PHYS 340: Modern Physics Lab

Purdue University

Department of Physics

W. Lafayette IN 47907

Revised: Spring 2014

S. Savikhin

R. Reifenberger

M. Jones

M. Lister

D. Merrill

F. Wang / S. Jensen

PREFACE

The term “modern physics” is outdated. There is really nothing modern about the experiments you will perform since many of them were originally completed some 75-100 years ago. The lab PHYS 340 is a laboratory course focusing on important experiments in modern physics. The course is really a sequence of experiments designed to study the properties of electrons, photons and electron-photon interactions. Fundamental questions, raised in the late 1800's and early 1900's, centered on these important issues and led the way to the formulation of quantum physics.

During this course, you will have the opportunity to

- repeat some very important experiments and gain experience with a variety of experimental techniques,
- analyze and synthesize non-trivial experimental data, and
- gain experience in writing a lab report on the experimental work you have performed.

It is a pleasure to acknowledge the able and enthusiastic help of Carolyn Smith during the past three years. Her expert advice made it easy and fun to develop new experiments. Eric Dedrick and Josh Guffin worked during the summer months to test out the new experiments. One of us (RR) would also like to acknowledge many helpful conversations with Prof. Lazlo Gutay (Purdue), Prof. Terry Toepker (Xavier University) and Prof. John Hsieh (University of Arizona) during the initial phases of the upgrade. Their thoughtful comments provided a much-needed direction to the early work. Lastly, Prof. Andy Hirsch (Purdue) deserves much credit for helping to secure funds from the Purdue Administration to purchase new lab equipment for the lab.

S. Savikhin
R. Reifenberger
October, 2002

TABLE OF CONTENTS

Lab Procedures and Practices. 5
Notes on Analyzing Data. 21
Lab report – an example. 34
Getting Started with Computer Data Acquisition. 42

Experimental Write-ups

- Charge to Mass Ratio of Electrons. 51
- The Stefan-Boltzmann Constant. 60
- The Photoelectric Effect. 71
- Compton Scattering. 81
- The Franck-Hertz Experiment. 92
- Electronic Conduction in Solids. 104
- Light Interference. 119
- Electron Diffraction. 139

Lab Procedures and Practices

1. Notebook Keeping

One of the most important skills of an experimental physicist is keeping a good lab notebook. The details of laboratory notebook-keeping are patterned after standard notebook-keeping practices in professional physics research organizations. If these practices seem excessively legalistic, it is because they are, in part, to insure that notebook records of research and development will be legally binding to protect the inventor's claim to an original discovery. Countless legal contests have been decided by a properly kept, properly witnessed research notebook. Learn to do it correctly now and it will be of life-long value to you.

Type of Notebook — A Bound, "quadrille-lined" (coarsely gridded graph paper) computation notebook *Avery Dennison No. 43-648 or equivalent* is recommended for this course. The two important aspects are its bound nature, which makes it more lasting and prevents pages from being removed, and the quadrille ruling, which facilitates quick plotting of graphs and sketching of apparatus.

General Rules —
1. Every page of your notebook should be numbered.
2. All entries must be legible to others, written in ink and in serial order,
3. No blanks are to be left for filling in later (exception: Table of Contents as described below).
4. Any additional loose information sheets, such as ideas jotted down on separate paper at odd time, computer graphs or graphs on millimeter paper, occasional photocopies of pertinent tables or figures, etc., must be permanently affixed, preferably with glue or transparent tape. No erasures, white-outs or scribble-outs are allowed. Corrections are done by putting a line through the erroneous entry and writing the correct entry above or to the side. Large mistakes, *i.e.*, many lines or an incorrect drawing, can be marked by an "X" across the whole thing, with a dated comment on why it is X'd out.

Table of Contents —
1. The first few pages of the notebook should be reserved for a Table of Contents. At the very least it should contain the starting page number of each experiment. Greater detail would be useful in looking up material about which you or the instructor have questions.

Dates — All pages should be dated. This assures that the diary aspect of the notebook is maintained. It is also essential in cases where experimental priority needs to be established. Do not write on a page with an earlier date. Exceptions: i) if there is an interruption of thought or experiments, you can insert a comment on the old page giving directions to the continuation. ii) If you discover later that some material on an old page is wrong, you can line it out and give directions to a current page explaining why it is wrong. If you wish to add columns to an earlier table, you can photocopy the table (reduced in size if you wish), glue it to the current page (state the page from which it was copied), and write in your new columns. If you summarize previous results into neater form, be sure to state which pages are being summarized.

Units — All dimensioned quantities **must** have units associated with them.

Uncertainties and significant figures — Not only should a numerical value be given but also a justification for that value; *e.g.*, "the smallest division is 0.1 mm and I can interpolate to 1/4 division so error is about .03 mm". If there is a parallax error, include that as well. If there is a systematic error, note that too; *e.g.*, "end of tape was broken off so starting point was 2.5." Regarding significant figures, every number should be quoted to a value commensurate with its associated measurement error. A recorded value such as $(3.04175 \pm .021) \times 10^8$ m/s is poor form. These concepts are discussed in further detail later in this section.

Pre-Class Preparation —
- Refer to books you have already read (including lab handouts)
- Outline goal of experiment – make sure this is clear
- Outline the basic theory: derive equations, this is preferable to simply copying. You may use more than one source. If you cannot get the same results as the lab manual, consult with your TA or professor.
- Outline (roughly) experimental setup and procedures.

Note: You can cut and glue schematics or diagrams from lab handouts and paste them into your notebook, though drawing by yourself is preferable as the process of drawing will often force you to understand how something works.

Your instructors may want to examine your preparation before you start an experiment. *Pre-class preparation is one of the most important parts of the experiment.* In the real world of science, preparation and planning for an experiment take a large share of the time. Building and debugging the apparatus, data analysis, and writing the results take much of the rest of the time. Actual data taking often

occupies very little of the time. This is why Advanced Lab students are expected to devote a considerable amount of time to work done outside the lab.

In-class Lab Work — Do not collect a lot of data in meaningless tables. After you get some preliminary set of data – analyze it: calculate correlation coefficient(s), plot it if necessary. As soon as you get some results, graph them on the spot and make a rough calculation to see if your results are reasonable. Many people have spent hours collecting beautiful lists of worthless numbers because they omitted this step. Make sure things are going right before you continue too far. The preliminary graph need not be a work of art, but nonetheless should be included in your lab notebook. When you do your final analysis at home, you should re-draw the graph on good graph paper or do a computer plot, with (possibly) more suitable scales, better captions, etc. More about this later!

Tables — When it is possible, try to incorporate tables into your notebook. Good tables are both compact and easy to follow. They call attention to the numbers and relations between them. Head each column with a name and/or symbol. Don't forget to label the table columns with units. Tables are especially effective in collecting together the important results of an experiment. Don't be afraid to use them. However, avoid lengthy multi-page tables, which are better suited for computer file storage.

Sample Calculations — Demonstrate your method of data analysis by showing at least one complete sample calculation. The rest of the data may be analyzed a calculator or computer, but all of the results should be tabulated. The sample calculation should not be just a string of numbers and combinatorial signs but should give a description of what you are doing and why. The latter is especially important if there is more to be done than substituting into an expression derived in the introduction. You should indicate which data table your values came from so that the procedure can be quickly verified. This will be most helpful when you get unreasonable results.

Diary Function — A little before the end of each lab session you should write a brief diary or journal entry describing in a few sentences what has been done during the period.

End-of-Class Check — At the end of each lab session you are required to have your notebook entries briefly examined and initialed by your TA. Although time will not allow them to thoroughly review your work, certain helpful comments and suggestions may be made at this time. This is not simply a legalistic watch over what you are doing. In a sense it is proper research notebook procedure followed by all good research organizations in one form or another. In order to insure legality of the notebook priorities, notebook entries are signed and dated by an independent person or persons who understand the contents.

Final Write-up — This should be done before starting a new experiment, or at least in the week that you start the next experiment. It should be a summary of the experiment you just completed in not more than two pages. It should include your best value for the experimentally determined quantity and your best assessment of the uncertainty on this value.

There should be a comparison with the "accepted" value taken from handbooks or other sources. It should include a discussion of the appropriateness of the procedure as you see it, difficulties you experienced, sources or error and their magnitude, and your major source of error with possible ways to reduce it. It is here that you should apply any known corrections to your raw value, giving adequate justification for the corrections. You are offering the world an experimental value and you are now justifying its merit, accuracy and precision. It is frequently appropriate to suggest, based on your experience in doing the experiment, how one could improve on the experiment in the future.

Although the final write-up should not exceed two pages, it may make reference to notebook pages where more details can be found. If done properly, the final write-up paves the way toward easy production of a formal paper on the experiment. It will also be examined first and with the most scrutiny by graders who may not have time to examine every other detail of your notebook.

2. Notebook Problems

It is easy to list some common problems when trying to keep a good lab notebook. The items listed below are indicators that your efforts to succeed in this course may not be paying off.

- You spend no time writing in your notebook before starting a new lab.
- The entries in your lab notebook are essentially dated one week apart.
- You make no attempt to re-derive important equations in your notebook.
- You staple many graphs (or spreadsheets) onto one page of your notebook.
- No error estimates are included when you quote your final result.

\• There is no final write-up. **Remember**, you should analyze and organize everything in your notebook **before** writing a final report. Feel free to appear disorganized and make mistakes; but be prepared to reorganize your thoughts in a better way on the next page!

• You attach graphs (or data tables) to your notebook without giving them an identifying number, e.g. 'Graph 2' or 'Plot 4' or 'Table II'.

• You spend time copying data from 'loose' data sheets into your notebook. (Note: If your lab partner records data during a lab period while you adjust the

apparatus, we can make you a photocopy of the data before you leave the lab that can be pasted into your notebook.)

• You are overly concerned when your values do not match accepted results. In fact, when this happens, you actually have a real opportunity to learn something! By tracking down the source of the discrepancy, you will learn how to design effective experiments.

In most instances, the habit of keeping a good notebook is just as important as the acquisition of good data. In fact, these two activities are closely related.

.

Notebook Grading Guidelines

Formatting

All pages must be numbered. Entries must be written in ink and dated. Information should be entered in chronological order without blank pages left for filling in later. Graphs and printouts must be permanently fixed to a blank page in the appropriate location. Each experiment should start a new page and notebook sections must be titled as outlined in this guide.

Prelab

Theory: Theory for the experiment must be summarized in paragraph form using equations where appropriate. Lists of equations are not acceptable. The summary should relate the quantity of interest to the physical measurements taken in the lab.

Procedure: The experimental setup must be described with references to appropriate figures if necessary for clarity. The summary must be specific in detailing which values were chosen for all independent variables. A diagram of the experiment must accompany the description of procedure.

Data: All data should be tabulated in an orderly fashion. For large data sets taken by computer, it is sufficient to record the location (file name, email, etc.) of the data. All measurements must have accompanying units. Experimental uncertainties are to be quantified and justified.

Postlab:

Analysis: A partial analysis can and should be done after the first day of an experiment to ensure that data have been correctly recorded and procedures have been correctly followed. A complete analysis must follow the final day's data collection. Analysis must present the results including the related uncertainties. A sample calculation of results must be recorded. Sources of uncertainty should be

discussed. Raw data must be presented in graphical form unless inappropriate. A discussion of final uncertainty calculations should be included.

Final Write-up: The experiment should be briefly summarized, including experimental goals and methods. Results must be discussed and any deviations from expected outcomes must be addressed. A discussion of how measurement uncertainties affected the results should be included. Any statements made must be justified with logical arguments and experimental observations. Improvements to the experiment may also be suggested. The implications of the results should also be addressed.

Sample Notebook Grading Rubric

1st Lab		__/50pt
Formatting	__/5pt	
Theory	__/10pt	
Procedure w/ Experimental Diagram	__/10pt	
Data	__/5pt	
Analysis	__/10pt	
Final Write-up	__/10pt	
2nd Lab		__/50pt
Formatting	__/5pt	
Theory	__/10pt	
Procedure w/ Experimental Diagram	__/10pt	
Data	__/5pt	
Analysis	__/10pt	
Final Write-up	__/10pt	
Late Assignment Penalty	___ school days late x 2pts/day = -- __ **pt**	
Total points		**__/100pt**

Comments:

3. Formal Paper Requirements

All students are required to submit two or more formal lab reports (see your syllabus for details). These formal reports should resemble scientific papers to be submitted for publication. Examples of such papers are available in the Advanced Lab file at the main desk in the Physics Library. In addition, you are encouraged to read a few published papers in the *American Journal of Physics*. It is important that you gain an understanding of the contents and style of these papers. When in doubt, you should follow rules for manuscripts submitted to the *American Journal of Physics*. A copy of these rules is also available in the Physics Library.

The formal paper should have the following items in the order shown:

I. Title page containing:
   1. Title
   2. Authors' name
   3. Abstract

II. Text containing divisions such as:
   1. Introduction
   2. Theory
   3. Experimental apparatus and procedure
   4. Data, analysis, results and uncertainties
   5. Discussion of results
   6. Conclusion
   7. Acknowledgments

III. List of references in the required order and style

IV. Additional information containing:
   1. Appendices (if necessary) containing detailed theoretical derivations, error analysis, etc.
   2. Tables with table captions
   3. Figures and captions

Formatting Guidelines: Here are some suggested text formatting guidelines. Note that MS-Word and LaTeX templates are available from your TA and/or the class website.

- **Margins:** 1" on all sides
- **Font:** consistent throughout: Garamond (preferred) or Times New Roman
- **Body Text**: Size—12pt; Alignment—Justify; Spacing—1.5 space; Body Text includes the Abstract but does not include section headings.
- **Report Title:** Size—28pt; Alignment—Center; Followed by a line running across the page.

- **Author:** Size—14pt; Alignment—Left; Text Decoration—Bold, Italics; Do not give your address or school affiliation.

- **Abstract:** Indent the text by 0.5" (~5 spaces) on both the left and right side.

- **Major Section Headings:** Size—18pt; Alignment—Left; Text Decoration—Bold, Small Caps.

- **Minor Section Headings**: Size—14pt; Alignment—Left; Text Decoration—Bold.

Grammar: Correct English grammar and spelling must be used throughout the report. The use of commas, colons, semi-colons, dashes, and parentheses should be generally correct so as to not distract the reader. Reports should have few or no spelling mistakes. The past tense should be used for most of the report.

Page Limits: Although there is no page limit on your formal papers, they should be complete, organized and concise. In the real world of science, many journals have page limits and others require you to pay page charges (more pages = more $ out of your research grant!). Use Strunk and White's *Element of Style* to learn how to be concise.

References: References are listed at the end of the paper in the order in which their citations appear in the body text. Reference "1." will be the first citation in the body text, etc. References are cited in the text as "[1]" after the cited information.

Equations: Equations should be well typeset. Use a typesetting program like LaTeX or an equation editor like MS Equation Editor, Maple, Mathematica, or Wolfram Alpha to write your equations; then paste them into your report. Equations do not need to be numbered unless you want to refer to them in the text.

4. Formal Lab Report Section Contents

Abstract

Summarize your lab report in one paragraph. Your abstract should be limited to roughly 250 words. It should contain the essence of your paper, including the major numerical results (with uncertainties). Give a sentence or two of background, summarize your procedure, and state your results. Be sure to state any inferences you might make from your results. The purpose of an Abstract is to aid the reader to decide whether the paper is relevant to their research interests and is worth reading. Write this section last, after you know exactly what is in your report.

Introduction

Introduce the subject matter, give some context and then introduce your specific experiment. Define "what" and show "why", but leave "how" to the main text.

Provide a brief historical and conceptual background with a set of helpful references to help your reader dig deeper. Emphasize the significance of the subject matter in science and its relation to your present work. Conclude your introduction with a brief outline of your paper so that your reader knows what to expect and could skip and select more easily, especially if it is a long paper.

Theory

Experiments are meant to 1) test the predictions or extend the realm of predictions of theories, or 2) measure indirectly some physical quantities by way of theoretical relations between quantities that can be measured directly. Purpose 1) may be accomplished by comparing the result of 2) with other known or verified quantities or values. Display the relevant theory. Express what is to be tested or measured in a working equation or a set of working equations. Derive these equations. Relate the equations to the variables or parameters to be measured. Note that all equations must be punctuated like sentences, i.e., they must end with commas or periods. Use conceptual or schematic diagrams to help clarify your derivations. Cite references to any facts or arguments you have adopted. Move the derivation to an Appendix if it is too long.

Experimental apparatus and procedure

Describe the means through which each of the variables and parameters you have identified in the working equation(s) is to be measured. Procedure should be summarized in paragraph form. Bullet-point or chronological listing of steps (i.e., Then . . . Next . . .etc.") is not acceptable. The summary must be specific in detailing which values were chosen for all independent variables. Use geometrically and conceptually accurate diagrams and figures to help your written description. Give all necessary quantitative information with proper units and reliable uncertainties. Use tabular form if appropriate. Cite references for each piece of special equipment, so a reader could procure the same if so desired. Include any references to design schematics or tables of parts. Be sure to mention any anomalies in the way you took your data. (e.g., you had to repeat a data set because of discovered calibration issues.)

Data, analysis, results and uncertainties

Describe what you did with the data and present the results. Analyze the uncertainties, present any sources of error, and explain how you arrived at the final uncertainty in your measurement. You might briefly mention any anomalies in the data, but save the discussion of their implications for the Discussion section. Use graphics to make your work more easily understood in view of what you have presented in the previous sections. Tabulate only important information that cannot be presented graphically. There is no need to waste your time to reproduce endless tables of numbers in a formal lab report! Avoid this procedure at all costs. Short tables of final results are acceptable, especially if you think they might be of use to other researchers. In real science journals, long tables are typically published only in electronic form, as computer-readable files. Move detailed analysis and sample calculations into appendices. Be very careful with significant figures and uncertainties.

Discussion of results

Discuss the implications of your results. Discuss how the sources of error affected your data and give justification for your statements. Compare your results to accepted or expected values. Evaluate any discrepancy in terms of accuracy and precision of your measurements. This is where you really get to show how much you have thought about your results and what they mean. Propose plausible consequences of your results on the theory. Suggest possible improvements for the next generation of experiments.

Conclusion

The conclusions section should summarize the results and their implications in the context of the experiment topic. . Be realistic here and don't make extraordinary claims. No new information can be presented that didn't appear earlier in the report.

Acknowledgment

Give credit to institutions and individuals that made your work possible and successful; e.g., who provided you with the necessary equipment, the setting up and execution of the experiment, discussions, and encouragement. Normally, your lab-mate and you should be co-authors of your formal report, but for this course, each writes his or her own paper. So, do not be afraid to acknowledge your lab-mate. Please remember, however, that this is a scientific paper: acknowledging your dog for keeping you company while writing up the report is not acceptable. Here is an example of an acknowledgment:

Acknowledgment: I would like to thank Prof. J. Hsieh from the Physics Department at the University of Arizona for providing a copy of the University of Arizona's advanced lab manual. Much of the material found in Secs. 4-6 of the PHYS 340 document 'Procedures and Practices' were adapted from the write-up which Prof. Hsieh provided.

References

The citing of references serves two purposes. First, it's a good way to give credit to the workers who came before you, and second, it provides a valuable lead to workers who come after you. Many journals have their own specific format for citing references to published literature, so considerable variability can be found from journal to journal. For many physics-related journals, an appropriate format for citing references is as follows:

Journal article:

[citation number]. Author1, Author2, Author3, Journal's abbreviated name, Journal volume (usually in bold face print), page (year). For example, an acceptable journal reference might look like this:

1. J. Jones, K.L. Mack and K.B. Stuart, Phys. Rev. **77**, 2084 (1962).

Book:

[citation number]. Author1, Author2, BookTitle (usually in italics), publisher, city where published, year when published, page numbers (if appropriate). An acceptable book citation might look like this:
2. S. Datta, *Electronic Transport in Mesoscopic Systems*, Cambridge University Press, Cambridge, 1995, pgs. 20-37.

In the case of URLs, there is no need to underline them.

Appendices

Details that require extensive space to derive or issues that might break the logical flow of the paper belong in the appendices. Appendices MUST be referred to in the main text of the paper.

Figures and tables.

Graphs of data must be neatly drawn, preferably using a computer. Make sure the axis labels are easy to read and in a large enough font to be legible. Examples of acceptable and unacceptable graphing procedures are given in Appendix A. Graphs of few (i.e., < 20) data points must have error bars on all data points. A line of best fit should be included on all linear data. Figures and tables are traditionally attached to the manuscript on separate pages, one figure/table per page to make future copying easier. You may also choose to insert figures/tables into main text at the appropriate place as it is done in lab manuals. Tables should be used only where graphs would be inappropriate. Place captions on the same page as the corresponding figure or table. Each caption should begin with the figure/table number. They should be numbered in the order that they first appear in the text. All figures and tables must be referred at least once in the body text. (e.g. "see Figure 1")

5. What is a Writing-Emphasis Course?

The writing-emphasis portion of *PHYS 340* is inherent in the required formal papers and notebook keeping that you will perform during the semester.

The formal papers will be judged in two ways with approximately equal weight: i) the quality of the scientific content and ii) quality and clarity of the writing. The scientific content will be judged by the quality and appropriateness of the theory, experimental procedures, data analysis, and error analysis. The writing will be judged by the organization into a proper format, acceptable to a scientific journal (like *American Journal of Physics*), by details such as spelling, grammar, sentence structure, figures, tables, etc. but most of all by whether the student has **clearly** and **concisely** presented the information and arguments required to interpret and report the results of the scientific work. It is important to realize that the most important aspects of scientific writing are conciseness and clarity. Other researchers will greatly appreciate your papers if they are short, to the point, and easy to follow. Flowery prose and lengthy digressions are to be avoided at all costs.

To be honest, you will be repeating some ground-breaking experiments that have been performed in the past. Your research will not be original and you should make no attempt to claim that it is. It is a fair question to ask 'Why should I repeat old experiments? ' Repeating classic experiments is an excellent way to better place time-tested results into a context that has meaning for you. In addition, you will learn some useful experimental techniques that will be of value to you in the future, especially if you maintain an interest in highly technical subjects.

The writing of a good formal report and the keeping a good notebook is perhaps the most difficult aspect of this course. It requires a careful, thoughtful approach that is difficult to teach, and is best learned through experience. You will always be able to obtain some experimental data, but properly analyzing the data and understanding the context in which the data should be viewed is challenging. It is important for you to realize that performing an excellent experiment can only be achieved by a long process that requires you to make many decisions along the way. Each decision by itself may seem insignificant, but a thoughtful sequence of decisions will lead to a good notebook and an excellent lab report. The only way to learn this process and become proficient in it is to immerse yourself in the task at hand. Those willing to commit the time to PHYS 340 will learn this valuable research skill.

Grades in the notebook(s) will not intentionally be divided into separate parts. Quality of the writing (clarity, grammar, sentence structure, and spelling), however, will account for a non-negligible portion of your notebook grade. Write so that if you opened your notebook 30 years from now, you could still understand what you did. The text, *Elements of Style* by Strunk and White, is recommended for the purpose of learning how to write **clearly** and **concisely**. Another excellent (and easy to read!) book is *The Craft of Scientific Writing* by Michael Alley.

6. Use of Spreadsheet Programs

This course will require you to extensively analyze data and fit data to theoretical expressions. The use of computer spreadsheet programs with graphing capabilities like Excel is invaluable in this regard. Tutorials in the use of spreadsheet programs will not be included in this course, since this skill is often taught in a high school curriculum. If you have any questions about how to use spreadsheet programs to analyze or plot your data, please see your T.A. or professor as soon as possible.

7. Some Useful References Worth Consulting

1. Philip R. Bevington and D. Robinson, *Data Reduction and Error Analysis for the Physical Sciences 3rd Ed.* , McGraw-Hill (2002).
2. Louis Lyons, *A Practical Guide to Data Analysis for Physical Science Students*, Cambridge University Press, Cambridge (1991).
3. John R. Taylor, *An Introduction to Error Analysis*, 2nd Edition, University Science Books; Sausalito, CA, (1982).

4. G.L. Squires, *Practical Physics*, 3rd edition, Cambridge University Press, Cambridge (1985).
5. G.P. Harnwell and J.J. Livingood, *Experimental Atomic Physics*, McGraw-Hill, NY (1933).
6. A.C. Melissinos, *Experiments in Modern Physics*, Academic Press, New York, 1966.
7. D.C. Baird, *EXPERIMENTATION: An Introduction to Measurement Theory and Experiment Design*, 3rd edition, Prentice Hall, New Jersey (1995).
8. M. Alley, *The Craft of Scientific Writing 3rd Ed*., Springer (1996)

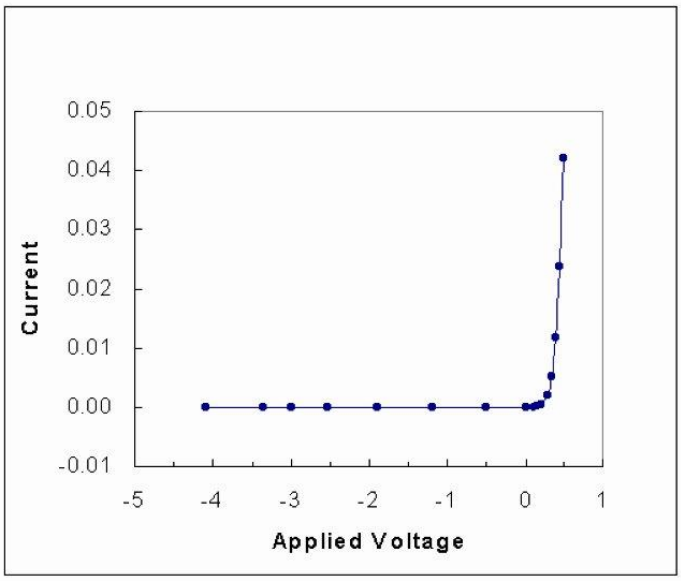
Appendix A: PHYS 340 Graphing Standards

A graphical comparison between data and theory is a very convincing way to demonstrate that you have correctly taken a set of data and that you understand its significance. Sloppy graphs casts doubt on your ability to carefully take data and analyze it correctly. In addition, given the capabilities of today's spreadsheet programs for graphing, there is really no excuse for a poorly drawn graph.

Figure 1 shows examples of an unacceptable and an acceptable graph. The unacceptable graph lacks a graph title, does not have units on the x and y axis, uses too small of a symbol to represent the experimental data, does not include error bars on the data, and does not indicate the meaning/origin of the line through the data. Formal reports containing such untidy graphs will be marked down accordingly.

In the case when the number of data points is very large and/or error levels are small compared to data values, error bars may be omitted. In this case, the scattering of data points on the graph (statistical 'noise') provides a visual measure of the error level. Any additional error should be, however, clearly mentioned in the text or figure caption (see Fig. 2 for example).

a) Unacceptable example of a PHYS 342L Graph:



b) Acceptable example of a PHYS 342L Graph:

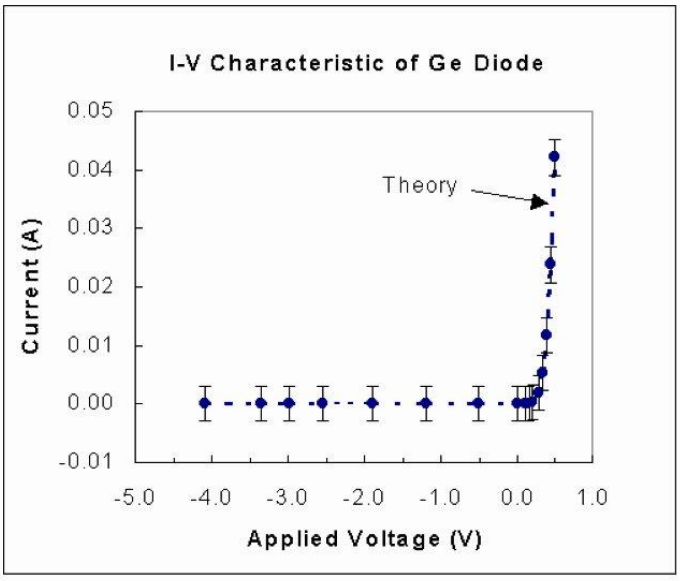


Figure 1: Examples of (a) unacceptable and (b) acceptable graphs.

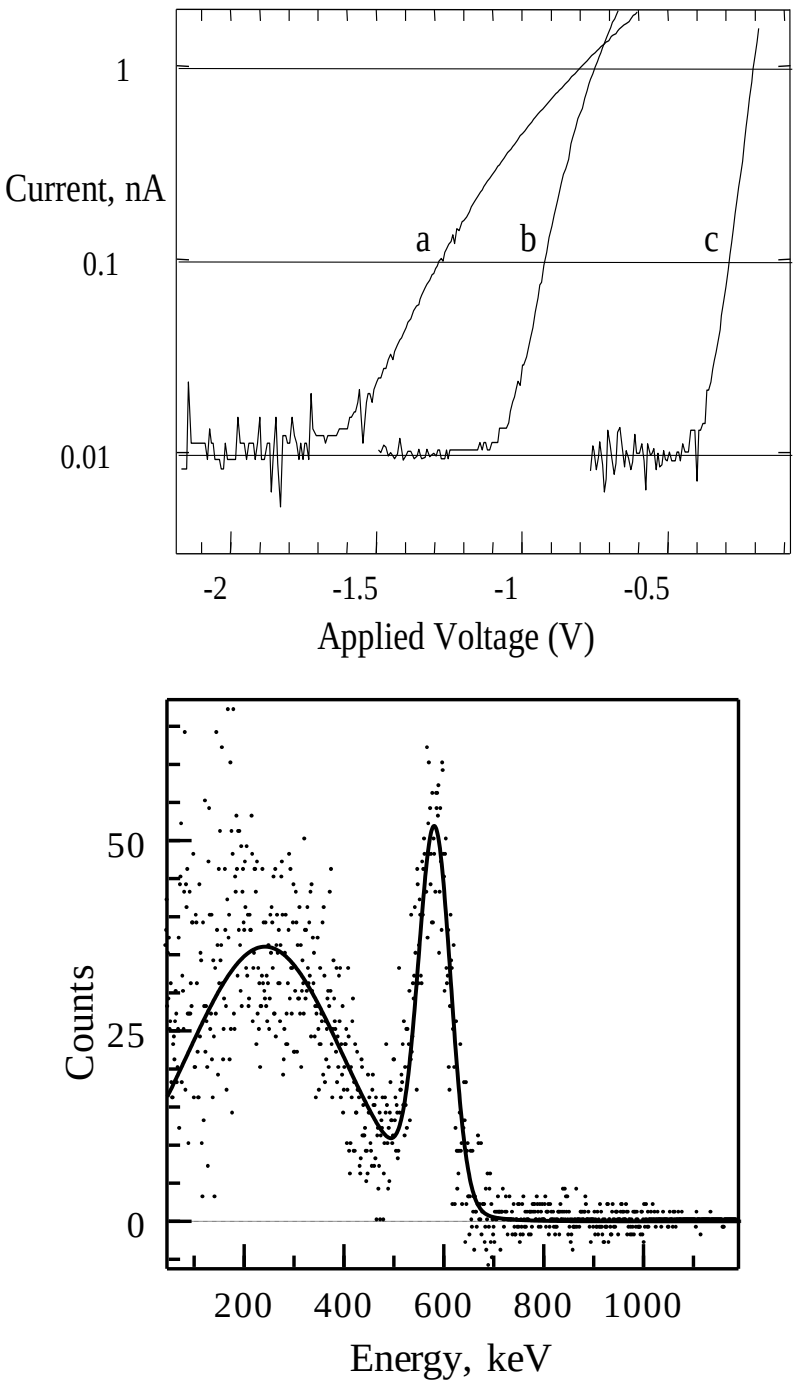


Figure 2. Examples of acceptable graphs containing more than 100 data points. In the first graph, multiple curves are plotted on one graph. In this case, each curve is denoted by a letter which must be explained briefly in the figure caption. In this particular example, the data points comprising each curve are so closely spaced, they appear as a continuous line. Due to the large number of data points, 'noise' in the measurement in the second graph is visible without the need to include individual error bars.

Notes on Analyzing Data

A major aspect of experimental physics (and science in general) is the measurement of physical quantities and analysis of experimentally obtained data. While there are a lot of books devoted to this problem, in the next paragraphs we will summarize some of the important ideas that will be needed to successfully analyze data acquired in PHYS 340. Students are advised to consult one or more of the references listed in the previous section for more detailed discussions of the topic.

1. The importance of estimating errors.

Suppose you are asked to measure the length of a piece of notebook paper. This seems pretty straightforward. You grab a ruler and proceed with a measurement. The ruler shows 276 mm. Does it mean the length is 276.0000 mm? Most probably not. Why? Because the distance between the neighboring marks on your ruler is 1 mm. By saying 276 mm you cannot exclude, for example, length 276.2 or 259.9 mm. Thus any time you use a measuring device you assume a particular precision (or error) in your measurement. In this case it is probably ~0.5 mm, as the distance between the closest marks is 1 mm. The result of the measurement is not just the length of the paper but also the error of this measurement: (276.0±0.5) mm. In a scientific experiment, both parts of measurement are important. Suppose you measure the length of the next sheet of paper to be (275.5±0.5) mm. Within the error of your measurement these two sheets of paper have the same length.

2. Precision (or Accuracy) of a Measurement.

Here's what's meant by absolute uncertainty and relative uncertainty:

absolute uncertainty: 27.6 ± 0.1

relative uncertainty: $\pm \dfrac{0.1}{27.6} = \pm 0.003623188$

All those digits don't mean much when calculating the relative uncertainty, so round off to ± 0.004, or, expressed as a percent, ± 0.4%.

3. Combining Uncertainties.

Suppose that you measure two quantities A and B. Suppose you measure *A* to an accuracy of $\pm\delta A$ and *B* to an accuracy of $\pm\delta B$.

How do you algebraically combine these uncertainties?

a) When adding:

$$(A \pm \delta A) + (B \pm \delta B) = (A + B) + ?$$

there are four possibilities:

$(A + \delta A) + (B + \delta B) = (A + B) + (\delta A + \delta B)$
$(A + \delta A) + (B - \delta B) = (A + B) + (\delta A - \delta B)$
$(A - \delta A) + (B + \delta B) = (A + B) - (\delta A - \delta B)$
$(A - \delta A) + (B - \delta B) = (A + B) - (\delta A + \delta B).$

Clearly, the worst case scenario will be

$$(A+B) \pm (\delta A+\delta B) \quad . \tag{1}$$

This is often called the case of *correlated errors*.

b) When subtracting:

$$(A\pm\delta A) - (B\pm\delta B) = ?$$

Again consider four cases. From above, it should be obvious that the worst case (correlated error) will be given by

$$(A-B) \pm (\delta A+\delta B) \tag{2}$$

c) When multiplying, the correlated (worst case) scenario is given by

$$(A \pm \delta A) \times (B \pm \delta B) = AB + A(\pm\delta B) + B(\pm\delta A) + \underbrace{(\pm\delta A)(\pm\delta B)}_{small, neglect}$$

$$\approx AB \pm (A\delta B + B\delta A) \tag{3}$$

$$\approx AB\left[1 \pm \left(\frac{\delta B}{B} + \frac{\delta A}{A}\right)\right]$$

d) When dividing

$$\frac{A \pm \delta A}{B \pm \delta B} = ?$$

After some algebra, you find that

$$\frac{A \pm \delta A}{B \pm \delta B} \approx \frac{A}{B}\left[1 \pm \left(\frac{\delta A}{A} + \frac{\delta B}{B}\right)\right] \quad . \tag{4}$$

Remember:

- **relative** uncertainties add when multiplying or dividing.
- **absolute** uncertainties add when adding or subtracting

4. Systematic and random errors.

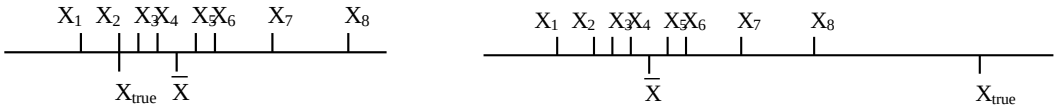


Figure 1: Spread in the measurement of some quantity x in the absence of systematic error (left) and presence of systematic error (right).

If error in your measurements is random, then the average value should be close to the actual value. In the case of systematic error, that is not true. This situation may occur when, for example, using a clock that is running slow to measure some time period. Random errors are inevitable, while systematic errors can be taken into account or eliminated.

5. Average value and standard deviation.

In order to decrease the influence of random error multiple measurements x_i are taken and averaged:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (5)$$

How close is this average value $\bar{x}$ to the actual value $\bar{X}$? If we have a set of measurements we can find an average error for a single measurement. The commonly accepted value to characterize error is called *standard deviation* σ, or *root mean square* (rms):

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2 \qquad (6)$$

Since the actual value $\bar{X}$ is usually unknown, we must use $\bar{x}$ instead. It can be shown [1] that in this case:

$$s^2 \approx \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \qquad (7)$$

The value s characterizes the error in a *single* measurement of value $\bar{X}$ and is a good approximation to σ. If we take several measurements of the same value *x* and average them, the resulting value $\bar{x}$ must on average be closer to the actual value $\bar{X}$

than a single measurement. It can be shown [1] that standard deviation σ_n for the average value $\bar{x}$ of *n* measurements is:

$$\sigma_n = \frac{s}{\sqrt{n}}. \qquad (8)$$

This quantity is often referred to as *the standard error of the mean*. It implies that if every measurement has a random, uncorrelated error, you will need to make four times as many measurements to double the precision of your mean value.

Distribution of measurements

A series of measurements may be represented as a histogram (Fig. 2). It is usually very difficult to see any trends after taking just a few data points.

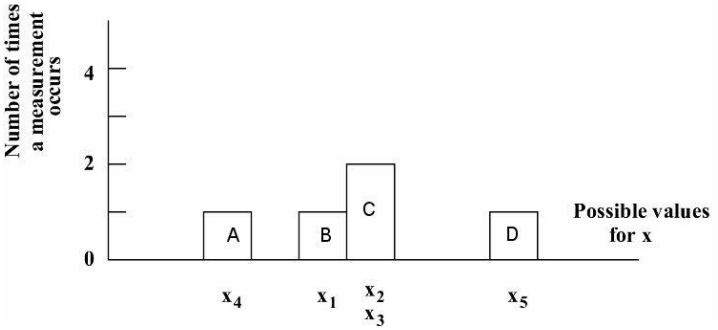


Figure 2. A simple histogram after taking just five data points (n=5). There was only one data point falling into range of *x* marked as A, B and D, and two measurements in region C.

If you make more measurements and use smaller bins, you'll eventually get a data set when analyzed using a histogram that might look like this.

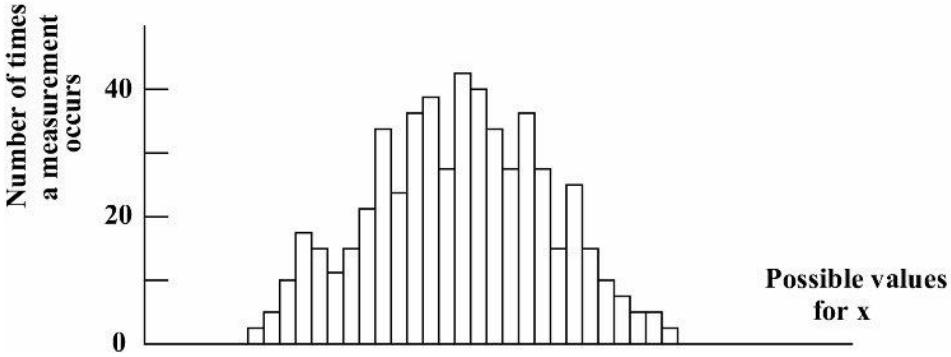


Figure 3. A histogram after taking hundreds of measurements.

