Freie Universität Berlin · Department of Physics

Basic Laboratory Course in Physics

GPII

Two Semester basic laboratory course for students of Physics, Geophysics, Meteorology and for Teacher Candidates with physics as first or second subject.

Two Semester basic laboratory course for students of Physics, Geophysics, Meteorology and for Teacher Candidates with physics as first or second major.

Aim of the Laboratory Course

Introduction to the fundamental techniques of quantitative experimental- and scientific methods in physics (measurement methods, measurement techniques, documentation, mathematical-statistical und practical evaluation methods / error calculations, critical discussion and scientific conclusion, written report and presentation). Dealing with selected topics in physics in a deeper and complementary way.

Core Rules

- Preparation based on lectures and text books according to information contained in the script.
- The experiments begin c.t. and students arriving more than 15 minutes later will be excluded from taking part.
- The two page introduction (intended as part of the experimental report) is to be presented at the beginning of the experiment.
- The tutor introduces the students to the experiment and makes sure that they are sufficiently prepared and if not, whether the work should be repeated at a later date.
- The experiment and documentation of the results is made as quick as possible under the guidance of the tutor, whereby, time for further discussions of the physical background should be taken into account.
- Evaluation of the experiment by means of tables and graphs takes place after about 3 hours with the help of the tutor. Thereafter, further work is to be done on the report (protocol).
- The 4 hours are to be fully used to complete the protocol and can then only be cut short when the tutor hands out an attestation.
- The total number of experiments (as a rule 11) must be completed within the laboratory course, whereby a maximum of 2 experiments can be repeated at the end of the course.
- Attestations for all experiments must be noted at the latest on the last day of the course, otherwise the course can not be assessed and becomes invalid.

Integration with the Physics Curricula

Two laboratory courses (GP I and II) are scheduled after the respective lecture courses (Physics I and II). Restrictions with respect to the contents of the lectures are unavoidable due to the timescale and the placement of the laboratory course. This is especially evident for students taking part in the vacation laboratory courses where subjects must be handled in advance without prior lecture material (Optics, Atomic Physics Quantum Phenomena).

Organization

Semester Course (weakly, 4 h) and *Vacation Course* (4 weeks, 12 h per week).

Laboratory course in small groups. Pairs of students performing and evaluating an experiment. A tutor assists a group of 3 pairs on the same or related experiments. Good preparation before the experiment is important. A two page introduction to the subject matter is handed out before each experiment and is intended as part of the evaluation .

Course Schedule with Experimental Work, Evaluation and (as a rule) start of the written report (protocol).

Work on the two page introduction to the subject matter (prepared beforehand), presentation of the experimental findings with summary and critical discussion of the results.

Course Material: Description of the experiment (script) containing information on the relevant physics, experimental set-up and the tasks to be performed. Report book for the written experimental protocol – to be bought by the student.

Evaluation

Experimental certificate with grades according to ECTS (European Credit Transfer System). Point system for the individual experiments. No tests or final seminar.

Experiments

Experiments with various grades of difficulty from simple experiments in GP I, to give a basic feeling for the methods involved in experimental physics, to experiments with deeper physical background, which, for a fuller understanding, require higher lecture courses in physics.

Note

A sensitive indicator for physical understanding is the application of gained knowledge. The physical principles and the connections between phenomena should be demonstrated by dealing with the problems involved and by critical observation.

As a part of scientific training, it is not the intention of the laboratory course to only impart „mechanical knowledge“ but it should lead to scientific thinking, i.e., answering questions of a physical nature or drawing conclusions from findings and laws through critical discussions in small groups and final evaluation of the observations and quantitative results.

Edition: 22.12.006

Revision: Rentzsch

BASIC LABORATORY COURSE IN PHYSICS

Introduction to the fundamental techniques of quantitative experimental- and scientific methods in physics: Measurement methods, measurement techniques, documentation, mathematical-statistical und practical evaluation methods (error calculations), critical discussion and scientific conclusion, written report and presentation. Dealing with selected topics in physics in a deeper and complementary way.

Two laboratory courses (GP I and II) scheduled after the lecture courses Physics I and II, however, with reference to the complete material handled in lecture courses Physics I-IV.

Experiments and reports done in team work consisting of a group of 6 (3 pairs) under the assistance of a tutor.

Completion of introductory reports on the subject matter and physical background, presentation of the experimental findings with a summary and critical discussion of the results as an exercise in scientific writing.

Introductory text books provided the basic knowledge in a clear and connected manner, but only in passing, mention the way to the working methods of physics. Physical knowledge comes about either through quantitative observation of the natural processes, i.e., by means of experiments or by mathematical formulations of physical phenomena – theoretical work.

Laboratory courses give a feeling for the experimental methods of physics. The aim of the basic course is to introduce the students to elementary experimental and scientific working methods and critical quantitative thinking. This includes setting-up and conducting an experiment (measurement techniques and methods), documentation, evaluation (error calculations), discussion of the findings and scientific conclusions and finally presentation of the written report.

The basic course intentionally places the **scientific method** in the foreground. The physical questions presented in the course have long been answered, and the experiments are to be understood as providing classical examples for methods and techniques which recur in current research. Yet physics is always behind the work and does not differentiate between simple and difficult. It is the physicist, whether “professional” or in training, who asks the questions and thus determines the standard.

The laboratory course allows the student to tackle the work in an individual way so that the learning process is strongly self determined. Elementary and important prerequisites are curiosity and the ambition to understand.

Error Calculations

A fundamental phenomena of experimental work is the fact that the evaluation of natural processes is never absolute and all results must be considered as approximate. As a consequence, the empirical experimental data must be handled statistically in the form of error calculations.

An important aspect of the laboratory course is to introduce the student to the basic methods of error calculations. The first steps and basic exercises in error calculations are found in Annex I of this script (under the heading „ERROR CALCULATIONS”). (Practical exercises in error calculations will be given out before the laboratory course begins and must be handed in at the date of the first experiment). Learning the skills of error calculations is then the aim of the subsequent experimental work.

Two Semester basic laboratory course for students of Physics, Geophysics, Meteorology and for Teacher Candidates with physics as first or second subject.

Topics and Experiments

The topics of the laboratory course are coordinated with the contents of the lecture course. The experiments range from simple to demanding.

In some cases, due to organizational problems (especially in vacation courses), the topics handled in the experiments have not yet been discussed in the lectures. This requires intensive self-preparation by the student.

Preparation

Successful experimental work requires good physical- and mathematical preparation using text books and the experimental script. The laboratory course has the specific aim of deepening ones knowledge of physical processes and must be seen as complementary to the material handled in lectures and work done in tutorial exercises.

Report

The written reports serve not only as proof of experimental work but also as an exercise in the method of scientific writing. Contents and form must be such that the interested reader is introduced to the topic and the questions to be answered in an efficient and concise way and is able to follow and understand the work and conclusions. This aspect must be kept in mind and it should not be limited to a mere presentation of measured data and calculations.

Rules of the Laboratory Course

Laboratory Report Book

Laboratory regulations require that all experimental work from description to data recording and evaluation be presented in bound exercise books. Please bring suitable books (DINA4-chequered, no ring bound books) to the course. You should buy 2 – 3 books. Work done on loose or tacked paper leads to uncertainty as to its origin or loss of pages i.e. data.

Additional pages (e.g. graph paper) must be glued to a thin strip of the inside edge of a book page so that both sides of the additional page can be used. Attaching pages with paper clips is not permitted.

Graph Paper

Graphs must be drawn on graph paper (mm paper, log-paper; available in the laboratory).

Written Preparation

A written introduction to the topic and experimental task (as part of the report) must be presented before beginning the experiment. This must be prepared by each student. Since, as a rule, one of the report books of a pair of students is in the hands of the tutor for correction, the affected student must write the introduction on loose paper and later glue it into his/her report book.

The students must be able prove that they have prepared the work through discussions with the tutor.

Insufficiently prepared students will not be permitted to take part in the experiment. The experiment is noted as failed and must be repeated at a later date. If a student is rejected because of insufficient preparation, a colloquium can be set up by the head of the course to test the student. (The rules stipulate that no more than 2 failures are allowed).

Times of the Laboratory Course

The courses begin punctually at 9.15 or 14.15 h.

$3^{3/4}$ hours (9.15-13 h and 14.15-18 h respectively) are set for the work. After the experiment is completed, the remaining time is used to evaluate important parts of the data under the direction of the tutor (e.g., graphical presentations).

Structure and Form of the Report

The report is structured in two sections: Experimental documentation (measurement protocol) and the presentation (basic theory, evaluation, conclusion and discussion). The form is such that an interested reader can follow and understand the contents, results and conclusions (and allows the tutor to make corrections in a reasonable time).

The measurement protocol must be hand written and checked by the tutor for completeness and correctness. Thereafter the tutor gives an attestation. Measurement protocols without attestation will not be recognized.

Handing Over the Report

The reports should be started during the respective experiment and must be handed over at the date of the next experiment.

Failure to hand over the report punctually leads to exclusion from the next experiment.

Missing- and Failed Experiments

If a student misses or is expelled from an experiment then his/her partner must complete the experiment alone.

The excluded partner must repeat the experiment on his/her own at a later date. (The date is set by the head of the laboratory course).

Working in Partnership

Normally students work in pairs, so that each is dependent on other. Work in conjunction with your partner and discuss each experiment so that no problems occur in completing the report and the handing out of attestations.

Attestations; Handing Out the Course Certificates

The handing out of the course certificates only t - 3 -akes place after presenting the complete attestations. Attestations can only be given by the responsible tutor.

Point System

Each experiment is graded according to a point system. At the end of the course, the summed points serve to measure the total performance according to the rules of the ECTS (*European Credit Transfer System)*.

The grading is given in **% of the maximum number of points.**

[100% – 81%]	= A (very good)
[80% – 61%]	= B (good)
[60% – 41%]	= C (satisfactory)
[40% – 27%]	= D (sufficient)
[– 27%]	= E (fail)

Each experiment is individually graded, whereby a maximum of 5 points can be given. The performance points for each experiment corresponds to the ETCS grades.

5 - 4.3 points	= A (very good)
4 - 3.3 points	= B (good)
3 - 2.3 points	= C (satisfactory)
2 - 1.0 points	= D (satisfactory)
< 1.0 point	= E (sufficient)

(successful completion of an experiment requires, as a minimum, a grade of 1 point).
The assessment of the work done is based on the following categories:

A: Basic knowledge and understanding of the physics involved, preparing for the experiment.

B: Experimental ability (practical and methodical work and evaluation).

C: Scientific discussion and report (evaluating the experiment and the results, written report).

The points are noted on the group cards, report book, attestation certificates and the file cards by the tutor.

CONTENTS GP II

General Information

Experiments

Annex

REPORT GPII

The report serves as an exercise in scientific writing and presentation. It should, on the one hand, be complete and on the other concise and efficient. As an orientation, refer to the model report below.

The report consists of a *measurement protocol* and *elaboration*:

- The measurement protocol is a documentation of the experimental procedure.

  It must contain all information with respect to experimental set-up, data and observations from which one can completely understand and evaluate the experiment even after the equipment is dismantled.

- Elaboration refers to presentation and communication.

  It contains a short presentation of the basic physics involved and the question posed, evaluation, summary and critical discussion of the results and the scientific conclusions.

One of the most important aspects of a written report is its organization, i.e., how it is structured. The following describes a standard structure obligatory for the laboratory reports.

Measurement Protocol

The measurement protocol is structured as follows:

Title (Experimental Topic)

Name; Date

- Names of the students carrying out the experiment and of the tutor; date the experiment was done.

Experimental Set-Up and Equipment

- Drawing of the set-up; list of the equipment used and equipment data.

Measured Values

- Values with dimensions and units, error limits. Commentary on the error estimates. Data in the form of tables.
- Other Observations.

Elaboration

The elaboration must also be handwritten in the report book (machine written sections or formulae are glued onto the pages of the report). The elaboration is structured as follows:

Title

- (Experimental topic; name of the authors and the tutor; date of the elaboration)

Basic Physics

- A concise presentation of the basic physics with respect to the topic and the questions involved, the measurement method and the equations (copying directly from the literature is not allowed).
- The presentation must give a short but complete overview of the essential aspects of the physical quantities studied and the laws governing them. It is not required to go into details as found in text books.
- A description of the practical experimental methods is out of place here.

Evaluation

- A presentation of the evaluation in graphical form (on graph paper glued onto the appropriate page of the text), evaluated parameters, intermediate results, final results and error limits. Error discussion.
- The derivation of the results must be simple to understand and check (no scribbled notes).

Summary and Discussion of the Results

Concise Presentation:

- What was measured and how the measurements were made?

(1) **MODEL REPORT** GPII

(2) SPRING PENDULUM

(3) Albert Ach, Paula Puh

(4) Physical Basis

(5) **With an ideal spring, the restoring force is proportional to the displacement (Hook's Law):**

(6)

(A) $F = -D\,x$

The proportionality factor D is called the spring constant. This law is examined in exercise 1.

With (A) and using Newton's Law of motion, we have, where m is the total mass displaced:

(B) $-D\,x = m\,\ddot{x}$

A solution is:

(C) $x(t) = x_0 \cos(\omega t + \Phi)$

where x_0 is the amplitude, ω the frequency and Φ is a phase constant. Substituting (C) in (B) we have for the frequency:

(D) $\omega = \sqrt{\frac{D}{m}}$ and $T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{m}{D}}$

The relationship (D) for the period T is examined in exercise 2.

Equation (B) assumes that the total mass experiences the same acceleration. This is not true for the spring itself. At the attachment point, the amplitude and the acceleration are zero. At the free end, they have the values of the attached mass.

(1) The adjacent model report serves as an example for the form and presentation of scientific writing required for the basic laboratory course.

(2) The physical groundwork must be prepared and worked out before the experiment begins. (If the report book is not available, the work must be hand written on loose pages and latter glued in the report book).

(3) Each report begins on a new page commencing with the title of the experiment.

(4) Headings must be used to clearly structure the report.

(5) The presentation of the physical groundwork gives a short and concise introduction to the topic and the questions involved:

Which phenomena or principles are to be studied?

(6) Which measurement methods are used?

The presentation must refer to the subject matter in a short and precise way. Long textbook-type discourses and mathematical derivations of formulae for elementary facts are not required.

The presentation must be independently written. Literature references alone or the word for word copying of text is not permitted.

(7) The mass of the spring is accounted for by an effective mass at the free end which experiences the same acceleration and thus possesses the same kinetic energy as the spring itself.

The velocity at the spring is linear:

(E) $$v(x) = v_0 \frac{x}{x_0}$$

The mass distribution along the spring is constant and for a spring element dx we have:

(F) $$dm = m_F \frac{dx}{x_0}$$

Therefore the total kinetic energy is given by:

(8) $$E_{kin} = \frac{1}{2}\int_0^{x_0} v^2\, dm = \frac{1}{2}\frac{v_0^2}{x_0^2}\frac{m_F}{x_0}\int_0^{x_0} x^2\, dx = \frac{1}{2}\frac{v_0^2}{x_0^2}\frac{m_F}{x_0}\frac{1}{3}x_0^3$$

(G) $$= \frac{1}{3}\left[\frac{1}{2} m_F\, v_0^2\right]$$

i.e., the mass of the spring is taken as a third of the original mass.

Exercises

1. Calculate the spring constant by measuring the displacement.
2. Calculate the spring constant by measuring the period of the spring pendulum.

(7) Special facts and formulas must be explained or derived respectively.

(8) Repeating the purpose of the exercises serves as an orientation and helps to make clear the aim of the experiment.

Measurment Protocol

(9) Paula Puh, Albert Ach;
(10) Tutor: Peter Pi;
(11) 3.3.1981; Begin 10.15 am, End 12.20 pm.

(12) Equipment

(13) Stand with mirror scale (300 mm; scale divisions 1 mm).
Spring with marker and dish (Apparatus 3).
Weights(5/10/20/20´/50 g).
Stop watch (accuracy 0.1 s).
Balance "Sartorius"; (accuracy 0.05 g).

Weights

(14) ~~m_5 = 4.99 g~~ (all mass errors with 0.05 g precision)
~~m_{10} = 9.92 g~~
~~m_{20} = 19.92 g~~

(15) (Measurements discarded because of zero-point readjustment).

(16) m_5 = 5.00 g
m_{10} = 9.90 g
m_{20} = 19.90 g
$m_{20'}$ = 19.95 g
m_{50} = 49.90 g

Mass of spring m_F = 15.15 g

(17) Mass of marker and dish ms = 8.50 g

Measurement of the period (exercise 2)

Amplitude approx. 30 mm.
The period of the unloaded spring could not be measured since it did not oscillate in a regular manner.
The times were measured at the point were the displacement reverses.
Measurements at intervals of 10 T were made to reduce reaction errors.

(9)

(10) The names of the authors and tutor are important in order to know who the report belongs to and who is responsible.

(11) The date is standard information. The time can be important for subsequent discussions on disturbing influences (temperature changes, mains voltage fluctuations, ...).

(12) For the reconstruction of the experiment and the interpretation of the data (e.g., error information) a listing of all the equipment with their important nominal data must be presented (type, manufacturer; error specifications).

(13) All equipment specifications must be noted as given (measuring range, sensitivity coefficients, scale divisions, error information, ...).

(14) Information as to where the experiment was conducted and which devices or probes were used is important for later reconstruction and comparison of results.

(15) Discarded values must be recognizable (e.g. by crossing out), but readable. Do not rub out or otherwise destroy data.

(16) Zero's are also numbers; e.g. do not write down 5 g for the measurement, but the correct value of 5.00 g. The number of digits in a value contains implicit Information on the accuracy and resolution of the value.

(17) A sketch of the experimental setup is descriptive and helps to understand the connection between the equipment and quantities to be measured. In electrical experiments this is a circuit diagram, in optical experiments the ray path with the position of the optical components as an essential prerequisite for the physical understanding of the measured data.

Write down all considerations and sundry information with respect to the measurements.

(18) Displacement of spring under load conditions (exercise 1)

	Weights	Pos. Marker	M / g	x / mm
(19)	none	2.5		0
	5	17.0	5.00(5)	14.5
(20)	10	31.0	9.90(5)	28.5
	10+5	45.5	14.90(7)	43.0
	20	59.5	19.90(5)	57.0
	20+5	74.0	24.90(7)	71.5
	20+10	88.3	29.80(7)	85.5
	20+10+5	102.5	34.80(9)	100.0
	20+20'	116.0	39.85(7)	114.5
	20+20'+5	131.5	44.85(9)	129.0
	50	145.5	49.90(5)	143.0
(21)	none	2.5		

Period of spring pendulum (exercise 2)

	Weights	10 T / s	M / g	T^2 / s^2
(22)	10	5.4	9.90(5)	0.29(4)
	20	6.2	19.90(5)	0.38(4)
(23)	20+10	7.0	29.80(7)	0.49
(24)	20+20'	7.9	39.85(7)	0.62(5)
	50	8.7	49.90(5)	0.76
	50+10	9.0	59.80(7)	0.81
	50+20	9.7	69.80(7)	0.94(6)

Evaluation

Displacement as a function of load

See figure on the next page: displacement x vs. weight m. The measurement gave the expected straight-line curve with gradient:

$$D \cdot g = (0.345 \pm 0.003)\ \text{kg m}^{-1}$$

(18) Each table must have a heading in order to see which measurements are involved.

(19) Do not confuse the units of length mm and cm. Many scale divisions (straight edges, rules, callipers) are calibrated in cm. The scale in exercise 1 has a mm-division.

(20) A scale can be read to an accuracy of better than one scale division (div.) by estimating between two divisions (estimation position; here the position after the decimal point). The error is found from the reading conditions; here the upper limit is taken as the whole scale interval between two divisions (± 0.5 div).

(21) The last measurement was made as a control of the initial value (zero point).

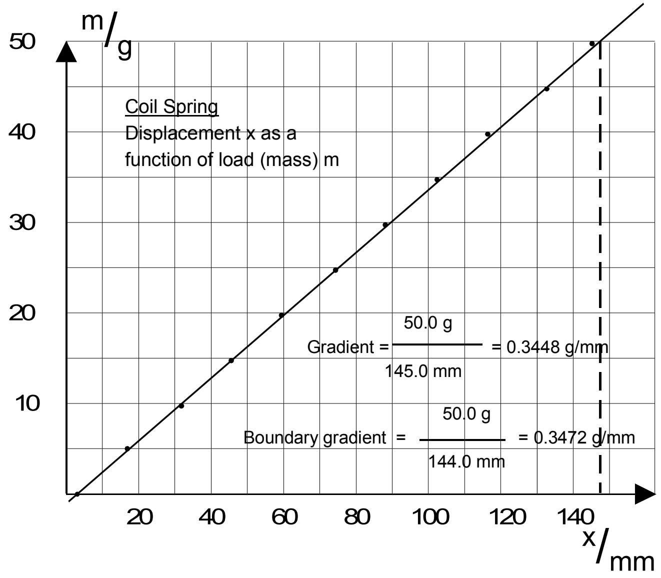
(22) For a better overview, integrate calculated values in the table.

(23) Note all measurements as they are made; hence do not write "15 g" or only the calculated value 14.90 g but which weight was used.

(24) Only cite error calculations and error values without further comment when they are formally calculated according to the error propagation laws and include all initial errors.

(25) To exercise 1: Displacement of a spring under load

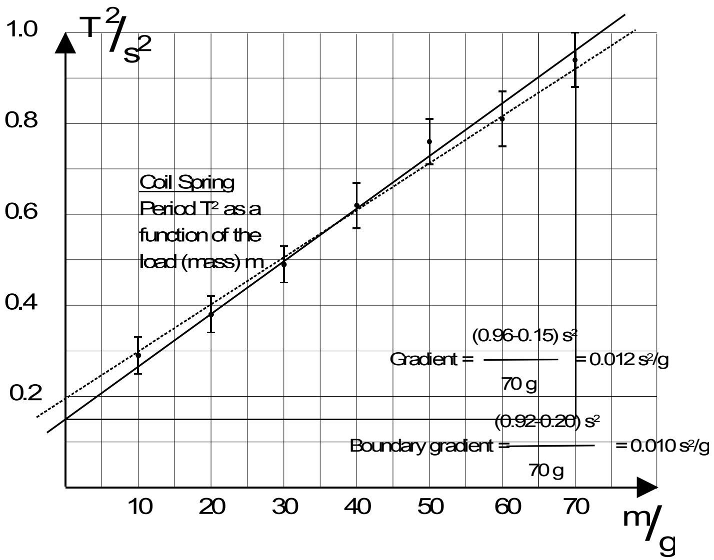
(26)
(27)
(28)
(29)
(30)
(31)
(32)



(33)

To exercise 2: Period of a Spring

(34)
(35)
(36)
(37)



(25) See annex II *GRAFISCHE DARSTELLUNGEN UND AUSWERTUNG VON FUNKTIONEN* in this script for information concerning graphical representations.

(26) Graphical representations are to be drawn on original graph paper (mm paper, log-paper; obtained in the lab). (The adjacent graphs are shown only as examples). Graphs drawn for the lab courses are to be glued in the report protocol.

(27)

(28) For reasons of accuracy, the graphs shall not be drawn too small; The standard format in the lab course is DIN A4.

(29) For optimal use of the graph paper and for control purposes ensure that the scales selected for the axis are simple and do not contain odd- or
(30) complicated divisions (e.g., 1 unit per 30 mm or similar).

Each graph must have a heading (what is represented under which con-
(31) ditions).

The axes must be completely labelled (scale, quantities, units).

(32)

All measured points must be shown. Error bars are sufficient for some representative values especially when they are constant.

(33) If the scatter of data is very small, then one must select a more sensitive
(34) representation or one must take into consideration a numerical evaluation method. In border-line cases the accuracy of the resolution and the individual values for error estimation must be taken into account.

(35) If variables are transformed, then this must be taken into consideration when labelling the axes.

(36) It is sufficient to only consider one of the two possible boundary lines for error estimation. The errors results from the differences between the
(37) lines of best fit and the boundary lines.

Draw the triangles used to calculate the gradients. They should be selected as large as possible (axes intercept points), to minimize drawing and reading errors.

The calculated gradients should be presented in the graph as a check.

When writing down the gradients note that in general these are dimensional quantities and their units must be given.

(38) The errors Δm are negligible, the errors Δx lie at the limit of representation (point diameter).To estimate a boundary gradient an error of 1 mm in displacement was assumed taking into account drawing accuracy and scatter.

(39) With g = 9.8128(1) m/s² (value taken from script, error negligible) we find for the spring constant:

(40)

$$D = (3.383 \pm 0.024)\ \mathrm{N\ m^{-1}}$$

Exercise 2: Displacement of a spring under load

With the mass of dish and marker m_S, the effective mass of the spring according to (G) and the variable hanging mass m we then have according to (D):

$$T^2 = \frac{4\pi^2}{D}\left(m_S + \frac{1}{3}m_F\right) + \frac{4\pi^2}{D}m$$

The transformed representation of T^2 against m (see previous page) resulted in the expected straight line with the gradient:

$$\frac{4\pi^2}{D} = (12 \pm 2)\ s^2\ kg^{-1}$$

Thus, the spring constant D is:

$$D = (3.29 \pm 0.55)\ \mathrm{N\ m^{-1}}$$

The axis intercept is

$$\frac{4\pi^2}{D}\left(m_S + \frac{1}{3}m_F\right) = (0.15 \pm 0.05)\,s^2$$

Therefore, with the above value for the spring constant, the effective mass of the spring is:

$$m_F = (13 \pm 13)\ g$$

(38) The error calculation (error propagation, error estimation of parameters) is an integral part of the evaluation. Error values and comments concerning the error calculations belong directly to the results.

(39) Writing down formal error calculations (equations for error propagation) is not required. However, error values must always be explained when individual or local aspects were taken into consideration in the estimation of errors.

(40) Error intermediate values noted during the evaluation must be written down as a two-digit number.

(41) Whereby, we have taken the D-value from exercise 1 because of the better accuracy. The large error results partly from the poor quality of the measured data (error of the axis intercept 34 %), whereby the increase to 100% results from forming the difference in the evaluation the relationship.

(42) In estimating the error, one must essentially take into consideration that the gradient and the axis intercept are correlated. This, however, was left out of consideration here since the error in D does not contribute and just like the mass of the scale was neglected.

(43)
(44)
(45)
(46)
(47)
(48)
(49)

Summary and Discussion

The determination of the spring constant of a coil spring by measuring the displacement as a function of load and by measuring the period gave comparable values. For the determination of the spring constant we used the values from the displacement exercise because of their better values:

$$D = (3.38 \pm 0.03)\ \mathrm{N\ m^{-1}}$$

The accuracy of the determination from the period could have been increased by increasing the number of periods.

The qualitative and quantitative agreement of the measurements among each other and with the expected characteristics demonstrates the validity of Hooke's Law (A) and the law of motion (C) with (D).

In addition, from the measurement of the period as a function of the added weights one was able to determine the contribution of the mass of the spring itself. The result for the spring mass m_F of (13 ± 13) g is formally (still) in agreement with the value of (15.15 ± 0.05) g determined by weighing and at least does not contradict the approach of an effective spring mass of 1/3 m_F. However, the large error must be considered as unsatisfactory since at the other limit it allows the conclusion that the spring itself does not contribute to the mass.

(41) In special cases errors must be discussed. Which errors make the largest contribution? What influences error propagation?

(42) With correlated quantities the maximum error must at least be calculated. (It is better to investigate the correlation).

(43) The summary and discussion must present the subject, aim and results of the experiment in a clear and essential way: What was investigated? How or according to which methods were the measurements made? Which results were found? How are the results to be scientifically assessed?

(44)

(45) The results are to be compared qualitatively and quantitatively with one another or with literature values.

Values are in agreement when the error intervals overlap.
(46) Values are compatible when the threefold error intervals overlap. Values are considered significantly different when the threefold error intervals no longer overlap.

(47) A weighted mean value must be given for results with different errors. When one has results with very different errors, the final result is taken as that which is the most accurate.

(48) Results are given with the absolute error as the basis for comparison. The relative error is a measure of the accuracy.

(49)

Errors of final results are rounded off and given as a single-digit.

Errors must also be a part of the discussion. How can the accuracy be essentially improved?

STANDARDLEHRBÜCHER GP

Die folgenden Lehrbücher werden verbreitet zur Vermittlung physikalischen Grundwissens herangezogen, wie es für das Physikstudium und die Vorbereitung der Praktikumsarbeit erforderlich ist. Eine Reihe von Lehrbüchern wurden in verschiedenen Auflagen bzw. Jahren herausgegeben, so daß auf eine Angabe des Erscheinungsjahrs verzichtet wurde. Alle Bücher sind in der Lehrbuchsammlung der Fachbereichsbibliothek vorhanden.

obligatorische Literatur

[1]: Gerthsen·Kneser·Vogel; Physik; Springer-Verlag

[2]: Bergmann-Schaefer Band 1 (11. Auflage)

[3]: Bergmann-Schaefer Band 2 (8. Auflage)

[4]: Bergmann-Schaefer Band 3 (9. Auflage)

[5]: Eichler Kronfeld Sahm

Das neue Physiklaische Praktikum

Zusatzliteratur

Alonso·Finn;
Physik;
Addison-Wesley bzw. Inter European Editions

Atkins;
Physik;
de Gruyter

Kittel·Knight·Rudermann;
Berkeley Physik Kurs
(1: Mechanik, 2: Elektrizität und Magnetismus,
3: Schwingungen und Wellen, 4: Quantenphysik,
5: Statistische Physik);
Vieweg & Sohn

Demtröder;
Experimentalphysik 1-4;
Springer-Verlag

Dransfeld·Kalvius·Kienle·Lucher·Vonach;
Physik (I: Mechanik, II: Elektrodynamik,
IV: Atome-Moleküle-Wärme);
Oldenbourg

Feynman·Leighton·Sands;
Vorlesungen über Physik (I: Mechanik-Strahlung-Wärme, II: Elektromagnetismus und Struktur der Materie);
Oldenbourg

Hänsel·Neumann;
Physik 1-3;
Spektrum Akademische Verlagsanstalt

Kohlrausch;
Praktische Physik (3: Tafeln);
Teubner

Tipler;
Physik;
Spektrum Akademische Verlagsanstalt

Martienssen;
Einführung in die Physik (I: Mechanik,
II: Elektrodynamik, III: Thermodynamik,
IV: Schwingungen-Wellen-Quanten);
Akademische Verlagsgesellschaft

Otten;
Repititorium der Experimentalphysik;
Springer-Verlag

PSSC;
Vieweg

Pohl,
Einführung in die Physik (1: Mechanik-Akustik-Wärme,
2: Elektrizitätslehre, 3: Optik-Atomphysik);
Springer-Verlag

Zinth·Körner;
Physik I-III;
Oldenbourg

Westphal;
Kleines Lehrbuch der Physik;
Springer-Verlag

Optik

Born·Wolf;
Principles of Optics;
Mac Millan

Fowles;
Introduction to Modern Optics;
Dover Publication Inc.

