

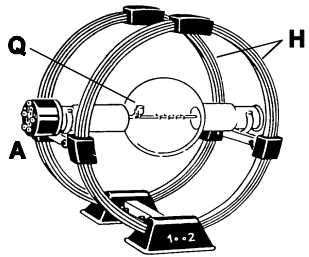
## Exp. B 8: Determination of the specific Charge e/m of the Electron

### 1. Literature:

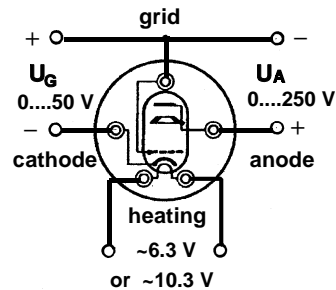
**Keywords:** Electron tubes, forces of electric and magnetic fields on electric charges, magnetic field strength and magnetic induction, ionization and recombination, emission of light, Millikan experiment

### 2. Basics

The aim of this experiment is the determination of the specific electric charge e/m of the electron by the deflection of an electron beam in a magnetic field. Similar experiments had been performed first by H. Busch (1922).



**Fig.1:** Beam tube with Helmholtz coils H, electron source Q, base and connection socket A



**Fig.2:** Connection socket with circuit scheme for beam generation unit

Fig. 1 shows the beam tube with Helmholtz coils, Fig. 2 shows the end side of the beam tube with circuit scheme and connection sockets for the hot cathode, the grid, and the anode with a hole. The electron beam is generated by thermal emission of electrons from the heated cathode and by their subsequent acceleration and focussing. If an accelerating voltage U is applied across the cathode and anode, the electrons (mass m, charge e) leave the anode through the hole with a certain velocity  $\underline{v}$ . According to the law of energy conservation, the kinetic energy

of the electron is equal to the electric work of acceleration:

$$\frac{1}{2}mv^2 = eU \quad (1)$$

Having left the anode, the electrons fly in a straight beam with constant velocity  $\underline{v}$  in free space in the absence of electric or magnetic fields. If, however, a magnetic field with induction  $\underline{B}$  is applied, the *Lorentz force*  $\underline{F}_L$  acts on the electrons, whereby

$$\underline{F}_L = -e(\underline{v} \times \underline{B}). \quad (2)$$

Since by definition of the vector product  $\underline{v} \times \underline{B}$  the force  $\underline{F}_L$  always acts in a direction perpendicular to the common plane of the vectors  $\underline{v}$  and  $\underline{B}$ , only the direction of the velocity  $\underline{v}$ , but not the absolute value of  $\underline{v}$  is changed.

Considering the path of an electron moving with velocity  $\underline{v}$  perpendicular to the B-vector of a homogeneous magnetic field, the absolute value of  $\underline{F}_L$  is simply

$$F_L = e v B. \quad (3)$$

The Lorentz force acts permanently with constant value  $F_L$  perpendicular to  $\underline{v}$ , and the electron is forced into a circular path with centripetal force equal to  $F_L$ , which means that

$$e v B = m \frac{v^2}{r}, \quad (4)$$

with r the radius of the circle. Combining eqs. (1) and (4) gives

$$\frac{e}{m} = \frac{2U}{r^2 B^2}. \quad (5)$$

The acceleration voltage U and the path radius r are to be measured directly in this experiment. The B-value of the magnetic field has to be determined from the used geometry for the field generating coil system and the current I through the coils. In free space (vacuum), the relation between magnetic induction B and magnetic field

strength  $H$  is

$$B = \mu_0 H, \quad (6)$$

where  $\mu_0 = 4\pi \times 10^{-7} \text{ VsA}^{-1}\text{m}^{-1}$  is the *magnetic permeability constant*. In principle one can calculate the field strength  $H$  of any particular coil system using *Biot-Savart's law*. For the field strength or induction in the center of a ring coil with  $n$  windings and radius  $R$ , one gets, if the length  $l \ll R$

$$H = \frac{nI}{2R} \quad \text{and} \quad B = \mu_0 \frac{nI}{2R}. \quad (7)$$

To create a homogeneous magnetic field, one uses two circular coils (*Helmholtz coils*) on a common axis with coil center distance equal to the coils radius  $R$  (see Fig. 1). Provided that the current through both coils with equal number  $n$  of windings is the same, the magnetic induction  $B$  within the center region of the coil system can be easily calculated from the mean radius  $R$ , the number  $n$  and the current  $I$  from the equation