In a limit of large *n* the distribution of data is given by continuous *distribution function f(x),* so that *f(x)dx* is the probability that a single measurement taken at

random will lie in the small interval between x and $x+dx$. The average value for the data set can be then found as:

$$\langle x \rangle = \int_{-\infty}^{\infty} x \cdot f(x) dx , \qquad (9)$$

and the standard deviation of the data set is defined as:

$$\sigma^2 = \langle x^2 \rangle = \int_{-\infty}^{\infty} (x - \overline{X})^2 f(x) dx \approx \int_{-\infty}^{\infty} \left(x - \langle x \rangle\right)^2 f(x) dx . \qquad (10)$$

In many cases, the error distribution function is well described by a Gaussian (also called a *normal* distribution) (see Fig. 4) given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\overline{X})^2}{2\sigma^2}} \qquad (11)$$

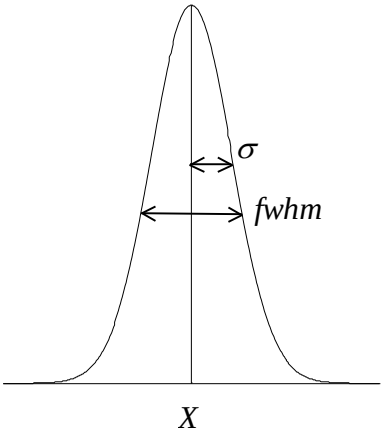


Figure 4. Gaussian distribution function.

The standard deviation σ for Gaussian distribution can be also expressed as:

$$\sigma = \frac{fwhm}{2\sqrt{2\ln(2)}} \approx 0.425 \times fwhm \qquad (12)$$

where *fwhm* stands for full width at half maximum, which can be easily estimated graphically.

Suppose now that you perform a single measurement of a quantity that results in a value x, and you also know that the standard deviation of this measurement is σ. Since the Gaussian distribution is a continuous function which approaches zero only

at infinity, in principle your measured value x can lie anywhere between -∞ and +∞. What is the probability that your measured value x lies within an interval σ of the actual value X that we are trying to measure? Since $f(x)dx$ is the probability of measuring a value between x and $x+dx$, the probability of measuring x between $\overline{X}$ - σ and $\overline{X}$ +σ is given by the integral $\int_{X-\sigma}^{X+\sigma} f(x)dx$. Substituting the functional form of the normal distribution (Eq. 11) for $f(x)$ gives a probability of 0.68.

Using this method, you can show that the probability that your measurement of x lies between

$\overline{X} \pm \sigma$ is 68%
$\overline{X} \pm 2\sigma$ is 95%
$\overline{X} \pm 3\sigma$ is 99.7%
$\overline{X} \pm 4\sigma$ is 99.994% .

Therefore, in a situation where your measurement errors are uniform and random, and there is no systematic error, there is very little chance of measuring a value that is several σ away from the actual value. Such measurements are often called *outliers*, and usually warrant further investigation.

6. Combining σ's.

Let us now go back to the case when we need to add two values A and B, where the standard deviations are σ_A and σ_B, correspondingly. What would be the standard deviation of the sum σ_{A+B}? As we already know, the errors should be added for the worst case scenario (Eq. 1). However**, if errors in *A* and *B* are random and mutually uncorrelated,** they tend to cancel to some extent, as there is a 50% probability that they have different signs in one measurement set. It can be shown, that the standard deviation of the sum (or difference) is:

$$\sigma_{A\pm B} = \sqrt{\sigma_A^2 + \sigma_B^2} \qquad \text{(uncorrelated errors)} \quad . \qquad (13)$$

Notice that $\sigma_{A\pm B} \le \sigma_A+\sigma_B$. If $\sigma_A=\sigma_B=\sigma$, then $\sigma_{A\pm B} = \sigma\sqrt{2}$.

Similarily, the *relative* standard deviation σ_C/C for product (or ratio) of A and B is:

$$\frac{\sigma_C}{C} = \sqrt{\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2} \qquad \text{(uncorrelated errors),} \quad (14)$$

where $C=AB$ or $C=A/B$.

Note, that this is true only if errors are *uncorrelated* and *not systematic*.

7. A simple example

Suppose you need to evaluate the charge to mass ratio of an electron. This is to be done using the following equations

$$\frac{q}{m} = \frac{2V}{B^2 R^2}, \tag{15}$$

where $B=kI$. How would you analyze the error in $\frac{q}{m}$?

Suppose you measure I, V and k as follows:

$I=(1.4\pm0.1)\ A$
$V=(140\pm2)\ V$
$k=coil\ constant=(7.5\pm0.5)\times10^{-4}\ T/A$

You also measure R, the radius of the electron's orbit, by measuring its diameter D. Since the smallest marking on the ruler is 1 mm and you have to determine the positions of both sides of the electron orbit, the precision of such a measurement is not better than 2mm = 0.2 cm. To reduce random error you may want to take several measurements. Suppose you make three measurements of D:

$D_1=6.0\pm0.2$ cm	$R_1=D_1/2=3.0\pm0.1$ cm
$D_2=5.8\pm0.2$ cm	$R_2=2.9\pm0.1$ cm
$D_3=5.7\pm0.2$ cm	$R_3=2.85\pm0.1$ cm

Using Eqs. 5 and 7, the average value of R and the standard deviation σ for one measurement will be

$$\bar{R} = \frac{3.0+2.9+2.85}{3} = 2.92\,\text{cm}$$

$$s = \frac{\sqrt{0.08^2+0.02^2+0.07^2}}{\sqrt{2}} = 0.08$$

According to Eq. 8, for the average of 3 measurements the standard deviation σ_3 will be

$$\sigma_3 = \frac{0.08}{\sqrt{3}} \approx 0.05$$

However, the ruler we use has a precision of only ~0.1 cm. Using Eq. 13 we can account for both errors:

$$\Delta R = \sqrt{\sigma_3^2 + 0.1^2} \approx 0.11\,\text{cm}$$

and we have

$R=$ (2.92±0.11) cm

Now we calculate:

$$\frac{q}{m} = \frac{2V}{B^2R^2} = \frac{2V}{k^2I^2R^2}$$
$$\frac{q}{m} = \frac{2(140 \pm 2)}{\left[(7.5 \pm 0.5) \times 10^{-4}\right]^2 \left[1.4 \pm 0.1\right]^2 \left[0.0292 \pm 0.0011\right]^2} \quad (16)$$

Omitting errors we get the value of $\frac{q}{m} = 2.98 \times 10^{11} C/kg$

Using Eq. 14 we can write:

$$\frac{\sigma_{q/m}}{q/m} = \sqrt{\left(\frac{\sigma_V}{V}\right)^2 + \left(\frac{2\sigma_k}{k}\right)^2 + \left(\frac{2\sigma_I}{I}\right)^2 + \left(\frac{2\sigma_R}{R}\right)^2} \quad (17)$$

Note the factors "2" in the equation above. These stem from the fact that corresponding values are squared in Eq. 16, i.e. we have products $k \times k$, $I \times I$ and $R \times R$. Since $k \times k$ is a product of two *correlated* values, we must use Eq. 3 – the relative errors simply add up.

$$\sigma_{q/m} = 2.98 \times 10^{11} \sqrt{\left(\frac{2}{140}\right)^2 + \left(\frac{2 \times 0.5}{7.5}\right)^2 + \left(\frac{2 \times 0.1}{1.4}\right)^2 + \left(\frac{2 \times 0.0011}{0.0292}\right)^2}$$
$$\sigma_{q/m} = 0.63 \times 10^{11} C/kg$$

And the final result for the ratio can be written as:

$$\frac{q}{m} = (2.98 \pm 0.63) \times 10^{11} C/kg \quad \text{(uncorrelated errors).}$$

Here we used Eq. 14 since values *V, k, I* and *R* are uncorrelated. If we assume, just for example, a correlated case, we must use Eq. 3 and 4 – i.e. add relative errors:

$$\frac{\sigma_{q/n}}{q/n} = \left(\frac{\sigma_V}{V}\right) + \left(\frac{2\sigma_k}{k}\right) + \left(\frac{2\sigma_I}{I}\right) + \left(\frac{2\sigma_R}{R}\right)$$

and $\sigma_{q/m} = 1.10 \times 10^{11} C/kg$ (correlated, worst case errors)

The accepted value is 1.76×10^{11} C/kg. It's easy to make a simple plot including error bars to graphically illustrate this result.

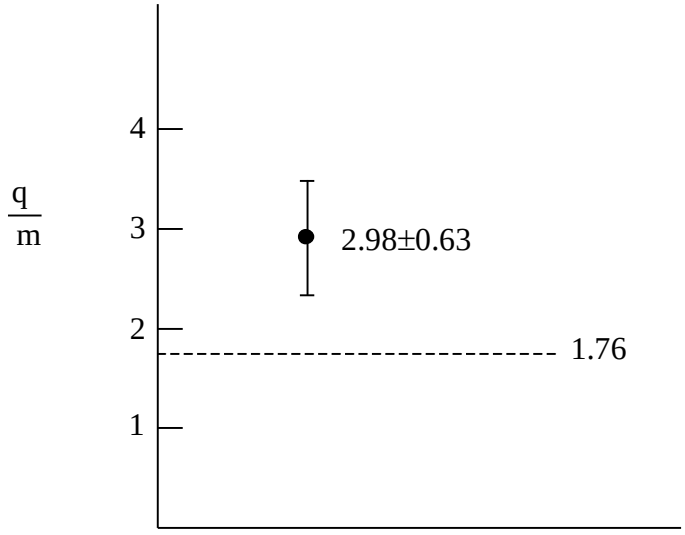


Figure 6: A plot showing the measured value with error bar. The dotted horizontal line represents the accepted value. The q/m axis has units of $\times 10^{11}$ C/kg.

The measured value of q/m is ~2σ higher than the accepted value. The probability for that to happen is only ~5%, which strongly suggests the presence of some systematic error in our measurement.

8. Least Squares Fit.

Suppose you measure some data points y as a function of a variable called x. After the measurements, you will have a set of data points

x_1, y_1
x_2, y_2
……
x_N, y_N

Sometimes you might know that the data should fit a straight line (e.g., from theoretical considerations). The equation of a straight line is

$$y=mx+b$$

where the slope m might equal 'a certain quantity of interest' and the intercept b might equal 'some other quantity of interest'.

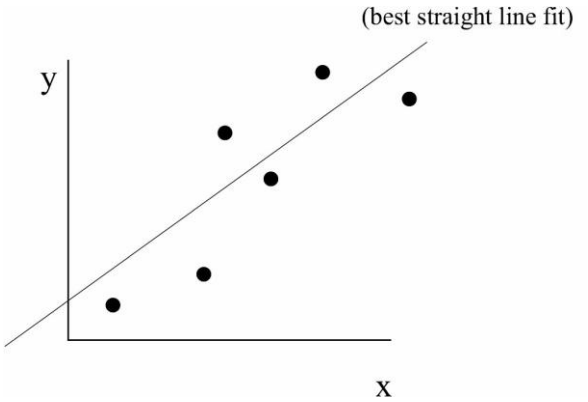


Figure 7: A plot showing the best straight line fit to a collection of data points.

In short, if you could determine *m* and *b*, these values may contain estimates for useful quantities. One way to determine *m* and *b* is to plot the data and use a ruler to draw a straight line through the points. Then, by calculating *m* and *b* from the straight line drawn, you have produced some weighted average estimate of *m* and *b* from all your data.

A simple example: Suppose you are asked to determine π experimentally and suppose you already know that for circles

$$\text{circumference} = \pi\ (\text{diameter})$$

One way to proceed might be to make a variety of circles of different diameters and then measure the circumference of each one. You might plot the data as follows:

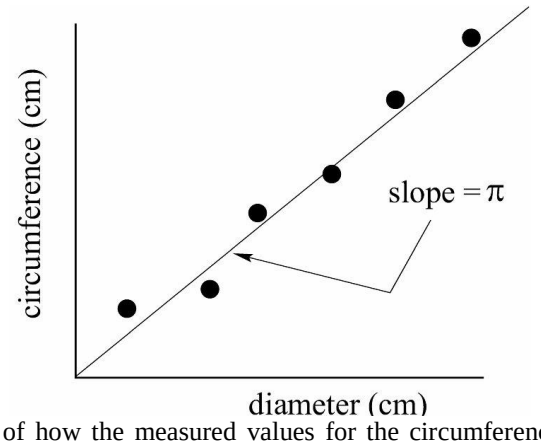


Figure 8: A plot of how the measured values for the circumference of different circles might vary as a function of the measured diameter. Note that in this example, the intercept of the best straight line through the data MUST pass through the co-ordinate origin.

Clearly, the slope of a straight line through the data contains useful information since π=slope.

Q: How can you determine the 'best value' for the slope and intercept without prejudice or personal judgment?
A: Use the principle of **least squares**.

Assume you draw N circles and make measurements of each circumference and diameter. Let the independent variable (the diameter) be represented by the symbol *d*. Let the dependent variable (the circumference) be represented by the symbol C. Also assume the *d* values are accurate. After the measurement process, you'll have a set of numbers (*d*,*C*):

$$d_1, C_1$$
$$d_2, C_2$$
$$\ldots\ldots$$
$$d_N, C_N$$

It is conventional to map these numbers into the parameters (*x*,*y*) as follows

$$x_1=d_1, y_1= C_1$$
$$x_2=d_2, y_2= C_2$$
$$\ldots\ldots$$
$$x_N=d_N, y_N= C_N$$

Let the difference between the 'best line' through the data and each individual data point be represented by (δy_i). One unambiguous way to specify the 'best line' through all the data can be defined by the condition that the sum of all the $(\delta y_i)^2$ have a minimum value.

How are the individual δy_i defined? Graphically, they are indicated in the plot below. Note that at this point of the analysis, the straight line drawn need not be the best straight line through the data.

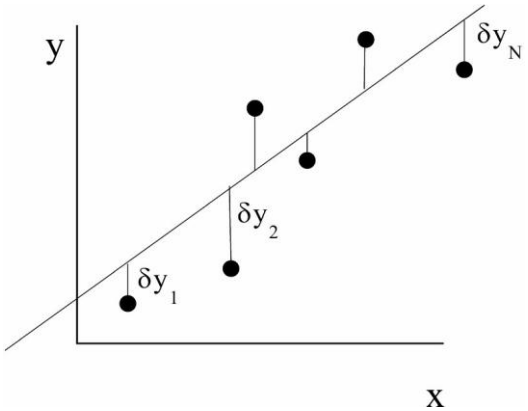


Figure 9. A least squares analysis requires you to calculate the deviation of each data point from the 'best' straight line.

Mathematically, you can calculate the δy_i as follows. Suppose you define a quantity Y_i such that $Y_i=mx_i+b$ where the symbols m and b are somehow chosen to represent the 'best' straight line through the data, whatever that means. Calculate

$$\delta y_i=y_i-Y_i=C_i-(md_i+b)$$

Least squares fitting requires that (where the switch in notation from (C,d) to (y,x) has been made)

$$\sum_{i=1}^{N}(\delta y_i)^2=\sum_{i=1}^{N}\left[y_i-(mx_i+b)\right]^2=\text{minimum}$$

Write $\Sigma\,[y_i-(mx_i+b)]^2=M$. The conditions for M to be a minimum are

$$\frac{\partial M}{\partial m}=0,\qquad \frac{\partial M}{\partial b}=0$$

Performing the derivatives and setting them equal to zero gives, after some algebra, two unique equations for the 'best' m and b.

$$m=\frac{N\sum_{i=1}^{N}(x_iy_i)-\sum_{i=1}^{N}x_i\sum_{i=1}^{N}y_i}{N\sum_{i=1}^{N}x_i^2-\left(\sum_{i=1}^{N}x_i\right)^2}$$

$$b=\frac{\sum_{i=1}^{N}x_i^2\sum_{i=1}^{N}y_i-\sum_{i=1}^{N}x_iy_i\sum_{i=1}^{N}x_i}{N\sum_{i=1}^{N}x_i^2-\left(\sum_{i=1}^{N}x_i\right)^2}$$

Once we have m and b, then also calculate the intermediate quantity σ_y:

$$\sigma_y=\sqrt{\frac{\sum_{i=1}^{N}\left[y_i-(mx_i+b)\right]^2}{N-2}}$$

It can be shown that the uncertainty in the slope m and intercept b is given by

$$\sigma_m=\sigma_y\sqrt{\frac{N}{N\sum x_i^2-\left(\sum x_i\right)^2}}$$

$$\sigma_m=\sigma_y\sqrt{\frac{\sum x_i^2}{N\sum x_i^2-\left(\sum x_i\right)^2}}$$

You can conclude that the best values for m and b are

$$m \pm \sigma_m \quad b \pm \sigma_b$$

This means that there is a 68% chance of the real m lying between $m-\sigma_m$ and $m+\sigma_m$. Likewise, there is a 68% chance of the real b lying between $b-\sigma_b$ and $b+\sigma_b$.

Since the least squares fitting formulae involve sums over various combinations of measured data, the least squares fitting procedure is especially easy to implement in spread sheets like Excel. In fact, most spread sheet programs have pre-programmed least square fit routines available as analysis tools.

In cases where your data points have different individual error values associated with them, it will be necessary to perform a *weighted least squares fit*, where the formulas are more complicated (see e.g., Bevington & Robinson, p 103). A sample spreadsheet for this purpose is provided on the course website.

References

1. G. L. Squires. *Practical Physics*, 4th Edition, Cambridge University Press, Cambridge (2001)
2. P. Bevington and D. Robinson, *Data Reduction and Error Analysis for the Physical Sciences 3rd Edition*, McGraw Hill (2002)

PHYS 340

Lab Report – An Example

Example experiment report for PHYS 340

The following report is written to help students in compiling their own reports for PHYS 340 class. Note that this report does not represent a real experiment and thus should be used only as an example of style and form. The actual experiment reports will usually be longer as there is more material to cover.

Disclaimer: The attached report has no scientific value, any resemblance of the theory, referred names or experiments presented in it with known theories, names or experiments are absolutely coincidental. It is expected, however, that students will describe real theory, use real names and present data from real experiment in their reports.

Measuring g, the magnitude of the gravitational field of earth

Boris Yolkin

Department of Physics, Purdue University, W.Lafayette, IN 47907

Abstract.

In this work we determined the magnitude of Earth's gravitational field g by measuring free fall times for various objects released at different heights and using the Newton's 2nd law. We found that g=9.81±0.08 N/kg, which is in good agreement with the commonly accepted value of 9.8 N/kg.

Introduction.

Humankind has known from historical times that all material objects fall towards Earth if not supported. Moreover, all live creatures on Earth consciously or unconsciously use this phenomenon in everyday life. However, it has long been believed that heavier objects fall faster than lighter ones, and there was no exact theory to describe the motion of an object. Only in the 17th century, when Isaac Newton stated his famous motion laws and gave an exact relationship between force, mass, and acceleration, the quantitative description of motion became possible. In the following paper, we have used the laws of motion to measure the magnitude of the gravitational field of Earth.

Theory.

Newton's 2nd law states that any object would accelerate at a constant rate a if subjected to a constant force F [1]:

$$a=F/m \qquad (1)$$

where m, the proportionality coefficient, is called *mass* of an object. The mass of an object should not be confused with its weight, as it is an intrinsic property of an object to resist acceleration (inertia), while weight refers to a force which attracts one object to Earth or some other usually larger object. While mass is constant[1], weight may be different on different planets, or even in different places on the same planet.

We can rewrite Eq. 1 in a more conventional form:

$$F=ma \qquad (2)$$

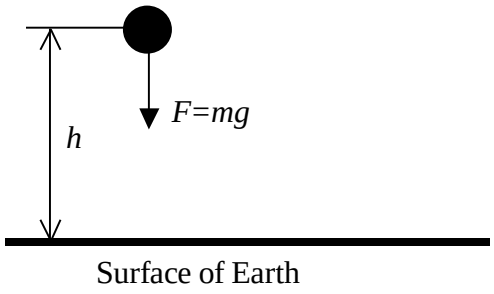


Fig. 1. Any object is a subject to force $F=mg$ toward Earth

[1] We consider only nonrelativistic case.

The force at which an object is attracted to Earth, or its *weight*, is also proportional to its mass (Fig. 1) [2]:

$$F_{weight}=mg \qquad (3)$$

where g is proportionality coefficient. By comparing Eq. 2 and Eq. 3, one can see that g has the same units as a. Moreover, if an object is subjected to gravitational force we can combine Eqs. 2 and 3:

$$ma=mg \qquad (4)$$

or

$$a=g \qquad (5)$$

The last equation states that an object in free fall will accelerate toward earth with constant acceleration equal to g. The units of g are therefore m/s^2, the same as for acceleration. Sometimes, however, units of N/kg are used to reflect the nature of g.

Since acceleration is a second derivative with respect to time of an object's position , we can rewrite Eq. 5 in the following form:

$$\frac{d^2h}{dt^2} = g \qquad (6)$$

where h is a height of an object from the surface of Earth. Integrating Eq. 6 we get:

$$\frac{dh}{dt} = \int_0^t g dt = gt \qquad (7)$$

$$h = \int_0^t gt dt = \frac{gt^2}{2} \qquad (8)$$

$$g = \frac{2h}{t^2} \qquad (9)$$

The last equation shows that g can be easily determined if one measures free fall times t as a function of height h. In fact, a single pair of (h,t) is sufficient to uniquely determine the value of g. In this work, however, we will measure fall times for various heights to test the validity of Eq. 9 as well.

Experimental apparatus and procedures

In the first part of the experiment a 0.5 kg solid aluminum ball was dropped down from different floors of a 10-story building and fall times were measured by a stop watch. The heights to different floors were measured by a conventional ruler.

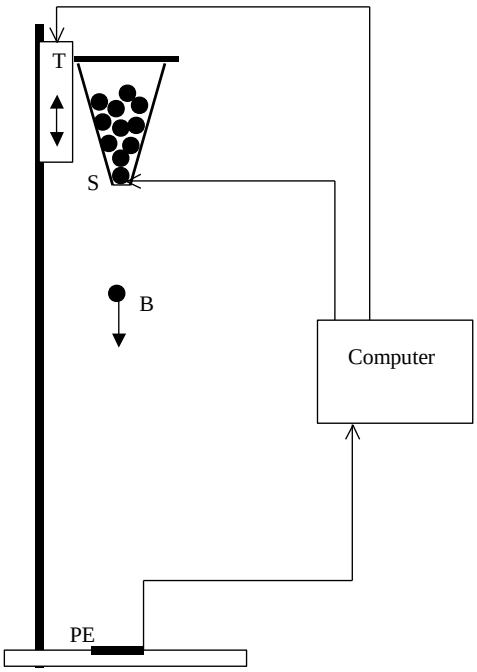


Figure 2. Experimental setup for measuring fall times as a function of height. T – translation stage, S – automated shutter, PE – piezo-electric detector, B – metal ball.

In the second experiment fall times from heights up to 2 meters were measured by an automated setup (Fig. 2). Here, a container filled with small metal balls *B* and equipped with computer controlled shutter *S* is attached to a caret of a vertical translation stage *T*. A computer program slowly moves the caret T and releases metal balls at specified heights, one at a time. Simultaneously, the computer timer is started when the ball is released. The ball falls onto a piezo-electric detector *PE* that sends a stop signal to the computer timer. The precision of the timer is 1 ms, and the height is determined to an accuracy of ±5 mm. The experimental apparatus is described in more details in [3].

1. Data analysis, results and uncertainties.

Part 1. The results of the first experiment are shown in Fig. 3. Here, the error in fall time measurement was estimated to be 0.2 s and is mostly determined by the reaction time of the experimentalist. The height was measured by a ruler to ±2 cm precision. Error bars for height are not shown in Fig. 3 since they are smaller than the size of the dots that represent the measured points.

The general shape of this graph is well described with the quadratic dependence defined by Eq. 8, the solid line in Fig. 3 represents the expected dependence with the value g=9.8 m/s^2. To analyze these data further, we used Eq. 9 to calculate the value of g for each data point, and the result of this calculation is shown in Fig. 4.

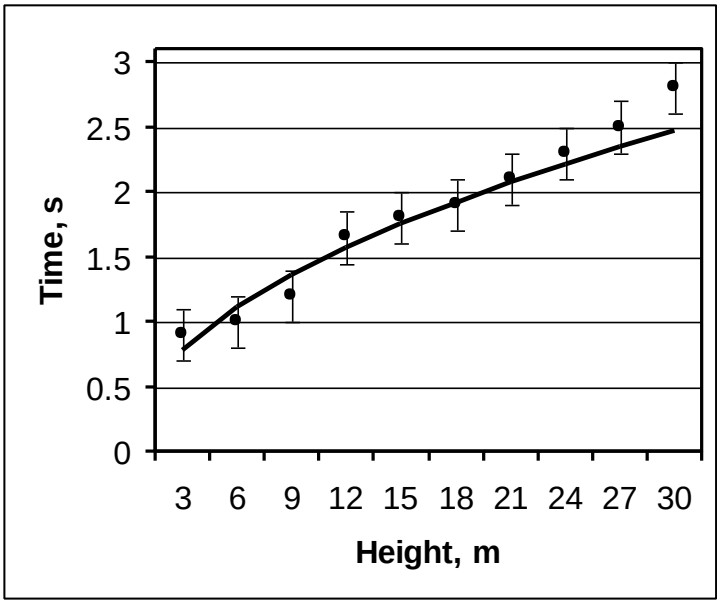


Figure 3. Dependence of fall time on height. The solid line is theoretical simulation using Eq. 8 and g=9.8 m/s^2. Error bars for measured values along X-axis are not shown as they are smaller than the size of the circles.

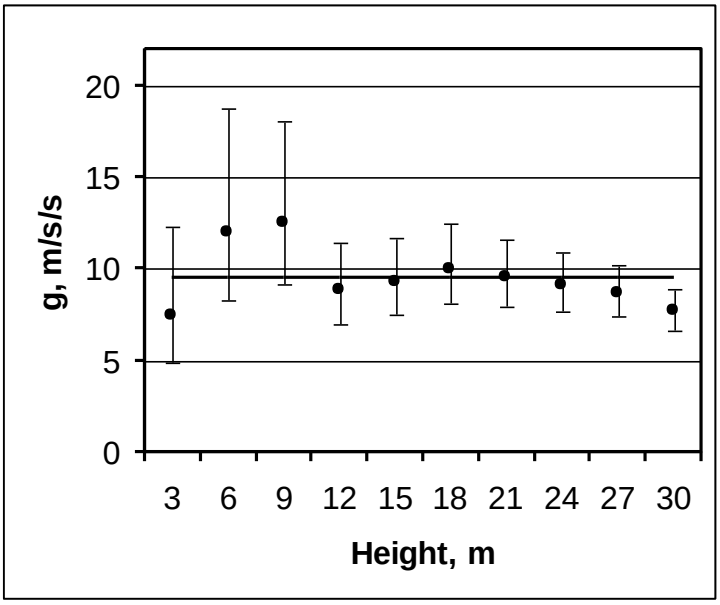


Figure 4. The calculated value of g for different heights. The solid line corresponds to the average value 9.45 m/s^2.

The indicated error bars were calculated using the worst case scenario for each data point using the following equations:

$$\Delta g^{+} = \frac{2(h+\Delta h)}{(t-\Delta t)^2} - g$$
$$\Delta g^{-} = g - \frac{2(h-\Delta h)}{(t+\Delta t)^2} \qquad (10)$$

In Eq. 10, g is the value calculated using equation (9), Δh=0.002 m and Δt=0.2 s are errors in measuring height and time, correspondingly. Δg^{+} and Δg^{-} are error levels in positive and negative side of the calculated g value. Based on these results, the average value of g can be calculated using least square fit as (9.5±2) m/s^2. Corresponding straight line at g=9.5 m/s^2 is also shown in Fig. 4.

Part 2. In part II the computer measured 200 (*h*,*t*) points and the resulting dependence is shown in Fig. 5. The error level in these measurements is much smaller and is reflected in the visible noise in the curve.

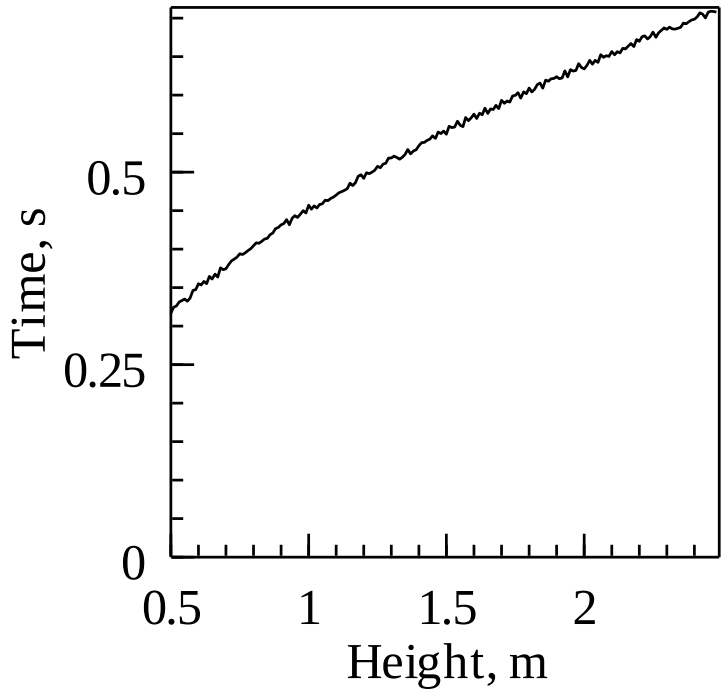


Fig. 5. Fall time versus height measured by computer controlled apparatus.

The data obtained by using automated data acquisition setup was analyzed in two ways.

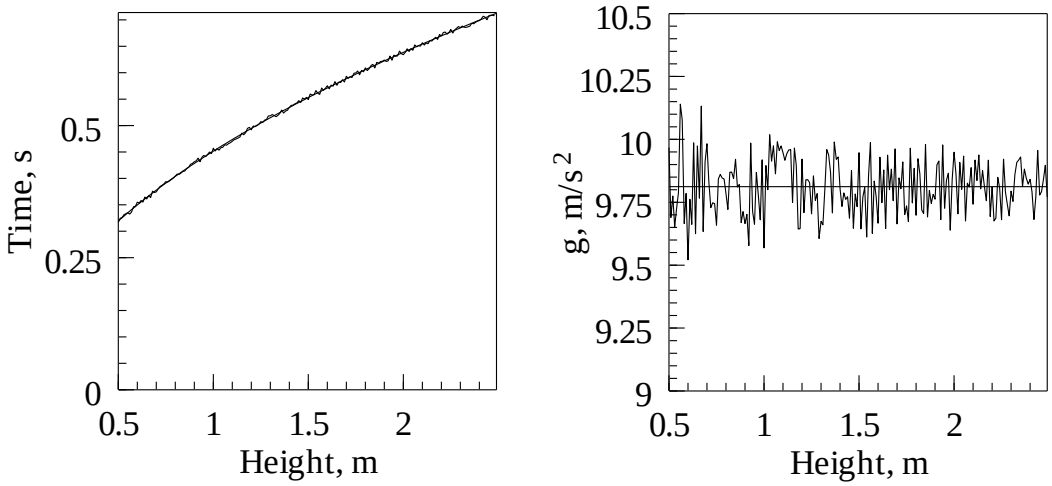


Figure 6. Left: the data *h(t)* fitted using least square by Eq. (9), the best fit was obtained for $g=9.81$ m/s^2. Right: the g was calculated for every height h resulting in *g(h)* graph, which was then averaged (solid line) resulting in $g=9.81$ m/s^2.

First, the data was fitted as is by Eq. (9) using linear least square fitting method. The best fit was obtained for $g=9.81$ m/s^2 as shown in Fig. 6, left. In order to analyze the precision of this measurement we calculated the values of g for every height (Fig. 6, right). That allowed us to find the average value for $g=9.81$ m/s^2. Along with the average, we also calculated standard deviation $\sigma_1=0.14$ m/s^2, which reflects error level in one measurement. Since we averaged results of 200 independent measurements, standard deviation for the average is $\sigma = 0.14/\sqrt{200} = 0.01$ m/s^2. Using the declared errors for timer (1 ms) and position (5 mm), we can estimate measurement error using the following equation:

$$\frac{\Delta g}{g} = \sqrt{\left(\frac{2\Delta t}{t}\right)^2 + \left(\frac{\Delta h}{h}\right)^2}$$

For the highest and lowest drop, Δg is 0.03 and 0.08 m/s^2, respectively. Since this error may be systematic, it cannot be averaged out and has to be combined with σ, resulting in error level 0.08 m/s^2.

1.1 Discussion

While the data obtained in these experiments is well described by the Newton's 2nd law, there are some deviations for larger heights as can be seen in Fig. 3. We believe that this is due to the friction of the air, which becomes larger at higher speeds and for

which our theory does not account. Indeed, qualitatively the presence of friction must slow down objects and make fall times longer, and that is exactly the effect we see for height 30 m (Fig. 4). For the same reason for the highest drop (30 m) the value of *g* turned out to be smaller than average, and smaller than expected beyond error level (Fig. 5). Therefore, measurement for heights ≥30 m are not reliable. On the other hand, error levels for heights below 10 m become increasingly larger (Fig. 5) due to very short fall times, which are hard to measure manually. The optimal height for such measurements is 10-25 meters.
The second experiment minimizes the effect of air friction by using drop heights below 3 meters. Using computer controlled automatic timer makes precision far superior to manually controlled timer and allows us to measure *g* with extremely high precision - σ~1% versus 20% in the first experiment.

1.2 Conclusion

Using fall times we were able to measure the magnitude of Earth's gravitational field *g*. The first method, where we used manual timer, resulted in $g=(9.5\pm2)$ m/s^2, while the second measurement using computer controlled experiment refined this value to (9.81 ± 0.08) m/s^2. Both measured values are in excellent agreement with the commonly accepted value of $g=9.8$ m/s^2, though the computer driven experiment is clearly more precise.

1.3 Acknowledgements

I would like to thank Anya Smith, my lab partner, for her help in performing these experiments, Alex Ivanov for helpful discussions and suggestions on data analysis, and Bill Doors for developing the software for the computer controlled experiment.

References

1. I. S. Newton, *Philosophie Naturalis Principia Mathematica,* Londini, London (1686).

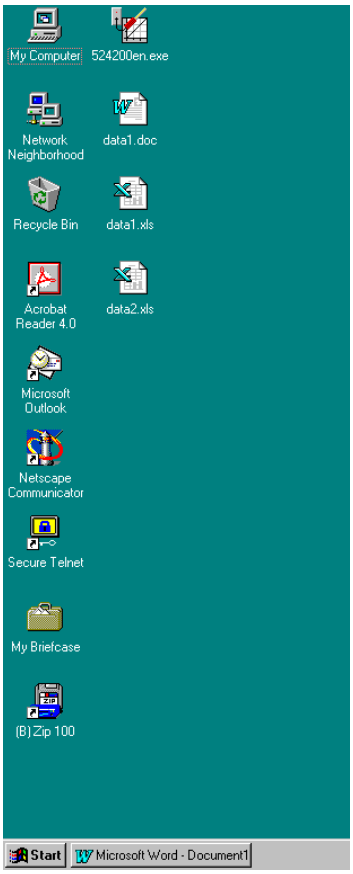
2. R. Chabay, B. Sherwood, *Matter & Interactions*, vol. 1, John Wiley & Sons, New York (2002)

3. Measuring the magnitude of the gavitational field of Earth g. PHYS 340 manual, Purdue University, W. Lafayette (2002).

Getting Started with Computerized Data Acquisition

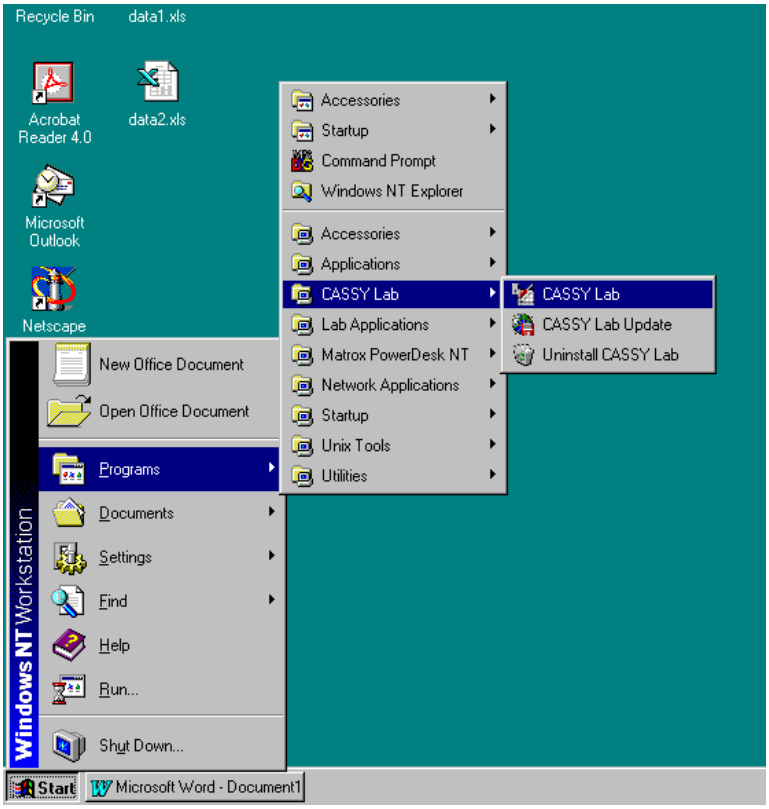
Login:

When you prepare to use the 340 computers, you will find a login screen. Use your Purdue Career account to log in.

