KING'S College LONDON

UNIVERSITY OF LONDON

Department of Physics

Second Year Physics
Laboratory Manual
2008/2009

Name: . . . .

Department of Physics

Second Year Physics
Laboratory Manual
2008/2009

Contents

Student Feedback

This manual is provided to help students carry out the experiments in the Second Year Laboratory class. We welcome suggestions for improvements to the style and content of the manual. Please pass your suggestions on to Dr Suhling.
In addition, it would be very helpful if you could tell the demonstrators of any errors or omissions that you discover within the manual, so that future versions of the manual can be corrected.

List of Experiments

Electronics

Optics

Others

Computer laboratory

Course organisation

1   Aim of the Course

The Practical Physics courses 5CCP211A (Second year Laboratory physics A) and 5CCP211B (Second year Laboratory physics B) are intended to provide single-honours and joint-honours students with experience of a broad range of techniques in experimental physics, and to develop skills in experimental data collection, analysis and interpretation. In addition, the course places a significant emphasis on teamwork skills and communication skills, requiring clear and concise descriptions of the experimental work, observations and interpretation, both in written and oral form.

2   Course Structure

The laboratory class runs on Tuesdays from 10:00 to 16:30. The experimental work is carried out in the undergraduate physics laboratory 4.66 in the Franklin Wilkins Building on the Waterloo Campus throughout both semesters. The computing experiments and computational simulations will be carried out in the Physics Department's Computer Laboratory on the Strand. Students usually work in pairs on the experiments to which they are assigned.

The class will be divided into two approximately equal groups and each group is expected to carry out an experiment every second week. The intervening weeks are to be used to complete the analysis of the experimental data, so **every student is expected to attend the lab class on every week that the course runs**. The usual pattern of laboratory work is therefore an alternating sequence of *Experiment* weeks and *Analysis* weeks, with one group of the class doing experiments while the other group completes the analysis of their data.

> **Students must submit their Laboratory Notebook for marking by the end of the Tuesday Lab Class of the *'Analysis'* week immediately following their experiment.**

Experiments will normally be assigned by a demonstrator one week in advance, so that there is time to read the description of the experiment in the Laboratory Manual, to do whatever background reading is necessary, and to make any other preparations that will ensure the experiment can be carried out efficiently in the scheduled laboratory time.

3   Lab Safety

In any scientific laboratory the most important aspect of its operation is the safety of the people working there.

> **All students should be familiar with the School notice, 'Safety in Laboratories', and the Departmental Safety Regulations. Both of these are included at page 31 of this manual. No student should start work in the laboratory before reading these documents.**

4   Course Requirements

- **5CCP211A & 5CCP211B: Single Honours students**
  Students taking 5CCP211A & 5CCP211B are expected to carry out 8 experiments in the year. In addition to the record of each experiment in their Laboratory Notebook, students will be expected to prepare formal written reports on 2 experiments of their choice, and will give an oral presentation on 1 other experiment.

- **5CCP211A: Mathematics & Physics Joint Honours students, Physics with Management or Computer Science**
Students taking 5CCP211A are expected to carry out 4 experiments in the year. In addition to the record of each experiment in their Laboratory Notebook, students will be expected to prepare a formal written report on 1 experiment of their choice, and will give an oral presentation on 1 other experiment.

Note that in order to progress to the 3rd year, it is compulsory to pass the laboratory class.

5   Laboratory Notebooks

Students usually work in pairs, but it is essential that every student keeps a full record of the experiments carried out in the laboratory in their own "Laboratory Notebook", **not** just one member of each pair. "Laboratory Notebooks" which may be in the form of a note book should contain both lined and graph paper. Computer generated output of results e.g. EXCEL graphs must be **firmly attached** to your note book or clipped into your file. Normally at least **two "notebooks" will be needed**, so that one may be in use while the other is being marked. Do not record the experiments on loose sheets of paper.

Before leaving the laboratory at the end of an *Experiment* day each student must have their "notebook" signed by a member of staff. Without this signature a Notebook cannot subsequently be accepted for marking. At the end of an *Analysis* day your "notebook" must either be handed in for marking or signed by a member of staff to show that satisfactory progress has been made.

6   Attendance sheet

Students **must sign an attendance sheet** on *every day* that they attend the Laboratory class, irrespective of whether it is scheduled to be an *Experiment* or an *Analysis* day. **At the beginning of an *Experiment* day each student must make sure that she/he is registered for a particular experiment. Only experiments for which you have been registered can be counted.**

7   Outline Laboratory Schedule for 5CCP211A & 5CCP211B

Week No.	Date	Group 1	Group 2	Notes
1	23 Sep 2008	10h–14h	14h–18h	Class Registration and Graph plotting exercises
2	30 Sep 2008	Error analysis		Introduction and revision of error analysis
3	7 Oct 2008	E-1		
4	14 Oct 2008	A-1	E-1	
5	21 Oct 2008	E-2	A-1	
6	28 Oct 2008	A-2	E-2	
7	4 Nov 2008	Reading Week		
8	11 Nov 2008	E-3	A-2	
9	18 Nov 2008	A-3	E-3	
10	25 Nov 2008	E-4	A-3	
11	2 Dec 2008	A-4	E-4	
12	9 Dec 2008	Class talks		Report due Friday 12 Dec 2008
13	6 Jan 2009	Exam Week		
14	13 Jan 2009	E-5	A-4	
15	20 Jan 2009	A-5	E-5	
16	27 Jan 2009	E-6	A-5	
17	3 Feb 2009	A-6	E-6	
18	10 Feb 2009	E-7	A-6	
19	17 Feb 2009	A-7	E-7	
20	24 Feb 2009	E-8	A-7	
21	3 Mar 2009	A-8	E-8	
22	10 Mar 2009	Class Talks	A-8	
23	17 Mar 2009		Class Talks	Report 2 due Friday 20 Mar 2009

Table 1: Outline Laboratory Schedule for 5CCP211A & 5CCP211B for 2008/2009

Note: An entry of the form **E-3** in Table 1 means that the relevant subgroup of the class must carry out their 3rd experiment on that particular Lab day.

An entry of the form **A-5** means that the relevant subgroup of the class must complete the analysis of their 5th experiment on that day, and hand it in for marking.

Computer laboratory organisation

1 Aims and objectives

This course presents some of the techniques, practices and pitfalls in the computational modelling of physical systems. The students will learn how to choose the most appropriate method for the solution of various problems. They will put some of this into practice using a spreadsheet or by writing computer programs. They will learn about the problems that computational techniques are subject to: rounding errors, stability, slow convergence and inefficient methods. They will learn the use of numerical techniques: simultaneous equations, numerical integration, minimisation, Monte Carlo techniques, simulated annealing, finite difference equations, Runge-Kutta.

2 Learning outcomes

By the end of the course, the students should be able to choose the most efficient technique to model physical systems, to formulate the problem, and to implement a reliable solution.

3 General notes

3.1 Advantages of different packages

When do you use Excel, Maple or program the problem yourself?

Package	Advantages	Disadvantages
Maple/Mathematica	Analytical maths. Good numeric packages Excellent plotting.	Slow and inflexible for complex numeric calculations. Doesn't tell you how it is doing it!
IDL, MATLAB	Very sophisticated and quick programming for complex mathematical/physical problems, good libraries, excellent/easy plotting.	Other packages (of limited outside usefulness) to learn. Don't tell you how they are doing it!
Excel	Easy to use for simple problems, good plotting, some sophisticated programming.	Slow and inflexible for complex numeric calculations. Very limited serious numeric libraries.
C/C++ etc.	Totally versatile. Good libraries. Already know the language (?)	Graph plotting more difficult.

Table 1: Advantages and disadvantages of different packages.

3.1.1 Other considerations

It is always quicker, and more reliable, to use someone else's numerical solution than to write your own. For this reason, you should use libraries of numerical methods whenever possible. Because Excel is not designed for physics, its library of functions is not very extensive. There are a number of numerical libraries which you may be able to get at, for example the NAG library is excellent. IDL has a huge library of numeric functions. Matlab is also good, but like IDL, is almost a separate programming language. So, when you come to solve **real problems**, use what you have learned in this course to select a suitable method, but use a library function for any but the most simple problems. **In the computer laboratory, though, you will be programming the solutions yourself using a programming language or Excel**.

3.1.2   For this course

Maple, Mathematica, IDL, MATLAB do a very good job of solving problems numerically, but they do not give you the choice of how it is to be done. Since the whole object of this course is to understand the numerical techniques, **Maple, Mathematica, IDL and MATLAB solutions to the problems are NOT acceptable in the the computer laboratory**. You can use these packages to check your solutions which you derived using Excel or C/C++, Fortran, Basic etc.

3.2   Programming languages

You should have some memory of the first year C course, or if you did not do that, some familiarity with Fortran, C++, Java or Pascal (or perhaps Basic), but that is not a prerequisite. You should all be able to use Excel. The exercises are not language-specific, but I recognise that most of you will be using C or Excel.

3.3   Units

S.I. units give a nice self-consistent system, in which all the constants are tabulated and all the equations work. However, in many problems the magnitudes of the constants make the calculations quite difficult to manage. For example there is a factor of $\hbar^2$ in the Schrödinger Equation, which immediately brings a factor of 10^{-68} into the calculation (see above). Similar problems arise in astronomical calculations. It is always easier to deal with numbers that are close to 1. Therefore, many fields have their own systems of constants, which are much more manageable than SI units, using a set of units of the right sort of size for the problem.

Examples:

1 eV $= 1.602 \times 10^{-19}$ J

1 Å $= 10^{-10}$ m

1 light year $= 9.47 \times 10^{15}$ m

1 parsec $= 3.26$ light years

1 mass unit for nuclear physics $(\text{MeV}/c^2) = 1.763 \times 10^{-30}$ kg

1 length unit for nuclear physics (fm) $= 10^{-15}$ m

1 atomic mass unit $= 931.5 \,\text{MeV}/c^2 = 1.661 \times 10^{-27}$ kg

Sometimes they define more specialised units such that key constants are equal to one.

Example: Atomic units

$\hbar = e = m_e = 1$

Then:

1 a.u. (of energy) $= 27.21$ eV $= 4.36 \times 10^{-18}$ J

1 a.u. (of length) $= 0.0529$ nm

1 a.u. (of mass) $= 9.11 \times 10^{-31}$ kg

You should take extreme care with the units you use in any calculation, if possible keeping the numbers you use near unity.

3.4   Accuracy and rounding errors

All computers store integers and real (floating point) numbers in different ways. The standard size for both types of number is 4 bytes (32 bits), but this can be extended. Excel uses 8 byte floating point numbers (*double* in C). Whether you use this precision or the 8 or even 16 byte length, you need to understand how they are stored to appreciate the pitfalls.

Integers (*int's*) are stored as exact binary numbers. One bit is used for the sign, 31 bits for the number (for 4 byte integers). All operations are **exact**, except that the result of division has to be an integer (rounded down). However, 4 byte integers cannot have magnitudes greater than $2^{31} - 1 \approx 2.14 \times 10^9$.

Reals (*float's*) are stored as a sign, s, (one bit), an exponent, e, (one byte including 1 bit for its sign), and a mantissa, M, (23 bits), that is $s \times M \times 2^e$. That's for a 4-byte float. Note that the largest exponent to be stored in 1 byte (less 1 bit for the sign) is $2^7 - 1 = 127$, and 2^{127} 1.7×10^{38}, so the largest magnitude stored by a 4-byte float is 1.7×10^{38}, and smallest in magnitude is 1.7×10^{-38}. The 23 bit mantissa corresponds to about 7 s.f. so that limits the precision of your calculation.

To illustrate the main problem, e.g. add 3 and 0.000000237. Stored as (in decimal, but the idea's the same) $(0.)3 \times 10^1 + (0.)000000237 \times 10^1$, but it can only fit 7 s.f. into the mantissa, so the least significant get lost, and the result of the addition is: 3.0000002.

This may not matter, but suppose the next step is to subtract 3 from the result giving 2×10^{-7} instead of the 2.37×10^{-7} you might have expected.

These are called **rounding errors**. Usually, they only affect the last few significant figures of your result, but they can become troublesome. Improve the situation by:

- Using increased precision of your real numbers (double precision).

- Being very careful that you algorith

Rounding errors are about storing numbers in binary notation, which are not exact binary. You have the same problem with decimal—just think of 1/3.

A similar problem is due to **truncation errors**. Computers calculate many functions as the summation of a series. An exact result can only be obtained by summing an infinite number of terms of the series. Usually, it is sufficient to sum a convergent series until the terms are considerably smaller than the precision of the answer required, but if the series is very slowly converging, more care must be employed.

Example

Sum the series $\sum_{n=1}^{N} n^{-2}$ until the next term is smaller than: 10^{-2}, 10^{-3}, 10^{-4}. This series is convergent to a sum of $\pi^2/6$ but it is very slowly convergent.

	Full Excel 8 byte precision	Only 6 decimal places	
Actual answer =	1.644934		
10 terms	1.549768	1.549768	Rounding errors
100 terms	1.634984	1.634983 ⟵	creeping into final
1000 terms	1.643935	1.643917	decimal place
10000 terms	1.644834	1.643917	already
65335 terms	1.644919	1.643917	

If only 6 decimal places, all terms after the 1000$^{\text{th}}$ are zero, so sum does not get better.

See the file `rounding-truncation.xls` on filespace. Run the C program `series-cp2720.cpp` which is a better illustration because it only uses 4 byte floats so the rounding errors are more apparent.

What do you actually see?

The errors come up as "overflow" errors for floats in C which will crash your program when the number is too big (i.e. try to store a number $> 1.7 \times 10^{38}$) and "underflow" when you try to store a number smaller than 1.7×10^{-38}, which will be rounded down to 0 with no warning, usually; in Excel, the overflows come up as #NUM and the underflows just get rounded to 0 (but since Excel uses 8-byte numbers, these problems are rarer). Worse, if you try to store an int $> \sim 10^9$ in C, there is no error message, it just forgets the most significant part of the int (leading to bizarre results!). However, for most errors of this type (rounding errors, truncation errors, underflows, careless use of integer arithmetic), you don't see anything—the results are just wrong!

References

[1] William Press, Saul Teukolsky, William Vetterling, and Brian Flannery. *Numerical Recipes in C*. Cambridge University Press, Cambridge, UK, 2nd edition, 1992.

[2] William J. Orvis. *Excel for Scientists and Engineers*. SYBEX Inc., Alameda, CA, USA, 1995.

Assessment and marking

1   General Policy

The mark for the Laboratory Physics courses is based on a continuous assessment of the work carried out throughout the course.

5CCP211A : The error exercise and the graph plotting exercise will count **4%** each towards the final mark. One write-up in the **Laboratory Notebook** will count for **12%** of the available marks, so 4 experiments are worth **48%**. The formal report counts for **20%** of the mark, and **4%** are awarded for computer skills. The remaining **20%** are based on a **Class Oral Presentation** at the end of the year.

5CCP211B : One write-up in the **Laboratory Notebook** will count for **12%** of the available marks, so 4 experiments are worth **48%**. The formal report counts for **20%** of the mark, and **4%** are awarded for computer skills. The mark will then be scaled up to 100%.

In Sections 5, 7 and 9 there are more detailed descriptions of how Lab Notebooks, Formal Reports and Oral Reports will be assessed.

2   Plagiarism

Students will normally carry out experiments in pairs, so it is inevitable that the data collected in the course of an experiment will be common to that pair. Similarly, the free discussion of the *methods* of data analysis is encouraged, but this collaboration and cooperation should not extend to the final stages of preparing work that is submitted for assessment. **All work that is handed in for marking should be the student's own work**, and **any instance of copying or paraphrasing another student's work will be considered** an attempt at **cheating** and will be dealt with accordingly. All students should have signed a "Statement on Plagiarism" when joining the Physics Department. A more complete statement of the Department's policy in dealing with plagiarism can be found in the current edition of the Physics Department's Undergraduate Handbook.

3   Leica Microsystems Prize

The **Leica Microsystems Prize** is awarded annually to "*the student with the best performance in the 2nd year physics laboratory class for a single or joint honours degree involving physics as a main subject*". Normally the student with the highest mark is proposed for the prize. The prize is worth **£200**. Leica Microsystems have given the prize since 2005 in recognition of fruitful association with the Physics Department and is an indication of the importance that they attach to the training received by students in the practical laboratory.

4   Laboratory Notebooks

Guidelines for the use of Laboratory Notebooks

The Laboratory Notebook forms an essential and complete record of the work that you carry out in the Lab class.

1. It is *not* necessary to include in this record any general information about the experiment that is already available in the Laboratory Manual (e.g. manufacturer of equipment, or insignificant statements about plugging it in or switching it on), but you **must include** the theoretical background and a brief context. State why the experiment is important and what it illustrates.

2. It should contain all the specific experimental details that would be needed to let you recall clearly and accurately how you carried out the experiment. For example, it should include:

   (a) the **title** of the experiment and the **date** on which it was performed;

   (b) an introductory sentence stating the **main objectives and purpose** of the experiment;

   (c) a **risk assessment**;

   (d) a **sketch** showing the layout of the **apparatus**, with dimensions if necessary, or a **circuit diagram**, as well as a note about the appropriate settings for the instruments;

   (e) the basic theoretical background and context of the experiment in the field of physics;

   (f) **comments regarding the operation of equipment**, particularly if you used it in a way that differed from the procedure recommended in the Lab Manual;

   (g) **tables of raw data, graphs** (with error bars), and perhaps a summary table showing the **final results** produced by your analysis, along with calculated values for the **experimental uncertainties**;

   (h) details of the algebra and arithmetic used in the **calculations**; if the data plotted in a graph are expected to follow a straight line, then a **weighted least-squares fit** should be carried out **to determine the slope and intercept** of the line, the standard deviation of gradient and intercept, as well as the associated experimental errors;

   (i) short paragraph **commenting on your results** and identifying the **conclusions** that you can draw from the experimental results, including a **comparison** of your results **with literature values**.

3. The *'Conclusions'* paragraph should be brief, but should try to address the following points:

   (a) **Did the experiment achieve the objectives** set out at the start of the experiment?

   (b) If you measured the same experimental quantity more than once, or in more than one way, **did the different measurements agree** within experimental error? *If not, then suggest reasons why not.*

   (c) **Did the measured values agree with accepted values**, either theoretically predicted values, or other experimental values published in reference books? *If not, then suggest reasons why not.*

   (d) What were **the main sources of experimental error**? How large were these errors in relation to the overall uncertainty in the final results?

   (e) Was there any evidence for a **systematic experimental error**? (e.g. were all the measured values higher than the accepted values?)

   (f) **How could the experimental errors be reduced**?

   (g) **Could the experiment be improved** in some way?

4. Wherever possible, **graphs should be plotted as the experiment proceeds**. This will indicate whether the spacing of data points is appropriate, whether data are sufficiently accurate, and help you to spot problems with your experiment at an early stage, so that you do not waste time collecting spurious data.

5. The Lab Notebook should be the first and only place where you record experimental data and the subsequent analysis. As it is intended to record the work you do in the Lab as you do it, it is not necessary for it to be extremely neat, although it is important for you to annotate your work sufficiently for the contents to be clear.

   (a) **Do not re-write** sections of your notebook to try and make them neater.

   (b) **Do not record your original data or analyses on loose sheets of paper**, with the intention of later transcribing them into your notebook.

   (c) **Do not spend time** outside the Lab **writing long accounts** of every experiment.

5 Marking scheme for laboratory notebooks

Each record of an experiment in the Laboratory Notebook will contribute about 12% towards the final mark awarded for the Lab class. In assessing the experiment record, marks will be awarded for each of the following categories. Some indication of the points the markers will bear in mind is also given below.

Note that the weighting applied to each of the categories is usually the same, so the final assessment of an experiment record depends at least as much on the analysis as it does on the acquisition of the data.

1. *Quantity of data*

    (a) Did the student complete all parts of the experiment?

    (b) Were sufficient data recorded (**and signed off by a demonstrator**) to allow a satisfactory analysis?

    (c) Do the graphs show any obvious gaps in the range of data collected?

    (d) Is there enough information to allow the experimental conditions to be reproduced at a later date, using only the Lab Notebook? Is the basic theory included? Is there a diagram of the experimental set up?

2. *Quality of the data*

    (a) Have the correct measurements been made?

    (b) Do the experimental data show the behaviour expected?

    (c) Were there any problems with the functioning of the apparatus?

3. *Analysis of the data*

    (a) Has the analysis been carried out correctly?

    (b) Are the calculations clearly set out so that the steps are easy to check?

    (c) Are there any arithmetical errors?

4. *Assessment of errors and uncertainties*

    (a) Have experimental uncertainties been assessed quantitatively, using the standard methods for propagation of errors?

    (b) Are there comments identifying the most important sources of error?

    (c) Have any cross checks or tests of the consistency of the data been carried out?

5. *Concluding discussion of the results*

    (a) Have the results been compared with accepted and/or theoretical values?

    (b) Are the conclusions justified by the results?

    (c) Are there suggestions for improving the experiment?

Feedback is provided in written and in oral form by the demonstrators when the marked lab notebook is returned—**read it, listen to it and act on it** for the next experiment.

All lab notebooks are to be handed in at the end of the course in March.

6   Formal Written Reports

Guidelines for preparing a Formal Report

The primary purpose in preparing Formal Reports is to allow you to develop a concise and accurate style in writing factual accounts of the work you have carried out and to communicate the results and conclusions—a skill that will be essential not just in a scientific career but in almost any future context where you have to prepare written summaries of your activities. A good scientific report is one that enables the reader to understand clearly what was done, and to repeat the experiment being described.

The formal report is on one of the experiments of your choice. The information needed to begin preparing the report is available in the Lab Manual, and in your own record of the experiment in your Laboratory Notebook. The report will be produced using a word processor, with graphs and diagrams also prepared in electronic form.

1. **Do not** simply **copy** or paraphrase **the Laboratory Manual**.

2. Use **grammatically correct** sentences, use a **spell-checker** or a **dictionary** to ensure your spelling is correct, and **check carefully for typographic errors**. **Proof read the report before you submit it.**

3. Try to use a **matter-of-fact** and **impersonal style** of writing.

   (a) The report should not be in the form of a diary, nor written as a narrative. **Avoid** statements such as: "*First we assembled the apparatus, then we switched it on and we started to make measurements…*" or "*We were very happy with the results we got…*"

   (b) Use the **passive voice** to describe most of the events: e.g., "*The current was measured as a function of temperature…*" instead of "*I heated the sample and measured how the current changed…*"

   (c) Remember that the report is written with the **benefit of hindsight**. You do not need to report the results in the order that you made the measurements if they can be presented more logically in a different order.

4. Organise the report into logically distinct sections, such as, for example, ***Introduction***, ***Theory***, ***Experimental Procedure***, ***Results and Discussion*** and ***Conclusions***. It is usually helpful to number the sections.

5. Include the **Title** of the experiment, the **Author**'s name, and a short **Abstract** (i.e. a summary, in less than 100 words) of the report on the first page.

6. In the introduction, briefly state the **aims** of the experiment, indicate its significance in the context of physics and outline any relevant background information.

7. Symbols and abbreviations should be defined the first time they are used in the report. For example: "*The high voltage* (HV) *setting of the photomultiplier tube* (PMT) *was adjusted…*", or "*The wavelength* λ *is given by the formula…*"

8. Include only sufficient theory for the report to be understood.

   (a) **Do not reproduce pages of standard theory**. If necessary, cite a reference to this work, or use an appendix.

   (b) Do **include** the **key equations** that are needed as part of the analysis, making sure that you **number** all **equations** that are referred to later in the text.

   (c) If you use a novel theoretical analysis, this should be explained in full.

9. Describe in your own words the **apparatus** and the **experimental procedure**, in enough detail for a reader to understand what you did and how you did it.

10. All diagrams should be labelled as ***Figures***, should have captions stating what is shown and should be numbered consecutively.

11. A **clear and logical presentation of the results** is very important.

   (a) **Tables of raw data should not be included in a report**. A graph is a much more efficient and easily interpreted way of presenting the same data, if they need to appear in the report.

   (b) Graphs should have clearly **labelled axes**, and data points should be plotted with **error bars**.

   (c) All graphs are ***Figures***, should have captions stating what is shown and should be numbered appropriately.

   (d) Make clear the formulae and methods being used in the analysis, but details of the intermediate **arithmetic steps should not be included** in the report.

   (e) Always include an estimate of the **experimental error** associated with measured values, and with results derived from these values. Use appropriate numbers of significant figures: for example, 3.14 ± 0.03, **not** $3.1415926535 \pm 0.0271828183$

   (f) A **summary table of the final results** is often useful. It should be numbered and have a caption stating what is shown.

12. State clearly the main **conclusions** that can be drawn from the results.

   (a) Did the experiment achieve its aims?

   (b) Do the results agree with theory or with accepted values? If not, explain why not.

   (c) What are the most significant sources of error?

13. Suggest ways to **extend or improve the experiment**.

    (a) What additional measurements would increase the accuracy of the results?

    (b) Is there a better way of collecting the same data?

    (c) Is there another (better) way of measuring the same experimental quantities?

    (d) Can the data analysis be improved?

14. Always include a list of **references** at the end of the report.

    (a) The references must be cited in the report.

    (b) Make sure you include the sources of any accepted values that you made use of.

    (c) References should always specify the authors, the book or journal title (and volume number, if appropriate), page numbers, and year of publication. For books you should also specify the publisher, place of publication and edition if appropriate. For journal articles, you can optionally include the title of the specific article referred to.

7 Marking scheme for Formal Reports

Each Formal Report will contribute 20% towards the final mark awarded for the Lab class. In assessing the report, marks will be awarded for each of the following categories. Some indication of the points the markers will bear in mind is also given below.

1. *Introduction and Theory*

    (a) Is there an adequate introduction and description of the relevant theory, referenced appropriately?

    (b) Has there been an attempt to place the experiment in a wider context?

2. *Apparatus and Experimental Method*

    (a) Is there a diagram of the apparatus, labelled and with a caption?

    (b) Has the experimental method been described clearly and concisely?

    (c) Is there sufficient information to allow the reader to repeat the experiment?

3. *Presentation of Data*

    (a) Are the results presented clearly and logically?

    (b) Do all figures (diagrams and graphs) have captions and are they numbered in sequence?

    (c) Are the graphs neat, with labels on each axis and error bars on the data points?

4. *Analysis of Results*

    (a) Is there a sensible discussion of the results and their associated experimental errors?

    (b) Have any cross-checks or tests of the consistency of the experimental data been carried out?

    (c) Have the experimental results been compared with accepted values, referenced appropriately?

    (d) Has the student stated the conclusions that can validly be drawn from the data?

5. *Novel Information*

    (a) Has the report provided any additional information that was not in the Lab Manual?

    (b) Are there suggestions for ways to improve the experiment? For example to make it more accurate or easier to carry out.

    (c) Are there suggestions for ways in which the experiment could be extended?

6. *Late Penalties*
   Unless there are valid medical or other reasons (with supporting evidence) for late submission, Formal Reports that are submitted for marking after the published deadlines will not be marked.

8   Assessment of computer skills

It is assumed that all students will have been instructed in the use of software packages such as `Word` and `Excel` in the First Year, and are now reasonably proficient with them. If you have entered the Second Year without having received this instruction, you should inform a staff demonstrator at once.

In addition to the mark awarded for the content of a formal report, a mark of 4% will be awarded for the use of computer skills demonstrated in the production of each Formal Report. Specifically, a well-presented report should include evidence of some or all of the following skills:

1. *Word processing* (50% of the available marks)

   (a) Is the document neatly formatted ("justified"), with a consistent layout and use of fonts? Are the pages numbered? (30%)

   (b) Have Greek letters and Maths symbols been used? Have subscript and superscript been used appropriately? (10%)

   (c) Have the equations been typeset (e.g. using `Word`'s Equation Editor)? (10%)

2. *Data analysis and graph plotting* (30% of the available marks)

   (a) Have any graphs been plotted by computer, with appropriate labels and legends, and have they been inserted into the text appropriately?

   (b) Have summary tables of final results (*not raw data*) been produced by `Word` or `Excel`?

   (c) Have least squares fits to experimental data been carried out by computer?

   (d) Were any theoretical simulations carried out using for example `Excel`?

3. *Diagrams* (20% of the available marks)

   Have any diagrams been drawn by a computer drawing package (such as the drawing package included with `Word`)?

Feedback is provided in written and in oral form by the demonstrators when the marked lab notebook is returned—read it, listen to it and act on it for the next Formal Report (if applicable).

The Class Talk

In this formal presentation the student is expected to describe one of the laboratory experiments to the rest of the class. In this case it is expected that the material will be presented by both students together using a presentation on a computer, or an overhead projector, and that the necessary visual aids will have been prepared by the students in advance of the talk. A whiteboard is also available. The talk is expected to last about 15–20 minutes, at the end of which questions will be invited from the audience, including both other students and the demonstrators. The talk counts for 20% of the overall mark.

Guidelines for the Class Talk

While doing the presentation, students should keep in mind the following advice from previous lab demonstrators:

- Smile! You all have lovely smiles and it would be nice to see them while you are speaking, it cheers your audience up.

- Do not chew chewing gum while presenting, it looks awful.

- Try not to read straight from a piece of paper. It is fine to have a script, but you should look up at your audience every now and make eye contact.

- Think about how you are standing and try not to have your back to the audience all of the time.

- When speaking about the slides try to add extra information rather than just reading everything straight from the slide.

Whereas it is perfectly acceptable to use transparent visual aids on a overhead projector, the students are strongly advised to use a `PowerPoint`, `OpenOffice Impress`, `Apple Keynote`, `KOffice KPresenter` or `PDF` presentation. Please note that the `PDF` format can also include videos, animations and slide transitions, just as other presentation format. Should the students whish to use LaTeX to prepare their presentation, they should consider using the `beamer` document class [1], which easily produces high quality presentation.

Here are some general guidelines for creating presentations (from the Beamer user guide [1], with permission of the author. Students are strongly advised to read the Section I.5 of this guide, even if they are not using LaTeX and `beamer` to produce their slides):

1. **Know the Time Constraints**
   When you start to create a presentation, the very first thing you should worry about is the amount of time you have for your presentation. Depending on the occasion, this can be anything between 2 minutes and two hours.

   - A simple rule for the number of frames is that you should have at most one frame per minute.

   - In most situations, you will have less time for your presentation that you would like.

   - *Do not try to squeeze more into a presentation than time allows for.* No matter how important some detail seems to you, it is better to leave it out, but get the main message across, than getting neither the main message nor the detail across.

   In many situations, a quick appraisal of how much time you have will show that you won't be able to mention certain details. Knowing this can save you hours of work on preparing slides that you would have to remove later anyway.

2. **Global Structure**
   To create the "global structure" of a presentation, with the time constraints in mind, proceed as follows:

   - Make a mental inventory of the things you can reasonably talk about within the time available.

   - Categorize the inventory into sections and subsections.

   - Do not feel afraid to change the structure later on as you work on the talk.

   - Ideally, a table of contents should be understandable by itself. In particular, it should be comprehensible *before* someone has heard your talk.

   - Keep section and subsection titles self-explaining.

   - Both the sections and the subsections should follow a logical pattern.

   - Begin with an explanation of what your talk is all about. (Do not assume that everyone knows this. The *Ignorant Audience Law* states: Someone important in the audience always knows less than you think everyone should know, even if you take the *Ignorant Audience Law* into account.)

   - Then explain what you or someone else has found out concerning the subject matter.

   - Always conclude your talk with a summary that repeats the main message of the talk in a short and simple way. People pay most attention at the beginning and at the end of talks. The summary is your "second chance" to get across a message.

3. **Frame Structure**
   Just like your whole presentation, each frame should also be structured. A frame that is solely filled with some long text is very hard to follow. It is your job to structure the contents of each frame such that, ideally, the audience immediately seems which information is important, which information is just a detail, how the presented information is related, and so on.

   - Put a title on each frame. The title explains the contents of the frame to people who did not follow all details on the slide.

   - The title should really *explain* things, not just give a cryptic summary that cannot be understood unless one has understood the whole slide. For example, a title like "The Poset" will have everyone puzzled what this slide might be about. Titles like "Review of the Definition of Partially Ordered Sets (Posets)" or "A Partial Ordering on the Columns of the Genotype Matrix" are *much* more informative.

   - Ideally, titles on consecutive frames should "tell a story" all by themselves.

- In English, you should *either always* capitalize all words in a frame title except for words like "a" or "the" (as in a title), *or* you *always* use the normal lowercase letters. Do *not* mix this; stick to one rule. The same is true for block titles. For example, do not use titles like "A short Review of Turing machines." Either use "A Short Review of Turing Machines." or "A short review of Turing machines." (Turing is still spelled with a capital letter since it is a name).

- In English, the title of the whole document should be capitalized, regardless of whether you capitalize anything else.

- Use columns.

- *Never* use footnotes. They needlessly disrupt the ow of reading. Either what is said in the footnote is important and should be put in the normal text; or it is not important and should be omitted (*especially* in a presentation).

- Do not use long bibliographies.

4. **Writing the Text**

- Use short sentences.

- Prefer phrases over complete sentences. For example, instead of "The figure on the left shows a Turing machine, the figure on the right shows a finite automaton." try "Left: A Turing machine. Right: A finite automaton." Even better, turn this into an itemize or a description.

- Punctuate correctly: no punctuation after phrases, complete punctuation in and after complete sentences.

- *Never* use a smaller font size to "fit more on a frame". Never ever use the evil option shrink.

- Do not hyphenate words. If absolutely necessary, hyphenate words "by hand".

- Break lines "by hand". Do not rely on automatic line breaking. Break where there is a logical pause. For example, good breaks in "the tape alphabet is larger than the input alphabet" are before "is" and before the second "the". Bad breaks are before either "alphabet" and before "larger".

- Text and numbers in figures should have the *same* size as normal text. Illegible numbers on axes usually ruin a chart and its message.

5. **How Much Can I Put On a Frame?**

- A frame with too little on it is better than a frame with too much on it. A usual frame should have between 20 and 40 words. The maximum should be at about 80 words.

- Do not assume that everyone in the audience is an expert on the subject matter. Even if the people listening to you should be experts, they may last have heard about things you consider obvious several years ago. You should always have the time for a quick reminder of what exactly a "semantical complexity class" or an "ω-complete partial ordering" is.

- Never put anything on a slide that you are not going to explain during the talk, not even to impress anyone with how complicated your subject matter really is. However, you may explain things that are not on a slide.

- Keep it simple. Typically, your audience will see a slide for less than 50 seconds. They will not have the time to puzzle through long sentences or complicated formulas.

- Lance Fortnow claims: PowerPoint users give better talks. His reason: Since PowerPoint is so bad at typesetting math, they use less math, making their talks easier to understand. There is some truth in this in my opinion. The great math-typesetting capabilities of TeX can easily lure you into using many more formulas than is necessary and healthy. For example, instead of writing "Since $|x \in 0,1|x \sqsubseteq y| < \infty$, we have…" use "Since y has only finitely many prefixes, we have…"

  You will be surprised how much mathematical text can be reformulated in plain English or can just be omitted. Naturally, if some mathematical argument is what you are actually talking about, as in a math lecture, make use of TeX's typesetting capabilities to your heart's content.

6. **Using Graphics**
   Graphics often convey concepts or ideas much more efficiently than text: A picture can say more than athousand words. (Although, sometimes a word can say more than a thousand pictures.)

- Put (at least) one graphic on each slide, whenever possible. Visualizations help an audience enormously.

- Usually, place graphics to the left of the text. (Use columns.) In a left-to-right reading culture, we look at the left first.

- Graphics should have the same typographic parameters as the text: Use the same fonts (at the same size) in graphics as in the main text. A small dot in a graphic should have exactly the same size as a small dot in a text. The line width should be the same as the stroke width used in creating the glyphs of the font.

- While bitmap graphics, like photos, can be much more colorful than the rest of the text, vector graphics should follow the same "color logic" as the main text (like black = normal lines, red = hilighted parts, green = examples, blue = structure).

- Like text, you should explain everything that is shown on a graphic. Unexplained details make the audience puzzle whether this was something important that they have missed. Be careful when importing graphics from a paper or some other source. They usually have much more detail than you will be able to explain and should be radically simplified.

- Sometimes the complexity of a graphic is intensional and you are willing to spend much time explaining the graphic in great detail. In this case, you will often run into the problem that fine details of the graphic are hard to discern for the audience. In this case you should create anticipated zoomings of interesting parts of the graphic.

7. **Using Animations and Transitions**

- Use animations to explain the dynamics of systems, algorithms, etc.

