

Exp. B07 - Frank-Hertz experiment

1 Basics

According to the Rutherford model, an atom consists of a positively charged core and negatively charged electrons, which orbit the core. The nucleus is very small compared to the size of the atom, but it contains almost the entire atomic mass. Since atoms themselves are not electrically charged, the positive charge of the nucleus is compensated completely by the negative charge of the electrons. The predictions of this model regarding the core were validated by scattering experiments with α -particles on metal foil. But the model could not explain the stability of atoms or the nature of the light emission (the occurrence of line/discrete spectra). According to the laws of the classical electrodynamics, the (by coulomb-interaction between core and shell of the atom) accelerated electrons should emit continuous radiation and collapse into the core due to their energy loss.

Bohr postulated, that electrons can orbit the core on trajectories of certain energies (shells) without radiating. During the transition of an electron between shells, electromagnetic radiation (light) is emitted or absorbed. The energy of the emitted photons (what 's this?) is $h\nu$, where ν is the frequency of the electromagnetic radiation. This energy corresponds to the energy difference between the shells, namely

$$\Delta E = h\nu. \quad (1)$$

The electron collision experiments, realized by Franck and Hertz in 1913, were a direct experimental confirmation of this quantum theory of the electron. They have shown, that atoms can get energetically excited by collisions with (free) electrons. The kinetic energy of the electron is transferred to the atom in energy quanta ΔE , which correspond to the characteristic excitation energy.

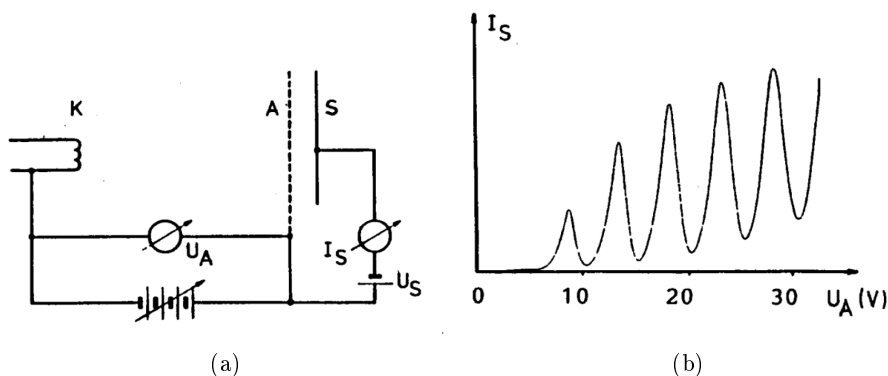


Figure 1: (a) Setup of the experiment of Franck and Hertz (schematically) with cathode K, grid-shaped anode A, counter electrode S, counter voltage U_S (b) measuring results for $I_S(U_A)$, the locations of the minima and maxima differ in each case by a constant voltage difference of about 4.9 V.

Figure 1a shows the setup used by Franck and Hertz schematically: In an evacuated glass bulb, filled with one mercury drop, a Hg-steam pressure of some mbar is generated by external heating of a temperature of 150 °C to 200 °C. The bulb

contains the electrodes K, A and S, to produce¹ free electrons and for the determination of the energy transfer by collisions with Hg-atoms. Electrons are released from the heated cathode K due to its thermal energy and accelerated by the voltage U_A towards the grid-shaped anode A.

The electrons collide with the Hg-atoms while traveling to the grid-shaped anode A (the mean free path of the electrons between two collisions in accordance with the Hg-steam pressure is about some μm). A part of the electrons is directly conducted to the anode grid. The other part of the electrons is braked due to a opposing field caused by the voltage U_S on its way to the counter electrode S, that lies behind anode A. Thus only electrons with a sufficient large kinetic energy can reach the electrode S. Figure 1b shows the current I_S to the counter electrode plotted versus the acceleration voltage U_A with a brake voltage $U_S = 1.2\text{ V}$. If the voltage U_A is increased continuously from zero, the current I_S for $U_A > U_S$ will increase monotonously at first in accordance with the Hg-steam pressure and corresponding to the tube-like characteristics of the diode in the space-charge region ($I_S \approx U_A^{3/2}$). After an exceedance of a certain threshold voltage U_{A1} the current I_S decreases and passes minima and maxima, whose locations differ in each case by a constant voltage difference of approximately 4.9 V, while U_A rises further. This behavior is explained by the type of the collisions, that the electrons suffer on the way between cathode and anode by the Hg-atoms. For voltages $U_A < U_{A1}$ all collisions are elastic, thus practically no kinetic energy is transferred because of the quite different masses of the particles brought to collision ($m_{\text{Hg}}/m_e = 3.7 \times 10^5$).

