would only get a rough approximation at the most. It is superior to determine the slope and calculate the value to be extrapolated.

2.2 Fitting lines, slope of a line

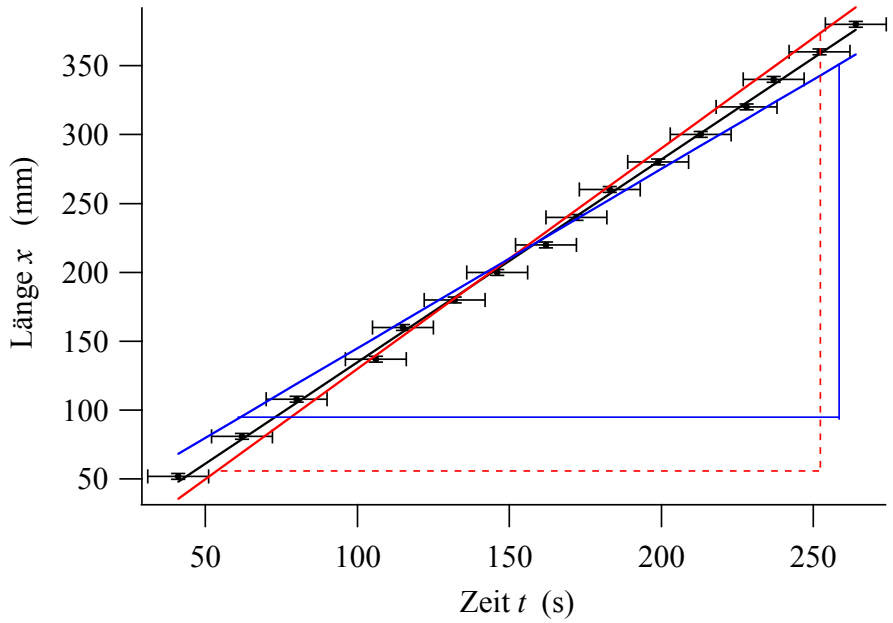
In our example, we presume that the velocity is constant and the function should represent a straight line. We want to determine the velocity. From the equation above, one can identify the velocity with the slope of the curve.
In order to extract the slope, one draws the straight line which matches best all data points including their error bars. This line is used to draw the slope triangle such that one side of this rectangular triangle runs parallel to the abscissa (x-axis) whereas the other one rises parallel to the ordinate (y-axis). The slope triangle should be drawn as large as possible. From the side lengths of the triangle one can determine the slope according to

$$\text{slope} = \frac{\text{side length along the ordinate}}{\text{side length along the abscissa}} \quad \rightarrow \quad v = \frac{\Delta x}{\Delta t} = \frac{275 \text{ mm}}{188 \text{ s}} = 1.46 \text{ mm/s}.$$

Hereby, the side lengths refer to the corresponding axes (units!!!).
To estimate the **error** of the determined slope, one draws the flattest and steepest lines, which are compatible with the data points, next to the best fit. The difference of the two extremal slopes yields a good estimate for the error on the slope:

$$a \pm \Delta a \text{ where} \quad \Delta a = \frac{a_{max} - a_{min}}{2},$$

where a, a_{max} and a_{min} are the slopes of the best fit and the extremal lines.



2.3 Linearization

Frequently, one has to face the problem that the measured points follow a non-linear but well-defined curve. As an example consider the count rate of x-rays or radioactive radiation behind a shielding as a function of layer thickness (see lab course Ab) or the absorption of light in a solvent as a function of concentration or thickness. Note that in such cases either

1. one adapts the graphic representation (logarithmic or semi-logarithmic scaling) or

2. one converts the quantity into a linear quantity by taking, e.g. the logarithm.

It is important to take the data processing into account when extracting the slopes of the graphs in such a linearized representation, in particular, if (semi-) logarithmic paper is being used: one has to bear in mind that the values used for the actual calculations have to be transformed in the same way as the curve itself!
An example for two possible representations of a non-linear function are being shown in the following figure. You will use both methods in the lab courses, in particular graphs on semi-logarithmic paper.