- Do *not* use animations just to attract the attention of your audience. This often distracts attention away from the main topic of the slide. No matter how cute a rotating, flying theorem seems to look and no matter how badly you feel your audience needs some action to keep it happy, most people in the audience will typically feel you are making fun of them.

- Do *not* use distracting special effects like "dissolving" slides unless you have a very good reason for using them. If you use them, use them sparsely. They can be useful in some situations: For example, you might show a young boy on a slide and might wish to dissolve this slide into slide showing a grown man instead. In this case, the dissolving gives the audience visual feedback that the young boy "slowly becomes" the man.

8. **Colors and fonts**

- Use colors sparsely.

- Be careful when using bright colors on white background, *especially* when using green. What looks good on your monitor may look bad during a presentation due to the different ways monitors, beamers, and printers reproduce colors. Add lots of black to pure colors when you use them on bright backgrounds.

- Maximize contrast. Normal text should be black on white or at least something very dark on something very bright. *Never* do things like "light green text on not-so-light green background".

- Inverse video (bright text on dark background) can be a problem during presentations in bright environments since only a small percentage of the presentation area is light up by the beamer. Inverse video is harder to reproduce on printouts and on transparencies.

- Choose carefully the size of the font. Depending on how far your audience is removed from the projection and on how large the projection is, between 10 and 20 lines should fit on each slide. The less lines, the more readable your text will be.

- In a presentation, you will want to use a very small font for text in headlines, footlines, or sidebars since the text shown there is not vital and is read at the audience's leasure. Naturally, the text should still be large enough that it actually can be read without binoculars.

- In the other direction, you can use larger text for titles. However, using a larger font does not always have the desired effect. Just because a frame title is printed in large letters does not mean that it is read first. In essence, large text at the top of a page signals "unimportant since I know what to expect". So, instead of using a very large frame title, also consider using a normal size frame title that is typeset in bold or in italics.

- One of the most important rules of typography is that you should use as little fonts as possible in a text. In particular, typographic wisdom dictates that you should not use more than two different families on one page.

9 Marking scheme for Oral Reports

In assessing the oral presentations, the markers will bear in mind the following points. Use them as guidelines for preparing your talk.

1. *Style of presentation*

    (a) Did the student speak clearly and audibly, facing the audience?

    (b) Did the student make efficient use of the time available for the talk?

    (c) Was the sequence of the presentation easy to follow?

    (d) Were the visual aids clear and effective in assisting the presentation?

2. *Technical understanding*

    (a) Did the student contextualise the experiment and state its significance?

    (b) Did the student show a confident knowledge of the material?

    (c) Did the student describe clearly the aims of the experiment?

    (d) Did the student describe clearly how the experiment was carried out?

    (e) Did the student describe clearly how the data were analysed?

    (f) Did the student explain clearly the conclusions that could be drawn from the results?

    (g) Could the student deal adequately with questions at the end of the talk?

References

[1] Till Tantau. User's Guide to The Beamer Class, Version 3.06.
   http://www.ctan.org/tex-archive/help/Catalogue/entries/beamer.html.

Error Analysis for Practical Physics

1 Introduction

The estimation and calculation of errors plays an important part in the analysis of experimental data and it is important that your treatment of errors should be clear, both in laboratory note books and in full reports. During the first year laboratory and experimental methods courses you should have covered in detail the necessary theory and practice to enable you to do this. The purpose of this note is briefly to revise the important relationships and considerations.

2 Estimating Errors

The estimation of experimental errors, or even just identifying the sources of experimental error, are common sources of difficulty. The most likely errors may be classified as follows.

2.1 Systematic errors

These errors are usually concerned with the reliability of the calibration of a measuring instrument. A systematic error will **consistently** give a measurement that is wrong in a well-defined way. For example,

1. When a metal scale is used to measure distances at temperatures higher or lower than the temperature at which it was calibrated, the expansion or contraction of the metal would cause the measured length to be either shorter or longer than the true value.

2. An ideal ammeter should have zero resistance, so that there is no potential difference across its terminals. When a real ammeter, with non-zero series resistance, is introduced into a circuit it will increase the effective resistance of the circuit, and may affect the value of the current being measured.

3. An ideal voltmeter should have infinite resistance, so that no current flows through it. When a real voltmeter, with finite resistance, is placed in parallel across an electrical component some current will flow through the voltmeter, affecting the potential difference across the component under test.

2.2 Scale-reading errors

Scale-reading error refers to the uncertainty associated with reading a measuring scale. The 'error' can frequently be taken as half of the smallest marked calibrated division on the instrument being used. For example, ± 0.5 mm with a meter rule, ± 0.01 mm with some types of vernier callipers or micrometer. With some forms of electrical multimeter the uncertainty may vary with the measurement range being used, in which case the appropriate error needs to be recorded for each setting that is used.

2.3 Random statistical errors

Systematic and scale-reading errors are instrumental errors and should normally be prevented by careful design of the apparatus used in the experiment. *If instrumental errors are the dominant sources of uncertainty, then the apparatus is not well-suited to the measurements being made*. A more important, and unavoidable, source of error is random variations (sometimes known as 'noise') in the parameter to be measured. Normally these variations should be larger than the instrumental uncertainty, and they are usually characterised by a statistical distribution of values. Very often the quantity of interest is given by the mean of this distribution, and the spread of the random variations is represented by the standard deviation from the mean. In some cases where the standard

deviation for a measurement is less than the instrumental uncertainty (for example, when repeated measurements give the same value), the instrumental uncertainty can be taken as an estimate of the error in the measurement. This would also apply when only a single measurement is made, since then there can be no statistical estimate of the random variations.

2.4 Standard Deviation and Standard Error in the mean value

When a large number of measurements are made of the same experimental quantity x, the variation in measured values is often described by a Gaussian or normal distribution. The **standard deviation** σ is estimated from a set of n measurements using the formula

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

where $\bar{x}$ denotes the mean value of the distribution. For obvious reasons, the standard deviation is also known as the **root-mean-square deviation** of experimental values from the mean value. Statistically, 68% of measurements should lie within the range $\bar{x} \pm \sigma$ while 95% of measurements should fall within $\bar{x} \pm 2\sigma$. Most modern calculators will calculate the standard deviation from a set of measurements, but it should be remembered that around ten measurements are required for the value of σ to be meaningful. Indeed, with such small sample sets, a better estimate of σ is usually obtained with a divisor of $(n-1)$ instead of n. Clearly the value of $\bar{x}$ should be a better estimate of the 'true' value than any of the individual measurements. The **standard error** is a measure of the uncertainty of using $\bar{x}$ as an estimate of the 'true' value when only n repeated measurements are made, and can be found from the expression $\frac{\sigma}{\sqrt{n}}$. In a sense, the standard error estimates the standard deviation in a set of repeated measurements of $\bar{x}$ rather than repeated measurements of x.

In practice, the standard deviation should be calculated in experiments where a single value is required, and the same measurement is repeated many times. For example, the radius of a wire in a Youngs Modulus experiment, or the measurement of the period of a pendulum using a stop watch.

2.5 The Poisson distribution

In some experiments, where the measured values take discrete integer values (for example in experiments that count the number of radioactive decays that occur randomly in a given time), the distribution of values is better described by a Poisson distribution rather than by a Gaussian distribution. In this case, for a repeated set of measurements that give a mean value of N counts, the standard deviation is given by $\sqrt{N}$.

Note that in experiments where randomly occurring events are being counted, and where only one measurement of N counts is made, it is better to assume that Poisson statistics apply and assign a measurement uncertainty of $\pm\sqrt{N}$ to this measured value. It would be nonsensical to assume that a 'scale-reading' or instrumental error of ± 1 count was appropriate in these circumstances.

In practice, when N is very large, the Poisson distribution and the Gaussian distribution are very similar to one another, but in counting experiments that involve only modest values of N, there are significant differences between the two distributions.

3 The Combination of Experimental Errors

Usually in an experiment two or more measured quantities are used to calculate a final result. As all measured quantities will have an experimental uncertainty associated with them, it is essential to know the errors in the measured values propagate through to produce an uncertainty in the final result.

3.1 The general case

Consider the case where there are 2 measured values x and y, and the final result U is some general function of x, y, so that $U = f(x,y)$.

The uncertainty in U is given by

$$(\Delta U)^2 = \left[\frac{\partial f}{\partial x} \Delta x\right]^2 + \left[\frac{\partial f}{\partial y} \Delta y\right]^2 \tag{2}$$

The extension to functions of more than 2 variables is straightforward. **The combination of errors should be considered a <u>routine but essential</u> activity in the analysis of any experiment in the second year physics laboratory**. As all the main ideas were introduced in the first year, only a summary of a few particular cases derived from equation (2) will be given here. However, it is a useful and instructive exercise to confirm that you can derive these results for yourself. In all the cases below, the parameters a and b are assumed to be constants, with no associated error.

1. Linear combination of x and y

$$U = ax + by \quad \Rightarrow \quad \Delta U = \sqrt{a^2 (\Delta x)^2 + b^2 (\Delta y)^2} \tag{3}$$

2. Product or quotient of x and y

$$U = axy \quad \text{or} \quad U = \frac{ax}{y} \quad \Rightarrow \quad \frac{\Delta U}{U} = \sqrt{\left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2} \tag{4}$$

3. Raising x to a constant power

$$U = ax^b \quad \Rightarrow \quad \frac{\Delta U}{U} = b\frac{\Delta x}{x} \tag{5}$$

4. Other common functions

$$U = \ln x \quad \Rightarrow \quad \Delta U = \frac{\Delta x}{x} \tag{6}$$

and

$$U = \sin(ax) \quad \Rightarrow \quad \Delta U = a\cos(ax)\Delta x \tag{7}$$

3.2   Assessing the relative importance of different errors

In all cases where more than one measured quantity is involved, the calculation of the error in the result will involve an expression similar to equation (2). From this it is clear that the different error terms are combined in *quadrature* (that is, as a sum of squares, similar to the calculation of the hypotenuse of a right-angled triangle). To determine which error is most important, simply compare the relative sizes of the various error terms, and identify which terms are the largest. As an example, consider the case $U = axy$ in equation (4), and assume that the fractional errors in x and y are 8% and 3% respectively. In other words, $\frac{\Delta x}{x} = 0.08$ and $\frac{\Delta y}{y} = 0.03$.

It is immediately obvious that the uncertainty in x is larger, and when the combination in quadrature is performed it is clear that this error completely dominates the final result, since the fractional error in U is $\frac{\Delta U}{U} = \sqrt{(0.08)^2 + (0.03)^2} = 0.085$ or 8.5%.

The consequence of this for anyone wishing to improve the experiment is that there is almost no point in reducing the uncertainty in y until a significant reduction is made in the error associated with a measurement of x.

4   Graphs

4.1   Error bars

Any experimental data that are displayed on a graph should have appropriate error bars associated with every point plotted. Often it is only the dependent variable (the y coordinate) that requires an error bar, although in some case both coordinates of the plotted points will have error bars associated with them.

In general, the error bars could be of different length for every plotted point. However, in the second year laboratory class, it is not usually necessary, and sometimes not possible, to carry out a time-consuming statistical analysis of the errors for each point on the graph. Indeed it is common to find that the length of the error bar is almost the same for the whole range of data. It is often adequate to calculate the error bars for a representative set of data, and to assume that the other error bars are similar. As noted in section 2.3, when no repeated measurements are available to allow the standard deviations for each measurement to be calculated, the best estimate of the instrumental error should be used to determine the size of the error bars.

In some cases, where a large number of closely-spaced data-points are plotted, it is also unnecessarily confusing to plot error bars around every point, and it is better to plot bars around only a representative set of points.

4.2 'Least squares' fits to data

The error bars are intended to indicate the degree of uncertainty associated with the measured values on the graph, and to provide a statistical significance to these points when trying to find the line of best fit through the points. Large error bars indicate that the line of best fit need not pass close to the plotted points, whereas small error bars indicate that the line of best fit would be expected to pass close to these points.

Assume that the measured data consists of a set of (x_i, y_i) coordinates, where x is the independent variable, y is the dependent variable, and σ_i is the uncertainty on the value of y_i. When the measurements are expected to lie on a straight line, the variables are related by an equation of the form $y_i = mx_i + c$. The slope and intercept of the line of best fit[1] can be found by minimising the function

$$\chi^2(m,c) = \sum_{i=1}^{n} \left(\frac{y_i - mx_i - c}{\sigma_i} \right)^2$$

with respect to the parameters m and c.

It is also possible to get error estimates for these parameters at the same time. *Note that it is essential to over-plot the line of best fit on top of the experimental data so that you can confirm whether the line passes near to most of the measured values.* This reduces the chance of a simple computational error passing undetected, and indicates whether the calculated line position is being unduly influenced by one or two outlying points on the graph.

4.3 Unweighted linear 'least squares' fits

When all the error bars are of the same size, σ_i is a constant, and an unweighted least-squares fit can be carried out to find the line of best fit through the data.

In this case the theoretical expressions for the gradient m and intercept c for the linear least squares fit through a set of n data points are

$$m = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} \tag{8}$$

$$c = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n\sum x_i^2 - (\sum x_i)^2} \tag{9}$$

Equations (8) and (9) can be evaluated using some calculators, but it is easier to use the functions built into `Excel` to calculate the values of m and c. Once a set of data has been entered into the spreadsheet in the form of columns of values for the x and y coordinates, the `Excel` function **=SLOPE(y,x)** gives a value for the gradient m while the function **=INTERCEPT(y,x)** gives a value for intercept c.

The standard deviations in the slope and intercept can be found from the expressions

$$\sigma_m = \sqrt{\frac{n\sum D_i^2}{(n-2)\left(n\sum x_i^2 - (\sum x_i)^2\right)}} \tag{10}$$

$$\sigma_c = \sigma_m \sqrt{\frac{\sum x_i^2}{n}} \tag{11}$$

where $D_i = y_i - mx_i - c$ represents the difference between the measured value y_i and the line of best fit.

`Excel` can again be used to evaluate equations (10) and (11), although the calculation is a little more involved than for the slope and intercept, as there are no built-in functions that give σ_m and σ_c directly. First create another column representing the differences D_i. Then use the expressions

Sigma_m=SQRT(n*SUMSQ(d)/((n−2)*(n*SUMSQ(x) − SUM(x)^2)))
Sigma_c=Sigma_m*SQRT(SUMSQ(x)/n)

---

[1]The term *best fit* as used here implies that it is best in the 'least squares sense'. In other words, the line is chosen so that its slope and intercept minimise the squared distance of all data points from the line. In some contexts, a different criterion is used to define the best fit (for example, the least absolute distance of all points from the line could be minimised instead), but the least-squares criterion is the one most commonly adopted.

One note of caution should be introduced: if you plan to write your own computer program to evaluate equations (8), (9), (10) and (11) then do not use them in the exact form displayed here, as the calculations will be prone to round-off error. The formulas can be written in a slightly different form that avoids this problem (see reference [1] for further details).

4.4   Weighted linear 'least squares' fits

When all the error bars are not the same size, the values of σ_i must be included in the calculations and a weighted least-squares fit must be carried out to find the line of best fit through the data.

In this case the theoretical expressions for the gradient m and intercept c for the linear least squares fit through a set of n data points are more complicated than for the unweighted fit, and will not be given explicitly here. They can be found in references [1] and [2]. The latter is available for consultation in the Lab class.

A program to carry out weighted linear least-squares fitting is also available on the Physics PC network. The program file is `wlsquare.exe`, and there should be an icon entitled Least Squares on the standard student desktop.

References

[1] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C*, chapter 15.2. Cambridge University Press, Cambridge, 2nd edition, 1992.
`http://www.library.cornell.edu/nr/bookcpdf.html`.

[2] L. Kirkup. *Experimental methods. An introduction to the analysis and presentation of data.* John Wiley & Sons, 1994.

Safety Notes

List of emergency telephone numbers

In case of FIRE, or if MEDICAL AID is required, dial 2222

Medical Centre		**2613**
College Safety Officer	**Mr Tony Scott**	**1505**
Physics Department Safety Officer	**Mr Ray Moslin**	**2045**
Physics Laser Safety Officer	**Mr David Parker**	**2578**
Physics Radiation Protection Officer	**Mr Ray Moslin**	**2045**
Departmental Superintendent	**Mr Ray Moslin**	**2045**
Technical Supervisor, Teaching	**Mr David Parker**	**2578**

Physics Department First Aider

Ms Sophia Dimoudi	4284, 2320
Ms Slawka Pfauntsch	2883

OTHER USEFUL NUMBERS

Moorfields Eye Hospital	**(020) 7253 3411**
Guy's Hospital (Poison Unit)	**(020) 7635 9191/7771 5350**
Middlesex Hospital	**(020) 7636 8333**

General Considerations

A safe laboratory requires the co-operation of ALL personnel. It is the responsibility of everyone in the department to ensure that they carry out their work and use the apparatus in a manner which avoids putting themselves, their colleagues and the laboratory at risk. Students are allowed to work in the laboratory only when a member of staff, technical or academic, or a demonstrator is present.

Eating, drinking or smoking in laboratories is prohibited.

The College Safety Committee issues the main directives on safety, but Senior Staff and Supervisors should ensure that all members in the groups are made aware of any special hazards and draw their attention to the safety code and other regulations of the laboratory.

Emergency Procedures

All personnel should familiarise themselves with:

1. Procedures to be followed in the event of a fire, explosion, radioactive spillage, injury or electric shock.

2. Fire marshals and first aid personnel.

3. The nearest fire exit.

4. The location of fire-fighting equipment.

5. The location of the nearest fire alarm point.

6. The medical centre, first aid boxes and eye wash facilities.

FIRE is the greatest single hazard.

Make sure that you know the nearest escape route and also alternative routes from those parts of the College where you work.

If you discover a fire

1. Attack the fire if possible with the appliance provided but without taking personal risks.

2. Break the glass in the nearest fire alarm call point.

3. During working hours, inform the switchboard — EMERGENCY NUMBER 2222.

4. After working hours inform the Security Officer in the Porter's Lodge — dial 2222.

Fire alarm and evacuation procedure

On hearing the alarm

- Turn off gas supplies.

- Switch off electric power in use.

- Close all windows and doors.

- Vacate the building without delay — DO NOT await further instructions.

- Follow the instructions of the fire marshals.

- Proceed to an assembly area.

Use the nearest available exit and

- **DO NOT** use the lifts.

- **DO NOT** stop to collect personal belongings.

- **DO NOT** re-enter the building or evacuated areas.

- **DO NOT** run.

**Note: The fire marshal for the area will take charge and will ensure that the area is evacuated.
Carbon dioxide extinguishers may be used safely on most types of fire.
Water or foam extinguishers should NOT be used on electrical equipment fires.**

COSSH Regulations

As part of the regulations governing the Control of Substances Hazardous to Health the risks associated with the various procedures carried out in all research laboratories have been assessed. A risk assessment proforma describing the procedures and the research activities that have been addressed should be displayed in all research laboratories. Anyone who is new to the laboratory should read the risk assessment before starting work. If new procedures or research is to be carried out a COSSH risk assessment covering the work must be submitted to the Departmental Safety Officer.

Electrical Equipment

Electrical Appliance Regulations require mains operated equipment to be protected with a safety fuse from damage caused by overheating. It is therefore essential that the rating of the fuse is of the correct value. Failure to meet this requirement could result in serious damage to valuable equipment, or in fire. Do not uprate a fuse to keep equipment operating unless advised to do so by qualified staff.

If a blown fuse is not clearly marked and the equipment gives no indication of the current required, seek advice from qualified staff.

Any mains-operated equipment housed in a metal case must have an earth connection as protection for the operator against a serious (possibly fatal) shock. It is essential to ensure that the earth connection on the apparatus is in continuity with the mains earth pin. The practice of using mains adapters in tandem is against regulations. The weight of combined adapters and plugs is sufficient to cause the earth pin of the adapter to become disconnected from the mains earth line and in that event all personnel and the equipment operating from that socket would be at risk

All mains-operated equipment should now be tested for electrical safety on a regular basis. This will be indicated by a label on the apparatus stating when it was last tested and when it will need to be tested again. Ensure that all apparatus you are using carries a current test label.

Fuses

Mains Cartridge Type Fuses — Capacity Guide.

2 amp	small low wattage equipment, lamps, soldering irons etc.
3 amp	small electrical heaters, high power lamps.
5 amp	vacuum pumps, compressors, diffusion pumps.
10 amp	machines, larger capacity stills.
13 amp	ovens, furnaces, other specialised equipment.
Slow blow type	these handle surge currents in certain types of equipment.

Before fitting a mains supply cable to equipment ensure that the current carrying capacity is adequate. The cable should not be excessively long so that it trails across the floor, nor so short that the cable is put under stress. Choose the appropriate fuse rating following the manufacturers specification. It is essential that the earth wire is in good electrical contact with the earth pin of the plug, and not clamped by insulation. Ensure that the cable is held firmly by the clamp provided. The colour code of standard electrical cable is

Brown	Live
Blue	Neutral
Green/yellow	Earth

ONLY special waterproof plugs (metal covered) should be used in cold rooms or elsewhere where condensation might be expected. Ordinary domestic plastic covered plugs are inadequate and dangerous.

The plugs and sockets of all equipment which is left on overnight MUST be clearly marked so that quick action can be taken in an emergency.

When servicing electrical equipment you are responsible for unplugging it from the mains supply.

Do not depend on someone else

Experiments left running overnight must have the approval of the supervisor. Shut down or emergency procedures must be clearly displayed and the night porter's office should be notified. All water shut off valves and relevant mains switches should be clearly labelled.

Mechanical Equipment

The following precautions should be observed when working with machinery or within the vicinity of machinery.

- Permission must be obtained from Mr Kevin Smith (Workshop Supervisor) to use machinery in the workshop area.

- All machinery is potentially dangerous. The result of catching fingers, loose hair or clothing in moving parts, even of low power equipment can be serious. Such equipment should be guarded wherever possible and must *NEVER* be left running unattended. Workshop staff must where possible use protective guards, and any necessary protective clothing. Eye shields or glasses must be worn when required by safety regulations; they are *NOT* an alternative to guards.

- Machine operators should ensure that their actions do not put other personnel at risk.

- Any parts of machines which necessarily protrude into a walkway should be made clearly visible with a coloured marker.

- Students and any other non-workshop staff may only use workshop machines after appropriate training and instruction, and guards must always be used.

- Machinery used for laboratory services, for example, a lift truck, cylinder carrier, or chain hoist are all potentially dangerous if not used with care. Operators of this type of equipment should observe any safety instructions on the equipment and must not take risks by handling loads which may be outside the capacity of the equipment.

- Personnel required to move heavy loads by hand, often in an awkward or restricted space, can cause serious injury to themselves if they adopt the wrong posture for lifting. The correct way to lift a heavy object is to bend the knees while keeping the back erect, and to return to a standing position without bending the body forward.

Chemical hazards

You should regard every new chemical with which you have to deal as a potential hazard until you have satisfied yourself to the contrary. Information regarding the toxic hazards of most chemicals is available in the College Safety Office or in the Physics Stores (Q29).

You should make yourself familiar with the code of symbols found on many reagent containers, and the colour coding for gas cylinders. Charts are to be found in Q29.

In general highly toxic or otherwise dangerous chemicals have warning notices on the labels. Read the labels carefully before opening the container. Special precautions may be needed in handling the material. In case of doubt consult the chemical hazard literature kept in Room Q29.

The ordering of **ALL** chemicals must be sanctioned by the Departmental Safety Officer; students must **NOT** place orders on their own initiative.

The College Safety Officer must be informed when the use of any specially hazardous material is contemplated.

A fume cupboard with the hood lowered as far as possible should be used if dangerous substances are to be used in gas liquid or powder form.

Do not pipette toxic liquids by mouth.

Any operations with caustic or corrosive liquids should not be undertaken unless you are wearing safety glasses or a visor and preferably gloves and a protective apron.

All chemical operations, for example, distillation or vacuum transfer require safety glasses to be worn, and everyone entering a laboratory in which potentially hazardous operations are in progress must also wear safety glasses. If you do not know how to set up an experiment involving an operation such as distillation take advice before you attempt it.

No potentially hazardous experiments should be done in the evening, or at weekends, when there may be no-one on hand to help if an accident occurs.

Be careful to clean up any spillage of corrosive or toxic chemicals. Do not pour any organic solvents whether flammable or not down the sink. Consult the Departmental Safety Officer about the disposal of unwanted chemicals.

Ensure that Chemicals are sensibly stored. Keep heavy bottles below eye height and always transport them in the baskets provided. Keep bottles of incompatible chemicals (strong acids and organic solvents; oxidizing and reducing agents; oxidizing and flammable materials) separate from each other. Always ensure that all bottles are properly labelled.

If you are handling liquid nitrogen in large quantity, wear safety glasses and leather gloves. Make sure that liquid nitrogen is never confined in a stoppered container. When removing vials from liquid nitrogen wrap them in a cloth to allow trapped liquid to boil off before opening.

The key to your safety and that of other people in the laboratory is never to lose concentration and to think before you act.

Ionizing Radiation

College Rules

Rules governing all work with radioactive substances and with X-rays are issued by the College, and must be obeyed. Every worker using ionizing radiation should possess a copy of these rules. Further copies may be obtained from the Assistant Safety Officer.

Radioactive Substances

The use of radioactive substances in the laboratory is controlled by the Radioactive Substances Act 1993. The purchase, use and waste disposal of radioactive substances in the laboratory is controlled by the Safety Officer of the College with local supervision by Departmental Radiation Protection Supervisors, and no radioactive substance may be brought into the laboratory without permission of the Safety Officer, or Assistant Safety Officer.

All groups using radioactive materials must keep a record of the purchase, use and disposal of radioactive substances which must be kept up to date for possible examination by the Department of Environment Inspector.

All workers using radioactive substances must attend a course on radiation protection before commencing work or work under the supervision of a trained member of staff, unless exempted by the Assistant Safety Officer on the basis of previous experience.

The details of series of experiments and or projects using radioactive substances must be submitted to the Assistant Safety Officer on the forms obtainable from him after consultation with the Departmental Radiation Supervisor.

Almost all radioactive substances used in the department are of the sealed type. In the event of a mishap **DO NOT ATTEMPT** to deal with the matter yourself. Notify the staff member responsible for the area in which you are working, and the Assistant Safety Officer.

All users of sealed radioactive sources must be given appropriate training and instruction in the use of any sealed source or equipment containing a sealed source and must be informed of the hazards involved. A working procedure satisfying the requirements of the local rules [III–7] must be followed, unless these requirements are modified in a written system of work as recommended by the Radiation Protection Advisor in special circumstances.

X-Ray Protection

X-rays are potentially the most hazardous radiation source used in these laboratories. It is then essential that the following precautions be adhered to at all times.

All research users must be registered with the College Safety Officer.

In research work using X-rays all users must be given appropriate training and instruction in the use of X-ray equipment and associated safety systems and be informed of the hazards involved. An operating procedure

satisfying the requirements of the Local Rules [II–7] must be followed, unless these requirements are modified in a written system of work as recommended by the Radiation Protection Advisor in specific circumstances.

All research users must wear personal monitors whenever they are working in X-ray rooms.

Always check apparatus for radiation leakage; if in doubt shielding must be arranged. No X-ray apparatus may be energized unless a beam stop is in position.

You must not enter areas where X-ray warning signs are displayed unless you have permission. Such permission can be granted for research work only by the Assistant Safety Officer, and in the case of teaching laboratories by the staff member in charge of the laboratories.

Laser Safety

These notes are intended to provide preliminary information about laser hazards and to be used as a guide to the safe use of existing laser systems. Reliance must be placed on minimizing the hazard at source rather than on protection but suitable safety goggles must be available for all work with lasers in Classes 3B and 4.

Before setting up a new laser system the Departmental Laser Safety Supervisor, Dr K Powell should be consulted.

- **CLASS 1 LASERS** are safe for direct viewing. (Max output: 0.98 mW)

- **CLASS 1M LASERS** are safe for direct viewing, but potentially hazardous with optical instruments like magnifying glasses (for LEDs and bare laser diodes), binoculars or telescopes (for wide, collimated beams).

Radiation in classes 1 and 1M can be visible, invisible or both.

- **CLASS 2 LASERS** emit visible radiation, safe for accidental viewing under all operating conditions. However, it may not be safe for a person who deliberately stares into the laser beam for longer than 0.25 s, by overcoming their natural aversion response to the very bright light. (Max output: 1 mW)

- **CLASS 2M LASERS** emit UV or IR radiation, safe for accidental viewing with the naked eye, as long as the natural aversion response is not overcome as with Class 2, but may be hazardous (even for accidental viewing) when viewed with the aid of optical instruments, as with class 1M.

- **CLASS 3R LASERS** are considered of low risk radiation, but potentially hazardous. Depending on wavelength, the class limit for 3R is $5\times$ the class limit for Class 1 (for invisible radiation), or $5\times$ the class limit for class 2 (for visible radiation) and can cause eye injury with long term intra-beam viewing.

- **CLASS 3B LASERS** are likely to cause eye injury even for accidental viewing. Their diffuse reflection is normally safe. (Max output: 500 mW)

- **CLASS 4 LASERS** are high power lasers, which present the likelihood of serious eye injury even for exposure to close up diffuse reflections as well as of skin injury.

CLASS 3B and CLASS 4 LASERS present real hazards which should be clearly understood before such lasers are used. The College Safety Officer should be consulted before such a laser is brought into operation.

Pulsed lasers can produce extremely high power densities and extra care must be exercised in their use.

> **Note: It is important to realise that the electrical power supply for lasers can present the greatest potential hazard and can be lethal.**
> **A class 4 or focused class 3 laser is also a fire hazard and an adequate beam terminator is essential for these more powerful lasers.**

Display Screen Equipment

A KCL code of practice describing safety standards for work with display screen equipment (*DSE*) or work stations has been produced. This is intended to apply primarily to equipment used by College staff who use DSE for college purposes for periods of more than ten hours a week. The code of practice contains a check-list which it is recommended should be used to identify significant risks with all existing units. All new installations should be designed to comply with the recommendations set out in the code of practice. Copies of the code of practice and check list may be obtained from the College Safety Office, School Office or Departmental Safety Officer.

EXPERIMENTS

The Operational Amplifier

1 Aim of the Experiment

This experiment is intended to illustrate two important characteristics of an operational amplifier: the **gain-bandwidth product**, and **slew-rate distortion**.

2 Apparatus

The operational amplifier (op-amp) type 741 used in this experiment is already wired up in a metal box to minimise noise pick-up, with the necessary resistor network to obtain gains of 1, 10, 100, 1000 and 10,000. The amplifier is powered by an external power source that supplies $-15, 0, +15$ V. A battery powered signal generator is connected to the input, and a battery powered AC voltmeter is connected to the output. A switch allows either the input to, or the output from, the op-amp to be directed to the AC voltmeter. Additional connectors are available to allow the input and output waveforms to be monitored on a dual-trace oscilloscope (which should be triggered on the input signal), and for a phase meter to determine the phase difference between the input and output signals.

3 Theoretical Background

Signal gain is defined to be the ratio of the output signal to the input signal. The ideal operational amplifier has infinite gain at all frequencies, infinite input impedance and zero output impedance. A real op-amp does not achieve this ideal, but is designed to provide a very high gain, usually with a very high input impedance, and a very low output impedance. However, the precise values of these important parameters are difficult to control during the manufacturing process, and so vary from device to device. This device variability is less important than it might seem at first sight, since by using negative feedback and a few external components (such as resistors) it is possible to build an amplifying circuit whose properties are largely independent the operational amplifier.

The op-amp normally has two input terminals, the inverting input $(-)$ and the non-inverting input $(+)$. Without feedback, an input voltage V_i applied to the input terminals is multiplied by A_{OL}, the **'open-loop' gain** of the amplifier. This is normally very large (10^5 to 10^6) and all practical applications use the amplifier under "closed-loop" conditions where a fraction β of the output signal is fed back to the inverting input. This negative feedback
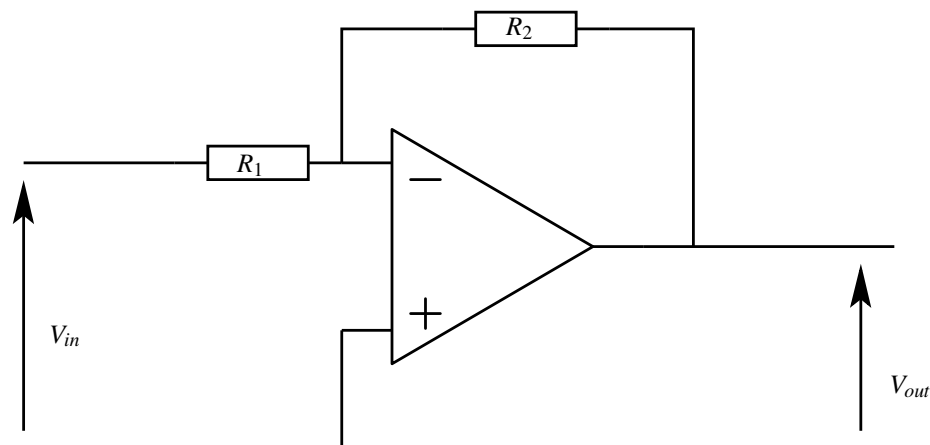

Figure 1: Diagram of an inverted amplifier

results in a **closed-loop gain**:

$$A_{CL} = \frac{A_{OL}}{1 + \beta A_{OL}} \tag{1}$$

The quantity βA_{OL} is often just called the **loop gain**, and usually $\beta A_{OL} \gg 1$

Hence

$$A_{CL} \approx \frac{1}{\beta} \tag{2}$$

and the closed loop gain is determined solely by β (which is set by the resistors in the feedback circuit).

For example, the inverting amplifier in Figure 1 has a gain given by $\frac{v_{out}}{v_{in}} = -\frac{R_2}{R_1}$, so it is set by two resistors, whose values are reliable and can be determined accurately. The books listed page 45 provide a more detailed theoretical treatment of the use of feedback in amplifier circuits.

3.1 Frequency compensation

Although at low frequencies modern operational amplifiers come close to meeting the ideal description, the gain must be limited at high frequencies otherwise the circuit becomes unstable (and would probably oscillate). All op-amps therefore employ some form of frequency compensation which may be internal to the integrated circuit (as for the type 741, used in this experiment) or external, in the form of a resistor-capacitor network.

In its simplest form the frequency compensation may be regarded as a low-pass RC filter following an ideal op-amp.

3.2 Representing the Data

The family of curves representing the frequency response of an op-amp covers a wide range of gains and frequencies, and is conventionally plotted on a log-log graph. The voltage gain is normally expressed in decibels where

$$Gain(dB) = 20\log_{10}\left[\frac{V_{out}}{V_{in}}\right] \tag{3}$$

3.3 Modelling the frequency response of an op-amp

3.3.1 Open loop frequency response

An operational amplifier with frequency compensation may be regarded as a voltage source A_0V_i with a low-pass RC filter on the output, as indicated in Figure 2, where A_O is the open-loop gain of the amplifier at zero (or very low) frequency.