$$B = 0.715 \mu_0 \frac{nI}{R}. \quad (8)$$

The factor 0.715 results from the geometry of the coil system. Using eqs. (5) and (8) we finally get

$$\frac{e}{m} = C \frac{U}{(rI)^2}, \quad (9)$$

with the constant

$$C = 2 \left( \frac{R}{0.715 \mu_0 n} \right)^2. \quad (10)$$

The radius of the coil system used is  $R = 20 \text{ cm}$ ; the number of windings in each coil is  $n = 154$ . From the experimental results for  $U$ ,  $I$  and  $r$ , one can calculate the specific charge  $e/m$  using eq. (9).

The path of the electron beam becomes visible, because the electrons collide on

their way with inert gas molecules filled into the tube at some reduced pressure  $p \approx 10^{-1} \text{ Pa}$ . The ionization of the molecules results from the impact. After recombining with electrons, the gas molecules are found in energetically excited states. During subsequent transitions into energetically lower states or into the ground state, the molecules emit electromagnetic radiation mostly in the visible range (light).

Since the value of the electronic charge  $e = -1.6021 \times 10^{-19} \text{ C}$  can be determined independently (e.g. in the experiment of R. A. Millikan (1910), charged oil droplets within a capacitor), the electron mass  $m_e = 9.1091 \times 10^{-31} \text{ kg}$  can be calculated from the ratio  $e/m$ .

The deflection of electron beams by magnetic fields is technically used in TV or conventional computer monitors as well as in magnetic lenses of electron microscopes. The technique can be applied also to other charged particles (protons, ions, etc.). This is done, e.g., in mass spectrometers and acceleration installations in the field of elementary particle and nuclear physics (cyclotron, synchrotron).

### 3. Experimental Tasks

- 1. Task:** Using the electron beam tube, the specific charge  $e/m$  has to be determined from several measurements of the coil current and the acceleration voltage for different given radii of the circular electron beam.
- 2. Task:** Possible errors in the determination of  $e/m$  have to be critically discussed and a calculation of the statistical error of  $e/m$  due to the errors in the single measured quantities  $U$ ,  $I$ ,  $r$ ,... has to be performed.

### 4. Carrying out the Experiment

To provide the magnetic field, both coils (see Fig. 1) are connected in series with a stabilized variable DC current source. The current  $I$  through the coils is measured using a DC Ammeter.

The beam tube is electrically connected with the DC voltage sources for the anode voltage  $U_A$  and the grid voltage  $U_G$  according Fig. 2 (*Check the polarity!*) and the AC voltage source for the cathode heating. This AC voltage may be different for the different tubes in operation, i.e., 6.3 V or 10.3V, as indicated on the tube socket. The acceleration voltage  $U$  to be determined is the sum of  $U_A$  and  $U_G$ .  $U$  is

measured using a DC voltmeter connected across the cathode and the anode of the tube.

*Important:* Before the voltage and current sources are switched on, the circuits have to be checked by the assisting student! Furthermore, the potentiometers for the variation of  $U_A$  and  $U_G$  have to be checked for zero voltage positions before the voltage source is put into operation. Wait about 1 min. after the start of cathode heating before varying  $U_A$  between 0 and 250V and  $U_G$  between 0 and 50V.

The sharpness and brightness of the electron beam is optimized by varying the voltage  $U_G$ . During longer breaks,  $U_A$  and  $U_G$  have to be set to zero, but not the cathode heating voltage. If after the heating up time the beam is visible, a certain acceleration voltage  $U (= U_A + U_G)$  is adjusted, the current through the coils switched on, and the tube cautiously rotated until the beam leaves the anode plate perpendicular to the direction of the magnetic field to complete a full circle. The current through the coils is now varied until the beam completes a circle with radius  $r = 2, 3, 4$  and  $5$  cm, as indicated by different fluorescent marks within the tube.

The following are to be considered for the measurements:

For a given acceleration voltage (e.g. 150V), the current  $I$  through the coils has to be measured 4 times for each of the 4 radii. Afterwards, the mean current for each radius has to be determined.