Atom- und Quantenphysik

Eisberg·Resnick;
Quatum Physics of Atoms, Moleculs, Solids, Nuclei and Particles;
Wiley & Sons

Finkelnburg;
Atomphysik;
Springer-Verlag

Haken·Wolf;
Atom- und Quantenphysik;
Springer-Verlag

Beiser;
Atome, Moleküle, Festkörper;
Vieweg & Sohn

Fehlerrechnung

Taylor;
Fehleranalyse;
VCH Verlagsgesellschaft

MICROSCOPE	GP II

Key Words

Geometrical Optics; Imaging with Lenses. Resolution and diffraction limit; Abbe's Theory, numerical aperture.

Aim of the Experiment

Understanding the working principles of a microscope and handling optical components and instruments.

Literature

Standard literature (see list of standard text books).

Exercises

1. Determining the focal length of a lens using the *Bessel Method*.
2. Constructing the ray path of a microscope. Determining the magnification for three different tube lengths and comparing the results with the theoretical expectations.
3. Calibrating an ocular micrometer (measurement ocular). Determining the grating constant and the thickness of the wires of a wire grating (cross grating).
4. Verification of *Abbe's Theory*. Observing the resolution limit of the microscope using the wire grating. Determining the numerical aperture for this limiting case and comparing the expected smallest resolvable point separation with the measured grating constant.
5. Calculation exercise: Specifying the smallest resolvable point separation for the strongest objective (numerical aperture 1.4 with immersion fluid) and hence the achievable meaningful limit of magnification of a microscope.

Physical Principles

Imaging through Lenses

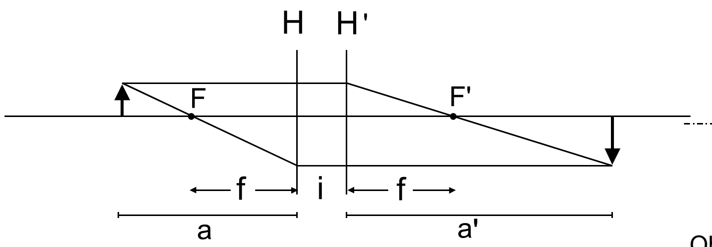
The image of an object through a lens or a centered lens system for rays of light in the neighborhood of the symmetry axis (optical axis) is given by the thin lens formula:

$$\frac{1}{a} + \frac{1}{a`} = \frac{1}{f} \qquad (1)$$

with the image scale

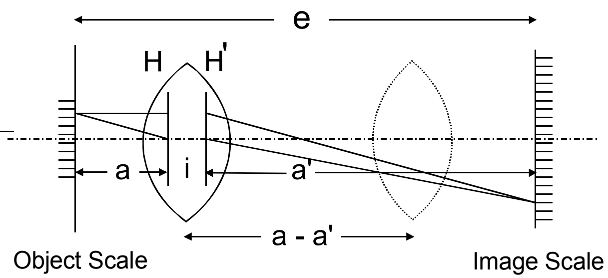
$$\beta = \frac{a}{a`}$$

The *image scale* ß specifies the linear size ratio between the real or virtual image and the object (see diagram below).



Here, *a* is the object distance, *a'* the image distance and *f* is a characteristic quantity of the system, the *focal length*. The system is also determined by both *principal points H* and *H'* on the axis. The associated *focal points F* and *F'* are at a distance of the focal length to the principal points. Similarly, the object distance and the image distance are the distances from the associated principal points. The principal points do not always lie in or close to the lens; lens systems exist with principal points several focal lengths away from the axis of the last lens.

In *thin lenses,* the principal point interval *i* is small and can be virtually neglected. In general, however, this is not the case, so that a simple measurement of the focal length according to (1) is not possible due to the unknown principal interval. In the *Bessel Method* (*Friedrich Wilhelm Bessel*; 1784-1848; German astronomer and mathematician) a lens in moved along the optical axis between a fixed object and a fixed image screen. The object and image positions are separated by a distance more than 4 times the focal length of the lens (e > 4f). Two positions of the lens are found for which the image is in focus on the screen, magnified in one case and reduced in the other. One measures the distance *e*, the difference *a-a'* (from the two lens positions) and the image scale ß so that three independent quantities are available to determine, according to (1), the two unknowns *f* and *i* (see diagram below).



Solving the thin lens formula for *a'* and *a* and substituting the ration *a'/a* by ß gives the focal length as:

(2) $$f = \frac{a - a`}{\frac{1}{\beta} - \beta}$$

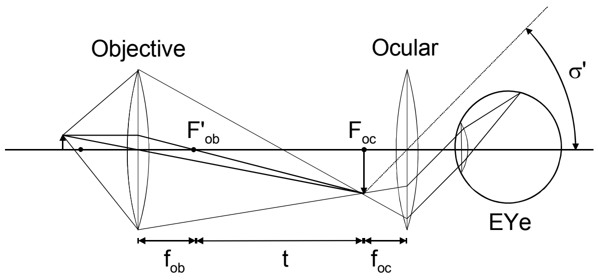
and for the principle interval i from the distance e (see diagram) we get:

(3)

$$i = e - (a + a`) = e + (a - a`)\frac{\beta + 1}{\beta - 1}$$

Microscope

A microscope is a two stage imaging system for the magnification of objects. An *objective lens produces a real image* which is viewed through an *ocular* functioning as a loupe. The distance *t* between the focal points of the objective lens and ocular is called the *optical tube length* of the microscope.



The *magnification* of an optical instrument is specified as the magnification of the *angle of vision*, which determines the apparent size of the object being observed. The magnification Γ is defined as:

(4) $$\Gamma = \frac{\tan(angle\ of\ vision\ with\ instrument)}{\tan(angle\ of\ vision\ without\ instrument)} = \frac{\tan(\sigma)}{\tan(\sigma_0)}$$

The angle of vision is referred to the so called *conventional near vision accommodation* a_o = 250 mm. The terms magnification and image scale must be distinguished.

The total magnification of a microscope is given the product of the image scale of the objective lens and the magnification of the loupe of the ocular:

(5) $$\Gamma = \beta_{Ob}\ \Gamma_{Ok}$$

The angle of vision giving an (virtual) image of an object in front of the loupe (ocular) is given by:

(6) $$\tan \sigma_{Ok} = \frac{y'}{a'} = \frac{y}{a}.$$

Hence, the magnification is:

(7) $$\Gamma_{Ok} = \frac{a_0}{a}$$

For an eye accommodated to infinity one has $a = f_{Ok}$, i.e.:

(8a) $$\Gamma_{Ok}(\infty) = \frac{a_0}{f_{Ok}}$$

In the experiment, a comparison scale at a distance a_o is observed simultaneously with the image in the microscope so that the eye is accommodated to the near vision distance ($a'=a_o$). From the thin lens formula one can calculate the object distance *a* (note sign!) and hence it follows that:

(8b) $$\Gamma_{Ok}(a_o) = \frac{a_o}{f_{Ok}} + 1$$

The loupe magnification increases with decreasing focal length, and because of the necessarily small radius of curvature results in a diminution of the lens diameter and thus a reduction in the resolving power. Furthermore, with a strong loupe, the free working distance between eye and object is uncomfortably small. Because of these disadvantages, one uses loupes with at most a 30-fold magnification.

The total magnification of the microscope, depending on the accommodation of the eye, is found from (5) and (8a or 8b) and is:

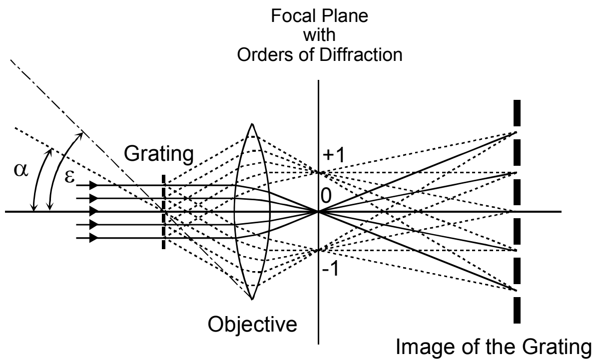
(9a) $$\Gamma_\infty = \frac{t}{f_{Ob}}\frac{a_o}{f_{Ok}}$$ or

(9b) $$\Gamma_o = \frac{t}{f_{Ob}}\left[\frac{a_o}{f_{Ok}} + 1\right]$$

Resolution of the Microscope

In the framework of geometrical optics it should be possible, according to (9a, b), to arbitrarily increase the magnification with sufficient tube length and small objective focal length. In contrast, however, one observes a limitation in the resolution of the system which can only be explained with wave optics, i.e., interference and diffraction must be taken into consideration.

The *Abbe's Theory* of Imaging (*Ernst Abbe*; 1840 – 1905; German physicist) uses a diffraction grating (with grating constant *d*) as the object which is illuminated by plane waves of light (see diagram below).



According to *Huygens' Principle* (*Christian Huygens*; 1929 – 1695; Dutch physicist, mathematician and astronomer) a pattern of diffraction slits appears in the focal plane of the lens due to interference effects. Intensity maxima occur when all partial waves interfere constructively, i.e., when the phase difference of the waves of neighboring slits are whole multiples of the wave length λ. This is fulfilled when:

$$\frac{d \sin \alpha}{\lambda} = 0, \pm 1, \pm 2, \ldots \tag{10}$$

A characteristic feature is that the diffraction angle becomes larger α the smaller the separation *d* of the grating slits.

In the image plane *Z* one observes the image of the grating (see diagram). For an image with periodic intensity distribution the partial waves of at least two neighboring orders of diffraction must be gathered by the lens, i.e., the angle for the first order of diffraction may be at most equal to the aperture angle on the incidence side of the grating. Hence, the minimum resolvable grating constant is:

$$d_{\min} = \frac{1}{\sin \varepsilon} \tag{11}$$

The resolvable separation becomes smaller when the space between the object and the objective lens is filled by a medium of refractive index *n* (immersion fluid) thereby reducing the wavelength by a factor *n*. The quantity

$$A = n \sin \varepsilon \tag{12}$$

which together with the wavelength determines the resolution of the microscope is called the *numerical aperture* of the objective lens.

The diagram with the limiting case discussed above corresponds to a grating with a $\sin^2$-type of transmission. The only information –aside from the orientation- is the grating constant. In order to determine, for example, the ration of the slit width of the grating to the grating constant, further orders of diffraction must contribute to the image. The considerations made for the slit grating are also of importance for arbitrarily structured objects, which can be formally described by a superposition of gratings with different grating constants (Fourier decomposition; see also experiment *DIFFRACTION AND INTERFERENCE*).

Apparatus and Equipment

Two 40 mm lenses (ls objective and ocular). Two illuminated 1 mm scales as object and comparison scale.
1/10 mm on a glass slide (ocular micrometer). Half-reflecting mirror to image the comparison scale. Wire grating. Various apertures.

Experiment and Evaluation

The experimental setup was designed according to didactical aspects, in order to provide a clear and visual arrangement with open experimental possibilities. As a consequence, however, the optical images are of low quality and formed under difficult measurement conditions so that the usual measurement accuracy in the lab course is not achievable.

Exercises 1-3 can be performed under normal lighting conditions. The room must be darkened only for exercise 4 (observing the limit of resolution).

To Exercise 1

The object scale is an illuminated mm scale and the image scale is a 1/10-mm scale on a glass slide. The image scale is mounted on the back side of the glass slide so that it faces the objective side when it is correctly imaged in the field of view of the loupe (of the ocular). A convenient distance for both scales is *e* = 250 mm. An adjustable iris aperture is placed directly in front of the lens to be measured in order to adjust the brightness of the image and to reduce aberrations by limiting the beam of light close to the optical axis. The image and the image scale are observed with the second lens functioning as a loupe. The criterion for an acceptable image is freedom of *parallax.*

When making error calculations take into account that the image scale ß according to (3) goes into the evaluation equation a number of times. Either one uses *Gauss's law of error propagation* in the general form (partial derivatives) or another meaningful method of error estimation.

To Exercise 2

Using both lenses a microscope ray path is constructed with tube lengths of *t* = 150, 200 and 300 mm. To reduce aberration and to adjust the brightness, a circular aperture is again placed directly behind the objective lens (seen from the object).

The magnification is determined by placing a half-reflecting mirror in front of the ocular lens (seen from the eye) at an angle of 45° and which is used to image a second mm-scale at a distance of a_0 = 250 mm so that this image and the magnified image can be simultaneously observed. The magnification is found by comparing the largest possible intervals and taking into account the thickness of the scale divisions.

To Exercise 3

For this experiment the tube length must be 300 mm to obtain a sufficient magnification to measure the wire mesh. A 1/10 mm scale is placed in the real intermediate image plane of the microscope and calibrated by comparison with the object scale (ocular micrometer). Finally, one measures the grating distances and the thickness of the wires.

To Exercise 4

The aperture holder is placed 40 mm behind the objective lens so that it lies in the focal plane of the lens. It is sufficient to measure the position with a rule. One observes the wire mesh and reduces the effective opening of the microscope by selecting smaller apertures until the periodic structure of the grating vanishes. The aperture angle ε can be determined from the diameter of the aperture opening *B* and the focal length of the objective lens f:

$$\tan \varepsilon = \frac{B/2}{f} \tag{13}$$

OPTICAL SPECTROSCOPY | GP II

Key Words

Dispersion; Prisms. Diffraction and Interference; Diffraction Grating. Spectral Equipment and Spectral Analysis.

Aim of the Experiment

Phenomenological and experimental introduction into the fundamentals of optical spectroscopy as an important scientific and applied analytical tool in many areas of the natural sciences.

Literature

Standard literature (see list of standard text books).

Exercises

Performing experiments either with the prism spectrometer or grating spectrometer:

Prism Spectrometer

1. Setting up and adjusting the spectrometer (illumination, collimator, telescope).
2. Measuring the angle of the refracting edge of a prism.
3. Recording the spectrum of a mercury lamp to calibrate the spectrometer.
4. Performing one of the following experiments.
5. Plotting the dispersion curve $n(\lambda)$ and determining the differential dispersion $dn/d\lambda$ for the 577/579 nm line of mercury.
6. Determining the resolving power of the prism and comparing the result with the theoretical expectation.
7. Qualitative observation and discussion of the diffraction spectrum of a grating.

Grating Spectrometer

2. Recording the spectrum of a mercury lamp in the first and second order and determining the grating constant.
3. Performing one of the following experiments.
4. Determining the resolving power of the grating in the first and second order and comparing the result with the theoretical expectations.
5. Qualitative observation and discussion of the dispersion spectrum of a prism.

Spectroscopic Tasks

Spectroscopic analysis of an unknown lamp and determining its gas content.

Physical Principles

Prism

The transmission of light through transparent media represents a resonance phenomena with a frequency- or wavelength dependence of the refractive index *n* known as *dispersion*. Consequently, light of different wavelengths is refracted differently at a boundary surface and thus resolved into its spectral parts. The total deflection angle when light passes through both boundary surfaces of a prism depends not only on the refractive index but also on the direction of the incident light. Simple conditions result for the special case when a light ray passes through the prism parallel to the base and is thus symmetric with respect to entrance and emerging angles. In this case, the total deflection angle is minimal (*minimal deflection*). The entrance and emerging angle at one of the boundary surfaces follows from the geometrical ratios (see diagram below):

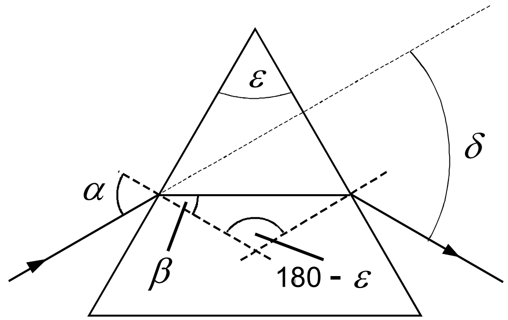


Fig. 1: Refraction of Light at a Prism

(1) $$\beta = \frac{\varepsilon}{2}$$ (inner triangle) and

$$\delta = 2\,(\alpha - \beta) \quad \text{or} \quad \alpha = \frac{\delta + \varepsilon}{2} .$$

From the law of refraction we then have

$$\frac{\sin\alpha}{\sin\beta} = \frac{c_0}{c_P} = \frac{n_P}{n_0}$$ or

(2) $$n_P = n_0 \frac{\sin\frac{\delta+\varepsilon}{2}}{\sin\frac{\varepsilon}{2}}$$

where ε is the angle between the refracting surfaces and n_p and n_o are the refractive indices of the prism and the surrounding medium respectively (for air n_o = 1.0003).

Prisms find application is spectroscopy and light filtering. The dispersion power and the refractive index are independent of one another. For example, the refractive index of flint glass is only slightly higher than that of crown glass, however, the dispersion power is almost twice as high. The different behavior of various types of glass allows the construction of prisms with strong deflection properties but do not disperse (deflection prism, *achromatic prism*) or prism with strong dispersion properties
but do not deflect (*direct vision prism*).

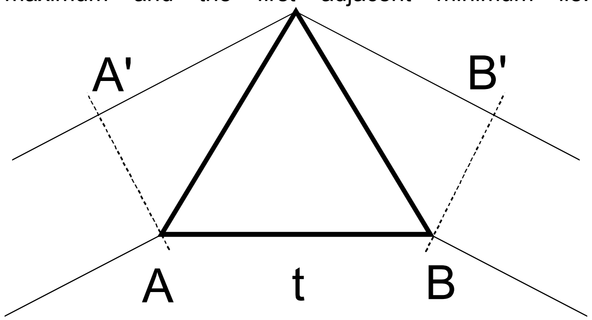
Resolution Criterion

The determination of the resolving power of spectral equipment requires a conventional agreement as to when two spectral lines can be considered as separated. In general, the most practical criterion is the *Sparrow criterion*, whereby two lines are seen as separated when they possess a relative minimum. Quantitative more accurate is the *Rayleigh criterion* (*John William Strutt*, since 1873 *Baron Rayleigh*; 1842-1919; Engl. physicist), stating that lines can then be considered as separated the diffraction maximum of the one line coincides with the first diffraction minimum of the other (see figure below). The intensity in the minimum of this double line then falls to the value of $8/\pi^2$ of the maximum.

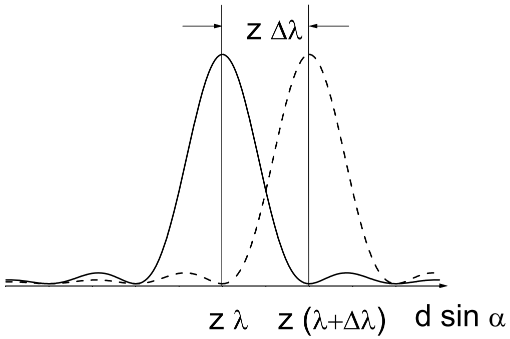
Resolving Power of a Prism

The finite resolution of a prism is conditional on the diffraction since it represents a limitation for the ray path. It is derived from a consideration of the optical path length in the prism (see figure below).

A-A' and B-B' represent two wave fronts ahead and behind the prism belonging to a direction of deflection under which, corresponding to the *Rayleigh criterion* for the wavelength λ and for the wavelength $\lambda+\Delta\lambda$, the main maximum and the first adjacent minimum lie.



Resolution of the Prism



Rayleigh Criterion

For the main maximum (to λ) the rays should not exhibit a path difference, whereas the first adjacent minimum originates in the diffraction pattern (to $\lambda+\Delta\lambda$) when the rays at the edges exhibit a path difference of just one wavelength (see experiment *DIFFRACTION AND INTERFERENCE*).

For small differences in wavelength, the dependence of the refractive index on λ is approximately given by a linear relation:

(2) $n(\lambda) = n$ and $n(\lambda + \Delta\lambda) = n + \frac{dn}{d\lambda}$

Since the optical path from A' to B' is the same for both wavelengths because $n \approx 1$, a path difference of λ must arise at the base of the prism (base length *t*):

$$\left(n + \frac{dn}{d\lambda}\Delta\lambda\right)t - n\,t = \lambda \qquad (3)$$

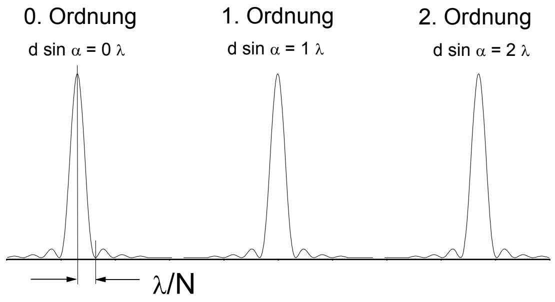
Consequently, the resolving power of the prism is determined by the base length *t* (representing the opening of the prism) and the *differential dispersion* $dn/d\lambda$.

Diffraction Grating

A grating can be simply conceived as an aperture with a periodic sequence of sharp and impermeable bounded slits. Normal gratings used in practice, e.g., made by scribing on a glass plate do not correspond to this picture. In general, one speaks of a grating when at an object a transmission (or reflection) recurs periodically at a spacing *d*, the *grating constant*.

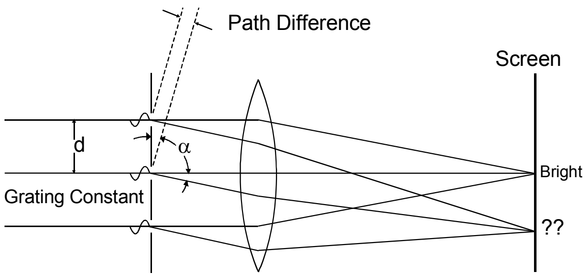
If one shines coherent light onto a grating or, as a model, monochromatic plane waves, one observes behind the grating a periodic intensity distribution explained by diffraction and interference effects. Comparatively simple relations are found in the *plane of observation* at infinity which can be realized in practice by placing a convex lens behind the grating. The resulting intensity distribution is termed the *Fraunhofer Diffraction Pattern* (*Joseph Fraunhofer*, 1787-1826, German optician and physicist).

The patterns are in the form of sharp *diffraction maxima* (main maxima), separated by wide extinction zones (see diagram below; in extinction zones in the diagram are not to scale in order to clearly show the adjacent maxima):



The main maxima can be simply derived from the condition that the path difference of rays of neighboring slits must be a whole multiple of the wavelength for *constructive interference* to occur see diagram below):

(4) $$d \sin \alpha = z\,\lambda \qquad \text{with} \qquad z = 0, 1, 2, \ldots$$



The number *z* labels the *Order* of the diffraction maxima. Since for a given grating constant, the position of the maxima, aside from the order, is dependent on the wavelength, a grating can be used to perform (absolute) wave measurements.

Resolution of a Grating

Besides the main maxima given by (4) there exists a series of adjacent maxima, whose intensity rapidly approaches zero with increasing distance from the main maxima. The position of the first adjacent minimum of order *z* is given by:

(5) $$d \sin \alpha_{\text{min}} = \left(z + \frac{1}{N} \right) \lambda$$

where *N* is the total number of contributing slits. (For even N this relation can be derived rigorously, by imagining the grating composed of two equal parts with half the number of slits and allowing pairs of slit to interfere destructively, i.e., with a path difference of $\lambda/2$).

If one sets, according to the *Rayleigh criterion,* condition (4) for the main maximum with a wavelength $\lambda+\Delta\lambda$ and for the adjacent minimum condition (5) with a wavelength λ then for the resolution of the wavelength difference or for the resolving power it follows that:

(6)

$$\Delta\lambda = \frac{\lambda}{z\,N} \qquad \text{or} \qquad \frac{\lambda}{\Delta\lambda} = z\,N$$

i.e., the resolving power increases with the growing number of slits and with increasing order number.

Hydrogen Spectrum and the Rydberg Constant

In 1885 the Swiss mathematician and physicist *Johann Jakob Balmer* (1825-1898) found, while conducting an empirical analysis of the characteristic line series of hydrogen (*Balmer-Serie*), that the wave number of the lines could be described as the difference of two terms:

(7)

$$\bar{\nu} = \frac{1}{\lambda} = R \left[\frac{1}{2^2} - \frac{1}{n^2} \right] \qquad \text{with} \qquad n = 3, 4, 5, \ldots$$

This discovery was later to become an important support for the *Bohr model of the atom* (*Niels Bohr*; Danish physicist; 1885-1962), according to which the radiation of atoms is the result of electron transitions between atomic levels. The constant *R* in the relationship is the *Rydberg constant* (*Johannes Rydberg*; 1854-1919; Swedish.physicist):

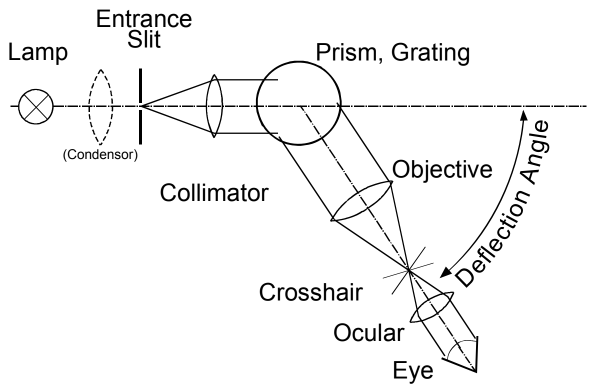
(8) $$R = \frac{2\pi^2\, m_e\, e^4}{h^3\, c}$$

Apparatus and Equipment

Goniometer assembly (angle measuring equipment) on rails, including a tiltable and rotatable table for a prism or grating.
Optical components: Entrance slit, collimating lens, objektive lens, ocular with crosshairs. adjustable crosshairs (for angular measurements on the prism), measurement slit (to determine the resolution).
Spectral lamps with power supply Hg lamp, unknown lamps.

Experiment and Evaluation

The spectrometer consists of a goniometer for angular measurements, an entrance slit with a collimator to produce parallel light, the dispersive element (prism, grating) and a telescope. Usually a condenser lens is placed between the lamp and the entrance slit. See diagram below.



Adjustment

The prerequisite for quantitative proper results in optical experiments is very careful adjustment of the optical assembly. Furthermore, this also conveys an understanding of the fundamentals of geometrical optics

The collimator lens and the objective lens are fixed at a certain height. This sets the height axis of the assembly. The other components must be aligned to this height

Illuminating equipment is not necessary because of the high light density of the lamps and the Hg lamp is placed directly behind the entrance slit. The slit represents (in one spacial direction) an approximately point shaped, *secondary light source*.

Adjustment of the collimator is done by *autocollimation*. The light from the collimator is reflected by a mirror placed on the prism table. When the collimator is correctly adjusted one sees a sharp image of the slit reflected back at the slit. The image can be slightly shifted to one side of the slit mechanism to get a better view.

The crosshairs of the ocular are focused against diffuse background lighting with the eye relaxed. The ocular is then placed in the swivel arm of the goniometer.

Finally, the objective lens is adjusted so that a sharp image of the slit is seen in the image plane of the ocular. The criterion for proper adjustment is freedom of parallax error, i.e., that the slit image and the crosshairs are not shifted against each other by a sideways movement of the eye. Causes for a poor slit image, aside from incorrect adjustment of the collimator- or objective lens, may be that the slit is opened too wide or is too narrow or the slit is not positioned vertically or is dirty. A slight over-illumination of the slit is unavoidable under the given circumstances.

The angle between the collimator- and telescope axis is read on an angle scale with 1/100 degree divisions (vernier scale). To ensure correct reading it is recommended to practice using the vernier scale before performing the measurements.

Error Estimations

In particular with optical experiments, one should make control measurements of the settings in order to check the reproducibility or observed deviations, thus allowing statements to be made on possible errors (estimated errors).

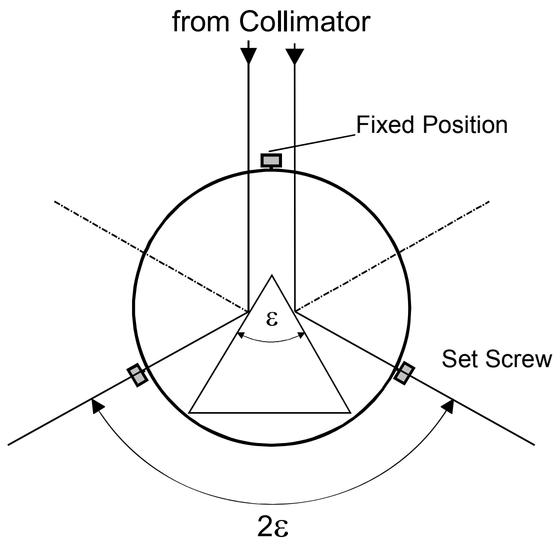
Prism Spectrometer

To Exercise 1 (Angle of the Refracting Edge; see diagram below)

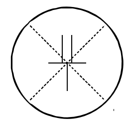
The angular measurement is performed by placing a ground glass screen with crosshairs in the ray path where the entrance slit is normally positioned (crosshairs facing the telescope), the image is focused and adjusted in height, so that both crosshairs (in the collimator and in the ocular) can be made to cover each other.

The prism table is first visually aligned to a horizontal position so that one of the three set screws points to the collimator and the other two point in the respective directions of observation.

The prism is then placed on the table so that the light reflected from both side faces can be observed with the crosshairs in the telescope.



Collimator Cross



Telescope Cross

The prism must be placed as far as possible towards the telescope because the illumination is very narrow and the side faces meet only at the tip of the prism.

Now fine alignment is made of the prism side faces perpendicular to the optical axis. This is done by successively observing both sides of the crosshairs and adjusting the respective opposite set screw to achieve the same height. The adjustments are performed until the height on both sides coincides.

Finally, the angular difference of the intercept points of both reflected crosshairs is measured, thus giving the angle of the refractive edge.

To Exercise 2 (Calibration Curve)

The prism is now positioned in the deflection position, again ensuring good illumination. The minimal deflection for the 546 nm is now set and the complete spectrum of the mercury lamp recorded. The positions of the optical components on the rail and, in particular, that of the prism must not be changed for this and subsequent measurements, otherwise the assignment between wavelength and deflection angle would be lost.

Evaluation is made by plotting wavelength against deflection angle on DIN-A4 mm paper to match the accuracy of the measurements!

To Exercise 3 (Spectroscopic Experiments)

See the following notes on the exercises.

To Exercise 4 (Dispersion Curve and Differential Dispersion)

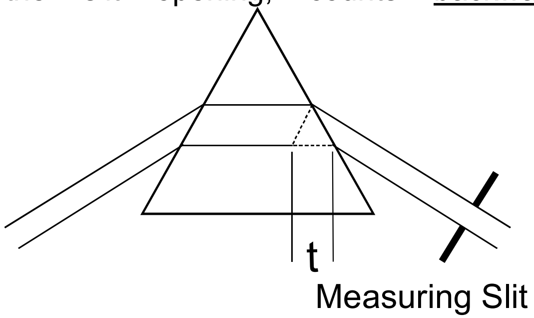
The minimal deflection is set and the deflection measured for each of the main lines of the Hg-spectrum (579, 577, 546, 492, 436 and 405 nm). From the measurements one can calculate the refractive indices and plot the dispersion curve $n(\lambda)$ employing equation (2). The differential dispersion for the 579/577 nm lines is determined by constructing a tangent to the dispersion curve at these wavelengths.