At frequency $f = \frac{\omega}{2\pi}$ the output voltage V_o can be written, in a complex form that takes into account the phase shift introduced by the RC filter:

$$V_o = \frac{A_O V_i}{1 + j\omega RC} \tag{4}$$

where $j^2 = -1$

If we set $f_1 = \frac{1}{2\pi RC}$ then the complex open loop gain at frequency f can be written

$$\tilde{A}_{OL}(f) = \frac{V_o}{V_i} = \frac{A_O}{1 + j\left(\frac{f}{f_1}\right)} \tag{5}$$
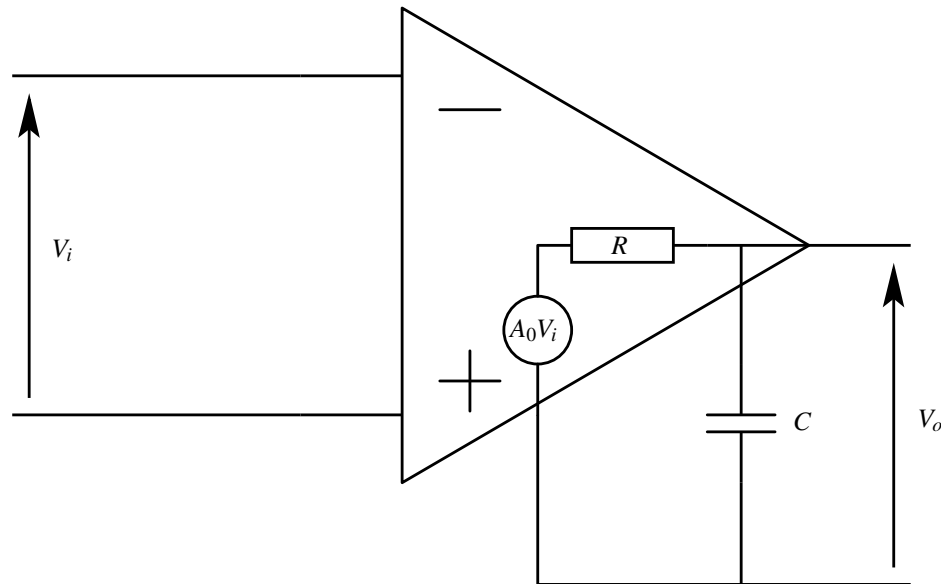
Figure 2: Diagram of a modelised operational amplifier

The magnitude and the phase shift of the open-loop gain are therefore

$$A_{OL}(f) = \frac{V_o}{V_i} = \frac{A_O}{\sqrt{1 + \left(\frac{f}{f_1}\right)^2}} \tag{6}$$

$$\theta = \arctan\left(-\frac{f}{f_1}\right) \tag{7}$$

If we express the gain in decibels

$$A_{OL}[dB] = 20\log_{10}(A_0) - 20\log_{10}\left[\sqrt{1 + \left(\frac{f}{f_1}\right)^2}\right] \tag{8}$$

then we see that f_1 is the frequency at which the response is 3 dB lower than the gain at $f = 0$ (sometimes referred to as '3 dB down'). When $f \gg f_1$, the gain in decibels is

$$A_{OL}[dB] \approx 20\log_{10}(A_0) - 20\log_{10}\left(\frac{f}{f_1}\right) \tag{9}$$

and we see that the gain drops off at about 20 dB per decade (i.e. for every factor of 10 increase in frequency, the gain falls by a factor of 10).

It is a reasonable approximation to say that A_{OL} is almost constant for frequencies less than f_1, which is known as the corner frequency, and then rolls off at 20 dB per decade, as shown in Figure 3. The bandwidth of the amplifier is therefore said to range from 0 to f_1.

At the crossing frequency, f_0, the gain is unity (=0 dB), and no amplification is possible at frequencies higher than f_0.

3.3.2 Closed loop frequency response

The closed-loop gain of the amplifier was given by equation (1). Substituting this equation into equation (5) for the open-loop gain gives

$$\tilde{A}_{CF}(f) = \frac{A_0}{1 + \beta A_0 + j\left(\frac{f}{f_1}\right)} \tag{10}$$
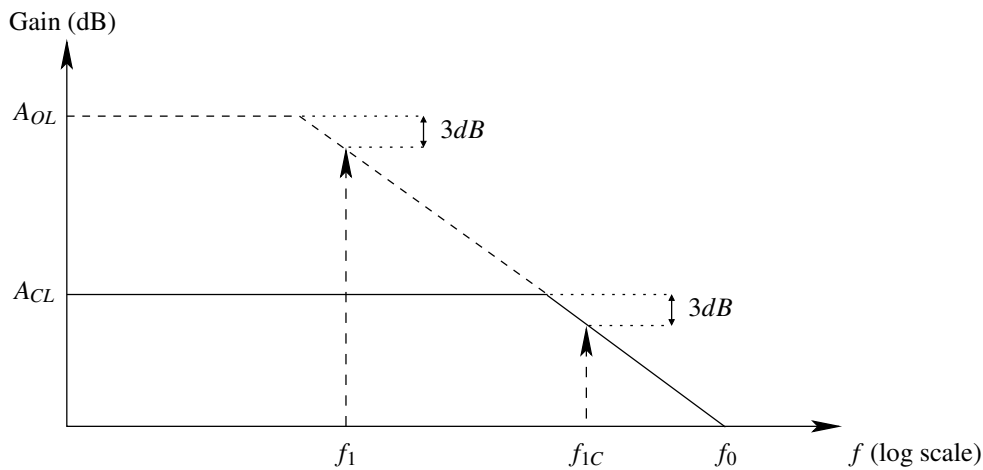
Figure 3: Frequency response of an ideal operational amplifier

The significance of equation (10) can be made clear by noting that the denominator can be written as

$$1 + \beta A_0 + j\left(\frac{f}{f_1}\right) = (1 + \beta A_0)\left[1 + j\left(\frac{f}{f_1(1 + \beta A_0)}\right)\right] \tag{11}$$

If we now define the **zero-frequency closed-loop gain**

$$A_C = \frac{A_0}{1 + \beta A_0} \tag{12}$$

and the closed-loop corner frequency

$$f_{1C} = f_1(1 + \beta A_0) \tag{13}$$

then the closed loop gain is given by the expression

$$\tilde{A}_{CL}(f) = \frac{A_C}{1 + j\left(\dfrac{f}{f_1}\right)} \tag{14}$$

which is now in exactly the same form as equation (5) for the open-loop gain. The general behaviour of $A_{CL}(f)$ will thus be similar to $A_{OL}(f)$, except that the maximum gain in the presence of closed-loop feedback is now reduced to A_C, as given by equation (12), while the bandwidth is increased to f_{1C}. At frequencies above f_{1C} the gain falls off at 20 dB per decade, just as for the open-loop gain. An example of the closed-loop gain as a function of frequency is shown in Figure 3.

Another important point to note from equations (12) and (13) is that the product of the gain and the bandwidth is given by

$$A_C f_{1C} = A_0 f_1 \tag{15}$$

From this we see **that the gain-bandwidth product is a constant for a given op-amp**.

3.4   Slew rate distortion

The op-amp contains capacitors, and there is a maximum rate at which these can be charged, depending on the characteristics of the integrated circuit. This means that there is a maximum rate of change of signal voltages that can be handled without distortion, and this maximum rate is know as the **slew-rate** limit for the amplifier.

Suppose the output signal as a function of t is $V_p \sin 2\pi ft$ where V_p is the peak voltage and f is the frequency. The maximum rate of change can be found by differentiating, to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}V_p \sin 2\pi ft = 2\pi f V_p \cos 2\pi ft \tag{16}$$

which has a maximum value of $2\pi f V_p$. This quantity must be less than the limiting slew-rate S for the op-amp, otherwise the signal will be distorted. When $2\pi f V_p \gg S$ the output signal will be a triangular wave, not a sine wave. It is clear that for large-amplitude signals slew-rate distortion will occur at lower frequencies than for small-amplitude signals.

4 Experimental Procedure

The frequency range covered in the experiment is from 20 Hz to 1 MHz. Data should be taken at a sequence of frequencies 20, 50, 100, 200, 500, ... Hz, and the data points will then be approximately equally spaced on a $\log_{10}$ (frequency) axis. There are some important points to bear in mind when making measurements:

1. The **AC voltmeter** (use Kenwood VT181) used to measure the signals **may have a limited bandwidth** over which it works reliably. On some voltmeters of the older type this bandwidth can be adjusted by a switch on the front panel, and in this case you should always ensure that the bandwidth setting of the meter includes the frequencies being used for measurements. The upper frequency of the meter's bandwidth should be set to 1 MHz or greater.

2. The **maximum output voltage range from the op-amp is set by the power supply** (± 15 V in this case). When working with large amplifier gains you should always ensure that the amplified signal at the output will lie well within this range, or the signal will be 'clipped' when the amplifier output reaches its limiting value.

3. You should **always monitor the output signal on the oscilloscope screen**, even if you measure the signals using the voltmeter, as the oscilloscope allows you to spot immediately if the output is no longer a sine wave.

4.1 Check that the op-amp is working properly

Connect the components together, but do not include the phase meter at this stage. Set the amplifier to gain range 1 and the signal generator to provide a 1 kHz sine wave at an amplitude of about 1 volt (equivalent to about 0.7 volt RMS). Check on the oscilloscope that the input and output signals have the same amplitude and are exactly in phase. If not there is something wrong, or the oscilloscope is incorrectly set, and the problem must be rectified before proceeding.

4.2 Slew-rate distortion

Before you can measure the gain of the amplifier as a function of frequency it is important to determine the slew-rate limit for the op-amp, so that all subsequent measurements are not affected by slew-rate distortion.

Set the op-amp gain to 1 and Increase the signal level to the maximum possible, to get a sine wave with a peak voltage $V_p \approx 7.5$ volts (5 V RMS), and then increase the frequency from about 100 Hz upwards. If you display the input and output waveforms simultaneously on the oscilloscope, you will find that at some frequency the output starts to become distorted compared to the input, and as the frequency is increased still further the output waveform becomes more like a triangular wave. At the same time, the amplitude of the output signal will reduce.

1. Measure the frequency at which the output first becomes distorted and note the amplitude of the output waveform. Calculate the slew rate limit $S = 2\pi f V_p$ from these data. The normal units for slew rate measurements are $\text{V·}\mu s^{-1}$.

2. Adjust the frequency to a higher value, so that the output waveform is triangular, and measure the slope of the triangular wave directly from the oscilloscope. This will give another estimate of the slew rate S that the op-amp is capable of achieving.

For the rest of the experiment, you must ensure that all measurements that you make **satisfy the requirement** $V_p f \ll \dfrac{S}{2\pi}$.

4.3 Frequency response as a function of gain

With the op-amp gain set at 1, reduce the input signal to 30 mV RMS and measure the gain of the circuit as a function of frequency between 20 Hz and 1 MHz. You should find that the gain is approximately unity (or 0 dB) right up to about 500 kHz. Plot the gain (in decibels) against $\log_{10}$(frequency). Suitable graph paper may be obtained from a lab technician if required.

Repeat these measurements for the other four gain settings of the op-amp. The input voltages must be carefully chosen so that at all times the amplitude of the output waveform is not clipped by the range of the power supply, and that the slew-rate limit is not exceeded. In general this means that the input signal will need to be reduced as the op-amp gain is increased, and as the frequency is increased. You will find that the measured gain falls at high frequencies, and in these circumstances it is possible to increase the input signal level once more to make it easier to measure very small output signals.

Note that it is highly desirable to plot the data as the experiment proceeds.

4.4   Phase shift

You should have noticed from the oscilloscope traces that as you approach and exceed the corner frequency ('*-3 dB point*') the input and output signals are no longer in phase, but that the output lags in phase behind the input.

This should be measured in detail on the gain 100 range using the phasemeter. Use an input signal of 10 mV RMS and measure the phase angle between the input and output signals from 20 Hz to 200 kHz.

When you have finished the experimental work, remember to switch off the battery-powered equipment.

5   Data Analysis

5.1   Slew rate

Compare your measured values with each other, and with a typical manufacturer's specification for a type 741 op-amp. Comment on the size of the experimental errors.

5.2   Frequency response

You should have a family of curves, similar to the two curves shown in Figure 3. Each curve should show the gain to be almost constant up to some corner frequency, and then to fall off at a rate of 20 dB per decade at frequencies above the corner frequency. All the gain curves should lie within the same 'envelope' set by the open-loop gain of the amplifier.

From the gain curves, determine the corner frequencies, and show that the gain bandwidth product is approximately the same for each curve. *Use the true value of the gain, **not the value in decibels**, for this calculation.* Compare your measured values with the manufacturer's specification for the op-amp.

5.3   Phase shift

Confirm that when the gain is "flat" (almost constant) the input and output signals are approximately in phase, and that when the gain has dropped by 3 dB the phase lag between the output and input is approximately $45°$.

5.4   The distributed amplifier: a thought experiment

Consider trying to amplify a signal by a factor of 10^6.

Using the information and insight provided by your experimental data, comment on the differences in performance that you would expect between a single 741 op-amp wired up to provide a gain of 10^6, and a set of three 741 op-amps, each with a gain of 100, connected in series so as to provide an overall gain of 10^6.

References

[1] M. Kahn. *The Versatile op-amp*. Holt, Rinehart and Winston, 1970.

[2] R.F. Coughlin and F.F Driscoll. *Operational amplifiers and Linear Integrated Circuits*. Prentice Hall, 4th edition, 1991.

[3] H. Ahmed and P.J. Spreadbury. *Analogue and Digital Electronics for Engineers: an Introduction*. Cambridge University Press, 1984.

Series LCR circuit

1   Aim of the Experiment

The aim of this experiment is to study the behaviour of damped harmonic oscillations in a resonant circuit consisting of an inductor L, a capacitor C and a resistor R connected in series. From measurements of the resonant behaviour it is possible to deduce values for the resistance, capacitance and inductance of the circuit elements.

2   Apparatus
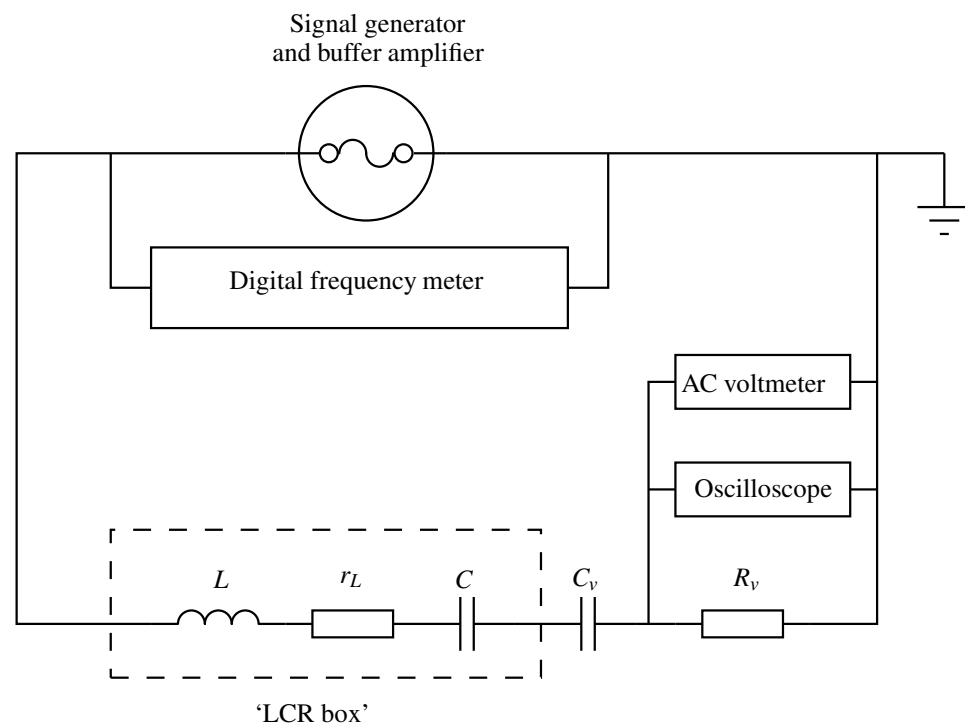
Figure 1: Schematic of the experimental set-up

The circuit used for the first part of the experiment is shown in Figure 1. An unknown inductance L, (which will have its own internal resistance r_L) is connected in series with an unknown fixed capacitance C. A variable capacitor C_v and a variable resistor R_v are also connected in series and the circuit is driven by a signal generator. The generator is connected to a buffer amplifier with a very low output impedance ($\ll 1\ \Omega$) so that it acts as an almost ideal voltage source, and the frequency of the generator is measured with a digital frequency meter.

3   Theoretical Background

The model of an harmonic oscillator plays an important role in many physical theories and concepts, and therefore also the physics curriculum. You have already been introduced to it in Dr Alexandre's 1$^{\text{st}}$ year course

CP1400 *"Classical Mechanics and Special Relativity"* [1], and you will encounter it again in Prof Davies 2nd year course CP2201 *"Introductory Quantum Mechanics"* [2].

Here, in its incarnation in a resonant circuit, the combination of an inductor L, a capacitor C and a resistor R connected in series with an alternating voltage source has a direct analogy with a mechanical system undergoing forced or damped harmonic vibrations. The motion of a particle of mass m attracted to the origin by a force kx and subjected to a resistive force $b\dfrac{\mathrm{d}^2 x}{\mathrm{d}t^2}$ is governed by the differential equation

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + b\frac{\mathrm{d}x}{\mathrm{d}t} + kx = 0 \tag{1}$$

The corresponding equation for a series LCR circuit driven by an alternating signal $V_0 \sin \omega t$ is

$$L\frac{\mathrm{d}^2 q}{\mathrm{d}t^2} + R\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{q}{C} = V_0 \sin \omega t \tag{2}$$

where q is the charge on the capacitor C at time t.

By comparison of equations (1) and (2) it can be seen that the inductance is the electrical analogue of mass, resistance is the analogue of a velocity-dependent friction, and the reciprocal of the capacitance is analogous to the spring constant k.

Solution of equation (1) gives the transient response of a damped, undriven oscillator, with 3 different types of damping (strong, weak and critical). Solution of equation (2) gives the steady-state response of a damped driven oscillator. This was shown in Dr Alexandre's 1st year course CP1400 *"Classical Mechanics and Special Relativity"* [1].

3.1   The resonant frequency

It can be shown that the current in the series LCR circuit reaches a maximum value at a frequency

$$f_0^2 = \frac{1}{4\pi^2 LC} \tag{3}$$

where f_0 is the resonant frequency of the circuit.

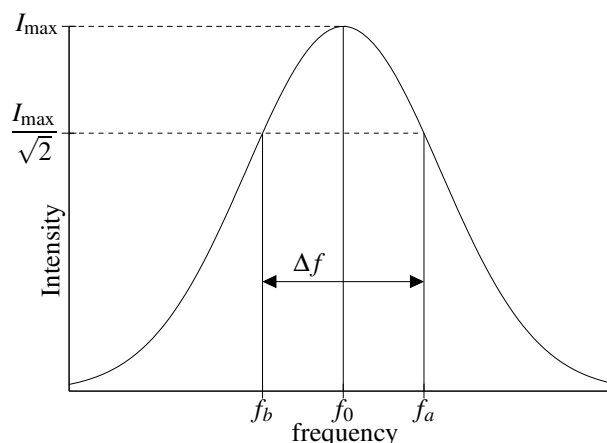## 3.2   The quality factor (Q) of the resonant circuit

Figure 2: Definition of the Q factor

The quality factor (or Q value) of a resonant series LCR circuit is a measure of how sharply-defined the current maximum is around the resonant frequency f_0.

If the current has a value $I_{\max}$ at f_0, and it falls to a value $\dfrac{I_{\max}}{\sqrt{2}}$ at frequencies f_a and f_b (respectively above and below the resonant frequency f_0), as shown in Figure 2, then the Q value may be shown to be

$$Q = \frac{f_0}{f_a - f_b} = \frac{f_0}{\Delta f} = \frac{2\pi f_0 L}{R} \tag{4}$$

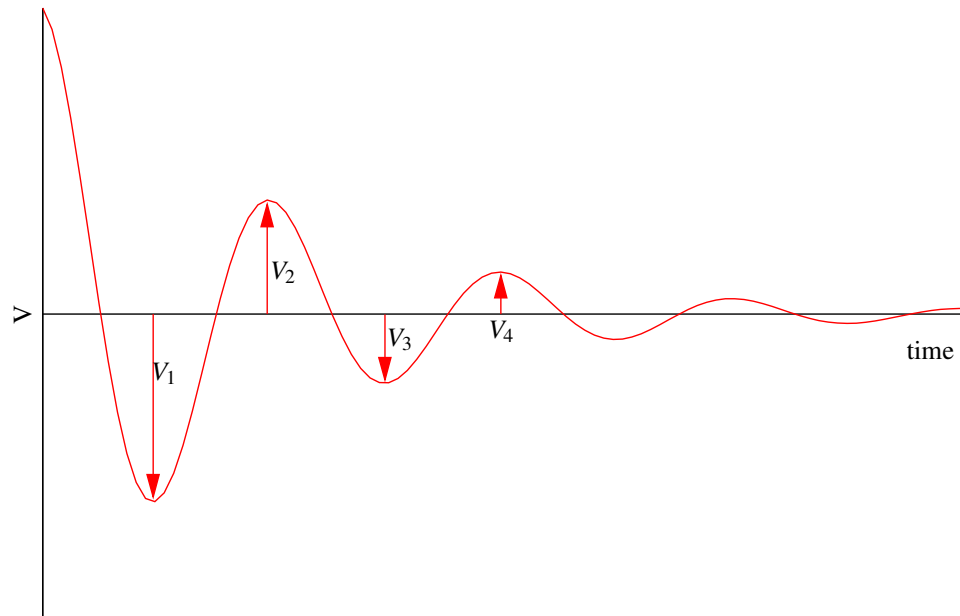## 3.3 Exponential decay of damped oscillations

Figure 3: The amplitude of a damped oscillation as a function of time.

The amplitude of a damped harmonic oscillator that is not driven by an external force will decay exponentially with time as shown in Figure 3. The **logarithmic decrement** λ is given by

$$\lambda = \ln\frac{V_1}{V_2} = \ln\frac{V_2}{V_3} = \ln\frac{V_j}{V_{j+1}} = \frac{1}{m}\ln\frac{V_j}{V_{j+m}} \tag{5}$$

where V_j represents the amplitude of the jth oscillation. Provided $R^2 \ll \dfrac{4L}{C}$ the **logarithmic decrement** is equal to

$$\lambda = \frac{R}{4Lf_0} \tag{6}$$

4 Experimental Procedure

4.1 The variation of resonant frequency with capacitance

Connect up the LCR resonant circuit as shown in Figure 1. The signal generator should be adjusted to give a 1 volt peak-to-peak sine wave, as measured at the output of the buffer amplifier. The variable resistance R_v is a resistance box that should be set to a fixed value of about 100 Ω. C_v is a capacitor box that should be varied over the range from 0.01 to 0.99 μF during this part of the experiment. **It is essential that the common** (or ground) **terminal is connected to same earth point**, as shown in Figure 1.

For each setting of C_v, adjust the frequency of the signal generator until a maximum voltage is observed across R_v. This corresponds to a maximum current in the circuit, which is said to be **resonant** at this frequency f_0. A graph of f_0^2 against $\dfrac{1}{C_v}$ should give a straight line (that does not pass through the origin).

The data points should be equally spaced. Note that it is wise to plot this graph as soon as you have collected the data, so that you identify whether any more data are needed before going on to the next part of the experiment.

4.2   The variation of the quality factor with resistance

Remove the variable capacitance C_v from the circuit shown in Figure 1, and find the new resonant frequency of the circuit. Next adjust the value of the variable resistance over a range of values from 10 to 100 Ω and in each case measure the width of the resonance $\Delta f = f_a - f_b$.

Δf can be measured by first noting the potential difference across the resistance R_v at the resonant frequency f_0, and then changing the frequency supplied by the signal generator until the potential difference has fallen by a factor of $\sqrt{2}$. This will occur at the frequencies f_a and f_b, which can be measured using the digital frequency meter. A graph of Δf against R_v should again yield a straight line.

4.3   The transient response of the LCR circuit

Using the same circuit as in section 4.2, adjust R_v to a value around 10 Ω. Set the signal generator to produce a square wave with a frequency of around 70 Hz and a peak-to-peak output voltage from the buffer amplifier of 10 V. Adjust the oscilloscope settings so that you can examine closely the voltage signal across resistor R_v as the square wave input signal jumps from one value to the other. You should see a damped harmonic oscillation of the signal across R_v.

Measure the logarithmic decrement λ for the damped oscillation for a range of different values of R_v (from 10 to 100 Ω), and plot a graph of λ against R_v.

Make sure that the oscilloscope timebase is set to the *Calibrated* position, and measure the period τ of the decaying oscillations.

Also, measure r_L directly using a multimeter and compare the result to the two previous values.

5   Data Analysis

The three sets of data acquired in section 4 allow estimates of the inductance, the resistance of the inductor windings, and the capacitance to be obtained. Consider whether the various estimates for the values of L, C and r_L agree within experimental error, and identify which method of measurement leads to the most accurate values.

5.1   The variation of resonant frequency with capacitance

If C_t is the total capacitance of the circuit, then $\frac{1}{C_t} = \frac{1}{C} + \frac{1}{C_v}$, and on making use of equation (3) it is clear that the resonant frequencies measured in section 4 are

$$f_0^2 = \frac{1}{4\pi^2 L C_t} = \frac{1}{4\pi^2 L C} + \frac{1}{4\pi^2 L C_v} \tag{7}$$

A graph of f_0^2 against $\frac{1}{C_v}$ should thus give a straight line with a slope of $\frac{1}{4\pi^2 L}$ and an intercept of $\frac{1}{4\pi^2 L C}$. A least squares fit to these data will allow values of both L and C to be determined, as well as the associated experimental error.

5.2   The variation of the quality factor with resistance

If R is the total resistance of the circuit, then $R = r_L + R_v$, and on making use of equation (4) it is clear that the values of Δf measured in section 4 are given by

$$\Delta f = \frac{R}{2\pi L} = \frac{r_L}{2\pi L} + \frac{R_v}{2\pi L} \tag{8}$$

A graph of Δf against R_v should thus give a straight line with a slope of $\frac{1}{2\pi L}$ and an intercept of $\frac{r_L}{2\pi L}$. A straight line least squares fit to these data will allow values of both L and r_L to be determined, as well as the associated experimental error.

5.3   The transient response of the LCR circuit

If R is again the total resistance of the circuit, then equation (6) shows that the logarithmic decrements λ measured in section 4 are given by

$$\lambda = \frac{R}{4 f_0 L} = \frac{R_v}{4 f_0 L} + \frac{r_L}{4 f_0 L} \tag{9}$$

A graph of λ against R_v should thus give a straight line with a slope of $\frac{1}{4 f_0 L}$ and an intercept of $\frac{r_L}{4 f_0 L}$. A straight line least squares fit to these data will allow values of both $f_0 L$ and r_L to be determined. The period of the decaying oscillations determines the resonant frequency, since $f_0 = \frac{1}{\tau}$, so the value of L can be found, and using equation (3) **get another value of the capacitance** C.

The values of the LCR circuit used are of the following order of magnitude: R$\simeq$10-100 Ω, L$\simeq$10-100 mH, C$\simeq$50-200 nF.

Draw a table comparing the values you obtained using the three different methods.

References

[1] Alison Mainwood and Jean Alexandre. Mathematics and mechanics for physics i.
    http://kcl.ac.uk/schools/pse/physics/courses/cm1350details.html.

[2] Gordon Davies. Introductory quantum mechanics.
    http://kcl.ac.uk/schools/pse/physics/courses/cp2201details.html.

[3] Bleaney B. L. and B. Bleaney. *Electricity and Magnetism*, volume 1. Oxford University Press, Oxford, 3rd edition, 1989.

Digital Electronics

1 Introduction

This experiment consists of two parts. The first part studies logic gates and digital circuit design [1]. It is recommended that chapters 1–3 of Digital Logic Design by B. Holdsworth [2], available in the laboratory, is used as background reading. A more recent edition of this book [3] can be found in the library. The second part of this experiment concerns machine code programming of a simple microprocessor and requires no background reading, although the experiment is easier to perform if the concept of machine code programming is well understood. A short but concise introduction can be found in Wikipedia [4].

2 Digital logic design

2.1 Aim of the experiment

In this experiment you are required to first test the behaviour of the Transistor Transistor Logic (*TTL*) Integrated Circuits (*IC's*), and then construct various circuits using the IC's and test their logical behaviour. You should draw a circuit diagram of each circuit you construct, and use truth tables [2] to describe their behaviour. For the construction of circuits you are mainly asked to use NAND gates, as these are the most commonly used circuits commercially—why? (Answer in your lab book.)

2.2 Apparatus

The experiments are based around the use of the Limrose Electronics IC patchboards. These boards allow the use of standard TTL IC's for building and testing combined logic functions. Each patchboard has three rows of four sockets in the central part of the board which will accept standard IC's (usually 14 or 16 pin). An IC can be plugged into each of the sockets and the adjacent pins on the board may be used to connect it to other components and the power supply using the interconnection leads supplied with the board.

TTL IC's are the standard way of implementing digital logic functions (AND, OR, NAND, NOR, XOR etc., referred to as gates) in discrete form. Details of the pin connections on TTL IC's can be obtained from the Texas Instruments TTL data book [5], which is available in the laboratory. Each TTL IC will require a power supply (+5 volts and 0 volts) connected to it before it will function at all. Any given TTL IC performs a specific logical function according to its internal design; this is specified by a code number of the form 74XXX where XXX is a two or three digit number. It is possible that you may encounter numbers of the form 74LSXXX or 74HCXXX and others, these represent different internal constructions in the device and are not important here. It is the numeric code that is important. For example 7400, 74LS00 and 74HC00 are all, four (or quad), 2-input NAND gates in one package—see TTL data book for details. (Note: Vcc is 5 V and GND is 0 V).

On the patchboard, there are also five input switches to introduce switched logic levels into an IC or combination of IC's by connecting the adjacent pins to those adjacent to the IC with connection leads. Similarly there are ten logic level test LED's (light emitting diodes) which can be used to indicate the logic level of the output of an IC or combination of IC's. A logic level 1 lights the LED, a logic level 0 extinguishes it.

There is also a free running square wave clock with fast and slow settings which produces a five volt peak-to-peak square wave output. You may determine its frequency with an oscilloscope. This can be used to drive circuits that require a constant clocked input.

2.3 Experimental procedure

2.3.1 Truth tables

Using a 74LS00 (or equivalent) IC, test the output of a two input NAND gate and draw a truth table. If one wished to use a three input NAND gate such as a 74LS10 (or equivalent) as a two input gate, what would one do with the unused input to preserve the logic? Test this experimentally. Using 74LS08, 74LS32, 74LS02 and 74LS86 IC's determine experimental truth tables for AND, OR, NOR and XOR logical functions.

2.3.2 Logical sums and products

Using AND (74LS08) gates and OR (74LS32) gates construct a circuit for the sum of products expression $X = AB + CD$, where A, B, C and D are the inputs and X is the output. Construct also a NAND (74LS00) only circuit for the same function and test the logical equivalence of the two circuits for all combinations of A, B, C and D.

2.3.3 Race hazards and optional products

Consider a circuit for the sum of products expression $X = AB + BC + \overline{A}C$ using NAND (74LS00 and 74LS10) gates only and prove that BC is an optional product (the circuit should behave the same without BC). TTL gates have finite switching times which can lead to a phenomenon known as 'race hazard'. In a sequence of cascaded gates, it is possible for unstable or ambiguous states to arise for example when the inputs to the final gate are derived from sequences with differing numbers of gates. What is the reason for this?

The product BC is optional in the above circuit but is present to exclude the race hazard. It would be difficult to see the race hazard in a TTL circuit as above since gate switching times are very fast, you may be able to see it with a very fast oscilloscope if the circuit is driven with a fast clock.

2.3.4 Adder circuits

An adder circuit can be used to add two binary inputs, as described in [2, p 71]. Construct both a half adder and a full adder using NAND gates, and test their operation. What is the purpose of the carry input and carry output in the full adder?

2.3.5 Decoders

A 2-bit decoder has two binary inputs to decode to 4 outputs since there are four possible combinations of the 2 inputs, such that one and only one of the outputs is high (logical 1) depending on the input combination. Construct a 2-bit decoder using NAND only logic and use this to light 4 LED's in turn using only 2 switches as inputs.

3 Machine code programming

3.1 Aim of the experiment

The aim of this experiment is to understand the basic operation of a computer by giving the computer instructions using machine code programming and examining how these instructions are stored and executed in the computer.

3.2 Apparatus

The 8080 Microtutor is a self contained microprocessor driven computer [6]. The Central Processing Unit (CPU) is an 8080 microprocessor. Data and programs are stored in 1 kByte memory (1024×8-bit words). The Microtutor has a 16-bit (a binary number with 16 digits) address bus with which it accesses with memory. The address bus is viewed by a sequence of 16 red LED's, which indicate in binary form the address being currently examined. This is presented as two 8-bit words (bytes), the HIADD—the topmost significant eight bits of the

sixteen bit number and the LOADD—the bottom or least significant 8 bits of the 16-bit number. The first few addresses are shown in Table 1. It is easy to see that the address is simply a binary number presentation of the position in memory.

	HIADD	LOADD
First address:	00000000	00000000
Second address:	00000000	00000001
Third address:	00000000	00000010
Fourth address:	00000000	00000011
Fifth address:	00000000	00000100
Sixth address:	00000000	00000101

Table 1: The binary representation of the first few memory addresses of the Microtutor.

The data bus, viewed by the 8 red LED's, represents the binary value of the eight bit word (byte) stored at the location specified by the address on the address bus. The contents of the data bus can be viewed modified—this is how to program the Microtutor. At each address in the memory specified by any given address bus location you will have an associated value stored in the byte at that location in the memory. The value stored at each location usually contains an instruction. The instructions are executed in logical order, i.e. the instruction stored in the first memory address is executed first, the instruction in the second address after that and so on.

The Microtutor also has an input port and an output port. This is how we can communicate with the Microtutor. The input port is a sequence of 8 switches, representing an eight bit binary number input. Each switch may be set in the 1 or 0 position. The output port consists of 8 red LED's representing an 8 bit binary number and is the means of getting output from the Microtutor.

The basic operation of the Microtutor is that the microprocessor loads instructions (the program) and data into itself by putting an address out on the address bus (over which it normally has control) and reading in data from the data bus (which then has the instruction or data on it that is at the address pointed to by the address bus).

3.2.1 Programming the Microtutor

There are two types of Microtutor in the laboratory, which are functionally identical apart from the method by which we enter programs.

One has a keyboard of push-buttons which are labelled 0–9 and A–F, with additional information (programs or data) in hexadecimal notation, i.e. in base 16. It is convenient to represent a 16-bit binary number as a number in base 16 instead. As there are not enough digits to count from 0 to 15, the letters A to F are used instead in the sequence:

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F$$

16 is then 10 in base 16 notation, 32 is 20 in base 16 notation and for example FF in base 16 would be 255 in base 10 (F = 15 in base 10, $FF_{16} = 15 * 16 + 15 = 255_{10}$). When we wish to enter data into the Microtutor we must convert it to base 16 (hexadecimal) notation. You will usually find this notation used in reference manuals for the Microtutor when referring to the instruction set for the microprocessor.

The other Microtutor has a sequence of 8 switches as inputs to the data bus which are used to input eight bit binary numbers into the machine. In this case data is entered in binary (or base 2) notation. For example the binary number

$$0000001_2$$

is equal to 1 in base 10, as it is also in base 16, and for example

$$10101111_2 = AF_{16} = 175_{10}$$

When using the Microtutor with 8 switches instead of the keyboard it will be necessary to convert all instructions to binary (base 2) notation.

The microprocessor in the Microtutor operates according to instructions (equivalent to commands in a high level programming language such as Basic, Fortran or C) stored in the memory. There may also be data stored in the memory. The instruction set for the 8080 is well defined and can be found in the reference manual for the Microtutor as a list of Mnemonics (these are the names of the instructions for example MVI, NOP) which tell you

what the instruction does and what operation the 8080 will carry out when this instruction is executed. Against each mnemonic there is a hexadecimal code and binary code which represents that instruction. A sequence of these instructions, in their binary form, forms a program in the memory of the Microtutor. Programs are entered either on the hexadecimal keypad or the sequence of 8 switches depending on which machine you are using.

The microprocessor operates on one instruction at a time by putting out an address on the address bus, loading the instruction from the memory, carrying out the instruction then moving to the next address location and so on. In a computer, this process would be very fast as the microprocessor would be driven by a clock (with a rate of the order of MHz), one clock cycle corresponding to one address/load cycle. On the Microtutor it is possible to run the machine in single step modes.

3.3   Experimental procedure

3.3.1   A simple example program

Entering the program   Set up the Microtutor as indicated on the side of the box, and the machine is ready for programming. Put the machine into programming mode using the MMLE (Machine Memory Load and Examine) switch set to MMLE (not RUN). Reset the machine using the reset button. The machine should now indicate the number

<div align="center">0000 0000 0000 0000</div>

on the address bus LED's (all LED's extinguished). If this does not happen repeat the reset until it does. This is the starting location in the memory for any program run on an 8080 microprocessor. The data bus LED's will almost certainly show a random binary number. Ignore this, it is the result of previous activity or a random sequence that formed when the machine was switched on. You can now begin entering data into the machine.

Program the sequence of instructions shown in Table 2 by entering their hexadecimal value on the keypad or their binary value on the data bus switches depending on which machine you are using. Press the step button (marked S) on the machine with the keyboard and the load button on the machine with binary switches only followed by the step button. This will then enter the instruction given by the hex or binary number into the memory and move to the next memory location. The address bus should now show:

<div align="center">0000 0000 0000 0001</div>

Repeat this procedure until all instructions are loaded into the machine. The machine is now ready to run your program.