The experiment has to be performed for at least 4 different acceleration voltages (e.g. 150 V, 200 V, 250 V and 300 V).

## 5. Analysis and Discussion

First, the constant  $C$  is to be calculated according eq. (10), then  $e/m$  according eq. (9) for each set of corresponding values  $U$ ,  $r$  and the mean of  $I$ . Finally, the mean value of  $e/m$  is to be calculated and compared with the value known from literature.

An error calculation for  $e/m$  is to be performed (see appendix). Use the following errors:  $\Delta r = \pm 0.5\text{mm}$ ,  $\Delta R = \pm 2\text{mm}$ ,  $\mu_0$  and  $n$  without error.  $\Delta U$  and  $\Delta I$  depend on the grade of the instruments used (Please ask the assisting student). In addition,  $\Delta I$  is determined from the standard deviation of the average value (statistical uncertainty as calculated for every 4 measured  $I$  values). The calculation is to be continued using the larger of both values  $\Delta I$ . The maximum relative error of  $e/m$  is to be calculated according to the law of statistical error propagation for each

radius  $r$  and voltage  $U$ . Since  $e/m$  is the product of different powers of  $R$ ,  $U$ ,  $r$  and  $I$ , the calculation of the maximum relative error is quite simple (see appendix).

The results are to be discussed, and possible systematic errors (e.g. electrostatic charging of the tube, electrical potentials of the metal pieces) are to be taken into account.

## 6. Questions for Self-checking

- 1) How can free electrons be generated (except by thermal emission)?
- 2) How does the radius and the orbit frequency of the electron beam depend on the velocity or acceleration voltage?
- 3) How can one prove that the circular beam consists of negatively charged particles?
- 4) Which path do charged particles follow when entering the magnetic field in an angle?
- 5) Why do the values of  $e/m$  become smaller for very high acceleration voltages?
- 6) What is understood by the terms *impact ionization* and *recombination*?
- 7) Describe the microscopic mechanism of atomic light emission.

## Appendix: Analysis of Errors

### 1. Systematic and Statistical Errors

Every result of a measured physical quantity inevitably contains an error. To evaluate the experimental result obtained, it is, therefore, necessary to give an estimate of the numerical error(s) inherent to the experimental quantity.

There are different kinds of errors: systematic and random, i.e., statistical. Systematic errors are caused by the measurement system and can be recognized from the fact that the measured numerical value is strictly too large or too small as compared to those obtained when using other methods of measurement or theory. To minimize systematic errors, one has to change the experimental setup, i.e., the apparatus or the measuring procedure. Alternatively, the numerical result has to be corrected properly to account for the systematic errors involved in the measurement.

A statistical error arises due to random positive and negative deviations of the actually measured value from the mean or precise experimental value. If, e.g., the measured length of a distance is not exactly that of a certain number of scale divisions, one has to estimate a more accurate value by interpolation, thus a statistical error may result. On the other hand, the displayed (analog or digital) value of, e.g., a measured voltage can vary with time, i.e., fluctuating around an unknown (average) value. Thus, by evaluation of a time averaged value, or by the choice of the measurement moment, a statistical error may arise as well. Statistical errors are characterized by a probability distribution, which determines the probability of a measured deviation from the precise (true), i.e., most probable value (expectation value). The more a measurement is repeated, the more precisely are the probability distribution and the most probable value determined and the less becomes the statistical (measuring) uncertainty (see below).

### 2. Average Value, Standard Deviation, Statistical Uncertainty

The best evaluation for the most probable value of a measured quantity  $x$  out of  $n$  different single measurements  $x_i$  is obtained by taking the *arithmetic average value*  $\bar{x}$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i . \tag{1}$$

Very often a shortened notation is used in the literature containing statistical equations with expressions of sums, where the lower and upper summation limit and the summation index is suppressed:

$$\bar{x} = \frac{1}{n} \sum x . \tag{2}$$

This shortened notation will be used in the following text.