To Exercise 5 (Resolving Power)

Since the wavelengths of the lines cannot be changed, the optically effective base length *t* of the prism must be shortened. This is done by placing an additional measuring slit in the collimator ray path directly in front of the prism and closing the slit so far until one observes two lines adjusted to minimal deflection that can just be separated under the subjective resolution criterion (*Sparrow* or *Rayleigh*).

The investigation is carried out on the 579.1/577.0 nm pair of lines of mercury. The experimentally observed resolution results from the ratio of the mean value of the lines to the difference, and the theoretically expected value from equation (4).

The effective base length *t* is calculated from the set slit opening *b*, the deflection angle γ and the prism angle ε. Note that when reading the scale on the micrometer to determine the width of the slit, take into account that the zero point is arbitrarily shifted and that the scale, with respect to the slit opening, counts backwards.



To Exercise 6 (Grating Spectrum)

The prism is replaced by a grating. The characteristic differences of the spectra are to be observed and recorded and a short discussion presented in the report.

Grating Spectrometer

To Exercise 1 (Grating Constant)

The grating is placed in the ray path (see lab bench script for the orientation of the grating). The grating is carefully adjusted perpendicular to the ray path by autocollimation, i.e., observing the surface of the grating reflected back on the slit.

The grooves of the grating are asymmetrically scribed (*blazed grating*), whereby, for a certain range of wavelengths the largest intensity is available for a certain direction of deflection and a certain order. The complete observable spectrum of the Hg - lamp is to be recorded in the 1.order and in the 2.order the main lines (579, 577, 546, 492, 436 and 405 nm). Because of the high accuracy of the measurement, graphical evaluation is unsuitable in this case.

To Exercise 2 (Spectroscopic Experiments)
See the following notes on the exercises.

To Exercise 3 (Resolving Power)

Since the wavelengths of the lines cannot be changed, the effective width of the grating must be shortened to determine the limit of resolution. This is done by placing an additional measuring slit in the collimator ray path directly in front of the grating and closing the slit so far until one observes two lines that can just be separated under the subjective resolution criterion (*Sparrow* or *Rayleigh*).

The investigation is carried out on the 579.1/577.0 nm pair of lines of mercury. The experimentally observed resolution results from the ratio of the mean value of the lines to the difference, and the theoretically expected value from equation (7).

The number of grating slits can be calculated from the effective grating width governed by the limitation due to the measuring slit and from the grating constant. Note that when reading the scale on the micrometer to determine the width of the slit, take into account that the zero point is arbitrarily shifted and that the scale, with respect to the slit opening, counts backwards.

To Exercise 5 (Prism Spectrum)

The grating is replaced by a prism. The characteristic differences of the spectra are to be observed and recorded and a short discussion presented in the report.

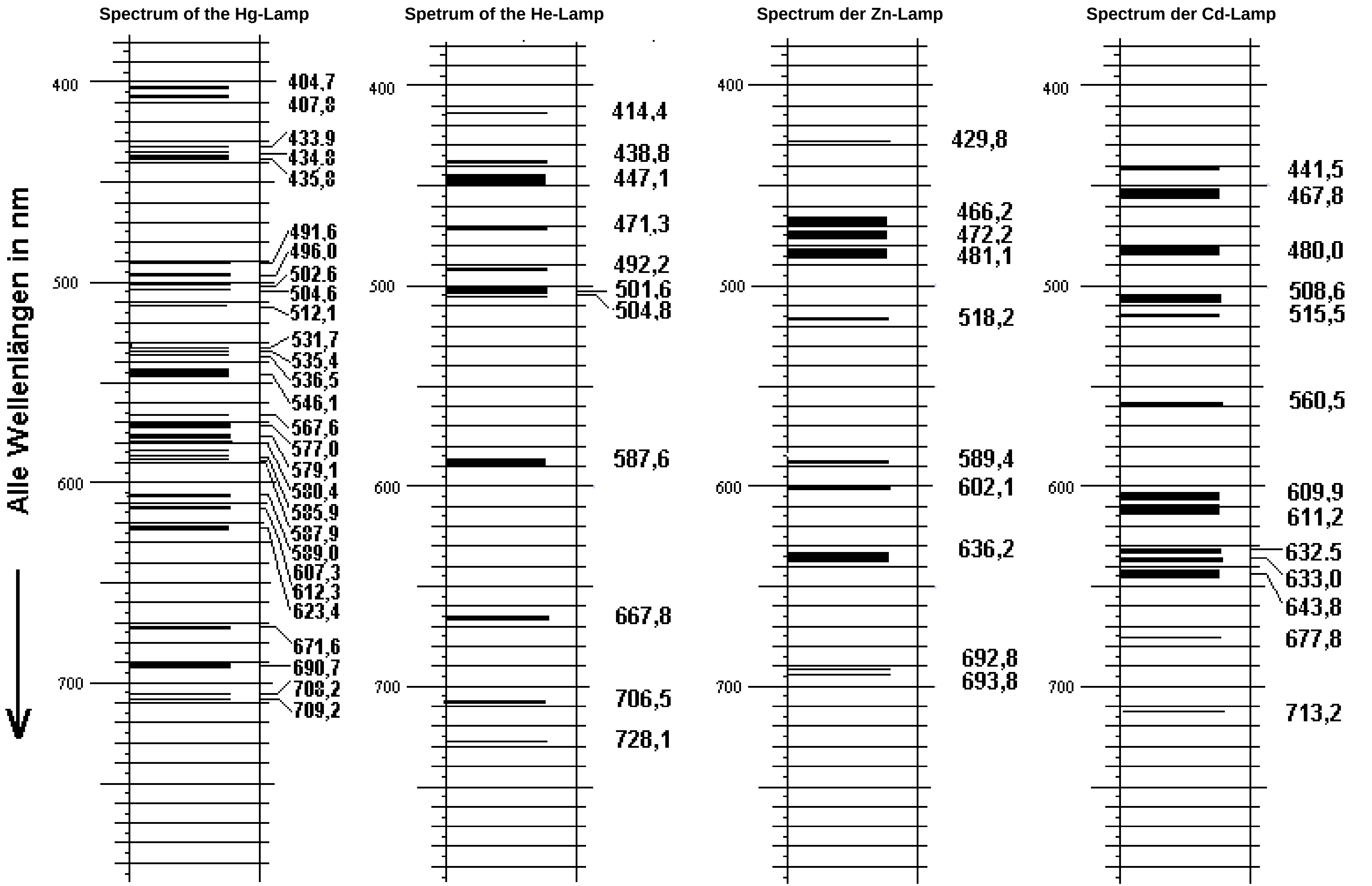
Spectroscopic Experiments

Unknown Lamps

The spectrum of one (of the three available) unknown lamps is recoded and the observed wavelengths determined from the calibration curve or the grating constant. The results are analysed using the table of selected spectral line attached to this script.

Spectral Lines

See the following page for the spectrum of the Hg-lamp and the lines of Cd, He and Zn.


Spectrum of the Hg-Lamp
Spetrum of the He-Lamp
Spectrum der Zn-Lamp
Spectrum der Cd-Lamp
Alle Wellenlängen in nm
400
500
600
700
404,7
407,8
433.9
434.8
435,8
491,6
496,0
502.6
504,6
512,1
531,7
535,4
536,5
546,1
567,6
577,0
579,1
580,4
585,9
587,9
589,0
607,3
612,3
623,4
671,6
690,7
708,2
709,2
414,4
438,8
447,1
471,3
492,2
501,6
504,8
587,6
667,8
706,5
728,1
429,8
466,2
472,2
481,1
518,2
589,4
602,1
636,2
692,8
693,8
441,5
467,8
480,0
508,6
515,5
560,5
609,9
611,2
632.5
633,0
643,8
677,8
713,2

DIFFRACTION AND INTERFERENCE | GP II

Key Words

Wave optics; *Huygens Principle*. Coherence. Diffraction and interference at slits and gratings.

Aim of the Experiment

Experimental introduction to diffraction phenomena. Wave treatment of optical images (*Abbe's Theory*) and the interconnection between diffraction patterns and the image of an object. Exemplary investigation of *image filtering*.

Literature

Standard literature (see list of standard text books).

Exercises

One can select between exercises employing a thermal spectral lamp (exercise A) or a He-Ne laser (exercise B).

Exercises using a Na-spectral lamp (exercise A)

A1. Constructing the ray path and determining the image scale of the microscopic image.

A2. Determining the width of a single slit from the image of the slit and from the *Fraunhofer* diffraction pattern. Comparison and discussion of the results.

A3. Determining the widths and separation of a double slit as in A1.

A4. Determining the spacing of a diffraction grating as in A1.

A5. Investigating the image of a grating by blocking out different orders of diffraction from the pattern (image filtering).

Exercises using the He-Ne Laser (exercise B)

B1. Recording the *Fraunhofer* diffraction pattern for three different slit widths. Comparison and discussion of the results.

B2. Recording the diffraction pattern of a double slit and determining the widths and separation.

B3. Determining the scale divisions of a metal rule from the diffraction pattern of the divisions at glancing incidence (Reflection grating).

Physical Principles

Huygens Principle

The geometrical optical treatment with a linear propagation of light fails when boundaries in the wave field or structures of the order of the wavelength of light appear in the ray path transverse to the propagation direction. *Huygens Principle* (*Christian Huygens*; 1629-1695; Dutch physicist, mathematician and astronomer) is a useful aid to completely describe diffraction phenomena and resolution in optical imaging occurring under these conditions. It states that all points of a wave front are the origin of coherent spherical wavelets with amplitude and phase of the incoming wave (elementary wave). The calculation of amplitude and phase at any point is the superposition (summation) of all elementary waves reaching this point.

Fraunhofer Diffraction Patterns

Comparatively simple relationships result from the important limiting case of a plane wave and with a parallel plane of observation points at infinity. The intensity distribution produced by an object in the wave path, in this arrangement, is referred to as the ***Fraunhofer Diffraction Pattern*** (*Joseph Fraunhofer*; 1787-1826; German optician and physicist). It can be calculated by integration of the elementary waves emanating from the opening of the object, whereby for each elementary wave, the path difference to the respective observation point must be taken into consideration. Experimentally, the conditions mentioned above (plane wave, observation plane at infinity) can be realized with convex lenses to place the source and observation plane at infinity. Simpler still is the use of a laser source which virtually produces plane waves and because of the high intensity one can place the observation plane sufficiently far away.

Forming an Image of an Object

When parallel light falls on an object, the image thus formed can be completely described by applying Huygens principle twice. At first, a *Fraunhofer diffraction pattern* is formed in the focal plane of the lens. This can be again seen as the source of elementary waves producing the image of the object in the image plane. The decomposition of the imaging process in these two stages is referred to as the *Theory of Abbe* (*Ernst Abbe*; 1840-1905; German physicist), with which, in particular, the resolution of the microscope can be estimated. The second stage of Abbe's Theory, from diffraction pattern to image, leads to extensive calculations even with simple objects; however, the aim of this experiment is to explain qualitatively, the relationship between diffraction pattern and image.

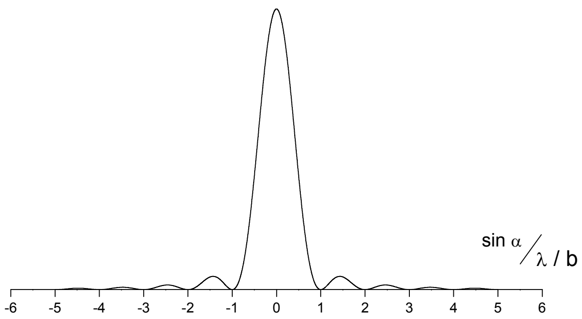
Diffraction Pattern of a Slit

The calculation of the diffraction patterns is difficult, so that here we will only present the results and a short discussion. In general, the intensity distributions are a function of the angle α or $\sin\alpha$ against propagation direction. For a single slit of width *b*, where λ is the wavelength of light, one gets:

(1)
$$I_{Sp}(\sin\alpha) \approx b^2 \frac{\sin^2\left(\frac{\pi b}{\lambda}\sin\alpha\right)}{\left(\frac{\pi b}{\lambda}\sin\alpha\right)^2}$$

Function (1) becomes zero for:

(2) $$\sin\alpha = \pm n\frac{\lambda}{b} \quad \text{for} \quad n = 1,2,3,...$$



Nullstellen der Intensität bei sin α = n · λ / b

The record of the amplitude (square root of (1)) can be derived in a simplified manner by imagining the slit as decomposed into many partial slits, then taking into account the interference of neighboring partial wave packets and finally letting the number of partial slits go to infinity.

Diffraction Pattern of a double Slit

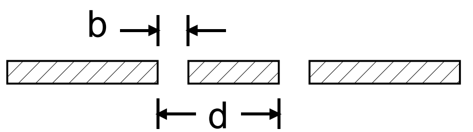
The intensity distribution for a double slit is composed of a factor fort he diffraction at a single slit as in (1) and a factor for the effect of both slits separated by a distance d:

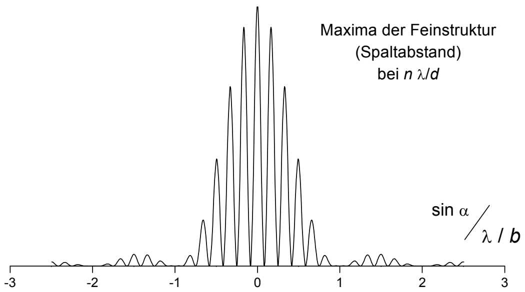
(3)

$$I_{DSp} = I_{Sp}\, I_d = I_{Sp}\cos^2\left(\frac{\pi\, d}{\lambda}\sin\alpha\right)$$

The function I_d has maxima for:

(4) $$\sin\alpha = \pm\, n\,\frac{\lambda}{d} \quad \text{for} \quad n=1,2,3...$$





Nullstellen der Einhüllenden (Spaltbreite) bei $n\,\lambda/b$

The diffraction pattern contains a series of maxima separated by λ/d, whose height is determined by the envelope of the slit width.

Diffraction Pattern of a Grating

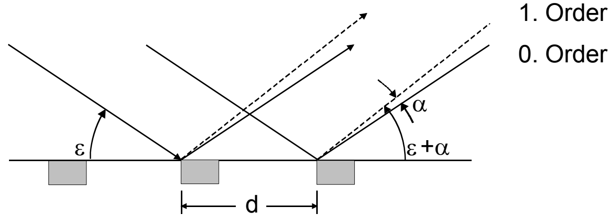
The *Fraunhofer diffraction pattern* of a grating is similar to that of a double slit, and the position of the maxima are described by (4) where, the slit separation is then termed the grating constant. The crucial difference to the double slit lies in the width of the maxima. With the double slit it is of the order of half the distance between maxima, with the grating it is of the order $1/N$ of the distance, where N is the number of slits. Hence, gratings with a high number of slits produce very sharp diffraction maxima making them an important tool in spectroscopy.

Diffraction Pattern by Reflection at a Grating

In exercise B.3 the scale divisions of a metal rule shall be determined, by illuminating the rule with laser light at a glancing angle ε. The scale divisions of the rule are then imaged as a reflection grating d sin ε, where d is the division of the metal rule.

The relationship for the position of the maxima, compared to (4) must be modified because (4) is only valid for the special case of perpendicular incidence. The effective path difference of neighboring rays is found from simple geometric considerations (see figure below):

(5) $$d\left[\cos\varepsilon - \cos\left(\varepsilon + \alpha\right)\right] = n\,\lambda$$



He-Ne Laser

In part B the same tasks are to be performed only with different experimental equipment. The source for coherent light is a He-Ne laser, with which the diffraction patterns can be directly observed, due to the high intensity, at a large distance (ca. 5 m) thus, giving a good approximation to ∞. The physical principles and the mode of operation of the laser are found in annex II *HE-NE-LASER* in the laboratory script.

Apparatus and Equipment

Experiments Part A (Na Spectral Lamp)

Optical Bank (rail). Sliding mounts for: Na spectral lamp, entrance slit, collimating lens, filter holders for the objects, objective lens and measuring ocular.
Objects (slit, double slit, grating). Sliding aperture to mask out orders of diffraction in the diffraction pattern.

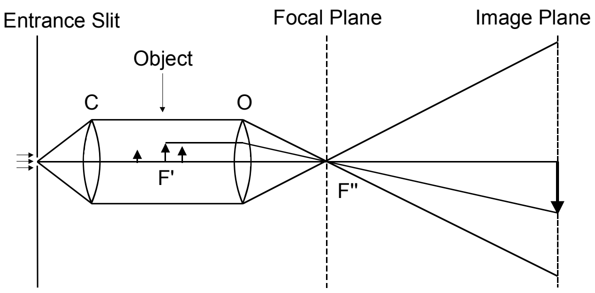
Experiments Part B (He-Ne Laser)

Components on magnetic holders: He-Ne-Laser, filter holder with double slit, Precision measuring slit, slanted metal rule.
X-Y plotter with a photo sensitive resistor (LDR), paper role to record the diffraction patterns. 5 m measure, 1 m metal rule.

Experiment and Evaluation

Experiment A (Na Spectral Lamp)

The optical bank is arranged and adjusted according to the diagram below. (C = collimating lens, Object = filter holder to support the object, O = objective lens). More detailed information on the setup is found in the lab bench script.



The final adjustment of the optical setup must be carefully performed to obtain optimal imaging and observation conditions since the quality of the measurement results strongly depends on proper alignment.

A measurement ocular is used to measure the objects or diffraction patterns. This contains a line which can be moved perpendicular to the optical axis by turning a micrometer screw.

The positions of the measurement ocular to observe the diffraction – and the image plane for the experimental tasks are given by the respective image scale and the positions fixed by two stop mounts. When performing the measurements take note that the ocular stands at these positions.

The focal length of the objective lens needed to carry out the calculation to determine the image scale in found in the lab bench script. The wavelengths of the Na-D double lines are λ_1 = 589.0 nm and λ_2 = 589.6 nm.

Exercise A.1 (Image Scale)

A slide with a mm scale is available to determine the image scale. The position of the objective lens and the ocular at the end of the optical bank must be maintained for the following measurements sine the calibration depends on these positions. Focusing must be carried out by shifting the respective object.

Exercise A.2 (Slit) and Exercise A.3 (Double Slit)

To measure the images and the diffraction patterns, the brightness ratio can be suitably matched to the conditions by adjusting the width of the entrance slit. In particular with the double slit it is of advantage to select different brightness's for the different structures (envelope, fine splitting).

Exercise A.4 (Grating)

Focusing the grating is difficult and must be made with great care since under certain defocusing conditions the image can also show sharp, periodic structures.

Exercise A.5 (Image Filtering (masking))

A sliding aperture is available which is placed in the focal plane of the objective lens (plane of the diffraction pattern). The sliding aperture must be carefully adjusted so that a sharp image of the entrance slit is formed on the partition of the second aperture when no object is placed in the object holder. The sliding aperture must be engaged in the snap lock position (ball stop) and the lateral alignment of the sliding aperture is adjusted by loosening the knurled screws.

The apertures refer to the grating as object and have the following functions:

Aperture 1: no masking.
Aperture 2: masking the maximum of the 0. order.
Aperture 3: passing (only) the 0. order.
Aperture 4: passing the 0. and a maximum of the 1. order
(tilted field illumination).
Aperture 5: passing the 0. and the ± 1. order.
Aperture 6: passing the 0., 2. and 4. order.
Aperture 7: passing the 0., 3. and 6. order.

After inserting the grating in its holder, observe and describe the images seen with the various apertures.

In particular with aperture 2 also use the mm scale as an object.

Experiment B (Laser)

Exercise B.1 (Slit) and B.2 (Double Slit)

The diffraction patterns are recorded with the aid of a light sensitive resistor (LDR = Light-Dependent Resistor) in a voltage divider circuit, whereby the partial voltage from the LDR is passed as the intensity signal to the Y-input of an X-Y recorder. The LDR is mounted on the X-sliding carriage of the recorder. If the carriage runs transverse to the optical axis of the experimental setup it directly plots the diffraction patterns. A saw-tooth voltage (time deflection of the recorder) is used to run the carriage through the set range.

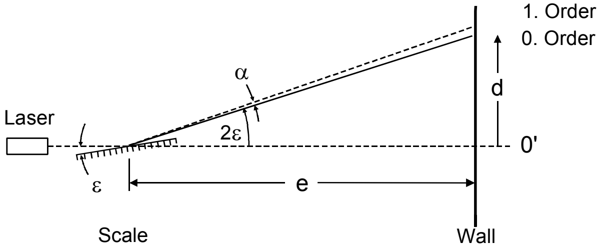
X-Y recorders are sensitive instruments and must be carefully handled! Special information on the use of the recorder is found in the lab bench script.

The input sensitivity (Y) must be set such that the first adjacent maxima of the diffraction patterns utilize the whole format of the recording paper. The intensity of the central (main) maximum is a number of times higher than the adjacent maxima. This is however, irrelevant for the experiment since the recorder is equipped with a limiting circuit which prevents overloading the unit. The rate of movement in the X-direction (time base) must be sufficiently slow (5 s/cm), so that the plots are not influenced by the electronic and mechanical time constants of the recorder.

The distance from the diffracting object is measured with a 5 m rule (with distance markings) and an additional 1 m rule. The 5 m rule is fixed between the object and LDR (entrance slit) at the same height and the 1 m rule is used to measure the additional distances to the diffracting object and to the LDR.

Exercise B.3 (Scale Divisions of a Metal Rule)

The metal rule is mounted on a wedge-shaped block and placed in the path of the laser light giving glancing deflection. The (extended) diffraction pattern is imaged on the opposite wall on a strip of paper. The diffraction maxima (including estimation errors) are marked on the paper (see figure below).



For subsequent evaluation, the position of the 0. maximum must be marked and to determine the angle of incidence ε one must mark the image point of the laser beam without the diffracting object. The angle of incidence ε is relatively large and shall not be approximated by sin ε = tan ε in the later evaluation but must be exactly taken into account. The distance between the object and the wall is again measured with the 5 m and 1 m rule.

Evaluation is performed by plotting the measurement cos(ε+α) of the diffraction maxima against the order number. The wavelength of the red He-Ne laser beam is λ = 632.8 nm.

Consider and discuss under which conditions this method is suitable for the absolute determination of the scale divisions.

FABRY-PEROT ETALON | GP II

Key Words

Diffraction and Interference; Multiple-ray interference. Interference spectrometer; *Fabry-Perot* Interferometer or *Fabry-Perot r*esonator; Optical Resonators

Aim of the Experiment

Experimental introduction to the *Fabry-Perot* interferometer as an important component in high resolution spectral equipment and in laser technology.

Literature

Standard literature (see list of standard text books); furthermore:

Bergmann·Schaefer; Lehrbuch der Experimentalphysik Band III (Optik); de Gruyter.

Exercises

1. Assembling and adjusting the apparatus.
2. Determining the plate separation of a *Fabry-Perot* etalon with the red 643.9-nm line of cadmium and calculating the (approximate) interference order.
3. Relative determination of the wavelength of the green and dark blue lines of the cadmium spectrum.
4. Estimating the line width of the interference maxima for the red line and comparison with the expected instrument line width.

Physical Principles

Multiple Ray Interference

Diffraction objects with periodic structures, as e.g., a diffraction grating, produce *multiple ray* interference with very narrow interference maxima, enabling high resolution in technical and spectroscopic applications.

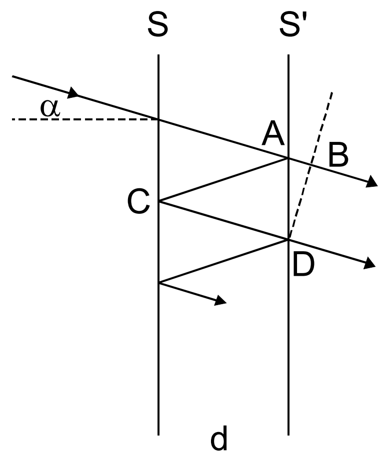
In diffraction and interference at a single or double slit, the intensity maximum appears at a point of emission where all elementary waves concerned are in-phase. A deviation in direction (detuning) from this point causes differences in the optical path length and hence portions of destructive interference, leading to a decrease in intensity compared to the maximum and the distinctiveness of the corresponding diffraction pattern (see also experiment *DIFFRACTION AND INTERFERENCE*).

If the diffraction structure repeats itself periodically, the optical path difference increases and multiplies according to the separation and number of structures and effects a much faster decrease in the intensity seen from the maximum. Multiple ray interferences are characterized by narrow maxima separated by wide dark zones (or vice-versa in multiple ray interference with reflected light.

Fabry-Perot Etalon

A *Fabry-Perot etalon* is an optical resonator formed by two plane parallel, partially reflecting surfaces and an enclosed optical medium. An incident plane wave is split into multiple, coherent partial waves due to "zigzag reflections." These interfere with one another and lead to multiple interferences in reflected or transmitted light. The ray path through an etalon is sketched in the diagram below. (Diffraction at the boundary surfaces was not taken into account since it only causes a parallel displacement of the rays).

The diagram shows the wave front of two rays emerging from the splitting of the incident ray. The path difference δ results from simple geometric considerations:



(1)

$$\delta = \overline{AC} + \overline{CD} - \overline{AB} = 2\,d\cos\alpha$$

where d is the distance between the boundary surfaces and α is the angle of incidence of the radiation.

Accordingly, the path difference is smaller the larger the angle of incidence α is. An additional path difference due to phase jumps caused by reflections at the boundary surfaces can be neglected, since, for transmission, it amounts to a whole multiple of the wavelength. The condition for constructive interference in transmitted light is then:

(2)

$$\delta = 2\,d\cos\alpha = z\,\lambda \qquad \text{with} \qquad z = 1,2,3,...$$

whereby interference maxima appear when the angle of incidence α or the wavelength λ satisfy condition (2). The path difference in units of wavelength is called the *phase value* ϕ ($\phi = \delta/\lambda$), the whole number values z of ϕ are called the *interference order* of the maximum.

The high quality of such *optical resonators* and hence the high resolution (over a small *spectral range*) is based on multiple-ray interference and the high interference order at correspondingly large resonator dimensions. With a plate separation of 5 mm and wave length of 500 nm, the interference order is z = 20 000. At such a large optical path length even very small wavelength difference sum to give destructive interference.

Free Spectral Range

The uninterrupted region of the wavelength to be investigated using spectral equipment is called the *free spectral range* or *Dispersion range*. With the Etalon one can distinguish between two neighboring maxima by their difference in order (Δz = 1) at the same wavelength, or by small differences in wavelength $\Delta\lambda$ at the same order of interference. The interference condition (2) for a definite order in both cases is:

(3) $$(z+1)\lambda = z(\lambda + \Delta\lambda)$$

Thus the free spectral range of the Etalon is:

(4) $$\Delta\lambda = \frac{\lambda}{z} \quad \text{or} \quad \frac{\Delta\lambda}{\lambda} = \frac{1}{z}$$

Because of this comparatively small dispersion range at large z, the Fabry-Perot-Etalon is preferably suited for the investigation of virtually monochromatic light of for precise measurements of narrow wavelength ranges after spectral decomposition.

Fabry-Perot Spectrometer

When used as a spectrometer, divergent light is passed through the Etalon and the parallel rays belonging to a certain order of interference are collected by a convex lens and focused on an image plane. Because of the rotational symmetry of the optical arrangement one observes in the focal plane of the lens concentric rings of equal inclination (*Haidinger Rings*; *Wilhelm Ritter von Haidinger*, 1795-1871; Austrian Geologist and Mineralogist). If one sets fort he angle of inclination α and cos α approximately:

(5)

$$\alpha = \frac{r}{f} \quad \text{and} \quad \cos\alpha = 1 - \frac{1}{2}\alpha^2$$

where f is the focal length of the lens, we get for the interference condition (2):

(6)

$$z = \frac{2\,d}{\lambda}\left[1 - \frac{r^2}{2\,f^2}\right] \quad \text{or} \quad \frac{2\,d}{\lambda} \approx z\left[1 + \frac{r^2}{2\,f^2}\right]$$

With known f, equation (6) contains the quantities λ, d and z. If one measures at least two radii of the ring system, one obtains two equations of type (6) and can thus eliminate the order z. If one specifies the innermost observable ring with the index 0 (in general the ring center does not represent an interference maximum) and the following rings with the index i (i=1, 2, 3,...), then from (6) we have

(7) $$d = i\,\frac{\lambda\,f^2}{r_i^2 - r_0^2}$$

Relative wavelength measurements are also possible without knowledge of d or z. Let r and r' be the order to a wavelength λ and $\lambda' = \lambda + \Delta\lambda$, then from (6) we get approximately:

(8) $$\Delta\lambda \approx \frac{\lambda}{2\,f^2}\left(r - r^{`2}\right)$$

The accuracy (apart from the error in the focal length f) is determined by the measurement accuracy of the radii and the above mentioned high resolution of the Etalon

does not come to bear. This is due to the fact that in (6) the order z was eliminated although it represents an essential element in the relationship (the term in the bracket: $r^2/2f^2 \ll 1$).

The Etalon must be calibrated for the exact and absolute determination of wavelength, i.e., the plate separation must be as accurately as possible, and the order exactly determined. This requires spectral analysis of several very well known wavelengths and setting the measured radii of the rings of a definite order with initial values for d and z in (6)to give a system of equations for the different wavelengths. For evaluation, the initial values for d and z are varied until the system of equations shows the largest agreement (minimum for the sum of mean square errors, i.e. the variance).

Resolution of the Fabry-Perot Etalon

The exact intensity distribution in the vicinity of an interference maximum as a function of the phase ϕ is dependent on the transmission T and reflectivity R of the plate and is described by the *Airy formula* (*Sir George Bidell Airy*; 1801-1892; Engl. mathematician and astronomer):

(9) $$\frac{I}{I_0} = \left[\frac{T}{1-R}\right]^2 \frac{1}{1 + \frac{4\,R}{(1-R)^2}\sin^2\phi\pi}$$

Hence, from this relation, one can determine the full-width-at-half-maximum $2\Delta z$ of the interference maxima, giving for small angles approximately:

(10) $$2\,\Delta z = \frac{1-R}{\pi\sqrt{R}}$$

With a plate distance of 5 mm and a reflectivity of about 92 % one gets $2\Delta z$ = 0.026, hence, a width of approx. 3 % of the ring separation. The wavelength difference corresponding to the difference in order (free spectral range) is given by (4). Inserting the full width $2\Delta z$ as the criterion for the resolution gives

(11)

$$\Delta\lambda = \frac{\lambda}{z}\,0.026 = 0.0007\,\text{nm} \quad \text{or} \quad \frac{\Delta\lambda}{\lambda} = 2\cdot 10^{-6}$$

Line Width of Optical Transitions

Seen classically, a radiating atom is a damped harmonic oscillator continuously loosing energy. The associated frequency spectrum has the form of a bell-shaped curve whose width is determined by the damping constant (*decay constant*). From the aspect of quantum mechanics, the decay constant corresponds to the *mean lifetime* of the system and the finite line width is explained from *Heisenberg's Uncertainty Principle* (*Werner Heisenberg*; *1901; German Physicist). The decay constant of optical transitions is around 10^{-8} s, giving a *natural line width* of approximately 10^{-5} nm.