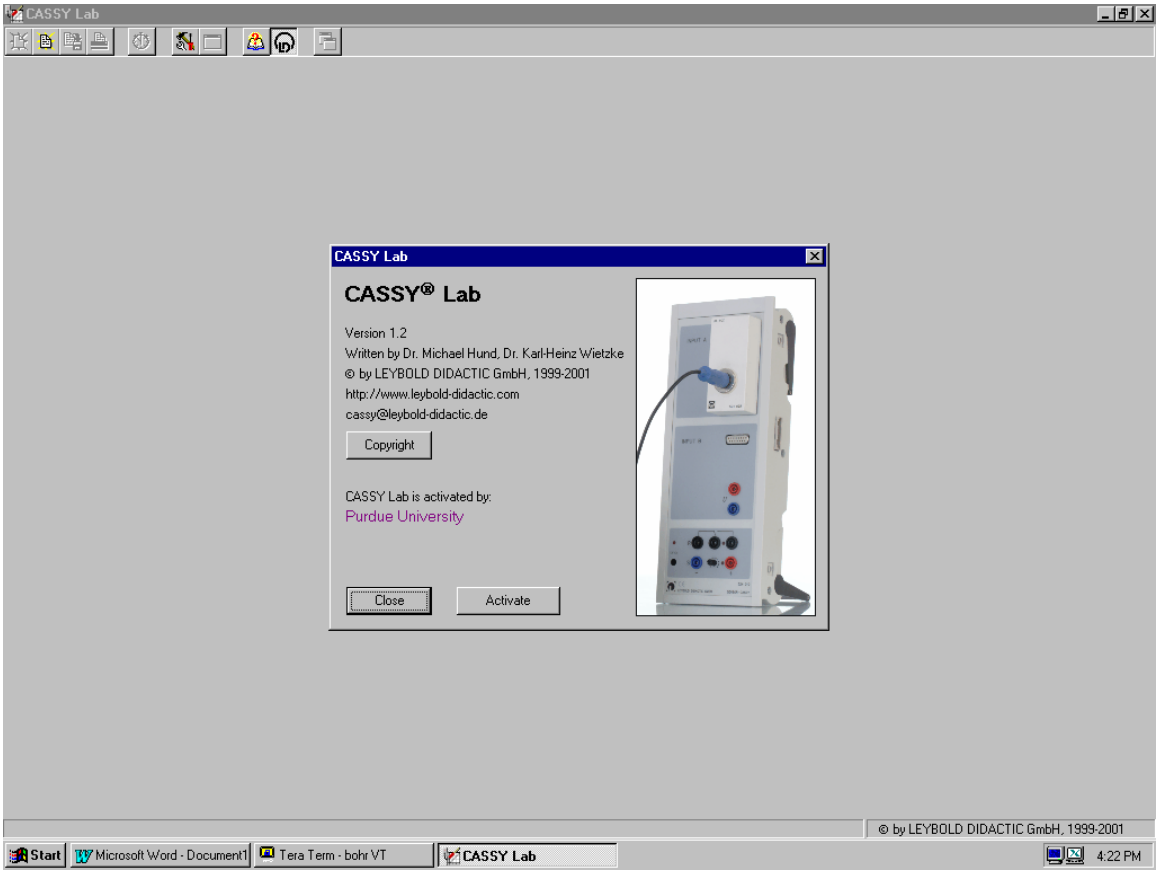


After logging in, the screen should look something like the picture to the right (only fragment is shown)

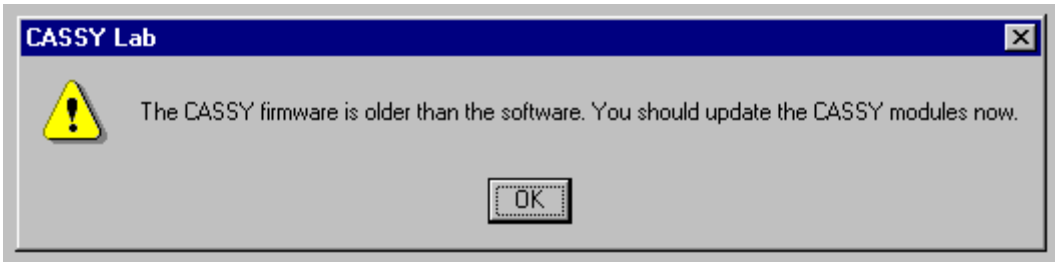
CASSY Lab is the name of the software we will be using to perform data acquisition. CASSY is started via the familiar Start-Programs-CASSY Lab-CASSY Lab, as shown in the picture below.



The CASSY program will then start opening main window:

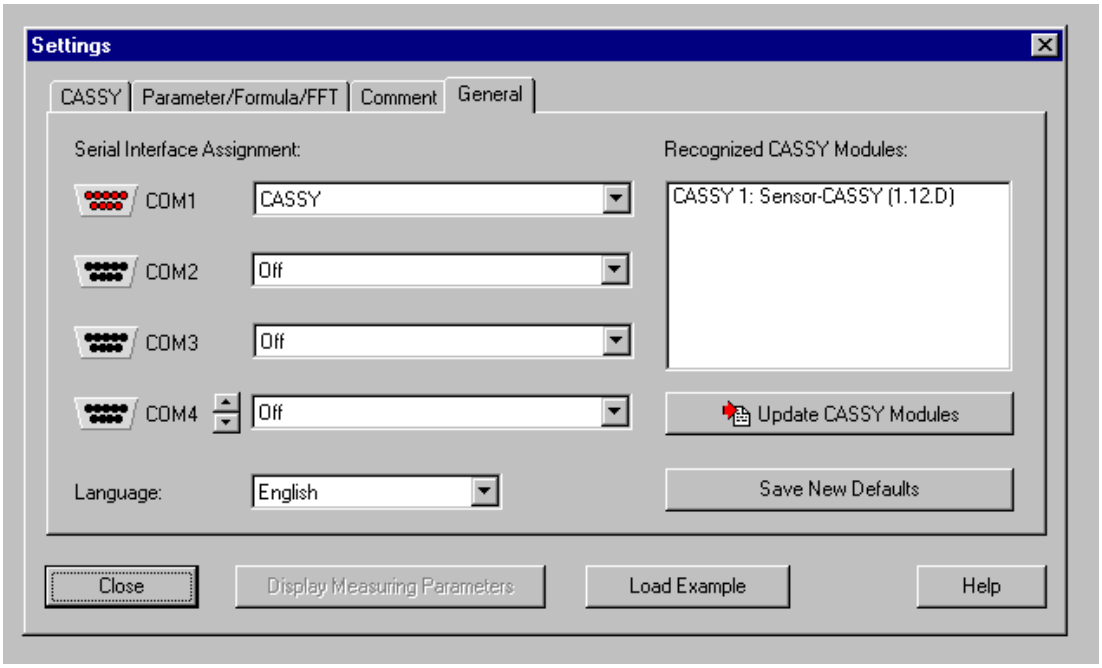


Then after a few seconds one may get an error message:



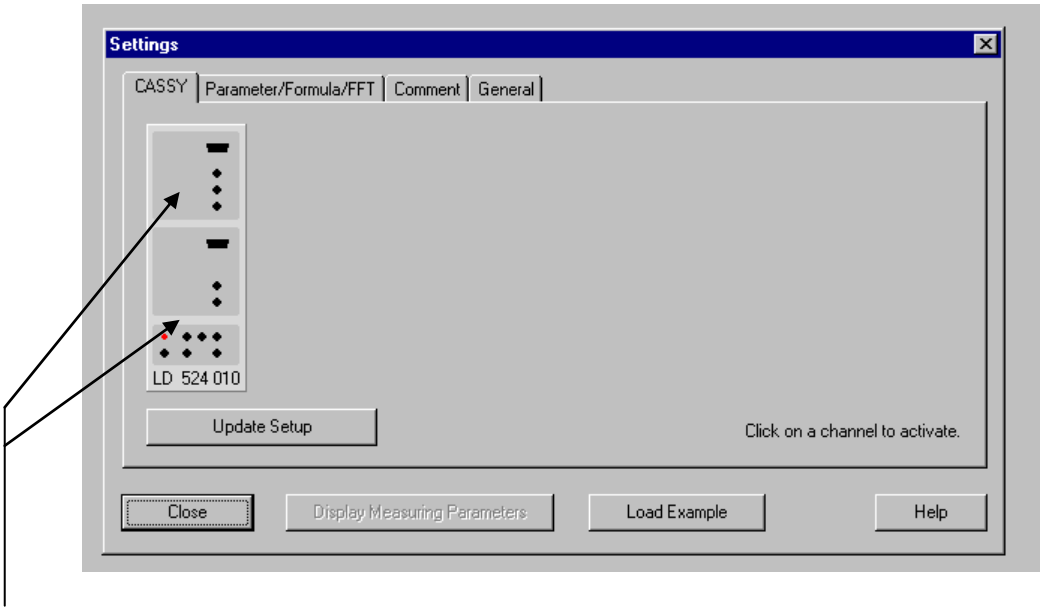
If you get this error message, just ignore it and click OK.

The following window then appears:



At this point things should be ready to go. Note that the CASSY sensor has been located on COM1. The other COM ports should remain turned off.

After clicking on the CASSY tab, you will see a picture of the CASSY box. This picture may vary depending on what equipment is being interfaced to the CASSY sensor:

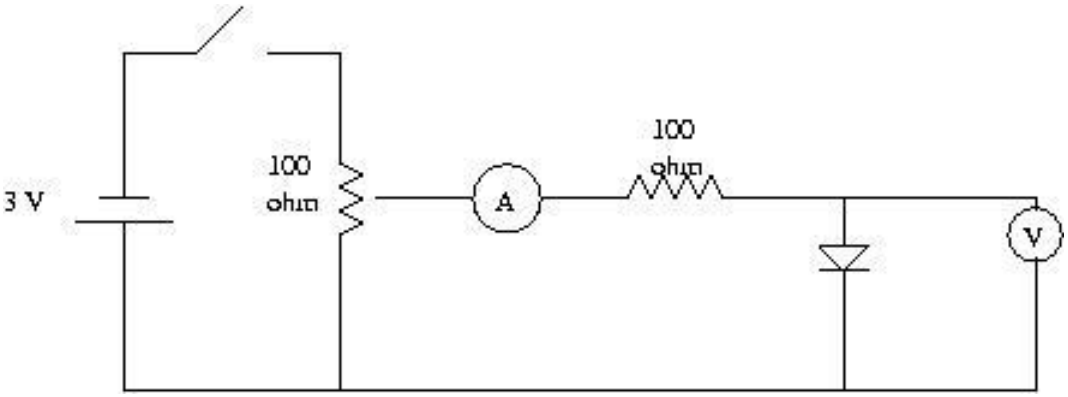


Click on each of the 'interface' boxes to activate them.

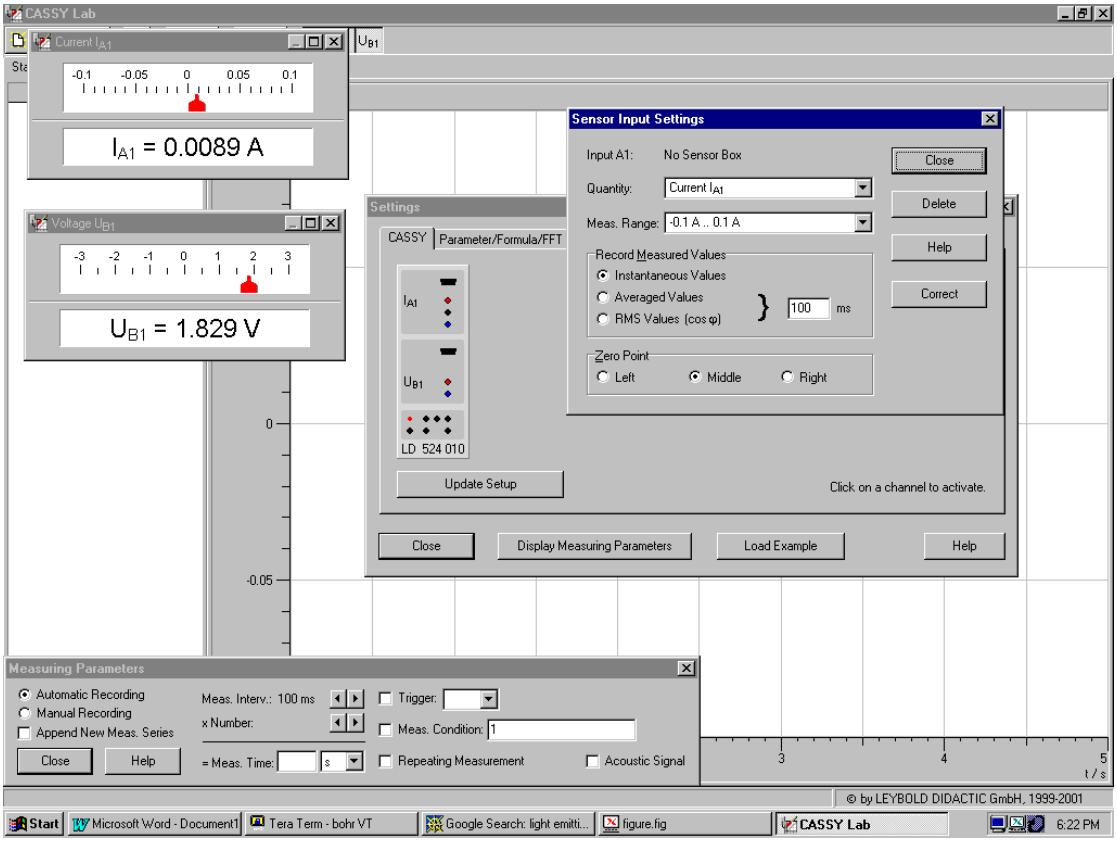
Let's now begin a simple demonstration.

A light emitting diode is made from the junction of a P-type and an N-type semiconductor, with light being produced by the recombination of elections and holes at the junction. There are many references which explain how solid state devices work. For our purposes, we need only to know that the current (and emitted radiation) grows exponentially with the applied voltage. There also exists a threshold voltage below which no power will be radiated. We want to find this threshold voltage.

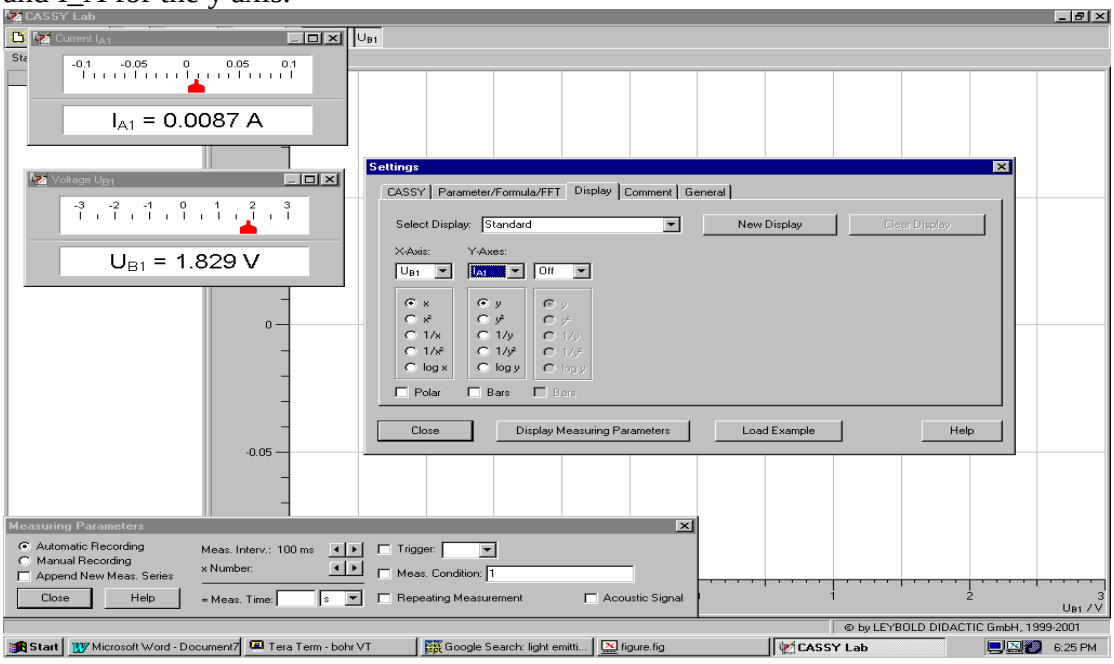
We being by constructing a circuit as follows



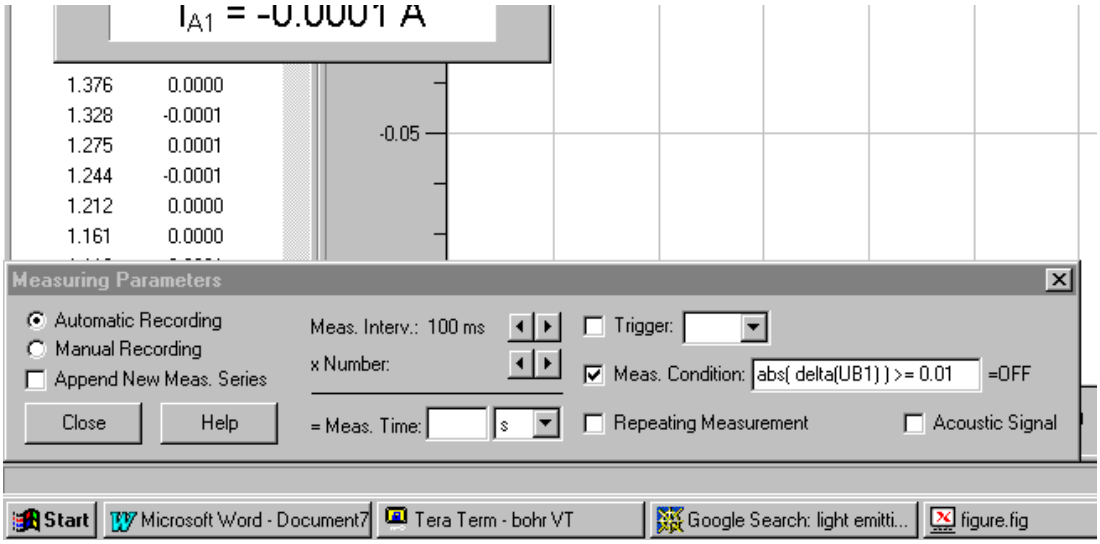
Now we click on the Input A section of the CASSY and select current (I_A) as the type of measurement to make, and then click on Input B and select voltage (U_B) as the measurement parameter. The following window will appear:



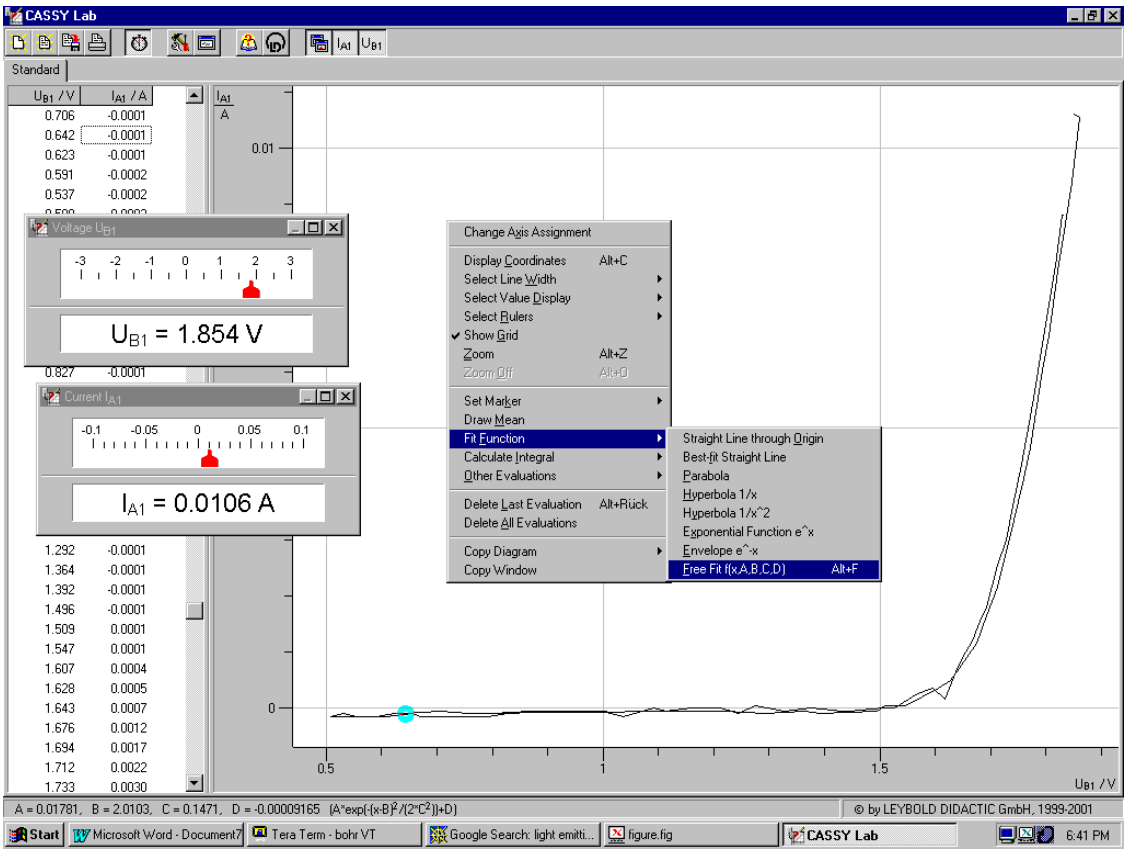
Current should be set for the smallest parameter possible (-0.1 A to 0.1 A) and voltage should be set for -3V to 3V. (We may also zoom in later in the graph window. Now we click on the DISPLAY tab and set the voltage U_B for the x axis and I_A for the y axis.



Let's now put in a "measurement condition" in the bottom box. Let's say we want to read a data pair for every 0.01 volt change in the applied potential, the absolute value of the change in the voltage parameter should be greater than or equal to 0.01; the measurement condition should be abs(delta(UB1)) >= 0.01. Note that when typing in the commands that they are case sensitive.

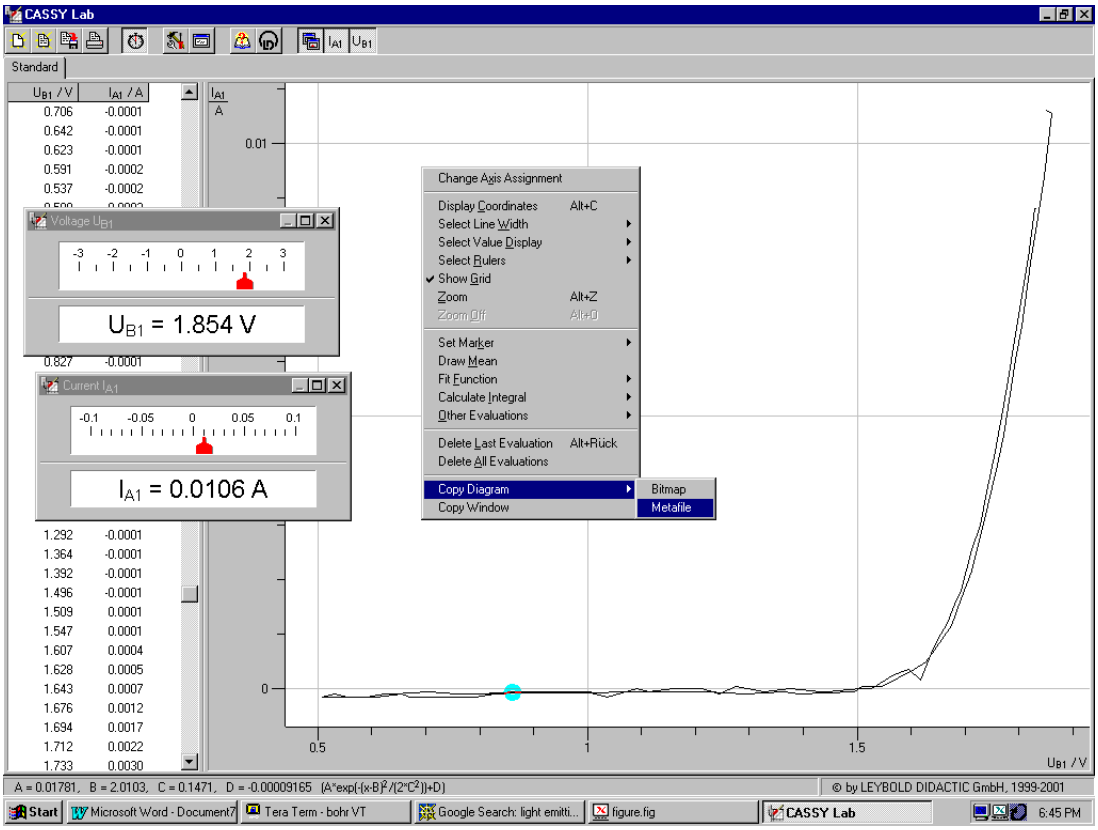


We now vary the voltage and can begin to see a rise in the current passing through the diode (as well as begin to see a glow through the LED). Let's see if we can fit a function to this data.

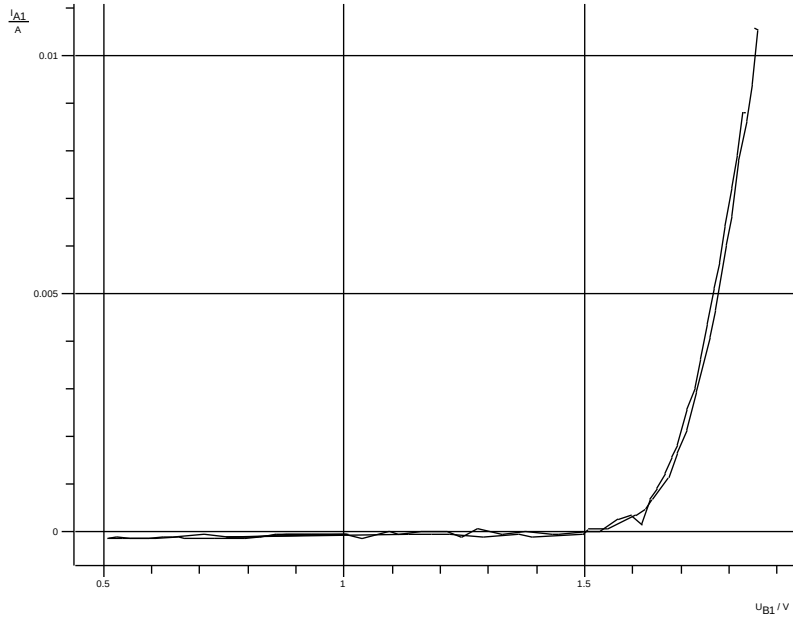


We right click, choose fit function, and then select free fit. We then select the type of curve we want to fit and highlight the region of the graph we want to fit. Constants and the function fitted are near the bottom of the screen.

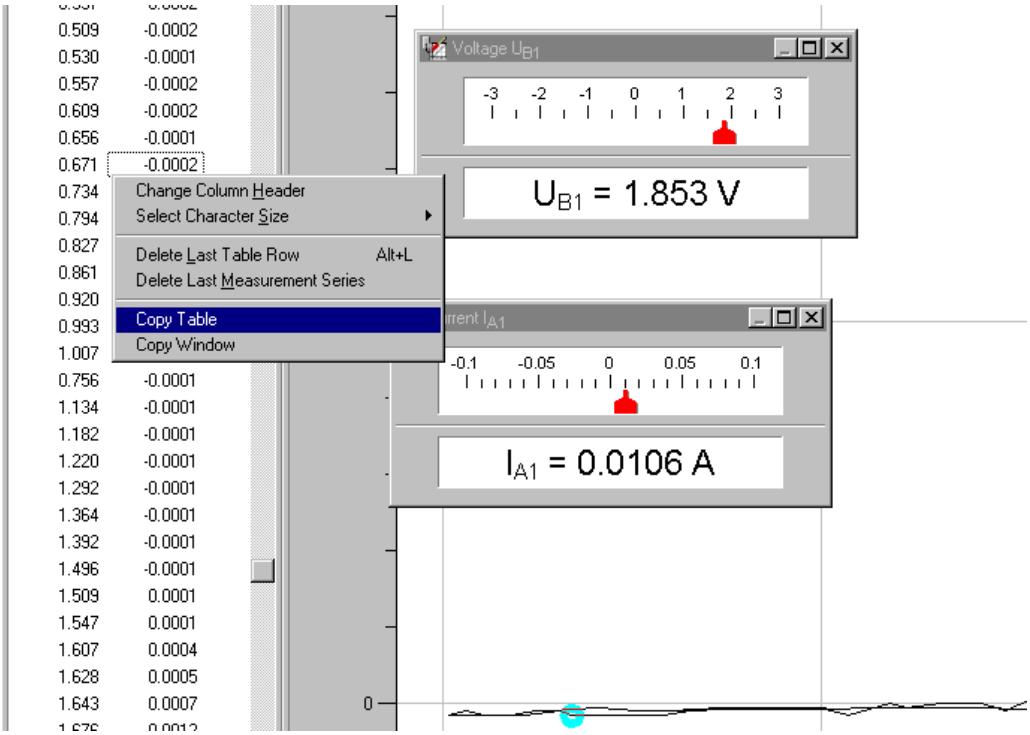
Now let's copy our graph and our data to a word document so we can paste it into our lab notebooks and lab reports.



We copy the graph as a metafile so that it may be edited later, and the pasted version follows,



We can also copy our data as the list of (x,y) numbers for further processing (in Excel, for example):



And paste it into this word document (or just a simple text editor like Notebook) so we can print it out and save it should we need it later.

U_B1 / V	I_A1 / A
1.829	0.0088
1.817	0.0079
1.806	0.0072
1.793	0.0064
1.781	0.0057
1.770	0.0051
1.755	0.0044
1.740	0.0036
1.728	0.0030
1.712	0.0025
1.692	0.0018
1.680	0.0016
1.667	0.0012
1.649	0.0009
1.637	0.0007
1.619	0.0002
1.596	0.0004
1.569	0.0003

etc…

Physics 340 Laboratory
A New Piece of Matter: The Determination of e/m for Electrons

Objective: To determine the charge to mass ratio of electrons.

Apparatus: Vacuum tube (Bainbridge); Helmholtz coil pair; calibrated meters to measure current and voltage; 300 V power supply; 6.3V AC power supply; Sargent-Welch interface box.

References:
1. J.J. Thomson, Phil. Mag. **44**, 295 (1897).
2. K. T. Bainbridge, The American Physics Teacher **6**, 3 (1938).
3. Common Apparatus AAPT Novel Experiments in Physics (Am. Inst. of Physics 1964) pp. 237-41.
4. J.R. Reitz and F.J. Milford, *Foundations of Electromagnetic Theory*, (Addison Wesley; Reading, Massachusetts; 1969), pp. 156-57.

Introduction

The first experimental evidence for the granular nature of electricity can be found in Michael Faraday's work on electrochemical processes in 1839. Prior to Faraday's work, electricity was considered a 'fluid' that could be added or subtracted in a continuous fashion from objects. It was subsequently discovered that metals emitted negative electrical current when heated, illuminated by light, or subjected to a strong electric field. It was theorized that the negative current was comprised of particles each carrying a negative charge (now known to be 1.602×10^{-19} C). These negative particles were thought to be universally present whenever a negative current was emitted from an electrode. Surprisingly, they were not apparently related to the particular metal from which the emitting electrode was fabricated, thus providing strong evidence for their fundamental nature. In 1874, George Johnstone Stoney suggested they be called 'electrons', although the general scientific community remained skeptical about their existence for several decades.

Early attempts to measure the mass of electrons proved futile. To address this issue, various direct experiments were devised. For instance, if a metal sphere of 1 meter radius is charged to a potential of -1×10^6 V, you can quickly estimate that about 7×10^{14} electrons must be added. Yet attempts to directly measure the mass increase of such an electrified sphere from the added electrons yielded no conclusive results. As a consequence, it was argued that the mass of an electron must be very, very small compared to any atomic masses that were known in the late 1800's.

Indirect methods were therefore sought to measure the mass of these negatively charged particles. J. J. Thomson in 1897 was the first to partially succeed in this

endeavor by measuring their charge-to-mass ratio. Robert Millikan completed the story in 1910 by measuring the quantized electron charge in his famous oil-drop experiment. These two experiments eliminated any lingering doubts about electrons as bona fide sub-atomic particles, and both men eventually received Nobel prizes.

Thomson's experiments combined both ingenious insight and the use of newly improved vacuum pumps and charge measuring devices (electrometers). Using clever calorimetric techniques, he measured the temperature rise of thin metal targets inserted into a glow discharge tube (first produced by Faraday, when two metal plates were inserted inside an evacuated glass tube and raised to a high electrostatic potential). This strategy enabled Thomson to calculate the energy imparted to the metal targets by the invisible particles responsible for producing the glow. By assuming the invisible particles followed Newton's laws of motion (at the time, there was no reason to suspect that this might be the case), he was able to estimate the velocity of the particles responsible for the glow discharge. He also used crossed electric and magnetic fields to further study the motion of these invisible particles.

As a result, Thomson concluded (within the accuracy of his measurements) that a negatively charged particle with a mass to charge ratio of $1.3\pm0.2\times10^{-11}$ kg C^{-1} was responsible for producing the glow discharge. This mass to charge ratio was roughly 10^3 times smaller than any ratio previously recorded for atoms or molecules. For this reason, he felt confident in announcing laboratory confirmation of electrical particles ('corpuscles', in his words). The modern, accepted value for the mass to charge ratio of an electron is $0.5685629653\pm0.0000000001\times10^{-11}$ kg C^{-1} (or more commonly used: the charge to mass ratio = 1.758 820 088(39) x 10^{11} C kg^{-1}).

In this experiment, you will perform an experiment similar to Thomson's original work. By measuring the deflection that a magnetic field produces on a beam of electrons having a known energy, you will deduce a value for the charge to mass ratio of electrons.

Theory

A charged particle moving in a magnetic field experiences a force $\vec{F}$ given by

$$\vec{F} = q\vec{v} \times \vec{B}, \qquad (1)$$

where q is the charge of the particle, $\vec{v}$ is the particle's velocity, and $\vec{B}$ is the magnetic field. If $\vec{v}$ is perpendicular to $\vec{B}$, the resultant trajectory is circular. Using Newton's laws of motion, the radius of the electron orbit is given by

$$R = \frac{m|\vec{v}|}{q|\vec{B}|}, \qquad (2)$$

where m is the mass of the charged particle.

Although R and $|\vec{B}|$ in Eq. 2 can be measured experimentally, determining the magnitude of the velocity is more problematic. Progress can be made by using conservation of energy considerations. If a charged particle is initially at rest and is accelerated through an electric potential difference V, then the kinetic energy after acceleration is equal to the change of the potential energy qV. From the conservation of energy principle we know that

$$\frac{m|\vec{v}|^2}{2} = qV \,. \qquad (3)$$

By eliminating the velocity $|\vec{v}|$ from Eqs. (2) and (3), we obtain

$$\frac{q}{m} = \frac{2V}{|\vec{B}|^2 R^2} \,. \qquad (4)$$

Since the charge of an electron is particularly important, we often use its special symbol e instead of q.

To calculate the e/m ratio for an electron, we need to know the accelerating potential V, the value of the magnetic field B and the radius R of the circular path of the electron beam. Since it is hard to detect a single electron, we will use a beam of electrons in which all have approximately the same kinetic energy.

Experimental Method

In this experiment, a beam of electrons is produced by an electron gun (Fig. 1) composed of a filament surrounded by a coaxial anode (i.e., an electrode with a positive charge). Electrons thermally emitted from the filament are accelerated by a known potential difference V between the filament and the anode. The source of electrons in this experiment is a metal plate called the cathode. The cathode is often coated with metal carbonates having a low work function. This enhances electron emission from a material in which the free electrons are only slightly bound, producing measurable emission currents for modest filament temperatures.

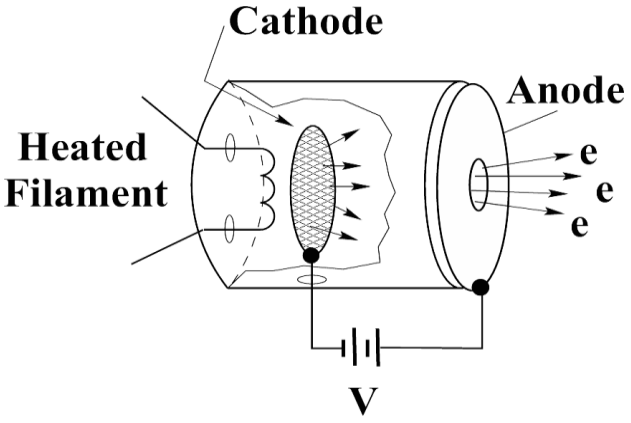


Figure 1: A schematic diagram of a simple electron gun showing the filament (heating power supply not shown), the cathode, the anode and the accelerating voltage V.

The cathode can be either in thermal contact with the filament, or be designed to perform the function of both the cathode and filament combined. Thus by heating the cathode, electrons gain enough energy via thermal excitation to emerge from the cathode surface as free particles. This process is called thermionic emission. The kinetic energy and velocity of the electron beam may be calculated using Eq. (3).

During its manufacture, the air in the *e/m* tube is evacuated and the tube is backfilled with a small quantity of helium gas at a pressure of ~1 Pa (approx. 100,000 times smaller than atmospheric pressure) before permanent sealing is performed. The electron beam leaves a visible trail in the tube because some of the electrons collide with helium atoms and promote them into an electronically excited state. These excited He atoms thereafter emit visible light as they return into ground state (i.e., they *fluoresce*). The visible light observed in 1838 by Faraday was the faint glow produced when electrons collided with residual gas atoms left inside the tube by the inefficient vacuum pumps that were then in use.

Experimental Equipment

A side view of the *e/m* apparatus is shown in Fig. 2. The important aspects of this apparatus are the vacuum tube containing the electron gun (see Fig. 3), which produces a narrow beam of electrons of known energy having a velocity nearly perpendicular to an applied magnetic field. The latter is produced by a pair of coils arranged in a Helmholtz configuration.

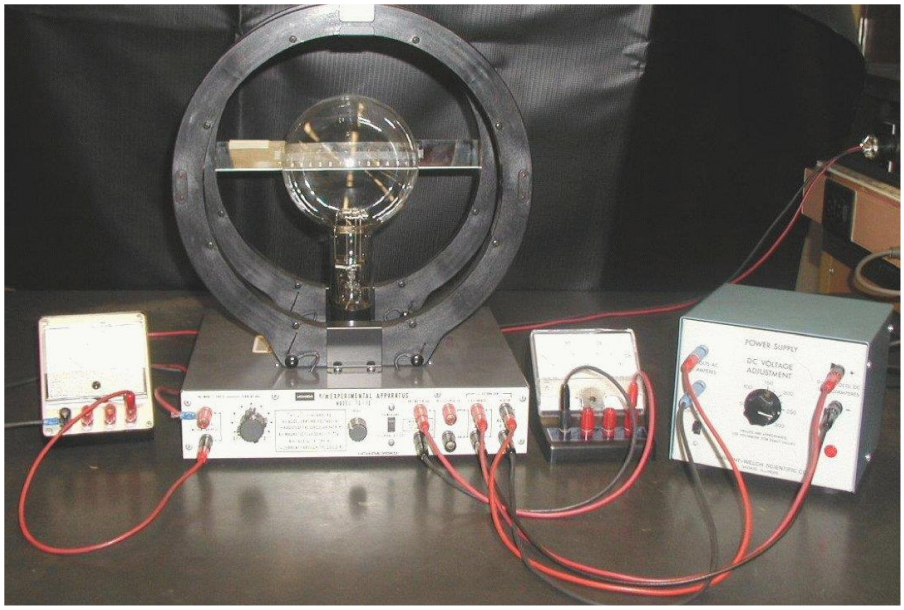


Figure 2: A photograph of the *e/m* apparatus (note that the setup depicted here uses a different illuminated scale device).

A pair of identical Helmholtz coils separated by a distance equal to their radius produces a uniform and measurable magnetic field along the central horizontal axis of the coils. If the electron beam is directed at a right angle to the field, this magnetic field then deflects the beam into a circular path via the Lorentz force. By measuring the accelerating potential, the current to the coils, and the radius of the circular path of the electron beam, *e/m* can be calculated from Eq.(4).

Figure 3: A close-up picture of the *e/m* electron gun. From this view, the electron beam emerges to the right.

The Helmholtz coils for the *e/m* apparatus you will use each have 130 turns. The magnitude of the magnetic field *B* produced at the axial mid-point of the Helmholtz pair is proportional to the current through the coils *I* and can be calculated using Ampere's law, yielding:

$$B = \frac{8N\mu_0}{5\sqrt{5}a} I = kI \,, \qquad (5)$$

where *N* is the number of turns in each Helmholtz coil, *a* is the radius of each coil in meters, and μ_0 is the permeability of free space ($\mu_0 = 4\pi\times10^{-7}$ T·m/A).
The field constant for these coils, $k = B/I$ (in T/A), gives a measure of how many Teslas result when 1 A of current passes through the coils. A DC power supply provides the current to the Helmholtz pair. Using the current adjust knob of the *e/m*

apparatus, you can adjust the value of the current through coils to be between roughly 1 and 2 A.