Once the average value is known, the moduli of differences  $|x_i - \bar{x}|$ , denoting the deviations of single measuring results from the average value, give some estimate for the precision of the measurements. Because of numerical reasons, the moduli of the differences are replaced by the squares  $(x_i - \bar{x})^2$  and the *standard deviation*  $s$  is defined by

$$s = \sqrt{\frac{\sum (\bar{x} - x)^2}{n - 1}} , \quad n > 1 . \tag{3}$$

The standard deviation gives the statistical average error of a *single* measurement. By introducing the (positive) square root, the quantity  $s$  has the same units as the measured quantity and is therefore comparable with it. The division by  $n - 1$  instead by  $n$  takes into account, that for only a single measurement ( $n = 1$ ) no statistical statement can be given, i.e.,  $s$  is not defined.

The equation defining the standard deviation is usually not that given by eq. (3). All pocket calculators use instead the fully equivalent equation

$$s = \sqrt{\frac{\sum x^2 - \frac{1}{n} (\sum x)^2}{n - 1}} , \quad n > 1 , \tag{4}$$

because, when using eq. (4), not the single measured values  $x$ , but only the sum  $\sum x$  and the sum of the square  $\sum x^2$  has to be stored.

Besides the calculation of the average value and the standard deviation, it is often interesting to consider the value of the statistical uncertainty of the average value as well. This is because  $\bar{x}$  is just a guess of the result  $x$  according eq. (2), which for a small number  $n$  of single measurements, can be very unprecise. The

statistical measuring uncertainty  $u$  is a measure of the (statistical average) error of the average value  $\bar{x}$ :

$$u = \frac{s}{\sqrt{n}} = \sqrt{\frac{\sum x^2 - \frac{1}{n}(\sum x)^2}{n(n-1)}} \quad , \quad n > 1 \quad . \quad (5)$$

While the standard deviation  $s$  as a measure of the statistical spread of single measured values  $x_i$  approaches a finite value  $> 0$  with increasing  $n$ , the statistical measuring uncertainty  $u$  of the average value  $\bar{x}$  decreases with increasing  $n$  and approaches zero at large  $n$ .

Very often the measured values  $x_i$  are so called "normally" distributed, i.e., their relative probabilities are given by the *Gaussian distribution function*  $\varphi(x)$ :

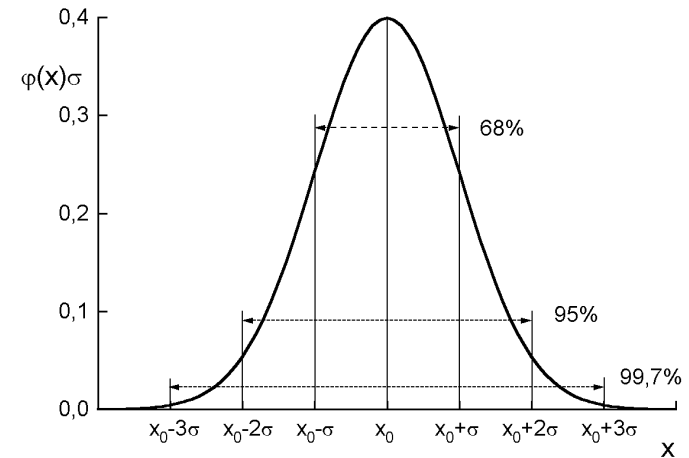
$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right) \quad . \quad (6)$$

The integral

$$P(x_1, x_2) = \int_{x_1}^{x_2} \varphi(x) dx \quad , \quad (7)$$

gives the probability that the values  $x_i$  (for a large number of measurements  $n \rightarrow \infty$ ) lie within the interval  $(x_1, x_2)$ . As shown by Fig. 1, the function  $\varphi(x)$  is symmetric around the most probable value  $x_0$  (the expectation value) and has the shape of a bell with a full width at half maximum of somewhat more than  $2\sigma$ .