Under real conditions, the lines become broader due to a number of influences. In an ensemble of atoms, the mean lifetime is shortened by collisions of atoms with one another or with impurity atoms (*pressure broadening*). Atomic thermal motion and the *Doppler Effect* (*Christian Doppler*, 1803-1853; Austrian physicist and mathematician), results in a line shift when the atoms radiate, leading to a broadening of the lines due to the statistical distribution of the velocities (*Doppler broadening*). Moreover, a number of further complex interactions may occur under various conditions resulting in line broadening.

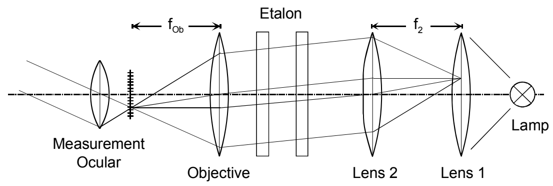
Apparatus and Equipment

Optical bank; Cd spectral lamp with power supply unit; optical lens, collimator, iris diaphragm, Fabry-Perot etalon, objective lens, ocular. Color filters (red, blue, green.)

Experiment and Evaluation

Exercise1 (Setting-up and Adjustment)

The ray path is presented in the following diagram.



The optical lens (lens 1) corresponds to the entrance slit of conventional optical ray paths and fulfils two functions. Firstly, it presents a uniformly illuminated *secondary plane light source* to generate, with the help of the collimator (lens 2), parallel light beams of different inclination. Secondly, it forms an image of the lamp approximately in the Etalon to achieve a high intensity of the interference patterns. Furthermore, an iris is available to focus the rays in the vicinity of the axis to improve the image.

The objective lens forms an image of the interference patterns in the ocular (ocular micrometer), where one can directly measure the diameter of the interference rings.

Adjusting the ray path is relatively uncritical. Important, is a careful focusing of the diffraction patterns which is subjectively difficult due to their blurred intensity distribution and a correct height adjustment to center the measuring ocular on the ring system.

Color filters are available to select the desired lines of the Cd-spectrum.

Exercise 2 (Plate Separation) and 3 (Wavelength of the Green and Blue Lines)

With the respective color filters in place, one measures the positions of the left- and right hand border of the rings for decreasing interference order. Observation of the interference patterns of the blue and green lines is physiologically problematic because of the imperfect pre-selection of the color filters and needs some time getting used to.

Evaluation is made according to equations and (7) and (8) by plotting the square of the radii against the number of rings. From the gradient of the expected lines one can determine the plate separation for the „known" red line and hence determine the wavelength of the green and blue lines.

Exercise 4 (Line Width)

The width of the red line can only be estimated due to the non-linear response of the eye. The measurement is evaluated according to (8). Compare and discuss the result with that given by equation (10). The reflectivity R of the Etalon shall be taken as 80 %.

Which addition experimental factor influences the observed line width?

SPECIFIC CHARGE e/m_0 of the Electron — GP II

Key Words

Specific charge and charge quantization. **Lorentz force.**
Circular motion.

Aim of the Experiment

Exemplary experiment on the atomic structure of electricity and measurement methods by the deflection of charged particles in orthogonal electric and magnetic fields.

Literature

Gerthsen Physik (22. Auflage): 8.1.1, 8.2

Bergmann-Schaefer, Band 2 (8. Auflage): 3.4.4, 7.1.2, 11.1.7, 11.1.4 , 11.3.5

Suplementary:

Bergmann Schäfer; Experimentalphysik Band II: de Gruyter.

Geiger, Scheel: Handbuch der Physik XXII. (Historische, vertiefende Anmerkungen)

Exercise

Determining the specific charge of the electron e/m_0 by measuring the magnetic deflection of electrons in a narrow beam tube.

Physical Principles

At the beginning of the 19th. centaury results of physical experiments showed clear indications for a *corpuscular* nature of the electric charge. The experiments of *Faraday* (*Michael Faraday*; 1791-1867; Eng. physicist and chemist) and the laws of electrolysis gave a definite relationship between the amount of electricity and the mass of the carrier of electricity and allowed, together with the idea of an atomic structure of material, the conclusion of the quantized nature of charge. *Helmholtz* (*Hermann von Helmholtz*; 1821-1894; Ger. physicist and physiologist) talked in a *Faraday commemoration speech* 1881 about the "*the electric charge of the ion*" and said "*just the same certain amount, let it be positive or negative electricity moves with each single valued ion or with each valence of a multi-valued ion, and accompanies it inseparably in all movements the ions make through the liquid. If carried over to the electrical processes, this hypothesis, leads, in connection with Faraday's law, to a somewhat surprising conclusion. If we assume atoms of chemical elements, then we must conclude that also electricity, positive as well as negative, is divided into certain elementary quanta which behave as atoms of electricity*." The electrolytic measurements, however, only gave mol values for ions, so that the determination of the specific charge and the ion charge presupposed the knowledge of Loschmidt's number.

The story of the discovery of the electron began in 1860 with the investigation of gas discharges and the discovery of cathode rays by *Plücker* 1859 (*Julius Plücker*, 1801-1868; German physicist and mathematician). *Hittorf* (*Johann Wilhelm Hittorf*, 1824-1914; German physicist and chemist) discovered in 1869 the magnetic- and *Goldstein* (*Eugen Goldstein*; 1850-1930; German physicist) in 1876 the electric deflection of cathode rays, whereby, *Hittorf* already suspected negatively charge particles were the actual physical nature of the cathode rays. Around 1890 *Schuster* (*K. Schuster*, *1903), *Thomson* (*Sir Joseph John Thomson*; 1856-1940; Engl. physicist) and *Wiechert* (*Emil Wiechert*; 1861-1928; German Geophysicist) performed the first measurements of the specific charge in various deflection experiments. Even though the early results of these measurements had large errors, it was found that "*the mass of the cathode rays must be about 2000 times smaller than that of the lightest ions.*" Extensive measurements of the velocity (*Thomson*) and investigations on the velocity dependence of the mass and the agreement of the specific charge of the cathode rays with those of electrons released by photoelectric or thermal effects gave certainty to the existence of electrons. *Thomson* is named as the classical discoverer.

At the beginning of the 20th. centaury the value of the electron charge itself was measured (*Millikan* 1910-1925; see *MILLIKAN Experiment* in the laboratory script) and, especially due to the determination of the specific charge from the *Zeeman Effect* (*Pieter Zeeman*; 1865-1943; Dutch physicist) and from the *Rydberg Constant* (*Johannes Rydberg*; 1854-1919; Swed. physicist; see experiment *OPTICAL SPECTROSCOPY* in this laboratory script), the identity between the free electrons and the bound electrons as building blocks of atoms was proven.

The recommended values for the electron (CODATA 1999) are
(*Physikalische Blätter, March 2000*):

(1a) e/m_0 = $1.758820174 \cdot 10^{11}$ C/kg
relative error: $1.2 \cdot 10^{-7}$

(1b) e = $1.602176462(63) \cdot 10^{-19}$ C
relative error: $3.9 \cdot 10^{-8}$

(1c) m_0 = $9.10938188(72) \cdot 10^{-31}$ kg
relative error: $7.9 \cdot 10^{-8}$

If one interprets the total rest mass of the electron according to *Einstein's equation (E = mc²)* as the electrostatic energy content of spherical charge distribution of radius *r*, then we have:

(2) $$m_0\, c^2 = \frac{e^2}{8\,\pi\,\varepsilon_0\, r}$$

($E = CU^2/2$ with $C = 4\pi\varepsilon_0 r$). Solving for r and inserting the respective values gives $r = 1.4 \cdot 10^{-15}$ m. This value at least reflects correctly the order of magnitude of the electron radius as supported by deflection experiments, even when from the stand point of quantum physics it is basically unfounded to describe or explain the electron with classical concepts. In the literature it is usual to take the double value of (2) as the **classical electron radius**:

(3) $$r_0 = \frac{e^2}{4\pi\varepsilon_0 m_0 c^2} = 2.8179378(70)\cdot 10^{-15}\,m$$

Measurement Method

The ***Lorentz force*** acts on charged particles moving in a magnetic field ***B*** (*Hendrik Antoon Lorentz*; 1853-1928; Dutch physicist):

(4) $$\boldsymbol{F}_\mathrm{L} = q\left(\boldsymbol{v}\times\boldsymbol{B}\right)$$

where ***v*** is the velocity of the charged particle. The *Lorentz force* always acts perpendicular to ***v***; and thus alters the direction of the velocity but not its magnitude. Particles with constant velocity in a homogenous and constant magnetic experience a constant change of direction, i.e. they trace out a circular path. The equation of motion for an electron (point particle) with $q = -e$ is:

(5) $$-e\left(\boldsymbol{v}\times\boldsymbol{B}\right) = m\dot{\boldsymbol{v}}$$

In a constant magnetic field in the *z*-direction with ***B*** = (0/0/B) the components of the equation of motion are as follows:

(6) $$\begin{bmatrix} e\boldsymbol{B}\boldsymbol{v}_\mathrm{x} = m\dot{\boldsymbol{v}}_\mathrm{y} \\ -e\boldsymbol{B}\boldsymbol{v}_\mathrm{y} = m\dot{\boldsymbol{v}}_\mathrm{x} \\ 0 = \dot{\boldsymbol{v}}_\mathrm{z} \end{bmatrix}$$

A solution to the system of differential equations (5) is

(7) $$\begin{bmatrix} \boldsymbol{v}_\mathrm{x} = -\boldsymbol{v}\sin\omega t \\ \boldsymbol{v}_\mathrm{y} = -\boldsymbol{v}\cos\omega t \\ \boldsymbol{v}_\mathrm{z} = 0 \end{bmatrix}$$

and leads to the condition ($m = m_o$):

(8) $$\omega = \frac{\mathrm{e}}{\mathrm{m}}\boldsymbol{B}$$

The *precession frequency* (8) of a charged particle in a magnetic field is known as the *Larmor frequency* (*Sir Joseph Larmor*, 1857-1942; Brit. physicist) and is of significance in atomic- and nuclear physics. With $\omega = \boldsymbol{v}/\boldsymbol{r}$, we have for the specific charge:

(9) $$\frac{e}{m} = \frac{\boldsymbol{v}}{\boldsymbol{r}\,\boldsymbol{B}} \quad \text{or} \quad \left(\frac{e}{m}\right)^2 = \frac{\boldsymbol{v}^2}{\boldsymbol{r}^2\boldsymbol{B}^2}$$

where *r* is the radius of the circular orbit.

If one accelerates electrons through an electric voltage *U*, the velocity of the electrons can be derived from the energy:

(10) $$\mathbf{e}\boldsymbol{U} = \frac{1}{2}m\,\mathrm{v}^2 \quad \text{or} \quad \boldsymbol{v}^2 = 2\boldsymbol{U}\frac{e}{m}$$

and inserting in (9) gives:

(11) $$\frac{e}{m} = \frac{2\boldsymbol{U}}{\boldsymbol{r}^2\boldsymbol{B}^2}$$

Narrow Beam Tube

The production, acceleration and investigation of the motion of electrons takes place in the ***narrow beam tube***. Free electrons are generated at the cathode of an electrode system and accelerated by a voltage (70-300 V) to the anode. A focusing electrode (Grid, *Wehnelt cylinder*; *Arthur Rudolph Wehnelt*; 1871-1944; dt. physicist) is placed between cathode and anode. The electron beam escapes into the electric field-free room through a hole in the anode.

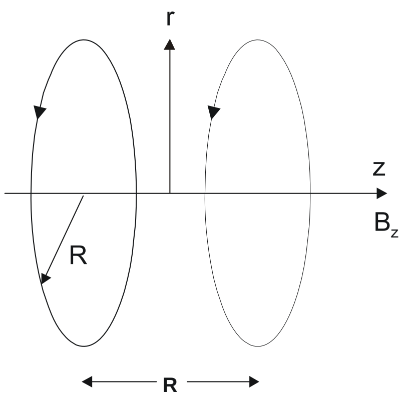
A homogenous magnetic field applied perpendicular to the direction of the beam forces the electrons into a circular orbit. The spherically shaped narrow beam tube is filled with hydrogen (about 1 Pa) so that the hydrogen molecules become optically excited by collisions with the electrons. The resulting emission of light makes the orbit of the electrons visible.

The reason why the electrons do not fan out due to mutual repulsion when entering the field-free space is connected with the focusing action of the space charge when electrons collide with the hydrogen molecules (narrow beam).

Helmholtz Coils

Helmholtz coils are used to generate homogenous and isotropic magnetic fields of medium strength. They consists of a pair of coils placed parallel and coaxial opposite each other. The magnetic field in the middle of a pair of windings in this arrangement is given by:

(12) $$\boldsymbol{B}_0 = \mu_0\,\frac{8}{5\sqrt{5}}\,\frac{\boldsymbol{I}}{\boldsymbol{R}}$$



Points with the coordinates *z* and *r* (see figure) have the following axial and radial field components:

(13)
$$\boldsymbol{B}_z = \boldsymbol{B}_0 \left\{1 - \frac{18}{125}\left[3\frac{\boldsymbol{r}^4}{\boldsymbol{R}^4} - 24\frac{\boldsymbol{r}^2}{\boldsymbol{R}^2}\frac{\boldsymbol{z}^2}{\boldsymbol{R}^2} + 8\frac{\boldsymbol{z}^4}{\boldsymbol{R}^4}\right] + ...\right\}$$

(14)
$$\boldsymbol{B}_r = \boldsymbol{B}_0 \left\{\frac{72}{125}\frac{\boldsymbol{r}}{\boldsymbol{R}}\frac{\boldsymbol{z}}{\boldsymbol{R}}\left[4\frac{\boldsymbol{z}^2}{\boldsymbol{R}^2} - 3\frac{\boldsymbol{r}^2}{\boldsymbol{R}^2}\right] + ...\right\}$$

Apparatus and Equipment

Narrow beam tube; Operating unit with adjustable anode voltage and auxiliary voltages. **Helmholtz coils**; Power supply unit for field current. Multimeters for accelerating voltages and field current.

Experiment and Evaluation

Attention: When preparing the tube for operation, take note of the information in the laboratory bench script!

Short wire pins are arranged in to tube to set the diameter of the electron orbit. The diameter specified by these pins is listed in the laboratory bench script.

A power supply unit is provided for the current supply of the ***Helmholtz coils***. When connecting the current circuit, ensure that both coils have their own connections and that the coils are connected in series. Coil data (radius, winding number) are also given in the laboratory bench script. The measurement of the accelerating voltage *U* and the magnetic current *I* is by means of two digital multimeters.

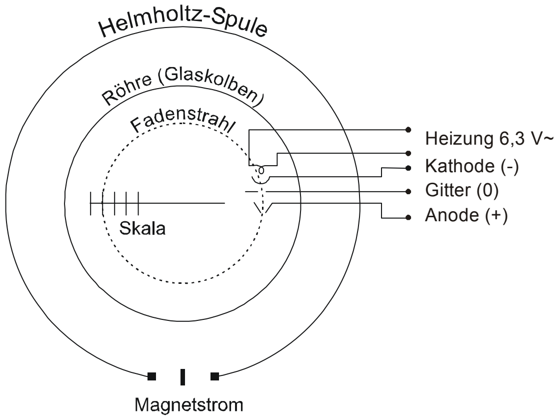
The measurements are to be carried out for the three largest orbital diameters. For evaluation (from (11) with (12)) the voltage *U* is plotted against the square of the current I^2. The specific charge is determined from the gradients of the straight-line curves.

A single final result is to be given (weighted mean value) from the three measurements.

Supplementary Questions

The experiment is carried out with electrons from a glow emission. What can be assumed about the initial energy of the electrons? How does this influence the measurement result?

How large is the linear velocity of the escaping electrons? Must a relativistic correction be taken into account?



The adjacent figure shows the basic circuit of the narrow beam tube. A special operating unit supplies the tube with voltage and current. The electrical connections for the tube are labeled on the operating unit and the sockets of the tube.

A focusing electrode (grid) is connected between the cathode and anode. The effective accelerating voltage, however, is the total voltage between cathode and anode.

MILLIKAN EXPERIMENT | GP II

Key Words

Charge quantization, elementary charge. *Coulomb force*; *Stokes' Law*.

Aim of the Experiment

Simple and exemplary experiment to verify and measure quantized values.

Literature

Bergmann-Schaefer, Band 2 (8. Auflage), Kapitel 7.1.2

Exercise

Determining the elementary charge from the measurements of the fall- and rise times of charged droplets of oil of various sizes, taking into account a radius dependent correction.

Physical Principles

Charge Quantization

After the discovery of electrical phenomena, people began, in the middle of the 18th. centaury, to make scientific enquiries as to the nature of electricity. The predominant theories at this time were *fluid theories*, although others argued for a more simple, mechanical world view, as e.g. *Franklin* (*Benjamin Franklin*; 1706-1790; Am. politician, writer and natural scientist), who imagined electricity to consists of small particles able to move through materials without hindrance.

Already in 1833 *Faraday* (*Michael Faraday*; 1791-1867; Eng. physicist and chemist) discovered, with his laws of electrolysis, the first experimental evidence for an atomistic nature of electricity without, however, (him and other physicist of the time) recognizing the quantization of charge. Only in the seventies and eighties of the last centaury did *Stoney* (*George Johnstone Stoney*; 1826-1911; Eng. physicist) and *Helmholtz* (*Hermann von Helmholtz*; 1821-1894; Ger. physicist and physiologist) formulate in all clarity that "*electricity is made up of certain, elementary quanta, which behave as atoms of electricity*".

Electrolytic measurements only gave values related to mass, such as e_o/m_I or $L \cdot e_o$ (elementary quantum of charge e_o, ion mass m_I, *Avogadro's number L*). The first values of the ion charge were found through independent experiments to determine the constant *L*. Since these experiments were fraught with difficulties, others were conceived in which the number of charge carriers could be directly counted. The charge carriers used were individual atomic ions (α-particles) as well as ions attached to liquid droplets (ion-cloud method). In addition, experiments were performed to determine the ion charge from radioactive decay data.

The further developments of the cloud- and droplet method lead to charge measurements on single, isolated particles. The leading working groups were those of *F. Ehrenhaft* (*Felix Ehrenhaft*; 1879-1952; Austrian physicist) in Vienna and *R. A. Millikan* (*Robert Andrews Millikan*; 1868-1953; Am. physicist) in Chicago. Both groups worked from 1910 onwards with the method described below. The two necessary measurements (in field-free space and under the influence of an electric field) were carried on one and the same particle.

While *Millikan* tried to prove the existence of the elementary charge qualitatively and quantitatively, *Ehrenhaft* was not convinced of the notion of an elementary quantum and the aim of his work was to disprove this concept. *Millikan* published his essential work in 1911 and 1913 and carried on till the middle of the twenties to systematically improve his results. *Millikan's* last, in 1925, published result was (converted to in SI-units):

(1) $$e_o = 1.5924(17) \cdot 10^{-19}\ \mathrm{C}$$

The importance of *Millikan's* work lies in a direct and simple qualitative and quantitative proof of charge quantization. In 1923 he received the Noble prize in physics for these measurements and for the experimental confirmation of Einstein's equation $E = h \cdot \nu$.

Today, the existence of the elementary charge is beyond doubt even when considerations are given to the existence of fundamental particles with the charge 1/3 e_o and 2/3 e_o (Quarks). Furthermore, the elementary charge seems to have the same magnitude for all known elementary particles. In experiments, the relative difference between electrons and protons was found to be less than 10^{-20}. At present, the elementary charge is determined indirectly from other experimental data. In 1999 CODATA recommended the following value (Physikalische Blätter 2000):

(1a) $$e_o = 1.602176462(63) 10^{-19}\ \mathrm{C}$$

relative error: $3.9\ 10^{-8}$.

Measurement Method

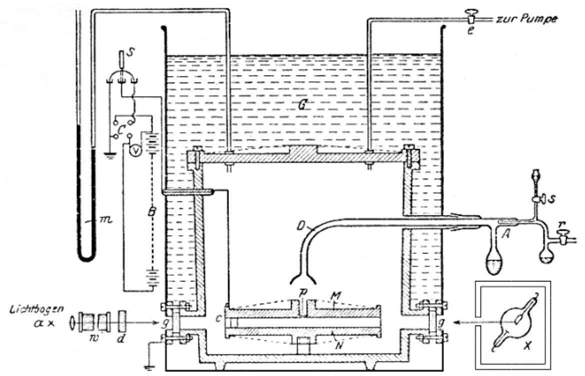
Four forces act on a spherical, charged body (here, a small droplet of oil) placed in a vertical, homogenous and constant electric field and within a homogenous medium: the force of gravity $\boldsymbol{F}_G$, the force of buoyancy $\boldsymbol{F}_A$, the *Coulomb force* $\boldsymbol{F}_C$ (*Charles Augustin de Coulomb*; 1736-1806; French physicist) and a frictional force $\boldsymbol{F}_R$. If laminar flow is assumed, the frictional force is given by *Stokes' Law* (*Sir George Gabriel Stokes*; 1819-1903; Engl. mathematician and physicist):

(2) $$\boldsymbol{F}_R = -6\,\pi\,\eta\,\boldsymbol{r}\,\boldsymbol{v}$$

η is the viscosity, $\boldsymbol{r}$ the radius of the droplet and $\boldsymbol{v}$ its velocity. Because of the velocity dependence, the frictional force compensates the other forces after a sufficient time and one observes a force-free uniform motion. One can thus calculate the radius of the droplet by measuring its velocity when falling in field-free space.

If one now applies an electric field (e.g. between two capacitor plates), so that the droplet rises, one can then calculate the charge *q* of the droplet from the radius and rise velocity.

The experimental setup consists of a plate capacitor, an atomizer to produce fine droplets of oil, illuminating equipment and a measuring microscope to observe the droplets (see figure).



R. A. MILLIKANS endgültige Versuchsanordnung. *MN* Kondensator. *AD* Zerstäuber. *G* Ölbad. *a w d g* Beleuchtung des Teilchens.

The liquid droplets generated by the atomizer are for the most part electrically charged due to friction effects. The measurements showed that the determined values of charge were small multiples of a smallest possible quantity, from which one can infer a quantized structure of charge and thus determine the elementary charge:

$$q = n \cdot e_0 \tag{3}$$

In order to determine the elementary charge as the smallest common integer multiple *n* of the measured value, one must measure not only an adequate number of different values but with respect to accuracy and resolution, also sufficiently small values (the error or resolution of the measurement must be at least small compared to e_o).

Millikan himself essentially measured the charge reversal of droplets brought about by light illumination via the photo effect. Another possibility of charge reversal comes from the attachment of gas ions and is a source of error in the measurements. This behavior can be recognized by the fact that the velocity of a droplet of oil jumps suddenly in the electric field.

Correction to *Stokes' Equation*

Millikan's measurements showed that the quantum of charge was always constant for one droplet but varied for different droplets. The systematic revealed deviations in the values for small droplets. From this, one concluded that the validity of *Stokes Law* in its simplest form (3) must be false. The basic assumption of the *homogeneity* of the surrounding medium is – just by small droplets, not well fulfilled. At atmospheric pressure, the radius ***r*** of the particles is no longer negligibly large with respect to the mean free-path of the gas molecules. Taking into account the so called *kinetic correction* introduced by *Cunningham* (*E. Cunningham*; Eng. mathematician), the frictional force is now expressed as:

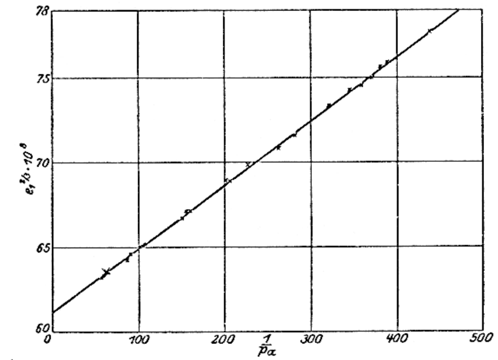
$$F_R = \frac{-6\,\pi\,\eta\,\boldsymbol{r}\,\boldsymbol{v}}{1 + A\,\dfrac{\lambda}{\boldsymbol{r}}} \tag{4}$$

Millikan devised a graphical procedure to take this correction into account since an exact value of the constant *A* ($A \approx 0.85$) was not known. If one designates ***e*** as the uncorrected elementary charge subject to the simple frictional force (3) and e_o the corrected value subject to (4), then one gets the following relationship by allowing for the correction factor (1 + *A*/*r*) in the expression of the velocities in the equation for *q*:

$$e = e_0 \left(1 + A\,\frac{\lambda}{\boldsymbol{r}}\right)^{3/2} \quad \text{or} \tag{5,6}$$

$$e^{2/3} = e_0^{2/3} \left(1 + A\,\frac{\lambda}{\boldsymbol{r}}\right)$$

If one plots the value of ***e*** found from the droplets of different radii *r* against 1/*r* one should get a linear relationship, where the corrected value e_o is read from the axis intercept (extrapolation auf $1/r \rightarrow 0$ or $r \rightarrow \infty$ respectively).



[Eine Serie aus MILLIKANS endgültigen Messungen; empirische Korrektion der Abhängigkeit vom Teilradius. Ölteilchen in Luft.

This procedure is an approximation since the correction must also be taken into account for *r*. On the basis of the evaluated results, a decision must be made as to weather a second correction is required.

Apparatus and Equipment

Millikan apparatus (plate capacitors, oil atomizer, microscope).
Operating unit with voltmeter. Stopwatch.

Experiment and Evaluation

Apparatus

Three Millikan setups are available (*LEYBOLD equipment*), which are simple and clearly arranged but only allow measurements with just sufficient accuracy and resolution. An atomizer to produce oil droplets, a microscope and an illuminating lamp are arranged around a large round plate capacitor. The voltage supply is provided by a power supply unit. The fall- and rise times of the droplets are measured with a stopwatch. The measurements are subjectively difficult and searching for the droplets and following the up- and down motion with the microscope is tiring on the eyes. Usually it takes a number of attempts with the atomizer to find the right droplet (see below).

An additional experimental setup is available from *Pasco-Scientific* which allows much better measurements.

Producing and Selecting the Droplets

Droplets of suitable charge and size must be selected from the mist produced by the atomizer. In general, large droplets carry a high charge and are therefore unsuitable for the detection of quantization and the determination of the elementary charge. Large drops can be recognized by a large falling velocity without an electric field.

When looking through the microscope, falling droplets "rise" in field-free space because of image reversal. When the voltage is switched on, the droplets must move in the opposite direction. With the *LEYBOLD apparatus* one selects from these droplets those that have a falling velocity of less than 0.5-1 scale division/s. This may require some effort since the atomizer produces mainly large droplets. More information is given in the lab bench script.

Evaluation of the measurements, taking into account the radius dependent correction according to equation (6) requires measurement data for droplets of different radii.

Measurements

The measurements consist of collecting fall-and rise time data of droplets traveling a given distance. The measurements must take into account the required accuracy and the possibility of checking and reproducing the results.

Of all the data needed for evaluation, the capacitor voltage contributes the largest and thus the decisive error. The voltage is adjustable and measured with an instrument of quality class 2.5.The time measurements should be made such that the total error does not increase significantly.

The fall- and rise times of a droplet must be repeatedly measured to increase the accuracy and to check the error and the reproducibility. A compromise must be made with respect to the voltage. A large voltage gives a small (relative) voltage error, but a larger time error due to the short rise time and vice versa.

FRANCK–HERTZ-EXPERIMENT	GP II

Key Words

Atomic model, structure of electron shells, quantum numbers, states (energy levels), transitions, electron collisions, excitations.

Aim of the Experiment

Historically significant and exemplary experiment proving the discrete structure of the excited states of atoms.

Literature

Gerthsen Physik; (22. Auflage) 12.4; 13.2.3

Eichler Kronfeld Sahm; “Das neue Physikalische Grundpraktikum“ 4.5

Finkelnburg; Atomphysik; Springer.

Tasks

1. Observing the electron collision-excitation curve (*Franck-Hertz Curve*) of mercury at an oven temperature of about 190 °C using an oscilloscope. Optimizing the curve by adjusting various experimental parameters (oven heating, cathode heating, accelerating voltage).
2. Quantitative recording of the curve with an XY-chart recorder. Determining the associated transition energy in mercury. Calculating the wave length and frequency of the transition.
3. Observing and recording further excitation curves for temperatures of 130-150 and 200°C. Qualitative discussion of the results.
4. Recording and evaluating a *Franck-Hertz curve* for neon at room temperature.

Fundamental Principles

Around 1911 *James Franck* (1882-1964; German Physicist) and *Gustav Hertz* (*1887; German Physicist) carried out experiments to investigate the interaction of electrons with gas molecules. In 1913 they proved the existence of discrete excitation states in mercury thus verifying the quantum hypothesis and the *Bohr-Sommerfeld model of the atom* (*Niels Bohr*; 1885-1962; Danish Physicist; *Arnold Sommerfeld*; 1868-1951; German Physicist).

The equipment they used was based on work done by *Lenard* (*Philipp Lenard*; 1862-1947; German Physicist). A tube containing a cathode, anode (in the form of a wire grid) and a collector electrode is filled with mercury vapor. The cathode is heated to emit electrons which are then accelerated towards the anode (accelerating voltage approx. 100 V). Behind the anode grid is a collector electrode with a small bias voltage (≈ 1 V) compared to the grid.

At first when the electrons are accelerated with increasing voltage only elastic collisions of electrons with mercury atoms occur. One observes a steadily increasing current in the collector circuit. However, as soon as the accelerating voltage reaches a certain threshold value or an integer multiple thereof, the current suddenly breaks down and then begins to rise again.

The tube used in this laboratory experiment contains an indirectly heated barium oxide cathode, a wire grid anode and a collector electrode. The electrodes are arranged plane- parallel to each other. To achieve high collision probability, the distance between the cathode and anode is large compared to the mean free path of the electrons. The distance between the anode and collector, on the other hand, is small compared to the electron mean free path.

The Hg-tube is mounted in an oven containing a regulated heating coil. The vapor pressure of mercury at a temperature of 170 °C is about 25 hPa.

Equipment

Franck-Hertz tube with mercury vapor. *Franck-Hertz tube* with neon gas.

Operating unit. Power supply units for additional auxiliary voltages.

Oscilloscope. Multimeter. XY-chart recorder. Cables.

Experiment and Evaluation

General Information

The voltage supply of the tubes (Hg and Ne) is delivered by a special operating unit. The connection and operation of the tubes is performed according to the information contained in the respective exercises.