Address (hex)	Instruction (binary)	Instruction (hex)	Instruction mnemonic
0000 0000	0011 1110	3E	MVI A,DB
0000 0001	0101 0110	56	data
0000 0010	0011 0010	32	STA Addr.
0000 0011	0000 1111	0F	Address (LOADD)
0000 0100	0000 0000	00	Address (HIADD)
0000 0101	0111 0110	76	HLT
0000 0110			

Table 2: A simple example program. All the addresses refer to LOADD. HIADD have been omitted and are all 0's.

Running the program   Run the program by switching the switch from the MMLE mode to the RUN mode, and press reset to get 0000 0000 0000 0000 on the address bus. This is the starting point of your program. Single step through the program using the step button. The single step mode should be set to SC as instructed on the machine. (N.B.: the machine may also be run in continuous mode by setting the SSTP/CONT switch to CONT. For some programs this may be necessary.)

The first instruction is $3E_{16}$ and translates to mnemonic MVI A,DB. This stands for MoVe Immediately into A, which is the internal register of the microprocessor (there are others: B,C, etc.) known as the accumulator, and can contain a data byte (one 8-bit word). The next entry in to the memory address 0000 0000 0000 0001 is the data byte to be stored by the MVI instruction (56_{16} in this example), so MVI is a two byte instruction. The following instruction at address 0000 0000 0000 0010 is STA Address: STore the content of the Accumulator (A) at

the following address. The next two entries in the contents of $0000\,0000\,0000\,0010$ and $0000\,0000\,0000\,0100$ are, respectively, the LOADD and HIADD of the location at which to store data, so STA is a three byte instruction. In this example the address is $0000\,0000\,0000\,1111_2$, $000F_{16}$. The last instruction is a HLT (Halt) instruction which stops the program. The program therefore simply moves a data byte from one location of memory to another via the internal register A (the accumulator) on the microprocessor. You can check your program by switching back to the MMLE mode and stepping through to location $0000\,0000\,0000\,1111$ on the address bus and you should see that the data bus will display 56_{16}, $0101\,0110_2$.

3.3.2 Using the input and output ports

Load the program shown in Table 3 into the memory using the same procedure as above and run it. You will find that the program uses the input and output ports to read in a number and write out a number. The instruction IN PORT 00 (a two byte instruction) reads from port 00, the sequence of switches marked 'input port' on the machine. This instruction is followed by two NOP instructions (one byte instructions). These are No-OPeration instructions and do nothing, however they do allow you to modify your program without typing it in again. These NOP's are followed by an OUT PORT 00 instruction (a two byte instruction) which outputs a number to the output port LED's on the machine. The final instruction is a JMP 0000 (a three byte instruction) which jumps to memory location 0000_{16}, i.e. to the beginning of the program. In continuous mode this program would continuously read the input port and write the contents of the input port to the output port before jumping back to the beginning. Try this, and should find that the binary number on the input port switches is repeated on the output port LED's.

Address (hex)	Instruction (hex)	Instruction mnemonic
0000	DB 00	IN PORT 00
0002	00 00	NOP, NOP
0004	D3 00	OUT PORT 00
0006	C3 00 00	JMP 00 00

Table 3: An example program that reads the input port and writes its contents to the output port.

Both the IN and OUT instructions operate on A (the accumulator), so that the IN reads into the accumulator and the OUT reads from it. In place of the NOP's in the program in Table 3 enter the following instructions (some 1 byte and some 2 byte), which operate on the accumulator:

(i) 3C, (ii) 3D, (iii) C6 10, (iv) D6 09, (v) 87, (vi) 27, (vii) AF, (viii) 07, (ix) 0F.

In each case run the program, record the outputs for inputs 0000 1001 and 0001 1011, and determine which operation the instructions carry out. You can use the Microtutor reference manual to help you interpret the instructions.

3.3.3 A more interesting example

The program shown in Table 4 is a program that reads the input port and tests if any bits are set to 1, It then outputs the highest bit only on the output port (for example, 00100011 would give 00100000). Enter this program and check that it works. Interpret its operation using the reference manual.

Address (hex)	Instruction (hex)	Instruction mnemonic
0000	3E 00	MVI A,00
0002	D3 00	OUT PORT 00
0004	DB 00	IN PORT 00
0006	C6 00	ADD 00
0008	CA 00 00	JZ 00 00
000B	4F	MOV C,A
000C	06 80	MVI B,80
000E	A0	ANA B
000F	C2 19 00	JNZ 00 19
0012	78	MOV A,B
0013	0F	RRC
0014	47	MOV B,A
0015	79	MOV A,C
0016	C3 0E 00	JMP 00 0E
0019	78	MOV A,B
001A	D3 00	OUT PORT 00
001C	C3 04 00	JMP 00 04

Table 4: A program that tests which is the highest input bit that is set to 1.

References

[1] Wikipedia. http://en.wikipedia.org/wiki/Digital_circuit.

[2] B. Holdsworth. *Digital Logic Design*. Butterworths, 2nd edition, 1987.

[3] B. Holdsworth and R. C. Woods. *Digital Logic Design*. Newnes, 4th edition, 2002.

[4] Wikipedia. http://en.wikipedia.org/wiki/Machine_code.

[5] Texas Instruments Limited. *TTL data book for Design Engineers*, 1973.

[6] http://www.limrose.com/MPT.htm.

Magnetic Circuits

1 Aim of the Experiment

Magnetic circuits are used in transformers, inductors, electric motors and magnetic recording heads. The overall aim of this experiment is to study some of the basic physical principles of such circuits and to provide familiarity with the factors that are important in their design. From the measurements you will make you will be able to compute the reluctance of magnetic circuits constructed using various materials and the relative permeability of these materials. You will also investigate the voltage-transforming capabilities of magnetic circuits.

2 Introduction and Theory

Just as an electromotive force (*EMF*) drives current through resistances that make up a closed electrical circuit, a magnetomotive force (*MMF*) drives a magnetic flux ϕ (unit Wb=Weber) round a closed magnetic circuit against the total reluctance R of the circuit. The *MMF*, ϕ and R are related by the equations (1, 2, 3). [1]

$$MMF = \phi R \tag{1}$$

A magnetic circuit is illustrated in Figure 1. The path is made up of a number of series elements through which the same flux passes; each has its own reluctance which is analogous to the electrical resistance of a series element in the electrical circuit.

An *MMF* is applied to a magnetic circuit by threading it with a driving current I. Ampere's Circuital Law states that:

$$MMF = \oint \overrightarrow{H} \cdot \overrightarrow{dl} = I \tag{2}$$

where $\overrightarrow{H}$ is the magnetic field strength (in A.m^{-1}) and $\overrightarrow{dl}$ is an element of the circuit length. [1] In the circuit the vectors $\overrightarrow{H}$ and $\overrightarrow{dl}$ are in the same direction.
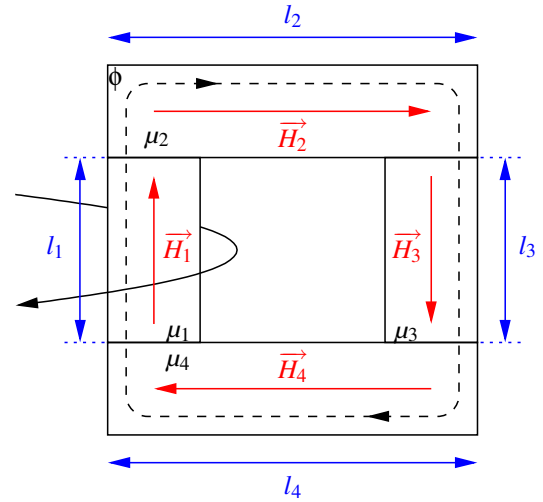

Figure 1: A schematic magnetic circuit.

Thus using the Circuital Law around a circuit with j limbs:

$$MMF = H_1 l_1 + H_2 l_2 + H_3 l_3 + \cdots + H_j l_j = I = Ni \tag{3}$$

where N conductors are threading the circuit, each carrying current i. $H_j l_j$ is the *MMF* 'dropped across' the jth limb.

Continuity of flux requires that ϕ is the same in every limb. Thus

$$\phi = H_1 \mu_1 A_1 = H_2 \mu_2 A_2 = H_3 \mu_3 A_3 = \cdots = H_j \mu_j A_j \tag{4}$$

where A_j is the cross-sectional area and μ_j is the absolute permeability of the jth limb. [2]

Thus, using (3) and (4),

$$MMF = Ni = \frac{\phi l_1}{\mu_1 A_1} + \frac{\phi l_2}{\mu_2 A_2} + \frac{\phi l_3}{\mu_3 A_3} + \cdots \frac{\phi l_j}{\mu_j A_j} \tag{5}$$

The total reluctance R of the magnetic circuit is thus

$$MMF = \frac{Ni}{\phi} = \frac{l_1}{\mu_1 A_1} + \frac{l_2}{\mu_2 A_2} + \frac{l_3}{\mu_3 A_3} + \cdots \frac{l_j}{\mu_j A_j} \tag{6}$$

i.e.:

$$R = R_1 + R_2 + R_3 + \ldots R_j \tag{7}$$

R is the sum of the limb reluctances. Equation (7) is analogous to the formula obtained for a series connection of resistances in an electrical circuit. [2] An electrical circuit can be used as an analogue model of a magnetic circuit. The driving MMF is divided between the elements in proportion to their reluctance, the element that has the highest reluctance needing the most MMF to sustain the same flux. One can also assemble parallel flux paths and calculate their equivalent reluctance as one would calculate the parallel equivalent resistance in an electrical circuit. [2, 3]

Equation (6) shows that the flux will be proportional to i in the coil that establishes the MMF. If a constant current i is used to establish the MMF that drives the flux around the circuit, the flux is constant and difficult to measure. It is much easier to detect a changing flux by allowing it to pass through a sense (or secondary) coil with a known number of turns n and to use Faraday's principle of magnetic induction to generate a voltage v in that coil equal to:

$$v(t) = n\frac{d\phi}{dt} \tag{8}$$

It is convenient to make i vary sinusoidally with a frequency f (Hz) so that $i(t) = i_0 \sin \omega t$ and (from equation (6)) $\phi(t) = \phi_0 \sin \omega t$ (with $\omega = 2\pi f$). It follows from (8) that

$$v(t) = v_0 \cos \omega t = n\phi_0 \cos \omega t \tag{9}$$

The peak amplitude v_0 of the cosinusoidally varying voltage is always proportional to ϕ_0 so long as the frequency is maintained constant. ϕ_0 is easy to calculate since

$$\phi_0 = \frac{v_0}{n\omega} \tag{10}$$
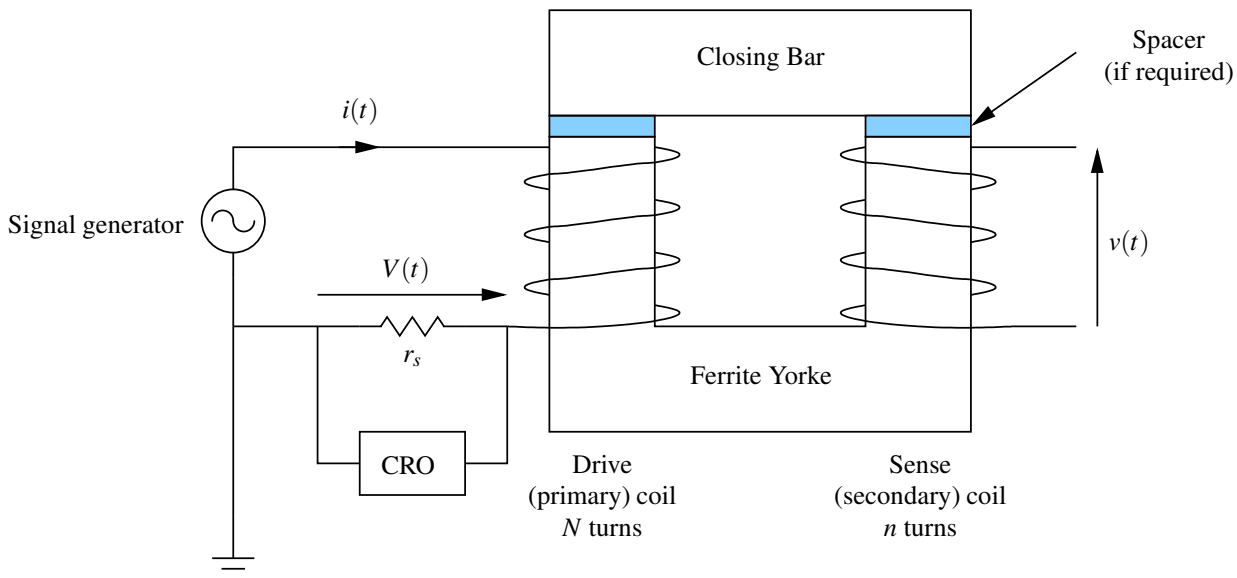
3   Apparatus

Figure 2: Experimental magnetic circuit set up. Looking from the top, if the turns of the primary coil are clockwise, they should be anti-clockwise in the secondary coil, producing a $90°$ phase shift.

You are supplied with a U-shaped piece of ferrite which forms three of the limbs of a magnetic circuit. It can be closed with a bar of the same material. Spacers of a different material can be inserted between the closing bar and the U to introduce further reluctances into the loop, as shown in Figure 2.

Four coils are provided each having 50 turns. Two or more of these coils can be used variously as drive and sense coils.

4 Experimental Procedure

4.1 Initial Investigations

The experimental set up is shown in Figure 2. Use the ferrite bar to close the loop. No spacers are to be inserted at this stage. Initially use one coil as a drive coil (N=50) and one coil as a sense coil (n=50). Set the signal generator frequency to 5 kHz.

Insert a resistor $r_s = 1$ kΩ, in series with the drive coil so that the coil current i can be easily related to the measured voltage across the resistor using $i = \frac{V_R}{r_s}$. Connect Y channel 1 of an oscilloscope (*CRO*) across this resistance and Y channel 2 across the sense coil, so that the current variation can be compared with the sense voltage using the two Y channels of the oscilloscope to display simultaneously signals which are proportional to i and ϕ.

> **Make sure that the two Y channels of the oscilloscope are set to 'Calibrated' for all of the experimental measurements.**

Set i to 1 mA (peak) using the CRO. Consider carefully how the best accuracy can be achieved when making the CRO measurements. Verify that the two signals v and I are respectively sinusoidal and cosinusoidal i.e. these two signals are in quadrature. Note the peak amplitude of the sense coil voltage. Explore the effect on the sense coil voltage of varying the frequency, being careful to keep the peak drive current, i.e. the peak flux, constant.

Demonstrate that the flux in the core is independent of position around the magnetic circuit, i.e. show that the position of both the drive and sense coils is not important.

Now use two coils connected in series as the sense coil and note that the sense coil voltage doubles as n is now 100. Notice that each coil has a red and a black lead to indicate the sense of the coil winding direction. Be careful to connect each coil in the same sense. Explore the effect of getting this wrong.

Now use one sense coil and one, two and then three drive coils **connected in series** (N = 50, 100 and 150) and note the results (making sure that **the same drive current is used for each value of** N). Explain the effect observed in terms of the net threading current $I = Ni$.

4.2 Measuring the Magnetic Flux ϕ

One drive coil and one sense coil should be used in this part of the experiment.

Choose a frequency in the range 1 to 10 kHz. This makes the product of n and ω a round number so that the induced sense coil voltage peak amplitude v_0 is easily related to the amplitude of ϕ using equation (10).

Set the peak drive current i_0 to 1 mA (peak $V_R = 1$ V) by varying the signal generator voltage. Measure the peak sense (secondary) voltage v_0, and hence deduce the peak flux ϕ_0 in the circuit. Repeat for at least eight other drive current values using the full amplitude range available from the signal generator. Plot a graph, to demonstrate that ϕ is proportional to i.

Explore the effect on the sense coil voltage v of varying the frequency, **keeping the peak drive current constant** at 1 mA. Plot a graph of v against frequency and calculate the peak flux for this current from the slope (making use of equation (10)).

4.3 The Circuit Reluctance and the Relative Permeability of Ferrite

For the magnetic circuit provided, all the limbs are made of ferrite (when no spacers are inserted) and have the same value of A and μ (=μ_F). Thus, if R_F is the reluctance of the whole ferrite circuit, (6) becomes

$$\frac{Ni}{\phi} = \frac{l_1 + l_2 + l_3 + l_4}{\mu_F A} = R_F \tag{11}$$

Measure the dimensions l and A of each limb using vernier callipers, and calculate an average value and standard deviation for each. From the slope of the graph obtained in Section 4.2, infer the reluctance R_F of the magnetic circuit. From this deduce the **relative** permeability μ_{rF} of the ferrite material. Note that $\mu_{rF} = \frac{\mu_F}{\mu_0}$ where $\mu_0 = 4\pi \times 10^{-7} H.m^{-1}$ is the absolute permeability of free space. Compare your value to the literature value.

4.4 The Effect of Inserting Spacers into the Circuit

Establish a value of peak flux in the closed circuit with the closing bar in place. Note the drive current. Remove the bar and observe the new value of the flux passing through the sense coil for the same drive current, altering the oscilloscope sensitivity if necessary.

Replace the bar and demonstrate that the original value of flux is obtained. Skew the closing bar and note how the flux is altered. Explain what has been observed.

Measure the total thickness Δl of each pair of spacers using a micrometer and calculate average value and standard deviation. Now insert the spacers, one pair at a time, into the circuit. For each spacer pair measure the sense coil voltage v_0 for a peak drive current of 1 mA and calculate the peak flux ϕ_0.

Equation (6) now becomes

$$\frac{Ni}{\phi} = R_F + \frac{\Delta l}{\mu_s A} = R_F + R_s \tag{12}$$

where μ_s is the absolute permeability of the spacers. Infer the reluctance of the circuit for each spacer pair.

Plot a graph of the total reluctance $R_F + R_s$ versus Δl. Under what circumstances would you expect this to be a straight line? You may assume that the effective area of the spacers A is the same as that of the ferrite limbs.

Use the spacer reluctance measurements to infer the **relative** permeability μ_{rs} of each spacer pair.

Show that the fraction of the total *MMF* which is 'dropped' across the spacer is $\frac{R_s}{R_F + R_s}$ and calculate this quantity for each spacer pair. Explain the fact that although Δl is small compared to the path length in the ferrite, most of the *MMF* is dropped across the spacer.

4.5 Transformer Action

Remove all the spacers and replace the closing bar. With reference to Figure 3 remove the resistor in the primary (drive) circuit and connect the signal generator and Y-channel 1 of the oscilloscope in parallel across a single drive coil. Use the three other coils at your disposal to make a transformer capable of simultaneously providing two independent secondary outputs, one with a doubling of the input voltage and another with the same magnitude as the input but of exactly opposite phase (see Figure 3). Verify that you have achieved this using Y-channel 2 of the CRO. Note that the primary and secondary *voltages* are not in quadrature as are the primary current and the secondary voltage (see Section 4.1).
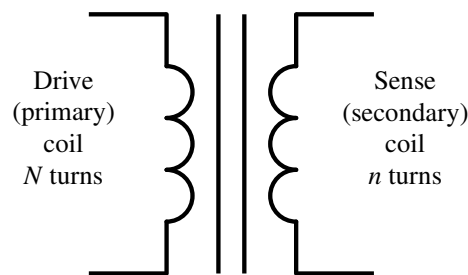

Figure 3: A Transformer.

5 Data Analysis and Conclusions

You should use a 'least squares' method to deduce the slopes and intercepts of the linear graphs obtained in the experiment and hence to obtain a best estimate for the various parameters which are measured, including their standard deviation. Take care to state the units of the measured parameters when you quote the result with an error estimate.

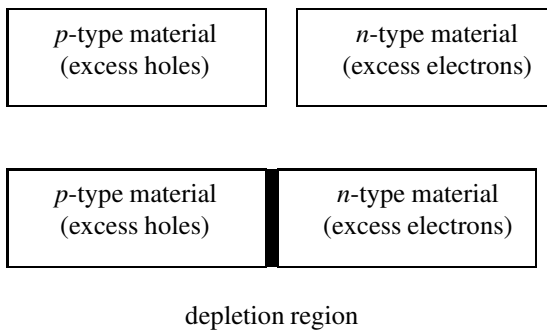The 'Conclusions' section should include a table summarising the results obtained.

References

[1] E. R. Dobbs. *Basic Electromagnetism*. Chapman and Hall, London, 1993.

[2] W. J. Duffin. *Electricity and Magnetism*. McGraw-Hill, London, 4th edition, 1990.

[3] B. L. Bleaney and B. Bleaney. *Electricity and Magnetism*, volume 1. Oxford University Press, Oxford, 3rd edition, 1989.

Characteristics of Semiconductor Diodes

1 Aim of the Experiment

The aim of the experiment is to investigate the variation of current with applied voltage for semiconducting *p-n* junction diodes under conditions of forward and reverse bias voltage, and to study how the reverse current varies as a function of the diode temperature.

2 Theoretical Background

Figure 1: Schematic of a *p-n* junction which is created when a *p* and *n*-type semiconductor are brought into contact.

A *p-n* junction is created by placing a *p*-type semiconductor, where a shortage of valence electrons leads to an excess of positive charge carriers ("holes") next to a *n*-type semiconductor with an excess of electrons (see Figure 1). Some of the free electrons diffuse across the junction to the *p*-type semiconductor and recombine with holes, thus creating the depletion region, a non-conducting layer between the *p* and *n*-type semiconductor material.

When a voltage is applied to a *p-n* junction, current will flow in one direction, but not the other.

The current *I* flowing through a *p-n* junction diode is given by [1]:

$$I = I_0 \left[\exp\left(\frac{eV}{kT} \right) - 1 \right] \tag{1}$$

where *V* is the voltage dropped across the junction region, *e* is the electronic charge, *T* is the absolute temperature and *k* is Boltzmanns constant (8.62×10^{-5} eV·K^{-1}). If *V* is negative and of magnitude greater than 0.5 volt the exponential term is $\cong 0$ and $I \approx -I_0$.

The term I_0 is known as the **reverse saturation current** and is a function of temperature. It can be shown that

$$I_0 \sim T^3 \exp\left(-\frac{E_g}{kT} \right) \tag{2}$$

where E_g is the energy gap of the semiconducting material from which the diode is made. [1]

Two factors limit the applicability of equations (1) and (2). Because of the resistance of the material from which the diode is made, the voltage *V* dropped across the junction is not equal to the potential difference V_0 measured across the diode as a whole, but is related to it by

$$V = V_0 - IR_1 \tag{3}$$

where R_1 is the effective series resistance of the diode (R_1 is very small). The characteristic curve of the diode, *I* plotted as a function of V_0, therefore departs from the form of equation (1) at large positive values of V_0. [1]

When V_0 is negative, equation (1) shows that the current tends to a saturation value of $-I_0$. A true saturation is not always observed in practice because there is also a stray current which flows round the edges of the junction and in some cases this may be comparable with or greater than I_0 (in magnitude). This stray current can be considered to flow through a very large resistance R_2 in parallel with the diode, and the equivalent circuit of the diode is represented in Figure 2. J represents the junction region which obeys equation (1); R_1 and R_2 are the series and shunt resistances of the diode, with typical values 0.1 Ω and 10^6 Ω respectively.

The stray current $\frac{V}{R_2}$ becomes negligible when V_0 is in the forward direction, and indeed may be negligible in the reverse direction for some diodes. However, if the leakage is significant in this region, the plot of current against voltage for $V_0 \sim -0.5$ volt will have a slope of $\frac{1}{R_2}$. If this is extrapolated back to $V_0 = 0$, the true value of I_0 may be found at the current diode temperature.
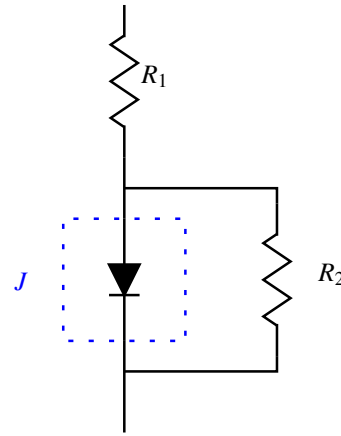

Figure 2: Equivalent circuit of the diode.

3   Experimental Procedure

Two *p-n* junction transistors, one silicon (OC200) and one germanium (OC71), are mounted on a thermally insulated copper block that can be heated electrically. The transistors are used as diodes (base emitter junction) and the forward and reverse characteristics of each diode are to be obtained at room temperature. The temperature dependence of the reverse current should then be investigated for the germanium diode. (The reverse current for the silicon diode is considerably smaller than that of the germanium diode and it is difficult to obtain reliable results in the time available.)
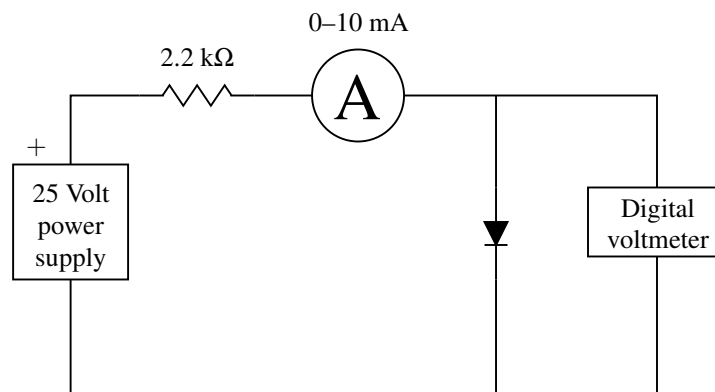
3.1   Forward characteristics

Figure 3: Circuit for measuring the forward diode characteristics.

Connect the circuit as shown in Figure 3. By varying the power supply voltage in the range 0 to 25 volts, the current through the diode may be adjusted in the range 0 to 10 mA. The voltage dropped across the diode at each current is measured on a digital voltmeter.

It is recommended that the curves are plotted during the experiment so that more readings may be taken where the characteristic changes rapidly.

3.2   Reverse characteristics

The reverse currents for these diodes are small and necessitate the use of a sensitive measuring instrument. A Levell type TM9BP Multimeter is used which will easily measure currents in the nanoampere range.
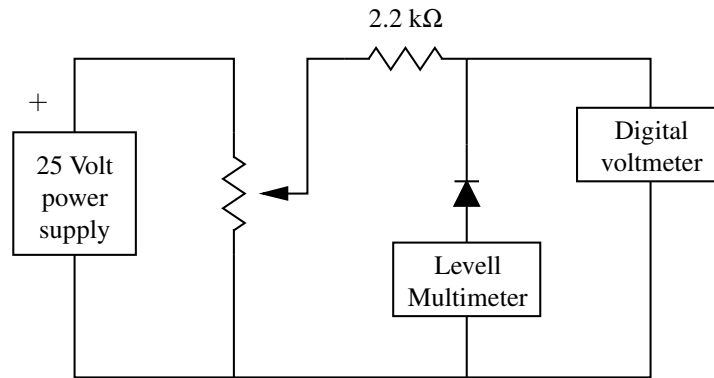
Figure 4: Circuit for measuring the reverse diode characteristics.

An appropriate circuit for measuring the reverse characteristics is shown in Figure 4. The power supply voltage should be set to about 0.5 volt and the voltage across the diode may then be finely adjusted in the range 0 to -0.5 volt using the 10-μm turn potentiometer.

The voltage recorded on the digital voltmeter (*DVM*) is the sum of the voltages across the diode and the Levell meter. The voltage across the diode must be calculated from

$$V\,(\text{diode}) = V\,(\text{DVM}) - IR \tag{4}$$

where I is the current recorded by the Levell meter and the meter resistance R depends on the range used, as follows:

For function switch on pA range $R = 1\ \text{M}\Omega$
For function switch on nA range $R = 100\ \text{k}\Omega$
For function switch on μA range $R = 100\ \Omega$

3.3   Band gap determination from temperature dependence of I_0 in reverse mode

We see from equation (2) that if I_0 is measured as a function of T a graph of $(3\log T - \log I_0)$ plotted against T^{-1} will yield a straight line of slope $\frac{E_g}{k}$, from which the energy gap of the semiconductor material may be estimated. The value of E_g itself is slightly temperature dependent, and an average value over the temperature range covered will be obtained. If the plot of reverse current against voltage does not exhibit a true saturation it should be extrapolated back to $V_0 = 0$ to obtain the true value of I_0 at a given temperature. Since R_2 (as defined in Figure 2) will probably vary with temperature, this procedure should be carried out at each temperature investigated. See Figure 5.

Slowly increase the temperature of the copper block by connecting the "Variac" AC power supply to the heater terminals. The temperature is measured using a thermocouple and an electric thermometer. When a steady temperature has been achieved a series of reverse voltage and current readings should be obtained, using the circuit of Figure 4. From these readings, values of I_0 may be obtained.

Do not exceed a temperature of 100 °C.

3.4   Band gap determination from temperature dependence of voltage in forward mode

Another way to determine the band gap is by measuring the voltage drop across the p-n junction as a function of temperature at constant current. [2]

For appropriate experimental conditions, the 1 in eq (1) can be neglected. Furthermore, if the T^3 term in eq (2) can be considered to be constant in comparison to the exponential term, combination of eqs (1) and (2) yields

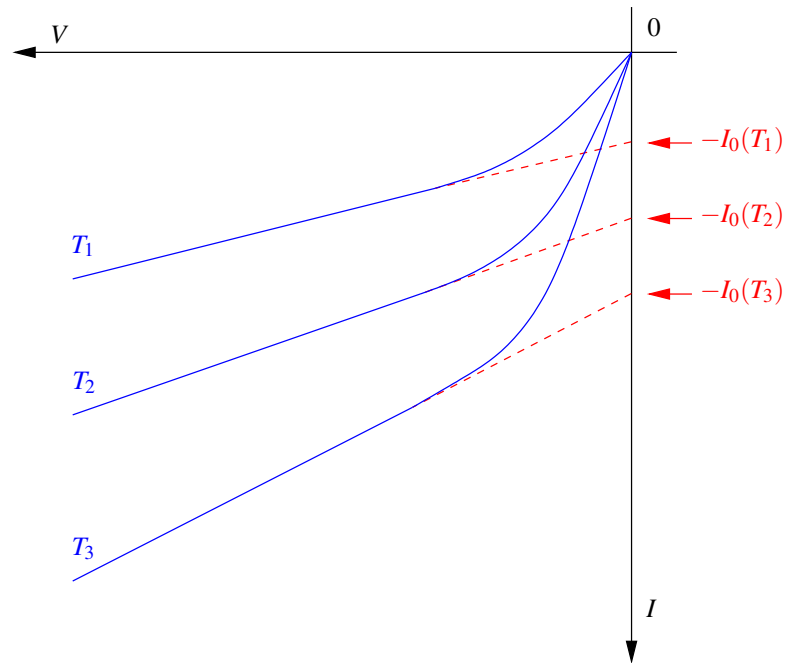$$I = B\left[\exp\left(-\frac{E_g}{kT} + \frac{eV}{kT}\right)\right] \tag{5}$$

Figure 5: Diagram showing how the reverse saturation current can be found by extrapolation of the linear part of the graph.

where B is a constant. Keeping the current I constant,

$$C = -\frac{E_g}{kT} + \frac{eV}{kT} \tag{6}$$

where $C = \ln\left(\frac{I}{B}\right)$. In terms of the temperature T

$$T = -\frac{E_g}{kC} + \frac{eV}{kC} \tag{7}$$

Therefore, a straight line fit of temperature versus voltage drop at constant current yields a gradient $\frac{e}{kC}$ and an intercept of $\frac{E_g}{kC}$. The bandgap E_g can be calculated from

$$E_g = e\frac{\text{intercept}}{\text{gradient}} \tag{8}$$

Conduct this experiment for both the Ge and Si diode, and compare your values for E_g of Ge with the previous reverse bias method. Also compare your values to the literature values. A copy of reference [2] is available from your demonstrators.

References

[1] J. R. Elliot and A. F. Gibson. *An Introduction to Solid State Physics and its Applications*. MacMillan, London, 1974.

[2] J.W. Precker and M. A. da Silva. Experimental estimation of the band gap in silicon and germanium from the temperature-voltage curve of diode thermometers. *American Journal of Physics*, 70:1150–1153, 2002.

Michelson Interferometer

1 Aim of the Experiment

This experiment demonstrates the phenomenon of interference with visible light, and illustrates how the Michelson interferometer can be used to make accurate distance measurements, and to measure unknown wavelengths and wavelength differences.

2 Apparatus

The components of the interferometer are housed in a storage box for safety. The principle reason for this is that the mirrors used to reflect the light beams into a path that allows them to interfere are front-surface aluminised and are extremely easily damaged if handled carelessly. **Never touch the mirror surfaces with your fingers.** Each of the main components should have a threaded support rod that allows it to be mounted securely on the interferometer base, using the pre-drilled holes to locate them in the correct position. A schematic diagram of the Michelson interferometer is shown in Figure 1.
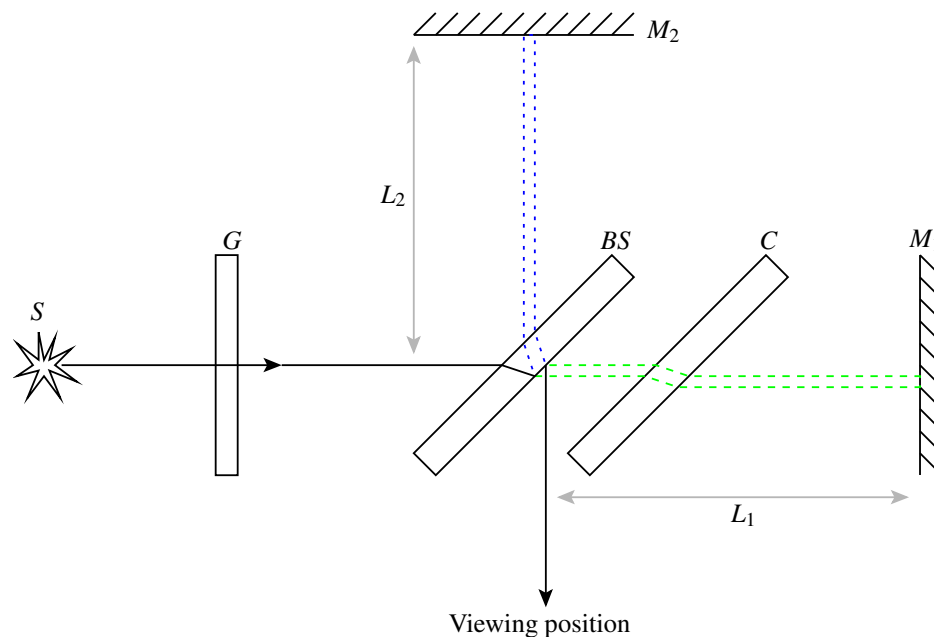

Figure 1: Schematic diagram of a Michelson Interferometer

The light source that is to be used first contains two separate lamps, one a conventional bulb providing white light (a continuous range of wavelengths over the whole of the visible spectrum), the other a mercury vapour lamp that provides a few discrete emission lines (characteristic of mercury) from the violet, blue, green, yellow and red parts of the visible spectrum. Use of suitable coloured filters in front of the mercury lamp allows light with only a very narrow range of wavelengths to be used for experiments. In front of the source S it is usual to place a ground glass screen G to ensure that the illumination is spread uniformly, and to prevent the internal structure of the source being visible during the experiment.