For very large  $n$ , the average value  $\bar{x}$  determined from the measuring series approaches the value  $x_0$  of the function  $\varphi(x)$ , and the standard deviation  $s$  approaches the value  $\sigma$ . The probability that the result  $x_i$  of a single measurement lies within the interval  $\bar{x} \pm s$ , i.e.,  $x_0 \pm \sigma$ , amounts according to eq. (7) to about 68 %, for the interval  $x_0 \pm 2\sigma$  to about 95 % and for  $x_0 \pm 3\sigma$  already 99.7 %. Similar relations are valid for the statistical measuring uncertainty  $u$  of the average value: the probability of the true value  $x_0$  lying within the so called (unit) *range of*



**Fig. 1:** Gaussian distribution function  $\varphi(x)$  with most probable value  $x_0$  and intervals  $x_0 \pm \sigma$ ,  $x_0 \pm 2\sigma$ , and  $x_0 \pm 3\sigma$  for relative frequencies 68%, 95%, and 99,7% of measuring value  $x_i$  confidence  $\bar{x} \pm u$  is about 68 %, for the twofold and threefold range of confidence  $\bar{x} \pm 2u$  and  $\bar{x} \pm 3u$  about 95 % and 99.7 %, respectively.

One has to take into consideration, however, that the evaluation of the statistical measuring uncertainty or the range of statistical confidence for the value  $\bar{x}$  is physically reasonable only in combination with the evaluation of possible systematic errors. The total error of a result obtained from a measurement is always the sum of moduli for systematic and statistical errors. It is, therefore, not useful to repeat a measurement very often just for minimizing the statistical error if the evaluated systematic error exceeds the former by orders of magnitude.

**Example 1:** In the experiment B10 (experiment with x-rays) the lattice constant  $d$  of NaCl is to be determined from the Bragg reflection using Bragg's law

$$n\lambda = 2d\sin\theta_n \quad ,$$

with the wavelength  $\lambda = 154$  pm and the measured Bragg angles  $\theta_n$  ( $n = 1, 2, 3$ ).

Typical values are given in the table below.

n	$\theta_n$ (°)	$d_n$ (pm)
1	16.0	279.4
2	33.2	281.3
3	55.15	281.5

Using these values the following are found:

the average value:  $d = 280.73$  pm,

the standard deviation:  $s = 1.16$  pm,

the statistical uncertainty  $u$  of the average value  $d$ ,  $u = 0.67$  pm,

Finally,

$$d = (280.73 \pm 0.67) \text{ pm} = 280.73 \text{ pm} \pm 0.2 \%$$

if no error (statistical or systematic) in the measured angles  $\theta_n$  is considered. (Otherwise see example 2 below!)

**Notice:** Any estimated error should not contain more than 2 or 3 decimal positions. The final result is to be rounded correspondingly. Very often it does not make any sense and only simulates precision to present the many decimal places displayed by a pocket calculator.

**Additional notice:** Comparison of the result for  $d$  given above with the value  $d_{\text{NaCl}} = 282.0$  pm, known from the literature, and with those obtained using other experimental setups in the laboratory, reveals a systematic deviation of  $\theta_n$  (of roughly 0.5% - depending on the system used - because of inaccurate adjustments). In this case, the value of the statistical uncertainty presented is not significant at all to estimate the (actual true) error of  $d$ . If  $u$  is replaced by the standard deviation  $s$ , the result

$$d = (280.73 \pm 1.16) \text{ pm} = 280.73 \text{ pm} \pm 0.4 \%$$

can account to a better extent the actual error but this presentation does not solve the problem principally, i.e., how the systematic error present in this case is properly taken into account. One possibility to solve the problem is to take into consideration that an inherent (constant) systematic error  $\Delta\theta^s$  (because of

misalignment of both the crystal and the detector) has to be taken into account when using Bragg's law for the evaluation of  $\theta$  from the experimental data, i.e.,  $\theta = \theta^m + \Delta\theta^s$ , where  $\theta^m$  is the measured value. Since in the experiment considered the task is not to validate Bragg's law but to determine the lattice constant  $d$ , we state that Bragg's law is valid in any order  $n$ , i.e., that the value of  $d$  is independent of  $n$ . Combining Bragg's law, e.g., for  $n = 1$  and 2,  $\Delta\theta^s$  can be calculated with the values in the table of example 1 from

$$\tan \Delta\theta^s = \frac{\sin \theta_2^m - 2 \sin \theta_1^m}{2 \cos \theta_1^m - \cos \theta_2^m}, \quad \Delta\theta^s = -0.2^\circ.$$

Using this systematic correction  $\Delta\theta^s$  of the angle  $\theta^m$ , the values  $d_1 = 282.8$  pm;  $d_2 = 282.8$  pm (i.e.,  $d_1 = d_2$ ) and  $d_3 = 282.2$  pm are obtained. Thus, the average of the corrected result for  $d$  is

$$d = 282.6 \text{ pm},$$

which is closer to the value  $d_{\text{NaCl}} = 282.0$  pm known from the literature than the value  $d = 280.73$  pm presented above, without taking into consideration the correction of the systematic error.