In principle, the magnetic field needed to produce the observed curved paths for the electron beam is small, so the possible effect of the Earth's magnetic field (roughly 3.1×10^{-5} T at the equator, 5.5×10^{-5} T in Indiana) should be considered. One way to minimize any effects due to the Earth's field is to rotate the apparatus so the local geomagnetic field is *parallel* to the motion of the electron beam. A small compass is provided to establish the direction of magnetic north. You should establish whether the Earth's magnetic field is important relative to the error margins of this particular experiment, and if so, how the tube and the coils should be rotated in order to achieve the proper orientation. Note that the effect of permanent magnets may be substantial - any permanent bar magnets should be far removed from the apparatus when this procedure is performed.

The power supply provides two necessary voltages:

- A fixed, low voltage, 6.3V AC voltage for heating the filament of the e/m tube. **CAUTION: The voltage to the heater of the filament should never exceed 6.3 volts. Higher voltage will burn out the filament and destroy the e/m tube.**

- An adjustable (0 - 300V) DC voltage used for acceleration of electrons between the cathode and the anode. You can adjust the value of that voltage by turning the knob labeled: "DC VOLTAGE ADJUSTMENT". The scale around the knob is only for estimation purposes. You will need to read the value of the accelerating voltage from a separate voltmeter using the 0 - 300V scale.

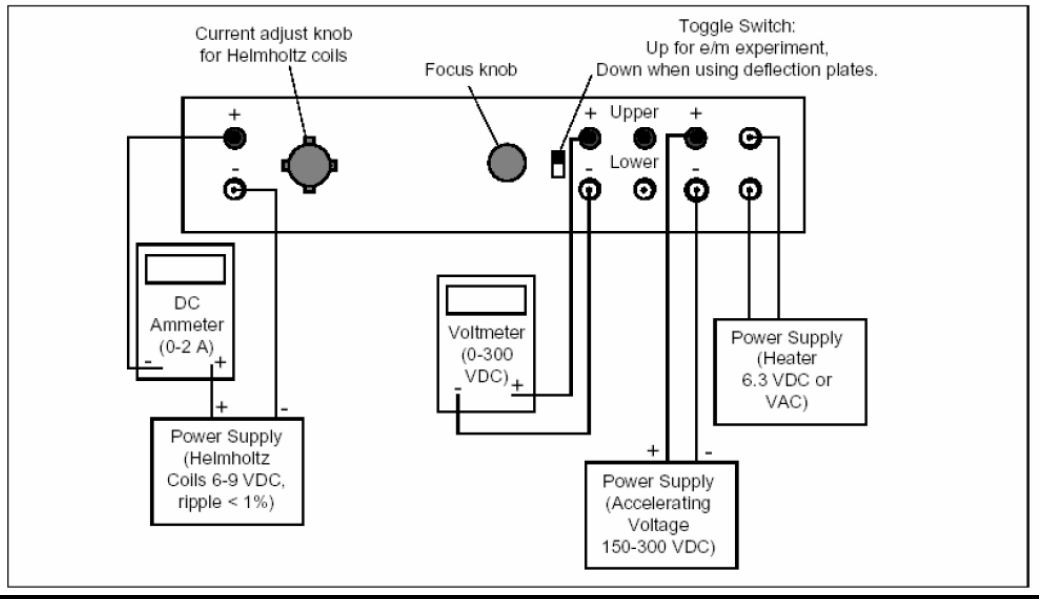


Figure 4: Electrical connections for e/m apparatus.

Fig. 4 shows the electrical connections for the e/m apparatus. The control panel of the *e/m* apparatus is straightforward. All connections are labeled. An illuminated scale is provided with the experimental setup. You will use it to place a virtual scale at the site of the electron ring by using the reflection off of the plate glass. By aligning the electron beam up with its image in the reflected scale, you can measure the radius of the beam path without parallax error.

Experimental Procedure

1. Flip the toggle switch up to the "e/m MEASURE" position.

2. Turn the current adjust knob for the Helmholtz coils and knob labeled "DC VOLTAGE ADJUSTMENT" on the POWER SUPPLY to the zero position (counterclockwise).

3. Before applying any power, check to make sure that all connections correspond to the wiring diagram shown above. **To avoid a possible shock hazard, check that shielded connector cables are used for the high voltage power supply connections.** If you do not have full confidence that the circuitry is OK - ask your TA to check it for you.

4. Calculate the field constant *k* of the Helmholtz coils and write it in your lab notebook now.

5. Turn the power supply on. The filament inside the tube should begin to glow red. Then, plug the low voltage (Helmholtz coils) cable into the wall socket labeled "A".

6. Slowly turn the current adjust knob for the Helmholtz coil clockwise. Watch the ammeter **and take care that the current does not exceed 2A.** Set the value at 1.4A.

7. Wait 2 minutes for the cathode to warm-up. Then apply an accelerating voltage of 200V. You will see the electron beam emerge from the electron gun and it will be curved by the field from the Helmholtz coils. You may need to increase the voltage to slightly above 200V if the beam does not initially emerge from the electron gun or initially makes only a very small ring (< 2 cm diameter).

8. **Check that the path of the electron beam is slightly helical so that it does not strike the back of the electron gun shield.** If it strikes the gun, the shielding box becomes charged and the electron velocity emerging from the gun can be noticeably lower as it requires extra energy to overcome the negative field produced by the shield. Why do you think it is possible to produce a helical path? If the trajectory is not helical and strikes the gun, then ask the TA to rotate the tube. As you rotate the tube, the socket will also turn – there is no need to

remove the tube from its socket. If needed, adjust the knob labeled "FOCUS" to get the sharpest trajectory.

9. Estimate the angle of the emerging electron beam with respect to the magnetic field. Resolve the velocity of the electrons into two components – one perpendicular to $\vec{B}$ and one parallel to $\vec{B}$. The perpendicular component is to be used in equations 2 - 4. Is the difference between the total speed and its perpendicular component significant? If it is, then the electron orbital radius is $|\vec{v}|\sin\alpha$, where α is the angle between magnetic field vector $\vec{B}$ and electron speed $\vec{v}$. You will then need to use a revised version of Eq. 4:

$$\frac{q}{m} = \frac{2V\sin^2(\alpha)}{B^2R^2}, \qquad (6)$$

10. You will next be making a series of (V,R) pair measurements which will be later used to calculate *e/m*. Think of the best way to record these data in your notebook.

11. Set the value of the current through the coils to 1.4A and the accelerating voltage to 200 V. Remember to read the value of the accelerating voltage from the voltmeter - not from the scale on the power supply. Carefully measure the radius of the electron beam. Look through the tube at the electron beam. To avoid parallax error, move your head to align the electron beam with the reflection of the scale.

12. Repeat measurements for the same value of current: I = 1.4A, but for the following values of accelerating voltage: V = 220V, 240V, and 260V.

13. Change the value of magnetic field B by setting the current through the coils to 1.6A. Repeat measurements for V = 220V, 230V, 260V and 290V.

14. Change the value of magnetic field B by setting the current through the coils to 1.8A. Repeat measurements for V = 220V, 230V, 260V, and 290V.

15. In order to obtain reasonable estimates for *e/m*, repeat the above set of measurements two more times.

16. Turn the current adjust knob for the Helmholtz coils and the knob labeled "DC VOLTAGE ADJUSTMENT" to the zero position (counterclockwise).

17. Turn the power off.

Data Analysis

The acquisition of data in this experiment is straightforward. The challenge is to interpret the data in a meaningful way and to arrive at realistic uncertainties for the various quantities that are used in the final computation of *e/m*. If the experiment is

performed as described above, you will have 12 sets of data from which to calculate *e/m*. By repeating the data acquisition three times, you will therefore have 36 independent data sets to estimate a best value for *e/m*.

Think about how best to present your data. Should you just perform a straightforward average of all 36 independent measurements? If you average the values of *e/m* obtained under the same experimental conditions, can you obtain an estimate for the accuracy of your experimental techniques? Are the *e/m* values calculated in your experiment the same as the known value of *e/m* within the error of experiment? Is there any evidence of systematic error? Note that there are only three measured values in this experiment: *R, B* (inferred from *I*) and *V*. The *R* measurement accuracy can be easily estimated. Using a gaussmeter, we have also checked the actual *B* values within the vacuum tube area. They were the same within 5% of the values predicted by Eq. 5. The only parameter which could not be independently verified was the energy of the electrons. If there is considerable discrepancy between the known and measured *e/m* values, estimate the actual energy of the electrons emerging from the electron gun as a function of accelerating voltage *V*.

Other issues that you might want to consider are estimates of the magnitude of the electron's velocity in this experiment. Are relativistic corrections required? Can you estimate the spread in the energy of the electrons in the beam by estimating the spread in radius of the beam's trajectory?

These are some of the issues you must consider when analyzing your data and writing your report. Whatever issues you choose to pursue, always try to be as quantitative as possible. At the end of the day, be sure to include a clear and concise discussion of your best estimate for *e/m* and how it compares to accepted values.

RR, SS 2001, ML 2012

Physics 340 Laboratory
Blackbody Radiation: The Stefan-Boltzman Constant

Objective: To measure the energy radiated by a blackbody cavity as a function of temperature.

References:

1. *Experimental Atomic Physics*, G.P. Harnwell and J.J. Livingood, McGraw-Hill, NY (1933), pgs. 50-58.
2. *Modern Physics, 2nd Edition*, Kenneth Krane, Wiley and Sons, NY (1996), pgs. 77-83 and pgs. 320-322.
3. *Radiation Processes in Astrophysics,* G. Rybicki and A.P. Lightman, Wiley-VCH, NY (1985), Chapter 1.

Apparatus: Electrical furnace, NiCr-Ni thermocouple, variac, CASSY interface, computer, water-cooled radiation shield and plastic tubing, Moll's thermopile.

Introduction:

Following James Maxwell's unification of electricity and magnetism in the late 1860's and his prediction of electromagnetic radiation, an intense effort followed to detect and generate this new type of radiation. After the realization that light itself was an electromagnetic wave, there was an explosion of interest to understand in detail how light was generated. This led to a series of fundamental studies characterizing many different types of light sources. The results of these studies were so puzzling that they eventually provided the groundwork for the formulation of quantum physics.

One subject of particular interest during this time was the characterization of light emitted from a hot object. As known from prehistoric times, any object heated to a high enough temperature emits visible light. As early as 1802, Count Rumford (Benjamin Thompson) suggested that blackening the surface of an object enhanced its thermal radiative properties. Consequently, the radiation emitted from a well-characterized object like a hollow cavity came to be known by a variety of names such as blackbody radiation, temperature radiation, or cavity radiation. It was discovered that such radiation depends only on the temperature of the cavity and this fact differentiates it from other types of radiation such as that emitted from a glow discharge tube. It was quickly realized that blackbody radiation emitted in the visible region of the electromagnetic spectrum only becomes appreciable when the temperature of the cavity is above 500-550 °C (about 800-850 K). Blackbody radiation emitted at lower temperatures had to be detected by other than optical means.

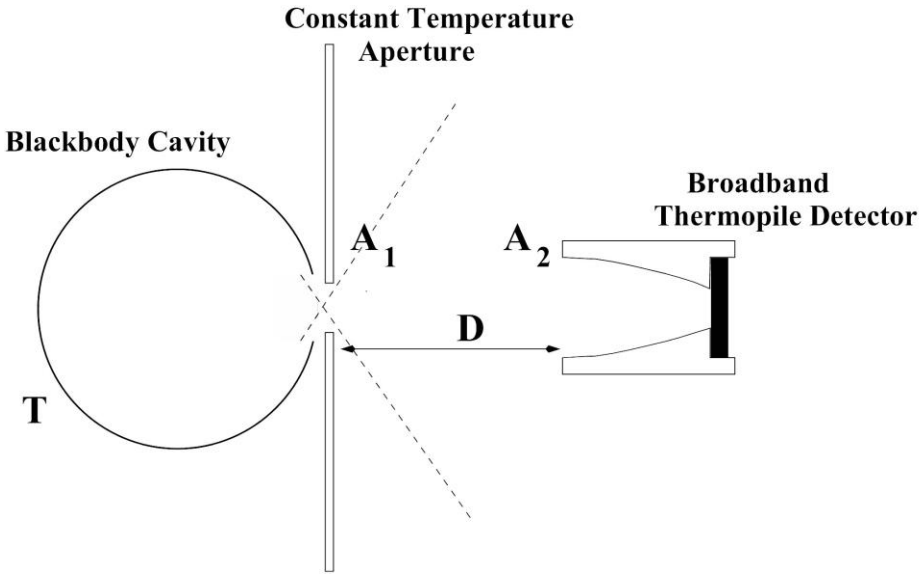


Figure 1: A schematic diagram showing the essential ingredients of the Stefan radiation experiment. A cavity held at temperature T is shielded by a constant temperature shield with an aperture area A_1 from a broadband radiation detector having an active area A_2.

The first quantitative measurements of the energy transferred by radiation between a body and its surroundings were reported by John Tyndall in 1867. Whereas traditional transfer of energy between two objects via convection and conduction process requires an intervening material medium, heat radiation does not. The energy is transferred by electromagnetic waves. A diagram illustrating the essential features of such a radiation experiment is shown in Fig. 1. In 1879, Josef Stefan's work suggested that the luminosity (total emitted power at all wavelengths, in Watts) of a blackbody cavity of surface area *A* varied with the temperature of the cavity as T^4, giving an empirical relationship known as Stefan's Law:

$$L = A\sigma T^4 , \qquad (1)$$

where σ is a proportionality constant now known as the Stefan-Boltzmann constant.

No fundamental understanding of this empirical result was achieved until five years later, when Ludwig Boltzmann used the theory of radiation pressure and the laws of thermodynamics to derive the correct temperature dependence inferred by Stefan. The significance of σ was not appreciated until 1900, when Max Planck realized that it could be expressed in terms of a combination of fundamental constants. Experimentally, σ was measured with increasing precision from the 1890's ($\sigma=5.45\times10^{-8}$ W m^{-2} K^{-4}) to the 1930's ($\sigma=5.737\pm0.017\times10^{-8}$ W m^{-2} K^{-4}). Thus knowing σ and the surface area of any object (assumed to be a blackbody), the power emitted into a vacuum can be calculated. Today, the Stefan-Boltzmann law plays a prominent role in astrophysics, allowing the surface temperatures of distant stars and

planets to be inferred from measurements of their distances and apparent brightnesses.

In this experiment, you will repeat Stefan's measurements using computer-assisted data acquisition techniques and you will obtain an estimate for the Stefan-Boltzmann constant σ.

Theory:

At the close of the nineteenth century, a major unresolved problem in classical physics involved a major discrepancy between the predicted spectrum of a blackbody and what was actually observed in the lab. By using Maxwell's equations and the laws of thermodynamics, Lord Rayleigh had derived that a blackbody spectrum (the energy emitted per unit frequency) should vary as $\nu^4 T$, where ν is the observed photon frequency. This prediction held true for low frequencies, but unfortunately it predicted that an ever increasing number of photons would be produced at the highest frequencies (the so-called 'ultraviolet catastrophe').

In late 1900, Max Planck came upon a brilliant solution to the problem, by making a crucial modification to the blackbody model. Whereas Rayleigh had described the photons inside a blackbody cavity as standing waves with nodes at the cavity walls, each with equal energy kT, Planck suggested that the energies of the standing waves were in fact quantized, and were integer multiples of hc/λ, where h is Planck's constant ($6.62606957(29)\times10^{-34}$ J s). Thus, at high frequencies (i.e., short wavelengths such as ultraviolet), each photon carried significantly more energy, and since photons will tend to occupy the lowest energy states first according to statistical mechanics, comparatively few UV photons will be produced for a given amount of thermal energy available in the blackbody cavity. This solution to the ultraviolet catastrophe cemented the failure of classical physics as a complete description of nature and was instrumental in launching the modern era of quantum mechanics.

Planck's formulation involved a consideration of the density of quantum states available to the photons in the blackbody cavity per unit frequency interval. To derive the average energy per photon state he used the Boltzmann factor ($e^{-E/kT}$) and his quantum wave energy $h\nu$. Multiplying all these quantities together, and accounting for two possible polarizations of the photons, his predicted blackbody spectrum was

$$B_\nu = \frac{2h\nu^3}{c^2(e^{h\nu/kT}-1)}\ , \qquad (2)$$

which is now known as the Planck function (a more comprehensive derivation can be found in the list of references accompanying this write-up).

In this experiment you will be measuring two main properties of a blackbody: its internal temperature (using the thermocouple), and its emitted radiation field (using the thermopile). The latter is traditionally described using such quantities as specific

intensity (brightness), flux, and luminosity. We have already defined luminosity as the total energy output of a blackbody over all frequencies. This is typically impossible to measure in the lab, as it would require you to capture and record every single photon emitted. Instead you will be observing the blackbody radiation (in our case, from an oven) that escapes through a small aperture of area A_1, using a thermopile having an opening aperture area A_2 , located a distance D away (Figure 1).

It is convenient to describe the measured energy *dE* from a small bundle of light rays crossing a small aperture of area *dA* in a time *dt* in terms of a *specific intensity* I_ν, which is defined such that

$$dE \equiv I_\nu \, dA \, dt \, d\Omega \, d\nu \, . \qquad (3)$$

In this formulation the directions of all the rays lie within a small solid angle $d\Omega$, are approximately perpendicular to the aperture *dA*, and have photon frequencies lying within a small range dν. The *solid angle* is the two-dimensional equivalent of a standard angle, and refers to the apparent area subtended by an object. For example, the Moon has an apparent radius of 0.25 degrees (4.36×10^{-3} radians) as seen from Earth, and thus subtends $\pi (4.36 \times 10^{-3})^2 = 6 \times 10^{-5}$ *steradians* on the sky.

Eq. (3) provides the definition of specific intensity I_ν (often called *brightness*) in units of W m^{-2} Hz^{-1} $steradian^{-1}$, and pertains to rays propagating only along a very narrow direction. In real life however, glowing objects emit rays in all directions, many of which will make it to our detector. Imagine you set up a small flat aperture of small area dA whose normal lies at some angle θ to the surface of a blackbody as in Figure 2. Then we can define the radiative *flux* from the blackbody as

$$F_\nu = \int I_\nu \cos\theta \, d\Omega , \qquad (4)$$

which in spherical coordinates becomes

$$F_\nu = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} I_\nu \cos\theta \sin\theta \, d\theta \, d\phi \, . \qquad (5)$$

You may notice that flux is an *observer-based* quantity, since to evaluate the right hand side, we integrate the brightness of the object over all possible directions – however I_ν is non-zero only along those directions where our line of sight intercepts the object. In other words, if we move the object farther away, its solid angle (and therefore its flux) decreases as $1/D^2$, giving rise to the familiar inverse-square law for light.

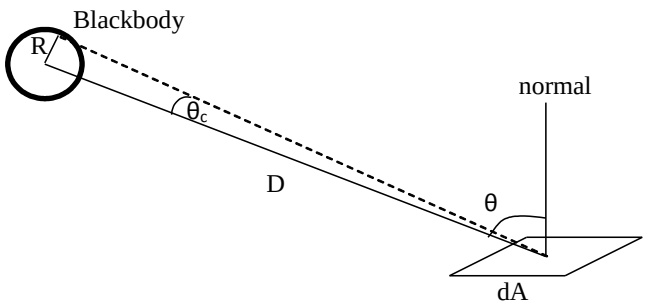


Figure 2. Sketch showing the geometry used in the definition of radiative flux. The effective area of the detector is proportional to the cosine of the observation angle θ. The blackbody subtends an apparent angular radius $\theta_c = \sin^{-1}(R/D)$.

To illustrate this effect, imagine locating your detector a distance *D* from a spherical blackbody of radius *R* and uniform surface intensity *I* (let's assume that your detector can receive all photon frequencies, and has its normal vector pointing directly at the blackbody). According to Eq. (5), the measured flux will be

$$F = I\int_{\phi=0}^{2\pi} d\phi \int_{\theta=0}^{\theta_c} \cos\theta \sin\theta \, d\theta, \qquad (6)$$

where from simple geometry, $\theta_c = \sin^{-1}(R/D)$ is the apparent angular radius of the blackbody. Performing the integration gives

$$F = \pi I \left(\frac{R}{D}\right)^2. \qquad (7)$$

Thus for a fixed spherical blackbody of surface intensity *I* and radius *R*, the flux varies as $1/D^2$, confirming the inverse-square law of light.

We can now use Eq. (7) to derive the flux per unit area radiating from the surface of a spherical blackbody by setting *R* = *D*:

$$F = \pi I \ . \qquad (8)$$

Multiplying this by the surface area *A* of the blackbody yields its total luminosity:

$$L = A\pi I \quad . \qquad (9)$$

Comparing this with Eq. (1), we can see that for a spherical blackbody

$$I = \frac{\sigma T^4}{\pi}. \qquad (10)$$

Now according to Planck's radiation law, the left hand side of Eq. (10) should be equivalent to Planck's function B_ν (Eq. 2) integrated over all photon frequencies. An integration of Eq. (2) by parts gives the predicted blackbody luminosity per unit surface area:

$$B = \frac{2\pi^4 k^4 T^4}{15c^2h^3} \quad . \qquad (11)$$

Comparing this to Eq. (10), we find that the Stefan-Boltzmann constant is in fact a combination of several other fundamental constants:

$$\sigma = \frac{2\pi^5 k^4}{15c^2h^3}, \qquad (12)$$

which has an experimentally determined value of $5.670373(21)\times10^{-8}$ W m^{-2} K^{-4}.

In practice, any object (not necessarily a hollow cavity) heated to a temperature T emits radiation. Experiment shows that the maximum power radiated per unit area comes from a cavity and is specified by Eq. (2). The emitted luminosity *per unit area* of any arbitrary object held at temperature T is specified by the non-ideal blackbody formula

$$L = \varepsilon\sigma T^4 , \qquad (13)$$

where ε is known as the emissivity of the emitting object. By definition, an ideal blackbody has $\varepsilon=1$. Every material has a characteristic emissivity; for instance, tungsten has $\varepsilon=0.2$.

Experimental Considerations:

It is difficult to accurately measure the **total** emissive power radiated from a blackbody. This would require a detector that collects the radiation emanating from the blackbody in **all** directions. In general, only a small fraction of the emitted energy can be collected. Therefore, careful attention must be paid to the geometrical arrangement of the blackbody and detector in order to allow a correct interpretation of the data. Also, careful measurements must take into account any energy absorbed by the intervening air column, which can be difficult to control if air conditioners or fans are continually circulating air through the laboratory room.

In addition, a broad-band detector must be designed to accurately measure the incident power over a wide range of wavelengths. In this experiment, you will use a thermopile to accomplish this function. A thermopile is a blackened disk of known dimension that is thermally anchored to a series of thermocouple junctions. A thermocouple is a junction formed when two wires, each made from a dissimilar metal, are joined together in an intimate fashion. A thermocouple junction is known to develop a voltage across it that depends on temperature. In this way, the temperature of the disk can be measured in terms of a voltage difference that

develops across a thermocouple junction. A thermopile refers to a number of thermocouple junctions connected in series (see Fig. 3). This serves to increase the voltage output of the device, allowing more accuracy in any measurement.

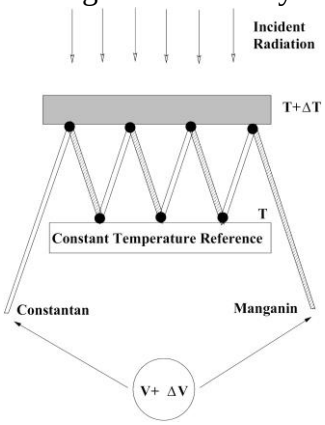


Figure 3: A schematic diagram showing the essential elements of a thermopile comprised of 4 thermocouple junctions made from constantan-manganin wire.

A thermopile must be calibrated by the manufacturer. Using this calibration, a measured thermopile voltage can be converted into an incident power (or received flux if the active area of the thermopile is known). For the thermopiles that you will use, the calibration lies between ~30 μV/(W m^{-2}) and ~70 μV/(W m^{-2}) for homogeneous irradiance of the front surface (see the Appendix). The Moll's thermopile has sixteen thermocouples comprised of constantan and manganin wires connected in series. It is sensitive to radiation spanning a range from 150 nm to 15 μm and has a response time of 2-3 seconds.

A schematic diagram of the relevant geometry employed in this experiment is sketched in Fig. 1. The blackbody radiation emerges from a circular hole of radius $r = (A_1/\pi)^{1/2}$. The detector is separated from the cavity opening by a distance D. In analogy with our earlier derivation of Eq. (7), the flux received at the thermopile is

$$F = \pi I \left(\frac{r}{D} \right)^2 . \qquad (14)$$

We now substitute $I = \sigma T^4/\pi$ (Eq. 10) and $r = (A_1/\pi)^{1/2}$ into Eq. (14) to give

$$F = \frac{A_1 \sigma T^4}{\pi D^2} . \qquad (15)$$

Thus a plot of the energy detected per unit time versus T^4 should be a straight line if Stefan's Law is correct. It follows that the Stefan-Boltzmann constant can be determined from the best slope ***m*** through the data using

$$\sigma = m\frac{\pi D^2}{A_1}. \qquad (16)$$

Experimental Technique:

First, familiarize yourself with the data acquisition software that you will use with this experiment. The CASSY system is designed to allow flexible and easy data acquisition from this important radiation experiment.

Fig. 4 is a photograph of the assembled Stefan radiation equipment and Fig. 5 is a schematic wiring diagram.

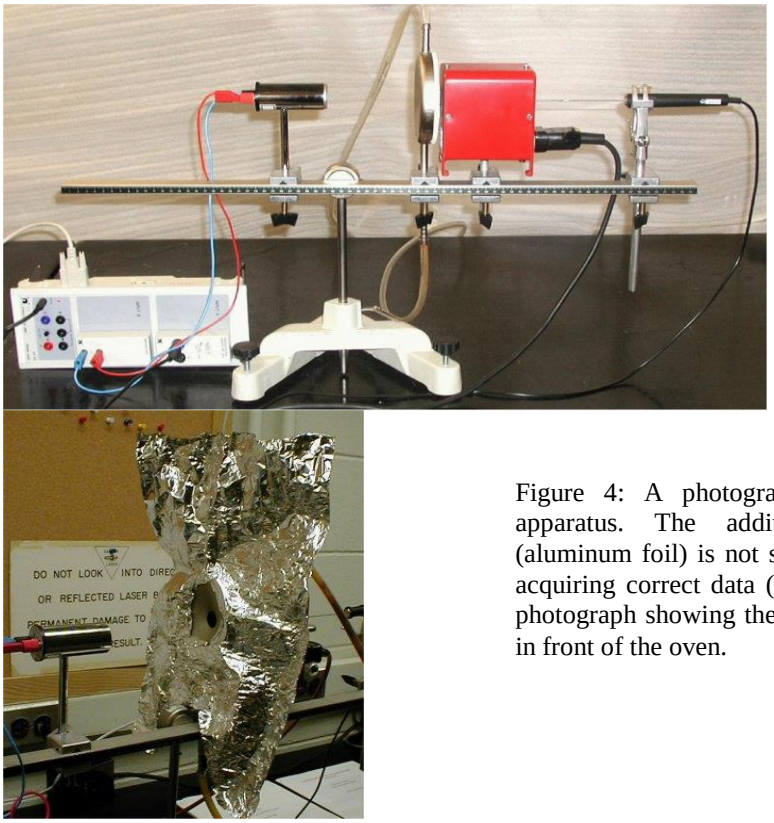


Figure 4: A photograph showing the Stefan apparatus. The additional radiation shield (aluminum foil) is not shown but is essential for acquiring correct data (see text). At the left is a photograph showing the radiation shield installed in front of the oven.

A special blackbody insert must be installed into the oven. The metal disk with the hole in front of the oven is cooled by water. Failure to keep it at room temperature will result in additional radiation emitted from the disk and will make the data acquired unusable. Use the small pump to circulate water through the disk. Since some of the parts of the oven are not covered by this disk, an additional radiation shield made of aluminum foil should be attached (Fig. 4, bottom).

Connect a variable voltage supply (variac) to the oven (the blackbody cavity) and insert a NiCr-Ni thermocouple into the back-side of the oven. Position the thermopile

approximately 0.15 m in front of the oven. *Make sure that the protective glass cover is removed from the entrance of the thermopile.*

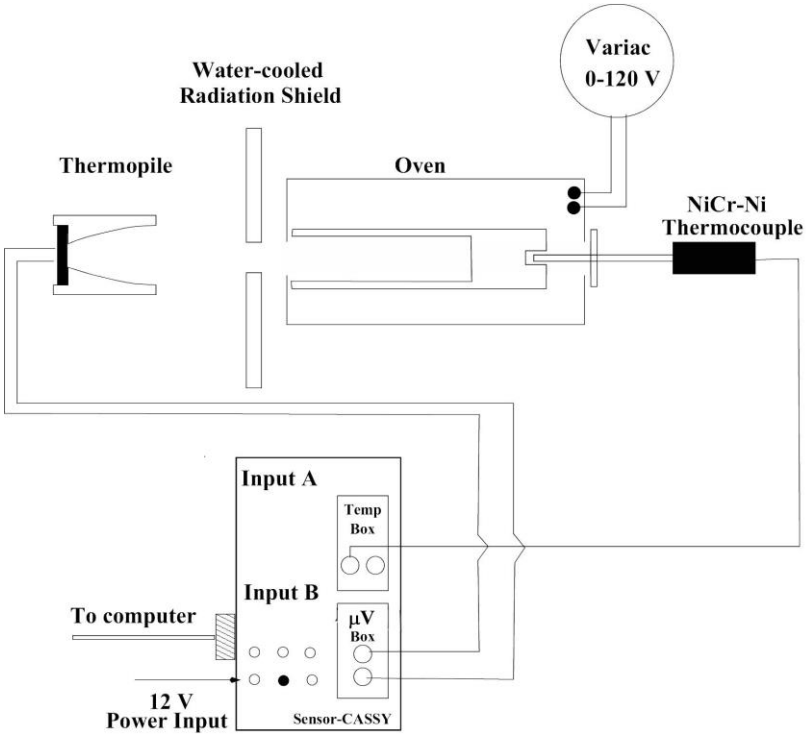


Figure 5: A schematic wiring diagram for the Stefan apparatus.

Data Acquisition Procedure:

(a) Note the manufacturer's serial number of your thermopile, and then record its output voltage with the oven at room temperature. At this time, make sure you know how to acquire data with the CASSY system. This measurement serves as the zero-point reference for all your future measurements. Set the distance D between the detector and the cavity opening to about 0.15 m. Measure the diameter of the cavity opening and calculate its area (A_1). Also, be sure you write down the ambient temperature of the room.

(b) Set up the CASSY system to record the oven temperature and the thermopile voltage. Set the display to show temperature as *x*-axis and thermopile voltage as *y*-axis. Set the measurement interval to 500 ms and the recording condition to *n=1 or delta(&JA11)>2* (the command is case sensitive). The latter condition means that the measurement should be taken *only* when the first point is acquired (*n=1*) or when temperature (&JA11) increases by 2. See that the number of points to be accumulated is left blank, or set it to maximum possible value. Also, set the measurement mode to *Average 100 ms* for both temperature and voltage detector. This serves to minimize the effects of short-timescale voltmeter noise in your measurements.

(c) Start the measurement and turn the variac up to ~100 V. At this setting it will take approximately 20 minutes for the temperature to reach 400 °C. At the highest temperatures, the outside of the oven is very HOT, so do not touch it without oven mitts. The data will be accumulated automatically.

(d) When the oven temperature reaches a constant temperature of approximately 400 °C, stop the measurement and decrease the variac voltage to ~90V. Save the data you just acquired.

(e) See that the temperature stabilizes at ~400 °C, tune the variac voltage slightly down if temperature keeps rising, or up if it goes down (by few volts). Note that it may take a minute or so before changes in voltage reflect in temperature change. Once the temperature is reasonably stable, make a series of measurements in which you vary the distance between the thermopile and the entrance to the blackbody. Be sure to cover distance between ~5 cm to the full length of the rail and acquire about 20 points. Remember to wait after each move before you take the measurement as the detector has long time constant.

(f) Choose another distance (greater than the 0.15 m used in step (b)) between the thermopile and the entrance to the blackbody. Set up the measuring condition to *n=1 or delta(&JA11)<-2.* Start new data acquisition (the previously accumulated data will be cleared automatically), set the variac to zero, and then turn off the variac. The new set of data will be recorded as the oven returns to room temperature.

(g) When you are finished, make sure you turn off the cooling water, replace the protective glass cover over the entrance of the thermopile, and copy any useful data onto a thumb drive or your network account for further analysis.

Data Analysis:

(a) Analyze your data as function of distance between the thermopile and the entrance to the blackbody. How well does it follow a D^{-2} relationship expected from Eq. 16? For what values of D would you expect a D^{-2} behavior to accurately hold? For each distance, calculate the value of σ using Eq. 16. Plot your calculated value of T as a function of distance. What can you say about this graph?

(b) Analyze your data as a function of temperature during the warming phase of the oven. Remember to correct your data for the energy radiated away from the thermopile detector at room temperature. This requires you to plot $F_{in}(T)$-$F_{in}(T_{RT})$ vs. $(T^4$-$T_{RT}{}^4)$, where T_{RT} is the temperature of the room. From this plot, measure the slope and extract an estimate for the Stefan-Boltzmann constant. Alternatively, you may analyze the original data using Eq. 16, and estimate the value of σ from that. Calculate the value of σ for several temperatures and plot σ versus T. Is it constant? If not, what might be the reason?

(c) In the same way, analyze your data as a function of temperature during the cooling phase of the oven.

(d) Tabulate your two estimates of σ in a clear way. What is your best estimate for the Stefan-Boltzmann constant? Make sure you include a realistic discussion of errors in your measurements of A_1 and D. What is the resulting uncertainty in σ?

(e) Discuss any significant sources of unaccounted error that you believe are relevant to this experiment. Is it possible that these errors account for the difference between your measured value of σ and the accepted value?

Appendix: Moll's Thermopile Calibrations

	Manuf. Specs.	**Calibrated by SS (Fall 2001)**
Serial No. 999415:	35.5 μV/(W m^{-2})	52±5 μV/(W m^{-2})
Serial No. 009561:	44.3 μV/(W m^{-2})	66±5 μV/(W m^{-2})
Serial No. 009562:	63.4 μV/(W m^{-2})	92±5 μV/(W m^{-2})
Serial No. 009576:	39.7 μV/(W m^{-2})	59±5 μV/(W m^{-2})

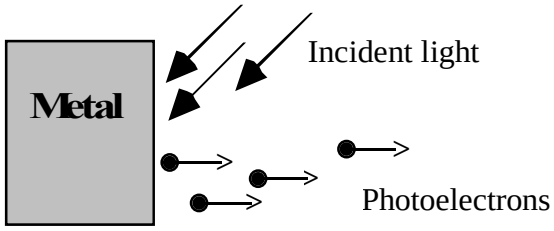
Physics 340 Laboratory
Quantization of the Radiation Field: The Photoelectric Effect

Objectives

- Understand how ejection of photoelectrons from a metal surface depends on the frequency and intensity of incident light.
- Measure Planck's constant and the work function.

Theory[1]

The photoelectric effect is one of several processes by which electrons may be removed from the surface of a metal. It is found that when electromagnetic radiation with sufficiently high energy is directed onto a metallic surface, electrons may be ejected from the surface. The ejected electrons are called photoelectrons.



The key experimental facts about the photoelectric effect are the following.

(1) The emission of photoelectrons will not occur at all, if the frequency of the incident electromagnetic radiation is less than a certain frequency called the cutoff frequency. Below the cutoff frequency there would be no photoelectrons regardless of the intensity of the incident light.

(2) The maximum kinetic energy of the ejected photoelectrons depends on both the frequency of the incident electromagnetic radiation and the metal itself.

(3) If the frequency of the incident radiation is larger than the cutoff frequency, then the number of emitted photoelectrons is proportional to the intensity of the radiation.

[1] Nicholas J. Giordano, *College Physics*, Ed. 2, Brooks/Cole, Cengage Learning, 2013, pp. 982-990.

Because classical electromagnetism is unable to explain these experimental facts, a quantum theory of light is required for a satisfactory explanation.

A. Electrons in Metals

Although the outermost valence electrons in a metal are free to move within a metal (that is why metals usually are good conductors), the electrons are constrained within the metal. The binding energy of the electrons least tightly bound to the metal is denoted by the work function W_0 of the metal. If energy greater than the work function is supplied to one of these electrons, that electron can be ejected from the metal surface. This process is known as the **photoelectric effect.**

B. The Incident Light Rays

The electromagnetic light rays that strike the metal surface can supply the outer electrons with the energy they need to leave the metal. According to quantum theory, these apparently continuous electromagnetic waves are actually quantized, consisting of discrete quanta called **photons**. Each photon has energy E that depends only on its frequency f (or wavelength λ) and is given by the following equation.

$$E = hf = \frac{hc}{l} \quad , \text{ where: } \quad h = \text{Planck constant} = 6.63 \times 10^{-34} \text{ J} \times \text{s} = 4.14 \times 10^{-15} \text{ eV} \times \text{s}$$

$$1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}, \qquad 1 \text{ J} = 6.242 \times 10^{18} \text{ eV}, \qquad c = 2.998 \times 10^{8} \text{ m/s}$$

The energy of photons is directly proportional to their frequency. In the photoelectric effect, a single photon interacts with a single electron at the metal surface; it cannot share its energy among several electrons. Photons with a frequency f greater than the cutoff frequency f_0 will cause the emission of photoelectrons, whereas photons with a frequency less than f_0 will not cause the emission of photoelectrons.

C. The Photoelectric Effect Equation

In photoelectric emission, light strikes a material, causing electrons to be emitted. The classical wave model predicted that as the intensity of incident light

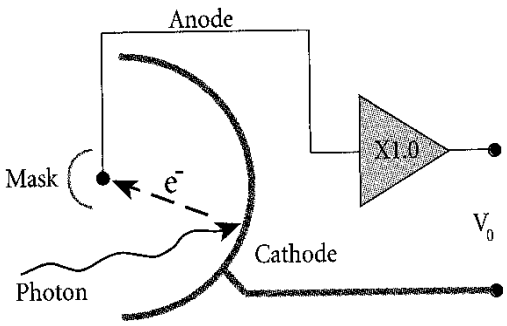
was increased, the amplitude and thus the energy of the wave would increase. This would then cause more photoelectrons to be emitted. The quantum model, however, predicted that higher frequency light would produce higher energy photoelectrons, independent of intensity, while increased intensity would only increase the number of the electrons emitted (or photoelectric current).

Einstein applied Planck's theory and explained the photoelectric effect in terms of the quantum model using his famous equation for which he received the Nobel Prize in 1921. The equation states that the maximum possible kinetic energy of ejected electron (KE_{max}) is equal to the energy of the incident photon (hf) minus the minimum work needed to eject electron (the work function W_0).

$$KE_{\max} = hf - W_0$$

The photoelectrons measured outside the metal surface will have kinetic energies ranging from 0 to $KE_{\max}$. One may notice that the photoelectric equation is basically another way to express the conservation of energy principle!

D. Measuring the Work Function and the h/e Ratio Using a Photocell



Some photoelectrons travel toward the stopping electrode (anode) and, upon reaching it, constitute a current, which flows through the circuit. As we increase the voltage V, some of the less energetic photoelectrons will be repelled from the anode and the current will decrease. At some voltage $V = V_S$ (the stopping voltage), the most energetic photoelectrons will be stopped just in front of the anode and current will cease to flow. For this condition:

$$eV_s = KE_{\max} \qquad \{e = \text{the charge of an electron} = 1.60 \times 10^{-19} C\}.$$

Therefore, using Einstein's equation,

$$hf = eV_S + W_0 \quad (1)$$

When solved for V_S, the equation becomes:

$$V_S = (h/e)f - (W_0/e) \quad (2)$$

If we draw V_S versus f for different frequencies of incident light, the graph will show the linear dependence, with the slope equal to h/e and the y-intercept equal to $-W_0/e$. The accurately measured value of the h/e ratio is equal to:

$$h/e = 4.136*10^{-15} \text{ V*s} \quad (3)$$

We are going to use a mercury vapor lamp as the light source for the photoelectric experiment. The light from the mercury lamp looks white. However, the mercury lamp does not produce a continuous spectrum (a rainbow) of all colors from red to violet. Instead, it produces only four lines (frequencies) in the visible range and one in the near ultraviolet. We will use a diffraction grating (like the one used in the diffraction grating experiment) to separate various frequencies (i.e., colors) of the light.