Light from the extended incoherent source G is split into two beams of equal intensity by beam-splitter BS. Mirror M_1 has tilt adjustments to ensure that it is accurately perpendicular to mirror M_2. The path length from BS to M_2 is controlled by a sensitive micrometer adjustment, with a lever mechanism providing a reduction of

about 5:1 between the motion of the micrometer and the motion of the mirror surface. The compensating plate C is situated between BS and M_1 so that the divided beams have equivalent paths though glass as well as through air. After reflection at the mirrors M_1 and M_2 the beams recombine at beam-splitter BS and may be viewed directly with the eye.
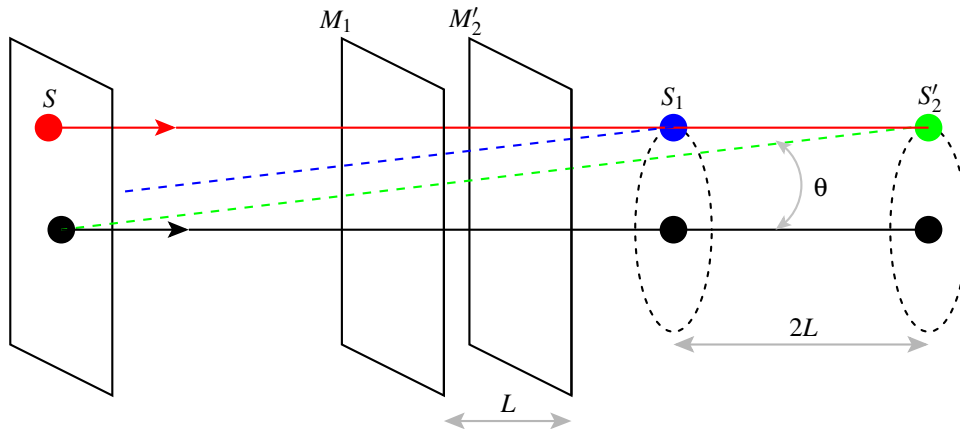

Figure 2: Virtual sources S_1 and S_2'

In practice it appears to an observer that the interference is between light from two virtual sources S_1 and S_2' reflected from two mirrors M_1 and M_2' (conjugate to M_2 by reflection symmetry) that are separated by a distance $L = (L_2 - L_1)$, as shown in Figure 2. A single source point S will have a virtual image S_1 after reflection in mirror M_1, and a virtual image S_2' after reflection in M_2'. These source images will be coherent with one another as they are derived from the same source point, so light from these virtual sources travelling in a common direction at angle θ to the optical axis can interfere to produce an interference pattern.

3   Theoretical Background

It is easy to see from Figure 2 that the path difference between rays from virtual sources S_1 and S_2' is $2L\cos\theta$, where the factor of 2 arises because the light must travel in *both* directions between the beam-splitter and either mirror M_1 or M_2. The condition for destructive interference[1] will be given by

$$m\lambda = 2L\cos\theta \tag{1}$$

where m is an integer and λ is the wavelength of the illumination.

The system has axial symmetry about the optical axis, so it is clear that when mirrors M_1 and M_2' are effectively parallel, the interference fringes will be a set of concentric circles. If the path difference L is decreased, $\cos\theta$ will increase, and thus the angular radius θ corresponding to fringe m will decrease accordingly, so the fringe pattern will appear to contract towards the optical axis. Conversely, when L is increased, the fringe pattern will expand, with new fringes appearing at the centre. As the path difference L approaches zero, the number of fringes visible in the field of view decreases (**make sure you understand why this is so**), and this can be a useful way of telling whether motion of the mirror M_2 is increasing or decreasing the path difference between the two beams.

If the mirrors are not exactly perpendicular to one another, the fringes will appear to be circular arcs, with a centre of curvature away from the centre of the field of view. It is possible to adjust the tilt of mirror M_1 to turn the fringes into straight lines, which makes some form of measurement easier to accomplish. The separation of these straight-line fringes depends on the relative tilt between the mirrors.

Since the distance L need change by only $\frac{\lambda}{4}$ to change a bright fringe at the centre of the pattern into a dark one, the fringe pattern is very sensitive to small changes in the path difference. This means that the interference fringes can be used to make very precise measurement of distance.

---

[1]Normally when the path difference is an integer multiple of λ this would give rise to constructive interference. However, a careful examination of Figure 1 shows that rays reflected off mirror M_1 undergo a total of two external reflections, whereas rays reflected from mirror M_2 undergo one external and one internal reflection. A phase change of π accompanies external reflections, but not internal reflections (see the texts listed in the References page 74 for a further discussion of Fresnel's Laws of Reflection), so even when distances L_1 and L_2 are equal, the rays would recombine with a π phase difference between them.

When two closely-spaced wavelengths λ and $\lambda + \Delta\lambda$ are present in the illumination each wavelength will produce its own set of interference fringes, and it is possible that the bright fringes of one pattern will occur at the same angular radius as the dark fringes from the other wavelength.

This type of spectrometer is named after Albert Abraham Michelson (1852–1931) who received the 1907 Nobel Prize in Physics "*for his optical precision instruments and the spectroscopic and metrological investigations carried out with their aid*" [1]. A Michelson interferometer was used to detect the motion of the earth through the ether [2], a hypothetical medium for electromagnetic waves to travel in. No motion could be detected, the ether concept was abandoned and Einstein (1879–1955), 1921 Nobel Laureate in Physics "*for his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect*" [3], proposed that the speed of light is the same in all reference frames. Large Michelson interferometers are used today to detect gravitational waves, i.e. ripples in space-time caused by (asymmetric) supernova explosions.

4   Experimental Procedure

4.1   Alignment of the interferometer

Interference fringes can only be seen by eye if the two interfering beams are almost parallel to one another (typically to within minutes of arc). This means that, in the Michelson interferometer, the two images of a pointer placed at the broad source (that is, the ground glass screen G) seen in reflection in mirrors M_1 and M_2 should appear coincident when the mirrors are accurately perpendicular to one another. When the images of the pointer coincide, the mirrors should be well enough aligned for the interference fringes to appear.

Place the source in position and switch on the mercury lamp. Align the eye with the viewing axis about 25 cm from the beam-splitter BS. Usually three reflected images of the pointer (the arrow drawn on the ground glass screen) will be visible in the beam splitter, but two will be of higher contrast than the other.

Adjustment of the tilt controls on mirror M_1 will show that one of the high-contrast shadow images can be made to move. The first adjustment is to superimpose this movable image on the other high-contrast shadow image. When this has been done, interference fringes should be seen. The nature and location of the fringes are determined by the position of the mirrors relative to each other, as seen in the beam-splitter. Further small adjustments of the tilt of mirror M_1 should be enough to make the fringe pattern circular.

4.2   Calibration of the lever reduction arm

Use the mercury discharge lamp as the light source for the interferometer, and place a green filter in front of the lamp so that all spectral lines except the green line of wavelength 546.1 nm are eliminated from the transmitted light. Ensure that the mirror alignment is adjusted to give circular interference fringes, and use a small pointer located on the screen G to mark a fixed point on the fringe pattern. Note the start position on the micrometer scale, and then slowly turn the micrometer dial, counting the number of fringes that pass by the fixed point in the field of view. Each time a complete period of the fringe pattern has passed, mirror M_2 will have moved by a distance equal to $\frac{\lambda}{2}$, so by noting the change in micrometer reading corresponding to the passage of a given number of fringes it is possible to calibrate the mechanical reduction R introduced by the lever arm between the mirror and the micrometer.

If the passage of m fringes corresponds to a distance x moved by the micrometer, then the reduction ratio is given by

$$R = \frac{2x}{m\lambda} \tag{2}$$

Since the micrometer has a smallest scale division of 0.01 mm, it will be necessary to count 50 or 100 fringes or so and record the micrometer reading. Continue to count another 50 or 100 and record the micrometer reading until you have counted several hundred fringes altogether. Plot the distance moved versus the number of fringes counted. Rearrange equation (2), and extract the value of R from the gradient of a straight line fit through the origin. It should be between 5 and 6.

4.3   Measurement of the mean wavelength of yellow sodium light

Replace the mercury lamp with a sodium discharge lamp. By far the strongest emission from this lamp comes form two closely-spaced yellow sodium lines known as the *D lines*, so it is not necessary to use a filter in front of

this lamp. This experiment is essentially the reverse of that in section 4.2, in that it is again necessary to measure the micrometer movement x corresponding to the passage of m interference fringes across a fixed point in the field of view. In this case however, equation (2) can be rearranged to allow the mean wavelength of the two D lines to be determined.

4.4 The observation of white-light interference fringes

Using the mercury lamp (without the green filter) obtain circular fringes, and turn the micrometer dial until you are sure you know which direction you need to turn the micrometer in order to reduce the path difference L between the mirrors. **Be sure you can explain why this is the good direction.**

Now adjust the tilt of mirror M_1 to observe fringes that are slightly curved in one direction, with the centre of curvature out of the field of view, and continue turning the micrometer drive to reduce the path difference L. Continue turning the micrometer until the curvature of the fringes has reversed, which indicates that you have moved through the position corresponding to $L = 0$.

Now replace the mercury source by the white light source, and very slowly turn the micrometer in the opposite direction, so that you are again approaching the position of zero path difference. Soon you should find *a very short range of micrometer positions* where a brightly coloured fringe pattern is visible with the white light source. **Explain why this range is so short.**

The centre of this pattern should correspond to a black fringe, which is the point where the path difference is exactly zero.

Clearly, when $L = 0$, all wavelengths in the white light source will simultaneously produce a dark fringe. Around this point there is a narrow band of colours, and you should record a description of these bands in your notebook. Try to explain, at least qualitatively, the reasons for the pattern of colours that you observe.

4.5 Measurement of the refractive index of a thin transparent film

Set up the interferometer to use white light fringes again, and note the micrometer setting that this corresponds to. Insert a thin transparent sheet (of glass or perspex, or other material supplied by a demonstrator) into the path between BS and M_2, ensuring that the sheet is perpendicular to the path the light travels. The white light fringes will now have disappeared, as the optical path in the two arms of the interferometer is no longer the same. Now, with the transmitting sheet still in place, turn the micrometer to once again find white light fringes, and note the micrometer reading that this corresponds to.

Finally use a micrometer to measure the thickness t of the thin transmitting sheet.

4.6 Measurement of the wavelength difference of the two sodium D lines

The yellow sodium light consists of two closely spaced wavelengths λ and $\lambda + \Delta\lambda$ so that, as the position of mirror M_2 is changed, the visibility of the fringes alternately increases and decreases depending on whether bright fringes associated with one wavelength coincide with the dark fringes from the other wavelength. Assuming the two wavelengths are of equal intensity, there can be almost complete cancellation of the fringe patterns when this happens, and the visibility of the two sets of fringes will be almost zero.

Set up the interferometer to produce white light interference fringes, and adjust the tilt of mirror M_1 to produce straight-line fringes in the centre of the field of view. Replace the white light source by the sodium lamp, but do not adjust the mirror positions, so there should be high contrast fringes from the yellow sodium light. Note the start position of the micrometer, $x_{\frac{1}{2}}$,[2] and slowly turn the micrometer. Continue turning the micrometer until the visibility of the fringes has reached a minimum, and the note the new micrometer reading x_p. Continue turning the micrometer in the same direction, when you will see that the fringe visibility first improves again, and then deteriorates. Record the micrometer reading corresponding to the next few positions of minimum visibility.

Since the measurements started at the position where the path difference L was zero, the path difference L_p at the pth position of minimum visibility corresponds to the condition

$$\left(m_p + p - \frac{1}{2}\right) = m_p\left(\lambda + \Delta\lambda\right) = 2L_p\cos\theta \tag{3}$$

---

[2]Why does the $L = 0$ position corresponds to $x_{\frac{1}{2}}$?

where m_p and p are positive integers, and $L_p = \frac{\left(x_p - x_{\frac{1}{2}}\right)}{R}$. If we consider only the visibility at the centre of the pattern, then we can set $\cos\theta = 1$. Since $m_p \gg p$, we can write $m_p \approx 2\frac{L_p}{\lambda}$ and then rearrange (3) to find

$$\left|x_p - x_{\frac{1}{2}}\right| = \left(\frac{R\lambda^2}{2\Delta\lambda}\right)\left(p - \frac{1}{2}\right) \tag{4}$$

4.7   Verification of the Fresnel-Arago laws

The Fresnel-Arago laws describe the interference of polarised light. They may be stated as follows:

1. Two beams of coherent light that are polarised in parallel directions will produce interference fringes in the same way as unpolarised light.

2. Two beams of coherent light that are polarised in perpendicular directions cannot produce interference fringes.

3. Two beams of coherent light that are derived from perpendicular directions of polarisation will not produce interference fringes even when the polarisation directions are rotated to be parallel with one another.

Using the three pieces of polarising filter that are available with the experiment, devise and carry out a set of experiments to test the Fresnel-Arago laws. Comment on whether or not your tests do indeed verify the laws stated above.

5   Data Analysis

5.1   Calibration of the lever reduction arm

Calculate the mean and standard deviation for your measurement of R in section 4.2, and comment on the main source of error in this measurement. Discuss briefly how you could test for the presence of systematic error in your value of R.

5.2   Measurements of the wavelength of yellow sodium light

Compare your measured value for the mean wavelength of the two D lines with the accepted value, and comment on any differences that you find.

Use your measurements of the positions of minimum visibility in section 4.6 to plot a graph of $\left|x_p - x_{\frac{1}{2}}\right|$ against p. From the slope of this graph, determine the wavelength separation $\Delta\lambda$ of the two D lines, and again compare your results with the accepted values.

5.3   Measurement of the refractive index of a transparent sheet

For a material of thickness t and refractive index μ the optical path length through the material is μt. Thus, if the positions of the micrometer that produce white light fringes with and without the sheet in place are x_t and x_0 respectively, the refractive index of the sheet is given by the expression

$$\mu = 1 + \frac{|x_t - x_0|}{Rt} \tag{5}$$

Make sure that you can derive this result for yourself, and use equation (5) to calculate a value for the refractive index of the sheet you used. Compare your value with the accepted value, and comment on the main sources of experimental error. Remember that the definition of the refractive index is the ratio of the speed of light in vacuum, c_0, to the speed of light in the medium, c, i.e. $\mu = \frac{c_0}{c} \geq 1$.

References

[1] http://nobelprize.org/physics/laureates/1907/index.html.

[2] Albert A. Michelson and Edward W. Morley. On the relative motion of the earth and the luminiferous ether. *The American Journal of Science*, 24(203):333–345, November 1887. http://www.aip.org/history/gap/PDF/michelson.pdf.

[3] http://nobelprize.org/physics/laureates/1921/index.html.

Fraunhofer Diffraction

1 Aim of the Experiment

The purpose of this experiment is to illustrate how the measurement of the Fraunhofer diffraction pattern produced by small apertures or obstacles can be used to determine the linear dimensions of the objects.
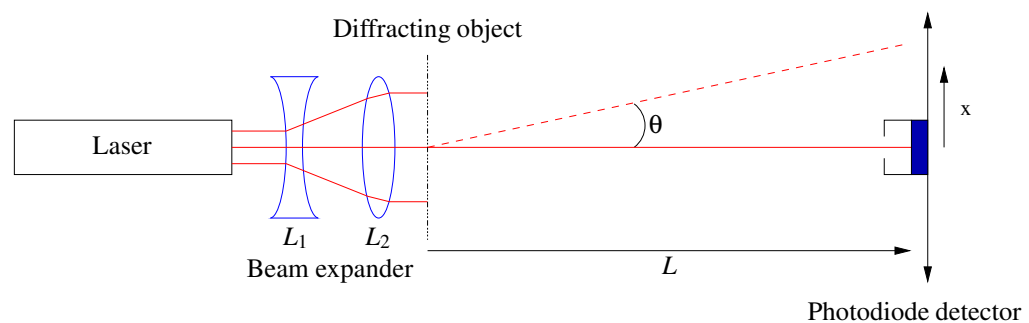
2 Apparatus

Figure 1: Schematic diagram of the set-up for the Fraunhofer diffraction experiment.

The equipment used in this experiment is shown schematically in Figure 1. The illumination is provided by a 2 mW He/Ne laser providing light of wavelength 632.8 nm.

> **Warning: Never look directly at the laser source, nor at a specular reflection of the light! Read carefully the section on laser safety in the safety notes issued with this folder. DO NOT SWITCH ON THE LASER unless you are sure that you understand the hazard. If in doubt, consult a demonstrator.**

The laser beam is diverged somewhat by lens L_1, a lens of power -10 D placed just in front of the laser, and is then collimated by lens L_2, a converging lens of power $+4$ D. *(The power of a lens, expressed in dioptres, D, is the reciprocal of the focal length in metres).* The diffracting object is either a single rectangular slit or a linear array of up to 6 identical slits, and is placed between L_2 and the photodiode, as close as possible to lens L_2.

The diffracted intensity is measured by scanning a narrow slit perpendicular to the optical axis of the system. Behind the slit is a solid-state photodiode that produces a voltage output proportional to the incident intensity. There is also a filter in front of the photodiode that allows only the red light from the laser to pass through, so that the experiment can be carried out under normal lighting conditions. The output from the photodiode is measured by directing the voltage to an analogue-to-digital converter (ADC) attached to a PC, and suitable software is used to record a time series of data as the detector slit is scanned across the diffraction pattern.

3 Theoretical background

Diffraction patterns arise as a result of interference between rays from different parts of an aperture, or from different points around an opaque object. Two kinds of diffraction may be distinguished. Fraunhofer diffraction patterns are observed when the distance between the point of observation and the source of light is effectively infinite so that the interfering waves may be treated as plane waves. This includes diffraction patterns observed in the focal plane of a lens placed on the other side of the aperture. Fresnel diffraction patterns are observed when

distances involved are sufficiently small that the plane wave approximation is not valid, and the curvature of the wavefronts must be taken into account.

The experiment is named after Joseph von Fraunhofer (1787–1826) who worked in the field of optics in the 19th century [1]. The experiment is central to optics, and will be extensively discussed and extended in Dr Morrison's 3rd year course CP3380 *"Optics"* [2].

3.1 Diffraction at a single slit

When a plane wavefront of wavelength λ falls normally on an opaque screen with a transparent slit of width a the distribution of the light intensity in the Fraunhofer diffraction pattern is given by

$$I = I_0 \frac{\sin^2 \beta}{\beta^2} \tag{1}$$

where $\beta = \frac{\pi a}{\lambda} \sin \theta$. It is clear that $I = 0$ whenever $\beta = n\pi$ (for non-zero integer n), that is when $\sin \theta = \frac{n\lambda}{a}$. The form of the diffracted intensity is shown in Figure 2. Further theoretical details can be found in references [3, 4].
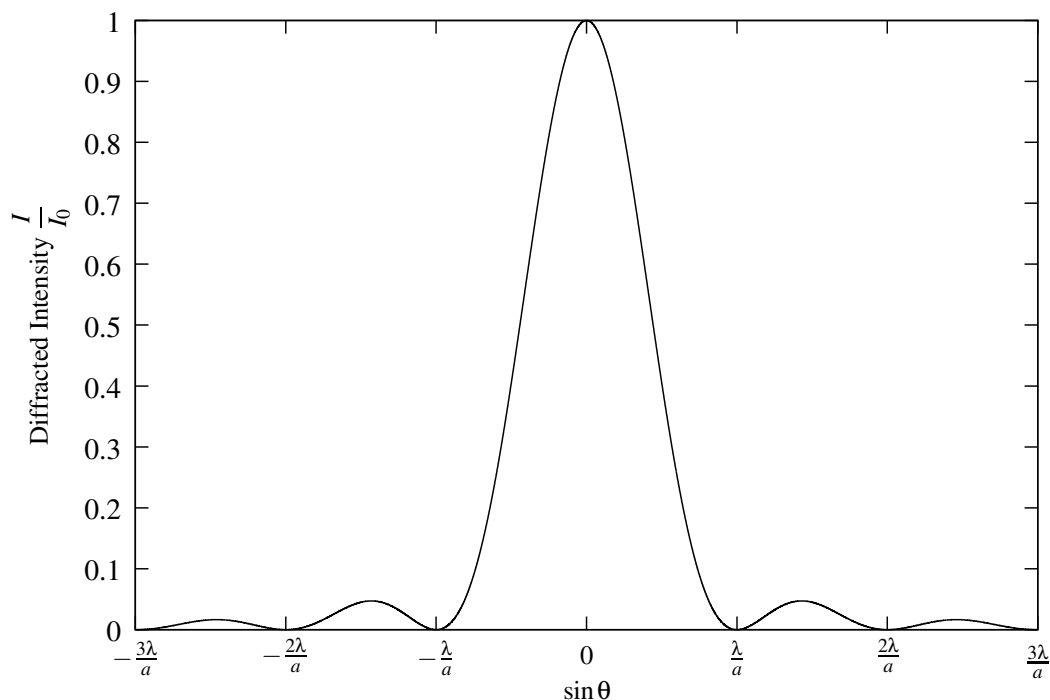

Figure 2: Fraunhofer diffracted intensity as a function of diffraction angle from a single slit of width a.

When $a \gg \lambda$ the small-angle approximation for sine means that the condition for a dark band (i.e. **an intensity zero**) may be written in the form

$$\theta = \frac{n\lambda}{a} \tag{2}$$

Thus if the angular separations of the first few dark bands are measured, the slit width can be calculated. Note that it is not so easy to use the positions of the intensity maxima. Theoretically the maxima also correspond to turning points of the function in equation (1), but it is easy to show that the maxima correspond to solutions of the equation $\tan \beta = \beta$, an equation that can only be solved numerically.

3.2 Diffraction at multiple slits

If there are two or more identical slits, each of width a that are separated by a centre-to-centre distance d, there is an additional term in the diffracted intensity distribution that arises from interference of the waves coming form

the different apertures. When there are N slits the intensity distribution is given by

$$I = I_0 \frac{\sin^2 \beta}{\beta^2} \left(\frac{\sin^2 N\gamma}{\sin^2 \gamma} \right) \tag{3}$$

where $\gamma = \frac{\pi d}{\lambda} \sin\theta$ and β is as previously defined. The bracketed term in this equation is known as the **interference term** for N slits. This factor has maximum values of N^2 whenever $\gamma = m\pi$. From the definition of γ and when $d \gg \lambda$, it follows that these principal maxima occur at

$$\theta = \frac{m\lambda}{d} \tag{4}$$

where m is an integer. The relative intensities in the different orders m are governed by the single slit **diffraction envelope** $\frac{\sin^2 \beta}{\beta^2}$. This is illustrated in Figure 3 for $N = 3$ and $d = 4a$. Note that a principal maximum of the interference term will coincide with a minimum of the diffraction envelope when

$$\theta = \frac{m\lambda}{d} = \frac{n\lambda}{a}$$

When this occurs, the principal maximum is suppressed, leading to a '**missing order**'. Missing orders occur when the condition $d/a = m/n$ is satisfied. In other words d/a must equal the ratio of two integers.

The numerator of the function $\frac{\sin^2 N\gamma}{\sin^2 \gamma}$ oscillates N times faster than the denominator, so there will be a number of minima (and consequently secondary maxima) between consecutive principal maxima at $\gamma = m\pi$ and $\gamma = (m+1)\pi$. These minima occur at $\gamma = p\pi/N$ for integer values of p excluding those values for which $p = mN$. Thus there will be minima whenever

$$d \sin\theta = \frac{\lambda}{N}, \frac{2\lambda}{N}, \frac{3\lambda}{N}, \ldots, \frac{(N-1)\lambda}{N}, \frac{(N+1)\lambda}{N}, \ldots \tag{5}$$

Between two adjacent principal maxima there are therefore $(N-1)$ points of zero intensity. There must be a maximum between consecutive minima, so there will be $(N-2)$ secondary maxima between two consecutive principal maxima. This is illustrated in Figure 3, which shows a calculation of the diffracted intensity for a 3-slit grating; between consecutive pairs of principal maxima it is clear that there are 2 minima, and 1 small secondary maximum.
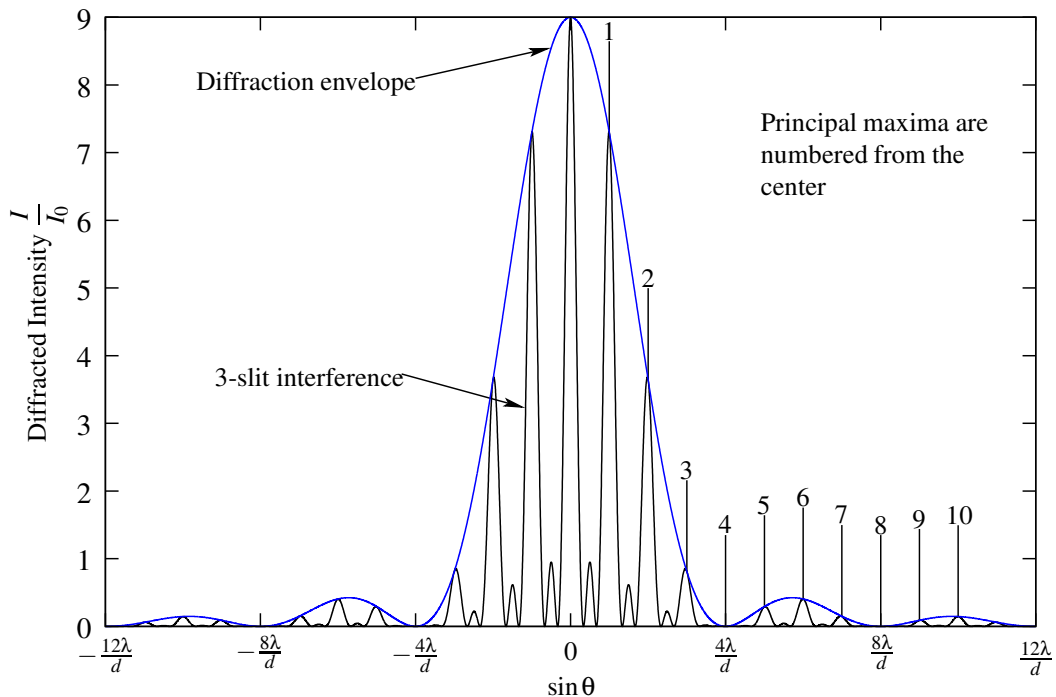

Figure 3: Fraunhofer diffracted intensity as a function of diffraction angle from three slits of width a that are separated by centre-to-centre distance $d = 4a$.

4 Experimental procedure

4.1 Alignment of the optical system

The apparatus shown in Figure 1 must be carefully aligned if the best results are to be obtained. Assuming that it is completely out of adjustment the procedure is as follows:

1. Remove all components from the optical bench except the laser and a small diaphragm aperture.

2. Sequentially adjust the tilting screws on the laser and the height of the pinhole to ensure that the laser beam passes through the aperture for all positions of the aperture on the optical bench and is at the same height as the detector slit. The laser beam is now parallel to the optical bench.

3. Move the diaphragm aperture to the end of the optical bench close to the photodiode detector and place the -10 D lens L_1 on the optical bench. Adjust the vertical and horizontal positions of the lens to make sure that the beam passes through the centre of the lens and that the image produced is concentric with the aperture. You can use the back reflection to verify this. Lens L_1 should now be aligned on the optical axis.

4. Place the $+4$ D lens L_2 just after L_1 and make sure that the beam passes through the centre of the lens. Obtain a collimated beam by sliding the $+4$ D lens L_1 up and down the optical bench. Collimated means that the beam diameter is independent of the distance from the lens, and we have a plane wavefront. Both lenses are now aligned on the optical axis, and the aperture may be removed.

5. Place one of the multiple slit objects in the position after lens L_2 so that a diffraction pattern is observed in the plane of the detector. The sample holder can be moved laterally and vertically to ensure that the slits are uniformly illuminated. The holder can also be rotated to ensure that the line of diffraction spots is horizontal, and lies parallel to the direction of motion for the detector scan.

4.2 Recording of Fraunhofer diffraction patterns

Two different sets of six diffracting objects are available; each set consists of a single rectangular slit and from two to six rectangular slits arranged in a regular linear array. The older set has the six diffracting objects mounted on separate slides. The newer set, which should be used if at all possible, has all six diffracting objects contained within a single glass slide, as shown in Figure 4.
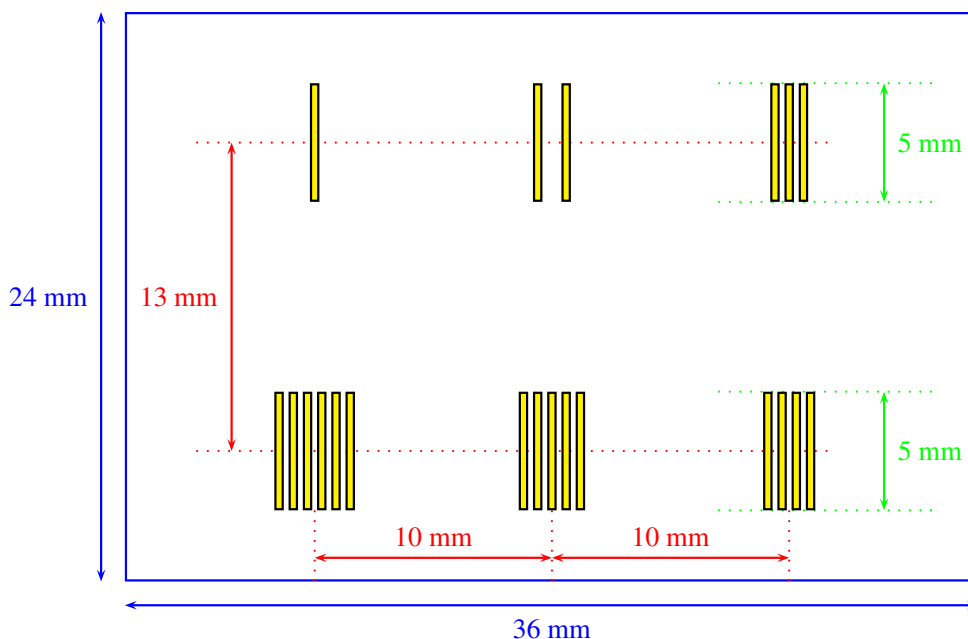

Figure 4: Diagram showing how all six diffracting objects are arranged on a single slide.

For each of the six diffracting objects, you must record a Fraunhofer diffraction pattern, and also make a direct measure of the slit size and spacing using a travelling microscope. The procedure for recording the patterns is the same in each case. Note that to be able to analyse the diffraction data you will need to know the distance from the diffracting object to the detector slit (distance L in Figure 1).

The detector slit is moved by a stepper motor that drives a lead screw with a pitch of 0.5 mm. The stepper motor controller can supply step pulses at rates up to 800 Hz, and 400 steps are needed to complete one revolution of the lead screw. Clearly the higher the rate that pulses are supplied, the faster the slit will scan cross the diffraction pattern. The detector signal is recorded digitally using a program on a PC.

4.2.1   Using a PC to collect the data digitally

It is assumed that you are familiar with the usual operations needed to run a program in the `Microsoft Windows` operating environment. If not, then consult a demonstrator. Make sure that the PC is running `Windows`, and double-click on the `Picoscope` icon to begin the program that emulates a cathode ray oscilloscope (*CRO*) display. Check that a coaxial signal cable connects the photodiode signal amplifier to a green BNC adapter connected to the parallel port of the PC.

Move the detector slit so that it is at the centre of the diffraction pattern, where the signal is highest. Adjust the sensitivity of the amplifier connected to the photodiode output so that the maximum signal lies in the range 0 to 5 V, and check that when you interrupt the laser beam (using a piece of paper or card, **not your hand**) the signal falls to a near zero value. The trigger setting on the CRO simulation should be set to *NONE*, so that the PC takes repeated samples of the data no matter what the signal level may be. Start with a time base setting of 5 s/div to 10 s/div so that a complete scan of the diffraction pattern will appear as a single trace on screen. The CRO trace can be started and stopped by pressing the SPACEBAR, or by clicking on the GO/STOP button at the bottom left of the `Picoscope` window. Note that after one sweep of the CRO trace over the 10 divisions of the time axis, the trace is cleared before a new one is started, so you **must manually STOP the CRO trace before it reaches the end of the sweep** if you wish to avoid losing any data that is recorded during the sweep.

Next move to the edge of the diffraction pattern and scan the detector through the whole diffraction pattern using a suitable time base settings and stepper-motor speeds. Note that for the multiple-slit patterns the detector may need to be scanned more slowly than for the single-slit pattern to ensure that the fine details of the interference term are recorded accurately.

Data collected during a sweep can be saved using the Menu option `File/Save As...` and then choosing to save the data as text. The file created will contain two columns of data: one with the time-axis value in milliseconds, the other containing the signal reading in volts. It is straightforward to import these data into other programs such as `Excel` if they are to be used for further processing, or for the preparation of reports.

Always ensure that you note the stepper-motor settings for each diffraction pattern that you record, as this information will be needed in the subsequent analysis of the data.

4.3   Direct measurements with a travelling microscope

As in any experiment, it is important to take steps to confirm the experimental measurements you make. Use one of the travelling microscopes that are available in the laboratory to make direct measurements of the slit widths and slit separations for each of the diffracting objects.

5   Data analysis

The analysis of the data should allow you to determine the slit width a, slit spacing d, and number of slits N from the diffraction patterns recorded.

The first step is to convert the time axis of each of the recorded patterns into an angular scale. If the stepper motor is being driven by pulses at a rate of r Hz, then in a time t the detector slit will move a distance $x = \frac{rt}{800}$ mm. The time variable t (in seconds) can be converted directly to an angle θ (in radians) using the relation

$$\theta = \frac{x}{L} = \frac{rt}{800L} \tag{6}$$

assuming the distance L is also measured in millimetres.

5.1   Calculate the slit widths a

Use the fact that the intensity minima of the diffraction envelope occur at angles given by equation (2) to calculate values for the slit width for each diffracting object. Combining equation (2) and (6) yields

$$a = \frac{n\lambda 800 L}{rt} \tag{7}$$

5.2   Calculate the slit spacings d

Measure the average angular separation of consecutive principal maxima for the multiple-slit patterns, and use equation (4) to obtain a value for the slit separation for each of the multi-slit objects.

5.3   Determine the number of slits in each object

Examine the number of intensity minima and secondary maxima between consecutive principal maxima in the multi-slit diffraction patterns, and deduce from this the number of slits present in the diffracting object.

5.4   Compare diffraction data with direct microscopic measurements

Compare the values found in sections 5.1, 5.2 and 5.3, with the direct measurements made by optical microscope in section 4.3. Do the results agree within experimental error?

Comment on the main sources of error for both the diffraction and the direct measurements of the slits, and identify the factors that would influence you if you had to choose between one method and the other.

5.5   Calculate theoretical Fraunhofer diffraction patterns

For each of the measured values of a and d it is possible to calculate the theoretical Fraunhofer diffraction pattern by using the `Maple V` program `fraud1.m` that can be found on the Physics PC network, in folder `P:\docs\courses\cp2110\`.

Make a careful and critical comparison between the form of the measured diffraction patterns and those predicted by the theoretical calculations, paying particular attention to the positions and relative heights of the intensity maxima and minima. Comment on any differences that you find.

References

[1] Wikipedia. `http://en.wikipedia.org/wiki/Joseph_von_Fraunhofer`.

[2] Graeme Morrison. Optics.
`http://www.kcl.ac.uk/kis/schools/phys_eng/physics/courses/CourseList/CP3380.htm`.

[3] E Hecht. *Optics*. Addison Wesley, 4th edition, 2002.

[4] M. Born and E. Wolf. *Principles of Optics*. CUP, 7th edition, 1999.

Fresnel Diffraction at a Slit

1 Aim of the Experiment

In this experiment a diverging wave-front from a coherent source is diffracted by a slit of variable width. The experimentally observed diffraction patterns obtained from several different slit widths are recorded quantitatively, by scanning the pattern across the entrance slit of a photodiode detector, and these are compared with the theoretical intensity variations calculated using the Fresnel diffraction integral.
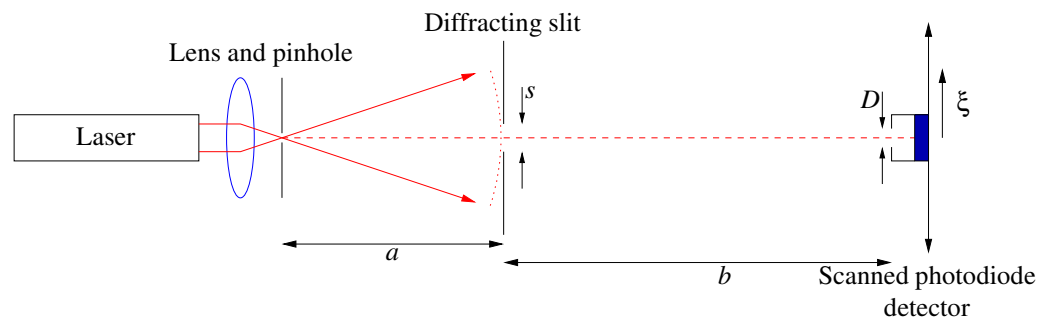
2 Apparatus

Figure 1: Schematic showing experimental set-up for Fresnel diffraction.

The equipment used in this experiment is shown schematically in Figure 1. The illumination is provided by a 2 mW He/Ne laser providing light of wavelength 632.8 nm.