The kinetic energy of the electrons, that pass the anode, is nearly equal for all electrons

$$E = e(U_A - U_K).$$

Where e is the elementary charge and U_K the contact voltage between cathode and anode ($eU_K =$ difference of the electronic work function from these electrodes).

For voltages $U_A - U_K > U_{A1} - U_K = U_1 = 4.9\text{ V}$ inelastic collisions between electrons and Hg-atoms happen, too. Whereby the electrons emit the energy eU_1 , that corresponds to the Hg-excitation energy between the ground state 6^1S_0 and the resonance level 6^2P_1 . After an inelastic collision for acceleration voltages bigger than U_{A1} , the electrons does not possess enough energy to overcome the opposing field between anode and counter electrode. This leads to the decrease of the current I_S , compare Fig. 1b. If the voltage U_A is growing further, the current I_S rises again, since the kinetic energy of the electrons for passing the opposing field is rising according to the increasing voltage. When the acceleration voltage U_A minus the contact voltage U_K reaches the n -fold of the excitation voltage U_1 , n inelastic collisions between the electrons and the Hg-atoms happen on average. This results in the occurrence of further maxima in the current I_S .

Remark: With a different experimental arrangement of the Franck-Hertz-Experiment, fig. 1a, it is possible to determine other excitation energies of mercury and other atoms.

¹Beware! Electrons cannot be produced (charge conservation)! But electrons can be released from a material, thus **free** electrons are produced.

The relaxation to the ground state of the Hg-atoms after the excitation caused by collisions is in the present experiment observed by the, as a result of the process, occurring light emission (at the anode, see [3](#) experimental setup). The direct transition ($\lambda = 254\text{ nm}$), according to the energy 4.9 eV , though can't be observed, since the wavelength is small.

2 Assignment of tasks

2.1 First task

The $I_S(U_A)$ -characteristic curves of a Franck-Hertz-tube at two different temperatures have to be recorded by a x - y plotter:

- (a) at $T_1 = \text{approx. } 175^\circ\text{C}$ for $U_A = 0\text{ V to } 30\text{ V}$ and $U_S = 0.5\text{ V, } 1.0\text{ V, } 1.5\text{ V}$ and 2 V
- (b) at $T_2 = \text{approx. } 200^\circ\text{C}$ for $U_A = 0\text{ V to } 60\text{ V}$ and $U_S = 2\text{ V}$.

2.2 Second task

The results should be evaluated graphically and discussed afterwards. Determine the excitation voltage U_1 at the location of the maxima of the characteristic curves and compare it to the literature values. Discuss possible errors in the measurement.

3 Experimental setup

Figure [2](#) shows the circuit diagram of the present experiment. The Franck-Hertz-tube is covered by a housing with observation-windows at the side and back wall. The housing is heated by an electrical driven oven, that is connected to the AC grid (220 V) by a bimetal switch (T) and a var. autotransformer STr. Thus constant temperatures can be set in a wide range. A thermometer serves for the measurement of the temperature. The source for the acceleration voltage U_A , the counter voltage U_S and the cathode heating voltage (6.3 V) is a tube-power supply with a special control gear (shown in fig. [2](#) on the left beside the tube): The 0 V to 12 V output of the power supply is (gear) reduced with the potentiometer $10\text{ k}\Omega/3.3\text{ k}\Omega$ in proportion 4:1 for the generation of the counter voltage U_S (0 V to 3 V). The voltage output 0 V to 60 V (respectively 30 V) generates a temporally increasing voltage $U_A(t)$ (charging voltage of the capacitor) when the switch S_1 is opened through the potentiometer $10\text{ k}\Omega/2200\text{ }\mu\text{F}$. The voltage $U_A(t)$ is determined by a voltmeter and is attached to the x input of the plotter for recording the characteristic curves.