### 3. Propagation of Errors

Generally, the determination of a physical quantity  $y$  requires the measurement of several single (different) parameters  $x_1, x_2, \dots$ . In this context, the question arises how far the single errors  $\Delta x_i$  resulting from the measurement of the individual parameters determine the uncertainty  $\Delta y$  of the quantity  $y$ . If the errors  $\Delta x_i$  are small as compared to  $x_i$ , the function  $y$  expanded into a power series of the errors  $\Delta x_i$  around the values  $x_i$  is approximately given by the terms linear in  $\Delta x_i$ . The error  $\Delta y$  resulting, e.g., from a single error  $\Delta x_i$  is given by

$$\Delta y \Big|_{x_{j \neq i} = \text{const}} = \frac{\partial y}{\partial x_i} \Delta x_i. \quad (8)$$

To estimate the *largest possible absolute error*  $(\Delta y)_{\text{max}}$  by taking into account all possible single errors  $\Delta x_i$ , the following relation is defined:

$$(\Delta y)_{\max} = \left| \frac{\partial y}{\partial x_1} \Delta x_1 \right| + \left| \frac{\partial y}{\partial x_2} \Delta x_2 \right| + \dots \quad (9)$$

**Example 2:** In the experiment B10 (experiment with x-rays, see also example 1) the lattice constant  $d$  of NaCl is to be determined from the Bragg reflection using the relationship

$$n\lambda = 2d_n \sin\theta_n, \quad d = \frac{1}{3} \sum d_n = \frac{\lambda}{6} \left( \frac{1}{\sin\theta_1} + \frac{2}{\sin\theta_2} + \frac{3}{\sin\theta_3} \right),$$

with the wavelength  $\lambda = 154$  pm and the measured Bragg angles  $\theta_n$  ( $n = 1, 2, 3$ ). Assuming a statistical error  $\Delta\theta = \pm 0.25^\circ = 0.0044$  of every measured angle  $\theta_n$ , the maximum statistical error  $\Delta d_{\max}$  is according to eq. (9)

$$(\Delta d)_{\max} = \left| \frac{\partial d}{\partial \theta_1} \Delta \theta_1 \right| + \left| \frac{\partial d}{\partial \theta_2} \Delta \theta_2 \right| + \left| \frac{\partial d}{\partial \theta_3} \Delta \theta_3 \right|$$

$$= \frac{\lambda \Delta \theta}{6} \left( \frac{\cos\theta_1}{\sin^2\theta_1} + \frac{2\cos\theta_2}{\sin^2\theta_2} + \frac{3\cos\theta_3}{\sin^2\theta_3} \right).$$

Taking the numerical  $\theta$ -values presented in example 1 we obtain

$$(\Delta d)_{\max} = 0.112 (12.65 + 5.58 + 2.55) \text{ pm} = 2.33 \text{ pm}; \quad (\Delta d)_{\max} / d = 0.83 \%$$

It is noted that the error contribution from the first Bragg angle  $\theta_1$  is the largest one, because the error  $\Delta\theta_1/\theta_1$  is the largest one, consistent with the observation in the table of example 1, where the deviation of  $(d_1 - d_{\text{NaCl}})$  is the largest one too.