> **Warning: Never look directly at the laser source, nor at a specular reflection of the light! Read carefully the section on laser safety in the safety notes issued with this folder. DO NOT SWITCH ON THE LASER unless you are sure that you understand the hazard. If in doubt, consult a demonstrator.**
>
> **Make sure you have identified all risks involved in the experiment and comment on them in your write-up.**

Before attempting to obtain a diffraction pattern the laser should be set up to provide a beam parallel to the axis of the optical bench. A diverging beam of laser light is then produced by placing a very short focal length lens (such as a microscope objective lens) in front of the laser, and a precision pinhole should be placed in the focal plane of this lens to act both as a spatial filter and, effectively, as a point source. A variable width slit can now be placed on the optical bench and a diffraction pattern observed on a screen beyond it; the actual pattern obtained will depend on the width s of the slit.

Quantitative intensity measurements of the diffracted intensity are made by removing the screen and scanning a narrow slit perpendicular to the optical axis of the system. Behind this slit is a solid-state photodiode that produces a voltage output proportional to the incident intensity. There is also a filter in front of the photodiode that allows only the red light from the laser to pass through, so that the experiment can be carried out under normal lighting conditions. The detector slit should be aligned so as to be parallel to the diffraction slit, and its width should be chosen so as to be smaller than any feature in the pattern being observed. The output from the photodiode is measured by directing the voltage to an analogue-to-digital converter (ADC) attached to a PC, and using suitable software to collect a time series of data as the detector slit is scanned across the diffraction pattern.

3 Theoretical Background

The type of diffraction pattern observed as a coherent wave from source A passes through an aperture depends on the phase of the waves that recombine at the observation point P. Figure 2 shows that the phase of the waves combining at P will depend on the optical path the light takes on travelling from A to P. In general, the phase of the waves reaching P can be expressed as a series of the form:

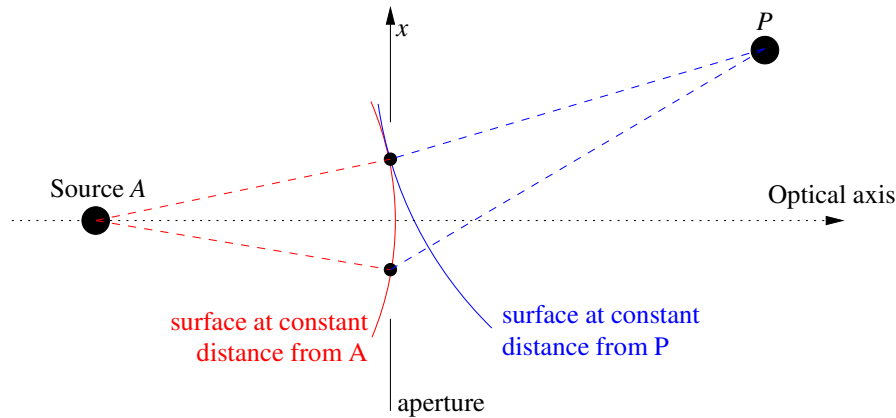$$C_0 + C_1 x + C_2 x^2 + C_3 x^3 + \dots \tag{1}$$

Figure 2: Schematic of the light path.

If the phase is a linear function of the coordinate x in the aperture plane, so that the wavefronts in the aperture plane can be treated as plane waves, and the second-order and higher-order terms with coefficients C_2, C_3, ... are all negligibly small, then it is known as Fraunhofer (far-field) diffraction. If the phase depends quadratically on x, then the second-order term cannot be neglected, and Fresnel (near-field) diffraction is said to take place. **Fresnel diffraction** thus corresponds to the case **when the curvature of the wavefronts at the aperture plane cannot be neglected**.

Augustin-Jean Fresnel (1788–1827) was a physicist who worked on diffraction and other problems in optics, and also made some contributions to mathematics [1]. The experiment is central to optics, and will be extensively discussed and extended in Dr Morrison's 3rd year course CP3380 *Optics* [2].

The total amplitude at P is given by the complex sum of the contributions from each elemental area of the slit, taking into account the phase differences that arise from the different paths taken from A to P. For the case of Fresnel diffraction from a single slit, the diffracted amplitude at P can be calculated by evaluating two integrals, the Fresnel Cosine integral and Fresnel Sine integral, with limits of integration that are determined by the geometry of the experiment.

The **Fresnel cosine integral** is $C(X_1) = \displaystyle\int_0^{X_1} \cos\frac{\pi X^2}{2}\, dX$,

while the **Fresnel sine integral** is $S(X_1) = \displaystyle\int_0^{X_1} \sin\frac{\pi X^2}{2}\, dX$,

and both of these are standard, tabulated functions [2, 3]. They are also available directly from within the `Maple` program available on the Physics PC network. The **Cornu spiral** is a graphical representation of these functions that can be useful when dealing with Fresnel diffraction. The Cornu spiral is the curve that is produced when the value of the Fresnel sine integral $S(X)$ is plotted against the Fresnel cosine integral $C(X)$ for all $X \in \,]-\infty;\infty[$.

The functions $C(X)$ and $S(X)$ can be thought of as the parametric equations of the Cornu spiral (Fig. 3), in terms of the parameter X. Both $C(X)$ and $S(X)$ are odd functions of X, and as $X \to \pm\infty$ both $C(X)$ and $S(X)$ tend to $\pm\frac{1}{2}$. The spiral has the rather surprising property that the length of the arc along the spiral from the origin to the point $(C(X_1), S(X_1))$ is just given by the value X_1 [3].

A real slit of width s corresponds to an arc of length $X_2 - X_1 = X = \left(\sqrt{\dfrac{2\,(a+b)}{\lambda ab}}\right) s$ along the spiral, where

a and b are the distances shown in Figure 1. It can be shown that the diffracted intensity at any given point in
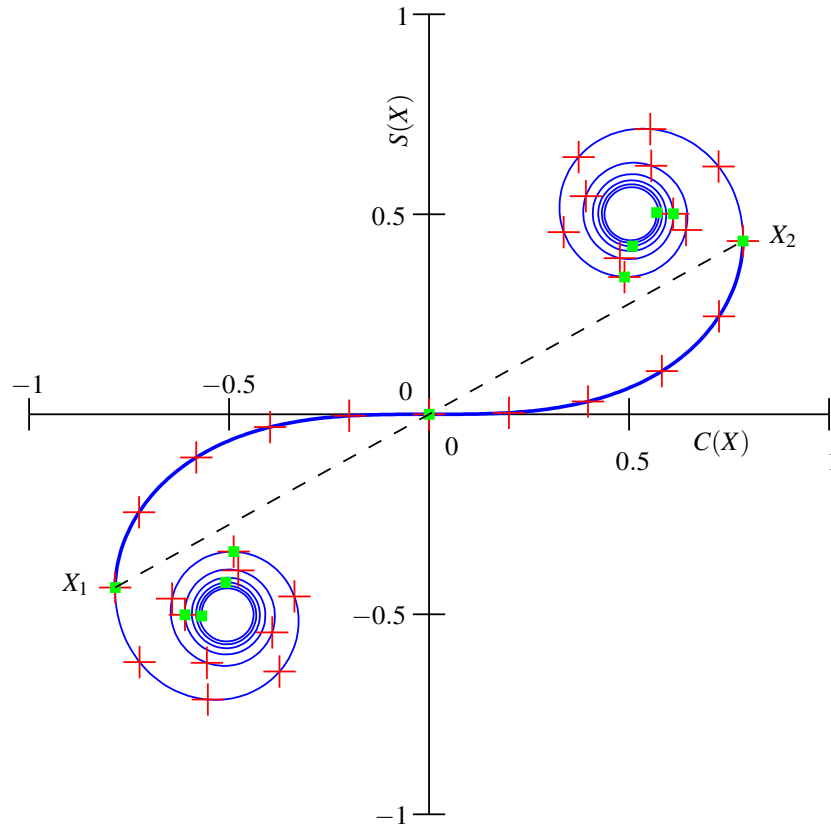
Figure 3: The Cornu spiral, with the arc-length from the origin marked off in odd multiples of 0.1 (small crosses), and in multiples of 1.0 (squares). The dashed line joins the end points of an arc from X_1 to X_2.

the Fresnel diffraction pattern is proportional to the square of the distance (along a straight line) between the end points of the arc of length $X_2 - X_1$. The intensity at different points is found by moving the end-points X_1 and X_2, while still keeping the arc length X constant, and again measuring the straight-line distance between the end-points.

4   Experimental Procedure

4.1   Alignment of the optical system

The apparatus shown in Figure 1 must be carefully aligned if the best results are to be obtained. Assuming that it is completely out of adjustment the procedure is as follows:

1. Remove all components from the optical bench except the laser and a small diaphragm aperture.

2. Sequentially adjust the tilting screws of the laser mount to ensure that the laser beam passes through the diaphragm aperture for all positions of the aperture on the optical bench, and that the beam is at the same height as the detector slit. *The laser beam is now parallel to the optical bench.*

3. Move the diaphragm aperture to the end of the optical bench close to the photodiode detector and place a short focal length converging lens (a microscope objective lens) on the optical bench. Adjust the vertical and horizontal positions of the lens to make sure that the image produced is concentric with the aperture. *The converging lens should now be aligned on the optical axis.*

4. Place the precision pinhole in the mount directly after the converging lens, and carefully adjust the position of the pinhole until it is accurately on the optical axis, and it lies in the focal plane of the converging lens. In this position, a bright, almost uniform disc of light should be seen on a screen placed downstream of the pinhole. *The pinhole is now acting effectively as a source of a diverging spherical waves.*

5. Place the variable-width slit about 100 to 200 mm downstream of the pinhole, and about 600 to 800 mm upstream of the photodiode detector. The slit holder should be moved laterally and vertically to ensure that the slit is uniformly illuminated. The holder can also be rotated to ensure that the line of the slit is accurately perpendicular to the direction of motion followed by the detector slit as it scans the diffraction pattern.

6. Place a ground glass screen or a white card in front of the detector slit, and observe the diffraction pattern from the slit. Vary the width of the diffracting slit, and confirm that the character of the diffraction pattern (the number of bright and dark fringes) changes as the width changes.

4.2   Recording of Fresnel diffraction patterns

A set of five or six perceptibly different diffraction patterns should be recorded, corresponding to different widths of the diffracting slit. For each one you must also measure the slit width *s* by means of a travelling microscope. *Note that to be able to analyse the diffraction data you will need to know the distances from the source to the diffracting slit, and from the diffracting slit to the detector slit (distances a and b in Figure 1).* The procedure for recording the patterns is the same in each case.

The detector slit is moved by a stepper motor that drives a lead screw with a pitch of 0.5 mm. The stepper motor controller can supply step pulses at rates up to 400 Hz, and 200 steps are needed to complete one revolution of the lead screw. Clearly the higher the rate that pulses are supplied, the faster the slit will scan cross the diffraction pattern. The detector signal is recorded digitally using a program on a PC.

4.2.1   Using a PC to collect the data digitally

Make sure that the PC is running `Windows`, and double-click on the `Picoscope` icon to begin the program that emulates a cathode ray oscilloscope (*CRO*) display. Check that a coaxial signal cable connects the photodiode signal amplifier to a green BNC adapter connected to the parallel port of the PC.

Move the detector slit so that it is at the centre of the diffraction pattern, where the signal is highest. Adjust the sensitivity of the amplifier connected to the photodiode output so that the maximum signal lies in the range 0 to 5 V, and check that when you interrupt the laser beam (using a piece of paper or card, **not your hand**) the signal falls to a near zero value. The trigger setting on the CRO simulation should be set to *NONE*, so that the PC takes repeated samples of the data no matter what the signal level may be. Start with a timebase setting of 5 s/div to 10 s/div so that a complete scan of the diffraction pattern will appear as a single trace on screen. The CRO trace can be started and stopped by pressing the SPACEBAR, or by clicking on the GO/STOP button at the bottom left of the `Picoscope` window. Note that after one sweep of the CRO trace over the 10 divisions of the time axis, the trace is cleared before a new one is started, so **you must manually STOP the CRO trace before it reaches the end of the sweep** if you wish to avoid losing any data that is recorded during the sweep.

Next move to the edge of the diffraction pattern and scan the detector through the whole diffraction pattern using a suitable timebase setting and stepper-motor speed. Note that when the diffracting slit is relatively wide, the detector may need to be scanned more slowly than for a narrow slit to ensure that the fine details of the diffraction pattern are recorded accurately.

Data collected during a sweep can be saved using the Menu option `File/Save As...` and then choosing to save the data as text. The file created will contain two columns of data: one with the time-axis value in milliseconds, the other containing the signal reading in volts. It is straightforward to import these data into other programs such as `Excel` if they are to be used for further processing, or for the preparation of reports. If you wish to obtain a hard copy of the trace on screen, then this can be achieved easily using the `File/Print` menu option.

> **Always ensure that you note the stepper-motor settings for each diffraction pattern that you record, as this information may be needed in the subsequent analysis of the data.**

4.2.2   Direct measurements with a travelling microscope

As in any experiment, it is important to take steps to confirm the experimental measurements you make. Use one of the travelling microscopes that are available in the laboratory to make direct measurements of the slit widths.

For each diffraction pattern that you record, measure the diffraction slit width *s* using a travelling microscope, and measure the distances from source to slit *a* and from slit to detector *b*. From these data the size $\Delta\xi$ (the

geometrical shadow of the diffracting slit on the detector plane) can be calculated. By using similar triangles it is clear that:

$$\Delta\xi = \frac{(a+b)s}{a} \tag{2}$$

As a check, compare $\Delta\xi$ with the size of the detector slit D, which can also be measured using the travelling microscope. (Make sure you understand why $\Delta\xi$ should be considerably greater than D which is approximately $300\,\mu$m and comment on this in your write-up.)

5   Data Analysis

5.1   Calculate theoretical Fresnel diffraction patterns

For each of the measured slit widths s, calculate the corresponding dimensionless arc-length

$$X = s\left(\sqrt{\frac{2(a+b)}{\lambda ab}}\right) \tag{3}$$

along the Cornu spiral. Use the `Maple V` program `fresd1.ms` that can be found on the Physics PC network, in folder `P:\docs\courses\cp2110\`, to calculate the Fresnel diffraction patterns for these values of X. V in the program is the dimensionless arc-length X.

5.2   Comparison of theory with experiment

Mark on the experimental plots the position of the geometric shadow of the diffracting slit. Make a careful and critical comparison between the form of the measured diffraction patterns and those predicted by the theoretical calculations, paying particular attention to the positions and relative heights of the intensity maxima.

5.3   Calculation of experimental errors

Each of the measured quantities a, b and s has an experimental error associated with it, so that there is an uncertainty in the value of X used to calculate the theoretical patterns. By noting that the factor $\frac{a+b}{ab} = \frac{1}{a} + \frac{1}{b}$, use the standard expressions for the propagation of experimental error to show that the uncertainty in X is given by the expression:

$$\frac{\Delta X}{X} = \sqrt{\frac{1}{4(a+b)^2}\left[\left(\frac{b\Delta a}{a}\right)^2 + \left(\frac{a\Delta b}{b}\right)^2\right] + \left(\frac{\Delta s}{s}\right)^2} \tag{4}$$

where the experimental errors are indicated by the prefix Δ. Identify which of the measured experimental errors, Δa, Δb, or Δs, makes the most important contribution to the error ΔX, and comment on whether or not the experimental traces agree with the theoretical plots within experimental error. If they do not agree, suggest other ways in which the discrepancy could be investigated further.

References

[1] Wikipedia. http://en.wikipedia.org/wiki/Fresnel.

[2] Graeme Morrison. Optics.
    http://www.kcl.ac.uk/kis/schools/phys_eng/physics/courses/CourseList/CP3380.htm.

[3] E. Hecht. *Optics*. Addison Wesley, 4th edition, 1987.

The Fourier Transform in Optics

1 Aim of the Experiment

The purpose of this experiment is to illustrate how diffraction, imaging and spatial filtering can be explained and understood using the theoretical framework of Fourier transforms.

2 Context and Background

2.1 The Fourier Transform

Many practical applications of diffraction, not only in optics, but also, for example, in X-ray crystallography, use the power and simplicity of the Fourier transform. Physically, the Fourier transform of a function gives you a quantitative idea of the frequency content of a function. The variable u has dimensions that are reciprocal to those of the variable x. Thus, if x represents time, in seconds, then u represents frequency, in Hertz. Likewise, if x represents length, in meters, then u represents spatial frequency (reciprocal length), in cycles per meter. This physical interpretation of the Fourier transform accounts, in part, for its usefulness in optics and other fields where diffraction phenomena occur. Note that the process of diffraction performs a Fourier transform operation. Fourier analysis and synthesis is very useful in one-dimensional signal processing, e.g. in sounds systems, radio etc. A second reason for its usefulness is that it has many convenient mathematical properties that you are about to explore.

A Fourier transform, $F(u)$, of a function $f(x)$ is defined by the following integral

$$F(u) = \int\limits_{-\infty}^{\infty} f(x)e^{-2\pi iux}\,dx \tag{1}$$

In general, $F(u)$ will be complex, even if $f(x)$ is real.

The inverse Fourier transform is defined by

$$f(x) = \int\limits_{-\infty}^{\infty} F(u)e^{2\pi iux}\,du \tag{2}$$

Such that

$$f(x) = \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{\infty} f(x')e^{-2\pi iux'}\,dx' \right) e^{2\pi iux}\,du \tag{3}$$

Thus, the inverse transform of the transform is the original function $f(x)$. Throughout these notes, the upper case notation will signify a Fourier transform, whereas the lower case notation will refer to the original function. An example of some one-dimensional Fourier transform pairs are in shown in Fig. 2.1. Note that equations (1) and (2) are definitions, they are not derived from other principles: they have to be learnt as facts. They are named after mathematician and physicist Joseph Fourier (1768–1830) who developed these ideas [1, 2]. Fourier transforms are discussed in more detail in Dr Kantorovich's 2nd year course CP2260 *"Mathematical Methods in Physics II"* [3].

The function $f(x)$ must satisfy certain mathematical conditions which are always satisfied by functions that exist physically. However, there are a number of functions which cannot exist physically, such as the function $f(x) = 1$ for all values of x, for which it would be convenient to have a Fourier transform.
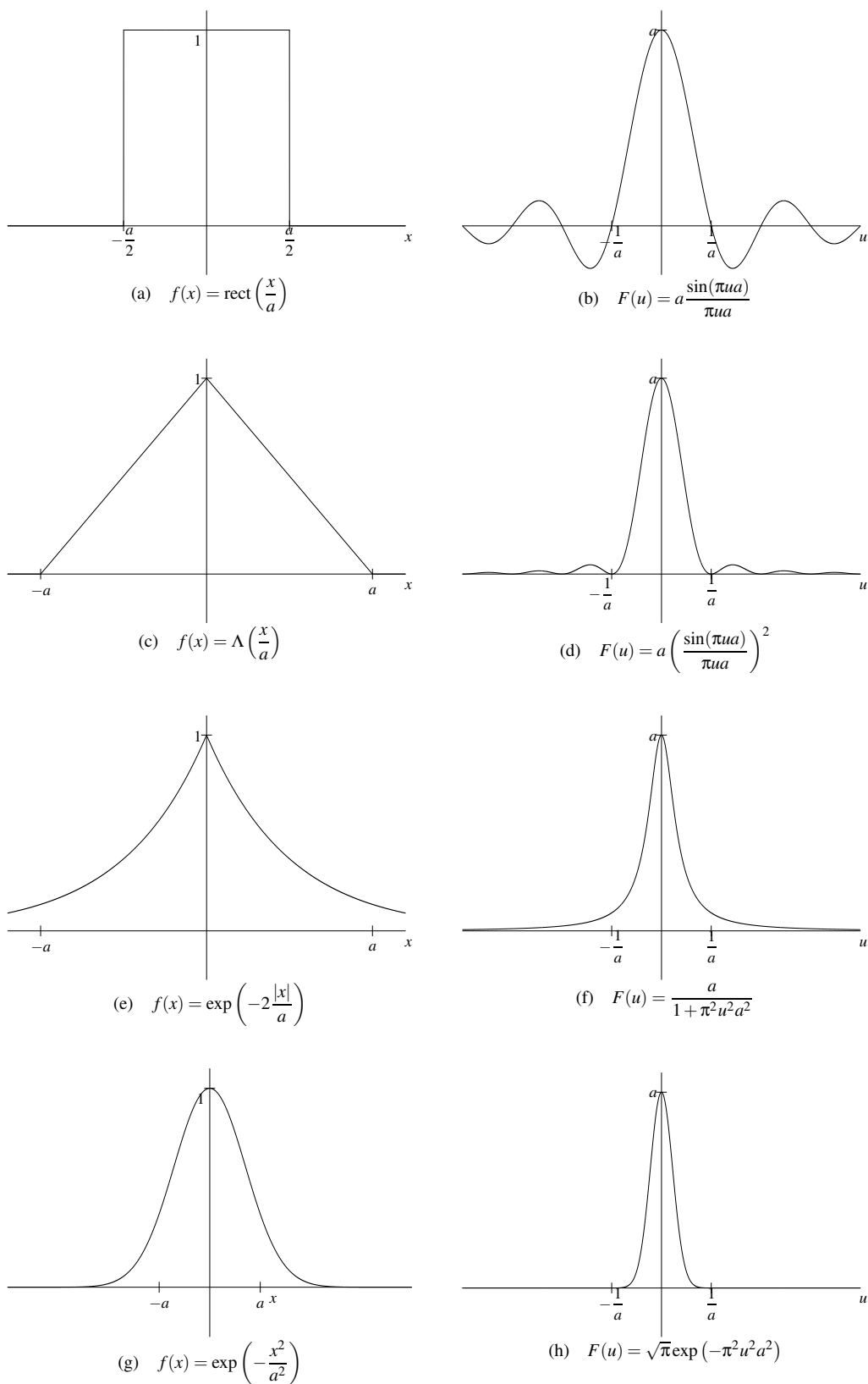
(a) $f(x) = \text{rect}\left(\dfrac{x}{a}\right)$

(b) $F(u) = a\dfrac{\sin(\pi u a)}{\pi u a}$

(c) $f(x) = \Lambda\left(\dfrac{x}{a}\right)$

(d) $F(u) = a\left(\dfrac{\sin(\pi u a)}{\pi u a}\right)^2$

(e) $f(x) = \exp\left(-2\dfrac{|x|}{a}\right)$

(f) $F(u) = \dfrac{a}{1 + \pi^2 u^2 a^2}$

(g) $f(x) = \exp\left(-\dfrac{x^2}{a^2}\right)$

(h) $F(u) = \sqrt{\pi}\exp\left(-\pi^2 u^2 a^2\right)$

Figure 1: One-dimensional Fourier transform pairs.

Consider function $f(x) = 1$ for all x. Clearly,

$$\int_{-\infty}^{\infty} |f(x)| \, dx = \infty \tag{4}$$

and, strictly speaking, the Fourier transform does not exist. However, this function could equally well be defined as

$$f(x) = \lim_{a \to \infty} \text{rect} \left(\frac{x}{a} \right) \tag{5}$$

And thus its Fourier transform is

$$F(u) = \lim_{a \to \infty} a \frac{\sin(\pi u a)}{\pi u a} \tag{6}$$

The function $a \frac{\sin(\pi u a)}{\pi u a}$ is shown in Fig. 1(b). In the limit of $a \to \infty$, this function equals the Dirac delta function $\delta(u)$, which is defined by the two properties

$$\delta(u) = 0 \text{ for } u \neq 0 \tag{7}$$

$$\int_{-\infty}^{\infty} \delta(u) \, du = 1 \tag{8}$$

(Paul Dirac (1902–1984), physics Nobel laureate 1933, was a physicist who worked on mathematical and theoretical quantum mechanics [4].)

We refer to $f(x) = 1$ and $F(u) = \delta(u)$ as a generalised Fourier transform pair.

Given the fact that the generalised Fourier transform of $f(x) = 1$ is a delta function, it is straightforward to find the generalised transforms of sine and cosine (neither of which strictly have a normal Fourier transform). Fig. 2.1 shows some generalised Fourier transform pairs.

2.2 Properties of the Fourier Transform

The Fourier transform has a number of useful properties. Here are five of the most useful ones, frequently encountered in optics:

2.2.1 Symmetries

$f(x)$	$F(u)$
even	odd
odd	odd
real	Hermitian[*]
real an even	real and even

[*]Hermitian means that $F(u) = F^*(-u)$, i.e. the real part is even and the imaginary part is odd.

2.2.2 Addition Theorem

If $f(x)$ has the Fourier transform $F(u)$, and $g(x)$ has the Fourier transform $G(u)$, then $f(x) + g(x)$ has the Fourier transform $F(u) + G(u)$.

2.2.3 Similarity Theorem

If $f(x)$ has the Fourier transform $F(u)$, then $f(ax)$, where a is real and positive, has the Fourier transform $\frac{1}{a} F(\frac{u}{a})$.

Since $f(ax)$, $a > 1$ is simply a version of $f(x)$ that is squeezed along the x-axis, $\frac{1}{a} F(\frac{u}{a})$ is simply $F(u)$ stretched along the u-axis and scaled to maintain a constant area under it.

(a) $f(x) = 1$
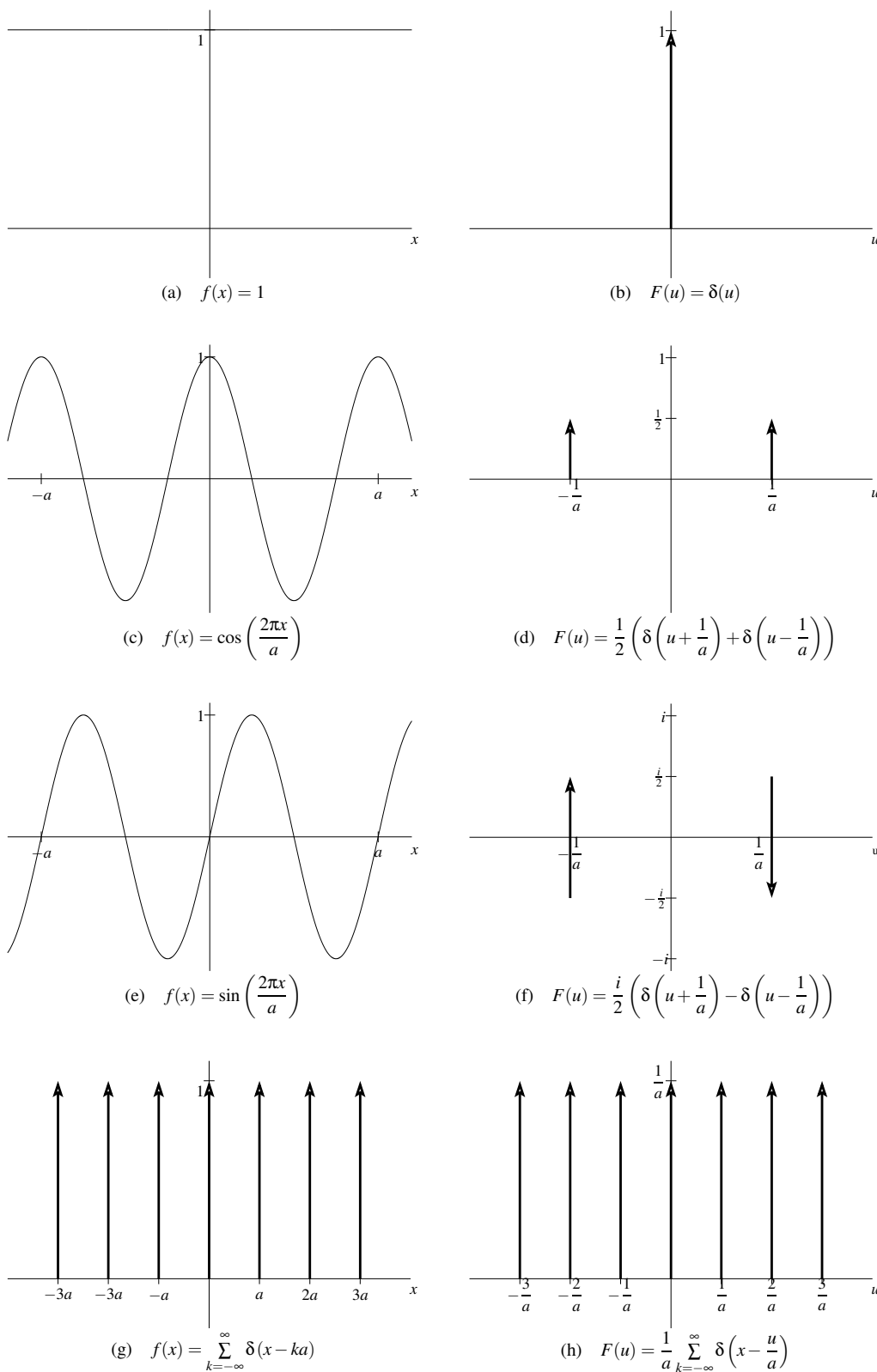
(b) $F(u) = \delta(u)$

(c) $f(x) = \cos\left(\dfrac{2\pi x}{a}\right)$

(d) $F(u) = \dfrac{1}{2}\left(\delta\left(u + \dfrac{1}{a}\right) + \delta\left(u - \dfrac{1}{a}\right)\right)$

(e) $f(x) = \sin\left(\dfrac{2\pi x}{a}\right)$

(f) $F(u) = \dfrac{i}{2}\left(\delta\left(u + \dfrac{1}{a}\right) - \delta\left(u - \dfrac{1}{a}\right)\right)$

(g) $f(x) = \sum\limits_{k=-\infty}^{\infty} \delta(x - ka)$

(h) $F(u) = \dfrac{1}{a}\sum\limits_{k=-\infty}^{\infty} \delta\left(x - \dfrac{u}{a}\right)$

Figure 2: Some generalised one-dimensional Fourier transform pairs.

2.2.4 Shift Theorem

If $f(x)$ has the Fourier transform $F(u)$, then $f(x-a)$ has the Fourier transform $e^{-2\pi iua}F(u)$,
and $e^{2\pi ixa}f(x)$ has the Fourier transform $F(u-a)$.

Note that $f(x-a)$ is simply $f(x)$ shifted to the right so that its origin lies at $x = a$. Shifting $f(x)$ does not change the modulus of the Fourier transform—a useful property in diffraction problems.

2.2.5 Convolution Theorem

Before we state and prove this theorem, we shall first consider the convolution integral. The convolution of two functions $f(x)$ and $g(x)$ yields a third function $h(x)$ defined by

$$h(x') = \int\limits_{-\infty}^{\infty} f(x)g(x-x')\,dx \tag{9}$$
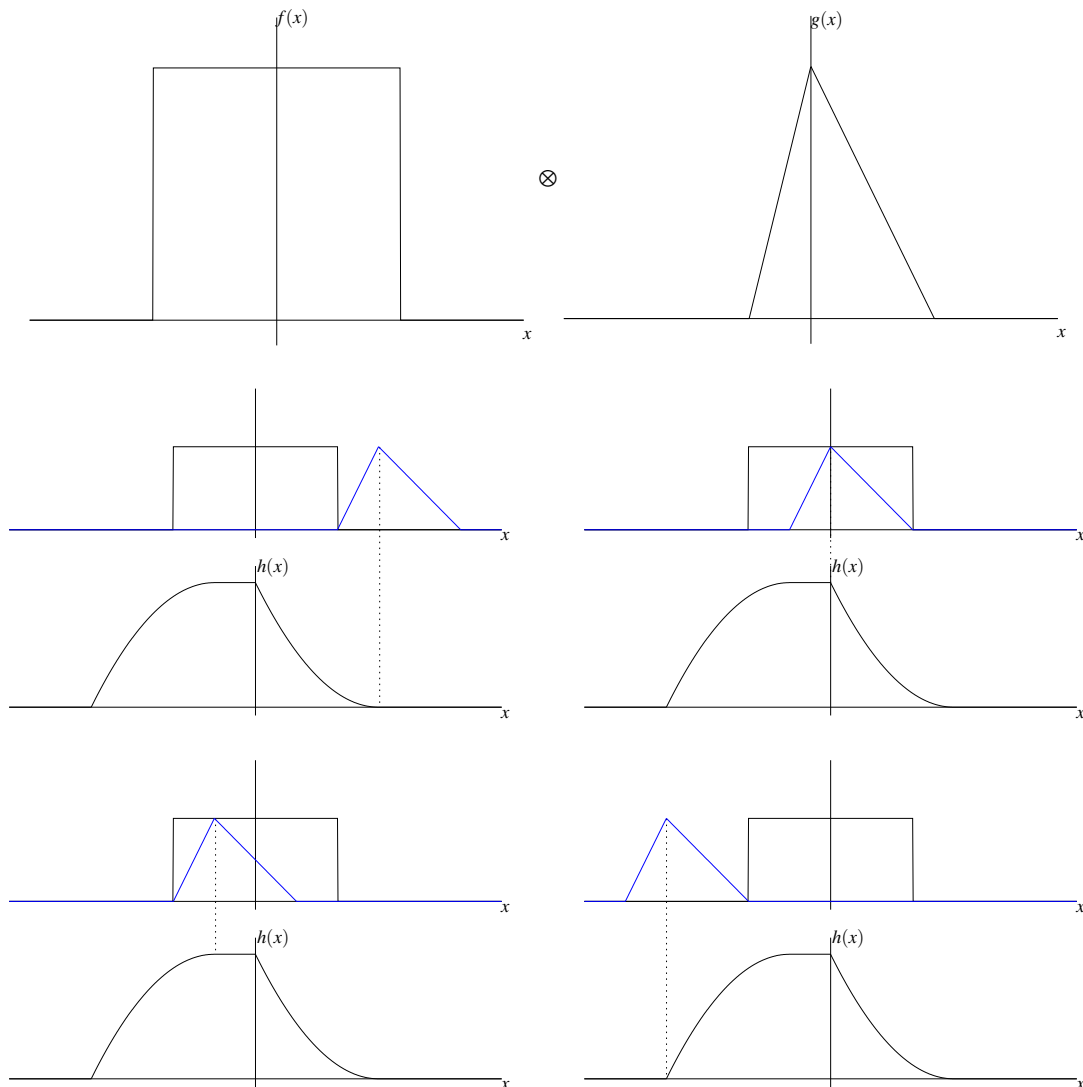

Figure 3: The convolution of two functions, $f(x)$ and $g(x)$, yields a new function $h(x)$.

Integrals of this form frequently occur in optics. For example, the intensity of an image equals the convolution of the object intensity and the so-called intensity point-spread function of the imaging system.

By variable substitution in (9), $x_1 = x - x'$

$$h(x') = \int_{-\infty}^{\infty} f(x' - x_1) g(x_1) \, dx_1 \tag{10}$$

giving a certain symmetry to the integral. The convolution integral is often abbreviated by the $\otimes$ notation

$$h(x) = f(x) \otimes g(x) \tag{11}$$

Physically, convolution can be thought of as the blurring or smearing of one function by another, as shown in Fig. 3.

The convolution theorem states:

The Fourier transform of $f(x) \otimes g(x)$ equals $F(u)G(u)$.

i.e. the Fourier transform of a convolution equals the product of the transforms.

Questions: Answer these questions now, in your lab book.

1. Prove the addition theorem

2. Prove the shift theorem

3. Prove the similarity theorem

4. Find the Fourier transform of the function $f(x) = \text{rect}\left(\frac{x}{a}\right)$, i.e. $f(x) = \begin{cases} 1 & \text{for } |x| < a/2 \\ 0 & \text{elsewhere} \end{cases}$.

   Show the calculation and sketch $f(x)$ and $F(u)$.
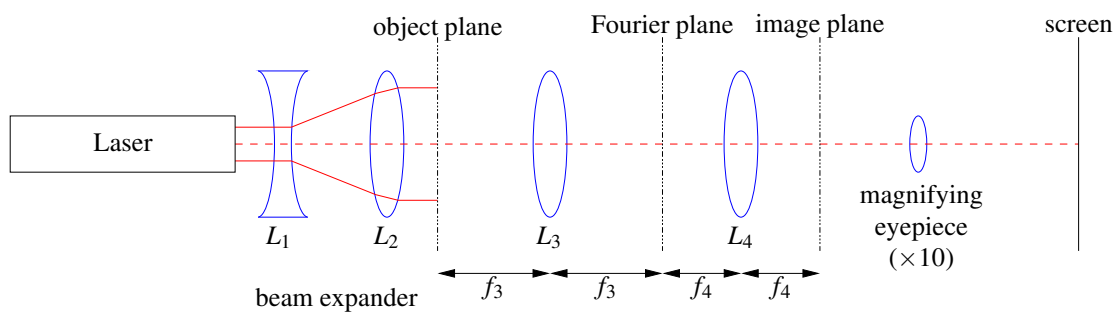
3 Experimental set up

Figure 4: Schematic set-up of the experiment. L_i are lenses, with f_i their corresponding focal length.