The current I_S of the tube (about 10^{-9} A to 10^{-8} A) will be measured by a DC measurement amplifier. Its output $U(I_S)$ (approx. 1 V) is connected to the y input of the plotter.

4 Experimentation

At first the Franck-Hertz tube must be heated. For this purpose the tube is driven by a turned on cathode heating and a var. autotransformer STr set to

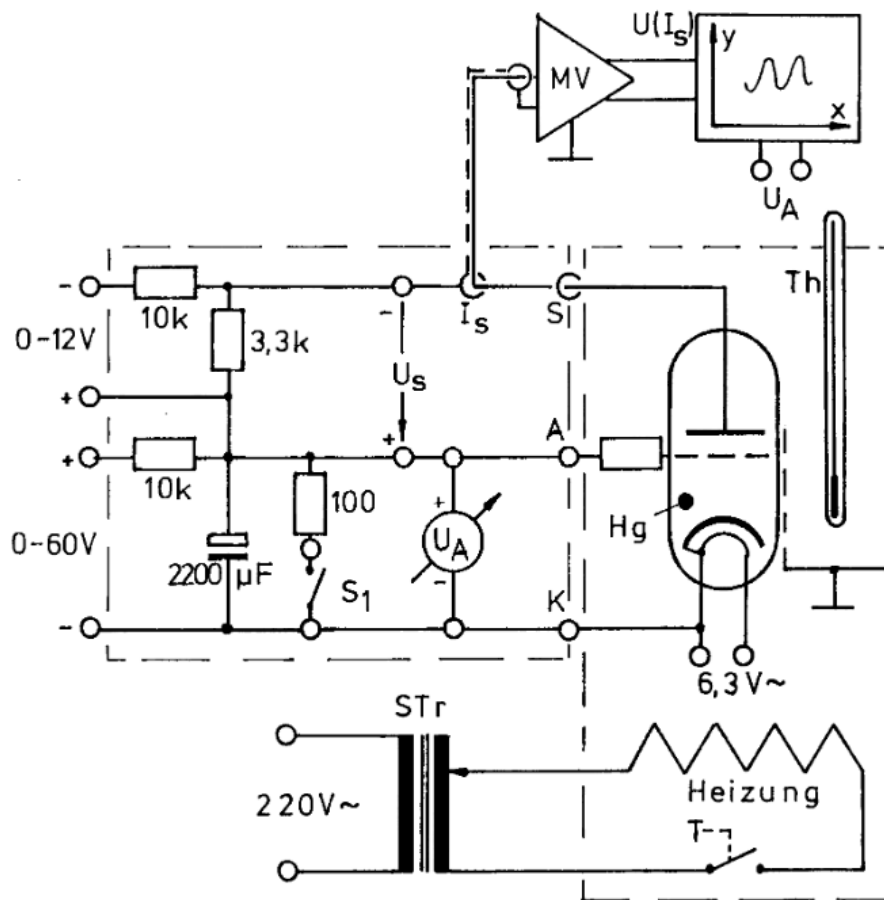


Figure 2: Circuit diagram of the Franck-Hertz experiment, with the potentiometer, the switch S_1 , a voltmeter U_A on the left, the Franck-Hertz tube (elektrodes S, A and K) and a thermometer Th on the right, a x - y plotter on the top and a cathode heating (Heizung) with a var. autotransformer STr on the bottom.

approx. 200 V. The bimetal thermostat on the side wall of the oven is set to approx. 175 °C. When the tube reaches a temperature of approx. 160 °C, the transformer is set back to 140 V. If necessary the bimetal thermostat is set to the desired temperature. The remaining setup is connected in accordance with the circuit diagram, fig. 2. The mass connections of the amplifier and the housing also need to be connected. Before you start the recording of the characteristic curves, become familiar with the handling of the amplifier and the plotter.

For recording the characteristic curves at about 175 °C the following settings are necessary:

Measurement amplifier range:	10^{-8} A
Plotter, y input:	0.1 V cm^{-1} var.: = ca. 30 mV cm^{-1}
Plotter, x input:	1 V cm^{-1} cal.

The 4 characteristic curves for all counter voltages U_S should be plotted together, if possible, after a test (start and end via opening and closing of S_1), into only one diagram.