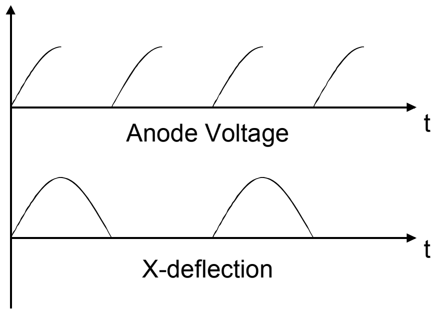
The operating unit incorporates a very sensitive measuring amplifier to amplifier the collector current which is detected as a voltage (in arbitrary units) at the output. The amplification factor and zero-point level are set with rotary knobs on the front panel. A bias voltage of 1.5 V between anode and collector is fixed inside the operating unit.

The operating unit allows static operation (-) for a point-wise (multimeter) or continuous slow tracing of the curve on the XY-chart recorder and dynamic operation (∠) for the periodic excitation and observation of the curves with the oscilloscope.

Dynamic Operation

In dynamic operation (∠) the anode voltage is periodically increased from 0 to the maximum value set by the rotary knob. The complete excitation curve is displayed in the oscilloscope when one connects the anode voltage to the X axis and the output of the measuring amplifier to the Y axis of the oscilloscope. The operating unit has two 4 mm sockets (X and Y) to provide the connection to the oscilloscope.

(The voltage at the "X output" of the operating unit is not identical with the anode voltage. It has the same time behavior but is independent of U and can be separately adjusted with the rotary knob. This has the advantage that a change in U results in a change of the voltage range of the X axis, however, the size of the curve on the screen does not change. The size of the curve is adjusted with the X-deflection).



Static Operation

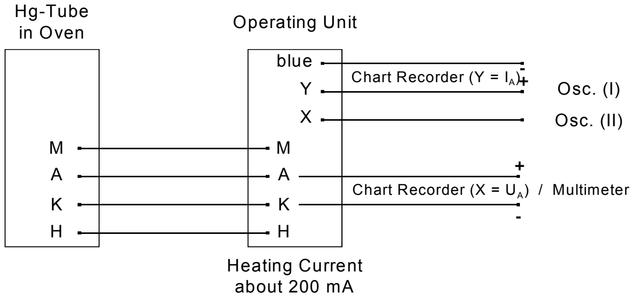
In static operation (-) the anode voltage remains constant and is adjusted with the rotary knob.

Cathode Heating

The functioning of the tube (shape of the excitation curve with the number and form of the observed minima and maxima) depends in a very sensitive way on the electron emission and thus on the heating power (heating voltage, heating current) applied to the cathode.

Exercise 1 (Optimizing the Hg-tube in dynamic operation)

Heat the oven to about 170 °C. Connect the Franck-Hertz tube according to the labeling on the equipment and as shown in the figure below:



ATTENTION: Turn on the cathode heating 10 minutes before applying anode voltage!

The cathode heating must be patiently and carefully regulated.

When the cathode emission is excessive and the anode voltage too high the tube ignites resulting in a gas discharge.

This is seen as a sudden over-amplification of the signal and a red glow inside the tube.

The heating current must also be continuously checked during the run of the experiment.

Switch the operating unit to dynamic mode (∠). Set the accelerating voltage first to 0 and then during the experiment slowly regulate to higher values.

Set the oscilloscope to XY-mode and the inputs to DC coupled. Set the input sensitivity to 1 V/cm. Adjust the Y- and X- position knobs to bring the point of light in the middle of the screen. Set the size and position of the curve with the *zero point*, *amplification* and *X-deflection* knobs on the operating unit. The observed voltage range is determined by the setting of U_B at the operating unit. After the curve is optimized write down all the associated parameters.

Exercise 2 (Quantitative recording of the curve with the chart recorder)

When the equipment and the experiment are fully understood, connect the X -inputs of the recorder to the anode voltage ("K" and "A") and the Y-inputs to the amplifier outputs (blue and red sockets) (see circuit diagram).

XY-chart recorders are delicate and expensive instruments. Operate the chart recorder with care and note the additional information in the lab bench script. Scan the spectrum slowly to avoid fast changes in voltage.

To record the spectrum, switch the operating unit to static mode (-) and carefully increase the anode voltage by slowly turning the U_B on the operating unit.

To optimize the useful measuring ranges select variable (*VAR*) on the recorder. Start with the less sensitive recorder settings and pen up to optimize the recorder settings.

The **X-axis is calibrated** by connecting a multimeter parallel to the X-input (see circuit diagram) and measuring the start- and end values of the spectrum and marking them on the chart paper (mark the values by a small movement of the Y-position knob).

Because of the contact voltage between cathode and anode the minima appear to be shifted. The excitation energy is evaluated by drawing a graph of voltage minima against ordinal number. The excitation voltage is then found by determining the slope of this straight line.

Exercise 3 (Excitation curves at other temperatures)

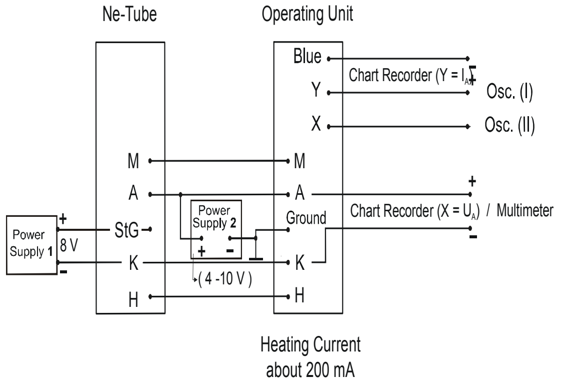
The excitation curves for other oven temperatures, e.g. 150 and 200 °C are observed and optimized on the oscilloscope. At lower temperatures one gets higher collector currents and thus only relatively small accelerating voltages can be set without saturating the measuring amplifier. The curves are to be recorded with the XY-chart recorder.

Exercise 4 (Recording the excitation curve for Neon)

The Ne-tube is connected essentially in the same way as the Hg-tube, with the exception of two additional auxiliary voltages for a control grid and bias voltage between

anode (+ pole) and ground of the amplifier (- pole)

[see following circuit diagram or three pages further on].



ATTENTION: Turn on the cathode heating 10 minutes before applying anode voltage!

The cathode heating (approx. 260 mA) as well as the control voltage (approx. 8 V) and the bias voltage (4-10 V) are optimized by checking the spectrum on the oscilloscope (best minima-maxima ratio).

PHOTO EMISSION	GP II

Key Words

Hallwachs-Effect; Light-/ Photo Electric Effect; Outer Photo Effect. Quantum Theory of Light, Quantum Optics, Photon.

Einstein Relationship. Work Function.

Aim of the Experiment

Reproduction of a classic experiment for the phenomenological introduction to the quantum nature of light.

Literature

Bergmann-Schaefer; Band 3 (Optik) (9. Auflage) Kapitel 7.1, 7.2, 7.4

Gerthsen Physik (22. Auflage) Kapitel 8.1.2, 12.1.1, 13.6.

Tasks

1. Setting up and adjusting the equipment.
2. Measuring the saturation current and the stopping potential of a calcium photo detector as a function of the intensity of the λ=436 nm line (indigo/blue) of mercury.
3. Measuring the current-voltage characteristic curve for all main lines of the mercury spectrum in the visible range. Evaluating the characteristic curve and determining *Planck's Constant* and the work function of calcium.
4. Theoretical work for the experimental report: Explaining the contradiction between the experimental results of photo emission and the classic wave theory of light.

Fundamental Principles

Hallwachs-Effect

Heinrich Hertz (1857-1894; German Physicist) said in a lecture "*On the Relationship between Light and Electricity*" in 1889: "The wave nature (of light) is, speaking from a human point of view, Certainty; (and) what follows from this, is also Certainty!"

At the same time *Hertz* and others investigated the influence light radiation had on the electric discharge processes. *Wilhelm Hallwachs* (1859-1922; German Physicist) discovered in 1888 that a metal plate radiated with light of short wave length emitted (negative) electric charges thereby becoming itself positively charged. This, however, only occurred up to a certain voltage, the *Stopping Potential.* From this point on, no charges were emitted (*Hallwachs-Effect*). The charges emitted from the metal by the action of light were identified as electrons (*Photo emission*).

Following this discovery, very intensive studies were carried out on the effect, since the observations were in contradiction to the ruling wave theory of light. The measurement of the quantity (photo current) and the kinetic energy of the emitted electrons as a function of the applied voltage and the wave length of light gave the following essential results:

- The effect occurs instantaneously.
- The saturation current is proportional to the intensity of the light.
- The kinetic energy of the emitted electrons is independent of the light intensity.
- The kinetic energy, however, is dependent on the frequency of light and increases at higher frequencies.
- There exists a limiting frequency for the effect (long wave length limit).

Quantum Theory of Light

Albert Einstein (1879-1955; German/American Physicist) proposed in 1905 a new corpuscular theory of light to account for the observations:

- Light of frequency ν consists of quantum particles (Photons) which move in a straight line at the speed of light ***c*** and carry the energy (Quantum Energy) $h\nu$ (*photo electric Einstein relationship*).
- In photo emission, a photon instantaneously transfers its total energy to an electron, whereby the energy $h\nu$ is taken up to eject the electron from the surface of the metal (work function) and to provide the kinetic energy of the electron.

The constant *h* is called ***Planck's Quantum of Energy*** (*Max Planck*; 1858-1947; German Physicist).

Work Function

The work function is the energy needed, to eject an electron from the surface of a solid (especially conduction electrons from metals). The alkaline metals possess a relatively low work function of about 2 eV and a long wave length limit in the range of visible light. The ease with which electrons are given up is also mirrored in their chemical properties. The other metals in the periodic table, further to the right of the alkaline metals, possess work functions of about 4 -5 eV with a long wave length limit in the ultraviolet region.

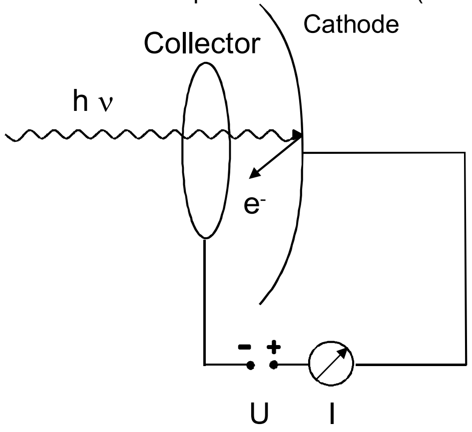
The work function is dependent on the crystal surfaces through which the electron passes and especially from the respective state of the surface. Literature values should therefore be considered as approximations. The work function of pure calcium is given as (*Kohlrausch*; Praktische Physik Band 3; Teubner):

(1)

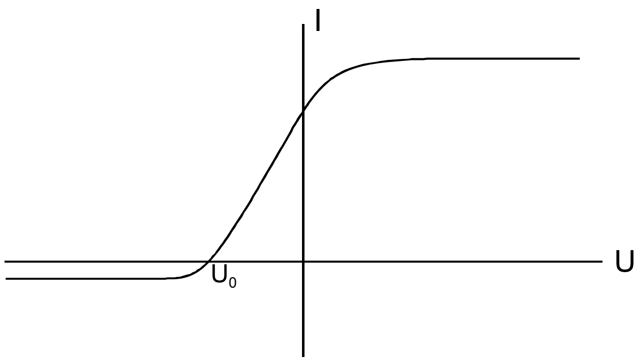
$$\mathbf{U_A} = 2.25\,\mathrm{eV} \qquad \text{with} \qquad \lambda_{\lim\mathbf{it}} = 551\,\mathbf{nm}$$

Experimental Setup

The experimental setup consists of a surface cathode coated with a layer of the metal to be investigated and a collector electrode to provide the bias potential and to collect the photo current (see Figure).



Basic Principle of the Photo Detector



Typical Characteristic Curve

Alkaline metals are mainly used for simple experiments, (in the present setup calcium). They have a low work function and a relatively high *quantum yield* (ratio of the number of photons hitting the surface to the number of ejected electrons).

The photo currents produced, however, are very small and lie in the range of a few nano- to pico-amperes. The measurement of such small currents is difficult and subject to many stray influences so that in the framework of the possibilities available with the modest experimental equipment one can expect large deviations and errors – which are to be discussed.

The characteristic curves of the photo detector (*I-U* diagrams) have approximately the form shown in the figure above. In stead of the (theoretically expected) cut-off of the photo current at the maximum energy of the electrons (stopping potential U_o), the characteristic curve proceeds with a negative photo current originating from photo emissions in the reverse direction, i.e. from the collector to the cathode.

A number of theoretical and empirical assumptions can be made for a precise discussion and evaluation of the form of the characteristic curves. As far as the experiment is concerned (see above), a starting point for the quantitative evaluation of the stopping potential shall be, on the one hand, the *real zero cross-over* U_o and as an alternative, the *extrapolated zero cross-over* U'_0 *from the linear part of the characteristic curve*. This is based on the consideration that the measurement conditions at $I_{Ph} = 0$ are more strongly subjected to interfering influences while the measurement settings at higher photo currents exhibit stabile values.

The work function, in reality, is not a very reliable parameter because it depends strongly on the state of the surface, whereby numerous irregularities are present in any given extended surface. Additional errors in the determination of the work function from measurements of the photo-electric effect come from contact potentials which can lead to a displacement of the stopping potential or the kinetic energy of the emitted electrons.

In this experiment, pay particular attention to critical error estimations because of interfering effects and the low accuracy of the measurements.

Equipment

Optical mounting rail; Hg spectral lamp, condenser, adjustable slit, collimator, direct vision prism, micrometer slide mount, achromatic (objective) lens, calcium photo detector; voltage source - pico-ampere measuring unit (SMU), cables.

Experiment and Evaluation

Optical Setup

The setup comprises the basic elements of optical spectroscopy and consists of a lamp and condenser, a collimator, a dispersion prism for spectral decomposition and an objective lens. A special feature of this setup is that the dispersion prism is a so called direct vision prism possessing high dispersion capability. It is formed by three prisms and only slightly deflects the light rays. The desired spectral line is focused into the photo detector by laterally moving the detector using the micrometer slide.

The optical setup must be carefully aligned before starting the measurements. Since the photo currents are relatively low, qualitative and quantitative acceptable results can only be achieved with very well adjusted apparatus. Pay attention to short- and well aligned ray paths and good light intensity and shielding from stray light. Check the adjustments during the experiment.

One can measure the photo current, without applying bias voltage, to determine whether the spectral lines properly pass through the entry slit into the photo detector.

Ensure that the respective line is sharply focused onto the entry slit of the photo detector so that no higher energetic spectral lines can pass through. Tuning to the maximum photo current for the desired spectral line is first done with the aid of the micrometer slide. In addition, the photo detector can be slightly turned with respect to the optical axis to achieve a maximum current.

Attention: Thereafter, check that no light from higher energetic photons pass through the slit onto the photo detector. This is done by a gradual, step by step readjustment of the micrometer slide and turning of the photo detector while reading the photo current.

Electrical Setup

The electrical setup is comparatively simple and consists of a series circuit of photo detector, voltage source and pico-amperemeter incorporated in the SMU.

Photo Detector

A coax cable leading out from the detector housing connects the calcium layer to the sensitive BNC input of the SMU.

The collector electrode is an open platinum ring, both ends of which connect to cables with 4 mm plugs (banana plugs). Both plugs must be connected together for the measurements.

The input at the back of the SMU is the BNC female connector. The output at the back of the SMU is the 4 mm socket at ground potential next to the BNC connector. Both plugs must be connected to this socket. The minus-pole socket is not used.

The polarity of the SMU can be reversed by a switch on the front panel.

Due of the small photo currents, the measurement is very sensitive to stray currents and zero-point errors. Ensure that the coax cable is not moved during the experiment. Also you should move as little as possible in the vicinity of the SMU (changes in capacity lead to display errors!)

The measuring range switches automatically between <1pA and 1000nA.
Attention: The lamps indicate (from bottom to top) the
current range pA , nA , 0.01 µA.

Voltage Source

The poles of the voltage source are labeled with + and - . The polarity of the voltage source can be reversed with switch (+/-).

The applied voltages lie in the range from mV to several volts and can be set with the step switch.

Exercise 1

(Measuring the saturation current and stopping potential as a function of the light intensity of the indigo 436 nm line).

Set a positive bias voltage of maximum 8 V and measure the saturation current and stopping potential of the indigo line λ= 436 nm as a function of the width of the slit up to $I_{Ph} = 0$

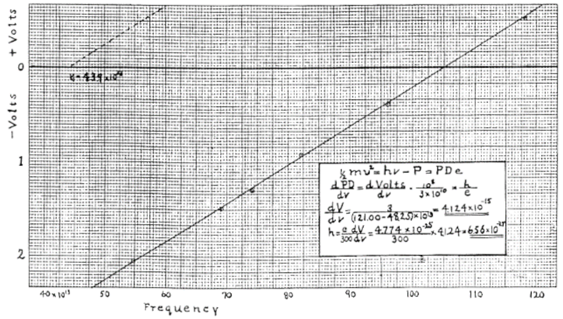
The measurements begin with the slit closed (zero point of the slit setting) and the slit must then be opened in 1/100 mm steps because of the sensitivity of the setup and to remain within the measuring range. This is at the limit of resolution of the slit setting and the settings must be done with special care. When setting and reading the width of the slit, take into account that the scale counts backwards. The results are to be presented graphically with saturation current and stopping potential as a function of the width of the slit.

Exercise 2 (Characteristic Curves)

When measuring the characteristic curves make sure that a sufficient number of measuring points (for example in 100 mV steps) are taken. For each characteristic curve measure the voltage at the zero cross-over point (*$I = 0$*).

For this exercise it is essential that the measuring points are graphically represented during the experiment to check the form of the curve and the reproducibility of the results.

From the characteristic curves note the measured zero cross-over points (*$I = 0$; maximum energy of the electrons*) and the zero cross-over points extrapolated from the linear part of the characteristic curves. Make a graph of the measured and extrapolated zero cross-over points as a function of the frequency of the lines. From the slope of the curves one can then determine the work function U_A and Planck's constant h.



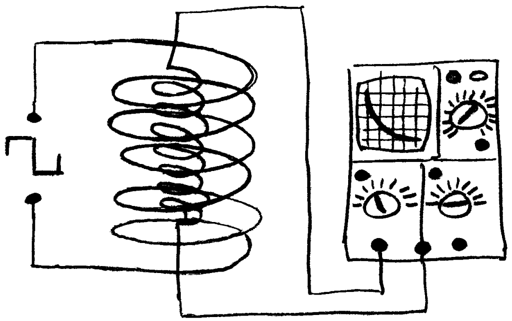
Original Diagram from *R.A.Millikan*; A Direct Determination of Planck's 'h'; Physical Review 7, 355 (1916).

When evaluating and discussing the results, take into account the different intensities of the spectral lines or eventual intensity fluctuations of the lamp respectively.

Main Lines of the Mercury Spectrum

Wave Length, λ		
577.0 / 579.1 nm	Yellow	strong
546.1 nm	Yellow/green	strong
491.6 nm	Blue/green	very weak
435.8 nm	Indigo	strong
407.8 / 404.7 nm	Violet	strong

INDUCTION	GP II



Key Words

Magnetic flux, magnetic induction, induction law. Magnet fields of coils, self induction. *R-L* circuits, transformer.

Ain of the Experiment

Investigation of induction phenomena in time dependent magnetic fields.

Literature

Induction: Tipler, Berkeley Physics Course, Bd. 2, Gerthsen, Kap. 7.3; *R-L*-Kreise: GP-Script „Wechselstromkreise“

Exercises

You are to construct two dc isolated circuits, one to generate a time dependent magnetic field with the aid of a *field coil*, a second with an *induction coil* to detect the induced voltage. Apply a square wave voltage to the field coil; observe the time dependent progression of the induced voltage U_{ind} with an oscilloscope and measure the maximum value of U_{ind}. Calculate the self induction L and compare this quantitatively with the observation.

Apply an alternating current $U=U_0\cos(\omega t)$ to the field coil and measure the root mean square (rms) value of the induction voltage as a function of ω.For this purpose you are to select five different voltage between 100 Hz and 1 kHz.

Measure the rms value of the induction voltage with an alternating current (100 Hz) through the field coil dependent on the orientation of the induction coil, i.e., for different angles between the axes of both coils.

In exercises1 to 3 make a quantitative comparison with the theoretically expected values. Discuss the most probable errors.

Physical Principles

Induction. Michael Faraday (1791-1867) discovered and investigated induction phenomena and formulated the induction law:

(1) $$U_{ind} = -\frac{d\,\phi}{d\,t}$$

U_{ind} is the induced voltage in a closed conducting loop of area A, when the **magnetic flux**

(2) $$\phi = \int_A \boldsymbol{B} \cdot d\,\boldsymbol{A}$$

through the loop varies in time. The unit of magnetic flux is T·m^2 (= V s) = 1 Wb (Weber). In equation (2) the direction of the surface elements dA at each point on the surface is perpendicular to the surface. The scalar product in (2) means that the flux not only depends on the magnitudes but also on the orientation of B and A. There exist two practical methods to generate induced voltages:

changing the magnitude of B with time and keeping A fixed (e.g.: transformer, self induction)

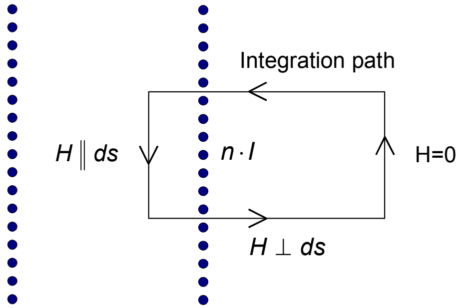
changing the relative orientation of B or A with time (e.g.: generator).

In the present experiment the change of the magnetic with time is used. This is generated by a switching process or alternating current.

The magnetic field of the coil. The magnetic field inside a long coil (infinitely long in the limiting case) can be calculated using **Amper’s Law**

(3) $$\oint \mathbf{H} \cdot \mathrm{d}\mathbf{s} = \int \boldsymbol{j} \cdot \mathrm{d}\mathbf{A} = \boldsymbol{I}$$

by selecting a suitable integration path (see diagram below). For a coil section of length ℓ with n windings we have from (3)



(4) $$H \cdot \ell = n \cdot I$$ or

$$H = \frac{n}{\ell} \cdot I \; .$$

For a real coil of finite length ℓ and radius r, the magnetic field becomes weaker near the ends; we then have for a point on the coil axis a distance a from the end

(5) $$H = \frac{n}{\ell} \cdot I \cdot \frac{1}{2} \cdot \left(\cos\alpha_1 + \cos\alpha_2\right)$$

.

If one expresses the angle by the length of the sides of a right-angle triangle and writes the **magnetic induction** as $B=\mu_0 \mu H$ (magnetic field constant μ_0 = $4\pi\cdot 10^{-7}$ Vs/Am, permeability in vacuum μ=1) one gets:

(6)

$$B(a) = \mu_0 \cdot H = \mu_0 \cdot \frac{n}{\ell} \cdot I \cdot \left[\frac{1}{2} \left\{ \frac{a}{\sqrt{r^2 + a^2}} + \frac{\ell - a}{\sqrt{r^2 + (\ell - a)^2}} \right\} \right]$$

(6a) $$= \mu_0 \cdot \frac{n}{\ell} \cdot I \cdot F(a) \; .$$

The geometry factor $F(a)$ describes the deviation from the ideal case of an infinitely long coil. For the experiment it is convenient to calculate a correction factor

$$\overline{F}(s) = \frac{1}{s} \cdot \int_{\frac{\ell - s}{2}}^{\frac{\ell + s}{2}} F(a) \cdot d\,a \; .$$

This allows the average (mean) induction to be expressed as

$$\overline{B}(s) = \mu_0 \cdot \frac{n}{\ell} \cdot I \cdot \overline{F}(s)$$

in an induction coil of length s placed symmetric to the middle of a field coil. For the integral, the following relationship applies:

$$\int \frac{a}{\sqrt{r^2 + a^2}} d\,a = \sqrt{r^2 + a^2} + C$$

Self induction. A wire-configuration which generates a magnetic field represents, at the same time, an antenna in which a voltage is induced because of the self produced change in flux. Since the magnetic field is proportional to the current and a change in flux can only be produced by a change in field (with a fixed geometric arrangement), there exists a simple relationship of the form

(9) $$U_{ind} = -L \cdot \frac{d\,I}{d\,t} \; .$$

The proportionality factor L is called the **self induction** (**inductivity**) of the arrangement. For a long coil with n windings we have according to (1) and (6):

(10)

$$U_{ind} = -n \cdot \frac{d\,\phi}{d\,t} = -\mu_0 \cdot \frac{n^2}{\ell} \cdot A \cdot \frac{d\,I}{d\,t} = -\mu_0 \cdot n^2 \cdot \frac{A}{\ell} \cdot \frac{d\,I}{d\,t}$$

By comparing with (9) it follows that:

(11) $$L = \mu_0 \cdot n^2 \cdot \frac{A}{\ell}$$

If one takes into account the correction factor (7) for the intrinsic magnetic field of a coil of finite length, one gets

(12) $$L = \mu_0 \cdot n^2 \cdot \frac{A}{\ell} \cdot \overline{F}(s) \; .$$

On/Off switching processes. Exercise 1 deals with induction processes appearing in a coil when the polarity of the voltage is reversed; this is accomplished here by applying a square-wave voltage. The equivalent circuit for the field coil can be taken as a series circuit consisting of an ideal coil of inductivity L and a loss resistance R_L. The voltage source (function generator) has an open circuit voltage U_G and an internal resistance of R_i.

The description of the current in the coil is given by a differential equation after applying **Kirchhoff's rules**. The solution, taking into account the boundary conditions is

(13)

$$I(t) = \frac{U_G}{R_L + R_i} \cdot \left[1 - 2 \cdot e^{-\frac{R_L + R_i}{L} \cdot t} \right] .$$

The switchover point from - U_G to + U_G was selected as t = 0 and the assumption made that the half period of the square-wave voltage is long compared to the decay time.

Induction due to alternating fields. If one generates an alternating magnetic in a field coil by applying a periodic current of the form

(14) $$I = I_0 \cos \omega\, t ,$$

the voltage induced in the coil grows proportional to the amplitude of the current I_0 and to the frequency ω.

Presentation of the Physical Principles

(As part of preparation for the report.)
Discussion of the induction law and the relationship between H and B.
Solution of the differential equation for the switching process.
Notation and discussion of the other equations needed for the measurements.

Apparatus and Equipment

Large coil with field windings (1000 windings) and induction windings (125 and 250 windings) on a common body. Exact coil data is found in the lab bench script.
Induction coil with adjustable orientation (tilt coil)
Function generator
Oscilloscope and multimeter.

Experiment and Evaluation

Operation of the equipment is self-explanatory.

<u>To Exercise 1</u>:
Large coil with field windings n = 1000 and induction windings n = 500. The terminals of the windings are located on the front ends of the coil next to each other.
The length of the windings is (190±1) mm.
The field windings are connected directly to the function generator (square-wave function, $\nu \approx$ 100 Hz, amplitude $\approx$ 3 scale divisions).

The oscilloscope is operated in dual mode: at input I the generator terminal voltage = field coil voltage and at input II the voltage in the induction coil is checked. Select the settings at the oscilloscope as well as the amplitude and frequency at the function generator such that the voltage-time curve sufficiently decays. Make optimal use of the screen with respect to resolution and accuracy.
<u>Attention:</u> when making sweep- and voltage settings take note that rotary knobs on the oscilloscope are set to <u>calibrate</u> (*cal*)! Make a qualitative sketch of the run of the curve in the report book.

With the decay curves $U_{Feld}(t)$ and $U_{ind}(t)$ we are only interested in the voltages with respect to $U_\infty = U(\infty)$.To compare them with one another and for the quantitative evaluation of the decay behaviour, both curves can be placed over each other by suitable settings of the y-deflection at the oscilloscope. Record the data in a table by reading a reasonable number of $U(t)$ values from the cm scale of the screen.
For further evaluation, one requires the open-circuit voltage of the generator U_G (field coil terminals disconnected) as well as the initial- and final voltage $U_0 = U(0)$ and U_∞ at the field coil. They are deter-

mined by difference measurements between the positive and negative half-periods to take into account the arbitrary position of the zero point of the generator (dc offset) and the oscilloscope.
Carefully measure and record the actual original measurement data (screen readings in cm and the associated oscilloscope settings), since from experience errors easily set in here. The dc resistance of the field coil $\boldsymbol{R_L}$ is directly measured with a multimeter (*METRA Hit 12S*).
Evaluation is made by plotting the measured voltage-time curve on half-log paper. $U_{ind}(0)$ is found by extrapolation of the measured values and L together with R_G+R_L (see below) from the gradient of the fitted line. The results are to be qualitatively and quantitatively with the theoretical expectations.
The internal resistance of the function generator R_G can be determined from the observed terminal voltage U_0 for $t = 0$ and U_∞ for $t > \infty$ with the help of (15):

(15) $$U_{0,\infty} = U_G - R_G \cdot I_{0,\infty}$$

(15a) $$= U_G - (R_G + R_L) I_{0,\infty} + R_L \, I_{0,\infty}$$

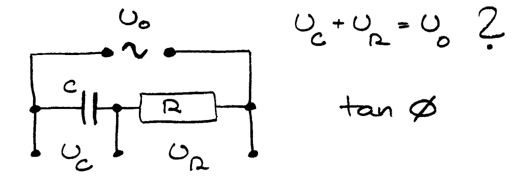
Error Calculations. In the propagation of errors it is advantageous to supplement the generator resistance R_G in (15) with R_L (see 15a) and from this directly calculate "closed" expressions for the required sums of the internal resistances R_G+R_L. The relation (15a) is to be evaluated for both cases (U_0, U_∞).

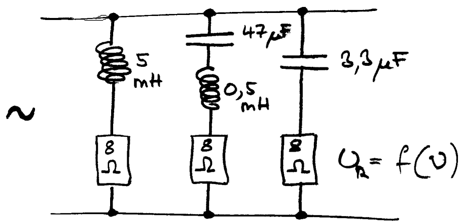
To Exercise 2:
The large coil with n = 1000 field windings and n = 500 induction windings is used.
The measurement of the rms values is performed with the multimeter *Voltcraft 5050-DB* (current measurement) and a milli-voltmeter *Grundig MV 40* (voltage measurement).
The field current should be about 30 mA.
Attention: The field current changes with frequency. (Why?) The frequency dependence of the induction voltage $U_{ind}(\omega)$ shall be measured in a range from 100 Hz to about 1000 Hz. Plot the results and discuss them quantitatively. Estimate the largest most probable error of the measurement.

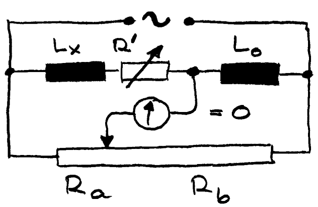
To Exercise 3:
Large coil n = 1000 as field coil and an additional coil with variable orientation (**tilt coil**) as the Induction coil. Make measurements of $\boldsymbol{U_{ind}}(\alpha)$ at 100 Hz. Evaluation is made by plotting $\boldsymbol{U_{ind}}$ against cos α. Discuss the results as well as the most probable error sources quantitatively; draw realistic error bars in the plots.