E. Why the *h/e* ratio is important?

The ***h/e* (Planck's constant / electron charge magnitude)** is the ratio of two fundamental constants. These constants are present in all theories of atomic, quantum, and elementary particle physics. It is essential to know their numerical values! The electron charge magnitude e was measured before the photoelectric effect was discovered. However, the photoelectric effect was one of the first measurements of h/e (and therefore the value of h).

How does this apparatus work?

The photodiode tube and its associated electronics have a small capacitor, which becomes charged by the photoelectric current. When the potential on this capacitance reaches the stopping potential of the photoelectrons, the current decreases to zero and the anode-to-cathode voltage stabilizes. This final voltage between the anode and cathode is therefore equal to the stopping potential of the photoelectrons.

To let you measure the stopping potential, the anode is connected to a built-in amplifier with ultrahigh input impedance ($>10^{13}$ W), and the output from this

amplifier is connected to the output jacks on the front panel of the apparatus. This high impedance, unity gain (V_{OUT}/V_{IN} = 1) amplifier allows us to measure the stopping potential with an ordinary digital voltmeter. That amplifier does not actually amplify the voltage. Its sole purpose is to provide the ultrahigh input impedance, i.e., to measure the potential difference (almost) without using any current.

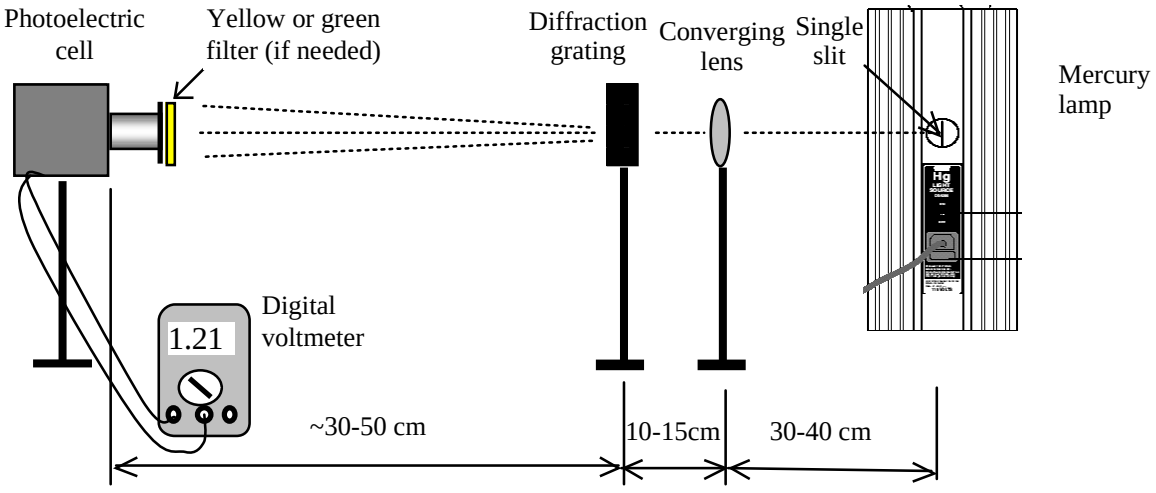
Due to the ultra high input impedance, once the capacitor has been charged from the photodiode current, it takes a long time to discharge this potential through some leakage. Therefore, a shorting button "PUSH TO ZERO" enables the user to quickly remove the charge and to reset the apparatus.

Procedure:

Activity 1: Measurements of the h/e Ratio and the Work Function.

1.1. You will need to use the following pieces of equipment: mercury lamp with a single slit; small converging lens; diffraction grating; the photoelectric cell apparatus (also called the "*h/e apparatus*") and a digital voltmeter.

1.2. First, assemble the pieces of the apparatus according to the drawing below:



1.3. The distances shown in the above picture are approximate. You will have to adjust them more precisely later during this experiment.

1.4. Make sure that a piece of the black cardboard with a single vertical slit is covering the front opening of the mercury lamp.

1.5. Turn on the mercury lamp and allow 2-3 minutes for warm up. **Do not turn off this lamp until the experiment is over**.

CAUTION! The light from the lamp contains some ultraviolet (UV) components, which can burn living tissue. Therefore, you should *avoid looking directly at the lamp*, and, if you do not already wear glasses, you should *wear the plastic safety glasses provided for you* (glass and plastic are opaque to UV light).

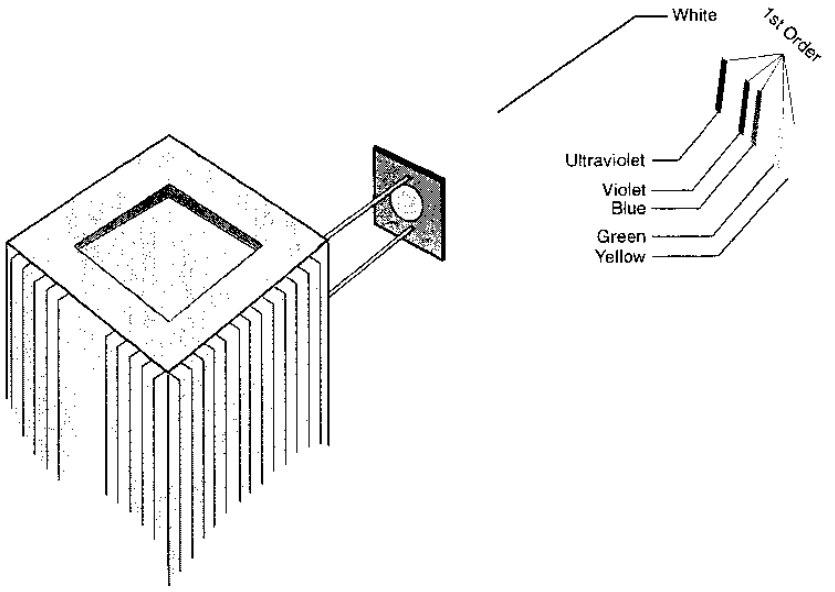
1.6. Make sure that the digital multimeter is connected to the photoelectric cell and that the multimeter is turned on and set to 2 V DC (V=) range.

1.7. Make sure that the photoelectric cell is turned on. Actually, the on/off switch turns on the high impedance amplifier.

1.8. Adjust the position of the converging lens to get a sharp, white image of the single slit (the one in front of the mercury lamp) on the small white reflective mask of the photoelectric cell.

1.9. The light from the mercury lamp looks white. However, the mercury lamp does not produce a continuous spectrum (a rainbow) of all colors from red to violet. Instead, it produces only four lines (frequencies) in the visible range and one in the near ultraviolet. This is called **line spectrum** for Mercury. The values of frequency and the corresponding wavelength of the line spectrum for Mercury are:

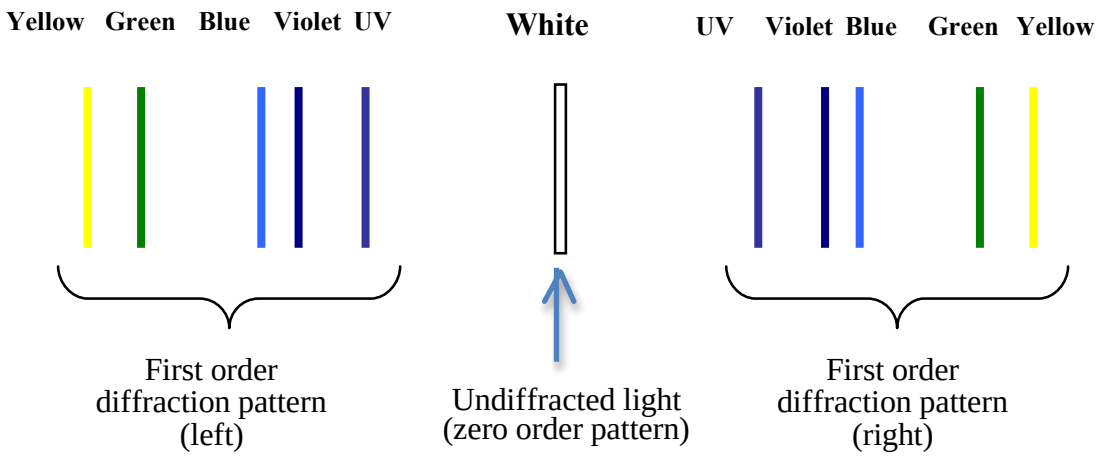
Color	Frequency (Hz)	Wavelength (nm)
Yellow (visible)	$5.187*10^{14}$	578.0
Green (visible)	$5.490*10^{14}$	546.1
Blue (visible)	$6.879*10^{14}$	435.8
Violet (visible)	$7.409*10^{14}$	404.7
Ultraviolet (hardly visible)	$8.203*10^{14}$	365.5



1.10. **The diffraction grating separates these lines**, so we will use only one color (i.e., one wavelength) at a time. Please recall your observations from the last part the "Diffraction Grating" experiment. To make the separation of lines even better, **use yellow and green filters for yellow and green lines, respectively**. These filters have magnets built in, so they can be easily attached to the front of the white reflective mask (see the text and the picture below).

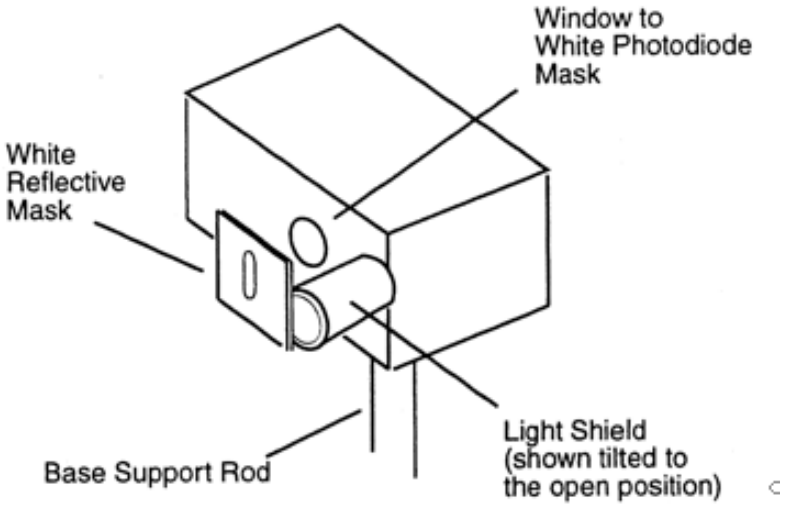
When making this experiment, it is essential that only a single color (i.e., single frequency) goes through the photocell window. With two colors reaching the photocell, we would not know the frequency of the incident light.

1.11. The diffraction grating creates two first-order patterns, one on the left side of the white, un-diffracted line, and one on the right side.



The second order patterns are also visible and they are located further away from the un-diffracted white line. The order of colors for the second order pattern is the same as for the first order pattern.

1.12. **Move** the photoelectric cell until the yellow light shines directly on the opening in the "White Reflective Mask" (see the picture below). **Rotate the photocell so that the same light that falls on the opening in the "White Reflective Mask" also falls on the window in the photodiode mask**. You need to open the "Light Shield" to be able to see the photodiode mask that is inside the black enclosure. Careful adjustment of the position and orientation of the photocell is essential to get good results!



1.13. **Close the "Light Shield"**. Then, **press** the red "PUSH TO ZERO" button on the side panel of the photocell box to discharge any accumulated potential in the unit's electronics. This will assure that the apparatus records only the potential due to the light shining on the photocell. Be careful not to change the position of the photocell when you push the red button!

1.14. **Wait** approximately 30 sec. to get a stable reading on the voltmeter. Then, write the voltage value on your data sheets. **It is a direct measurement of the stopping potential V_S for the photoelectrons.**

1.15. Repeat steps 1.12 - 1.14 for green, blue, violet, and ultraviolet light. Use filters only for yellow and green light.

Note: The white reflective mask on the photocell is made of a special fluorescent material. This allows you to see the ultraviolet line as a blue line, and it also makes the violet line appear bluer. You can see the actual colors of the light if you hold a piece of white paper in front of the mask.

1.16. Move the photocell to the other side of the zero order maximum (white line) and repeat measurements twice for all five colors. In other words, repeat steps 1.12 - 1.15 for the other set of the first order diffraction pattern.

1.17. Compare the values of the stopping voltage from both data sets. If your data from both runs are close to each other (ask your TA if you are not sure), then calculate the average values of the stopping voltage V_{S1}. *Hint:* for the UV line the stopping voltage should be ~2 V.

1.18. Make a graph of the average stopping voltage V_{SAV} vs. frequency f. Find and draw the best-fit straight line (do **not** just connect the points!) that approximates the behavior of your points (again, for both values of light intensity).

You should prepare the final version of the graph using a computer-graphing program (e.g., MS Excel that is available in all ITaP labs). These programs offer 'linear fit' or 'trendline' options to obtain the value of the slope and the y-intercept of the best-fit line.

1.21. Using the slope and the y-intercept values from the graph, calculate the values of the h/e ratio and the work function W_0. See the Theory section, part D. Be sure to include units.

1.22. **Turn off the mercury lamp, photocell and digital multimeter**.

Conceptual Questions

1. Assume that the frequency of the incident light is large enough to overcome the work function, i.e., the photoemission of electrons is already occurring. How does the number of photoelectrons emitted from a metal surface change as the frequency of the incoming light increases? Does the number of photoelectrons emitted from a metal surface depend on the work function of the metal? Does it depend on the light intensity?

2. Consider two photoelectric photocells that have cathodes made from different materials. The incident light has the same frequency for both photocells. Would the stopping voltage be the same for both photocells?

RR Oct 2001
SS Dec 2001
MJ Oct 2009
ML Oct 2012

Physics 340 Laboratory
Scattering of Photons from Electrons: Compton Scattering

Objective: To measure the energy of high energy photons scattered from electrons in a metal as a function of scattering angle.

References:

1. A.H. Compton, Phys. Rev. **21**, 715 (1923)
   A.H. Compton, *The Spectrum of Scattered X-Rays*, Phys. Rev. **22**, 409 (1923)
2. A.C. Melissinos, *Experiments in Modern Physics*, Academic Press, New York, 1966, p. 252-65.
3. K. Krane, *Modern Physics*, 2nd Ed., Wiley and Sons, New York, 1996, p. 83-87.

Apparatus:

- Set of low-activity (1 μCi) ɣ-ray calibration sources (^{22}Na, ^{54}Mn, ^{57}Co, ^{60}Co, ^{109}Cd, ^{133}Ba, and ^{137}Cs)
- Photo-multiplier tube attached to a NaI(Tl) scintillator crystal with a lead shield
- High voltage power supply
- Multi-channel analyzer PC peripheral interface
- High-activity (5 μCi) ^{137}Cs source encased in a lead shielding/collimator
- Aluminum rod for scattering ɣ-rays
- Movable carriage for changing observing angle

Introduction:

In 1923, Compton considered the problem of high energy photon (ɣ-ray) scattering from solids. Experimentally, he found that low energy (~ a few MeV) monochromatic photons scattered by metals change their frequency and that the frequency change depends on the scattering angle. This proved to be problematic, since at that time, light scattering was understood in terms of diffraction in which the scattered (diffracted) wave does NOT change frequency. Compton's experiments and his theoretical analysis of them came to be known as Compton scattering. Historically, his experiments are important because they provided further compelling evidence that photons do behave as particles which obey conservation of momentum and energy laws. Compton was awarded the Nobel prize in 1927 for his seminal work.

Compton's experiment can be understood by considering the interaction of the incident photons with the electrons that comprise a metal. If the quantized nature of

electromagnetic radiation is taken into account (electromagnetic radiation consists of photons, each of which has the same energy, $E = h\nu$, where ν is frequency and h is Planck's constant), and relativistic kinematics are used to describe the scattering process, the change in wavelength is understandable as a straight forward consequence of total energy and momentum conservation during a scattering process in which an incoming photon loses some of its energy to an electron with mass m_e. The basic kinematic diagram illustrating this interaction is sketched in Figure 1.

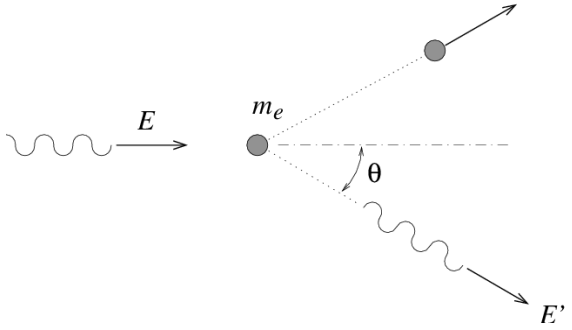


Figure 1: A schematic diagram showing the kinematic variables used to describe the scattering of an incident photon with energy E from an electron with mass m_e, initially at rest.

For a beam of incident photons, each of which has the same energy $E = h\nu$, there will be photons emerging at various angles θ with respect to the incident photon direction. The energy E' of a photon emerging at an angle θ can be calculated using relativistic kinematics and is described by the expression

$$E' = \frac{E}{1 + \left(E / m_e c^2\right)\left(1 - \cos\theta\right)}. \qquad (1)$$

From Equation 1 it can be seen that in order to obtain a large Compton shift (*i.e.*, a large value of $E - E'$), the incident photons should have an energy E that is comparable to the rest-energy of the electron: $m_e c^2 = 511$ keV. In this experiment, you will be investigating how an aluminum rod scatters a collimated beam of 662 keV gamma rays emitted by a ^{137}Cs source.

Experimental Considerations

NaI(Tl) Crystal Scintillator

The energies of gamma rays from the decays of radioactive isotopes can be measured using an inorganic crystal scintillation detector. A crystal of sodium iodide, doped with a small admixture of thallium, is used as the active detector element. An incident photon can scatter from the electrons in the crystal, which then deposit their energy in the crystal by ionizing other atoms in the crystal lattice. The electrons that are liberated in this way eventually recombine with the holes left in the lattice and emit photons with a range of wavelengths that peaks at about 400 nm, in the violet region of the visible spectrum. In this way, the number of photons produced is proportional

to the energy of the incident photon. Because sodium iodide is a hygroscopic crystal, it must be sealed in an aluminum can to prevent it from absorbing moisture from the air which would ruin its optical properties.

Photo-multiplier Tubes (PMT)

A photo-multiplier tube is a vacuum tube that produces an electrical pulse that has an amplitude proportional to the amount of light that is incident on a thin, semi-transparent glass window. The inside surface of the window has a very thin coating of metal alkali metals, which have low work-functions, allowing an incident photon to eject an electron via the photoelectric effect. This surface is held at a large negative electric potential relative to other elements in the photo-multiplier tubes and the ejected electrons are accelerated away from the photo-cathode and can gain several hundred eV before they impact the first dynode. Dynodes are coated with a material such as beryllium copper oxide, that will emit several low-energy electrons when hit by an incident electron. Several dynode stages with increasing electric potentials allows the charge of the electron initially ejected from the photo-cathode to be multiplied by a factor as large as 10^5 or 10^6, producing an electric pulse at the output that has an amplitude large enough to be easily measured by relatively unsophisticated electronics.

Photo-multiplier tubes require a high voltage power supply to provide the accelerating potentials across the dynodes. Typical operating voltages are of order 1 kV but the circuit used to provide the voltages to the dynodes usually draws less than 1 mA of current. The gain of a photo-multiplier typically varies with the applied voltage according to $G = V^{\beta}$, where β can be as large as 5 or 6. Therefore, even small changes in the operating voltage can result in large changes in the gain. For this reason, precise photon energy measurements need a high voltage power supply that is very stable in time.

Although the output of a photo-multiplier tube is proportional to the amount of incident light, the power supply may not be able to deliver enough current to produce very large pulses from, for example, high energy photons incident on a NaI(Tl) crystal. In addition, the available current may be insufficient even for moderate pulses at very high rates. Thus, it is possible that the gain of a photo-multiplier tube is slightly non-linear, becoming slightly less than expected for large pulses.

Multi-channel Analyzer (MCA)

A multi-channel analyzer (MCA) detects electrical pulses at its input, measures the amplitude (or charge), and stores the resulting measurements in a histogram. The signal from the PMT is connected to the input MCA where it can be amplified by a pre-amplifier with a selectable gain. A discriminator triggers the electronics to measure the amplitude of a pulse when the signal from the pre-amplifier exceeds a specified threshold. This threshold can be set to 1-2% full scale to ignore the large number of very small pulses due to electronic noise in the system. The amplitudes of triggered pulses are measured using a 10-bit analog-to-digital converter (ADC) and are stored in memory as a histogram.

The *dead-time* is the fraction of the time in which the MCA is measuring and analyzing pulses. If the dead-time is significantly larger than a few percent, then the probability that two photons could arrive at the same time becomes significant. This could degrade the energy resolution or bias the energy measurement. Therefore, keep the dead-time low by increasing the discriminator threshold or by reducing the intensity of the beam to which the PMT is exposed.

γ-ray Spectra of Radioactive Isotopes

During this lab you will be measuring γ-ray radiation from several different radioactive isotopes, including ^{137}Cs and ^{22}Na. These isotopes emit radioactivity via *beta decay*. In this process, the weak interaction in the atom can change a proton into a neutron or vice versa. In order to conserve mass and charge, this transformation is accompanied by the emission of a beta particle (electron or positron) and either a neutrino or antineutrino, (e.g., $p \rightarrow n + e^{+} + \bar{\nu}$, or $n \rightarrow p + e^{-} + \nu$). The decay scheme of ^{137}Cs is shown in Figure 2. The isotope has a half-life of 30.23 years, and 94% of the time it decays to an excited energy state of ^{137}Ba, while the reminder of the time it decays directly to the ground state of ^{137}Ba. In both cases it releases a neutrino and an electron, with the latter typically being immediately absorbed by the metal casing of the sample. The reason why ^{137}Cs makes a good γ-ray source for investigating Compton scattering is due to the fact that the ^{137}Ba isotope only lives for ~2.5 minutes on average in the excited energy state before dropping to the ground state, emitting a 661.7 keV γ-ray photon in the process. Thus ^{137}Cs produces a clean, steady γ-ray spectrum, with only a single emission line.

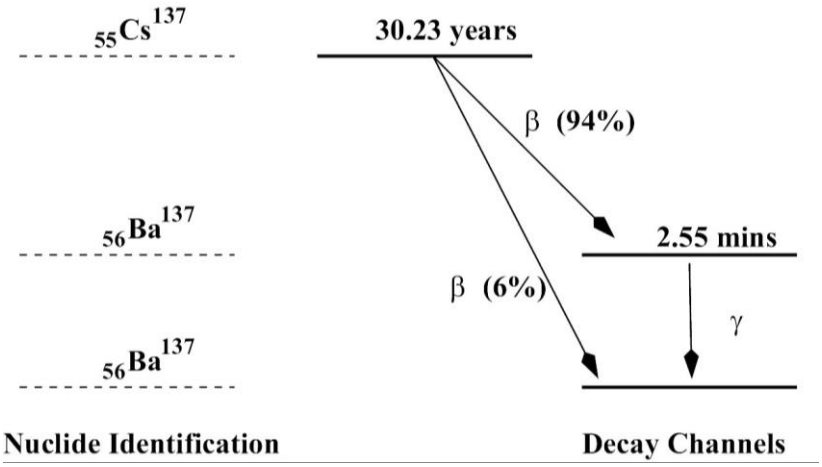


Figure 2: A schematic of the energy decay scheme of ^{137}Cs.

Before you begin your investigation of Compton scattering with ^{137}Cs, you must first calibrate your detector with one of the calibration sources provided by your T.A.

Most of these isotopes decay in a similar fashion to ^{137}Cs, producing one or more ɣ-ray emission lines. One calibration source, ^{22}Na, however, has a slightly more complicated decay scheme (Figure 3). It decays into an excited state of ^{22}Ne, but unlike ^{137}Cs, in this case a positron and antineutrino are emitted. The positron typically lasts only very briefly before combining with an electron in the sample or its metal casing. The two particles annihilate each other and release their rest mass energy ($E = m_e c^2$) in the form of two ɣ-rays, each with energy 0.511 MeV ($e^- + e^+ \rightarrow \gamma + \gamma$). Meanwhile, the ^{22}Ne remains only briefly in its excited state before dropping to the ground state, emitting a 1.274 MeV ɣ-ray in the process. The ɣ-ray emission spectrum of ^{22}Na thus consists of two peaks at 0.511 MeV and 1.274 MeV. In Figure 3 there is also a third peak, caused by instances when two of these different energy photons happen to hit the detector simultaneously.

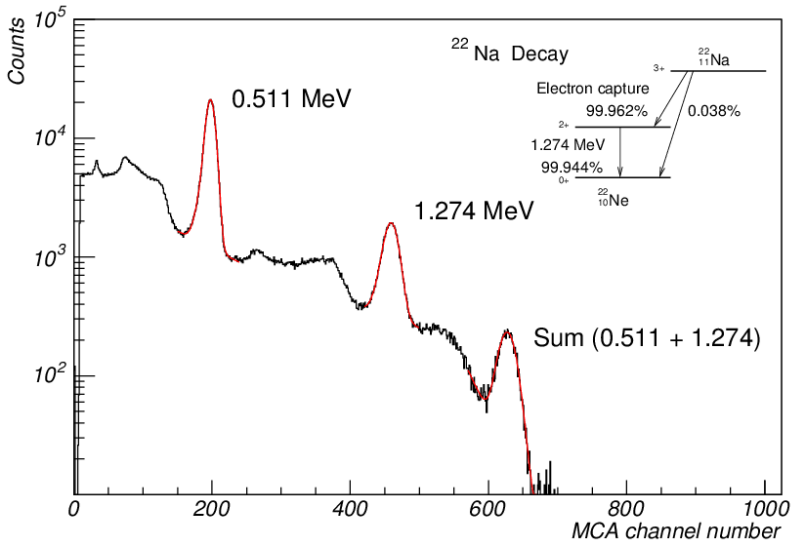


Figure 3: The observed energy spectrum for the decay of ^{22}Na (note the vertical log scale). The inset shows the ^{22}Na energy decay scheme.

The spectrum of Figure 3 is simply a histogram showing the number of counts the PMT system has accumulated in each energy channel. By using the known intrinsic energies of the ɣ-ray emission lines of the calibration isotopes, you will determine the relation between energy (in keV) and channel number for your scintillator+PMT+MCA system. A simple analogy would be if you were given a thermometer with equally spaced markings but no labelling –-- you might choose to calibrate it to measure degrees Celsius by noting its readings when it is first placed in an ice bath (0°C), and then a pot of boiling water (100°C) . A booklet of ɣ-ray spectrum charts for various isotopes is provided for your reference in the lab room, ask your T.A. for help if you cannot locate it.

Radiation Safety – IMPORTANT!

The calibration sources used in this lab have very low radioactivity and do not present an exposure risk. The stronger, ^{137}Cs source is contained in a lead shield which collimates the emitted photons into a narrow beam so that they are directed away from an experimenter. Nevertheless, it is useful to use good source handling practices. In particular,

• Maximize the distance between you and a source whenever possible (use the inverse-square law to your advantage!). For example, use tongs to handle and position the calibration sources.

• Reduce the time exposed to a source of radiation. Do not spend an unnecessary amount of time in front of the ^{137}Cs source beam when positioning the photo-multiplier tube.

• Use shielding to prevent exposure. Keep the lead shield in place on the high activity ^{137}Cs source when you are not taking data with it.

• Avoid unintentional ingestion of radioactive isotopes. Although all the sources used in this lab are sealed and cannot contaminate the surfaces in the lab, it is good laboratory practice to avoid ingestion of any contaminated material. Therefore, do not eat, drink, apply cosmetics, smoke, or chew gum or tobacco in the lab. Also, it is a good idea to wash your hands after working in the laboratory.

Calibration Procedure

1. Dependence of MCA channel number on PMT voltage

   The purpose of this step is to determine the possible effects of voltage fluctuations on the MCA channel output spectrum.

   - Turn on the high voltage using the software 'Adjust-HV' tab and click 'on'.

   - Set the high voltage to approximately 950 volts and the internal pre-amplifier gain to 8.

   - Select a calibration source such as ^{54}Mn, ^{60}Co, or ^{22}Na, which have well-defined peaks in the energy range 1-2 MeV.

   - Verify that the energy resolution is good enough to clearly resolve the peaks. Adjust the high voltage and coarse gain setting if necessary such that a peak in the 1-2 MeV energy range occurs somewhere between channels 400 and 800.

   - Measure the centroid of the peak as the high voltage is varied over the range ± 50 volts.

- Plot the measured channel position as a function of voltage on a log-log plot. Compare the relationship with $C = C_o(V/V_o)^{\beta}$, where C_o is the channel position at a reference voltage V_o.

- Calculate the percentage change in the centroid peak channel position that would result if the power supply output voltage fluctuated by 0.1%.

2. Calibration of MCA channel number

For the Compton scattering part of the lab you will study photons with energies less than 1 MeV. Therefore, the gain should be adjusted so that the photon peak from the ^{137}Cs calibration sample is located near channel 500. This is typically achieved with a high voltage of approximately **950 V**, and the pre-amplifier gain set to **16**.

- Set the high voltage and pre-amplifier gain so that the ^{137}Cs peak lies somewhere between channels 500 and 700.

- Record the spectra of several calibration sources, recording their peak positions estimated using the peak finding analysis provided by the MCA interface software.

- With the high voltage and gain selected above, not all of the peaks will be within the range of channels analyzed by the MCA. Identify which sources have well-defined peaks in the range of channels that you can observe and look up the energies of these peaks in the spectrum charts that are provided.

- You can save a spectrum to a thumb drive using 'Save-as' and selecting the 'TKA' file format. Each line of this file consist of the number of counts in each of the 1024 channels and is suitable for importing into Microsoft Excel.

- Plot a graph of channel number vs energy (in keV) for the set of isotope peaks that you have studied. Fit this curve first to a straight line and then a 2nd order polynomial. Which functional form provides the best fit to the data? Your final choice of fit is your calibration curve, which you will use to convert channel numbers into keV energy units.

Study of Compton Scattering

1. Estimate the angular resolution of the photo-multiplier tube as shown in Figure 4. By assuming an equal probability of detection across the face of the detector, the uncertainty in the angle measurements can be estimated from the

variance of a uniform probability distribution of width Δθ, which is $\sigma^2 = (\Delta\theta)^2/12$.

2. Remove the lead cover from the ^{137}Cs source and position the PMT on the movable carriage at approximately 5°. Locate and measure the position of the ^{137}Cs peak and compare this with your earlier results from the ^{137}Cs calibration source.
3. Place the aluminum scattering rod on the holder, making sure it is aligned directly in front of the beam. Repeat the measurement in step 2, recording sufficient data to obtain an accurate estimate of the peak position. Record the live-time over which data was accumulated, the angle and the measured centroid of the peak.

4. Repeat the previous step for an angle of 10°.

5. Return to the angle used in step 2 and measure the peak position again to determine if it has drifted for any reason.

6. Repeat steps 3 and 4 for angles up to 90°, always returning to the initial position of 5° to monitor changes in the PMT gain. Because the intensity of photons scattered at large angles decreases, you will have to accumulate data for longer periods at larger angles.

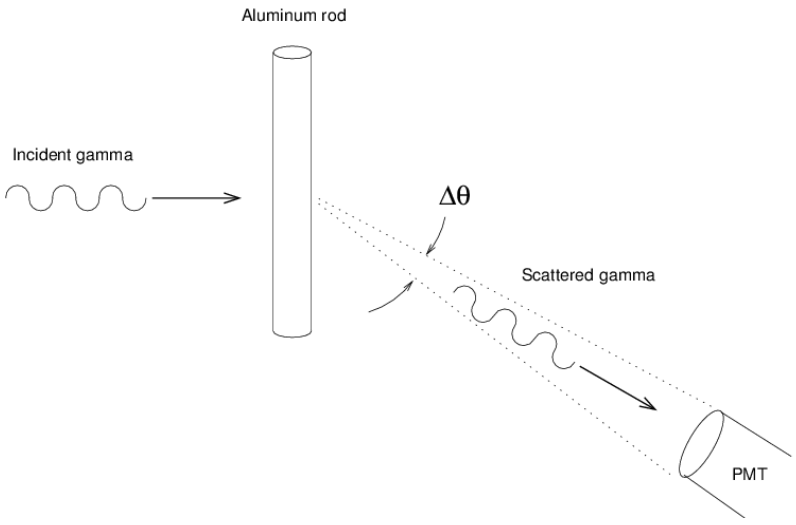


Figure 4: How to estimate the uncertainty in the scattering angle.

Data Analysis

1. Using your fitted curve to the MCA channel number as a function of photon energy, convert the channel numbers of the measured peak positions to energies.

2. Measure the mean and RMS of the distribution of photon energies recorded with the PMT at 5°. The RMS of this distribution can be used to estimate the uncertainty on any given measurement since it reflects the size of any uncontrolled fluctuations in gain.

3. Tabulate θ, peak position, peak energy (E') and their uncertainties for all angles at which data were taken. Include this table in your report. Make sure you explain and show sample calculations for the uncertainties in these quantities.

4. Plot $1/E'$ as a function of $1-\cos\theta$. State what can be concluded from this plot. Be sure to include both horizontal and vertical error bars in your plot.

5. Perform a least-squares fit of a straight line to the data plotted above. From the slope, determine a value of the rest energy of the electron $m_e c^2$. Be sure to show a detailed calculation of the uncertainty in this quantity in your lab book. Discuss how your value compares with the accepted value.

Appendix: Genie 2000 Reference Sheet

2. • *Introductory comments*

The PC contains a Multi-Channel Analyzer (MCA) board that accepts signals from a Photomultiplier Tube (PMT) and processes them in such a way as to make histograms of the counts vs. channel number. A program called Genie 2000 serves as a link between the MCA board and the computer. Genie 2000 program splits the monitor screen in half, with the upper half displaying the **MCA data** and the lower half comprising the **Report Window.**

• To start the computer MCA controller

From the **Start** menu, choose **Programs** and then **Genie 2000**. Select **Gamma Analysis and Acquisition**. When the program has loaded, select **Open Datasource** from the File menu and then click **detector**. In the file listing, choose *Compton*.

• *To set data acquisition time (Live Time)*

Menu Sequence:

MCA→Acquire Setup,
Select **Live Time**,
Input the time,
hit **OK**

• To use the peak locator to find the centroid of a peak

Menu sequence:

Analyze→Peak Locate→Unidentified 2nd Diff
Then, use the following settings:

Start Channel	1
Stop Channel	1024
Significance Threshold	3.00
Tolerance	1.00 keV (Energy)

Then, check *Generate Report* and click *Execute*.

• To Print the Scatterplot

Menu Sequence:

File→Data Plot

• *To save numerical data*

Menu Sequence:

Analyze→Reporting

Then use the following settings:

Start On	New Page
Template Name	340dump.tpl
Output to	Screen
Section Name	All
Activity Units	μCi
Multiplier	1

Then, click the *Execute* button. The counts are printed to the report window in channel order, starting from channel 1. Once the data is in the report window, it must be copied to the Clipboard and pasted into a text editor, where it can be saved.

Menu Sequence:

Options→Report Window→Copy Contents to Clipboard

The Clipboard can then be pasted into Notepad to create a permanent copy. For instance, the Notepad contents can be Saved as a file on the Desktop. Using Excel, the file just created can be imported using the Text Import Wizard. When doing this, it's important to start the import at Row 9. Once the data is in Excel, information about channel number (or γ energy) can be added, subtraction of background can be performed, and plots can be generated in the normal way.

• To clear your data

Menu Sequence:

MCA→Clear→Data,

Or Click the *Clear* button in the upper left part of the screen.

• *To clear the report window*

Menu Sequence:

Options→Report Window→Clear Contents

RR August 2000
SS December 2001

Physics 340 Laboratory
Discrete Electron States in an Atom: The Franck-Hertz Experiment

Objective: To measure the first excitation potential and the ionization potential of mercury atoms and to show that the energies of bound electrons are quantized.

References:

1. J. Franck and G. Hertz, Verhand Deut. Physik Ges. **16** ,10 (1914).
2. A.C. Melissinos, *Experiments in Modern Physics*, Academic Press, New York, 1966, pgs. 8-17.

Apparatus: Franck-Hertz tube, furnace, Kiethley Model 485 picoammeter, Wavetek Model DM2 digital voltmeter (optional), CASSY interface, 10 KΩ potentiometer, a.c. power supply for the furnace (Variac), two 1.5 V batteries, 0-100V d.c. power supply to accelerate electrons, 0-10 V d.c. Lambda power supply for heating the filament, Fluke Model 51 digital thermometer.

Introduction

An electron bound in an atom does not behave like a classical mechanical system, which can absorb arbitrary amounts of energy. Instead, as suggested by Bohr in 1913, an electron in an atom can exist only in definite discrete stationary states, with energies E_o, E_1,... . In this model, atomic excitations are represented by transitions of an electron, bound to the atom, from its ground state energy to a higher level. Excitation to increasingly higher energies is facilitated by energy levels that lie closer together. Eventually, excitation beyond the ionization energy of the atom produces an electron which is no longer associated with the atom. Such an electron enjoys a continuum of available energy states. The essential features of this scheme are represented by an energy-level diagram as shown schematically in Fig. 1.

Horizontal lines in this figure represent "allowed" values (measured along the vertical axis) of the total energy (E_{kin}+E_{pot}) of the most weakly bound electron in the atom. Notice that these discrete values are negative, indicating that these states are "bound" states of the electron; i.e. work has to be done in order to remove the electron from any of these states or "levels". In particular, the lowest lying level $E^{(0)}_{ground}$, called the "ground state", has the largest negative energy. When not excited, the electron and thus the atom stays in the ground state. Removal of an

electron from an atom is called "ionization". Thus, in order to ionize the atom in its ground state, an amount of work equal to $-E^{(0)}_{ground}$ (or larger) has to be supplied to the atom.

"Excitation" of the atom occurs when the electron in its ground state absorbs energy, after which it is promoted to one of the higher bound states $E^{(i)}_{excited}$. Electrons in atoms can be excited in a number of ways, such as bombarding atoms by free electrons or illuminating atoms by light.

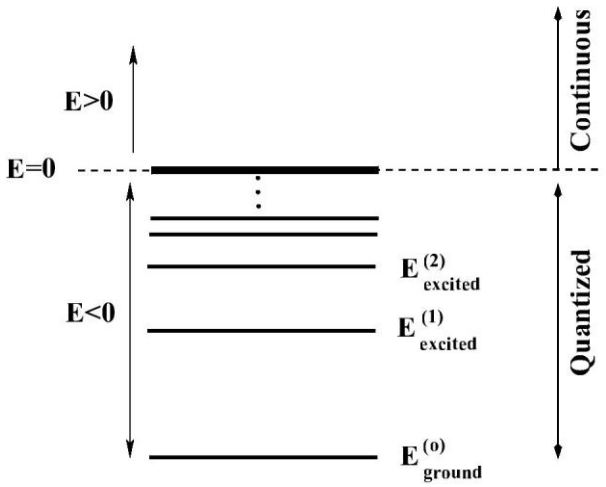


Figure 1: A schematic diagram showing the energy levels of an atom. The heavy solid lines represents the vacuum level and separates the quantized states from the continuum.