For recording the characteristic curves at about 200 °C the following adjustments need to be performed:

Measurement amplifier range:	10^{-9} A
Plotter, y input:	0.1 V cm^{-1} var.: = ca. 30 mV cm^{-1}
Plotter, x input:	10 V cm^{-1} var.: = ca. 2 V cm^{-1}

Remark: If the tube is driven at about 200 °C, the characteristic curve in the range of $U_A < 60 \text{ V}$ should exhibit approx. 10 maxima and minima. For smaller temperatures, the ignition voltage for the corona discharge of the tube (what's this?) is below 60 V. Thus a corresponding smaller range for U_A needs to be chosen. A recording of the characteristic curve during the corona discharge is not useful for this experiment. The beginning of the corona discharge can be recognized by a stage in the characteristic curve of the tube and by light emission of the Hg-atoms around the cathode. At the same time it causes a rise of the anode current. The protective resistor (directly at the anode grid, see fig. 2) limits the current and prevents the destruction of the cathode respectively increasing its service life.

5 Evaluation

The characteristic curves need to be labeled (scale and zero points of the axes, specification of the respective brake voltage as curve parameter). The maxima of the curves should be identified/labeled and their spacings should be determined. The mean values of the spacings of the maxima, as well as the standard deviations have to be determined separately for both temperatures T_1 and T_2 . The results need to be compared to the literature value. Discuss the course of the characteristic curves and possible errors in the measurement.

6 Questions for Self-checking

- Interpret the course of the $I_S(U_A)$ characteristic curve (fig. 1b).
- Why does the current I_S in the minima does not go back to zero?
- Why a counter voltage U_S is used?
- Which impact does the temperature of the Franck-Hertz tube have on the current I_S ?
- Which kind of excitation takes place in the Hg-atom?
- Which wavelength does light, emitted by the Hg-atom after excitation caused by collisions, have? How does one prove it?

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 2σ .

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 σ . 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*

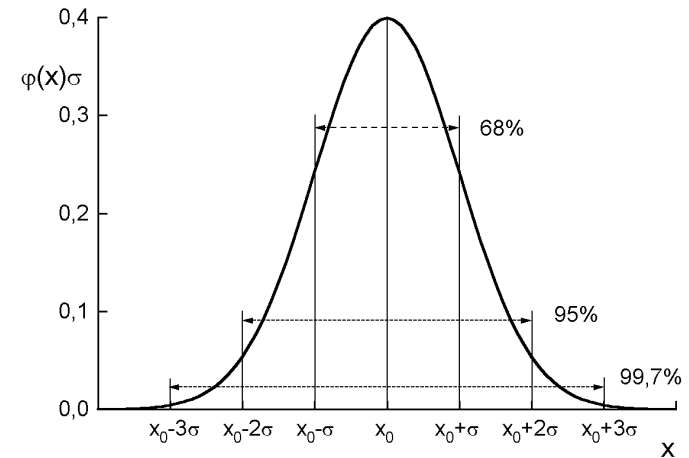


Fig. 1: Gaussian distribution function $\varphi(x)$ with most probable value x_0 and intervals $x_0 \pm \sigma$, $x_0 \pm 2\sigma$, and $x_0 \pm 3\sigma$ for relative frequencies 68%, 95%, and 99,7% of measuring value x_i confidence $\bar{x} \pm u$ is about 68 %, for the twofold and threefold range of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

Typical values are given in the table below.

n	θ_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 θ_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 θ_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 θ from the experimental data, i.e., $\theta = \theta^m + \Delta\theta^s$, where θ^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 θ^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 Δx_i resulting from the measurement of the individual parameters determine the uncertainty Δy of the quantity y . If the errors Δx_i are small as compared to x_i , the function y expanded into a power series of the errors Δx_i around the values x_i is approximately given by the terms linear in Δx_i . The error Δy resulting, e.g., from a single error Δx_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 Δx_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 θ_n ($n = 1, 2, 3$). Assuming a statistical error $\Delta\theta = \pm 0.25^\circ = 0.0044$ of every measured angle θ_n , the maximum statistical error Δd_{\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 θ -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 θ_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 ΔR , ΔU , Δr , and ΔI , 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 ^B	ln (Y)	ln (X)	ln (A)	B
A exp (BX)	ln (Y)	X	ln (A)	B
A X + B X ³	Y / X	X ²	A	B