The equipment used in this experiment is shown schematically in Figure 4. The illumination is provided by a 2 mW He/Ne laser providing light of wavelength 632.8 nm.

Warning: never look directly at the laser source or at a specular reflection of the light! Read carefully the section on laser safety in the safety notes issued at the back of the lab manual. DO NOT SWITCH ON THE LASER unless you are sure that you understand the hazard. If in doubt, consult a demonstrator.

Make sure you have identified all risks involved in the experiment and comment on them in your write-up.

Handle all components carefully. Never touch glass surfaces with your fingers.

3.1   Alignment of the optical system

The apparatus shown in Figure 4 must be carefully aligned if the best results are to be obtained. Assuming that it is completely out of adjustment the procedure is as follows:

- Remove all components from the optical bench except the laser and a small diaphragm aperture.

- Sequentially adjust the tilting screws on the laser and the height of the pinhole to ensure that the laser beam passes through the aperture for all positions of the aperture on the optical bench. *The laser beam is now parallel to the optical bench.*

  The laser beam is diverged somewhat by concave lens L_1, placed just in front of the laser, and is then collimated by lens L_2, a convex lens. This pair of lenses acts as a beam expander.

- Move the diaphragm aperture to the end of the optical bench and place concave lens L_1 on the optical bench. Adjust the vertical and horizontal positions of the lens to make sure that the beam passes through the centre of the lens and that the image produced is concentric with the aperture. You can use the backreflection to verify this. *Lens L_1 should now be aligned on the optical axis.*

- Place convex lens L_2 just after L_1 and make sure that the beam passes through the centre of the lens. Obtain a collimated beam by sliding L_2 up and down the optical bench. Collimated means that the beam diameter is independent of the distance from the lens, and we have a plane wavefront, i.e. the beam is parallel, neither diverging nor converging. *Both lenses are now aligned on the optical axis, and the aperture may be removed.*

Insert the remaining lenses into the beampath following the same principle.

3.2   Diffraction

Insert an adjustable slit into the object plane (focal plane of L_3) and observe the diffraction pattern on a screen. Change the slit width. What do you see? Is this what you expect from Fourier theory? Which theorem explains your observation? Make sure you note down all your observations.

Place a card or paper into the Fourier plane and carefully draw the diffraction pattern of your chosen slit width. Assuming the slit to be the function $f(x) = \text{rect}\left(\frac{x}{a}\right)$, use your calculation in the previous section to find the slit width. Remember the distance in the Fourier plane is equivalent to spatial frequency, but (taking into account the lens) it has to be scaled by a factor $\frac{1}{\lambda f}$, i.e. $u = \frac{x}{\lambda f}$, where λ is the wavelength of the light, and f is the focal length of the lens. Now place a fine glass grating into the object plane, and observe the diffraction pattern. Compare it to that of a coarse grating. What observation do you expect from Fourier theory?

A diffraction grating with period b can be expressed as a Fourier series:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi nx}{b}\right) \tag{12}$$

where

$$a_0 = \frac{2}{b} + \int_{-\frac{b}{2}}^{\frac{b}{2}} f(x)\,dx \tag{13}$$

and

$$a_n = \int_{-\frac{b}{2}}^{\frac{b}{2}} f(x)\cos\left(\frac{2\pi nx}{b}\right) dx \tag{14}$$

Using the information on the Fourier transform (FT) of a cos function on Page 90 of this experiment it can be found that maxima of the FT occur at $0, +\frac{1}{b}, +\frac{3}{b}, +\frac{5}{b}, \dots$

Try to familiarize yourself with this derivation. If you have difficulties speak to your demonstrator or use the Hecht Optics book [5].

Using the information for where the maxima occur and by again drawing the diffraction pattern in the Fourier plane, calculate the spacing of the grating. The distance in the Fourier plane is equivalent to spatial frequency, but (taking into account the lens) it has to be scaled by a factor $\frac{1}{\lambda f}$, i.e. $u = \frac{x}{\lambda f}$.

Move the grating sideways. What do you observe in the Fourier plane? Which theorem explains your observation?

3.3 Imaging

The process of diffraction performs a Fourier transform operation. With a diffracting screen in the front focal plane of a lens, the complex amplitude in the back focal plane is the Fourier transform of that in the front focal plane. Consider the arrangement shown in Fig. 4. With an object placed in the object plane, the Fourier transform is produced in the Fourier plane. The next lens, L_4, performs a second Fourier transform, so that in the image plane we see the transform of the transform of the object.

Question: If the object is $f(x)$, show that the transform of the transform is $f(-x)$, i.e. the image is an inverted version of the object.

Place the slide with the letter "A" in the object plane. What do you observe in the image plane? Make sure you note all your observations in your write-up.

3.4 Spatial Filtering

By introducing a mask into the Fourier plane, various spatial frequencies of the object may be selected, rejected or attenuated. The image thus becomes a modified version of the original object. The deliberate distortion of the diffraction pattern is called spatial frequency filtering, and this is one form of optical image processing.

Place the grating in the object plane. Using a mask in the Fourier plane, construct each of the following patterns in the image plane (Hint: Think about the spacing between the maxima in the Fourier plane):

1. a field of roughly uniform intensity

2. a fringe pattern the period of which is

    (a) $\frac{1}{2}$ period of the grating
    (b) $\frac{1}{3}$ period of the grating
    (c) $\frac{1}{4}$ period of the grating

Explain qualitatively how you achieve these patterns. Note that it is the intensity you observe, not the complex amplitude.

Place the slide with the circle into the object plane. Using spatial filtering, generate an image that is a square, not a circle.

Observe the diffraction pattern of an object, e.g. a slide with 3 slits or the letter "A" placed over a grating. Convince yourself that the element and grating are multiplied together, and hence explain the appearance of the diffraction pattern. Discuss the relative extent of the spatial frequency bandwidth of the object compared to that of the grating. Using this information, perform a spatial filtering operation that will form an image of the object without the grating visible.

Place the slide with the portrait of Fourier into the object plane. Using spatial filtering again, remove the dots from Fourier's face.

The convolution of a function with a Dirac comb yields a string of replicas of the function. How can you generate a series of replicas of the object in the image plane? What determines the spacing of the replicas?

References

[1] John J. O'Connor and Edmund F Robertson.
    `http://www-history.mcs.st-andrews.ac.uk/Mathematicians/Fourier.html`.

[2] Wikipedia. `http://en.wikipedia.org/wiki/Jean_Baptiste_Joseph_Fourier`.

[3] Lev Kantorovich. Mathematical methods in physics ii.
    `http://www.kcl.ac.uk/kis/schools/phys_eng/physics/courses/CourseList/CP2260.htm`.

[4] `http://nobelprize.org/physics/laureates/1933/dirac-bio.html`.

[5] E Hecht. *Optics*. Addison Wesley, 4th edition, 2002.

The Zeeman Effect

1   Aim of the experiment

The separation of a singlet fluorescence line into three components, when the source molecule is positioned in a magnetic field, is known as the Zeeman effect. It occurs due to the breakdown of degeneracy between electron orbitals due to the aligning of orbitals with a magnetic field. In the absence of a field the magnetic quantum number of an orbital has no effect on the energy of an electron within the orbital, however when a field is present, the orbitals can align with the field depending on their value resulting in a subtle shift in orbital energy. A photon emitted by fluorescence due to an electron shifting between orbitals therefore has a frequency dependent on which orbital it shifted to, resulting in the observed splitting of the singlet line.

The object of this experiment is to measure the separation in frequency of the components into which a singlet fluorescence spectral line (Ne 585.2 nm) is split when the source is in a magnetic field. From this separation a value for the electronic charge to mass ratio $\frac{e}{m_e}$ can be derived.

2   Introduction

The Zeeman effect refers to the splitting of spectral lines in a magnetic field, and is named after Pieter Zeeman (1865–1943), 1902 Nobel Laureate in Physics, who discovered it. [1]

When the source is observed in a direction perpendicular to the magnetic field (transverse effect), three lines, called the Lorentz triplet, are observed in place of the original singlet; the central undisplaced line, or π component, is polarized with the electric vector in a plane parallel to the field, while the two satellites, or σ components, are plane polarized perpendicular to the field. When the source is observed in a direction parallel to the magnetic field, the undisplaced line is absent and the two σ components are circularly polarized. The present experiment looks at the transverse effect.

3   Experimental Procedure

3.1   The Magnet

The coils of the electromagnet are connected in parallel to a supply unit via an ammeter. The current from the supply should not be allowed to exceed 2.2 A (1.1 A per coil) with water cooling to avoid overheating. Switch off the magnet current at times when you are not making observations of the Zeeman splitting of the lines.

3.2   Calibration of the Magnetometer

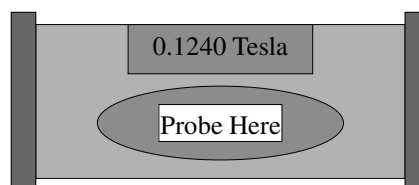Position the probe of the magnetometer as shown in Figure 1 and make sure the reading is 0.124 T.

Figure 1: Schematic of the magnometer set-up.

Fabry-Perot fringe pattern

Fringe pattern seen through slits
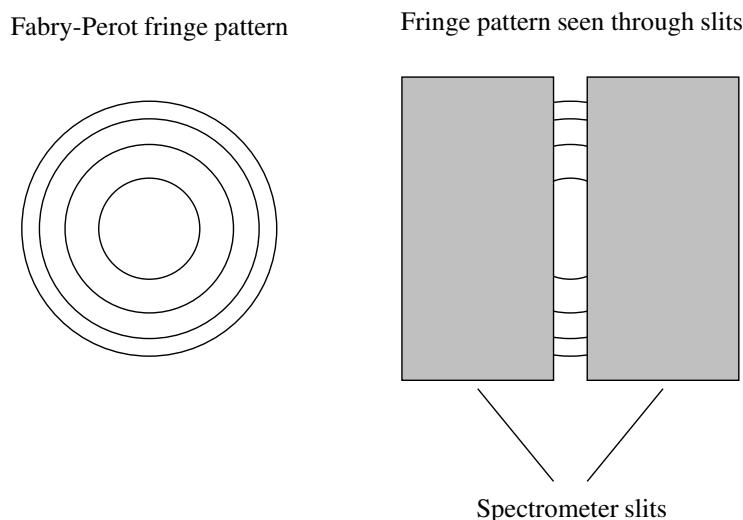

Spectrometer slits

Figure 2: Fabry-Perot fringe pattern. In the above diagram the Fabry-Perot pattern has only a few representative fringes. Normally only an outer section of the fringe system will be observed in the spectrometer.

3.3   Calibration of the Magnet

Use a Hall probe magnetometer to measure the flux density of the magnet as a function of the magnet current. Plot your results in a graph.

3.4   The Optical Equipment

Since the separation of the Zeeman components is small (~ 0.01 nm) a Fabry-Perot etalon (see references [2, 3]) is used in conjunction with a Hilger spectrometer to give sufficient resolving power. Remove the etalon from its table between the prism and the collimator and place the neon discharge tube between the pole pieces of the magnet (**magnet current off**). The neon tube runs at a high voltage so do not switch on the mains until the tube is in position. It is important to ensure that only light emitted from between the magnet poles is examined. This can be achieved by placing a small paper shield with a suitable hole around the discharge tube. The portion of the discharge lying between the poles of the magnet can then be focussed on to the entrance slit of the spectrometer with the lens. Handle the etalon with great care and **never touch the glass surfaces**. Check that the telescope is focussed so that a sharp image of the spectrometer entrance slit can be seen. Observe the neon emission spectrum and check that you understand the action of the spectrometer adjustments, e.g. slit width and height, shutter, wavelength drum and telescope focussing. Set the spectrometer to observe the Ne 585.2 nm line yellow line.

Place the etalon on its table and with a collimator slit of about 0.13 mm. The effect of the etalon is to produce a series of bright fringes along the lines in Ne spectrum (See Figure 2). Note the effect of the two etalon table adjusting screws. One screw (on the bottom of the spectrometer) controls the tilt of the etalon and enables the centre of the Fabry-Perot pattern to be displaced above or below the centre of the field of view. The other screw (on the side of the spectrometer) controls the lateral alignment of the etalon. The three screws on the body of the etalon itself control the parallelism of the etalon plates and **should not be adjusted**. Adjust the etalon table until about 10 fringes are visible in the field of view. With the Variac of the magnet supply unit at zero, switch on the magnet current and increase it until appreciable splitting of the Fabry-Perot rings occurs. Insert the polariser between the collimator and the neon tube and note the effect on the pattern as the polariser is rotated. Record the state of polarization of the various components of the spectral lines and use this to identify the σ and π components of the 585.2 nm line in the Fabry-Perot pattern.

3.5   Determination of $\frac{e}{m_e}$

If ν_0 is the frequency of the undisplaced singlet, the displaced σ components have frequencies

$$\nu = \nu_0 \pm \frac{eB}{4\pi m_e} \tag{1}$$

where B is the magnetic flux density in Teslas. The difference in frequency between the two components is therefore given by equation (2)

$$d\nu = \frac{eB}{4\pi m_e} \qquad (2)$$

For the Fabry-Perot etalon the line of order n obeys

$$n\lambda = 2\mu t \cos\theta \qquad (3)$$

where λ is the wavelength, μ is the refractive index of the material between the plates (~ 1) which are distance t apart and are viewed at an angle θ to their normal. Here θ is approximately zero. Re-arranging this formula in terms of frequency $n = 2\frac{\nu t}{c}$ and $dn = 2\frac{t}{c}d\nu$. The fractional order splitting of the σ lines is therefore given by equation (4). The value of t for this etalon is 5.54 mm.

$$dn = \frac{teB}{2c\pi m_e} \qquad (4)$$

By finding the magnetic field corresponding to a particular fractional order, equation (4) can be used to find a value for $\frac{e}{m_e}$. This will involve the use of your field and coil current calibration graph. Set the polariser so that only the σ components of the Fabry-Perot pattern are visible. Increase the magnetic field until the lines in the pattern move apart. Take a number of images with the charge-coupled device (CCD) camera and work out what the spacing dn of the lines is. Increase the field until the σ components are just overlapped. Show that this corresponds to a fractional order splitting of $dn = 1$.
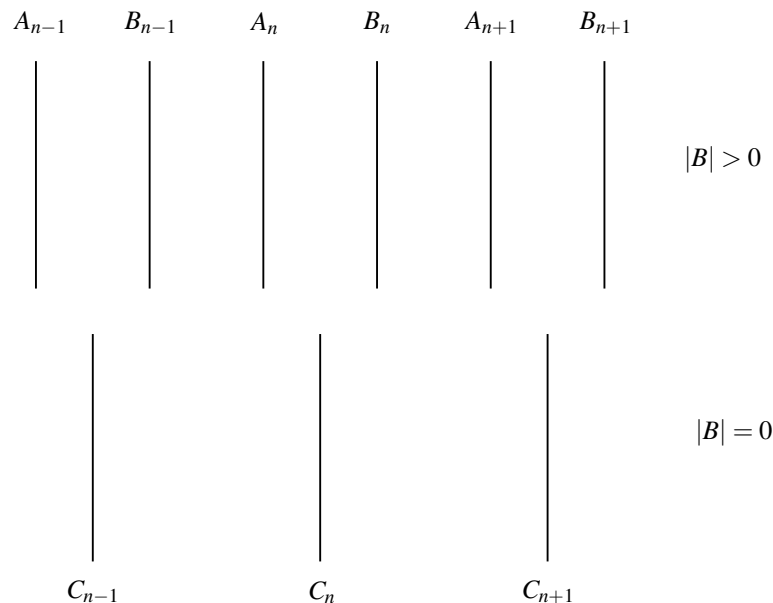

Figure 3: Demonstration of the Zeeman effect. Let the C_n values represent the position of the unsplit spectral lines and the A_n and B_n values be the positions of the split lines in the Fabry-Perot pattern, which results due to the breakdown of orbital degeneracy when the emitting molecule is in a magnetic field. The position of the unsplit or π lines is given by $C_n = \frac{A_n + B_n}{2}$.

Similarly $C_{n+1} = \frac{A_{n+1} + B_{n+1}}{2} \ldots$ and the fractional order splitting dn is given by a formula of the type

$$dn = 2\frac{B_n - A_n}{C_{n+1} - C_{n-1}} \qquad (5)$$

Adjust the polariser so that both the σ and π components of the Ne 585.2 nm line are visible in the Fabry-Perot pattern and adjust the field until all the lines are equally spaced. Show that this condition corresponds to $dn = \frac{2}{3}$ and use the condition to obtain a third value for $\frac{e}{m_e}$.

Plot dn versus B and calculate $\frac{e}{m_e}$ from the gradient of a straight line fit through the origin according to equation (4). You should in your discussion assess the accuracy of your value of $\frac{e}{m_e}$ in terms of uncertainties in the measurement of B, t and the evaluation of dn.

3.6   Calculation of the Separation

Calculate the separation in nm of the σ components of Ne 585.2 for the three fractional order splittings investigated.

3.7   CCD, etalon

Explain briefly how a CCD camera, and an etalon work.

References

[1] http://nobelprize.org/nobel_prizes/physics/laureates/1902/.

[2] F. A. Jenkins and H. E. White. *Fundamentals of Optics*. McGraw-Hill, New York, 4th edition, 1976.

[3] E. Hecht and A. Zajac. *Optics*. Addison-Wesley, Reading (MA), USA, 2nd edition, 1987.

Scintillation Counter

1 Aim of the Experiment

The aim of the experiment is to study the three principal processes by which γ (gamma) rays interact with matter. The processes are photoelectric absorption, Compton scattering and pair production, and the probability of each of these processes occurring depends on the energy of the γ-ray. Three radioactive sources will be used to provide γ-rays of different energy, so that the relative importance of the different interactions can also be assessed as a function of γ-ray energy.

2 Apparatus

Three different γ-ray sources can be obtained from the laboratory technicians. These are ^{137}Cs, ^{60}Co and ^{232}Th, giving γ-rays of energy 0.662 MeV, 1.33 MeV and 2.61 MeV respectively, so the energy of the γ-ray is increasing approximately by a factor of two as you go from one source to the next. You should note that the ^{60}Co source also produces a second, lower-energy γ-ray. You will be expected to determine the energy of this γ-ray as part of the experiment.

The different γ-ray interactions can be studied by allowing the γ-rays to interact with a crystal of sodium iodide (NaI) that has been doped with thallium. This crystal is known as a scintillator, since it gives off flashes of visible light when energy is deposited in the crystal. A photomultiplier is a very sensitive light detector that is used to measure the amount of light emitted by the scintillator, and hence to measure the amount of energy lost by the γ-ray as a result of its interaction with the NaI crystal. In each case the γ-ray interaction results in kinetic energy being given to an electron in the target material, and it is this kinetic energy that is then dissipated within the volume of the NaI crystal, causing the crystal to scintillate.

The light produced by the scintillator is detected by a photomultiplier tube (*PMT*). The PMT has a **photocathode** with a low work function that converts light pulses into electron pulses; a high potential difference accelerates the electrons towards of a series of **dynodes**, each of which liberates a further cascade of secondary electrons. The resulting charge pulse is then converted to a voltage pulse whose height is proportional to the intensity of the original light flash from the scintillator.

The essential components within a scintillation-counter system are:

i a mass of scintillating material (in this case the thallium-activated sodium iodide crystal) in good optical contact with the PMT,

ii a high-voltage power supply for the PMT and

iii a single channel analyser that incorporates the electronics needed to count and sort the number of electrical pulses of different sizes that emerge from the photomultiplier.

2.1 The scintillating crystal and photomultiplier tube

The sodium iodide crystal and photomultiplier assembly is sealed within a light-tight metal container with the crystal mounted behind a thin aluminium end plate. Sodium iodide is deliquescent and must not be exposed to the atmosphere, and the photomultiplier must be operated in almost total darkness otherwise the photocathode would be destroyed. A single coaxial lead supplies the high voltage for the photomultiplier and also returns the signal pulses to the amplifier unit.

2.2 The single channel analyser unit (ST7)

The ST7 scaler-timer is an example of a single channel analyser. It contains power supplies, amplifiers, counting and timing circuits. Signal pulses from the PMT are first amplified, and then sent to an analyser unit that passes only those pulses with a peak height lying in the range V to $V + dV$ to a counter-timer unit. V is known as the analyser **threshold voltage**, and dV is the **channel width** or **window voltage**.

In practice the analyser unit works as follows. The signal pulses sent to the analyser are fed to two circuits: the first emits a standard digital logic pulse (A) if the signal pulse is bigger than V, while the second emits a logic pulse (B) if the signal pulse is bigger than $V + dV$. Digital logic circuits then form an output logic pulse from the signal combination A AND (NOT B), so that a logic pulse is only sent to the counter if the original signal pulse was greater than V and not greater than $V + dV$. The output logic pulses can be counted for a preset time, or the time taken to accumulate a preset number of pulses can be recorded, depending on the mode of operation selected for the ST7.

In this experiment it will be better to count for a preset time that gives an adequate accuracy for the highest count rate of interest in that part of the experiment. By varying the threshold voltage V and recording the number of signal pulses in the range V to $V + dV$, a form of histogram known as a **pulse height spectrum** can be plotted for each of the radioactive sources under investigation.

3 Theoretical Background

There are three main ways in which γ-rays may interact with matter. The pulse height spectra that you record will show evidence for these various interactions, and it should be possible to make a quantitative comparison between the measured positions of features within these spectra and the positions that are predicted by theory.

3.1 Photoelectric absorption

In photoelectric absorption the gamma ray is completely annihilated, as it gives up all its energy in knocking a bound electron out of what is usually an inner atomic orbital. (Some of the energy may subsequently re-appear as an X-ray when the ionised species returns to its lowest energy configuration.) For a given γ-ray, photoelectric absorption is the **maximum energy event**; so it generates the largest flash of light within the scintillator, which in turn causes the largest signal pulse. Photoelectric absorption is relatively more important at low γ-ray energies, and becomes less likely as the energy of the gamma ray increases.

3.2 Compton scattering

In Compton scattering a γ-ray is involved in a "glancing collision" with a virtually free electron, and is scattered through an angle φ. By considering the conservation of energy and linear momentum, it can be shown that the energy E_e transferred to the electron during this process is a function of the scattering angle φ

$$E_e(\varphi) = \frac{E_\gamma}{1 + \dfrac{m_0 c^2}{E_\gamma(1 - \cos\varphi)}} \tag{1}$$

where E_γ is the energy of the incident γ-ray, m_0 is the rest mass of the electron and c is the velocity of light. (In using this expression, remember that the rest mass-energy of the electron is $m_0 c^2 = 0.511$ MeV). This suggests Compton scattering can result in a continuum of energies $E_e(\varphi)$, extending from 0, when the scattering angle $\varphi = 0$, up to a maximum of

$$E_e(\pi) = \frac{E_\gamma}{1 + \dfrac{m_0 c^2}{2E_\gamma}} = E_C \tag{2}$$

when $\varphi = \pi$ (the scattered γ-ray is reflected back along its original path). This maximum energy E_C is known as the "Compton edge", since the continuum distribution in the pulse height spectrum looks a like a broad plateau that drops rapidly to zero at the energy E_C.

3.3 The backscatter peak

γ-rays from the source that initially travel away from the scintillator detector can be Compton scattered by some matter in the surrounding environment and be redirected towards the scintillator. Clearly, such Compton-scattered γ-rays will have lower energy than the γ-ray originally emitted by the source, but may still interact with the scintillating crystal. By considering the geometry of the experiment, it can be seen that rays travelling away from the scintillator which are then Compton scattered through almost 180° have a good chance of subsequently reaching the scintillator detector. These γ-rays will have an energy $E_B = E_\gamma - E_C$, so a peak, known as the **backscatter peak**, is expected in the pulse height spectrum at this energy. When a steel plate is placed on the opposite side of the source from the scintillator, it will enhance the likelihood of *gamma*-rays being backscattered, and the peak should increase in height.

3.4 Pair production

If the incident γ-ray has energy greater than 1.022 MeV the formation of an electron-positron pair can occur. Clearly an energy of 0.511 MeV is required to create each of the positron and the electron, and the residual energy of $E_\gamma - 1.022$ MeV is equally divided as kinetic energy between the two particles. Both the electron and positron are charged particles, so their range within the NaI Crystal is very short; they soon give up all their kinetic energy to the surrounding crystal, causing it to scintillate. In addition, within a very short time (less than 10^{-10} s) the positron will annihilate with another electron, producing two γ-rays, each of energy 0.511 MeV.

If both of these γ-rays escape from the scintillator without further interactions, then the total energy deposited in the NaI crystal is just $E_\gamma - 1.022$ MeV, and a peak known as the **double escape peak** will appear in the pulse height spectrum at this energy.

If only one of the two γ-rays escapes, while the other is photoelectrically absorbed, then the total energy deposited in the NaI crystal is $E_\gamma - 0.511$ MeV, and a peak known as the **single escape peak** will appear in the pulse height spectrum at this energy.

When both γ-rays produced by positron annihilation are photoelectrically absorbed by the scintillator, the total energy absorbed by the scintillator will be E_γ, so a peak known as the **total conversion peak** will appear at exactly the same position in the pulse height spectrum as if the original γ-ray had been photoelectrically absorbed. With the apparatus available in this experiment it is not possible to distinguish a total conversion peak from a photoelectric absorption peak, but it should be remembered that as the γ-ray energy increases the probability of pair production will increase, whereas the probability of photoelectric absorption will decrease.

4 Experimental Procedure

The electronic equipment may be used for setting-up purposes as soon as it is switched on, but it should be allowed to warm up for an hour before the final measurements are taken. The detector assembly should be vertical with the scintillator crystal at the top. The first task is to set the high voltage (*HV*) supply for the PMT to a suitable value for the γ-ray source you are using, but before you can do this you should ensure that the ST7 controls are at the settings shown in Table 1.

ST7 Control	Setting
INT-DIFF	*DIFF*
HV	000
WINDOW	020
THRESHOLD	1000
COUNT	5×10^5
TIME	10
DISPLAY	*TREND*

Table 1:

Note: The HV, threshold and window voltage controls are ten-turn potentiometers and each turn is divided into 100 divisions. The range of values for these controls is therefore from 000 to 1000. The small lever on the top right edge of these dials is a locking mechanism. Adjust the dials with the lever in the fully anticlockwise position, but when you are at the correct setting, move the lever to the fully clockwise position to prevent any accidental change to its setting.

N.B.: Always use the tweezers provided when you are handling the radioactive sources. Never touch them directly with your hands.
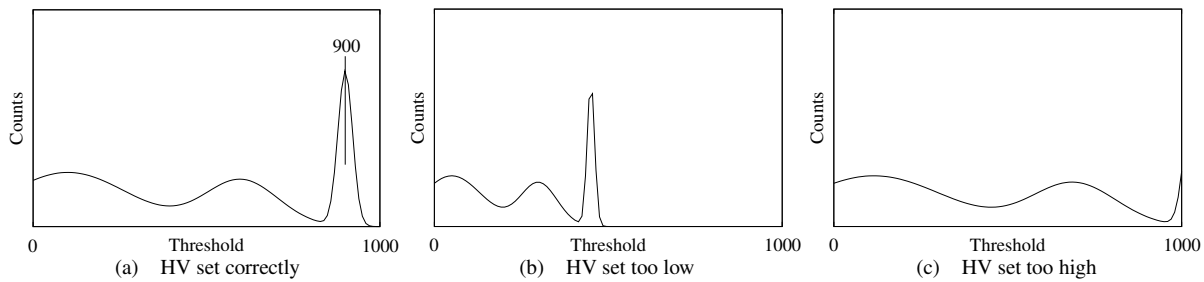
Always make sure that, when you are making measurements with one of the radioactive sources, the other two sources are located well away from the scintillation counter, so that the signals you detect are not affected by the γ-rays emitted by the other sources.

Figure 1: Examples of pulse height spectra from ^{137}Cs for different HV settings

4.1 Adjusting the HV for the ^{137}Cs source

Place the ^{137}Cs source on a support stand a few centimetres above the scintillator and PMT. Start the ST7 on *MULTICYCLE*, and with the DISPLAY set to *TREND*, the LCD panel will show the counts recorded in the previous one second. Slowly increase the HV until a reading of a few tens of counts per second starts to be seen. Keeping the same HV setting, now slowly reduce the THRESHOLD until you see that the recorded count rate goes through a well-defined maximum. This maximum should correspond to the photoelectric absorption peak for the γ-ray from the ^{137}Cs source, and the aim is to adjust the HV until the photoelectric absorption peak is found at a threshold setting of around 900 on the dial.

An example of a pulse height spectrum from ^{137}Cs with the HV set correctly is shown in Figure 1. The high voltage for the photomultiplier is a very non-linear gain control for the pulse heights so it needs to be adjusted very carefully. If the HV setting is slightly too low, then the pulse heights from the PMT will be significantly smaller and the pulse height spectrum will be compressed towards the low end of the threshold voltage axis. Conversely, if the HV setting is slightly too high, all the pulse heights from the PMT will be significantly increased, the spectrum will expand towards the high end of the threshold voltage axis, and it will become impossible to find the photoelectric absorption peak within the available range of threshold voltages.

If, after your initial adjustment of the HV, you find that the maximum corresponding to the photoelectric absorption is not near 900 on the threshold voltage scale, then reset the THRESHOLD to 1000, and adjust the HV setting as appropriate. Then check once more to see whether you have the maximum count rate at a threshold setting of 900.

When you have finished adjusting the HV setting, check that the count rate is small when you reset the THRESHOLD to 1000. If it is not, then the HV is not yet correctly set.

4.2 Obtain pulse height spectra from the ^{137}Cs source

Once the HV is set correctly, set the DISPLAY control to either *COUNT* or *RATE* and operate the ST7 in single cycle mode instead of multicycle mode. Make sure that the TIME of counting for each acquisition cycle is set to a value long enough for there to be at least 1000 counts when the threshold is at the maximum of the photoelectric absorption peak.

Starting at a THRESHOLD of 1000 record the number of counts (or, if you prefer, the count rate) in each acquisition cycle as a function of threshold voltage between 1000 and 0. You need to decide on a sensible interval between threshold values, remembering that there is no advantage in using an interval that is smaller than the WINDOW setting. Plot your data as soon as you have acquired them.

Collect a second pulse height spectrum once you have placed a horizontal steel plate a few centimetres ***above*** the ^{137}Cs source, and compare this spectrum with the first one.

4.3 Obtain a pulse height spectrum from the ^{60}Co source

Replace the ^{137}Cs source by the ^{60}Co source. This is a much stronger source so make sure that it placed at least 10 cm above the scintillator to avoid saturating the detector. Use the same procedure as in section 4.1 to adjust the HV for the PMT so that the photoelectric absorption peak for the higher energy γ-ray emitted by the ^{60}Co source occurs at a threshold setting around 900. Since the energy of this γ-ray is 1.33 MeV (compared to 0.662 MeV for the ^{137}Cs source) it is clear that the HV setting will be lower in this case.

Choose an acquisition time that will again ensure you measure at least 1000 counts when at the maximum of the higher-energy photoelectric absorption peak, and collect a pulse height spectrum between threshold settings of 1000 and 0. Plot the data as soon as you have acquired them, and confirm that your graph shows **two photoelectric absorption peaks**.

4.4 Obtain a pulse height spectrum from the ^{232}Th source

The photoelectric absorption peak for the γ-rays emitted by the thorium source is too small to be easily used for setting up the HV of the PMT in this part of the experiment. In this case, the best method is to use the ^{137}Cs source as described in section 4.1, but to set the maximum of the photoelectric absorption peak at a threshold setting of 220 (instead of 900). The maximum threshold setting of 1000 will now correspond to an energy of around 3MeV.

Using an acquisition time of at least 30 s per cycle, obtain a pulse height spectrum for the ^{232}Th source for threshold voltages in the range 1000 to 200. Plot the data as soon as you have acquired them.

4.5 Check the linearity of the scintillator-PMT detector

Keeping the same HV setting as you used in section 4.4, replace the thorium source by the cobalt and then the caesium sources, and measure the threshold values at which the photoelectric absorption peaks for these two sources can be found. Plot a graph of these threshold values against γ-ray energy, and see if the four points all lie on a straight line passing through the origin. *You will need to look up the accepted value for the second ^{60}Co γ-ray.* If the four points do not lie on a straight line, what consequence does this have for the analysis of the pulse height spectra?

5 Data Analysis

Examine the four pulse height spectra acquired in the experiment, and identify as many of the features as possible in terms of photoelectric absorption, Compton scattering, and pair production events.

When analysing the data, you should remember that radioactive decay is a random process, so that a standard deviation of $\sqrt{N}$ must be associated with any single recording of N counts (assuming no other sources of error). Plot vertical error bars of the correct size on your graphs, so that you can distinguish statistically significant changes in the measured count from random fluctuations.

You should also remember that the pulse height spectra are effectively histograms; a single channel analyser can only indicate that a count was measured somewhere between V and $V + dV$.

5.1 Calibrate the threshold axes in terms of energy

The scintillation counter relies on the proportionality between

 i the energy deposited in the scintillator by the gamma ray,

 ii the intensity of the flash of light generated by the scintillator and

iii the peak-height of the pulse which leaves the photomultiplier.

Assuming there is such a linear relation between the energy deposited in the scintillator by the γ-ray and the pulse height measured by the SCA, the position of the photoelectric absorption peak can be used as a fixed point to calibrate the rest of the threshold axis in terms of energy.

Using the known energies for the γ-rays incident on the scintillator, calibrate the threshold axes of the four pulse height spectra in terms of energy. Note that there will obviously be an experimental error associated with this calibration, because of the uncertainty in determining the threshold value at which the photoelectric absorption peak is found.

For the ^{60}Co source, use the energy calibration to determine the energy of the second γ-ray emitted by the source, and compare your measured value with the accepted value.

5.2 Predict the energies for Compton scattering and pair production events

Use the information given in section 3 to calculate the energies of the Compton edges, backscatter peaks, single and double escape peaks for the pulse height spectra that you have acquired. Mark the predicted positions of these features on the graphs, and make a quantitative comparison between the predicted positions and the measured positions, taking proper account of the experimental errors involved.

References

[1] K. S. Krane. *Introductory Nuclear Physics*. Wiley, 1987.

[2] F. Adams and R. Dams. *Applied Gamma-ray Spectrometry*. Oxford:Pergamon.

Microwave Experiments

1 Aim of the Experiment

Waveguides can be used to control the direction of propagation of electromagnetic waves by means of reflections from their internal walls. They therefore need to be constructed of materials giving a high reflectivity at the desired electromagnetic wavelength of operation, therefore metals are commonly used. The purpose of the following experiments is to illustrate the basic theory of microwave propagation in waveguides, with three specific objectives. The first task is to measure the wavelength of the electromagnetic field distribution in the guide by three different methods, and to compare this with the wavelength of propagation in free space. Additionally you will make measurements of the voltage standing wave ratio, the ratio of the maximum and minimum voltage recorded by a detector in the guide while a standing wave is oscillating. This is used to measure the reflectivity of a dielectric boundary inserted at the end of the guide. Finally, you will measure the permittivity of two different dielectric materials at microwave frequencies by measuring the shift in standing wave wavelength when they are placed in the guide.

2 Apparatus

The components in this set of experiments are shown in the block diagram of Figure 1. Ensure that there is no dielectric insert initially present in the standing wave detector. Connect the microwave components together as shown in the diagram and attach the adjustable short circuit, which has a micrometer on the end of it, to the terminal end of the waveguide. Connect the cathode ray oscilloscope (*CRO*) and millivoltmeter to the standing wave detector.