If an atom is supplied with energy by excitation from a free electron, then a bound electron can take up energy from the free electron only in quantized amounts ΔE equal to the difference in energy between the excited level and the ground state.

$$\Delta E_i = E_{kin}^{(before)} - E_{kin}^{(after)} = E_{excited}^{(i)} - E_{ground}^{(0)}$$

$E_{kin}^{(before)}$ - kinetic energy of the bombarding free electron before collision,

$E_{kin}^{(after)}$ - kinetic energy of the bombarding free electron after collision,

$E_{excited}^{(i)}$ - i-th excited state of the atom, and

$E_{ground}^{(0)}$ - ground state of the atom.

If an atom is bombarded with electrons whose kinetic energy are less than the first excitation energy of the atom, no exchange of energy between the bombarding electrons and the electrons bound to the atom can take place. (This of course neglects any small amount of energy that may be transferred in *elastic collisions* when the whole atom recoils without being electronically excited.) Thus, the electron in the

atom remains in the ground state $E_{ground}^{(0)}$. If $E_{kin}^{(before)}$ is equal to or greater than $E_{excited}^{(i)}$ - $E_{ground}^{(0)}$, the electron in the atom can be promoted into the first excited state.

If a free electron is accelerated through mercury vapor having an appropriate atom number density, the probability of exciting the ΔE_1 transition is much larger then exciting any other ΔE_i transition. Thus in a sequence of n collisions with n different mercury atoms, the bombarding electron can convert $n\Delta E_1$ energy into atomic excitations. Bohr's quantum ideas were well supported by many studies of electromagnetic radiation from atoms where photons with definite energies were either emitted or absorbed. The historical significance of the Franck-Hertz experiment is that it provided convincing proof that energies of atomic systems are quantized not only in photon emission and absorption but also under particle bombardment.

The energy levels of Hg

In this experiment, you will probe the energy levels of a Hg atom. A neutral mercury atom has 80 electrons. These 80 electrons are distributed in a configuration specified by

$$1s^2,2s^2,2p^6,3s^2,3p^6,3d^{10},4s^2,4p^6,4d^{10},4f^{14},5s^2,5p^6,5d^{10},6s^2.$$

It is convenient to divide these 80 electrons into two broad categories often referred to as inner shell and outer shell electrons. We know that 78 of these electrons reside in inner shells (1s,2s,2p, etc.) and 2 of these electrons reside in the outermost 6s shell.

At low energy bombardment, only one of the two outermost electrons in the 6s shell is promoted to an excited state referred to as a triplet 6 $^3P_1$ state as shown in Fig. 2. The most probable excitation to this triplet state requires a 4.86 eV energy transfer to the bound electron of the mercury atom. The probability of excitation to higher levels of the 6s electrons or the probability of excitations of any inner shell electron is very low and need not concern you in this experiment.

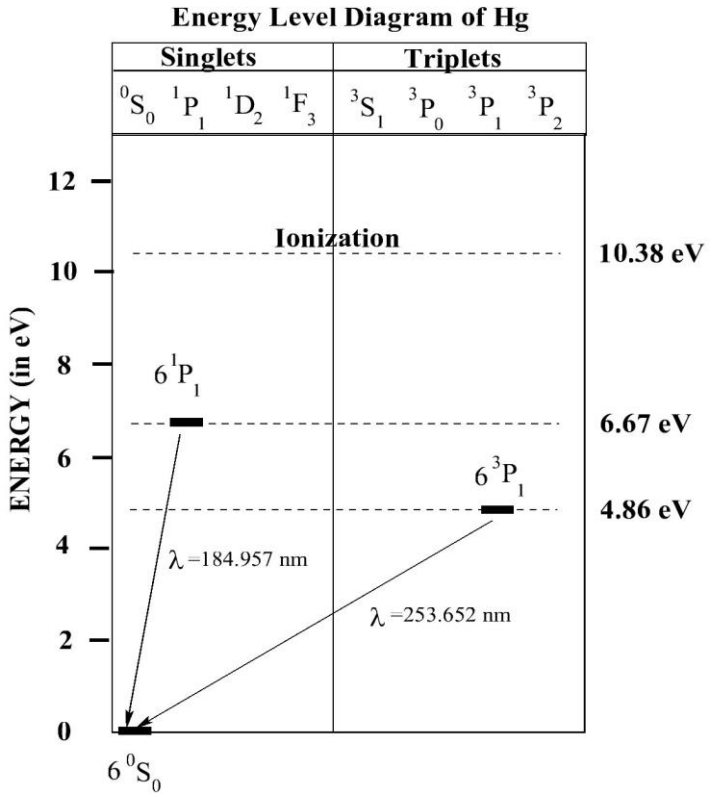


Figure 2: A term diagram showing the lowest lying energy levels for a mercury atom.

In general, the excited states are unstable, and the atom exists in that state only for a short time, typically 1 to 10 nanoseconds. When it returns to the ground state, an amount of energy $=E^{(i)}_{excited}-E^{(0)}_{ground}=\Delta E_i$ is released in the form of electromagnetic radiation. The wavelength of the radiation emitted when the first excited state decays into the ground state (ΔE=4.86 *eV*) is 255 nm.

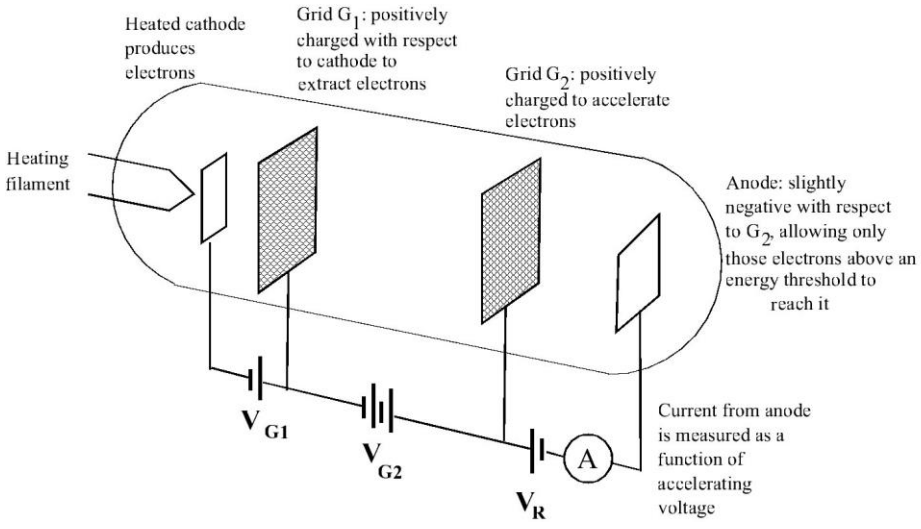


Figure 3: A schematic diagram illustrating the essential features of the Franck-Hertz experiment.

Experimental Considerations

The experiment described below is essentially the same as that performed by J. Franck and G. Hertz in 1914. Suppose a beam of electrons is produced between a heated cathode and an extraction grid G_1 that is biased positively with respect to the cathode. Electrons passing through G_1 are accelerated toward the grid G_2 in a vacuum tube that has been evacuated of air molecules (See Fig. 3) by a positive voltage V_{G2} applied with respect to G_1. At the grid G_2, the kinetic energy of the electron is $K_{kin}=eV_{G2}$, where e is the magnitude of the electron's charge and V_{G2} is the potential difference between G_1 and G_2.[2] Because of this kinetic energy, most electrons pass through the grid and reach the anode after being decelerated by the retarding potential V_R, provided that $|V_R|<|V_{G2}+V_{G1}|$.

Now suppose mercury vapor at low pressure is let into the tube. This can be achieved by inserting a drop of mercury into a heated vacuum tube. As long as $eV_{G2}<\Delta E_1$, the accelerated electrons may undergo many "elastic" collisions with Hg atoms, but they do not lose much energy because the Hg atom is much heavier than the electron. The electrons thus drift to the anode, through G_2, causing an anode current I_a.

However, when $eV_{G2} = \Delta E_1 = E^{(1)}_{excited} - E^{(0)}_{ground}$, it is possible to transfer the free electron's kinetic energy to an internal electronic excitation of the Hg electron. Thus the incident electron is left with zero kinetic energy, i.e. zero velocity. If the collision takes place close to G_2 the electron cannot re-gain enough speed to reach the anode.

To accomplish this, a slight retarding potential V_R must be maintained between the G_2 and the anode (see Fig. 3). Note that V_R is retarding only for negatively charged particles. This prevents electrons that suffer inelastic collisions close to G_2 from reaching the anode. The collective effect of many electrons suffering one such inelastic collision is the appearance of a minimum in the anode current I_a when plotted as a function of the accelerating potential V_{G2}.

As the accelerating potential becomes larger ($V_{G2}>\Delta E_1/e$), an electron can excite an atom at larger distances from G_2. Thus after an inelastic collision, the electron can accelerate towards G_2 and gain again enough kinetic energy to overcome the retarding potential and reach the anode. As a result, the anode current should then increase. Now suppose $V_{G2}=2\Delta E_1/e$. At a point approximately halfway between G_1 and G_2, the electron reaches an energy ΔE_1. This electron may undergo an inelastic collision, midway between cathode and grid, with a Hg atom. After colliding with the

[2] Since the energy of an electron accelerated in electric field is proportional to the potential difference, it is often measured in eV (electron-Volts) units, i. e. an electron accelerated by 1 V potential difference acquires energy 1 eV; 1 eV=1.602×10^{-19} J.

Hg atom, the electron is left with zero kinetic energy in the electric field, halfway between G_1 and G_2. Since the electron is half way between the grids, it will accelerate and gain ΔE_1 energy as it approaches the second grid G_2. In the vicinity of the grid it may again make a second inelastic collision with a second Hg atom, giving up its energy. This electron will be unable to reach the anode. This results in a second dip.

This reasoning makes it understandable that I_a not only decreases at $V_{G2}=\Delta E_1/e$ but in general at $V_{G2}=n\Delta E_1/e$. Between the values $(n\text{-}1)\Delta E_1/e$ and $n\Delta E_1/e$ the current I_a increases monotonically with increasing V_{G2}. The dips that occur at discrete potentials are an indication of the quantized character of the energy loss process. The first dip corresponds to the case when there is one inelastic collision in front of G_2 with one Hg atom, the second when there are two inelastic collisions with two different Hg atoms, one is half way between G_1 and G_2, the second at G_2, etc.

Experimental Technique:

A specially designed vacuum tube shown in Fig. 4 contains a small amount of Hg that is partially evaporated when the tube is heated inside a furnace. By heating the vacuum tube, the vapor pressure of of Hg can be adjusted. Besides the anode, the acceleration grid G_2 and the cathode, the tube has an extractor electrode - grid G_1. This grid helps to removes the charged cloud of electrons that forms in the region of the heated filament and cathode. The space charge cloud impedes electron emission from the cathode and thus reduces the flow of electrons towards the grid G_2.

Figure 4: A photograph of a modern Franck-Hertz tube.

The temperature of the oven is measured using a Fluke Model 51 digital thermometer. This instrument measures the voltage from a thermocouple junction and converts the voltage to a temperature using a calibration table stored inside the instrument. Make sure that the junction is in the proper position i.e. approximately 5 cm down from the lid between the tube and inner copper cylinder. If the thermocouple junction touches the heating element and it is not insulated, it may cause a short circuit.

A photograph of the apparatus is given in Fig. 5

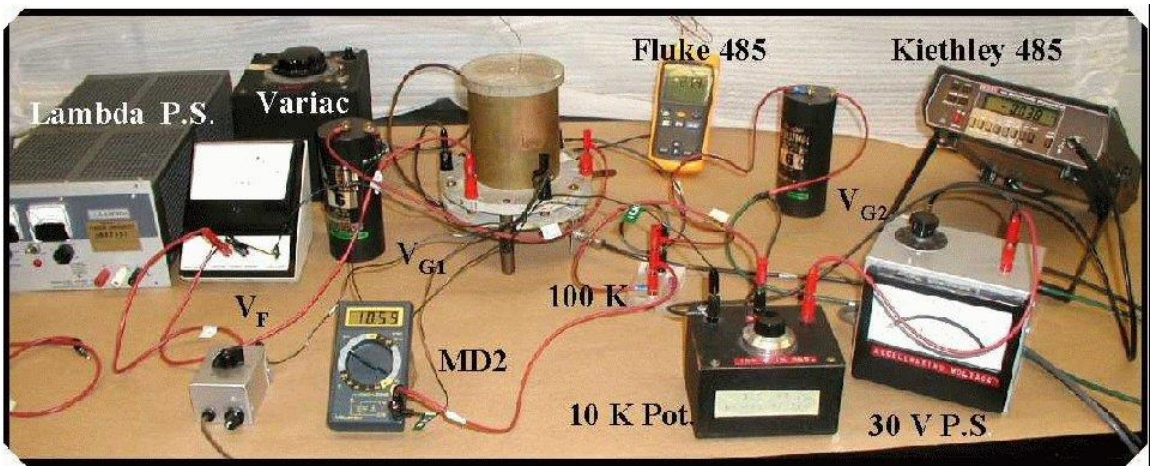


Figure 5: A photograph of the apparatus used to measure the first excitation energy for Hg atoms shown with optional Voltmeter MD2. To record data using computer attach CASSY interface (not shown) as follows: connect UB1 in parallel with (or instead of) voltmeter MD2, and connect analog output of the picoammeter Kiethley 485 to UA1 (voltmeter) input of CASSY interface.

Setting up Cassy Lab program for data acquisition.

Start Cassy Lab program and initialize both voltmeters. Set the display *x*-axis to show UB1 (V_{G2}) and *y*-axis to show UA1 (proportional to current). Set both voltmeters to measure *mean* signals, i. e. to average signals during 100 ms. This will dramatically reduce the noise by suppressing electrical noise at frequencies above 10 Hz. The main source of noise is induced by AC current in power lines (60 Hz). Set the data acquisition period also to 100 ms. Leave total data acquisition time blank. Next, check the *Condition* box in measurement window and type in the following condition:

n=1 or delta(UB1)>0.05

The above setting tells the program to take the next data point (UB1, UA1) *if* the first point is being measured (*n=1*), or when the UB1 value changes by more than 0.05V. The condition is checked every 100 ms, as specified by data acquisition period.

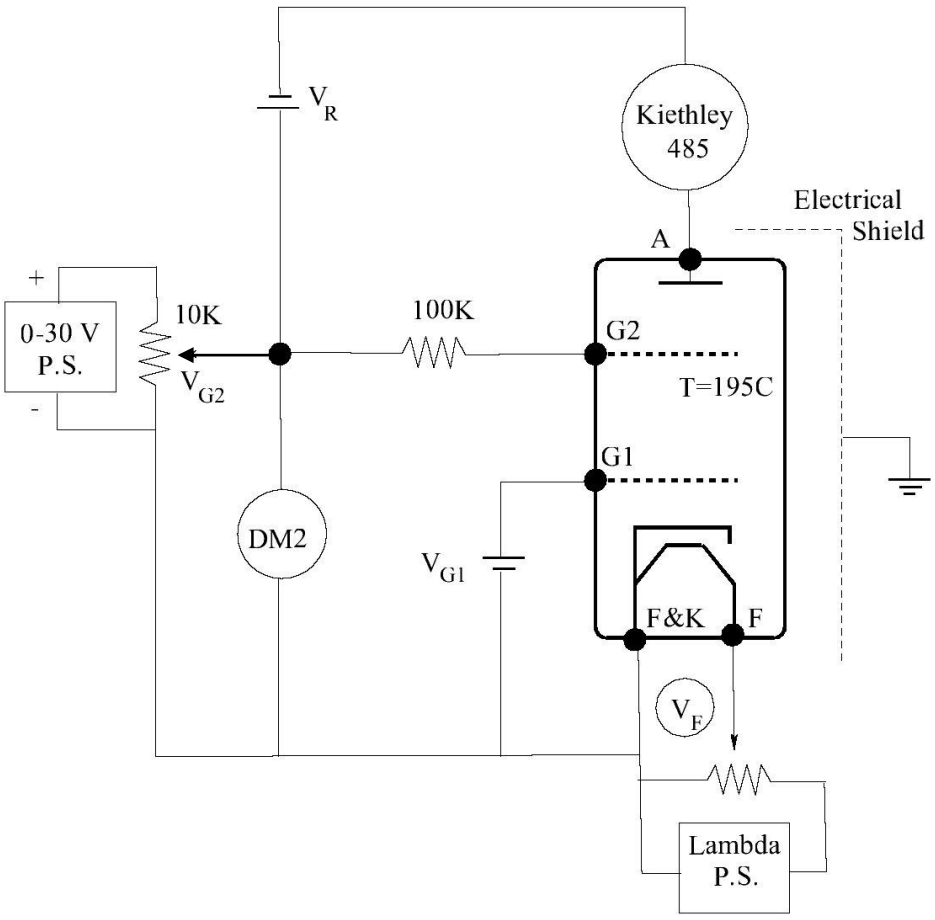


Figure 6: A wiring diagram to measure the first excitation energy for Hg atoms. For computerized data acquizition, CASSY interface UB1 is connected in parallel with (or instead of) DM2, and Kethley 485 picoammeter analog output is connectet to UA1 voltmeter input of the Cassy interface.

MEASUREMENT OF FIRST EXCITATION POTENTIAL OF MERCURY

Set-up Sequence:

(a) Turn on the Kiethley 485 picoammeter and the Lambda power supply. Make sure the Lambda power supply is set to the 7V range.

(b) Turn the oven Variac to 40 V and then switch it on. When the temperature reaches about 170° *C* reduce the setting to what is marked on each Variac by an arrow so that the temperature increases slowly to $195^{o}C$. Under no circumstances let the oven temperature exceed $210^{o}C$. The tube will explode. The final Variac setting and temperature should be such as to allow you to obtain three to five good peaks in the anode current as the grid voltage V_{G2} on G_2 is varied. Always vary the oven power gradually, not by more than 4 divisions on the Variac at a time.

(c) Wire the circuit as shown in Fig. 6. Do not connect the batteries and do not plug in the filament power supply at this time. Set the Wavetek Model DM2 (if used) to the 20 V d.c. scale, and set the UB1 Cassy voltmeter to ±30V range. Set the Kiethley 485 picoammeter to the 2 μA scale. Since the picoammeter is an extremely sensitive instrument, keep it on the 2 μA scale when taking preliminary data. Set the UA1 Cassy voltmeter (which monitors picoammeter output) to ±3V. Set the d.c. power supply which provides the accelerating voltage to 30 V.

(d) Have your wiring checked by the lab instructor. Make sure the initial setting of the 10KΩ potentiometer is such that V_{G2}=0 V.

(e) Connect the two 1.5 V batteries which are used for V_R and V_{G1}. Set V_{G2} to 20V. Raise the filament voltage V_F until the Kiethley 485 reads roughly 0.020 μA (20 nA). *Under no circumstances should filament voltage exceed 7V as it may be permanently damaged!* Take your time because the tube filament changes temperature slowly as the the voltage across it is adjusted. The final value of V_F will be somewhere between 4V and 7V, depending on the age of the tube.

(g) After adjustment of V_F the anode current should remain rather constant when the circuit is wired correctly, the contacts are tight, and the vapor in the tube is in thermal equilibrium. If the current fluctuates wildly, tighten the wire contacts on the banana plugs and panel-jacks and wait for thermal equilibrium. See that the heater shield wire is connected to common ground.

Data Acquisition:

(a) When the current has stabilized, increase the acceleration voltage V_{G2} to 30 V. Select the highest sensitivity range on picoammeter in which there is no signal overload. Check that the Cassy UA1 voltmeter range is suitable for measuring picoammeter output voltage. Record the conversion factor between UA1 reading and actual current, you will need it later to convert the data accumulated by Cassy Lab program to current.

(b) Decrease the acceleration potential V_{G2} to zero.

(c) Start data acqusition program. Immediately the first data point should appear (something like (0.001, 0.002). If more points appear then stop the data acquisition and ask your instructor to check data acquisition settings.
(d) Increase V_{G2} up to 30V at a moderate pace (~1 V/s). As V_{G2} rises by >0.05V, the next data point will be measured by the program automatically. You should be able to see the graph I_a versus V_{G2} as it is being measured. Stop data acquisition as you reach 30V. Rescale the Cassy graph if necessary for better view. You must be able to observe at least 2 oscillations in current. If you do not see them consult your instructor. If everything looks right proceed to the next step.
(e) Discard the data you have accumulated, set the V_{G2} to 0 again, start data acquisition and repeat the measurement. But this time increase V_{G2} *slowly* to 30V. See that not more than few points are measured every second (i.e. spend ~5 minutes for the whole range), as the condition is checked only 10 times a second. The anode current (I_a) as function of V_{G2} will be recorded automatically. Save your data.

Data analysis.
(a) Make a preliminary print of your data and insert it into your notebook. Don't forget to record all the conditions alongside the graph: oven temperature, V_{G1}, and the filament voltage V_F.
(b) Based on this graph, roughly estimate the first excitation potential of the mercury (ΔE_1) by measuring the distance (in Volts) between the local maxima preceding the first and second local minima in the anode current. Also, estimate the standard deviation in measuring ΔE_1 from your measurements. Check whether the accepted value for ΔE_1=4.86 eV falls within the range that you found experimentally.
(c) From the same graph, determine the contact potential $V_{contact}$ between the cathode and G_2 by subtracting the value of the first excitation potential ΔE_1 from the value of V_{G2} at the first local maximum that you measured. Estimate the error involved. The contact potential is caused by the difference in work functions between the materials used to fabricate the cathode and G_2.

Note: for final report you must convert UA1 voltage into actual current. Don't forget to record conversion factor in your notebook!

Repeat measurement for lower anode current:
(a) Set the V_{G2} to 20V and reduce the filament voltage until the Kiethley 485 reads roughly 0.005 μA (5 nA). Repeat the measurement of $I_a(V_{G2})$ as described above.
(b) Perform preliminary analysis of your data. Do the values for ΔE_1 and $V_{contact}$ depend on anode current (filament voltage)?

MEASUREMENT OF THE IONIZATION POTENTIAL OF MERCURY

If a free electron is given sufficient energy, it is possible to excite an electron from the ground state of a Hg atom into the continuum, thus producing a positive Hg ion. Since a Hg ion is quite a bit heavier than the bombarding electrons, it's velocity will be small and it will have a shorter mean free path between collisions than the bombarding electrons. In order to detect these ions, you must therefore reduce the vapor pressure of the mercury inside the tube by lowering its temperature. In this experiment, you will measure the onset of the ion current produced by Hg ions formed during the collision of electrons with mercury atoms. To measure this quantity, the anode is made slightly negative with respect to the cathode. Thus, the electrons cannot reach the anode, which becomes a positive ion collector (See Fig. 7).

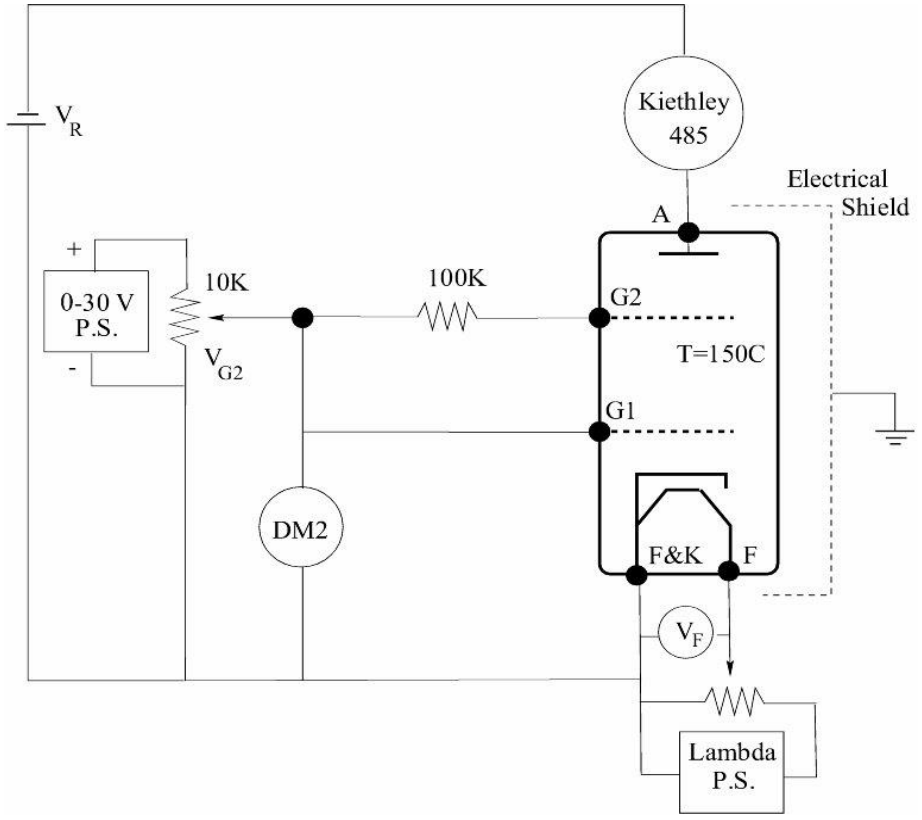


Figure 7: A wiring diagram for measuring the ionization energy of Hg atoms. Connection to CASSY interface is the same as in the previous experiment (not shown).

Set-up Sequence:

(a) Lower the Variac setting to obtain a stable oven temperature of 150^o C.

(b) Remove the battery V_{G1}.

(c) Reconnect V_R as is shown in Fig. 7.

(d) Change the measuring *condition* in the Cassy Lab program to:

```
n=1 or delta(UB1)>0.02
```

(accumulate data points every 0.02V).

Data Acquisition:

Set V_{G2} to zero and set picoammeter sensitivity to 2 nA. Record the ion current I_i versus V_{G2} as described in the first part. Increase V_{G2} slowly while observing I_i value. At ~12V I_i will suddenly rise very rapidly. Stop measurement when I_i reaches ~1.5 nA. Switch the picoammeter to 20 nA and repeat the measurement, this time let the current rise to 15 nA. Don't forget to record conversion factors for converting Cassy voltmeter values to actual current.

Data Analysis:

(a) You have measured the ion current I_i as a function of the accelerating potential V_{G2}. Plot these measurements on a I_i vs. V_{G2} graph. You do not need to show error bars since there are ~500 points in your graph, the error can be estimated visually by the scattering in data points.

(b) In most cases, the current I_i remains zero for small V_{G2} then suddenly increases dramatically. Devise a procedure to determine your best estimate for the value of V_{G2} at which I_i starts rising. From this value subtract the value of the contact potential, you obtained in part 1 for this experiment. The result is the ionization potential of mercury. Estimate the error in your result and then compare it with the accepted value of 10.38 eV. Sometimes, I_i as a function of V_{G2} behaves somewhat differently. At a value of V_{G2} smaller than the ionization potential you will measure a small increase in the current. When you increase V_{G2} further, the value of I_i will remain fairly constant until suddenly the current will start rising very fast. Develop a procedure to determine the value of V_{G2} for which the dramatic increase in I_i sets in. You obtain the ionization potential by subtracting the contact potential from this value. A good explanation of this phenomenon is given in the book by A.C. Melissinos.

(c) Discuss your errors. How important is the temperature of the tube? How sensitive are your measurements to V_F? Do you understand the purpose of the 100 KΩ resistor in the circuits you have constructed? Do you understand why you have used a shielded BNC cable when measuring the anode current? Did the current reverse polarity when measuring I_a and I_i. Why?

RR 9/00
SS 12/01

Physics 340 Laboratory
The Electronic Structure of Solids: Electrical Resistance as a Function of Temperature

Objective: To measure the temperature dependence of the electrical resistance of a metal and semiconductor and to interpret the observed behavior in terms of the underlying band structure of the solids.

Apparatus: Electrical furnace, NiCr-Ni thermocouple, variac power supply for furnace, CASSY power/interface, current module (524-031), thermocouple module (524-045), computer, Pt resistor, semiconductor resistor.

References:

1. W. Pauli, Z. Physik **31**, 373 (1925).
2. E. Fermi, Z. Physik **36**, 902 (1926).
3. P.A.M. Dirac, Proc. Roy. Soc. London A **115**, 483 (1926).
4. E. Wigner and F. Seitz, Phys Rev. **43**, 804 (1933) and Phys. Rev. **46**, 509 (1934).
5. N.F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys*, Oxford University Press, Oxford, 1936.
6. D. Halliday, R. Resnick and J. Walker, *Fundamentals of Physics; 5th Edition*, Wiley and Sons, New York, 1997; Part 5, pgs. 1053-69.
7. K. Krane, *Modern Physics*, 2nd Ed., Wiley and Sons, New York, pgs. 309-29 and pgs. 344-62.

Introduction

An understanding of how much current flows through a conductor for a given applied voltage resulted from Georg Ohm's thorough work in 1827. The empirical relationship known as Ohm's Law has remained valid over the years and is still widely used today. Although Ohm's work focussed primarily on metals, studies by Seebeck in 1821 and by Faraday in 1833 reported anomalies in current flow through a class of materials we now know as semiconductors. Interestingly, the temperature dependence of current flow measured by Faraday in semiconductors was quite different than the temperature dependence of current flow in metals first reported by Davy in 1820. The fundamental origin of this difference remained unexplained for about a century until the development of quantum mechanics.

Following the successful quantum theory of electronic states in isolated atoms, attention turned toward a better understanding of electronic states in molecules and solids. Only with the completion of this effort was it possible to understand the implications of the simple observations about the temperature dependence of current flow made in the early 1800s.

It is now well established that any property of a solid, including its electrical resistance, is in some way controlled by the electronic states of that solid. As a way

of introducing the important differences between the electronic structure of metals and semiconductors, you will measure the temperature dependence of the electrical resistance of samples made from these two important classes of materials. Before beginning these measurements, it is useful (without paying undue attention to many of the details) to review i) the modifications to electron states as we move from the atomic to the molecular to the solid state and ii) a simple physical model for current flow in solids.

Theoretical Considerations

A. Electronic Structure

The important features of an isolated atom are a nucleus surrounded by a complement of electrons that are associated with it in a specifically defined manner. The Pauli exclusion principle requires these electrons to be non-uniformly distributed around the nucleus in regions of space, forming 'shells' of charge known as atomic orbitals. The total negative charge of the electrons exactly balances the total positive charge contained in the nucleus. Most importantly, the electrons, because they are confined to a limited region of space, acquire quantized energy levels.

As atoms are brought together to form a molecule, the outermost electrons from one atom will interact with the outermost electrons of a neighboring atom. This interaction is subtle and a variety of theories have been devised to explain it accurately. The end result is a profound modification to the allowed energies and spatial arrangement of the electronic states.

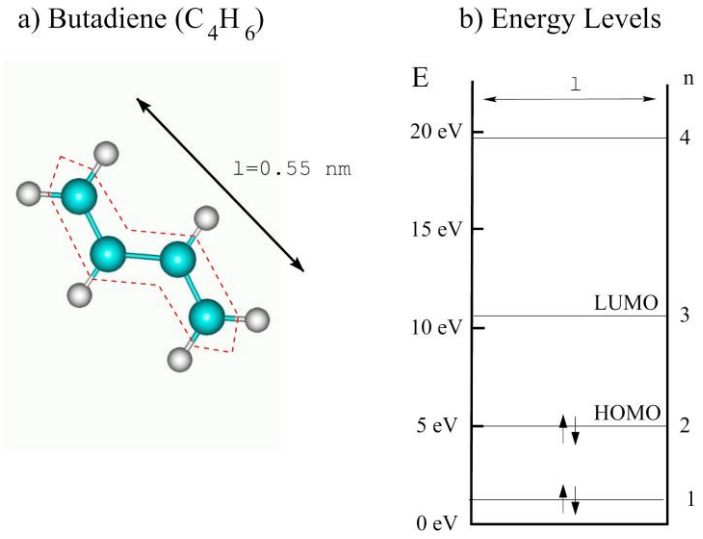


Figure 1: In a), a schematic diagram of a butadiene molecule C_4H_6. The bonding electrons are indicated by the heavy sticks between atoms. The delocalized electrons, which extend both above and below the plane of the diagram, are schematically indicated by the dotted path along the length l of the butadiene molecule. In b), the allowed energies for the electrons in the molecule assuming l is 0.55 nm. The two lowest states are filled. Higher vacant energy states (n=3, 4, 5 . .) are available for occupation. The HOMO (highest occupied molecular orbital) and the LUMO (lowest unoccupied molecular orbital) are also labeled.

To understand the nature of these modifications, it is useful to briefly consider a simple molecule like butadiene (C_4H_6). This molecule is a coplanar arrangement of 4

carbon atoms combined with six hydrogen atoms. Each carbon atom contributes 4 electrons; each hydrogen atom contributes 1 electron. During the synthesis of this molecule, interactions between electrons cause a significant rearrangement of negative charge. Many of the electrons become *localized* in regions of space that lie between two atoms, forming states known as σ bonds. These states are covalent in nature and are fully occupied, containing a charge equivalent of two electrons. The negative charge carried by these σ bonds effectively screens the electrostatic repulsion that is present between the atomic nuclei.

Each carbon atom brings one more electron than required to form the 9 σ bonds in butadiene. These extra electrons assume the lowest energy configuration possible which results in a *delocalized* occupied state referred to as a π orbital in the molecule. A schematic picture of these two different electron states is given in Fig. 1(a). As will become clear below, because of the delocalized nature of these π states, one might conclude that the butadiene molecule forms an extremely simple example of a tiny one-dimensional metal. If one could somehow connect clip leads to either end and apply a potential across it, one might expect current to flow through a single butadiene molecule in much the same way as it does through a copper wire!

As suggested in Fig. 1(a), the π electrons are confined to an extended region of space of length l by the attraction of the positively charged nuclei. Within this region of space, the π electrons are free to wander about. Whether this space has a zig-zag nature or is perfectly straight is not of much consequence here. The important point is that whenever electrons are delocalized over a finite region of space, they take on quantized energy values. The allowed energy levels E_n can be estimated using the well known particle-in-a-box result

$$E_n = n^2 \frac{h^2}{8ml^2} \qquad (1)$$

where n is an integer quantum number, h is Planck's constant and m is the electron's mass.

Let us now find the number of electrons which occupy the quantized π electron states. If each carbon atom brings 4 electrons and each hydrogen atom brings one, then the butadiene molecule has a total of 22 electrons. Of these 22 electrons, 18 are tied up in forming the σ bonds. Thus there must be four electrons occupying the π electron system. Furthermore, the Pauli exclusion principle allows only two electrons at each possible energy. A schematic of the resulting energy states and their occupation can be sketched as shown in Fig. 1(b).

This simple discussion has a number of similarities with the more complicated situation when $\sim 10^{23}$ atoms per cm^3 are brought together to form a solid. These similarities include:

- the existence of a pool of delocalized electrons,
- the existence of quantized energy states,
- the occupation of a certain fraction of these quantized states, and
- the presence of an energy gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO).

All of these basic principles are important when qualitatively discussing the electron states of a three-dimensional solid. To simplify the discussion, it is convenient to partition the solid into 'atomic-cells' known as Wigner-Seitz cells. Each atom in the solid will be surrounded by a Wigner-Seitz cell which takes on an interesting geometrical shape dictated by the exact arrangement of atoms in the solid as shown in Fig. 2.

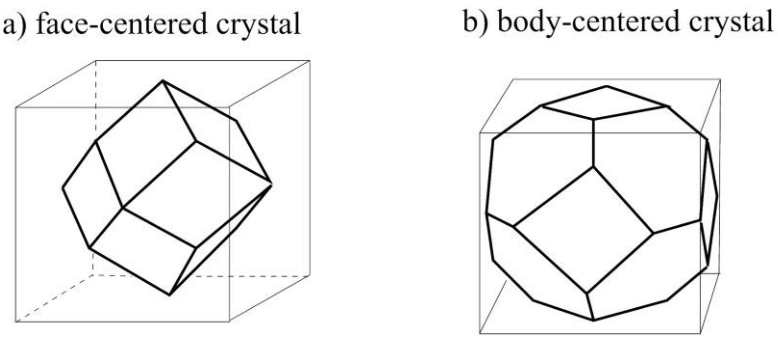


Figure 2: Wigner-Seitz cells for a) a face-centered cubic crystal structure and b) a body-centered cubic crystal structure.

The degree of interaction between all the electrons in the solid is now enormously complicated and depends not only on the shape, range and density of the relevant atomic orbitals but also the exact geometric arrangement of the atoms forming the solid. If it turns out that for a particular atomic shell configuration, certain electrons are localized to a region of space near the center of a Wigner-Seitz cell, then little modification to these electron states will result. These states will be dominated by the nucleus and will strongly resemble isolated atomic states. If, on the other hand, certain electrons become concentrated outside the nucleus, near the boundaries of a Wigner-Seitz cell, then these electron states will be governed by new boundary conditions and their allowed energy levels will change accordingly.

The ability to predict the modifications to different atomic orbitals when atoms are condensed into a solid is now well established, thanks to extensive work spanning a fifty year period beginning in the 1930's. The results of these studies indicate three predominant orbital types (s/p, d and f) that exhibit different behavior as more and more atoms are brought together. The trends exhibited by these three different orbital types are shown schematically in Fig. 3.

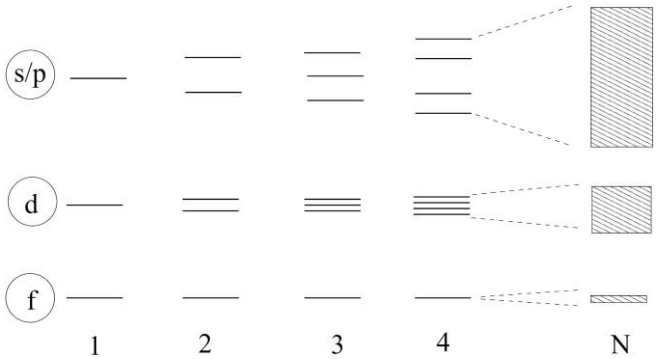


Figure 3: A schematic to illustrate the evolution of energy states in progressing

from 1, 2, 3, . . . to N atoms. When N is large, the separation between electron states is so small that a continuous band of energies is formed. The s/p, d, and f classification scheme is not a rigid one, but is useful for descriptive purposes.

For electrons having atomic s/p and and d character, there are appreciable interactions between electrons located in the edges of the Wigner-Seitz cells, causing the energy of each and every atomic orbital to shift slightly from its well defined atomic value. The resulting perturbed states are separated from each other by a very small energy difference (on the order of 10^{-8} eV). These perturbed states are said to form a continuous 'band' of energies between well-defined upper and a lower limits. These energy bands profoundly control the electronic properties of all solids. The width of each band is largely determined by the degree of interaction between the atomic states that populate them. Strong interactions result in wide s/p bands; weak interactions produce narrow d or f bands. The energy gaps between the bands are a reflection of the separation in energy between the discrete atomic states of an isolated atom. The population of each band is determined by the number of excess electrons left over after the bonding of each atom, one to the other, has been accomplished.

A consequence of this picture is that a simple phenomenon, like the passage of current through a material, will depend on what electronic states are available to carry the current. In turn, the available states are determined by whether a band is partially or completely filled. In addition, the statistical nature of exciting an electron from a filled to a vacant energy level must be properly taken into account. Surprisingly, the ability of a specific solid to carry current as a function of applied voltage and temperature is determined by all of these factors discussed above.

B. Ohm's Law

Electron conduction in a solid is governed by the empirical result discovered by Ohm in 1827 which states that the current density J is related to the applied electric field E by the relation

$$J=\sigma E \quad . \qquad (2)$$

The proportionality constant σ is known as the electrical conductivity of the solid through which the current flows. (To simplify the discussion, we neglect the inherent vector nature of J and E and the subsequent tensor nature of σ.) Ohms Law is often stated in terms of an applied voltage V and the resulting current I as

$$V = \frac{\rho L}{A} I = RI \qquad (3)$$

where A is the cross-sectional area (assumed to be uniform over the the length L of the solid) and $\rho=\sigma^{-1}$ is known as the resistivity of the material. The factor $\frac{\rho L}{A}$ is identified as the resistance R of the material under study.