Supplementary Questions
The following questions are of significance for the physics of transformers.
How does the induction voltage depend on the frequency when in the field coil the voltage is held constant and not the current?
In the oscilloscope observe the the phase between field- and induction voltage as a function of the frequency.
Why are the voltages in-phase at sufficiently high frequencies? Connect an addition (low resistance) load to the induction coil. Why does a phase shift now appear even at high frequencies? (Note: consider the energy.)

Alternating Current Circuits | GP II







Key Words

Alternating voltage and -current, impedance (alternating current resistance). Alternating current power.

Resistance operators and complex representation. Series- and parallel oscillating (resonant) circuits; filters (high-pass, low-pass); Equivalent circuits.

Aims of the Experiment

Investigations of resistors, capacitors and coils and their combinations in alternating current circuits.

Circuit description using complex resistance operators and equivalent circuits.

Literature

Standard literature

Exercises

1. Assembling an *RC* circuit. Setting the characteristic frequency with $U_R = U_C$. Measuring the generator- and partial voltages and determining the phase shift. Independent measurement of *R* and *C* with a multimeter and comparing the observed results on the RC circuit with the theoretical expectations.
2. Measuring the frequency response U_R/U_G (load voltage to generator voltage) on an audio frequency filter (three-way filter with *RL low-pass*, *RCL band-pass* and *RC high-pass*) and comparing the results with theory by making independent measurements of the resistances, capacitances and inductances with digital multimeters.
3. Measuring the inductance and the dissipation resistance of one of both coils from exercise 2 with a resistance bridge and comparing the results with independent measurements (digital multimeter) of *L* and the dc resistance *R* of the coil.

Physical Principles

Current and Voltage at *R*, *C* and *L*

The most important passive devices in electric circuits are resistors, capacitors and coils. Their model behaviour are characterized by the resistance *R*, the capacitance *C* and the self induction coefficient *L* (inductance)with the following relationships between voltage and current:

(1.1) Resistance (*R*): $$U_R = -R\, I_R$$

(1.2) Capacitance (*C*): $$I_C = -C \frac{d\, U_C}{d\, t}$$

(1.3) Inductance (*L*): $$U_L = -L \frac{d\, I_L}{d\, t}$$

R, *C* and *L* themselves are positive definite quantities fort he quantitative description of specific properties of models and the Orientation of voltage and current to each other in (1.1-3) must be taken into account by a minus sign. The resistance of a conductor causes an opposing voltage (voltage drop). In a capacitor, a positive current causes a drop in voltage and vice versa. A change in current in a coil induces a voltage opposing this change which endeavours to maintain the current. In the literature this discussion is often only carried on for coils and the minus sign in 1.3 is then explicitly referred to as *Lenz's Rule*.

Alternating Voltages at R, C and L and Impedance

Time varying alternating sine or cosine voltages or currents are referred to as are referred to as alternating voltages or alternating currents:

(2) $$U(t) = U_0 \cdot \cos(\omega\, t + \varphi_1)$$

or $$I(t) = I_0 \cdot \cos(\omega\, t + \varphi_2)$$

Alternating voltages in *R*, *C* and *L* circuits produce alternating currents of the same frequency (and vice versa), however, phase shifts appear between voltage and current. Without loss of generality, one can select the phase of the voltage as the reference phase and set it to zero (φ_1=0).

The *impedance Z* (alternating current resistance) of a device or circuit is denoted as the ratio of current- to-voltage amplitude:

(3) $$Z = \frac{U_0}{I_0}$$

Inserting (2) in (1.1-3) gives for the impedance and phase shift of the current to voltage at *R*, *C* and *L*:

(4.1) $$Z_R = R \quad \text{and} \quad \varphi = \pi$$

(At the resistor, the phases of voltage and current oppose each other and the product of voltage times current is always negative corresponding to the sign convention for work or power dissipated. In the literature the sign problem is not always handled consequently, and one also finds the statement that current and voltage at the resistor are "in phase", i.e., equal in phase)

(4.2) $$Z_C = \frac{1}{\omega \cdot C} \quad \text{and} \quad \varphi = -\pi/2$$

(At the capacitor the “current leads the voltage".)

(4.3) $$Z_L = \omega \cdot L \quad \text{and} \quad \varphi = +\pi/2$$

(At the coil the "the current trails the voltage".)

Alternating Current Networks

The impedances and phases of combinations of *R*, *C* and *L* can be calculated analogously to the dc case by using complex resistance operators **Z** of the devices, whereby, for a

series circuit:

(5.1) $$Z = \left|\sum \mathbf{Z}_i\right| \quad \text{and}$$

$$\tan\varphi = -\frac{\operatorname{Im}\left(\sum \mathbf{Z}_i\right)}{\operatorname{Re}\left(\sum \mathbf{Z}_i\right)}$$

and for a parallel circuit:

(5.2) $$\frac{1}{Z} = \left|\sum \frac{1}{\mathbf{Z}_i}\right| \quad \text{and}$$

$$\tan\varphi = \frac{\operatorname{Im}\left(\sum \frac{1}{\mathbf{Z}_i}\right)}{\operatorname{Re}\left(\sum \frac{1}{\mathbf{Z}_i}\right)}$$

A detailed presentation of the formalism of ac operators is found in annex IV *WECHSELSTROMOPERATOREN* of the lab script.

Alternating Current Power

In the complex representation (see *WECHSELSTROMOPERATOREN*), the ac power is given by the product of the real parts of **U** and **I**.

(6)

$$P = \mathrm{Re}(\mathbf{U}) \cdot \mathrm{Re}(\mathbf{I}) = \frac{1}{2}\left(\mathbf{U}+\mathbf{U}^*\right) \cdot \frac{1}{2}\left(\mathbf{I}+\mathbf{I}^*\right)$$

Setting an alternating voltage and a (phase shifted) alternating current for U and I respectively, one obtains for the time averaged power due to the vanishing periodic terms

(7)

$$\overline{P} = \frac{1}{4}\left[U_0\, I_0 \left(e^{i\varphi}+e^{-i\varphi}\right)\right] = \frac{1}{2} U_0\, I_0 \cos\varphi = U_{rms}\, I_{rms} \cos\varphi$$

where so called root mean square (rms) values were introduced as convenient measurement- and operation quantities, whose products, except for the phase factor, directly reproduce the average power. For alternating voltages and –currents we have according to (7):

(8) $$U_{rms} = U_0 / \sqrt{2}$$ or

$$I_{rms} = I_0 / \sqrt{2}$$

If in particular $R = 0$, it follows that $\varphi = \pm\,\pi/2$ and $\cos\varphi = 0$, i.e., the ac circuit, on average, does not draw power from the voltage source even though a current flows and voltages build up at C and L. Only R draws power from the circuit in the form of *heat dissipation*.

Power Losses

At ideal capacitors and coils alone, or at respective combinations with $R = 0$ or $R = \infty$, $\varphi = \pm\,\pi/2$ and $\cos\varphi = 0$, i.e., the ac power ($U_{rms} I_{rms} \cos\varphi$) is zero averaged over time, and the ac circuit does not draw power from the voltage source even though a current flows and voltages build up at C and L. Only R draws power from the circuit in the form of *heat* (loss resistance)

In contrast, real capacitors and coils are not loss-less and cannot be described by C and L alone but must be described by an equivalent circuit with additional loss resistance R.

The main reasons for power losses in coils are the resistance of the wire, eddy current losses in conducting material close to the coil, and remagnetization losses in coils with iron- or ferrite cores. Losses in capacitors result from the finite conductance of the dielectric and from dielectric losses brought about by a reversal of polarization, corresponding to remagnetization losses in ferromagnets. Capacitor losses are comparatively small and can be virtually neglected in most cases.

The *loss factor d* of a coil or capacitor is specified by the reciprocal of the tangent of the phase angle, which describes the ratio of the loss resistance to pure capacitive or inductive resistance:

(9) $$d = \frac{1}{\tan\varphi}$$

Equivalent Circuit Diagrams for real Capacitors and Coils

A real coil exhibits an experimental impedance Z (voltage-to-current ratio) and a phase angle φ, which must be reproduced by a model approach. The simplest models are series- or parallel combinations of L and R:

Series equivalent combination (R_s and L_s)

(10.1) $$Z = \sqrt{R_S^2 + \left(\omega\, L_S\right)^2}$$ and

$$\tan\varphi = -\frac{\omega\, L_s}{R_s}$$

Parallel equivalent combination (R_p and L_p)

(10.2) $$\frac{1}{Z} = \sqrt{\frac{1}{R_p^2} + \frac{1}{\left(\omega\, L_p\right)^2}}$$ and

$$\tan\varphi = -\frac{R_p}{\omega\, L_p}$$

Various loss resistances R_s or R_p, as well as various inductances L_s and L_p also belong to the two equivalent circuit diagrams which should reflect the observed physical circumstances (Z, φ).Hence we can speak of a certain inductivity of a coil only in connection with a certain equivalent circuit. Even when a series equivalent circuit with coils and a parallel equivalent circuit with capacitors give a basically better view of the situation, both equivalent circuits are effectively the same and can be selected in their application according to the simplest treatment of the problem.

In practice, it may be necessary to take into account the frequency dependency of Z and φ, which are based on the frequency dependency of the capacitive and inductive behaviour of conductor configurations and their losses. Their exact representation can lead to very complex equivalent circuits.

Filters (High-pass, Band-pass, Low-pass)

R-C, *R-L*- and *R-C-L* devices represent frequency dependent voltage dividers (filters) and their frequency response can be calculated with the help of resistance operators. A *R-C-L* band-pass filter is an oscillating system with an impedance minimum at resonance (series oscillating circuit; filter circuit) or an impedance maximum at resonance (parallel oscillating circuit; trap circuit).

Alternating Current (AC) Bridge

An ac bridge corresponds to a *Wheatstone bridge* and allows comparative measurements of capacitances and inductances (see diagram on the title page).

The balancing condition is, in the case of alternating currents, the agreement of the complex impedances, i.e., the simultaneous agreement of the partial voltages in both arms of the bridge and the phases (the loss factors of the capacitors or coils being compared must agree).

(11)

$$\frac{L_x}{L_0} = \frac{R_a}{R_b} \quad \text{and (!)} \quad \frac{R_L}{R_0 + R'} = \frac{R_a}{R_b}$$

Since this is generally not the case, an additional *phase balancing resistor R*' must be employed to shift the phase of that device with the lowest loss factor (see circuit diagram on the title page; here it is assumed that the coil to be investigated has the lowest loss factor and must be complemented by a phase balancing resistor).

Presentation of the Physical Principles

(as preparation for part of the report): Short summary of the formalism of complex resistance operators for the description of impedance and phase.

Calculation and discussion of the equations needed for the exercises below.

Apparatus and Equipment

Plug-in circuit board with various circuit elements (coils, capacitors, resistors).
10 turn potentiometer (balancing potentiometer), 10 turn variable resistor (phase balancing resistor) and comparison coil of known inductivity to set-up the bridge.
Function generator, multimeter, oscilloscope.

Experiment and Evaluation

Operation of the measuring instruments and the function generator follows largely from the labelling on the instruments. Please note the additional information in the lab bench script .

To Exercise 1:

Assembling a *R*-*C* circuit using a 1 kΩ resistor, a 1 μF capacitor and a function generator as driver (see circuit diagram on the title page). Setting the frequency at which the partial voltages at the resistor and capacitor agree (transfer frequency). Measuring the partial voltages and phase shift with the oscilloscope and comparing the results with the theoretical expectations.

Attention: With the function generator and the oscilloscope one input pole of each unit is grounded and the out- and inputs cannot be freely occupied. For one thing, one must ensure that no short to ground occurs. On the other hand, measurement circumstances arise, such as the interchange of poles with respect to a common “sense of rotation”, which must be taken into account when interpreting the results.

To Exercise 2:

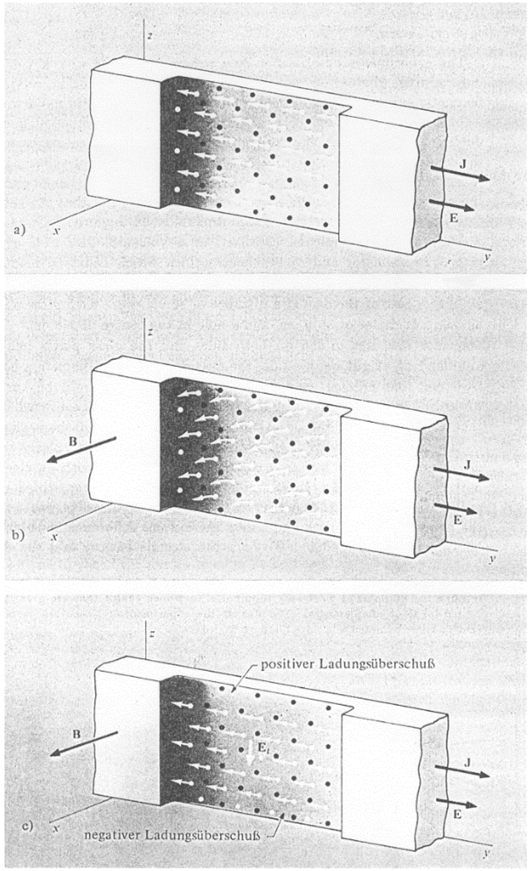
Assembling a filter network (*frequency filter*) corresponding to the circuit diagram on the title page. The 8 Ω resistors represent loads (loud speakers). The frequency response is determined by measuring the generator voltage and the voltage across the load resistors as a function of frequency in the range 50 Hz to 20 kHz (in suitable gradations of measured values). Take note that the generator voltage does not remain constant because of the comparatively high internal resistance of the function generator (50 Ω). Evaluation is made by plotting the measured voltage ratios (load/generator) as points against the frequency and fitting the points with a curve to represent the theoretical progression. The values are to be plotted on log-log paper.

To Exercise 3:

Assembling the bridge as shown on the title page.

L_x and L_o are the physical coils used in the circuit, which must be represented by a series circuit of an inductance ($L_{x,o}$) and a loss resistance ($R_{x,o}$) in the equivalent circuit diagram. A comparator coil L_o of 1.5 mH is available on a plug-in board.

Voltage balancing is made with a 1 kΩ rotary potentiometer and phase balancing with a 50 Ω variable resistor. Balancing is checked with the multimeter (PHILIPS PM 2505) and performed by successively changing voltage- and phase balancing. In effect, the bridge can only be balanced to a remaining rest voltage since the function generator produces, harmonics (distortions) which cannot be simultaneously balanced in phase.



Subjects and Terminology

Electric field; *Coulomb force*. Magnetic field; *Lorentz force*.

Band theory; Semiconductors, Valence band, Conduction band, Band gap; Self conduction.

Doping; negative and positive charge carriers (electrons and holes); Impurity conduction.

HALL EFFECT	GP II

Normal and anomalous *Hall Effect*.

Special Aims of the Experiment

Observing the *Hall Effect* as the common action of electric- and magnetic fields on moving charge carriers in solids.

Investigation of the conduction mechanism and determination of the type and concentration of the charge carriers in metals and semiconductors.

Literature

The physical principles of the *Hall effect* und of the electrical properties of semiconductors are presented in basic physics text books. In particular, the following are recommended:

Skript *HALBLEITER* im allgemeinen Teil dieser Praktikumsanleitung.

/1/ E.M. Purcell; Berkeley Physik Kurs Band 2, Elektrizität und Magnetismus; Stichwort Hall-Effekt (gute Einführung).

/2/ Gerthsen·Kneser·Vogel, Physik; Springer-Verlag; Stichwort Halbleiter (gute Behandlung der Eigenschaften von Halbleitern, der Bandlücke und der Effekte von Störstellen).

/3/ R.W. Pohl; Einführung in die Physik, Zweiter Band, Elektrizitätslehre, 21. Auflage; Springer-Verlag; Literaturwerte der *Hall-Konstanten*.

Exercises

1. Observing the *Hall Effect* on germanium (n- or p-Ge) as a function of control current and magnetic field. Calculation of the *Hall constant* of germanium. Determining the type and concentration of the charge carriers.
2. Investigation of the temperature dependence of the *Hall voltage* in germanium and calculation of the band gap.
3. Common exercise for all students in a group: Observation of the *Hall Effect* in Cu and Zn. Estimating the *Hall constants* as well as the type and concentration of the charge carriers.

Physical Principles

In addition to the fundamentals found in the literature, two aspects should be pointed out, which are often neglected or not presented:

- The charge carriers contributing to the *Hall Effect* can be negative (electrons) or also positive (holes), depending on the material. This leads to different signs for the *Hall voltage* and the *Hall constant R* :
- In doped semiconductors at low temperatures, the charge carriers of the impurities are responsible for conduction (impurity conduction). In addition, electrons from the host semiconductor (host lattice) are lifted from the valence band into the conduction band through thermal excitation.
  At sufficiently high temperatures this intrinsic conduction dominates over the impurity conduction. In the region of intrinsic conductivity, however, practically only electrons contribute to the *Hall Effect* due to the different mobility's of the electrons and holes.

Presentation of the Physical Principles

(as preparation for part of the report): Presentation of the *Hall Effect* and derivation of the *Hall voltage*. Short statement and discussion of the temperature dependence of the *Hall voltage*.

Apparatus and Equipment

Printed circuit (pc) board with metal probes (n-Ge, p-Ge, Cu and Zn) to determine the *Hall Effect* and various additional components for the experimental investigation.

Electromagnet formed by two coils and a U-shaped iron core with pole shoes. Various power supply units. Digital multimeter and a microvolt meter for the measurement of currents and voltages.

Experiment and Evaluation

Please handle the pc board with care. Ge crystals are brittle and sensitive to breakage and a fracture of the crystal by bending makes the pc board unusable (cost of a pc board ≈ 1000 €). Proceed carefully when placing the pc board in the magnet and connecting the cables. Compensate the pressure used to plug in and pull out the banana plugs by firmly holding the pc board in the area of the sockets.

Circuitry

The required connections are inscribed on the pc board.

Control Current

The current through the probe is called the *control current*. In semiconductors it is important to hold the control current constant during the experiment since the resistance is strongly dependent on the temperature.

For this purpose, the pc board incorporates a constant current source which delivers a constant current of about 30 mA independent of the external voltage (12 to 30 V). The constant current device is connected through outer socket of both (-) sockets for the control current. The constant current device is used to measure the *Hall voltage* as a function of the magnetic field sine it allows alignment of the resistance-voltage drop (see section on *Hall Voltage* below).

The constant current device cannot be used to measure the *Hall voltage* as a function of the control current (inner of both (-) sockets for the control current) and the power supply unit for the control current (*Voltcraft*) must be used as the constant current device (voltage limiter to maximum and the current adjusted with the current limiter). Proceed with care so that the maximum value of 50 mA for the control current is not exceeded (use the power supply unit with 200 mA limiting current).

With the metal probes, the control current is directly connected with a maximum current of 20 A.

Hall Voltage

The *Hall voltage* is tapped transverse to the probe dimension. Since both taps for the *Hall voltage* cannot lie directly opposite each other due to manufacturing constraints, a resistance voltage drop appears in addition to the *Hall voltage*. To compensate for this the voltage on one side is tapped slightly "above" and "below" the opposite point and aligned with the aid of a potentiometer. This zero-point correction of the *Hall voltage* is made in each case without applying the magnetic field.

This circuit is only effective in semiconductor probes when the constant current device (see above) is used. When the control current is directly connected to measure the *Hall voltage* as a function of the control current, the zero-point must be determined and the measured values accordingly computed.

Heater and Thermocouple

The probe (pc board) is equipped with a heater and thermocouple to investigate the temperature dependence of the *Hall voltage*.

Magnet

When connecting the magnet, pay attention to the correct series connection of the coils. A power supply unit (*Voltcraft* 0...30 V, 2 A) delivers the current for the magnet.

The lab bench script contains a calibration curve for the magnetic field in the center of the pole shoes as a function of magnetic current; the magnetic current must not exceed 2 A .

Attention: To check the polarity of *Hall voltage* with respect to the sign of the charge carriers one must very carefully take into account and document the orientation of all experimental quantities (magnetic field from the sense of the windings of the coil and the direction of the magnetic current, direction of the control current or polarity of the connection of the control current source, polarity of the *Hall voltage* or polarity of the connection of the voltage measuring instrument). A sketch of the experimental set-up is essential to clearly show the orientations and connections of the measuring instruments! When recording the values of the *Hall voltage* the sign must always be given!

To Exercise 1:

A second *Voltcraft* power supply unit ("200-mA" unit) is used as the control current source.

To investigate the *Hall voltage* as a function of the magnetic field one uses the constant current source on the pc board (outer (-) sockets) and the potentiometer (rotating pin) to align the resistance-voltage drop.

For the measurement as a function of the control current, the inner "direct" sockets must be connected; see the general information concerning the *Hall voltage* above.

- The control current must not exceed 50 mA.

The alignment potentiometer for the resistance-voltage drop is inhibited when the constant current source is not connected, and to correct the voltage values, the voltage drop for each control current value must be measured with the magnetic field switched off (switch off the magnetic current power supply at the mains switch) and the measured value computed with the correction.

The control current and the *Hall voltage* are measured with two digital multimeters.

To Exercise 2:

In addition to the circuitry for exercise 1, the heating current source and a further multimeter is connected to measure the thermo-voltage and the *Hall voltage* as a function of temperature.

The heater is operated with 6 V ac and the heating current is then about 5 A. The temperature coefficient of the thermocouple is 40 μV/K (temperature difference to room temperature).

- The temperature of the pc board must not exceed 150 °C corresponding to 5 mV thermo-voltage!

The heater (step transformer) should be used intermittently (momentary switch-on with pauses) so that the rate of heating is not too fast in order to ensure safe control and reliable recordings of the measured values.

The thermo-voltage can (still) be measured with the digital multimeters. The measuring accuracy is clearly limited by the low resolution (0.1 mV corresponding to 2.5 K) but may be considered as sufficient within the scope of the other measurement conditions (e.g., temperature gradients) and the aims of the experiment.

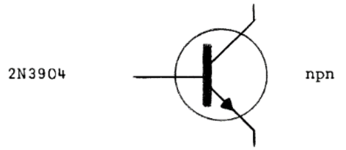
With a suitable logarithmic plot (see equation 1 in annex V „*HALBLEITER*“ in the lab script) the *Hall voltage* shows, in the high temperature range (intrinsic conduction), the expected linear progression from which one can determine the band gap ΔE .

To Exercise 3

A 10 A power supply unit is available to measure the *Hall Effect* on metal probes (copper and zinc) which require high control currents. The *Hall voltage,* however, remains comparatively small and must be measured with a sensitive microvolt meter (*KNICK*) or for comparison purposes with the digital multimeter (HP 3457A).

The measurements on the metal probes have only qualitative character with respect to the sign because of experimental difficulties (foils with large tolerances as thin probes; however, very small values for the *Hall voltage*). Since only one power supply unit for the high currents and one microvolt meter are available, the measurements shall be made as a common exercise for the whole group but reported and evaluated for each pair of students working together.

TRANSISTOR	GP II



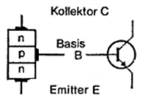


Abb. 9 Der Transistor mit seinen Anschlüssen

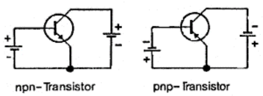


Abb. 10 Schaltzeichen für npn- und pnp-Transistoren mit richtig gepolten Batterien

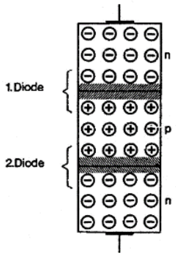


Abb. 4 Grenzschichten im Transistor ohne äußere Spannung

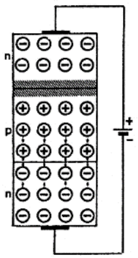


Abb. 5 Grenzschichten im Transistor mit einer äußeren Spannungsquelle

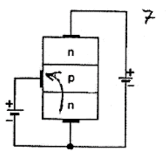
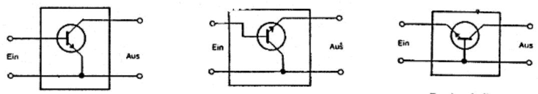


Abb. 6 Richtiger Anschluß der beid Spannungsquellen am npn-Transistor .



Emitterschaltung Kollektorschaltung Basisschaltung

Key Words

Semiconductor; Band model and conductivity; p-n junction, semiconductor diode; Transistor; amplifier circuits.

Aim of the Experiment

Introduction to the basic principles of transistors and elementary circuit techniques.

Literature

/1/ U.Tietz - Ch. Schenk; Halbleiterschaltungstechnik; Springer-Verlag

/2/ K.-H. Rohe; Elektronik für Physiker; Teubner Studienbücher

/3/ J.Pütz; Einführung in die Elektronik; Fischer-Taschenbuch-Verlag

HALBLEITER in annex V of the GPII script

Exercises

1. Recording and construction of the (static) characteristic curves of a npn-transistors (2N3904) for an operating voltage (supply voltage) of 12 V. Determining current amplification for the static case.

   Design an amplifier stage with negative feedback for stabilization.

2. Dimensioning the circuit: Estimating the working resistance of the base series resistance.
3. Experimental check of the collector resistance curves by varying the base series resistance and determining the current amplification.
4. Amplifying an input ac voltage as a signal. Measuring the voltage amplification and comparing the result with the theoretical expectation.

Physical Principles

Principles of Operation of a Transistor

Refer to the literature and *HALBLEITER* in annex V of the GPII script

Transistor Circuits

A transistor can be operated under different circuit configurations. Depending on whether the emitter (E), base (B) or collector (C) lies on the common reference potential of the circuit (ground), one differentiates between an *emitter-*, *base-* or *collector circuit*. In the scope of this experiment, only the emitter circuit will be handled.

Characteristic Parameters and Characteristic Curves

A transistor is specified by three currents and three voltages: I_B, I_C, I_E and U_{EC}, U_{BC} and U_{EB}. The sum of the three currents is zero, whereby the current flowing into the transistor is taken as positive and the out flowing current as negative:

(1) $$I_B + I_C + I_E = 0$$

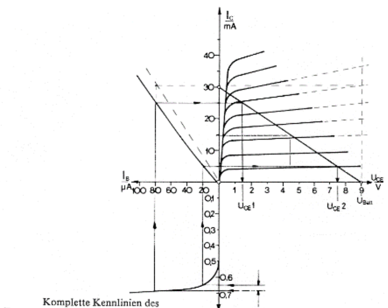
Correspondingly, for the voltage one has:

(2) $$U_{EC} = U_{BC} + U_{EB}$$

From the six variables, two are always dependent on the other four as seen in (1) and (2) and can be expressed in terms of these.

In the emitter circuit, the transmitter can be considered as a current amplifier in which a small change in base current ΔI_B causes in a large change in collector current ΔI_C . The current amplification factor and other parameters of the transistor or circuit can be taken from the *four-quadrant characteristic curves,* which show the interdependence of the four independent variables among one another. From the characteristic curves one can read:

- in the first quadrant the output resistance (U_{EC}/I_C),
- in the second quadrant the current amplification (I_C/I_B),
- in the third quadrant the input resistance (U_{EB}/I_B) und
- in the fourth quadrant the reverse voltage transfer ratio (U_{EB}/U_{EC}).



Komplette Kennlinien des Transistors und die des Kollektorwiderstandes

From the characteristic curves of the first quadrant it is clear that the value of the collector current only depends in a small way on the emitter-collector voltage. This is a useful property since such a voltage drop at the load only leads to a small reverse bias of amplification.

The second quadrant reproduces the current amplification ß which is practically constant over a wide range:

(3) $$\beta = \frac{I_C}{I_B} \quad \text{or} \quad = \frac{\Delta I_C}{\Delta I_B}$$

The third quadrant essentially corresponds to a "normal" diode characteristic in the direction of current flow; here the emitter-basis diode.

The fourth quadrant describes how a change in the emitter-collector voltage affects the base voltage (reverse voltage transfer, *Punch-through*).

Power Hyperbola

The current through the transistor, together with the non-vanishing internal resistance, leads to power loss and self heating which, at large values, can damage the transistor. The maximum permissible power loss $U_{EC} \cdot I_C$ (neglecting base power) can be plotted as a *power hyperbola* in the field for the output characteristics (1. quadrant).

Working Point Resistance and Voltage Amplification

With a given supply voltage U_o in the collector circuit, the collector current can be limited by a resistor R_A (working resistor). Depending on the current, a part of the supply voltage drops across the resistor so that the collector voltage U_{CE} is also limited. Since the voltage drop is dependent on the collector current, this boundary forms a falling straight line (*collector-resistance line*) in the field for the output characteristics and is fixed by the points $I_C = U_o/R_A$ for $U_{EC} = 0$ (short circuit case) and $U_{EC} = U_o$ for $I_C = 0$ (blocking).

The working resistor must be selected so that the resistance line does not cut the power hyperbola.

The emitter circuit with a working resistor represents a simple voltage amplifier. Due to the working resistor, a change in voltage occurs at the collector which is proportional to the change in current. The ratio $\Delta U_{EB}/\Delta U_{EC}$ is termed the *voltage amplification* v:

(4) $$v = \frac{\Delta U_{EC}}{\Delta U_{EB}} = \frac{R_A \cdot \Delta I_C}{\Delta U_{EB}} \left| \cdot \frac{\Delta I_B}{\Delta I_B} \right. = \frac{\beta \cdot R_A}{r_{EB}}$$

where r_{EB} is the *differential input resistance* $\Delta U_{EB}/\Delta I_B$.

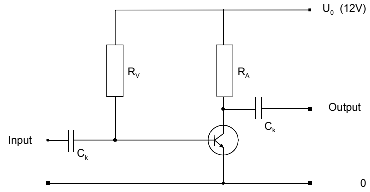
Such a simple amplifier works inverting, i.e., an increase in voltage or current at the input acts to lower the voltage at the output due to the increasing collector current and the larger voltage drop across the working resistor.

Working Point

A transistor only amplifies in the range of positive emitter-base currents. In order to transmit ac signals undistorted, a positive dc signal must be superimposed on the base. The associated point in the fields of the characteristic curves is called the *working point*. The working points are often selected as the half maximum permissible collector currents or half the supply voltages. The working point or the associated base quiescent current can be set by a so called base dropping resistor or a voltage divider ahead of the supply voltage.

In applications in amplifier circuits, setting up a working point has the disadvantage, that even in quiescent current operation without a signal at the input of the circuit, relatively high currents with power losses flow in the collector circuit (»Class-A-amplifier« in HiFi technology).

The following diagram represents a simple amplifier stage with base dropping resistor R_V, a working point resistor R_A and two coupling capacitors C_K:



Static and Dynamic Characteristic Curves

The characteristic curves described above under the assumption of freely specified variables, e.g., the collector-emitter voltage U_{EC}, are called *static characteristics*. However, the inclusion of a working point resistor results in considerable feedback of the dependent quantity (here the collector current due to the voltage drop across the working point resistor) on the independent variable. In this case one obtains so called *dynamic characteristics*, which can be substantially different from the static ones.