2.1 The Gunn Diode

A Gunn diode is a thin piece of gallium arsenide to which a uniform DC voltage of $\sim$7 V is applied. The relationship between current and potential difference is non-linear in GaAs as shown in Figure 2 and an electric domain sweeps rapidly across the GaAs, the time of transit depending on the thickness. The non-linear characteristics of GaAs produce about 8 mW of power at between 8 and 12 GHz (1 GHz = 10^9 Hz) in the microwave cavity. The oscilloscope is unable to respond to frequencies this high so the diode voltage is modulated by a low frequency square wave at about 1 kHz that switches the diode output off and on. A rectifying diode is permanently connected in the housing of the standing wave detector and the signal shown on the oscilloscope is the 1 kHz square wave produced by modulation of the power supply. Note that what you see on the CRO screen is the envelope of the microwave signal, not the signal itself.
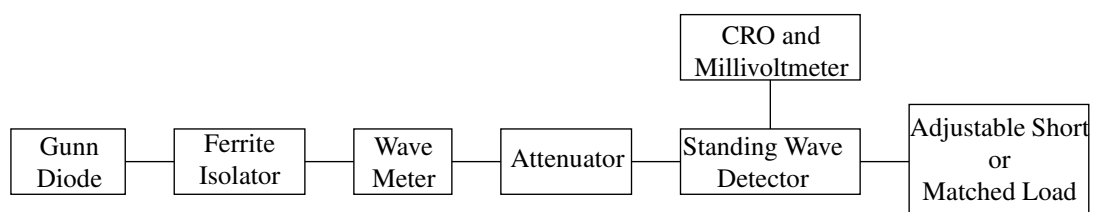
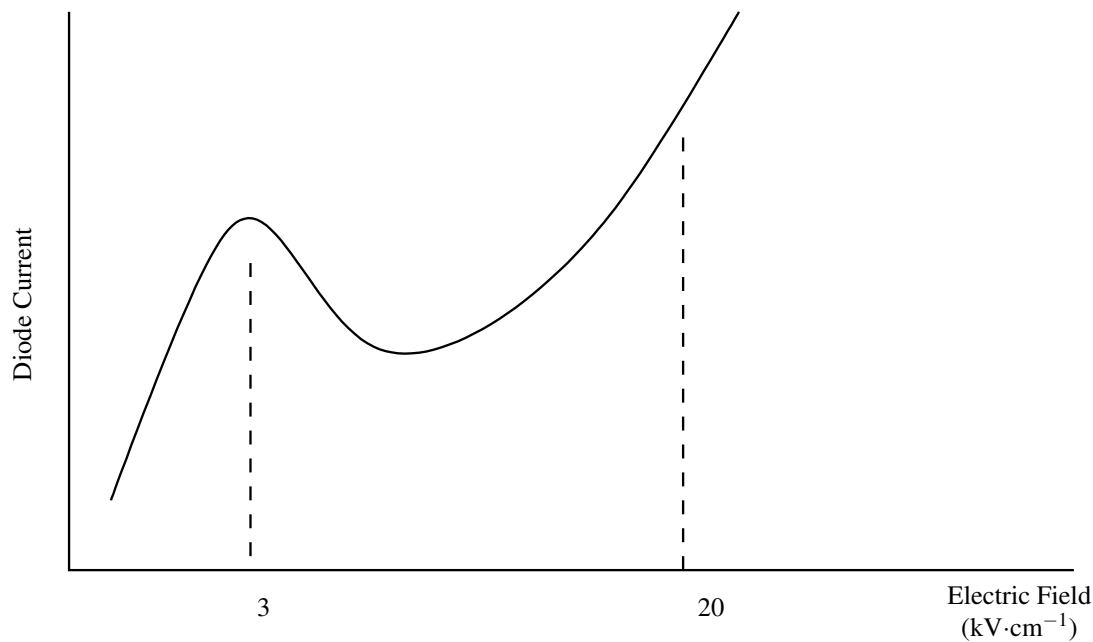Figure 1: Experimental set up.

Figure 2: The non-linear relationship between electrical current and voltage of GaAs.

3 Theoretical Background

The wavelength of the field distribution in the guide is not the same as the wavelength of the same frequency in free space, which is λ_0. The wavelength in the guide will depend on the dimensions of the wave guide, and on the mode of propagation of the radiation. For a rectangular guide the dominant mode is the TE_{10} (transverse electric) mode. The guide wavelength λ_g is related to λ_0 by equation (1)

$$\left(\frac{1}{\lambda_g}\right)^2 + \left(\frac{1}{\lambda_c}\right)^2 = \frac{\varepsilon_r}{\lambda_0^2} \tag{1}$$

where ε_r is the relative permittivity (or dielectric constant) of the material in the guide and λ_c is the critical wavelength for the TE_{10} mode of wave propagation in the guide. The critical wavelength has the value $2a$, where a is the larger dimension of the internal cross-section of the guide ($a = 22.86$ mm for the waveguide used in this experiment). When EM waves reach the end of the guide some are reflected back towards the source. This means that two waves of the same wavelength will be propagating in opposite directions along the guide, allowing a standing wave to be set up. The fraction of the amplitude reflected by the end of the guide is $|r|$, which determines the nature of the standing wave pattern that is set up.

4 Experimental Method

Make certain that the apparatus are correctly set up and switch on the power supply. The attenuator should be set for minimum attenuation, i.e. fully out. Move the standing wave detector down the waveguide and tune the adjustable short to obtain maximum signal on the oscilloscope. This should show a square wave alternating between zero voltage and a few mV, which may be negative. Expand the oscilloscope trace vertically and shift the zero point so that you can examine the detector voltage at a high sensitivity. Move the micrometer on the adjustable short if necessary to obtain as flat a trace as possible. You may further improve the display by inserting a small amount of attenuation with the attenuator.

4.1 Measuring the frequency of the wave

The wave-meter has a large cylindrical dial with a spiral scale. To measure the frequency, rotate the dial until the trace on the oscilloscope is distorted. Rotation of the dial is actually causing the length of a small resonant cavity in the wavemeter to change, and the distortion of the CRO trace occurs when microwave power is absorbed

by this cavity instead of propagating down the waveguide. The distortion is usually quite small, so the dial must be adjusted slowly and carefully. Record the frequency at which the distortion occurs, and calculate the wavelength in free space, taking the speed of light to be 2.9979×10^8 m·s^{-1}.

4.2   Measuring the wavelength in the guide

Move the probe of the standing wave detector (*SWD*) along the guide and observe the amplitude of the standing wave pattern on the oscilloscope. Locate all the nodes (minima) and all the antinodes (maxima), measuring their positions on the SWD scale, and decide whether the minima or maxima positions can be found with greater accuracy. Note that it is also possible to shift the pattern along the guide by adjusting the micrometer of the adjustable short circuit. The distance between 2 adjacent nodes (or 2 antinodes) is $\frac{\lambda_g}{2}$. Measure the half wavelengths for as many cycles as possible. One way of calculating a mean value of λ_g is to number successive minima and maxima, $n = 0, 1, 2, 3 \ldots$ and then to plot a graph of their positions against n. The slope of this graph (which you should determine by a least squares fit) gives, $\frac{\lambda_g}{4}$ and the free space wavelength λ_0 can then be found using equation (1).

4.3   Measuring the form of the standing wave pattern in the guide

The standing wave pattern will follow a $|\sin x|$ (mod sine) relation if the detector has a linear response. Plot the voltage produced by the probe at several positions between several minima, and decide whether your results do indeed follow a $|\sin x|$ curve. Note that it is also possible to measure the signal with a millivoltmeter in place of the oscilloscope if you find this a better method.

4.4   Measuring the Voltage Standing Wave Ratio and Reflectivity

With the adjustable short circuit terminating the guide, record the amplitudes of the maximum and minimum voltages observed with either the voltmeter or the oscilloscope. Switch off the Gunn diode, then replace the adjustable short with the matched load and again measure the amplitudes of the maxima and minima of the new pattern. The ratio of the $V_{\max}$ to $V_{\min}$ is called the voltage standing wave ratio (*VSWR*) and is related to the fraction of the amplitude reflected by the load by the formula

$$VSWR = \frac{V_{\max}}{V_{\min}} = \frac{1 + |r|}{1 - |r|} \tag{2}$$

Calculate the $VSWR$ and hence the reflectivity for the two guide terminations. Compare the values measured with those expected for a perfect short circuit, when $|r| = 1$, and for a perfect matched load, when $|r| = 0$.

4.5   Measuring the relative permittivity of Perspex and PTFE

Again switch off the Gunn diode, remove the matched load, insert the Perspex (*PMMA*) block into the standing wave detector, with the slotted side facing upwards, and terminate the guide with the adjustable short circuit. Adjust the plunger on the standing wave detector if necessary to obtain a large signal and again measure the wavelength in the guide using the same procedure as in Section 4.2 over as many cycles as possible. The wavelength in the air filled portion of the guide remains the same but the wavelength in the dielectric will be smaller. Repeat the procedure with the polytetrafluoroethylene (*PTFE*) block and calculate the relative permittivity ε_r using equation (1).

4.6   Measure the free-space wavelength directly using standing waves

Switch off the Gunn diode, remove the adjustable short circuit and the dielectric block from the wave guide, and replace these with the microwave horn. Reflect the microwaves that emerge from the horn back into the guide by using a plane sheet of metal that is perpendicular to the direction of propagation of the waves. Move the metal sheet in the direction of microwave propagation and observe the signal on the standing wave detector. As the sheet is moving in free space the distance between adjacent minima or maxima is $\frac{\lambda_0}{2}$. Measure the positions of the reflecting sheet for as many maxima and minima as possible. As in Section 4.2, one way of calculating a mean value of λ_0 is to number successive minima and maxima $n = 1, 2, 3 \ldots$ and then to plot a graph of their positions against n. The slope of this graph (which you should determine by a least squares fit) gives $\frac{\lambda_0}{4}$.

5 Data Analysis

Compare the values of λ_0 measured in sections 4.1, 4.2 and 4.6. Decide whether the values are in agreement within experimental error, and comment on which methods are likely to be the most accurate and reliable. Compare the values for the relative permittivity measured in section 4.5 with the accepted values for these materials. If there is any discrepancy, then try to suggest reasons for this.

5.1 Calculating the error in λ_0

Taking ε_r to be a constant, use the standard expressions for the propagation of experimental error to show that the uncertainty in λ_0 is given by the equation (3), where the experimental errors are indicated by a Δ prefix.

$$\frac{\Delta\lambda_0}{\lambda_0} = \left(\frac{\lambda_0^2}{\varepsilon_r \lambda_g^2}\right) \frac{\Delta\lambda_g}{\lambda_g} \tag{3}$$

5.2 Calculation of the error in $|r|$

Use the standard expressions for the propagation of experimental error to show that the uncertainty in $|r|$ is given by equation (4) where the experimental errors are indicated by the prefix Δ.

$$\Delta|r| = 2\frac{\Delta(VSWR)}{(VSWR+1)^2} \tag{4}$$

Taking experimental errors into account, comment on how well the loads provided for the experiment approximate to their ideal counterparts.

5.3 Calculation of the error in ε_r

The value of ε_r depends on the measured values of both λ_g and λ_0. Use the standard expressions for the propagation of experimental error to show that the uncertainty in ε_r is given by equation (5) where the experimental errors are indicated by the prefix Δ. Identify which of the measured values contributes most to the final error in ε_r.

$$\Delta\varepsilon_r = 2\sqrt{\varepsilon_r^2 \left(\frac{\Delta\lambda_0}{\lambda_0}\right)^2 + \left(\frac{\lambda_0^4}{\lambda_g^4}\right)\left(\frac{\Delta\lambda_g}{\lambda_g}\right)^2} \tag{5}$$

Compare the values you obtained with the litterature values at 5 MHz [1]: $\varepsilon_{r(perspex)} = 2.70 \pm 0.05$ and $\varepsilon_{r(PTFE)} = 2.00 \pm 0.05$.

References

[1] F. Thompson and J. K. Gagon. Fringe capacitance of a parallel-plate capacitor. *Physics Education*, 17(2):80–82, 1982.

[2] B. L. Bleaney and B. Bleaney. *Electricity and Magnetism*, volume 1. Oxford University Press, Oxford, 3rd edition, 1989.

[3] P. A. Matthews and I. M. Stephenson. *Microwave Components*. Chapman and Hall, London, 1968.

[4] E. L. Glinzton. *Microwave Measurements*. McGraw-Hill, London, 1957.

[5] A. W. Cross. *Experimental Microwaves*. W. H. Sanders, London, 1960.

Crystal Structure and X-ray Diffraction

1 Aim of the Experiment

The object of the experiment is to demonstrate the Bragg scattering of X-rays from crystalline materials and to show how the information gained by this technique is used to determine crystal structures.

2 Apparatus

The Bragg scattering measurements are performed with a *Tel-X-Ometer* X-ray generator. An X-ray beam is produced in a sealed tube by firing 30 keV electrons at a copper target. The beam is collimated by a narrow slit in front of the X-ray tube and allowed to fall on a crystal specimen clamped to a post at the centre of the diffractometer. The X-ray tube produces two strong emission lines corresponding to particular atomic transitions in the copper target. These are the copper K_α and K_β lines, and they sit on top of a continuous (or 'white') X-ray background caused by deceleration of the 30 keV electrons within the copper target (the background is known as *Bremsstrahlung* [='braking radiation']).

A Geiger-Müller (*GM*) tube is mounted on a carriage that moves around the circumference of the *Tel-X-Ometer* platform, and is used to measure the scattered radiation as a function of angle. The detector carriage is geared to move at twice the angular rate of the crystal mounted at the centre. The GM tube should be mounted on the carriage so that its entrance window is as far as possible from the crystal (while still allowing the lid of the apparatus to be closed), and a lead slit is mounted in front of the detector to define the angular range of scattered X-rays that will be collected by the detector.

If the white "POWER-ON" lamp does not come on when the mains are switched on the timer is probably set at zero—rotate it to the required setting. On pressing the "X-RAYS ON" button to switch on the EHT supply, the red warning lamp should light. However, safety interlocks prevent the high voltage supply to the X-ray tube being switched on while the *Tel-X-Ometer* cover is open; the cover being locked in its central position by a ball-ended spigot that locates in a key-hole slot behind the lead scatter shield.

Five different single crystal samples are provided for the experiment. These are the alkali halides lithium fluoride, sodium chloride, potassium chloride, and rubidium chloride, and the mineral periclase (MgO). When not being used for measurements these should be stored in the small vials provided. Note that to ease identification, one end of each crystal has been painted with the same colour as the marker in the bottom of each vial.

3 Theoretical Background

A crystal is an ordered assembly of atoms or molecules. It consists of a basic repeating unit, known as the *unit cell*, located at each point in a regular 3-dimensional array known as a *lattice*. The structure of the crystal is then determined by the structure and dimensions of the unit cell; in its simplest form the unit cell may contain just a single atom, or it could contain a large assembly of atoms, such as a molecule of organic or biological material. The distances between consecutive lattice points in the crystal are called the *lattice constants*; clearly the lattice constants are equal to the dimensions of the unit cell.

In this experiment all the crystals have the same type of cubic structure that is found in sodium chloride (NaCl); the unit cell is a cube that has an atomic arrangement illustrated by the simple structural model that can be found beside the experimental apparatus. Since the unit cell is cubic the three axes of the lattice are mutually perpendicular and the repeating distance along each axis is the same, so that a single lattice constant is enough to specify the dimensions of the unit cell. For cubic systems, this lattice constant is sometimes referred to as the 'cell edge'. The analysis of the diffraction information from such a crystal takes a particularly simple form.
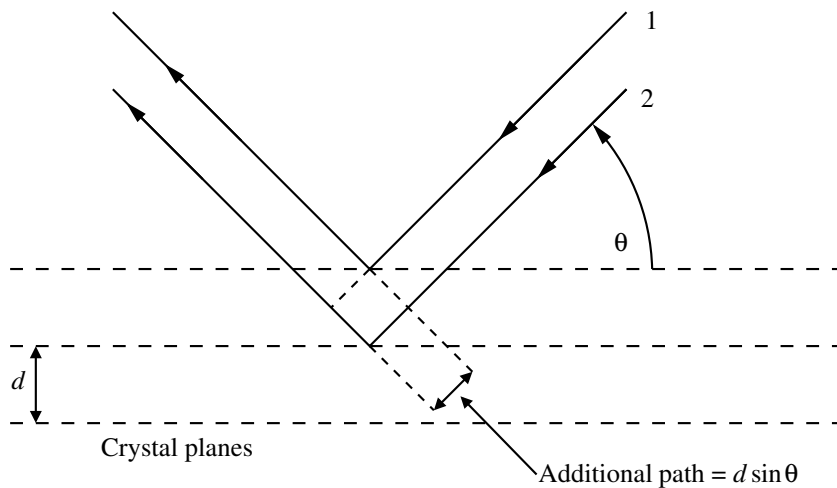
Figure 1: Bragg scattering from a set of parallel planes.

3.1 Bragg scattering from crystal planes

A crystal can be thought of as a set of atomic planes stacked one on top of another. By treating each of these planes as a partially reflecting surface, Bragg showed that there would be very strong scattering at certain well-defined angles, that is, at angles for which the path difference between waves reflected from successive crystal planes is an integer multiple of the wavelength λ. The diagram in Figure 1 shows a set of parallel crystal planes separated by a perpendicular distance d, with illumination of wavelength λ incident at an angle θ measured relative to surface of the reflecting planes. Clearly ray 2 will travel an additional path compared to ray 1, but the two reflected waves will remain in phase with one another provided the angle of incidence satisfies the Bragg equation (1). [1]

$$\lambda = 2d\sin\theta \tag{1}$$

Note that the convention in crystallography is simply to have λ on the left of equation (1), not $n\lambda$, where n is an integer, because the factor n is usually taken into account in the definition of the interplanar spacing d.

In the early part of the 20th century, William Bragg (1862–1942) and his son Lawrence (1890–1971) worked on the analysis of crystal structure by X-ray diffraction. The use of X-rays as a tool to investigate crystal structure was recognized by the award of the Nobel Prize jointly to father and son in 1915. [2].

3.2 Miller indices

In any crystal lattice there are an infinite number of sets of parallel crystal planes. Miller indices are an elegant way of identifying a whole family of parallel planes that have the same interplanar spacing. The Miller indices are usually denoted by the symbols $(h\ k\ l)$, where h, k and l are integers that may be interpreted as follows.

Consider a set of parallel planes with interplanar spacing d and Miller indices $(h\ k\ l)$. Taking one corner of the unit cell as the origin, you will find that as you move by one lattice constant along the x axis of the unit cell you will pass through h members of the set of planes (excluding any member that of the set that passes through the origin). Similarly, if you travel along the y axis of the unit cell for one lattice constant, you will pass through k members of the set of planes, and if you move along the z axis by one lattice constant you will pass through l planes belonging to the same set.

To confirm to yourself that you now understand the principle involved in determining the Miller indices, use the structural model provided to identify the sets of planes described by the following indices: (2 0 0), (4 0 0), (0 2 0), (1 1 0), (2 2 0), (1 1 1). *Note that the (4 0 0) planes do not always coincide with atomic planes.* If you succeeded with this exercise you should now realise that the three Miller indices $(h\ k\ l)$ are sufficient to specify both the orientation and the spacing of a set of parallel planes. For cubic crystals such as the ones used in this experiment the relationship between the lattice constant a, the Miller indices, and the interplanar spacing d is given by equation (2).

$$d = \frac{a}{\sqrt{h^2 + k^2 + l^2}} \tag{2}$$

Always remember that the Miller indices do **not** refer to a single plane in the crystal, but to a whole family of parallel planes that have the same interplanar spacing.

4   Experimental Procedure

4.1   Measurements on lithium fluoride

Mount the LiF crystal with the roughened side towards the crystal post of the diffractometer. Ensure that the slide mount containing the lead slit is in front of the GM tube and switch on the X-ray beam. Work on the left hand side of the main beam and scan the counter slowly through increasing scattering angles until you find the first of two closely spaced peaks in the intensity distribution of scattered X-rays. These peaks are due to the copper K_α and K_β wavelengths being scattered from the (2 0 0) planes in the LiF crystal. The larger of these is the K_α peak, and should be around 1000 counts per second. If you cannot find a peak of this magnitude, consult a demonstrator.

Measure the count rate for different angular settings as you scan through the peaks and, from the angles at which you observe the maxima, calculate the wavelength of the K_β radiation, given that the K_α radiation has a wavelength of 0.1542 nm. Remember that the angular scale on the *Tel-X-Ometer* gives the angle 2θ through which the X-ray beam has been deviated. It is the angle θ (the **Bragg angle**) that is needed when using the Bragg equation.

Using equations (1) and (2) and the measured Bragg angle for the (2 0 0) peaks, predict where you would expect to measure the (4 0 0) peaks, and then measure the actual positions at which you find them.

4.2   The effect of a nickel filter

The K_β radiation can be effectively filtered out of the X-ray beam by using a nickel filter, since nickel has an absorption edge in the X-ray region between the copper K_α and K_β wavelengths. Insert the Ni foil into the path of the X-ray beam and observe the effect that it has on the intensities of the K_α and K_β peaks.

4.3   Measurements on NaCl, KCl and RbCl

Measure the positions of the (2 0 0) and (4 0 0) peaks for each of the other alkali halide crystals. You need only use the stronger K_α radiation for these measurements.

4.4   Measurements on periclase (magnesium oxide)

4.4.1   Single crystal MgO

Measure the positions of the (2 0 0) and (4 0 0) peaks for the MgO single crystal sample, using the K_α radiation for these measurements.

4.4.2   Powder MgO

Measurements on single crystal specimens are often complemented by measurements on powdered samples, where all orientations of crystallites are simultaneously present and able to diffract. This has the advantage of speed, in that all possible Bragg reflections can be recorded simultaneously, and sometimes it is necessary to study powder samples because it is impossible to prepare sufficiently large samples in single crystal form.

A simple geometry for recording powder patterns is shown in Figure 2. The powdered sample is placed on the axis of a cylindrical chamber and a monochromatic X-ray beam enters the chamber through hole *C*. The undiffracted beam leaves through hole *B*, while X-rays diffracted by the sample will diverge from the sample as a set of hollow cones with half-angles 2θ, for all allowed Bragg angles θ. A rectangular strip of photographic emulsion is wrapped around the inside of the sample chamber so that the cones of diffracted radiation appear as circles (or parts of circles) on this film. Since the distance from the centre of hole *B* to the centre of hole *C*
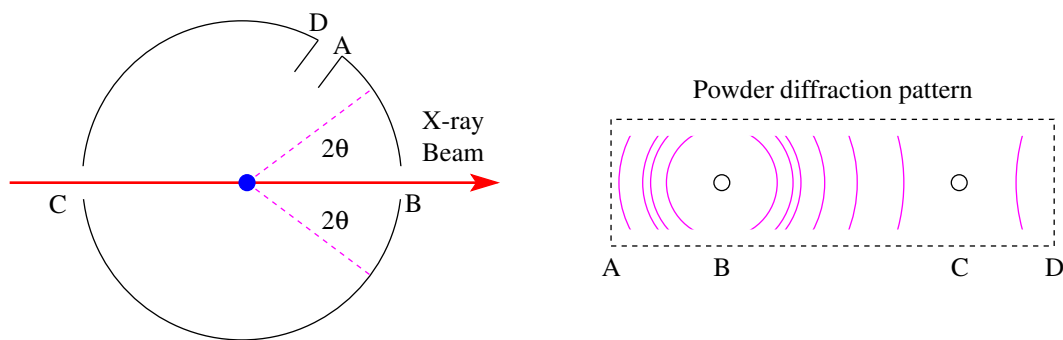
Figure 2: A simple geometry for recording powder patterns.

corresponds to an angle of $180°$ it is straightforward to determine the Bragg angles for all the lines visible in the powder diffraction pattern.

A copy of the diffraction pattern from powdered MgO is provided with the experiment. It was recorded with copper K_α radiation, and the radius of the diffraction chamber is marked on the print. Measure the Bragg angles for all the lines visible on this photograph, and use the information gained from the measurements on single crystal MgO to identify which lines correspond to reflections from the (2 0 0) and (4 0 0) planes.

5   Data Analysis

5.1   Alkali halide crystals

Calculate the values of the lattice constants for the alkali halides and compare these measurements with the accepted values. Do they agree within experimental error? Comment on the main sources of experimental error, and explain why the measurements made with the reflections from the (4 0 0) planes are expected to be more accurate than those made using the (2 0 0) planes.

Ion or Atom	Radius (Å)
F^-	1.23 ± 0.04
Cl^-	1.66 ± 0.08
Ne	1.6
Ar	1.9

Table 1: Accepted values for ionic radii. [3]

The accepted values for the ionic radii of the fluoride and chloride ions, and for the atomic radii of the noble gases neon and argon are given in Table 1. Use the information for F^- and Cl^- to derive values for the radii of the Li^+, Na^+, K^+ and Rb^+ ions, and again compare these calculated values with the accepted values. Try to explain the results in terms of the electronic structures of the ions.

5.2   Magnesium oxide powder pattern

Equations (1) and (2) can be combined to give equation (3)

$$h^2 + k^2 + l^2 = \frac{4a^2 \sin^2 \theta}{\lambda^2} \tag{3}$$

The lattice constant a can be found from the measurements of the Bragg angles for the (2 0 0) and (4 0 0) reflections, so it is possible to obtain a set of values for $h^2 + k^2 + l^2$ for each line in the powder diffraction pattern. (Note that h, k and l are always integers, so the sum of their squares must also be an integer.)

From the calculated values of $h^2 + k^2 + l^2$ it is possible to deduce Miller indices for each of the lines in the MgO powder pattern. For example, if $h^2 + k^2 + l^2$ is found to be 11, the only possible values of $(h\ k\ l)$ would be (3 1 1), (1 3 1), and (1 1 3), and for a system with a cubic structure these sets of planes would be equivalent, from the symmetry of the crystal.

Once you have found Miller indices for each of the lines on the powder pattern, see if you can identify a pattern in the numerical values for the indices. The pattern that is found is a characteristic of the particular cubic structure possessed by the crystals used in this experiment. (*The three common cubic structures are simple cubic, face-centred cubic, and body-centred cubic. Each produces a different pattern in the Miller indices found in a diffraction experiment*).

References

[1] J. R. Hook and H. E. Hall. *Solid State Physics*. 2nd edition.

[2] http://nobelprize.org/nobel_prizes/physics/laureates/1915/.

[3] V W Maslen. Crystal ionic radii. *Proceedings of the Physical Society*, 91(1):259–260, 1967.

Nuclear Magnetic Resonance in polystyrene, glycerine and teflon

1 Aim of the Experiment

The aims of this experiment are to observe Nuclear Magnetic Resonance of hydrogen and fluorine in liquids and solid samples. In addition, the g-factor of protons and fluorine will be determined.

2 Introduction

Nuclear Magnetic Resonance (NMR) is based on the fact that the nucleus of each element has a spin associated with it, I_S, which can be aligned in a magnetic field if I_S is non-zero. Application of a second appropriate magnetic field orthogonal to the first allows flipping the nuclear spins—this can be measured and yields the NMR signal.

The phenomenon of was developed by Felix Bloch and Edward Mills Purcell in 1945. In 1952, they were awarded the Nobel Prize in Physics "*for their development of new methods for nuclear magnetic precision measurements and discoveries in connection therewith*" [1].

Today, NMR spectroscopy is an important method to determine the structure of molecules in physics, chemistry and biology. In fact, it is so important that in 1991, Richard Ernst was awarded the Nobel Prize in Chemistry "*for his contributions to the development of the methodology of high resolution nuclear magnetic resonance (NMR) spectroscopy*" [2], followed in 2002 by another Nobel Prize in Chemistry to John Fenn and Koichi Tanaka "*for the development of methods for identification and structure analyses of biological macromolecules*" and Kurt Wthrich "*for his development of nuclear magnetic resonance spectroscopy for determining the three-dimensional structure of biological macromolecules in solution*" [3].

Moreover, spatially resolved NMR allows non-invasive and non-destructive imaging of flipping spins for clinical diagnosis. This is usually called Magnetic Resonance Imaging, or MRI, but it is effectively spatially resolved NMR. The invention of this method has revolutionized clinical imaging and diagnosis, and was rewarded, in 2003, with a Nobel Prize in Physiology or Medicine to Paul Lauterbur and Sir Peter Mansfield "*for their discoveries concerning magnetic resonance imaging*" [4]. It does not require ionizing radiation, as in x-ray imaging, and has a better resolution than ultrasound imaging.

This experiment demonstrates the basic phenomenon of NMR on liquid and solid samples and gives first insight into spectroscopic applications in chemistry and biology.

3 Theory

NMR is based on the magnetic properties of the nucleus. The nucleus of an atom is made up of protons and neutrons. Protons and neutrons are elementary particles classed as fermions (after the Italian physicist Enrico Fermi) with a half-integer spin. (Integer spin particles, on the other hand, are bosons, after Indian physicist Satyendra Bose.)

A nucleus that contains odd numbers of protons or neutrons has an intrinsic magnetic moment μ and angular momentum.

Atoms with a nucleus which possesses a nuclear spin I_S have a magnetic moment μ according the relation:

$$\mu = -q\mu_n I_S \tag{1}$$

where μ_n is the nuclear magneton ($\mu_n = 5.051 \times 10^{-27}$ J.T^{-1}) and g the gyromagnetic ratio (g-factor).
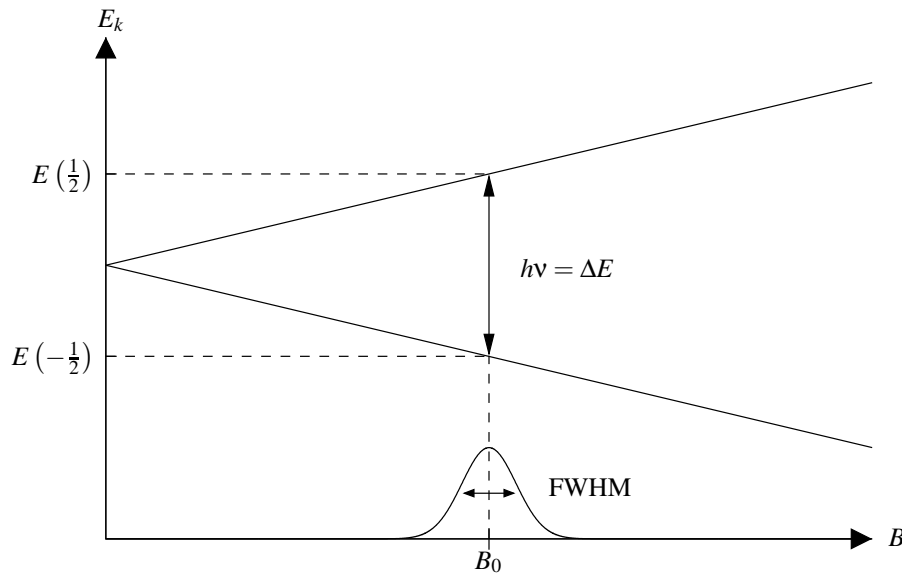
Figure 1: Energy splitting of a nucleus of $I_S = \frac{1}{2}$ in the magnetic field B_0. One line is observed in the NMR spectrum from which the full width at half maximum (FWHM) can be determined.

The nuclear magnets μ can orientate only in distinct directions in respect to a static magnetic field B_0. Each spin orientation corresponds to a particular energy level which is given by:

$$Ek = -g\mu_n B_0 k \tag{2}$$

with $k = -I_S, -I_S + 1, -I_S + 2, \ldots, I_S$

When a sample is placed in a static magnetic field B_0 the nuclear spins are distributed over the energy levels according the Boltzmann equation:

$$\frac{N_k}{N_{k+1}} = \exp\left(\frac{E_{k+1} - E_k}{k_B T}\right) \tag{3}$$

k_B: Boltzmann constant, N_k: Number of spins in level k

The spins may be excited to jump between adjacent energy levels E_k when the sample is exposed to a magnetic field B_1 of high frequency ν which is perpendicular to the static field B_0.

Therefore the frequency ν has to match exactly the energy spacing. This is called resonance ($\Delta k = 1$):

$$h\nu = E_{k+1} - E_k = -g\mu_n B_0 k \tag{4}$$

where h is Planck's constant, 6.626×10^{-34} J.s

4 Experimental Method

In this experiment the NMR signals of hydrogen nuclei (protons) and fluorine nuclei are observed. Both nuclei have a spin of $I_S = \frac{1}{2}$ and thus according to equation (2) they have only two possible orientations in respect to the magnetic field B_0 (Fig. 1). In order to detect the spin transitions between these two energy levels the sample is placed in an RF coil which is located between the gap of a homogeneous field B_0. Additionally, the static magnetic field B_0 is modulated at constant frequency ν by the use two modulation coils. Thus in this setup the resonance signal is observed in dependence of the magnetic field. Each time when the resonance condition (i.e. equation (4)) is fulfilled energy is absorbed from the RF coil due to the spin transitions. The NMR signal is proportional to the number of spins N in the sample. For example, in a hydrogen nucleus the resonance frequency in a magnetic field is about 42.5 MHz. The precise value depends on the chemical environment of the hydrogen atom as in addition to the external magnetic field B_0. The local internal field of electrons and nuclei in the near vicinity also acts on the hydrogen nucleus. This is called chemical shift and makes NMR to an important spectroscopic method for chemistry and biology.

The current I_{CC} through the 10 A coil is proportional to the magnetic field strength B_0:

$$B_0 = \alpha \times I_{CC} \tag{5}$$

Set up the experiment as shown schematically in Fig. 2

1. Set the oscilloscope to xy-mode.

2. Select fast sweep and set the modulation amplitude to a large value.

3. Set the value of the frequency to a maximum value.

4. Slowly enhance the HF amplitude and tune its frequency until the red LED lights up and a frequency of about 18.5 MHz is displayed.

5. Shift the O-ring of the water sample tube so that the sample will be located approximately in the center of the measuring chamber.

6. Carefully insert the sample tube into the measuring chamber. Take care with this procedure—if the sample tube is inserted at an angle with too much force, the RF coil can be damaged.

7. Set the power supply voltage to its maximum value, and slowly increase the current through the 10 A coils until an NMR signal appears on the screen of the oscilloscope.

8. Optimize the NMR signal by slowly varying the HF amplitude.

9. Shift the NMR signal to the center of oscilloscope screen by varying the magnet current or the frequency and adjust the modulation amplitude.

10. Adjust the phase of the signal until the signal of up and down sweep coincide.

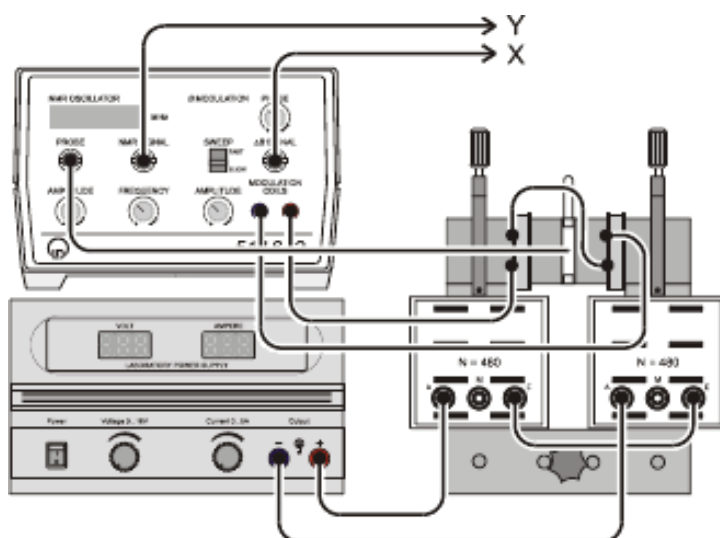11. Adjust the NMR oscillator frequency until the signal is centered on the CRO screen.

Figure 2: Electrical setup of the NMR experiment, schematically.

For around 10 different values of the current I_{CC} (ranging from 3 to 4 A), adjust the NMR amplitude and phase, and measure the resonnance frequency.

Repeat this procedure for glycerine, polytetrafluorethylene (PTFE) and polystyrene.

5 Analysis

Plot the resonance frequency as a function of current for water, and perform a straight line fit. Use equations (4) and (5) to determine α (the literature value for the gyromagnetic ratio for hydrogen is $g_H = 5.5857$), and plot the resonance frequency as a function of the magnetic field strength.

Plot the resonance frequency as a function of the magnetic field strength for both glycerine, PTFE and polystyrene, and perform a straight line fit.

Calculate the gyromagnetic ratio for the hydrogen in glycerine, and compare it to the literature value. Also calculate the gyromagnetic ratio for the fluorine and again compare it to the literature value.

References

[1] http://nobelprize.org/nobel_prizes/physics/laureates/1952/.

[2] http://nobelprize.org/nobel_prizes/chemistry/laureates/1991/.

[3] http://nobelprize.org/nobel_prizes/chemistry/laureates/2002/.

[4] http://nobelprize.org/nobel_prizes/medicine/laureates/2003/.

[5] LD Didactic GmbH. LD Physics Leaflet P6.5.3.1: Atomic and Nuclear Physics. Leyboldstrasse 1 . D-50354 Hürth, Germany.
http://leybold-didactic.de/literatur/hb/e/p6/p6531_e.pdf.