C. Toward a Microscopic Theory

The first question confronting anyone trying to construct a microscopic model for current flow is 'How do you treat the electrons? ' Are they particles or waves? Many models have been developed which answer this question in a variety of different ways. The most complete models are quantum mechanical and treat the electron as a wave. The more intuitive models treat the electron as a particle. In what follows, we adopt this latter approach.

At a microscopic level, current is ultimately related to the directed motion of charge carriers (electrons) having a charge q. An important question is the net number of the charge carriers crossing a fiducial plane per unit time. This question can be answered from rather elementary considerations.

Electrons have a velocity (~10^5 m/s) which is related to their energy in the solid. However, these velocities cause no net displacement of the electrons in a material since on average, there are as many electrons traveling in any one direction as in the opposite direction. Thus, the velocity related to the electron's energy is not effective in producing a net current flow. The situation changes when an electric field E is applied. Now electrons are accelerated by the electric field and each electron acquires an additional component of velocity, v_d - the so-called drift velocity (~5×10^{-3} m/s in a field of 1 V/m), due to the applied electric field.

The current density can be written as the product of the number of charge carriers per unit volume n and the mean drift velocity v_d imposed by the applied electric field:

$$J=nqv_d \ . \qquad (4)$$

Comparing Eqs. 1 and 2 gives the result that

$$\sigma = \frac{nqv_d}{E} . \qquad (5)$$

Treating the electron's as independent particles, the equation of motion for a charge carrier of mass m in an electric field is given by

$$m\frac{dv_d}{dt} = qE \ . \qquad (6)$$

Since there are many carriers participating in current flow, it is reasonable to expect that each charge carrier will experience many collisions as it travels through a solid. For this reason it makes sense to statistically define a mean time τ between collisions. Alternatively, you can define a drift mean free path $l_d=v_d\tau$ which is a measure of how far the charge carrier will drift between collisions.

With this definition, an estimate for the mean drift velocity of a charge carrier is given by

$$v_d = \frac{dv_d}{dt}\tau = \frac{qE}{m}\tau . \qquad (7)$$

One finally has

$$\sigma = \frac{nq^2\tau}{m} = \frac{nq^2 l_d}{mv_d} \tag{8}$$

or, equivalently,

$$\rho = \frac{mv_d}{nq^2 l_d}\,. \tag{9}$$

From this expression for ρ, the resistance R of a material can be calculated if the geometry of the sample is known (see Eq. 3).

The task at hand is to develop a model for the temperature dependence of the resistance of a material. This can be accomplished by considering the temperature dependence of each non-constant term in Eq. 9.

D. Temperature Dependence of Electrical Resistance for a Metal

To evaluate the various factors appearing in Eq. 9 for a metal, we must have a good model to calculate the relevant quantities. This is difficult when you must take into account ~10^{23} electrons per cubic centimeter. Under these circumstances, the best way to proceed is to use statistics.

Since electrons are fermions, Fermi-Dirac statistics must be used (see the Appendix). The mean number of electrons in any state with energy E is given by $2\times f(E)$ where $f(E)$ is the Fermi-Dirac distribution function and the factor of 2 is due to the two available values for spin of an electron. As the energy of the electrons increases above the bottom of a band, the number of available states increases. Each state can hold two electrons. These states are filled until all electrons in the band are used. A rather abrupt transition from filled to unfilled states then takes place at an energy called the Fermi energy. A consequence of Fermi-Dirac statistics is that at 0 K, all states less than the Fermi energy E_F are filled and all states above E_F are empty.

As the temperature is raised above 0 K, electrons just below E_F can be thermally excited to unfilled states just above E_F. States affected by this transition are located roughly within a $\pm 2kT$ range about E_F.

An important issue is the location of the Fermi energy with respect to the energy bands of a metal. What we know by counting available states is that for most metals, the Fermi energy is located somewhere inside a band. Furthermore, typical values of E_F are much greater than kT for temperatures easily attainable in a laboratory. This implies that only a small number of unfilled electron states within $\pm 2kT$ of E_F are readily accessible by thermal excitation. This important insight is indicated on the schematic diagram in Fig. 4. Under these circumstances, you can show that n is

essentially independent of temperature. Furthermore, v_d is essentially independent of temperature and very nearly the same for all electrons within $\pm 2kT$ of E_F. It follows that the temperature dependence of ρ is determined by the temperature dependence of the mean free path l_d.

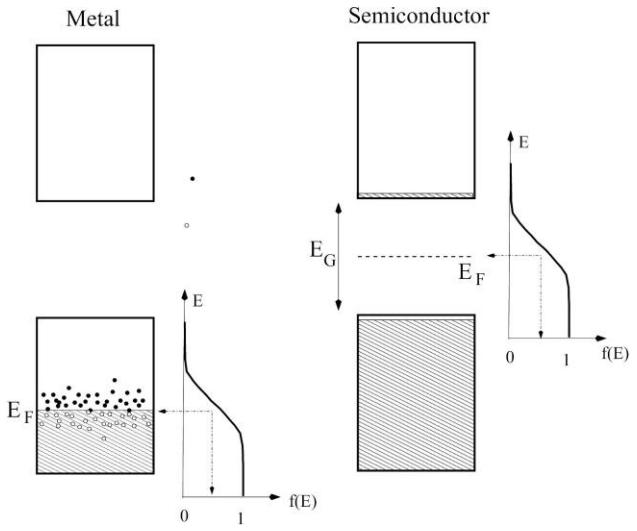


Figure 4: The location of the Fermi energy in a metal and semiconductor. The Fermi-Dirac distribution function at finite temperatures is also indicated. As the temperature increases more electrons occupy unfilled states above the Fermi energy. As suggested in the diagram, the main difference between a metal and a semiconductor is the location of the nearest unfilled states.

For temperatures near room temperature, it is reasonable to expect that l_d will be determined by scattering from atoms undergoing thermal motion. A simple model predicts that l_d is related to the cross-sectional area A occupied by atoms vibrating in the solid due to thermal motion. An estimate of this area can be obtained by assuming that an atom undergoes a rapid random vibration from its rest position by some amount r. It follows that

$$A \cong \pi r^2 . \qquad (10)$$

Treating the vibrating atom as an harmonic oscillator, the average potential energy of such an oscillator is proportional to the square of its displacement, r^2. From the equipartition theorem, the average potential energy is also known to equal $kT/2$. It follows that $A \propto r^2 \propto T$. Since the mean free path will decrease as A increases, we might expect that $l_d \sim 1/A \sim T^{-1}$. Thus in a metal, you might anticipate that the temperature dependence of the resistivity ρ will be given by

$$\rho(T) \propto T . \qquad (11)$$

E. Temperature Dependence of Electrical Resistance for a Semiconductor

For a semiconductor, the situation is a bit more complicated. Semiconductors are characterized by filled energy bands. The nearest unfilled states to carry current are separated from the filled states by an energy gap E_G (see Fig. 4). If $E_G>>kT$, then n, l_d, and v_d all become temperature dependent.

Under these circumstances, Fermi-Dirac statistics in principle still apply except the Fermi energy is now located at a distance $E_G/2$ above the edge of the filled band (often called the valence band). This situation is shown in Fig 4(b). For energy gaps such that $E_G>>kT$, the number of electrons in the unfilled states is quite small and the Fermi-Dirac distribution is well approximated by the Maxwell-Boltzmann distribution function as discussed in the Appendix.

Using Maxwell-Boltzmann statistics, the carrier concentration follows a thermal excitation model since no unoccupied states are available unless an electron is excited across the energy gap. Within this model, the probability of an electron being excited by an energy W is given by

$$n=n_o e^{-W/kT} \qquad (12)$$

where W, the activation energy, turns out to equal $E_G/2$.

Because the nearest unfilled states that carry current are now far up the tail of the Fermi-Dirac distribution curve, a Maxwell-Boltzmann analysis of the most likely drift velocity is required. It is well known that this analysis gives a most probable velocity that scales as $T^{1/2}$. Equivalently, for a dilute gas of electrons, the drift velocity v_d can be estimated from an equipartition of energy argument which also gives v_d proportional to $T^{1/2}$. The mean free path may again be taken proportional to T^{-1}, following the same arguments given above.

Putting all this together gives the temperature dependence of ρ for a semiconductor as

$$\rho(T) \sim T^{3/2} e^{E_G/2kT} \quad . \qquad (13)$$

Note that for small variations in T, the temperature dependence of the exponential term is usually the dominant factor.

It should be mentioned that this discussion is approximately valid for a pure semiconductor (often called an intrinsic semiconductor). If the semiconductor contains significant impurities, then additional energy levels are introduced and the above discussion must be modified to take them into account.

F. Summary

An understanding of the temperature dependence of the resistance of metals and semiconductors requires an appreciation of the energy states in a solid. In addition, a transport model for current flow must be in place. In the discussion above, a number of simple approximations have been used in order to give a framework for estimating the temperature dependence of the resistivity. The striking difference expected for

the temperature dependence of the resistance between metals and semiconductors forms the motivation for the measurements that will be performed.

Experimental Considerations

There are a number of techniques to measure electrical resistance R of an unknown resistor and each technique has its advantages and disadvantages. The two most common methods employed in accurate measurements of *R* are known as the 2-wire and 4-wire techniques as illustrated in Fig. 5. Although it may seem trivial, it is important to understand the differences between these two techniques.

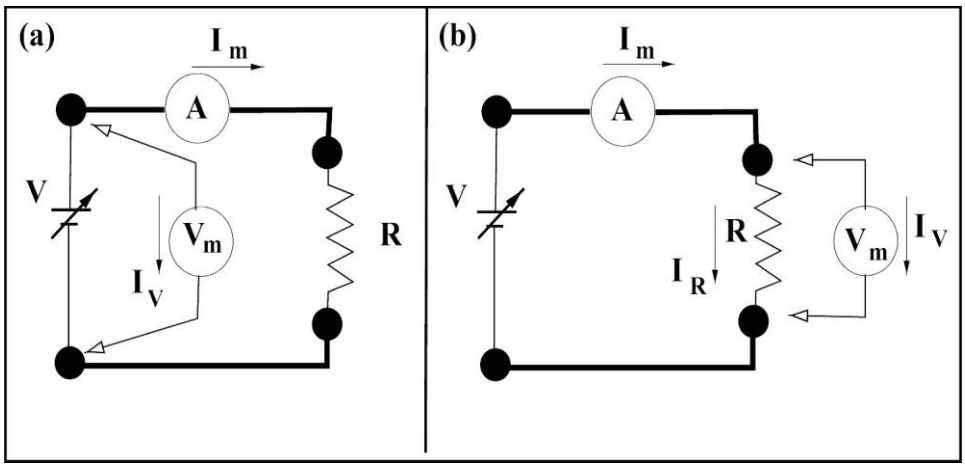


Figure 5: (a) A 2-wire measurement of resistance (note only TWO wires are connected to the resistor) showing the relative position of the ammeter and voltmeter with respect to the resistor R. The voltmeter reads a voltage V_m and the ammeter reads a current I_m. (b) A 4-wire measurement of resistance. Note that FOUR wires are connected to the resistor. Once again, the voltmeter reads a voltage V_m and the ammeter reads a current I_m. In both cases, the heavy lines represent the lead wires from the voltage source to the resistance being measured.

In the 2-wire method (see Fig. 5(a)), a voltmeter is connected across a voltage source and the current flowing to a resistor R is measured with an ammeter. If the voltmeter measures a voltage V_m and the ammeter measures a current I_m, then the measured resistance R_m is given by V_m/I_m. Ideally, R_m should equal R, but this may not be the case. The difficulty is that the ammeter has an internal resistance r_a and the wires between the voltage source and the resistor have a resistance r_w. Thus, V_m includes not only V_R, the voltage dropped across the resistor R, but also v_a, the voltage dropped across the ammeter, and v_w, the voltage dropped across the wires in the circuit. So, $V_m=V_R+v_a+v_w$, and the measured resistance will not be R but

$$R_m = \frac{V_m}{I_m} = R + (r_a + r_w) \ . \qquad (14)$$

From this we see that the resistance measured using a 2-wire method inevitably includes the resistance of the ammeter and lead wires. If the sum of these two resistances is comparable to R then serious inaccuracies will result. The common ohmmeter uses this 2-wire technique and gives an accurate measure of R only if R is large compared to r_w.

In a 4-wire scheme, the voltmeter is connected directly across the resistor R and now the reading on the voltmeter, V_m, **is** the voltage drop across the resistor. The ammeter measures I_m, the total current flowing through the circuit. However, because the voltmeter is in parallel with the resistor R, the measured current I_m splits (see Fig. 5(b)) into two parts: I_R, the current through the resistor, and I_V, the current through the voltmeter. Clearly, $I_m=I_R+I_V$. Since the voltmeter is connected in parallel across the resistor, the voltage drops across these two objects must be the same. This gives $I_V R_V=I_R R$, where R_V is the internal resistance of the voltmeter.

Again, the measured resistance R_m is given by V_m/I_m. This implies that when using a 4-wire technique, the measured resistance is not simply R but

$$R_m = \frac{V_m}{I_m} = \frac{V_m}{I_R + I_V} = \frac{R}{1 + R/R_V} \approx R(1 - \frac{R}{R_V} + ...) \qquad (15)$$

As long as R_V is considerably greater than R, the difference between R_m and R is small. Note that the 4-wire technique removes any inaccuracies introduced by the resistance of lead wires and ammeters (see Eq. 14).

From this discussion it should be clear that difficulties can arise when ammeters and voltmeters are connected to circuits. In principle, the internal resistance of the voltmeter should be much higher than the resistance of any resistor that it is connected across. In addition, the internal resistance of the ammeter should be much less than the resistance of any resistor that is under study. If these conditions are not satisfied, then the presence of the meters might seriously perturb the measurement of resistance. Modern digital voltmeters and ammeters have internal resistances that usually satisfy both these conditions.

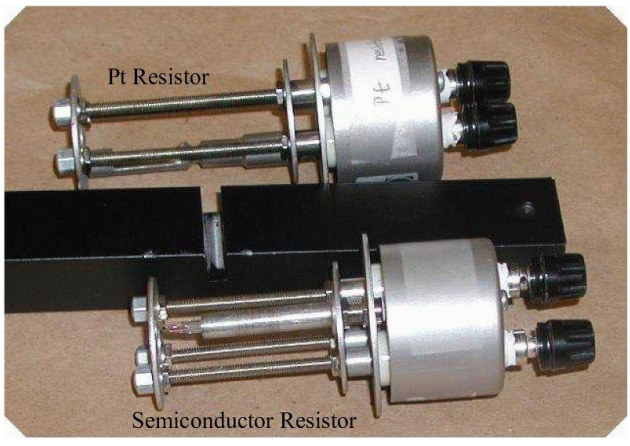


Figure 6: A photograph of the Pt and semiconductor resistors and sample holders.

Experimental Equipment

In this experiment, you will measure the temperature dependence of a Pt wire and a commercially available resistor made from semiconducting material. Fig. 6 is a photograph of the sample holders used in this experiment. These holders must withstand the high temperatures that will be used in this experiment. The Pt and semiconducting resistors are mounted at the left-hand end of these holders.

Fig. 7 is a photograph of the experimental apparatus showing the heating furnace with sample inserted, the thermocouple, a variac for heating the oven, and the CASSY interface that will be used to acquire data.

Make sure you know how the computer-controlled measuring apparatus works. What current is applied to the sample? Is it constant throughout the course of the experiment? Are you using a 2-wire or 4-wire technique?

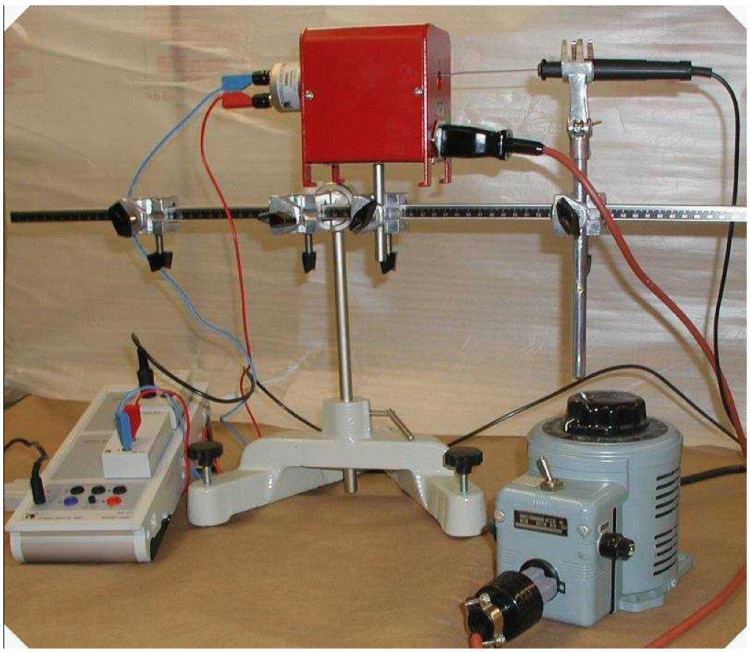


Figure 7: A photograph of the apparatus

Experimental Procedure

1. Measure the resistance of the Pt resistor as a function of temperature between room temperature and 350° C. Don't exceed this temperature as damage to the Pt sample might result.

a) Insert the Pt resistor assembly into the oven. From the other side insert a thermocouple. You may want to insert it at some angle so that its end would touch the inside of the metal tube which houses the Pt resistor. This way the temperature of the thermocouple end will be as close to the temperature of the resistor as possible.

b) Set the CASSY interface as follows: display x-axis to show temperature and y-axis to display resistance; data acquisition time to 500 ms and condition to "n=1 or delta(&JA11) > 2" (that will acquire data if it is the first point, or if the temperature

becomes at least 2 degrees higher than at the time of previous measurement). Do not define the number of data points, or define a maximum possible (16000). Set the CASSY ohmmeter limit to measure resistance of up to 300 Ω. The resistance of the Pt resistor should be about 100 Ω at room temperature.
c) Start the measurement and turn the variac ON and set voltage to ~100 V. The program will start taking measurements, one point for about every 2 C change.
d) Once temperature reaches 350 C turn the variac off, stop the data acquisition and save the acquired data.

B. Measure the resistance of the semiconducting resistor as a function of temperature between room temperature and 200^{o}C. Again, be careful not to exceed this upper temperature limit. The resistance of the semiconducting sample at room temperature should be about 200 Ω. The composition of the semiconducting material used in the fabrication of this resistor is unknown. It is very likely that it is not made from an elemental semiconductor like Si or Ge.
If you decide to measure the resistance as the oven cools down change the measurement condition to "n=1 or delta(&JA11) < -2". Expect the resistance to change between ~8 and 300 Ohm (set the CASSY ohmmeter limits accordingly).

Data Analysis

A. Analyzing data from the Pt resistor

You may want to analyze the data using your favorite data-fitting program, or perform data analysis using built-in CASSY-LAB fitting procedure.

1. It is recommended to convert all temperatures from ^{o}C to K. If you choose not to do so you must update the fitting equations accordingly.
2. Assume the resistance of the Pt resistor varies as $R=aT^{\beta}$ ($R=a[T+273]^{\beta}$ if T is in °C). Make a least square fit to your data leaving a and β free. How close is β to +1.00? Calculate the difference between your measured $R(T)$ and your best fit. Plot this difference as a function of T. Do you find any systematic differences in this plot? Is the data randomly scattered about zero? Explain what you observe.
3. Assuming temperatures near room temperature, the resistance change with temperature is usually specified by an equation of the form
$$R(T) = R_0\left[1+\gamma\left(T-273\right)\right] \qquad (16)$$
What is the best value for the coefficients R_0 and γ? When you quote these values, make sure you clearly specify the temperature range over which they are accurate. How does you value of γ compare to the value typically listed in introductory physics textbooks ($3.9\text{x}10^{-3}$ K^{-1}). Do you find any evidence that a higher order term proportional to T^2 is required?
4. The resistivity of Pt at 273 K is $9.60\text{x}10^{-8}$ Ωm. Using this value and the constant R_0 obtained above, determine the ratio of L/A for your wire.

3. B. Analyzing data from the semiconductor resistor

1. It is recommended to convert all temperatures from °C to K.

2. Assume the resistance of the semiconductor resistor varies as

$$R(T) = R_1\left(\frac{T}{273}\right)^{\frac{3}{2}} e^{\frac{E_G}{2kT}} \qquad (17)$$

   Fit your data using this equation with R_1 and E_G as free parameters.

   Alternatively, you may make a plot of $\ln\left[R(T)/\left(\frac{T}{273}\right)^{\frac{3}{2}}\right]$ vs T^{-1}, and use this graph to determine the best values of R_1 and E_G.

3. What is the band gap energy of the semiconductor? Express this energy in eV units which are conventionally used to measure band gap energy.
4. Make a more thorough analysis of your data. Calculate the difference between your data and best fit according to Eq. 17 as a function of T. Do you find any systematic differences between theory and data in this plot? Is the data randomly scattered about zero? Explain what you observe.

Appendix: Fermi Dirac Statistics

In the first studies (early 1900's) of the condensed state of matter, theories relying on a single average energy, a single average velocity, etc. were used to describe the physics of electrons in a solid. Later theories corrected this simplified approximation by using a Maxwell-Boltzmann distribution law to describe a gas of electrons. However, once it was realized that electrons obey quantum mechanics and the Pauli exclusion principle, a new quantum statistics was required. We now know that a gas of electrons (fermions) obeys Fermi-Dirac statistics.

For an ideal electron gas in thermal equilibrium with a heat bath of temperature *T*, the probability that an allowed state of energy *E* is occupied is given by

$$f(E) = \frac{1}{e^{(E-E_F)/kT} + 1} \qquad (18)$$

where the parameter E_F is defined by the energy at which $f(E)=\frac{1}{2}$. A plot of *f(E)* for three different temperatures is given in Fig. 8(a).

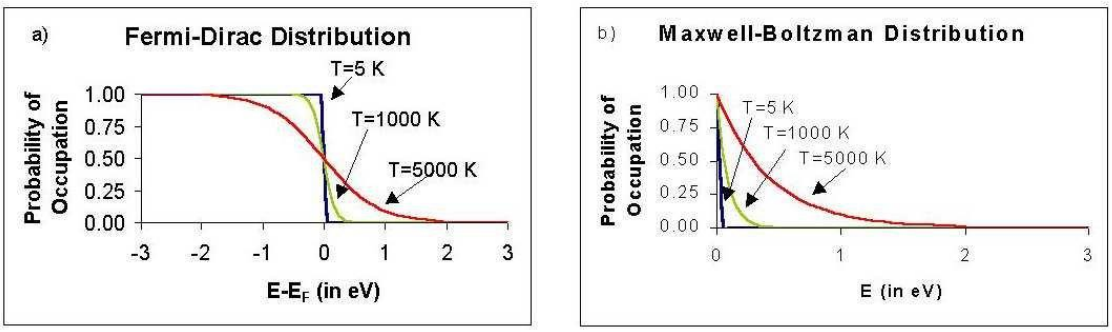


Figure 8. The Fermi-Dirac distribution function and b) the Maxwell-Boltzmann distribution function for temperatures of 5 K, 1000 K and 5000 K.

Note that when $E >> E_F$, the exponential term in the denominator of Eq. 18 dominates the additional factor of +1, and the Fermi-Dirac probability can be written as

$$f(E) = \frac{1}{e^{(E-E_F)/kT}} = A \times e^{\frac{-E}{kT}} \qquad (19)$$

Now *f(E)* closely approximates the form of the Maxwell-Boltzmann distribution function (see Fig. 8(b)), suggesting that the charge carriers can be viewed as a dilute gas of electrons. It turns out this approximation is valid when discussing electrons in the conduction band of an intrinsic semiconductor and provides a justification for the thermal activation model discussed in the Theory section above.

RR July 2000

SS Dec 2001

Physics 340 Laboratory
The Wave Nature of Light: Interference and Diffraction

Objectives: To demonstrate the wave nature of light, in particular diffraction and interference, using a He-Ne laser as a coherent, monochromatic light source.

Apparatus: He-Ne laser with spatial filter; photodiode with automatic drive, high voltage power supply for the laser, amplifier, computer with CASSY interface (no pre-amp boxes required), slits on a photographic plate, spherical and cylindrical lenses, diaphragm, and a razor blade.

References:

1. D. Halliday, R. Resnick and J. Walker, *Fundamentals of Physics; 5th Edition*, Wiley and Sons, New York, 1997; Part 4, pgs. 901-957.
2. E. Hecht, *Optics, 2nd Edition*, Addison-Wesley, Reading Massachusetts, 1974. Chapter 9 on interference, Chapter 10 on diffraction.
3. D.C. O'Shea, W.R. Callen and W.T. Rhodes, *Introduction to Lasers and Their Applications*, Addison-Wesley, Reading Massachusetts, 1978.

Introduction

In 1678, Christian Huygens wrote a remarkable paper in which he proposed a theory for light based on wave propagation phenomena, providing a very early theoretical basis for the wave theory of light. Because Huygens' theory could not explain the origin of colors or any polarization phenomena, it was largely discarded for over 100 years.

During the early 1800's, Thomas Young revived interest in Huygens theory by performing a series of now famous experiments in which he provided solid experimental evidence that light behaves as a wave. In 1801, Young introduced the interference principle for light which proved to be an important landmark and was hailed as one of the greatest contributions to physical optics since the work of Isaac Newton.

The interference principle was independently discovered by Augustin Fresnel in 1814. Unlike Young, Fresnel performed extensive numerical calculations to explain his experimental observations and thereby set the wave theory of light on a firm theoretical basis.

The interference and diffraction experiments performed by Young and Fresnel require the use of a coherent light source. While coherent light is difficult to produce

using conventional sources, the invention of the laser now makes intense coherent light readily available. In this experiment, you will reproduce some of Young and Fresnel's important discoveries using light from a He-Ne laser. In this way, you will become familiar with a few of the basic principles surrounding the wave theory of light. The remarkable successes of this theory explains why it was so prominent throughout the 1800's and why it was so difficult to challenge, even when convincing evidence for a quantized radiation field began to emerge in the 1890's.

Theoretical Considerations

Fraunhoffer diffraction, Fresnel diffraction

Diffraction phenomena are conveniently divided into two general classes:

1. Those in which the light falling on an aperture and the diffracted wave falling on the screen consists of parallel rays. For historical reasons, optical phenomena falling under this category are referred to as Fraunhoffer diffraction.
2. Those in which the light falling on an aperture and the diffracted wave falling on the screen consists of diverging and converging rays. For historical reasons, optical phenomena falling under this category are referred to as Fresnel diffraction.

A simple schematic illustrating the important differences between these two cases is shown in Fig. 1.

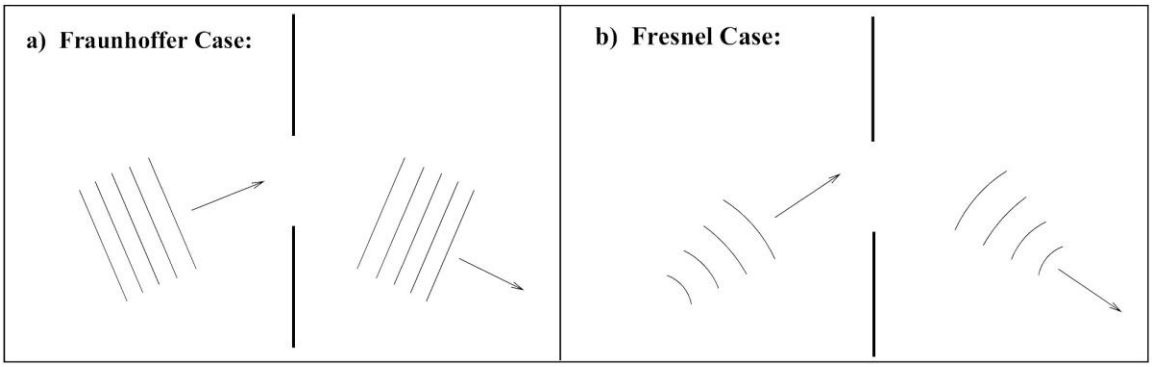


Figure 1: Qualitatively, Fraunhoffer diffraction (a) occurs when both the incident and diffracted waves can be described using plane waves. This will be the case when the distances from the source to the diffracting object and from the object to the receiving point are both large enough so that the curvature of the incident and diffracted waves can be neglected. For the case of Fresnel diffraction (b), this assumption is not true and the curvature of the wave front is significant and can not be neglected.

Fraunhoffer diffraction is much simpler to treat theoretically. It is easily observed in practice by rendering the light from a source parallel with a lens, and focusing it on a screen with another lens placed behind the aperture, an arrangement which

effectively removes the source and screen to infinity. In the observation of Fresnel diffraction, on the other hand, no lens is necessary, but here the wave fronts are divergent instead of plane, and the theoretical treatment is consequently more complex.

The important guiding principal of all interference and diffraction phenomena is the phase ϕ of a light wave. For light having a wavelength λ, the phase of the light wave at a given instant in time is represented by

$$\phi = \frac{2\pi}{\lambda} \times d \,. \qquad (1)$$

where d is distance traveled by light. If a light beam is equally split and the two split beams travel along two different paths *1* and *2*, then the phase difference $\Delta\phi$ between the two beams when they are recombined (after traveling distances x_1 and x_2) can be defined as

$$\Delta\phi = \phi_2 - \phi_1 = \frac{2\pi}{\lambda} \times (x_2 - x_1). \qquad (2)$$

In the wave theory of light, the spatial variation of the electric (or magnetic) field is described by a sinusoidal oscillation. When discussing interference and diffraction effects, $\Delta\phi$ appears in the argument of this sinusoidal function. Since the intensity *I* of a light wave is proportional to the square of its electric field vector, the intensity of two beams interfering with each other will be determined by factors proportional to $\sin^2(\Delta\phi)$ or $\cos^2(\Delta\phi)$. The exact expression for $\Delta\phi$ depends on the detailed geometry involved, but in general, $\Delta\phi = \pi/\lambda \times$ (*geometrical factor*).

A few important cases have been worked out in detail and the relative intensity variation $I(x)/I(0)$ produced by a coherent, monochromatic light beam as a function of position *x* along a viewing screen are given below. Because of the periodic nature of sinusoidal functions, they exhibit local maximums and zeros as the phase varies. The precise location of the maximums and zeroes can often be established by a calculation of the phase difference $\Delta\phi$.

Single Slit (Fraunhoffer limit)

If coherent light having a wavelength λ is made to pass through a long narrow slit of width *a*, then the relative light intensity as a function of lateral displacement *x* on a viewing screen located a distance R_0 from the slit is given by the expression:

$$\frac{I(x)}{I(0)} = \left[\frac{\sin\left(\frac{\pi a}{\lambda R_0} x\right)}{\frac{\pi a}{\lambda R_0} x} \right]^2 = \left(\frac{\sin u}{u} \right)^2 \qquad (3)$$

where:

a = width of the slit

λ = wavelength of radiation

R_0 = distance between cylindrical lens and viewing plane

$$u \equiv \frac{\pi a}{\lambda R_0} x$$

Double Slit (Fraunhoffer limit)

If coherent light having a wavelength λ is made to pass through two long narrow slits of width *a* with a center-to-center separation *b*, then the relative light intensity as a function of lateral displacement x on a viewing screen located a distance R_o from the slits is given by the expression:

$$\frac{I(x)}{I(0)} = \left[\frac{\sin\left(\frac{\pi a}{\lambda R_0} x \right)}{\frac{\pi a}{\lambda R_0} x} \right]^2 \cos^2\left(\frac{\pi b}{\lambda R_0} x \right) = \left(\frac{\sin u}{u} \right)^2 (\cos v)^2 \qquad (4)$$

where:

a = width of the slits

b = center-to-center separation between the two slits

λ = wavelength of radiation

R_0 = distance between cylindrical lens and viewing plane.

$$u \equiv \frac{\pi a}{\lambda R_0} x .$$

$$v \equiv \frac{\pi b}{\lambda R_0} x .$$

The first term in Eq. 4 is realted to diffraction through a single-slit as given by Eq. 3 above. The second term is due to interference from light passing through a double-slit.

Knife Edge (Fresnel limit)

If coherent light with intensity I_o and wavelength λ is made to pass across a sharp knife edge, then the relative light intensity as a function of lateral displacement x on a viewing screen located a distance R_o from the knife edge is given by the expression:

$$\frac{I(x)}{I_0} = \frac{1}{2}\left[\left(\frac{1}{2} - C(\omega) \right)^2 + \left(\frac{1}{2} - S(\omega) \right)^2 \right] \qquad (5)$$

where:

I_o is the intensity of the unobstructed beam

λ = wavelength of radiation

R_o = distance between the knife edge and viewing plane

R_s = distance between pinhole of spatial filter and knife edge

$$\omega = \left[\frac{2(R_S - R_0)}{\lambda R_S R_0}\right]^{\frac{1}{2}} \cdot x$$

C(ω) and S(ω) are the Fresnel integrals tabulated in Appendix B.

It should be evident from the above discussion that the detailed variation of intensity depends on the geometry of the experimental set-up. It should also be clear that the phase difference plays an important role. Representative plots for the intensity as a function of displacement are given in Fig. 2.

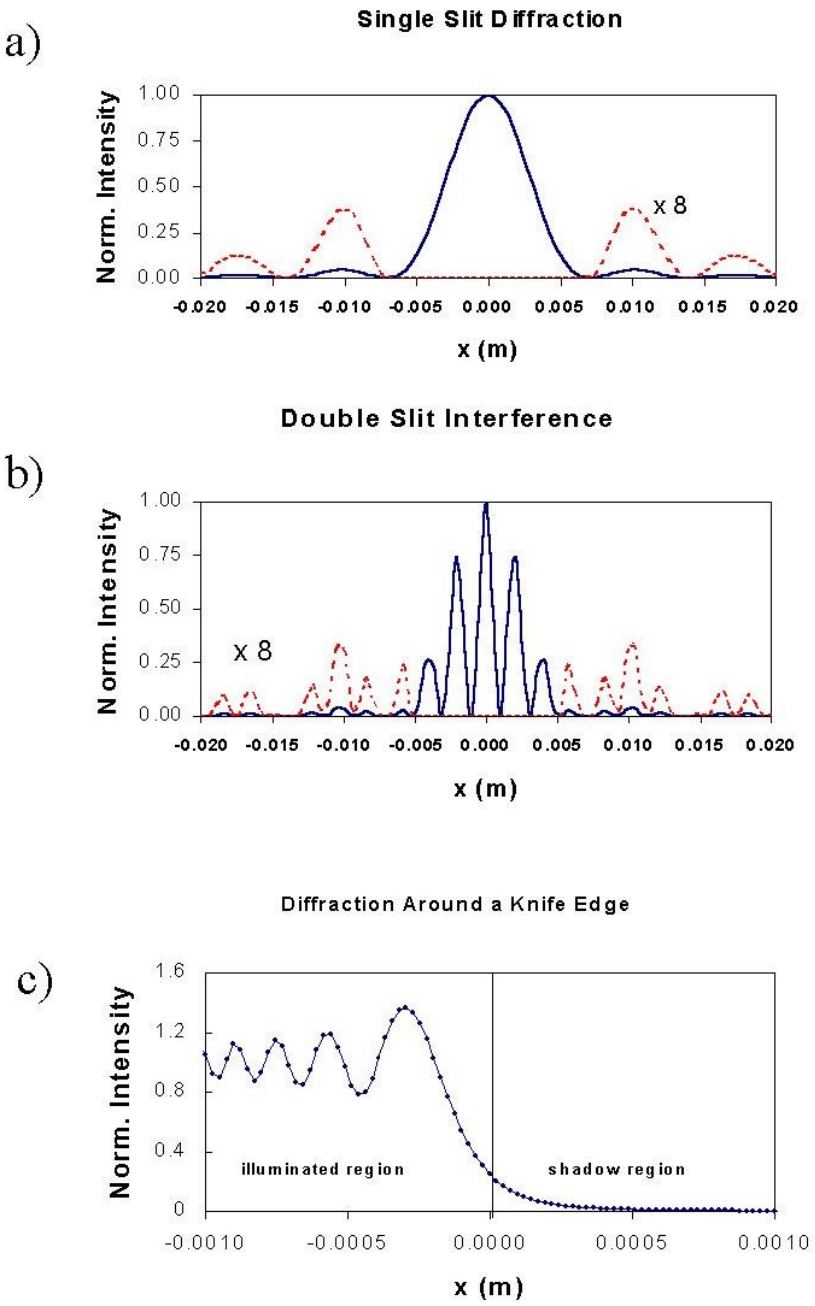


Figure 2: Representative intensity variations produced from (a) diffraction by a single slit (λ= 632.8 nm; a= 27 μm), (b) interference from a double slit (λ= 632.8 nm; a= 27 μm; b= 270 μm), and (c) diffraction from a knife edge. The dashed lines in (a) and (b) are the resulting intensity variation after multiplication by a factor of 8.0

Experimental Considerations

A photograph of the optical set-up of the equipment is given in Fig. 3. This photo shows the relative placement of the He-Ne laser with spatial filter, a focusing lens, an adjustable diaphragm, the slits, a cylindrical lens, and a viewing plane which consists of a scanning photodiode driven by a slow motor. The photodiode is apertured by an adjustable slit which controls the resolution of the detected intensity variations. A schematic diagram of this set-up is given in Fig. 4. A few of the more important details of the equipment are provided below.

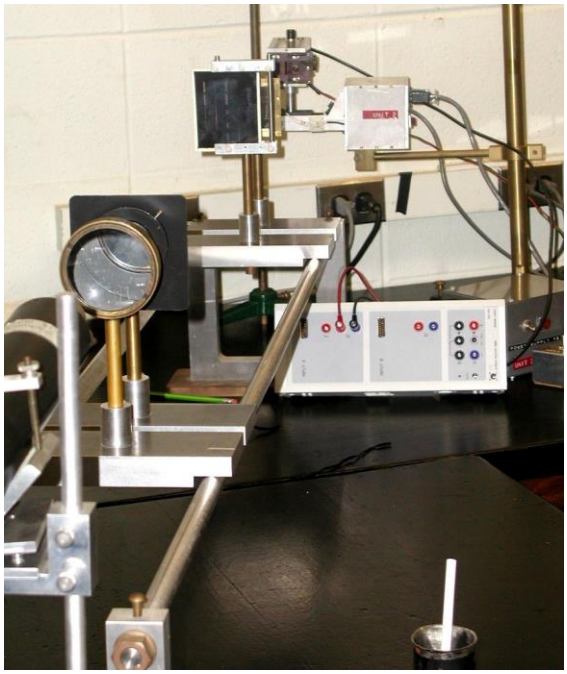


Figure 3: A photograph of the experimental set-up showing the He-Ne laser, the focusing lens, the adjustable diaphragm, the slit plate, the cylindrical lens, the scanning photodiode and the CASSY interface box. The spatial filter is not visible in this photo.

A. The He-Ne Laser

See Appendix A for a more detailed discussion.

B. Spatial filters

An ideal continuous laser produces laser beam that has gaussian intensity distribution in cross-section (see Fig. 11 in Appendix). In our experiment we need to produce plane or concentric waves with minimal intensity variation in the slit area.

Real laser systems often have internal apertures that produce diffraction pattern in output beam. This diffraction pattern resembles Newton's rings (i.e., a "bull's eye" variation in intensity) and may not be concentric with the main beam. It is a great nuisance for accurate diffraction measurements.