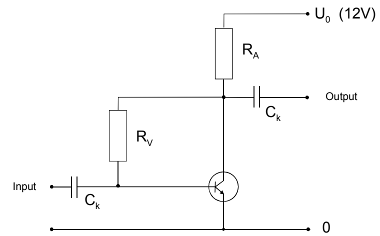
Stabilization

Since the conductivity of semiconductors is strongly dependent on the temperature, one must keep the influence of internal- and external heating on the properties of a circuit as low as possible by introducing special stabilization measures. The most important type of stabilization is *negative feedback*. Negative feedback means that a part of the amplified output signal is inverted and fed back to the input to counteract a change in the amplification factor. The price to pay is a reduction in total amplification.

There are many types of negative feedback. Which is suitable depends above all on the internal resistance of the stage driving the amplifier. In the present case so called *parallel negative feedback* will be investigated (see figure below).

If the amplification increases (with unchanged input signal) hence producing a rise in collector current, then this results in a drop in the collector potential due to the voltage drop across the working point resistor. Since the base dropping resistor R_V forms a voltage divider with the emitter-base resistance, the base potential and the base current also drop so that the collector current again decreases.

The measure of stabilization is described by the action via the base dropping resistance of the feedback ratio (feedback factor) $\alpha = \Delta U_{EB}/\Delta U_{EC}$ and the voltage amplification *v*. These determine the "*forward*" effect of a change in base potential on the collector potential. Both data are fixed by the dimensioning of the circuit elements.



If $\Delta U_{EC}'$ is an assumed change in collector potential without feedback, then with feedback the change is:

(5) $$\Delta U_{EC} = \Delta U_{EC}' - \alpha \ \Delta U_{EC} \ v$$

Solving for the actual change in output voltage gives:

(6) $$\Delta U_{EC} = \frac{\Delta U_{EC}'}{1 + \alpha \, v}$$

In other words, the higher the feedback factor and the higher the amplification the less is the actual output voltage fluctuation.

Presentation of the Physical Principles

(as preparation for part of the report): A summary of the functioning of a transistor. Describe and discuss the quadrants of the characteristic curves and the examples of the circuits to be investigated.

Apparatus and Equipment

Plug-in circuit board with transistor and other circuit elements (resistors, potentiometer).

Power supply unit 12 V; battery (1.5 V mono-cell) for base current. Various multimeters.

Experiment and Evaluation

General information

The open layout of the circuit and the comparatively high resolution of the digital multimeters gives rise to a certain instability of the measured values, resulting a number "jungle" which is a nuisance but unavoidable. In the scope of the accuracy to be achieved one should not be over fastidious in trying to set "smooth" values for the measuring variables.

The maximum ratings of the transistor are not to be exceeded (see lab bench script), to prevent overloading and damaging the transistor.

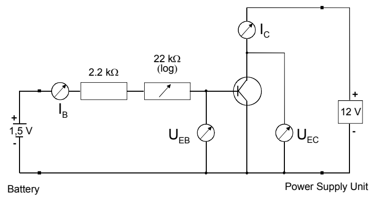
On the other hand, during the measurements, critically observe the measured data quantitatively (order of magnitude, qualitative behavior, relative stability), in order to recognize a damaged transistor in time.

To Exercise 1

The circuit is layed out according to the diagram below. Pay attention to an appropriate use of the measuring instruments (resolution) and in particular the correct voltage measurement in the base-collector circuit. Record the circuit construction and the use of the measuring instruments.

Make a check of the power loss ($U_{EC} \cdot I_C$) for each measurement setting to avoid overloading the circuit.

Record the measured data and as a control plot the values during the measurements.



Four data sets (I_C and U_{EB} as a function of U_{EC}) with I_B = 30, 60, 90 and 120 μA as parameters are to be recorded. The second and third quadrants of the field for the characteristic curves shall be constructed for an assumed voltage supply of 12 V. The static current amplification I_C/I_B is calculated from the second quadrant and the differential input resistance r_{EB} from the third. The determination of r_{EB} may be inaccurate because of the small voltage difference and the fluctuation of the data thus giving only a rough estimate of the value.

To Exercise 2.1

A small working resistance causes small voltage drops and is not suited for voltage amplification. A too higher working resistance could cause U_{EC} drop too low and thus act as a current limiter. A suitable *closed circuit voltage* U_{EC} lies at about half of the supply voltage for the circuit (here 6 V). The working resistance R_A is calculated from the required voltage drop across the working resistor and the quiescent current I_C at the working point and the base dropping resistance R_V is calculated from U_{EC} less the base threshold voltage and the base quiescent current I_B at the working point.

To Exercise 2.2

The amplifier circuit is built according to the circuit diagram on page 6 with the previously determined values for R_A und R_V, and with a supply voltage of 12 V the collector resistance curve (U_{EC}/I_C) is measured by varying the base dropping resistance. One requires fort he evaluation, the exact value of the working resistance.

Record the measured values (U_{EC}/I_C) together with the expected resistance curves in in the field for the characteristic curves. Construct the dynamic I_B/I_C characteristic curve and from this calculate the dynamic current amplification.

To Exercis 2.3

The circuit is complemented by two 0.1 μF coupling capacitors (see circuit diagram, page 6), and a sine signal (about 1000 Hz) applied to the input (function generator *Voltcraft* 7202). The signal can be suitably attenuated at the function generator by pressing the ATT 20-dB button.

The input circuit with the coupling capacitor and the emitter-base resistance represents an *R-C circuit* and thus a frequency dependent voltage divider (high-pass filter).

At first observe the input signal at the coupling capacitor and at the base of the transistor and the output signal on the oscilloscope (voltage ratio and phase as a function of frequency). Record the observations. AS a trial, select a larger base dropping resistor and increase the input signal.

Finally, determine the voltage amplification by measuring the input- and output voltage with the HC-5050-DB multimeter.

Anlage I FEHLERRECHNUNG	GP II

Alle realen Daten tragen unbestimmte zufällige und systematische Abweichungen, die als ***Fehler*** bezeichnet werden, und sind verteilt. Die Erhebung von Daten (Messungen, Berechnungen) stellen grundsätzlich Schätzungen dar, die durch

(1) Intervalle

wiedergegeben werden müssen (Intervallschätzungen, *Fehlerintervalle*; inhaltlich zutreffender, aber unüblich *Ergebnisintervalle*)!

Die Fehlerintervalle ermöglichen signifikante, schließende Vergleiche [siehe (28-30)], die mit singulären Zahlenwerten nicht gegeben sind, und

(2) die Angabe von Zahlenwerten allein ist unwissenschaftlich!

Diese Zusammenfassung enthält grundlegende Elemente und Methoden der Fehlerrechnung; eine ausführlichere Darstellung ist in der Anleitung zu Teil I des Grundpraktikums zu finden (GP I).

Darstellung und Eigenschaften der Fehlerintervalle

Die übliche Schreibweise für Fehlerintervalle in der Physik und Messtechnik besteht aus dem (zufälligen) Mess- oder Ergebniswert als *Zentralwert* des Fehlerintervalls und dem Fehler als *Intervallradius*:

(3) (17,4 $\pm$ 0,3) Maßeinheiten =


(4) Die Fehlerintervalle sind als homogen zu betrachten!

d.h. der zentrale Mess- oder Ergebniswert hat kein höheres Gewicht als jeder andere Wert des Intervalls auch!

Mathematisch-statistisch ist die ***Standardabweichung*** als Fehlermaß vereinbart, so dass Fehlerintervalle mit einer

(5) statistischen Wahrscheinlichkeit von (lediglich) 68 % ($\approx$ 2/3)

den Erwartungswert der betreffenden Größe erfassen, und demzufolge eine *Irrtumswahrscheinlichkeit* von (immerhin) 32 % $\approx$ 1/3 verbleibt.

Protokollierung und Notation

(6) Sämtliche Daten sind in der Form (3) zu notieren!

Ausnahmen sind dort möglich, wo es Gruppen von Daten gibt, die übereinstimmende Fehler tragen (wie Spalten oder Zeilen in Tabellen), oder deren Fehler sich aus gemeinsamen Sekundärangaben berechnen (wie Messdaten von elektrischen Multimetern). Die Fehler oder diese Angaben sind direkt bei den Daten zu notieren.

Oft tragen situationsbedingte oder subjektive Umstände zur Fehlerabschätzung bei, die dann zum Verständnis für Dritte im Messprotokoll zu dokumentieren sind.

Auch *Fehler tragen Fehler*, wobei die Genauigkeit der Fehlerwerte im allgemein weit unter der der Größen selbst bleibt und eher von der Größenordnung eines Faktors 2 anzunehmen ist. Fehler dürfen deshalb nur

(7) mit einer signifikanten Stelle

angegeben werden (z.B. δI = 2 % und nicht 1,3725... %), wobei die

(8) Fehler stets aufzurunden sind,

um (5) nicht zu verletzen (d.h. δI = 2 % und nicht 1 % bei dem vorhergehenden Beispiel)!

Abweichend von (7) sollen abgelegte und später wiederaufgenommene Fehler von

(9) Zwischenwerten zweistellig

protokolliert werden, um eine Akkumulation von Rundungsfehlern zu vermeiden (d.h. δI = 1,4 % in obigem Beispiel, wenn δI zur späteren Widerstandsberechnung notiert wird).

(10) Die Zahlwörter von Ergebnis und Fehler müssen in der gleichen Stelle enden!

Eine höhere oder geringere Zahlenauflösung wäre inkohärent und würde nichtsignifikante Genauigkeit vortäuschen oder vorhandene Aussage unterdrücken; also R = (1,70 $\pm$ 0,03) kΩ und nicht R = (1,7027 $\pm$ 0,03) kΩ oder R = (1,7 $\pm$ 0,03) kΩ.

Fehlerbeiträge bei Messgrößen

Die Fehlerbeiträge der Messwerte können grob drei Kategorien zugeordnet werden:

(11) *Statistischer Fehler (Streufehler)*
Praktische Schätzfehler
Nennfehler

Stichprobenschätzung und Streufehler

Bei einer signifikanten Streuung können eine ***Messreihe*** als Stichprobe aufgenommen und Ergebnis und Fehler mathematisch-statistisch berechnet werden.

Als ***Ergebnis*** wird der (einzelne) Mittelwert der Messreihe herangezogen:

(12) Ergebnis $x = \bar{x} = \dfrac{\sum x_i}{n}$

Der ***Fehler*** als Standardabweichung der Verteilung dieser Mittelwerte wird durch die Standardabweichung der Grundverteilung der Einzelwerte und durch den Umfang der Stichprobe bestimmt (statistisches Gewicht):

(13) Fehler

$$\Delta x = \sigma(\bar{x}) = \frac{\sigma(x)}{\sqrt{n}} \approx \frac{\sqrt{\dfrac{\sum (x_i - \bar{x})^2}{n-1}}}{\sqrt{n}} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n\,(n-1)}}$$

(Der Nenner n-1 berücksichtigt den Näherungscharakter bei der Berechnung von σ(x) aus einer Stichprobe).

Praktischer Schätzfehler

Ist das Auflösevermögen eines Messverfahrens deutlich geringer als die Standardabweichung, so wird eine Streuung nicht beobachtet. Dann muss als *praktischer Schätzfehler* ein

(14) Fehler nach praktischen Maßgaben unter kritischer Berücksichtigung aller Umstände (Ablesung, Anzeige) abgeschätzt werden.

Nennfehler (Gerätefehlerangabe)

Aufgrund des Funktionsprinzips und der Bauart zeigt jedes Messgerät typische, meist systematische Abweichungen, die vom Gerätehersteller in den Unterlagen angegeben werden müssen. Diese Fehler werden hier als *Nennfehler* bezeichnet.

Kontrollmessung und Messfehler bzw. Messergebnis

(15) Zur Feststellung des Streuverhaltens müssen jede Messung oder Messeinstellung wiederholt und auch das Ergebnis dieser Kontrollmessung protokolliert werden!

Kontinuierlich oder periodisch arbeitende Messverfahren (Multimeter als Beispiel) enthalten eine Messwiederholung implizit.

(16) Bei einer signifikanten Abweichung ist die Größe durch eine Messreihe zu untersuchen und nach (12) und (13) auszuwerten.

(17) Wird der Wert im Rahmen des Auflösungsvermögens reproduziert, muss ein Fehler praktisch geschätzt werden.

(18) Der gesamte Messfehler ergibt sich dann aus der Summe nach (24) des Streu- oder Schätzfehlers und dem Nennfehler, wobei das Prinzip nichtbeitragender Fehler (24) berücksichtigt werden kann.

Fehlerfortpflanzung

Für statistisch unabhängige Größen gilt das *Gaußsche Fehlerfortpflanzungsgesetz*:

(19) $z = f(a,b,\ldots)$ *mit*

$$\Delta z = \sqrt{\left(\frac{\partial f}{\partial a}\,\Delta a\right)^2 + \left(\frac{\partial f}{\partial a}\,\Delta a\right)^2 + \ldots}$$

Für die elementaren Rechenverknüpfungen folgt daraus mit den *absoluten Fehlern* Δx bzw. den *relativen Fehlern* δx:

(20) $z = a \pm b$ mit $\Delta z = \sqrt{\Delta^2 a + \Delta^2 b}$

(21) $z = a \overset{\times}{\underset{\div}{}} b$ mit $\delta z = \sqrt{\delta^2 a + \delta^2 b}$

(22) $z = a^r$ mit $\delta z = r\,\delta a$

(23) speziell $z = \dfrac{1}{a}$ mit $\delta z = \delta a$ und $\Delta z = \dfrac{\Delta a}{a^2}$

Die Rechenbeziehungen (20) und (21) mit der Wurzel aus der Summe der Quadrate führen dazu, dass auf den jeweiligen Verknüpfungsebenen kleinere Fehlerbeiträge als

(24) nichtbeitragende Fehler vernachlässigbar

im Rahmen der übrigen Aufrundungen sind. Ein Fehler ist als klein zu betrachten, wenn es vergleichbare Fehler gibt, die diesen um das Dreifache oder mehr übertreffen.

Sind Korrelationen zwischen Größen gegeben oder anzunehmen, muss das *Maximal-Fehlerfortpflanzungsgesetz* herangezogen werden:

(25) $\Delta_{max} z = \left|\frac{\partial f}{\partial a}\,\Delta a\right| + \left|\frac{\partial f}{\partial b}\,\Delta b\right| + \ldots$

Fehlerabschätzung bei der grafischen Auswertung von Funktionen

Bei der grafischen Auswertung linearer oder linearisierter Funktionen wird neben der Ausgleichsgeraden (Bestgerade) eine Grenzgerade eingetragen, die noch mit den Messwerten verträglich ist, und die zugehörigen Grenzwerte der Parameter berechnet (Grenzachsenabschnitte $a^{\pm}$, Grenzanstieg $m^{\pm}$). Die Fehler sind dann die Differenzen zu den Ausgleichswerten:

(26) $\mathit{Achsenabschnitt} = (a \pm \Delta a)$ mit $\Delta a = \left|a^{\pm} - a\right|$

(27) $\mathit{Anstieg} = (m \pm \Delta m)$ mit $\Delta m = \left|m^{\pm} - m\right|$

Es ist ausreichend, eine der beiden möglichen Grenzgeraden zu betrachten. Die Grenzgerade soll die Ausgleichsgerade etwa in der Mitte der Punktwolke schneiden (eine parallele Grenzgerade ist ungeeignet, da sie den Achsenabschnitt nicht variiert).

Die Lage der Grenzgeraden muss sich an der Streuung der Punkte und zusätzlich an den expliziten Einzelfehlern der Punkte orientieren.

Bei logarithmischen Darstellungen sind die Ausgleichsgerade und die Grenzgerade schwerpunktsmäßig an die größeren y-Werte mit den (im allgemeinen) kleineren relativen Fehlern anzupassen.

Bei der subjektiven Festlegung der Grenzgeraden bleibt typischerweise die Anzahl der Messpunkte als statistisches Gewicht unberücksichtigt, so dass die Fehler zu groß abgeschätzt werden.

Schließender Vergleich

Ein Vergleich bildet im Sinne eines *statistischen Tests* die kontinuierliche Menge der Ergebnisintervalle auf die diskrete Menge von drei Aussagen ab:

(28) Ergebnisse sind als gleich zu bewerten, wenn die Intervalle überlappen oder sich erreichen.

(29) Ergebnisse sind verträglich, wenn die dreifachen Intervalle überlappen oder sich erreichen.

(30) Ergebnisse sind (erst dann) signifikant unterschiedlich, wenn (28) oder (29) nicht zutreffen.

Anlage II He-Ne-LASER — GP II

Physikalische Grundlagen und Funktionsweise des Helium-Neon-Lasers

Elektronische Zustände; spontane Emission und Absorption

Physikalische Systeme befinden sich am Temperaturnullpunkt normalerweise im Grundzustand, d.h. dem Zustand geringster Energie. Die Elektronen des Systems besetzen dabei die tiefsten Niveaus, wobei die Besetzungszahlen für diese Niveaus sich aus quantenmechanischen Regeln ergeben. Durch Zufuhr von Energie (thermische Anregung, Absorption von Photonen, Stoßprozesse, etc.) können energetisch höher liegende Zustände besetzt werden. Die angeregten Zustände sind grundsätzlich nicht stabil und zerfallen spontan zu tiefer liegenden Zuständen (*spontane Emission*; vergleiche auch Versuche *OPTISCHE SPEKTROSKOPIE*, *RADIOAKTIVER ZERFALL* und *GAMMA-SPEKTROSKOPIE*). Der Übergang erfolgt in den meisten Fällen unter Aussendung von Lichtquanten (Photonen) mit einer für den Übergang charakteristischen Wellenlänge bzw. Frequenz. Die mittlere Verweilzeit in den angeregten Zuständen (Lebensdauer) ist dabei ebenfalls eine die Zustände kennzeichnende Größe. Die Untersuchung der Strahlung erlaubt Rückschlüsse auf die emittierenden Systeme (Spektroskopie).

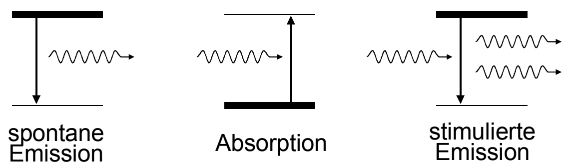
Befindet sich ein solches System in Wechselwirkung mit einem elektromagnetischen Strahlungsfeld (z.B. Licht), so finden neben den Emissions- auch Absorptionsprozesse statt. Dabei kann sich ein stationärer Gleichgewichtszustand einstellen, bei dem Emission und Absorption mit gleicher Rate erfolgen. Im thermischen Gleichgewicht mit dem Strahlungsfeld hängen die Besetzungszahlen der Zustände von der Temperatur und der Anregungsenergie E_i der jeweiligen Zustände ab und folgen einer Boltzmannverteilung der Form $\exp(-E_i/kT)$. In die genaue Verteilung gehen zusätzlich quantenmechanische Gewichtsfaktoren ein, sogenannte *g-Faktoren*.

Als Beispiel ergibt sich für einen Zustand mit einer Übergangswellenlänge zum Grundzustand von 500 nm bei Raumtemperatur eine relative Besetzungszahl von etwa 10^{-43}. Bei einer Teilchendichte von $6{\cdot}10^{23}$ Atomen pro Mol ist die Anzahl angeregter Atome damit praktisch gleich Null. Bei 2400 K, der Temperatur einer Lampenglühwendel, steigt die relative Besetzung auf etwa 10^{-5}

Stimulierte Emission

Sind in dem Strahlungsfeld Frequenzen vorhanden, die mit Übergangsfrequenzen zwischen Zuständen des Systems übereinstimmen, so tritt neben der spontanen Emission, die unbeeinflußt von dem Strahlungsfeld abläuft, als zusätzlicher Emissionsprozeß sogenannte *stimulierte* (oder *induzierte*) *Emission* auf, bei der ein Übergang vom angeregten Zustand zu einem tiefer liegenden Zustand durch ein passendes Photon des Strahlungsfeldes ausgelöst wird.

Das durch stimulierte Emission erzeugte Photon besitzt exakt die gleichen Eigenschaften wie das auslösende Photon, d.h. es trägt die gleiche Frequenz, Ausbreitungsrichtung, Polarisation und Phase. Bei spontan emittierten Photonen ist eine solche Übereinstimmung untereinander nicht gegeben.

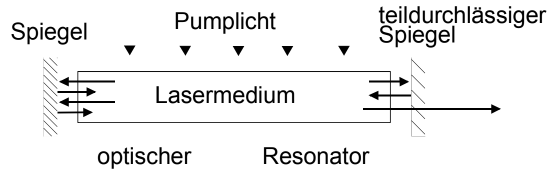


Das Laserprinzip

Wird durch irgendeinen Prozeß eine derartige Abweichung der Besetzung vom thermodynamischen Gleichgewicht erreicht, daß der angeregte Zustand stärker besetzt ist als ein betrachteter tiefer liegender Zustand (*Besetzungsinversion*), so werden durch stimulierte Emission mehr Photonen erzeugt als durch Absorption vernichtet. Man erhält eine Verstärkung des Strahlungsfeldes und spricht von einen *Laserprozeß* (LASER = light amplification by stimulated emission of radiation). Ein Laser ist ein Gerät, das Licht mit Hilfe stimulierter Emission von Strahlung verstärkt.

Ein Laser besteht aus einem Lasermedium, in meist länglicher Form, in dem ein Laserübergang bis zur Besetzungsinversion angeregt wird. Die Anregung ("*Pumpen des Laserübergangs*") kann bei Festkörpern und Flüssigkeiten optisch durch Beleuchtung mit intensiven Blitzlampen oder mit einem weiteren "*Pumplaser*" erfolgen, oder bei Gasen durch Elektronenstoßanregung in einer in dem Lasermedium gleichzeitig brennenden Gasentladung.

Um die Photonendichte des Strahlungsfeldes zu erhöhen, wird dem Laser ein Rückkopplungsmechanismus zugefügt. An den Stirnflächen des Lasermediums werden Spiegel angebracht, so daß ein *optischer Resonator* entsteht, der das Strahlungsfeld selektiv anreichert (*Fabry-Perot-Resonator*; siehe auch Versuch *FABRY-PEROT-ETALON*). Durch das Prinzip der optischen Vielstrahlinterferenz ist die Linienbreite eines solchen optischen Resonators um ein Vielfaches geringer, als die der Laserresonanz selbst, so daß der Resonator darüber hinaus eine Feinselektion bezüglich Frequenz, Ausbreitungsrichtung, Polarisation und Phase bewirkt. Ein Teil der so erzeugten Strahlung kann ausgekoppelt werden, indem einer der Endspiegel teildurchlässig gemacht wird.



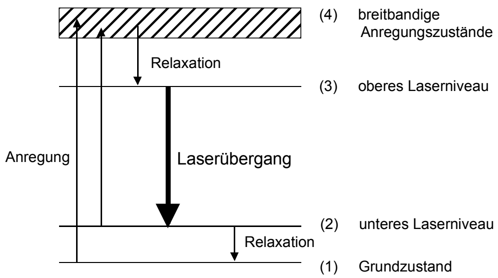
Eine tatsächliche Verstärkung durch stimulierte Emission kann nur dann erfolgen (das Medium nur dann *lasen*), und der optische Resonator nur dann stationär schwingen, wenn die Zahl der durch stimulierte Emission erzeugten Photonen die der Verluste ausgleicht (Schwellbedingung). Verluste entstehen durch die Auskopplung, aber auch an den Spiegeln und den Grenzflächen des Lasermediums, und vor allem infolge von *Selbstabsorption* innerhalb des Lasermediums. Die Inversion des Niveaus muß daher um so größer sein, je höher diese technischen und physikalischen Verluste sind.

Strahlungscharakteristik

Die Laserstrahlung zeichnet sich durch geringe Linienbreite, hohe Kohärenz, hohe Linienintensität (hohe spektrale Dichte) und eine sehr ausgeprägte Richtungscharakteristik aus. Um die hohe Linienintensität des Lasers deutlich zu machen, sei folgendes Beispiel erläutert: Ein Photon im sichtbaren Bereich repräsentiert etwa 10^{-19} J an Energie. Eine Laserausgangsleistung von 1 mW ergibt dann rund 10^{16} Laserphotonen pro Sekunde. Die Bandbreite der Strahlung beträgt etwa 10^5 Hz. Um mit einer thermischen Lichtquelle diese Linien-Strahlungsleistung zu realisieren, müßte die Temperatur der Quelle 10^{15} K betragen! Zum Vergleich: Die Temperatur im Sterninneren liegt bei etwa 10^8 K.

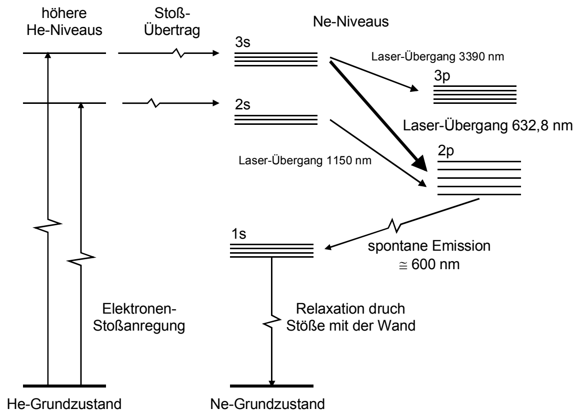
Der He-Ne-Laser

Der He-Ne-Laser ist ein sogenannter *Vier-Niveau-Laser*, dessen Funktionsprinzip an folgendem Termschema nachvollzogen werden kann.



Durch einen Anregungsprozeß wird eine Reihe eng benachbarter, höherer Niveaus (4) besetzt, die durch spontane Emission und strahlungslose Übergänge zum Niveau (3), dem oberen Laserniveau, relaxieren. Zwischen (3) und (2) findet der Laserübergang statt. Auch das untere Laserniveau (2) ist nicht stabil und relaxiert weiter zum Grundzustand (1). Bedeutsam für die Funktion ist die Besetzung des oberen Laserniveaus über die Zwischenzustände (4), die breitbandig mit hoher Effizienz angeregt werden können, und die Instabilität des unteren Laserzustands (2), wodurch dieser ständig entleert wird und so für die notwendige Inversion zwischen (3) und (4) sorgt. Ein solcher Vier-Niveau-Laser erfordert aus den geschilderten Gründen keine hohe Pumpleistung und läßt sich ohne großen technischen Aufwand realisieren.

Die folgende Abbildung zeigt schematisch einige Termgruppen und Übergänge des He- und des Ne-Atoms.



In einer Gasentladung werden durch Elektronen-Stoßanregung verschiedene Zustände des Heliums besetzt. Aufgrund quantenmechanischer Übergangsregeln sind optische Übergänge zum Grundzustand des He nicht erlaubt, und eine Rückkehr in den Grundzustand findet durch Stöße und Energieübertrag mit Neon-Atomen statt, für die wegen der energetisch benachbart liegenden 2s und 3s Niveaus im Ne eine hohe Wahrscheinlichkeit besteht. In dem vereinfachten Vier-Niveau-Schema entspricht die Anregung des Helium dem Schritt von (1) nach (4) und dem Stoßübertrag der Schritt von (4) nach (3).

Die rote Linie des He-Ne-Lasers liegt bei 632,8 nm. Es finden auch Laserprozesse bei den anderen Wellenlängen statt, die aber durch die Bauart des Resonators im allgemeinen unterdrückt werden. Dem Relaxationsprozeß von (2) nach (1) entspricht beim He-Ne-Laser ein zusammengesetzter Prozeß. Im ersten Schritt erfolgt durch spontane Emission ein Übergang in einen Zwischenzustand. Von diesem Niveau aus relaxieren die Ne-Atome über Stöße mit der Wand in den Grundzustand.

Der Vorteil dieses relativ komplizierten Laserzyklus liegt in der sehr effizienten Anregung der He-Atome in der Gasentladung. Es gibt auch reine Ne-Laser, die jedoch weit weniger wirkungsvoll arbeiten. Der He-Ne-Laser ist der erste im kontinuierlichen Betrieb realisierte Laser und gehört bis heute zu den zuverlässigsten und am häufigsten eingesetzten Lasertypen.

Der erste Betrieb eines Lasers gelang 1960 *H.Maiman* mit einem Rubinlaser. Der erste Gaslaser wurde 1961 von *A.Javan*, *W.R.Bennett* und *D.R.Herriott* realisiert.

Aufbau eines He-Ne-Lasers

Die Abbildung auf der folgenden Seite zeigt die typischen Elemente eines He-Ne-Lasers

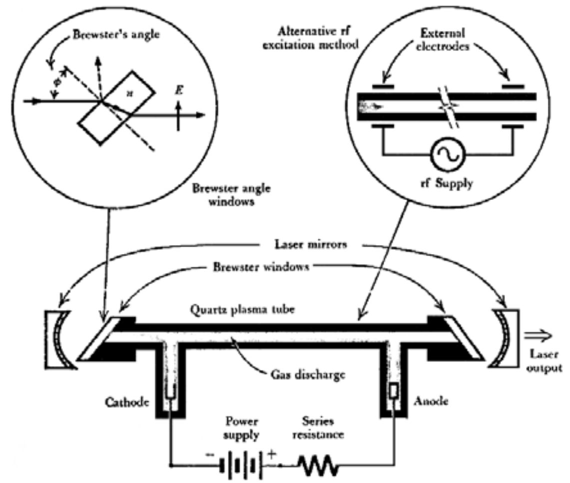


FIG. 1-33 The essential elements in a gas laser are the plasma tube, the laser mirrors, and the power supply. The gas discharge may be obtained with internal electrodes, or with external rf electrodes (upper right). The Brewster-angle windows transmit vertically polarized light with zero surface reflections (upper left).

In einem Quarzrohr (100-1000 mm Länge, einige mm Durchmesser) brennt in einem He-Ne-Gasgemisch (etwa 10 % He) eine Gasentladung. Die Endfenster des Quarzrohres stehen unter dem *Brewster-Winkel*, um Reflexionsverluste beim vielzahligen Hin- und Hergang der Strahlung gering zu halten. Zwei externe, konfokale Spiegel bilden den optischen Resonator. Der eine Spiegel besitzt einen möglichst hohen Reflexionsgrad. Der andere Spiegel ist teildurchlässig, um die Strahlung auszukoppeln. Bei einer Anregungsleistung von 5-10 W liegt die Laser-Ausgangsleistung bei 0,5 -50 mW. Die Bandbreite der Laserstrahlung beträgt etwa 10^5 Hz bzw. 10^{-7} nm.

Anlage III STROMLEITUNG	GP II

Die Beschreibung der Stromleitung in Festkörpern erfolgt im Rahmen der Festkörperphysik durch quantenmechanische Vorstellungen. Einige wesentliche Aspekte sollen hier phänomenologisch und qualitativ ausführlicher dargestellt werden, da das Thema einen Vorgriff auf den Stoff späterer Vorlesungen darstellt.