**Example 3:** In the experiment B8 (determination of the specific charge  $e/m$  of the electron) the ratio  $e/m$  is determined from measurements of the acceleration voltage  $U$  and of the current  $I$  through the coils with radius  $R$  for an electron beam with circular radius  $r$ .

$$\frac{e}{m} = 2 \left( \frac{R}{0.715 \mu_0 n} \right)^2 \frac{U}{(rI)^2}$$

Since  $e/m$  is a product of powers of the variables  $R$ ,  $U$ ,  $r$ , and  $I$ , the expression of the *largest possible relative error* of  $e/m$  is rather simple. Taking into account the errors  $\Delta R$ ,  $\Delta U$ ,  $\Delta r$ , and  $\Delta I$ , the largest possible relative error of  $e/m$  is given by

$$\left[ \frac{\Delta(e/m)}{(e/m)} \right]_{\max} = \frac{1}{e/m} \left( \left| \frac{\partial(e/m)}{\partial R} \Delta R \right| + \left| \frac{\partial(e/m)}{\partial U} \Delta U \right| + \left| \frac{\partial(e/m)}{\partial r} \Delta r \right| + \left| \frac{\partial(e/m)}{\partial I} \Delta I \right| \right) \quad (10)$$

$$= \left| \frac{2\Delta R}{R} \right| + \left| \frac{\Delta U}{U} \right| + \left| \frac{2\Delta r}{r} \right| + \left| \frac{2\Delta I}{I} \right|.$$

Typical estimates of errors and values of measured quantities are

$\Delta R = \pm 2$ mm	$R = 20$ cm	$\Delta R/R = 0.01$
$\Delta U = \pm 6$ V	$U = 200$ V	$\Delta U/U = 0.03$
$\Delta r = \pm 0.5$ mm	$r = 3$ cm	$\Delta r/r = 0.017$
$\Delta I = \pm 90$ mA	$I = 2$ A	$\Delta I/I = 0.045$

resulting in

$$\left[ \frac{\Delta(e/m)}{(e/m)} \right]_{\max} = 0.02 + 0.03 + 0.034 + 0.09 = 0.174.$$

#### 4. Graphs and Linear Regression

In experimental physics, the aim is often to validate a theoretically predicted functional dependence of two quantities  $x$  and  $y$  by a measurement. In simple cases, the quantities  $x$  and  $y$  are linearly connected, i.e.,

$$y = a + bx. \quad (11)$$

Since every measurement is inherent with an error, the data  $(x_i, y_i)$  in a graph  $y_i$  vs.  $x_i$  will scatter more or less around a straight line drawn through the data points. It is the task to find an optimum straight line balancing the deviations due to errors (i.e. that line which would represent the data if the errors were absent). This is the *regression line*. It can be simply done by a *visual estimate*, i.e., by drawing a suitable straight line with a ruler. A more objective way to construct the regression

line is the *least mean square fit*; the calculation of the parameters a and b of eq. (11) (a: intercept on y-axis, b: slope of line) with the condition

$$\sum (a + bx - y)^2,$$

being an absolute minimum. If this is the case, a and b fulfil the condition

$$\frac{\partial}{\partial a} \sum (a + bx - y)^2 = 0, \quad \frac{\partial}{\partial b} \sum (a + bx - y)^2 = 0. \quad (12)$$

Differentiation yields the equation system

$$an + b \sum x = \sum y, \quad a \sum x + b \sum x^2 = \sum xy, \quad (13)$$

with the solution

$$a = \frac{\sum y \sum x^2 - \sum x \sum xy}{n \sum x^2 - (\sum x)^2}, \quad b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}. \quad (14)$$

**Note:** For a regression line through the origin of the coordinate system, i.e. for the line  $y = b x$  ( $a = 0$ ) the corresponding solution is

$$b = \frac{\sum xy}{\sum x^2} \quad (15)$$

The linear regression algorithm using the method of the least mean squares is almost always applied to find an optimum fit to the data. It is (besides the (arithmetic average) the most frequently used algorithm and, therefore, is included in many pocket calculators.

When using the (critical) regression coefficient  $R^2$ , one has to be cautious. With  $R^2$  the quality of fit functions with several independent measuring series can be compared. The coefficient of a single measuring series, however, has no physical meaning.

In cases where the quantity  $Y(X)$  does not linearly depend on  $X$ , it may be possible to find a proper transformation  $Y(X) \rightarrow y(x)$  so that a linear relationship

$y = a + bx$  holds.

Some examples are given in the following table:

function Y(X)	transformation		y = a + b x	
	y(X,Y)	x(X,Y)	a	b
A X <sup>B</sup>	ln (Y)	ln (X)	ln (A)	B
A exp (BX)	ln (Y)	X	ln (A)	B
A X + B X <sup>3</sup>	Y / X	X <sup>2</sup>	A	B