<u>Atom</u>

Bei freien Atomen bewirkt der positiv geladene Kern eine Potentialmulde, in der im neutralen Fall sämtliche Elektronen des Atoms fest gebunden sind. Analog zu den diskreten Schwingungsfrequenzen einer beidseitig eingespannten Saite führt die räumliche Begrenzung der Elektronen auf das Atomvolumen dazu, daß nur ganz bestimmte Zustände mit diskreten Energien und Drehimpulsen eingenommen werden können (*Bohr'sches Atommodell*). Zustände mit unterschiedlichem Drehimpuls, aber eng benachbarten Energien, werden zu Schalen zusammengefaßt (Hauptquantenzahlen). Nach dem *Pauli-Prinzip* darf jeder Zustand nur mit (höchstens) <u>einem</u> Elektron besetzt werden, so daß mit steigender Kernladungszahl Z immer höhere Schalen aufgefüllt werden. Die am schwächsten gebundenen Elektronen in der äußersten Schale (d.h. die mit der höchsten Energie) bestimmen den chemischen Charakter des Elements (*Leuchtelektron*). Das Periodensystem der Elemente in der Anordnung mit steigendem Z spiegelt diese Schalenstruktur wieder.

<u>Elektronen in Festkörpern</u>

Bei der Kondensation von Atomen zu einem Festkörper (z.B. Kupfer mit Z=29 zu Cu-Metall) können sich die Potentiale benachbarter Atome deformieren und soweit überlappen, daß der höchste besetzte Zustand oberhalb der einzelnen Potentialmulden liegt. Damit sind die Elektronen in diesem Zustand nicht mehr an das Atom gebunden, und zurück bleiben Cu^+-Ionen und die ungebundene Elektronen als sogenanntes *Elektronengas*. Die Bildung eines derartigen Elektronengases hoher Dichte (etwa ein Elektron pro Atom, d.h. etwa 10^{29} m^{-3} für Kupfer) ist das Kennzeichen eines Metalls. Typische metallische Eigenschaften, wie z.B. die hohe elektrische und thermische Leitfähigkeit und das große Absorptions- und Reflexionsvermögen für Licht, sind Eigenschaften dieses Elektronengases. Allerdings sind die Elektronen nicht uneingeschränkt frei, sondern ihre Beweglichkeit unterliegt weiter dem Pauli-Prinzip, so daß Ortsveränderungen nur durch Platzwechsel auf freie Zustände erfolgen können.

<u>Bandstruktur</u>

Durch den Überlapp der Potentiale benachbarter Atome werden die Elektronen auf den äußeren Bahnen gekoppelt, und es entsteht eine Aufspaltung (Vermehrung) der möglichen Zustände analog der Aufspaltung der Eigenfrequenzen gekoppelter Pendel. Für einen Festkörper mit einer sehr großen Zahl von N gekoppelten Atomen liegen die Zustände so dicht beieinander, daß man sie für viele Betrachtungen als quasi kontinuierlich verteilt ansehen kann und von einem (*Energie*-) *Band* spricht. Atomare Zustände, die sich nur durch die Orientierung des Drehimpulses unterscheiden, bilden gemeinsame Bänder. Grundsätzlich sind jedoch auch die Zustände in den Bändern diskret und können, entsprechend dem Pauli-Prinzip, nur mit je einem Elektron besetzt sein, wobei die Anzahl der Zustände in einem Band gleich der Zahl der beteiligten Zustände pro Atom mal der Anzahl der Atome N ist. Zwischen den Bändern können *Bandlücken* (verbotene Zonen) liegen.

<u>Leiter</u>

Die äußeren Elektronen von Kupfer sind s-Elektronen mit zwei Drehimpulsorientierungen, so daß es $2N$ Leitungsband-Zustände gibt, die von den N Leitungselektronen (eins je Atom) also gerade zur Hälfte aufgefüllt werden. Derartige nicht vollständig besetzte Bänder werden als Leitfähigkeits- oder kurz *Leitungsbänder* bezeichnet, weil in diesen freie Plätze (Zustände) vorhanden sind, über die eine einseitige Elektronenbewegung, und damit ein Ladungstransport stattfinden kann.

Die Elektronen in den höchsten besetzten Bandzuständen an der *Fermigrenze* besitzen relativ hohe kinetische Energien mit Geschwindigkeiten von etwa 10^6 m/s (bei rein klassischer Betrachtungsweise eines Elektronengases würde dies einer Temperatur von 10^5 K entsprechen). Daran gemessen ist die zusätzliche Beschleunigungsenergie durch ein äußeres elektrisches Feld vergleichsweise klein, so daß nur die "obersten" Elektronen in freie Zustände angehoben werden und zum Ladungstransport beitragen können. Alle anderen dagegen (99% bei Zimmertemperatur) verbleiben in ihren Zuständen und tragen nicht zum Ladungsstrom (und auch nicht zum Wärmetransport und zur Wärmekapazität) bei.

<u>Nichtleiter</u>

Werden pro Atom genau so viele Elektronen freigesetzt (delokalisiert), wie Zustände in einem Band enthalten sind, so wird das Band vollständig gefüllt. Es gibt keine freien Zustände, und damit ist eine Aufnahme elektrischer Feldenergie und eine elektrische Leitung nicht möglich; der Festkörper ist ein Isolator. Derartige vollständig besetzten Bänder heißen *Valenzbänder.*

<u>Halbleiter</u>

Liegt das nächsthöhere, leere Band nur in geringem Energieabstand über dem gefüllten Valenzband, so können durch thermische Anregung Elektronen aus dem Valenzband über die Bandlücke ΔE gehoben werden, und das leere Band wird zum Leitungsband. Zusätzlich hinterlassen die angehobenen Elektronen Lücken (*Löcher*) im Valenzband, also freie Zustände, so daß auch darin ein Ladungstransport möglich wird. Die Substanz wird zum Halbleiter (genauer *Eigenhalbleiter* oder intrinsischer Halbleiter). Nach allgemeinen Gesetzen der Thermodynamik wird der Bruchteil der angehobenen Elektronen durch den Boltzmannfaktor $e^{-E_G/kT}$ (k = Boltzmannkonstante) bestimmt. (Wie groß ist dieser Faktor für Germanium mit E_G = 0,6 eV bei Zimmertemperatur?).

<u>Störstellenhalbleiter (Dotierung)</u>

Die spezifische Leitfähigkeit ist proportional zur Ladungsträgerdichte, bei einem intrinsischen Halbleiter also exponentiell mit der Temperatur wachsend. Höhere und bei Raumtemperatur praktisch konstante Ladungsträgerdichten erhält man durch *Dotierung* mit Atomen anderer Wertigkeit. Durch die Fremdatome entstehen besetzte Zustände in der Bandlücke des Halbleiters dicht unterhalb des Leitungsbandes bzw. freie Zustände dicht oberhalb des Valenzbandes, die ein zusätzliches Elektron abgeben (*Donatoren*) oder aufnehmen (*Akzeptoren*) können, und die wegen ihrer geringen Anregungsenergie bereits bei Zimmertemperatur vollständig angeregt bzw. besetzt sind.

Elektrischer Widerstand

Die Stromleitung in einem Metall und dessen elektrischer Widerstand lassen sich in einfacher Näherung klassisch verständlich machen. Wird an dem Metall durch Verbinden mit einer Spannungsquelle ein elektrisches Feld erzeugt, so versetzt dies die freien Elektronen in beschleunigte Bewegungen, die durch Stöße mit den Ion-Rümpfen immer wieder unterbrochen werden. Es stellt sich eine im statistischen Mittel gleichförmige Bewegung mit einer mittleren Driftgeschwindigkeit ein, die proportional zur Spannung ist, und die sich aus der Teilchenzahldichte, der Ladung und der Stromdichte berechnen läßt. Aus diesem Bild folgt das *Ohmsche Gesetz* mit R = const.

Dies klassische Bild besitzt aber Grenzen, und erst die Quantenmechanik erklärt, wie sich Elektronen in einem räumlich periodischen Potential ohne Streuung bewegen können, d.h. ohne (wie klassisch zu erwarten) mit den Atomrümpfen zusammenzustoßen. Ein ideal aufgebautes Metallgitter ohne Störungen der Periodizität durch Fremdatome oder durch thermischen Schwingungen der Ionen (am Temperatur-Nullpunkt) hat danach keinen elektrischen Widerstand (dies ist nicht die Supraleitung).

Bei Zimmertemperatur ist die mit der Temperatur zunehmende thermische Bewegung als Ursache für den Widerstand dominierend, der beim Abkühlen auf sehr tiefe Temperaturen für reine Metalle typisch um einen Faktor 100 sinkt. Der schließlich temperaturunabhängige Restwiderstand ist ein Maß für die Reinheit des Materials. Bei Legierungen dagegen kann der Fehlstellenbeitrag so groß werden, daß er schon bei Raumtemperatur überwiegt. Dies Verhalten erklärt den hohen spezifischen Widerstand von z.B. Manganin und Konstantan und dessen geringe Temperaturabhängigkeit.

Supraleitung

Grund der Supraleitung ist nicht die Verringerung der thermischen Schwingungen bei tiefen Temperaturen als eine der Ursachen des Widerstands, sondern eine Zustandsänderung des Elektronengases durch Bildung sogenannter *Cooper-Paare* durch eine sehr schwache Phononen-Austauschwechselwirkung zwischen Elektronen entgegengerichteten Spins, wobei sich die elektronischen Eigenschaften grundlegend ändern und die Gitterstörungen als Ursachen des Widerstands von den Elektronen nicht mehr wahrgenommen werden.

Im Normalzustand unterliegen Elektronen als Spin-1/2-Teilchen der Fermi-Statistik und sind als kleine Teilchen zu verstehen, die im Wellenbild eine kurze (de-Broglie-) Wellenlänge besitzen. Durch die Kopplung werden die Cooper-Paare spinlos. Sie folgen dann der Bose-Einstein-Statistik und nicht mehr dem Pauli-Prinzip, und können einen einheitlichen, *kohärenten* Zustand bilden, der sich durch eine große de-Broglie-Wellenlänge auszeichnet. In diesem Zustand nehmen die Elektronen die thermischen Gitterschwingungen der Atomrümpfe und auch Gitter-Fehlstellen als räumlich kleine Störungen nicht mehr wahr, und es kommt zu dem Effekt des verschwindenden Widerstands.

Die Kopplung zu Cooper-Paaren ist in einem Elektronengas innerhalb eines Gitters möglich, da dort die Coulomb-Wechselwirkung durch die allseitige Umgebung mit gleich geladenen Teilchen nahezu völlig ausgeschaltet ist, und unterhalb einer bestimmten Temperatur (der Sprungtemperatur), wenn auch die thermischen Phononen ausreichend abgeschwächt sind. Die innere Energie des Metalls ist dann umso kleiner, je mehr Cooper-Paare gebildet werden, so daß sie den energetisch günstigstens Zustand des Systems darstellen. Es handelt sich dabei um eine Phasenänderung des Systems, die nicht in Abhängigkeit der Temperatur verläuft, sondern sprungartig, und bei der sich die Entropie des Systems unstetig ändert.

Anlage IV WECHSELSTROMOPERATOREN	GP II

Für Kombinationen von R, C und L können die resultierenden Wechselspannungen und -ströme grundsätzlich durch Ansatz der *Kirchhoffschen Regeln* und Lösung der entstehenden Differentialgleichungen berechnet werden. Die allgemeinen Lösungen dieser Differentialgleichungen sind komplexe Exponentialfunktionen, und für den wichtigen Fall harmonischer Erregung (inhomogene Differentialgleichungen mit sin/cos als aufgeprägte Funktionen) ergibt sich ein einfacher Lösungsformalismus durch eine komplexe Ansätze von Spannung und Strom.

Komplexer Ansatz von Wechselspannung und -strom

Korrekt müßten Spannung und Strom als physikalisch reelle Größen durch eine Kombination komplexer Größer geschrieben werden, wobei man jedoch die gleichen Ergebnisse bei Ansatz einfacher komplexer Funktionen erhält (im folgenden **fett** geschrieben). Die Spannung wird ohne Beschränkung der Allgemeinheit ohne konstanten Phasenanteil geschrieben, während der Strom gegenüber der Spannung um φ phasenverschoben angesetzt wird:

(1.1) $$\mathbf{U} = \mathbf{U}(t) = U_0\, e^{i\,\omega t}$$

(1.2) $$\mathbf{I} = \mathbf{I}(t) = I_0\, e^{i\,(\omega t + \varphi)}$$

Spannung und Strom an R, C und L

Die Zusammenhänge zwischen Spannung und Strom an R, C und L sind gegeben durch (siehe Skript *WECHSELSTROMKREISE*):

(2.1) Widerstand (R): $$U_R = -R\, I_R$$

(2.2) Kapazität (C): $$I_C = -C\,\frac{d\,U_C}{d\,t}$$

(2.3) Induktivität (L): $$U_L = -L\,\frac{d\,I_L}{d\,t}$$

Reihenschaltung von R, C und L

Für eine *Masche* (Reihenschaltung) von R, C und L mit einer aufgeprägten Treiberspannung (Generator) liefert die *Kirchhoffsche Regel* für die Summe der Spannungen mit (2) und (1):

(3) $$\left[R + \frac{1}{i\,\omega C} + i\,\omega L\right] I_0\, e^{i\,(\omega t + \varphi)} = U_0\, e^{i\,\omega t}$$

bzw.

(4) $$\frac{U_0}{I_0}\, e^{-i\,\varphi} = \mathbf{Z} = \left[R - i\left(\frac{1}{\omega C} - \omega L\right)\right]$$

Der Ausdruck (4) stellt eine *komplexe Impedanz* dar, die die Lösung vollständig mit dem Betrag als physikalischer Impedanz und der Phase beschreibt:

(5.1) $$Z = \sqrt{R^2 + \left(\frac{1}{\omega C} - \omega L\right)^2}$$

und

(5.2) $$\tan\varphi = \frac{\frac{1}{\omega C} - \omega L}{R}$$

Parallelschaltung von R, C und L

Für einen *Knoten* (Parallelschaltung) von R, C und L mit einem aufgeprägten Treiberstrom (Generator) erhält man entsprechend aus der *Kirchhoffsche Regel* für die Summe der Ströme mit (2) und (1):

(6) $$\left[\frac{1}{R} + i\,\omega C + \frac{1}{i\,\omega L}\right] U_0\, e^{i\,\omega t} = I_0\, e^{i(\omega t + \varphi)}$$

bzw.

(7) $$\frac{I_0}{U_o}\, e^{i\,\varphi} = \frac{1}{\mathbf{Z}} = \left[\frac{1}{R} + i\left(\omega C - \frac{1}{\omega L}\right)\right]$$

Jetzt folgen als physikalische Impedanz und Phase:

(8.1) $$\frac{1}{Z} = \sqrt{\frac{1}{R^2} + \left(\omega C - \frac{1}{\omega L}\right)^2}$$

und

(8.2) $$\tan\varphi = \frac{\omega C - \frac{1}{\omega L}}{\frac{1}{R}}$$

Komplexe Widerstandsoperatoren

Aus (3) und (6) kann abgelesen werden, daß die Berechnung komplexer Impedanzen denselben Regeln wie bei der Kombination von Widerständen im Gleichstromfall folgt, wenn an Stelle der Widerstände die folgenden *komplexen Widerstandsoperatoren* eingesetzt werden:

(9.1) $$\mathbf{Z}_R = R$$

(9.2) $$\mathbf{Z}_C = \frac{1}{i\,\omega C}$$

(9.3) $$\mathbf{Z}_L = i\,\omega L$$

Als Kombinationsregeln für Serien- bzw. Parallelschaltung gelten entsprechend:

(10.1) Serienschaltung: $$\mathbf{Z}_{gesamt} = \mathbf{Z}_1 + \mathbf{Z}_2$$

(10.2) Parallelschaltung: $$\frac{1}{\mathbf{Z}_{gesamt}} = \frac{1}{\mathbf{Z}_1} + \frac{1}{\mathbf{Z}_2}$$

Phasenverschiebung

Bei der Diskussion von Phasenverschiebungen muß der mehrdeutige Verlauf der $\tan^{-1}$-Relation berücksichtigt werden. Aus $\tan\varphi = 0$ für ein Netzwerk ohne Kapazität und Induktivität folgt mathematisch $\varphi = \pm\, n\, \pi$. Physikalisch zutreffend sind aber nur die Lösungen mit ungeradem n, für die $\cos\varphi = -1$ ist, weil das System über den Widerstand Energie abgibt. Ebenso sind für Kombinationen von R, C und L nur die Lösungen korrekt, für die $\cos\varphi$ negativ ist, d.h. $\varphi > 90°$ oder $\varphi < -90°$.

Anlage V HALBLEITER	GP II

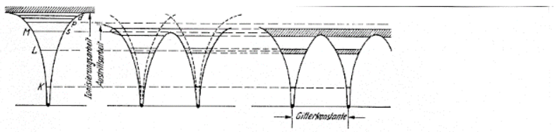


Abb. 14.50. Wenn die Einzelatome einander näherkommen, überlagern sich nicht nur ihre Potentiale zu einer Galerie von Rundbögen, sondern die ursprünglich scharfen Elektronenzustände verbreitern sich. Im rechten Teilbild ist rechts die Atomkette fortgesetzt zu denken, links liegt die Kristalloberfläche. Dort kann man auch die Austrittsarbeit für die Elektronen ablesen

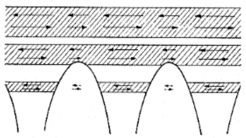


Abb. 14.51. Elektronenbänder im periodischen Potentialgebirge. Die Pfeile deuten die Beschleunigungsrichtung der Elektronen im äußeren Feld an: Die effektive Masse wechselt ungefähr in der Bandmitte ihr Vorzeichen

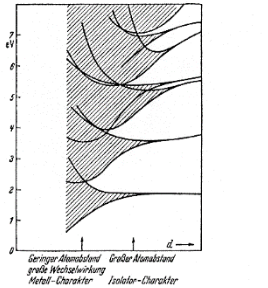


Abb. 14.52. Mit abnehmender Gitterkonstante *d* verbreitern sich die Elektronenzustände. Aus einem Isolator kann man durch hinreichende Kompression einen metallischen Leiter machen und umgekehrt (Rechnungen von *Slater* nach der Methode des selbstkonsistenten Feldes von *Hartree-Fock*). (Aus *W. Finkelnburg*)

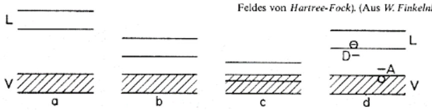


Abb. 648. Bändermodell (V = Valenzband, L = Leitungsband)
a) eines isolierenden Kristalls (großer Bandabstand)
b) eines Halbleiters ohne Störstellen mit Eigenleitung (kleiner Bandabstand)
c) eines Erdalkalimetalls (überlappende Bänder)
d) eines Halbleiters mit zwei verschiedenen Störstellen (A = Akzeptor, D = Donator)
Dieses Bändermodell ist stark vereinfacht. In Wirklichkeit ist die Breite der verbotenen Zone nicht konstant, sondern wird von der Gitterstruktur beeinflußt. Deshalb wird bei genauerer Betrachtung die Energie in Abhängigkeit vom Wellenvektor $\mathfrak{k}$ aufgetragen. Die Richtung des Wellenvektors ist die Richtung der Wellenfortpflanzung des Elektrons.

Themen und Begriffe

Stromleitung in Halbleitern, Bändermodell; p-n-Grenzschichten (p-n-Übergänge); Halbleiterdiode, Transistor.

Literatur

/1/ Bergmann-Schaefer; Lehrbuch der Experimentalphysik, Band IV, Teil 1; de Gruyter Berlin New York 1975

Physikalische Grundlagen

Bändermodell und Stromleitungsmechanismus in Halbleitern

Bei gebundenen Atomen treten infolge der Wechselwirkung mit den Nachbarn Verbreiterungen und Aufspaltungen der elektronischen Niveaus auf, die bei Molekülen zu komplizierten Bandenspektren und bei Festkörpern (Kristallen) schließlich zu quasikontinuierlichen *Energiebändern* führen, die von zustandsfreien, sogenannten *"verbotenen Zonen"* unterbrochen sind (siehe Abbildung auf der Titelseite).

Die Elektronen der inneren Schalen der Atome sind weiterhin an die jeweiligen Atomkerne gebunden (Atomrümpfe als Ionen). Die Bänder dagegen stellen eine Eigenschaft des gesamten Festkörpers dar. Die Elektronen in diesen Zuständen sind nicht lokalisiert und können sich unter Umständen quasi frei in dem Festkörper bewegen.

Auch die Bänder unterliegen, genau wie die diskreten atomaren Niveaus, dem *Pauli-Prinzip*, und können nur mit einer bestimmten, größten Anzahl von Elektronen besetzt werden. Voll besetzte Bänder können damit in einem äußeren elektrischen Feld keine weitere Energie aufnehmen, und ein einseitiger Ladungstransport und Strom als Träger von Bewegungsenergie bleibt ausgeschlossen (es bleiben aber Ortsveränderungen von Ladungen durch Platzwechsel möglich). Begrifflich bezeichnet man das oberste vollbesetzte Band in einem Festkörper, das nicht zur Leitfähigkeiten beitragen kann, als *Valenzband*, und das nächst höhere teilbesetzte oder leere Band als *Leitungsband*.

Eigenleitung

Isolatoren und Halbleiter sind Festkörper mit leeren Leitungsbändern. Bei Isolatoren ist die verbotene Zone zwischen Valenz- und Leitungsband sehr groß (> 2 eV), so daß bei Raumtemperatur entsprechend der Boltzmannverteilung für die thermische Anregung praktisch keine Elektronen aus dem Valenzband in das Leitungsband übergehen. Bei Halbleitern ist die verbotene Zone kleiner (0,56 eV bei Germanium), und schon bei Raumtemperatur wird ein zwar kleiner, aber merklicher Anteil von Elektronen durch thermische Anregung aus dem Valenzband in das Leitungsband angehoben. Diese Elektronen ermöglichen eine Leitfähigkeit des Festkörpers, die als *Eigenleitung* bezeichnet wird.

Bei dieser Eigenleitung hinterläßt jedes Elektron im Leitungsband ein Loch im Valenzband *("Defektelektronen")*, und die Dichten *n* dieser quasifreien Elektronen und *p* der Löcher sind gleich groß. Nach einer schwierigeren Rechnung, die im Rahmen des Praktikums nicht durchgeführt werden kann (siehe /1/, S. 674 ff) erhält man dabei als Abhängigkeit von der Temperatur:

$$n(T) \text{ bzw. } p(T) \propto T^{\frac{3}{2}} \cdot e^{-\frac{\Delta E}{2kT}} \qquad (1)$$

wobei ΔE der Abstand zwischen Valenz- und Leitungsband ist.

(Bei Metallen kommt es zu einem teilweisen Überlapp zwischen dem Valenzband und dem Leitungsband, so daß sich das Leitungsband unabhängig von der Temperatur bis zu einer bestimmten Höhe mit Elektronen auffüllt).

Störstellenleitung

Durch den Einbau geeigneter Fremdatome in einen Kristall, z.B. von drei- oder fünfwertigen Atomen in ein vierwertiges Wirtsgitter, entsteht ein sogenannter *Störstellenhalbleiter*.

Ein überschüssiges fünftes Elektron eines Fremdatoms (*Donator*; z.B. As) wird in einem vierwertigen Gitter nicht durch die Nachbaratome gebunden, und es genügt bereits ein geringer Energieaufwand, um dieses Elektronen von dem Atom abzulösen. Im Termschema befinden sich diese Elektronen in Niveaus dicht unterhalb des Leitungsbandes, von dem aus sie schon bei Raumtemperatur praktisch vollständig in dieses angehoben werden, und dort als negative, bewegliche Ladungsträger zur Verfügung stehen (*n-Halbleiter*).

Bei einem dreiwertigen Fremdatom (*Akzeptor*; z.B. Ga) kann die vierte Bindung der Nachbaratome des Wirtsgitters nicht abgesättigt werden. Es ist eine Fehlstelle mit einer Energie dicht oberhalb des Valenzbandes vorhanden, die bereits bei Raumtemepratur durch thermische Anregung besetzt wird. Im Valenzband bleibt ein Defektelektron (Loch) zurück, das als quasifreie positive Ladung für einen Ladungstransport zur Verfügung steht (*p-Halbleiter*).

p-n-Grenzschicht

(Siehe dazu folgende Abbildung). Stehen eine n- und eine p-Schicht in Kontakt miteinander, so diffundieren aufgrund der thermischen Bewegung an der Grenze Elektronen in das p-Gebiet und umgekehrt Löcher in das n-Gebiet. Sie treffen dort auf ihre jeweiligen komplementären Teilchen, mit denen sie zu neutralen Atomen rekombinieren. In der Grenzschicht entsteht ein Verlust an Ladungsträgern, wodurch sich in dem ursprünglich neutralen Material im n-Gebiet eine positive und im p-Gebiet eine negative Raumladung ausbildet, die mit dem dadurch entstehenden elektrischen Feld der Diffusion entgegenwirkt. Als Ergebnis stellt sich ein thermisches Gleichgewicht ein, bei dem der Diffusions- und der Feldstrom gleich groß sind.

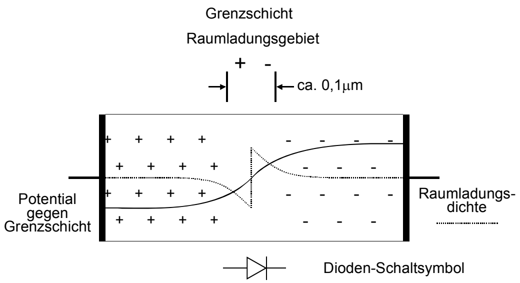


Abbildung p-n-Grenzschicht und Halbleiterdiode

Durch die Rekombination verarmt die Grenzzone an Ladungsträgern und bildet in dem Festkörper eine hochohmige *Sperrschicht*.

Durch Anlegen einer äußeren Spannung wird, je nach Polarität, die Sperrschicht durch weiteren Abzug von Ladungsträger verbreitert (*Sperrichtung*), oder durch Überfluten mit Ladungsträgern gleicher Polarität abgebaut, so daß auch der Übergang leitfähig wird (*Durchlaßrichtung*).

Halbleiterdiode

Eine solche Grenzschicht läßt bei Anlegen einer äußeren Spannung einen Stromfluß nur in einer Richtung zu. Legt man die äußere Spannung mit (+) an n und (-) an p, so werden weitere Ladungsträger aus den jeweiligen Gebieten abgezogen, und die Sperrschicht verbreitert sich; die Diode ist in *Sperrichtung* geschaltet. Polt man die äußere Spannung umgekehrt mit (-) an n und (+) an p, so werden der n- und p- Bereich von außen mit arteigenen Ladungsträgern überflutet, wodurch die Raumladungsbarrieren und die Sperrschicht abgebaut werden; die Diode ist in *Flußrichtung* geschaltet und leitet (siehe Abbildung auf Seite 4).

Die Funktion des Stromes in Abhängigkeit von der äußeren Spannung heißt Kennlinie der Diode. Im Rahmen einer allgemeinen Herleitung aus dem Bändermodell ergibt sich (*Shockley*-Diodengleichung):

(2) $$I = I_s \left(e^{\frac{e}{kT}U} - 1 \right)$$

wobei I_S der praktisch konstante Strom in Sperrichtung ist (Sperrstrom). Der Exponentialquotient e/kT wird als *Temperaturspannung* bezeichnet.

Experimentelle Kennlinien, insbesondere für Si oder GaAs, weichen z.T. von der *Shockley*-Beziehung ab und zeigen typischerweise ein Durchschalten in Flußrichtung erst ab einer bestimmten *Schwellspannung* oder auch zwei Bereiche mit unterschiedlichen Exponentialkoeffizienten (bzw. unterschiedlichen Anstiegs in der logarithmischen Darstellung).

Transistor:

Ein Transistor besteht aus einer dreifachen Halbleiter-Schichtfolge (p-n-p oder n-p-n), d.h. aus zwei "gegeneinander" geschalteten Dioden. Es gibt keine grundsätzlichen Unterschiede zwischen einer pnp- und einer npn-Schichtfolge. Jedoch werden aus technischen Gründen vor allem npn-Transistoren hergestellt, weshalb dieser Typ im folgenden dargestellt werden soll. Die drei Schichten bzw. Anschlüsse eines Transistors führen die Namen *Emitter*, *Basis* und *Kollektor* (siehe nebenstehende Abbildung).

Legt man bei einem npn-Transistor eine äußere Spannung mit (-) an den Emitter und mit (+) an den Kollektor, so fließt zunächst kein Strom wegen der Sperrschicht im Basis-Kollektor-Übergang. Mit einer zusätzlichen, positiven Spannung an der Basis ist die Emitter-Basis-Diode aber in Durchlaßrichtung geschaltet und es treten Elektronen aus dem Emitter in den Basisbereich ein. Ist die Basisschicht hinreichend dünn, so fließt aber nur ein kleiner Teil dieser Elektronen auch über den Basiskontakt ab. Der größte Teil dagegen diffundiert weiter in die Basis-Kollektor-Sperrschicht, wo sie in den Einflußbereich des Kollektorpotentials geraten und zum Kollektor hin abgesaugt werden.

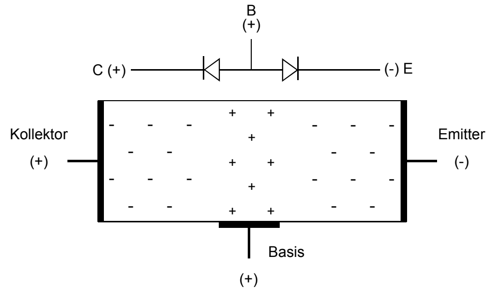


Abbildung Transistor

Ohne Basisanschluß ist der Transistor (Kollektor-Emitter-Strecke) wegen der Grenzschicht zwischen Kollektor und Basis gesperrt. Wird die Basis-Emitter-Strecke durchgeschaltet, so fließt nur ein kleiner Teil der Ladungsträger auch über die Basis ab. Der größte Teil dagegen diffundiert durch die dünne Basisschicht hindurch und wird dann durch den Kollektor gleichsam abgesaugt.

Die tatsächlich ablaufenden Vorgänge sind physikalisch kompliziert, und die obige Darstellung stellt nur eine grob anschauliche Vorstellung dar. Eine zentrale Rolle spielt die Dicke und die Dotierung der mittleren p-Schicht. Transistoren lassen sich so herstellen, daß 98 bis 99,9 % der Ladungsträger, die eigentlich über die Basis abfließen sollten, in den Kollektorkreis gelangen. Der Transistor stellt damit einen Stromverstärker dar, bei dem ein relativ kleiner Basis-Steuerstrom einen großen Kollektor-Laststrom regeln kann.