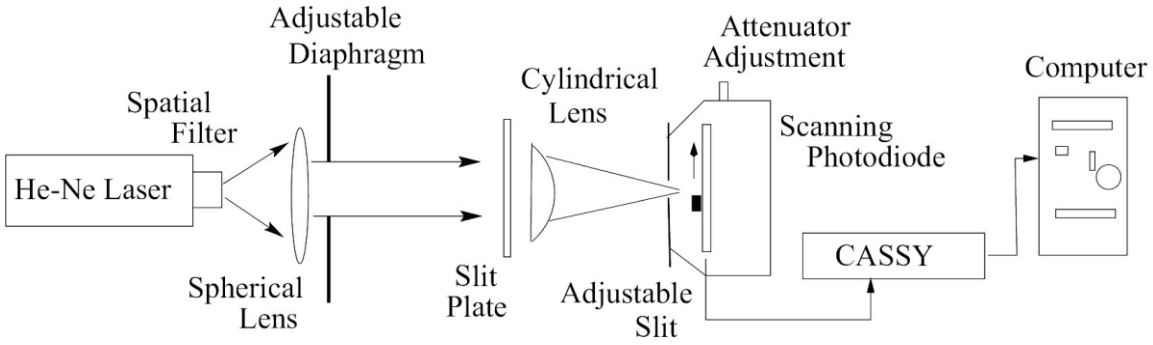


Figure 4: A schematic diagram of the experimental set-up.

This "bull's eye" pattern can be eliminated with a spatial filter. (See Fig. 5). The spatial filter is an adjustable arrangement consisting of a strongly converging lens and a small aperture pinhole located in the center of an opaque plate. The laser beam is focused by the lens through the pinhole aperture and the distorted segment of the laser beam is spatially blocked out by the small pinhole.

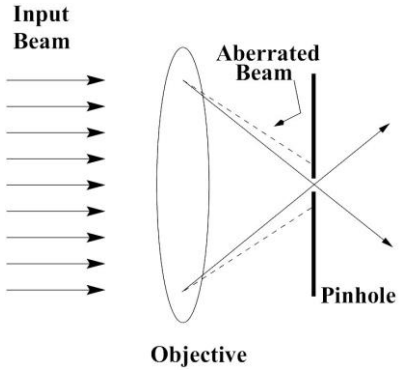


Figure 5: A schematic diagram of a spatial filter showing the incident laser beam, the microscope objective and the pinhole. The pinhole blocks the aberrated part of the beam if it is located precisely at the focal point of the objective lens.

Before you enter the lab, every attempt is made to adjust the spatial filter properly. There are two adjustments. One adjustment centers the converging laser beam onto the center of the pinhole. This adjustment is rather difficult to make. The second adjustment places the pinhole at the focal spot of the laser. Do not touch the screws on the spatial filter because readjustment can be time consuming. If problems arise, call the lab instructor.

Since the spatial filter is itself an aperture, it also creates a concentric "bull's eye" diffraction pattern at the edges of the beam spot. An adjustable aperture (~1cm diameter) is used to block all but the central maximum.

C. Description of the slit pattern.

In this experiment we use transparent slits etched on an opaque film and enclosed between two glass plates (Fig. 6). The single slit has a width on order of 100 μm and is located in the top-right corner of the plate. The double slit is located in the middle-right section of the plate, each slit is ~100 μm wide and the distance between two slits is about 300 μm.

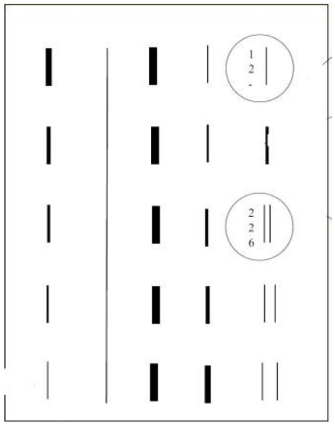


Figure 6: A schematic diagram showing the layout of the slit plate. For this experiment, use only the slits which are circled.

D. Description of the photodiode and data acquisition system.

The intensity variation of the diffracted wave is measured with a photodiode that is mounted on an automatic drive moving with a speed of 85×10^{-6} m/s. The amount of light reaching the diode is determined by the gap between a pair of jaws (slit) in front of it. One jaw is fixed, the other jaw is spring-loaded and thus adjustable. This input slit also determines spatial resolution of the experiment – the slit size should be significantly narrower than the sharpest features of the measured interference/diffraction pattern.

The signal is amplified and then recorded on the computer using the CASSY computer interface. The time interval between the acquisition of two data points digitized in the time mode can be adjusted through the software. Data points acquired roughly every 0.1 seconds will provide sufficient resolution in this experiment and will result in data files containing ~1000 data point pairs [t, I(t)]. Knowing the speed of the photodiode, these data can be converted to [x, I(x)] and compared with theoretical expectations given above.

E. Alignment of the optical elements.

Note: the alignment described in this section is usually done by the time you enter the lab.

1. Position the photodiode at right end of the optical rail. Make sure the adjustable slit to the photodiode is closed (gently!) and that the photodiode power is off.
2. Put the laser on its own stand at the other end of the optical bench and the photodiode at the other. Adjust the height and tilt of the laser support in such a

way that the laser is at the same height as the photodiode and that the laser beam hits the photodiode.

3. Set the adjustable diameter diaphagm as close to the photodiode as possible; make the hole ~1 mm in diameter and adjust its height so that laser beam which passes the hole hits the center of the photodiode. You may need to align the center position of the diode left and right as the diaphragm position can be aligned only vertically.
4. Without changing its height, move the diaphragm to the other end of the bench close to the laser. Realign the laser height and tilt so that laser beam passes the diaphragm and hits center of the photodiode.
5. Temporarily remove the diaphragm (with its stand) from the bench and position the spherical converging lens ~17.5 cm (focal length) from the laser end. Set the lens height so that center of the laser beam passes through it and still hits center of the diode. Fine-tune the distance from the lens to respect to the laser end so that laser beam which emerges from the lens does not change its size on its way to the photodiode. That would ensure that the beam front is parallel.
6. Place the stand with the diaphragm behind the converging lens (few centimeters apart). We use the diaphragm to remove all but the center maximum of the interference pattern created by the aperture of the spatial filter. Set the diameter of the diaphragm at about 1 cm.
7. Insert the plano-concave cylindrical lens between the photodiode box and the spherical lens with its plane surface towards the laser. Its exact location can be found by the requirement that the image of the laser beam should form a bright narrow vertical line on the slit mounted on the photodiode box. Thus the slit or more accurately the photodiode is located at the focal point of the lens: $R_0 = f_c \approx$ 32 cm.
8. Put the single slit denoted as (1; 2; -) in the beam between the diaphragm and cylindrical lens but as close as possible to the cylindrical lens. Make sure the single slit is in the center of the laser beam and the plate is normal to the laser beam. Adjust the plane of the glass plate until the reflected and incident light beams become parallel.

F. Final Adjustments.

1. Always open and close the jaws of the photodiode gently; otherwise they may get stuck shut.
2. Start up the data acquisition software and adjust the settings so one data point is collected about every 0.1 second. Position the scanning photodiode box so it is located at the midpoint of the diffraction pattern (brightest area). The scan has two end switches that automatically switch off the motor when the carriage reaches either end of the range of travel.
3. Switch on the amplifier. Slowly open the photodiode slit until CASSY voltmeter shows ~ 0.5-1V. Selecting too wide of a slit in front of the photodiode would compromise spatial resolution of your detector, selecting too narrow slit would compromise signal/noise ratio as too little light gets into photodiode.

4. The diffraction pattern is parallel with the vertical slit defined by the jaws on the photodiode box. This ensures that the diffraction pattern falls within the range of the automatic scan, and that the photodiode box is at the right height.

I. Measuring slit width.

Use a traveling microscope to measure both slit width (and separation). Don't forget to estimate an error of this measurement.

II. Diffraction produced by a single slit

1. With the set up described above, make the photodiode move to one of the ends of the scanning range. Then reverse the scan direction and start data acquisition to scan the diffraction pattern into computer. Your data is usable only if you record the first minimum, the secondary maximum, and in case of optimal line up, the second minimum.
2. Record all pertinent data in your lab notebook for future reference.
3. Your digitized data should resemble something like Fig. 2(a). Don't forget to save your data into a file.

Analysis of the single slit data

1. Knowing the speed of the automatic drive and the speed of data acquisition (85×10^{-6} m/s), you can convert digitization times measured on the computer into distances traveled by the photodiode.
2. Read your experimental data into a spreadsheet program and generate a plot of relative intensity vs. position. You must shift your data so that the central maximum coincides with x=0. You may want to normalize your data so that the relative intensity at x=0 is unity. You may need to discard some of your digitized data taken far away from x=0 since it may not contain useful information.
3. Write a computer program using Eq. 3 and calculate a theoretical fit to your data. Adjust the slit width, but hold the wavelength of the laser fixed. Make a plot of theory and experiment. Make sure your final plot comparing theory to experiment is clearly labeled. What is the best value of the slit width required to fit your data? How close is this measurement to the one done by microscope

III. Interference and diffraction from a double slit

1. Change the position of the glass slide in such a way that the center of the laser beam hits the (2; 2; 6) double slit.
2. Record the diffraction pattern following similar procedures which are given above. Make sure that the voltmeter is not overloaded in maximum of the interference pattern – since there are now two slits the light intensity in the middle of the pattern might be higher. You must record four secondary peaks on each side of the main central peak in order to observe the first zero of the first factor in Eq. 4.
3. Your digitized data should resemble something like Fig. 2(b).

Analysis of the double slit data

1. Read your experimental data into a spreadsheet program and generate a plot of relative intensity vs. position. Again, you must shift your data by identifying x=0 and you may want to normalize your intensity so that it is unity at x=0. You may need to discard some of your digitized data taken far away from x=0 since it may not contain useful information.
2. Write a computer program using Eq. 4 and calculate a theoretical fit to your data. To optimize the fit, adjust the slit width and slit separation, but hold the wavelength of the laser fixed. Make a plot of theory and experiment. What are the best values required to fit your data? What range of variation in the adjustable parameters can you make before you observe a significant discrepancy between theory and experiment? For that analysis, set the distance between slits to values 5%, 20%, 30% and 50% from its optimal values and repeat the fit changing only slit width. That will give you a measure of precision at which you were able to determine the distance between the slits based on your data. Repeat simila analysis for the width of the slit. Based on the above fits estimate the error range for double slit parameters.

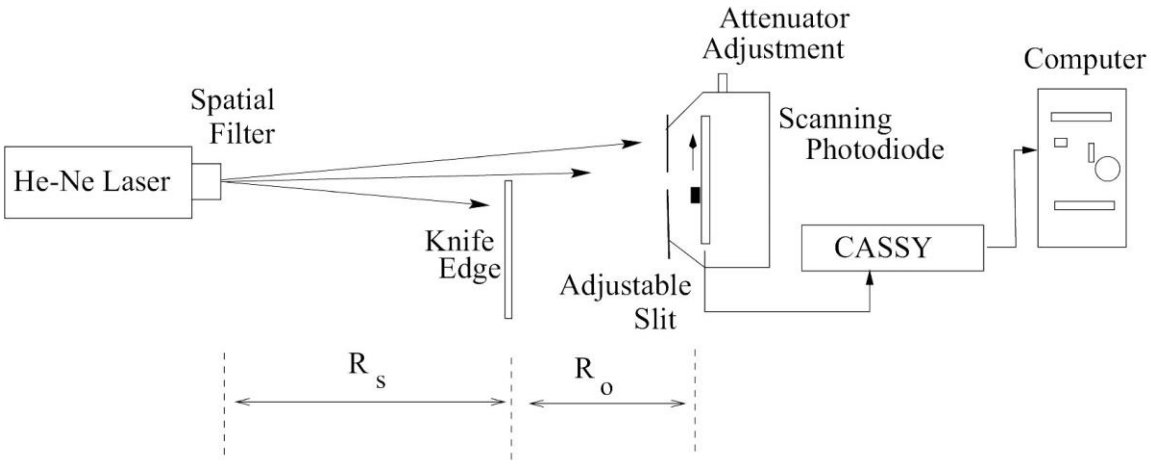


Figure 7: Set-up for Fresnel diffraction around a knife edge.

III. Diffraction from a knife edge

1. Remove both the converging lens and the cylindrical lens from the set-up for Fraunhoffer diffraction. This destroys the parallel ray approximation and places you in the Fresnel regime.
2. Remove the slit slide and insert a sharp edge of a new razor blade. Make sure the knife edge is parallel to the aperture of the photodiode. The edge should be in the center of the laser beam. Fig. 7 gives a schematic diagram of your set-up.
3. Record all relevant data in your notebook for future reference.

4. Record the diffraction pattern using the computer following procedures given earlier. You may need to reduce the slit width in front of the photodiode to avoid signal overload.
5. Your digitized data should resemble something like Fig. 2(c).

Analysis of the knife edge data.

1. Read your experimental data into a spreadsheet program and generate a plot of relative intensity vs. position. Again, the position of your data must be shifted and the intensity may be normalized to unity. You may need to discard some of your digitized data taken far away from x=0 since it may not contain useful information.
2. With the help of Table 1 in Appendix B and Eq. 5, calculate $I(x)/I_0$ as a function of x and plot it on top of your data. Make sure your final plot comparing theory to experiment is clearly labeled.
3. In your error analysis, only consider the errors in measuring the distances R_s and R_o. Compare your fit to theoretical expectations. Can you explain the reason for any differences between theory and experiment?

Appendix A: Theory of the He-Ne Laser

The laser is a modern light source with several interesting properties. The name is an acronym for Light Amplification by Stimulated Emission of Radiation. The light coming from a laser is unidirectional, monochromatic, intense and coherent. Let us try to understand these properties and the laser operation.

In 1916, Einstein worked on the blackbody problem, and in the process of solving this problem he also identified *spontaneous* and *stimulated emission rates.* Let us consider the simple system consisting of atoms which have two nondegenerate electronic states: so called *ground* state |0>, where the electron resides normally, and the *excited* state |1>, which is normally unoccupied.

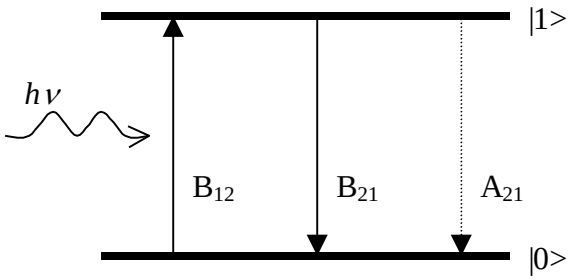


Figure 8. Electronic state diagram. |0> and |1> are ground and excited states, correspondingly. The A_{21}, B_{12} *and* B_{21} are Einstein's coefficients for possible electronic transitions.

When a light photon with energy $h\nu$ equal to energy separation between these two states strikes an atom, the probability W_{01} of the atom in the ground state to absorb light is proportional to B_{12} (see Fig. 8). Once in the excited state, an electron undergoes a spontaneous transition into ground state with a probability proportional to A_{21}. The latter process is often accompanied by radiation of a photon. The two processes described so far are responsible for absorption and fluorescence of the atom. However, if the atom is irradiated when it was already in the excited state, then the light photon can promote an electronic transition downward with probability W_{10} proportional to B_{21}. For this process, a second photon is emitted with the exact energy, phase, polarization and direction of the falling photon. Einstein called this phenomenon "negative absorption" (the modern term is *stimulated emission*), as it is the process opposite to absorption. For several decades it was considered more like a nuisance than a fundamental discovery that ultimately lead to the invention of lasers.

According to Einstein $B_{12}=B_{21}$. Therefore, in order to achieve amplification of incident light in a media, one needs to manipulate the ensemble of atoms or molecules in such a way that the number of atoms in excited state is larger than that in the ground state. In symbols, this means $N_eB_{21}>N_gB_{12}$, where N_e and N_g are populations of excited and ground states accordingly. This, however, turned out to be not an easy task. One can illuminate the ensemble of atoms with a strong light pulse

that is in resonance with electronic transition (optical pumping). However, one cannot excite more than half of atoms in such a way. The reason is simple – as the number of atoms in excited state reaches the number of atoms in ground state the same amount of transitions downward and upward will be produced by incident light, and the net change in population will be zero. In order to solve this problem atoms or molecules with 3 or 4 electronic states can be used.

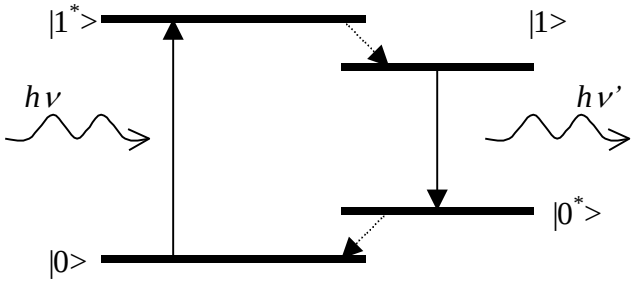


Figure 9. Electronic state diagram of 4-level system. |0> and |1> are ground and excited states, correspondingly.

In a 4-level system (see Fig. 9), absorption of a photon creates the nonequilibrium excited state $|1^*>$ which rapidly decays into intermediate state |1>. This state, in turn, decays into a nonequilibrium ground state $|0^*>$ with the emission of a photon $h\nu'$, and, consequently, the system returns rapidly into ground state |0>. The pump photon energy $h\nu$ is not in resonance with the laser transition $h\nu'$, and thus pumping light does not cause depopulation of the laser-active state |1>. Besides, the nonequilibrium ground state $|0^*>$ is initially not populated, and achieving population inversion (i.e. $N_e > N_g$) becomes easy.

Optical pumping is not the only way to prepare laser active media. One may use electrical current to excite optical states, or create molecules in excited state using chemical reactions.

The presence of light-amplifying media is not sufficient for laser operation. One also needs to create a *positive feedback* in a similar way as done in electronic oscillators. In positive feedback, part of the output signal from the electronic amplifier is send back into its input, and as a result the oscillation never stops. This effect is similar to the commonly observed phenomenon when a microphone is placed close to the speaker.

The most commonly used optical scheme that provides positive feedback for a light oscillator is shown in Fig. 10.

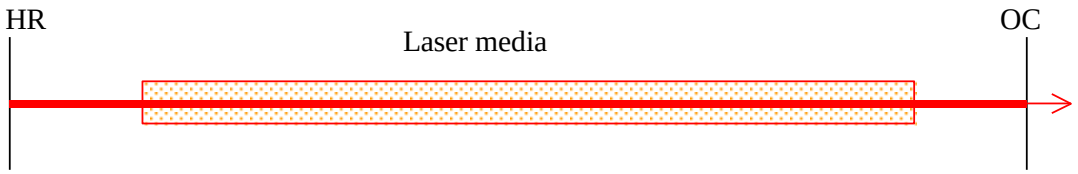


Figure 10. Basic laser design. HR – high reflector mirror, OC – output coupler mirror. The two-mirror design resembles a Fabry-Perot etalon, which is known for its unique optical properties since 1899.

Two parallel mirrors (HR and OC) reflect light back and forth through the laser media where the light is to be amplified. The output coupler (OC) lets some of the light through providing usable laser output beam. Initially, there is no laser light. As time passes, the laser media due to spontaneous transitions (fluorescence) emits some of the photons. Those photons that are not parallel to the laser axis escape from the media and are not further amplified. However, there is always a probability that a photon is spontaneously emitted along the laser axis. Due to the mirrors, this photon will bounce back and fourth between the mirrors. In each pass additional photons are created due to stimulated emission, these photons in turn are amplified. After a short period of time the intensity of the beam rises and we have laser operation. Ideally, all the photons in the laser would be created from a single initial photon, and therefore have the same energy, direction, phase and polarization.

Since it is impossible in principle to create absolutely parallel laser beam (due to diffraction), the laser mirrors are often curved to compensate for diffraction.

Let us now list the main properties of laser light.

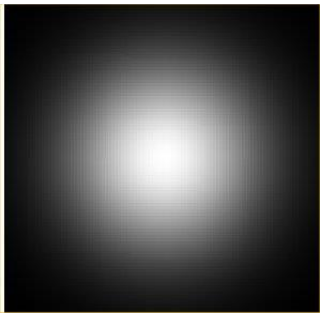


Figure 11. Gaussian beam profile

1. The laser light is highly directional. In best lasers the output beam has gaussian profile, i.e. the laser spot intensity varies as a gaussian function with maximum intensity in the middle as shown in Fig. 11. The advantage of this profile is in its diffration properties; while the overall size of the gaussian beam will change as it passes through space or optical elements, the functional shape will remain gaussian.

2. The laser light is highly monochromatic, i.e. most of the light intensity is concentrated in a very narrow spectral range. One can also produce monochromatic light by passing "white" light through a narrow spectral filter (like a monochromator). The latter, however, will have extremely low intensity and its spectrum will be still several orders of magnitude broader than that of a narrow-band lasers.

3. The laser light is coherent. That means that there is a fixed phase relationship between different parts of the laser radiation, both across and along the beam.

The lasing medium in a modern He-Ne laser used in our experiment is a mixture of about 85% He and 15% Ne, with Ne providing the lasing action. This mixture is excited in a glow discharge by passing a dc current through the gas. Transition of the Ne atoms to the excited state is not caused directly by the current but by indirect pumping accomplished by electron excitation of He atoms in the glow discharge. The excitation energy of the He is then transferred to the Ne atoms by way of atomic collision. The principle depends on the existence of excited levels in the Ne atom, which are close to the first excited states in He, and can therefore be populated by resonant neon-helium collisions. In addition, the He levels are metastable thus ensuring the most efficient energy transfer to the Ne atoms because the return of the excited electrons to the He ground state by radiative decay is very slow.

The level scheme for the He-Ne laser is shown in Fig. 12. To explain the operation of the He-Ne laser, the notation of levels in Fig. 12 must be understood. In case of light atoms the spins of the electrons are added vectorially to obtain the resultant spin $\vec{S}$. Next the orbital angular momenta of the electrons are added vectorially to obtain $\vec{L}$. The vector sum of the two is the total angular momentum $\vec{J}$. Levels are labeled then as $^{2s+1}L_j$. Thus the ground state and the two relevant excited states of He are denoted as 1 $^1S_0$, 2 $^3S_1$ and 2 $^1S_0$. The symbols S, P, D are denoted by L=0,1,2 respectively. This is called the Russell-Saunders coupling scheme notation. The numbers in front of $^1S_0$ and $^3S_1$ denote the principal quantum number of the excited electron.

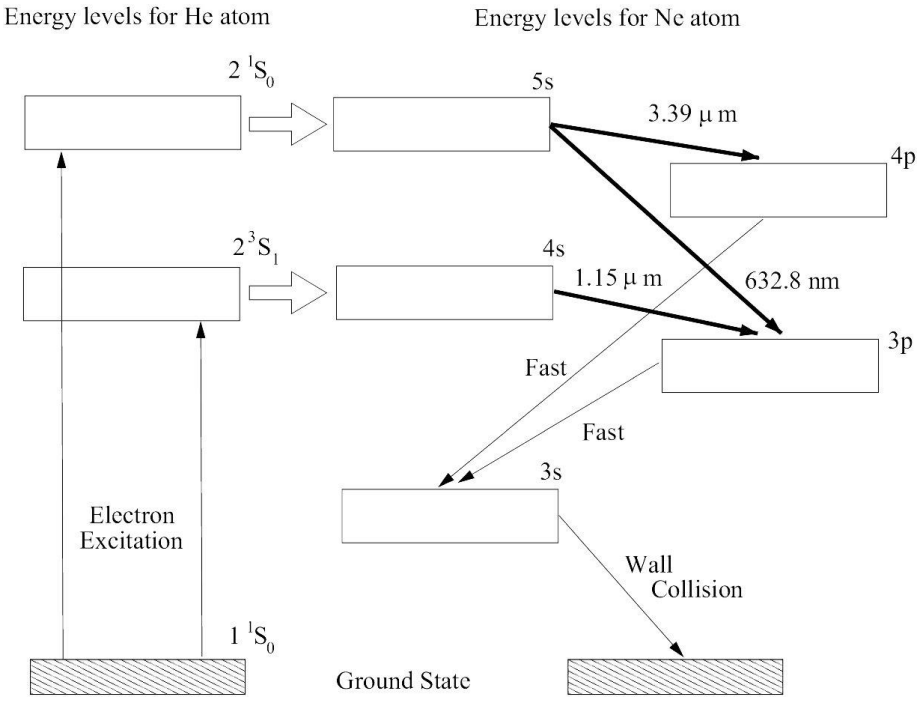


Figure 2: A schematic diagram of the energy level scheme for the He-Ne laser. The three darker solid lines indicate a laser transition.

To describe the excited states of Ne, note the electron configuration of the ground state is $1s^2 2s^2 2p^6$, where the superscripts denote the number of electrons in each state. The ground state is denoted by the usual spectroscopic notation $^1S_0$ (L=0,S=0,J=0). The configuration in which one of the $2p^6$ electrons is moved into the 3*s*,4*s* or 5*s* state are denoted by 3*s*,4*s* and 5*s*. Similarly when one of the $2p^6$ electrons is moved into either the 3*p* or 4*p* levels, the states are denoted by 3*p* and 4*p*.

Both the excited 2 $^3S_1$ and the 2 $^1S_0$ levels of He are metastable. The 2 $^1S_0$ state cannot decay by single photon emission since to do so would violate the conservation of angular momentum. This state decays by emitting two photons with a ~19.5 ms lifetime. By coincidence, the 2 $^1S_0$ level in He lies within 0.05 eV of the 5*s* Ne level. Thus excitation of Ne via the atomic collision:

$$He(2\ ^1S_0)+Ne(^1S_0)\rightarrow He(1\ ^1S_0)+Ne(5s)\text{-}0.05eV$$

can take place. Stating in words: when a Ne atom in the ground state collides with an excited He atom in the 2 $^1S_0$ metastable state there can be a reaction whose outcome is an excited Ne atom in the 5*s* state and a He atom in the ground state. The missing 0.05 eV energy is obtained from the thermal kinetic energy of the colliding atoms. As shown in Fig. 9, the 5*s* state can emit induced radiation in two ways. The energy level differences are:

$$E(5s)\text{-}E(3p)=1.96eV \ \ and \ \ E(5s)\text{-}E(4p)\approx 0.3eV$$

The laser you will be using emits the ΔE=1.96 eV visible red radiation (λ=632.8 nm). The five-level system (5s,4s,3s,4p,3p) of a He-Ne gas laser differs from the three-level system of chromium in that the emission of a photon does not return the Ne atom to the ground state. Transitions from the 3s state to the ground state are accomplished through a phonon (a quantized particle of sound) transition in which energy is transferred mainly through heat.

The He-Ne gas mixture is contained in a sealed tube. Excitation of the He is accomplished by a discharge of electricity through the tube, similar to a neon sign. Selection out of the possible five transitions of a single 1.96 eV transition is accomplished by having the mirrors in the laser tube made of excellent reflectors for the 1.96 eV photons and poor reflectors for the other stimulated (lasing) transitions in the green and infra red.

The laser tube and its mirrors forms an optical cavity which produces an interesting optical spectrum. The 632.8 nm laser transition (frequency = 4.74×10^{14} Hz) is not infinitely sharp but is broadened (roughly by ±750 MHz) by thermal motion (Doppler broadening) of the atoms inside the laser tube. This broadened transition in turn supports a number of discrete axial cavity modes which are separated in frequency by

(roughly 600 MHz), where c is the speed of light and L (~0.25 m) is the length of the laser cavity. It is possible to observe this mode structure in the laser beam output, but highly stable conditions are required to do so. Special optical components mounted on the laser tube can further tailor the laser beam output by tuning for a particular axial mode or by selecting a particular plane of polarization for the output laser beam.

Appendix B: Evaluation of the Fresnel integrals

Diffraction of coherent light around a knife edge requires the evaluation of two special integrals

$$S(w) = \int_0^w \sin\left(\frac{1}{2}\pi w^2\right) dw \qquad (1)$$

and

$$C(w) = \int_0^w \cos\left(\frac{1}{2}\pi w^2\right) dw\,. \qquad (2)$$

These integrals have the property that S(-w) = -S(w) and C(-w) = -C(w). A tabulation of these integrals is given below.

Table 1. Coordinates of the Cornu Spiral

w	C(w)	S(w)	w	C(w)	S(w)	w	C(w)	S(w)	w	C(w)	S(w)
0	.0000	.0000	1.3	.6386	.6863	2.6	.3890	.5500	3.9	.4223	.4752
.1	.1000	.0005	1.4	.5431	.7135	2.7	.3925	.4529	4.0	.4984	.4204
.2	.1999	.0042	1.5	.4453	.6975	2.8	.4675	.3915	4.1	.5738	.4758
.3	.2994	.0141	1.6	.3655	.6389	2.9	.5624	.4101	4.2	.5418	.5633
.4	.3975	.0334	1.7	.3238	.5492	3.0	.6058	.4963	4.3	.4494	.5540
.5	.4923	.0647	1.8	.3336	.4508	3.1	.5616	.5818	4.4	.4383	.4622
.6	.5811	.1105	1.9	.3944	.3734	3.2	.4664	.5933	4.5	.5261	.4342
.7	.6597	.1721	2.0	.4882	.3434	3.3	.4058	.5192	4.6	.5673	.5162
.8	.7230	.2493	2.1	.5815	.3743	3.4	.4385	.4296	4.7	.4914	.5672
.9	.7648	.3398	2.2	.6363	.4557	3.5	.5326	.4152	4.8	.4338	.4968
1.0	.7799	.4383	2.3	.6266	.5531	3.6	.5880	.4923	4.9	.5002	.4350
1.1	.7638	.5365	2.4	.5550	.6197	3.7	.5420	.5750	5.0	.5637	.4992
1.2	.7154	.6234	2.5	.4574	.6192	3.8	.4481	.5656	∞	.5000	.5000

Physics 340 Laboratory

Wave Properties of Matter: Electron Diffraction

Objective: To demonstrate the phenomenon of electron diffraction and to use it as a proof for the validity of the wave description of moving electrons.

Apparatus: Sargent-Welch Model 2629-A electron diffraction tube, micro-ammeter, clear plastic ruler, caliper, magnetized rod.

References:

1. Louis de Broglie, Nature **112**, 540 (1923).
2. C.J. Davisson and L. Germer, Nature **119**, 558-560 (1927).
3. C.J. Davisson, Franklin Institute Journal **205**, 597 (1928).
4. Weidner and Sells, Elementary Modern Physics, Sec. 5-1 through 5-5, especially note Fig. 5-7 for comparison with your data.
5. K. Krane, *Modern Physics*, 2nd Ed., Wiley and Sons, New York, 1996, pgs. 100-110.

Introduction

By the 1920's, the importance of Einstein's pioneering work in explaining the photoelectric effect as well as his seminal theory of relativity were widely recognized. The photoelectric effect pointed strongly to the particle properties of light. Only a few years before, it was widely believed that light was governed by wave physics. The dilemma between a wave or particle description for light was clearly a topic of considerable interest.

Amid this lively debate on the properties of photons, Louis DeBroglie wondered if it might be possible for the newly discovered sub-atomic particle, the electron, to exhibit wave properties. This was a completely foreign idea. After all, the charge to mass ratio of the electron was measured by J. J. Thomson in 1897. Millikan, using oil drops in 1911, measured the charge of the electron "e". The electron has mass, charge and therefore can move with any velocity v within the limits $0 \leq v < c$. Thus in a classical 19^{th} century framework, the electron was clearly a particle.

Nonetheless, DeBroglie proposed the revolutionary idea that particles might behave like a wave. As Einstein had shown, a photon of frequency f acts as if it carries a qunatized energy E_{photon} given by

$$E_{photon} = hf = hc/\lambda \ . \qquad (1)$$

Eq. 1 implies that

$$\frac{h}{\lambda} = \frac{E_{photon}}{c} \quad . \qquad (2)$$

A similar formula can be obtained starting from Einstein's formula for relativistic energy of a particle with rest mass of m_o and momentum p. According to Einstein, the total energy E_{total} of a particle is given by

$$E^2_{total} = p^2c^2 + m_o^2c^4 \quad . \qquad (3)$$

When the rest mass m_o is set equal to zero, E_{total} then must equal E_{photon} , and one quickly learns that the momentum of a photon is given by

$$p = \frac{E_{photon}}{c} \quad . \qquad (4)$$

Combing Eq. 2 and Eq. 4, the well-known expression connecting the wavelength of a photon to the photon's momentum is found

$$\lambda = \frac{h}{p} \quad . \qquad (5)$$

DeBroglie's hypothesis was that a similar formula might apply to an electron, leading to the prediction that an electron with velocity v and mass m_e would have a wavelength given by

$$\lambda = \frac{h}{m_e v} \quad . \qquad (6)$$

It would be extremely interesting if massive particles like electrons exhibit dual wave-particle behavior. This impelled de Broglie to suggest that massive particles may also behave like waves and thus exhibit interference and diffraction effects.

Electron Wavelength in the Nonrelativistic Energy Regime

In most everyday situations, electrons have low energy and, in this case, the kinetic energy of the electron is specified rather than the total energy. The kinetic energy (E_k) of an electron is easy to determine and is equal to E_k=eV where e and V denote the charge of the electron and V the accelerating potential respectively. Thus

$$E_{total} = E_k + m_e c^2 \qquad (7)$$

Substituting Eq. 7 into Eq. 3, we obtain an expression for an electron's wavelength given by

$$\lambda = \frac{hc}{\sqrt{E_k^{\,2} + 2E_k m_e c^2}} \quad . \qquad (8)$$

The low energy, nonrelativistic energy regime is specified by the condition

$$E_k << m_e c^2$$

In this case $E_k^{\,2}$ is negligible compared to $2E_k m_e c^2$ and we obtain

$$\lambda \cong \frac{hc}{\sqrt{2E_k m_e c^2}} \text{ (non-relativistic)} \qquad (9)$$

Using the constants hc=1239.852 nm eV, $2m_e c^2$=1.022 x 10^6 eV, we get

$$\lambda \cong \frac{1.239852}{\sqrt{1.022E_k}}$$

This can be rewritten in a simpler form

$$\lambda \cong \frac{1.226}{\sqrt{E_k}} \quad (\lambda \text{ in nm; } E_k \text{ in eV}) \qquad (10)$$

From this analysis, electrons accelerated through an electrostatic potential of 100 V should have a wavelength of approximately 0.1 nm. To observe interference and diffraction effects, this wave must encounter a slit or aperture having dimensions comparable to its wavelength, i.e ~0.1 nm. In 1923, it was difficult to fathom how anyone could fabricate such a small slit. DeBroglie's ideas were viewed with interest, but were thought to be impossible to confirm by experiment.

The situation quickly changed in 1926 when Clinton Davisson and Lester Germer working at the Bell Telephone Laboratories began studying the interaction of an electron beam with different metallic foils. Similar experiments were underway in England by G.P. Thomson. Quite by accident, Davisson and Germer noticed that when a nickel foil was heated and annealed, electrons having an energy ~ 55 eV were scattered in preferential directions. Upon closer examination, the preferred directions of scattered electrons resembled scattering from a diffraction grating. They realized that the lattice of Ni atoms, formed during the annealing process, must act as a

diffraction grating that effectively scatters electrons. This led Davisson and Germer to conclude that electrons do indeed have wave properties.

In 1929, DeBroglie received the Noble Prize in Physics for his revolutionary idea. In 1937, Davisson, Germer, and Thomson were also awarded the Noble Prize for their pioneering experimental work. The idea that electrons can behave as a wave is pervasive throughout contemporary scientific work and forms the underlying basis of our modern understanding of electrons in solids.

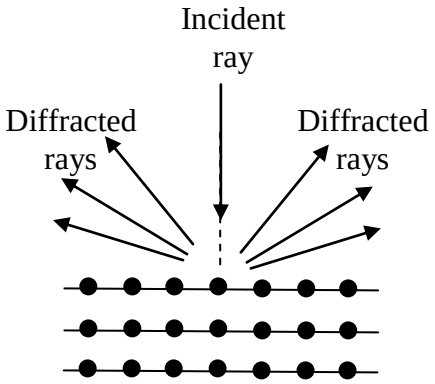


Figure 1: A schematic diagram showing the configuration of an electron diffraction experiment. At low incident energies, electrons interact with the uppermost plane of atoms and are diffracted in the backward direction as shown. As the energy is increased, transmission through the solid increases and diffraction in the forward direction becomes important (not shown).

Diffraction Physics

A crystalline solid is made up of regularly spaced atoms in three dimensions. Each atom acts a scattering center for an incident electron wave having a wavelength comparable to the inter-atomic spacing (see Fig. 1). At low energies (the case used by Davisson and Germer), the electron beam interacts only with atoms on the surface plane. The spacing between atoms on the surface is then analogous to the spacing between slits in an optical diffraction grating. Using the standard arguments for optical interference phenomena, the electron beam scattered from each surface atom must constructively interfere to produce a maximum in the scattered intensity in certain directions. The condition for constructive interference is that the path difference from adjacent atoms is an integer number of wavelengths: $n\lambda$. The path difference between adjacent atoms as a function of the scattering angle φ can be calculated from Fig. 2 (the length of the heavy line segment).

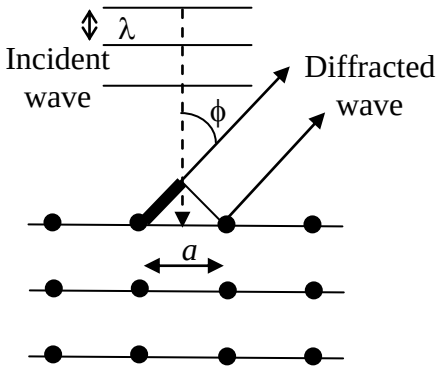


Figure 2: The geometry required to predict the diffraction angles φ for low energy electron diffraction.

The equation for intensity maxima is then

$$n\lambda = a \sin \phi. \qquad (11)$$

Davisson and Germer used 54 eV electrons and found a scattering maximum at an angle of about 50^o. This was the first peak in the scattered intensity, so it

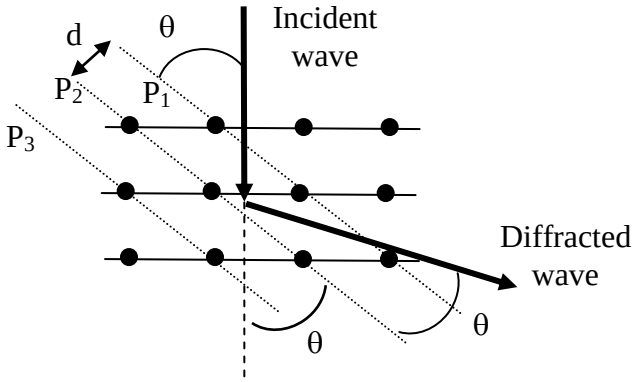


Figure 3: At higher energies, electrons penetrate further into the crystal lattice and diffraction from an arbitrary set of planes P_1, P_2, P_3, . . becomes possible. Note that the diffracted wave emerges at an angle of 2θ with respect to the undiffracted beam.

corresponded to the n=1 condition in Eq. 11. Using these values, it is easy to calculate that *a*=0.215 nm. From the known crystal structure of Ni, it was also known that in certain directions, the distance between Ni atoms is also 0.215 nm. The regime when electron scattering from the first plane on the surface of a crystal is the dominant contribution to diffraction is now known as Low Energy Electron Diffraction (LEED) and has become an important technique for determining the position of atoms on surfaces.

At higher electron energies, the situation is somewhat different. Now the electron beam can penetrate an appreciable distance into the substrate and interacts with atoms forming families of planes (see Fig. 3). If an electron is specularly scattered from a sequence of parallel planes, a constructive interference effect can arise along a certain direction. As a consequence, a physical electron will be scattered in that direction.

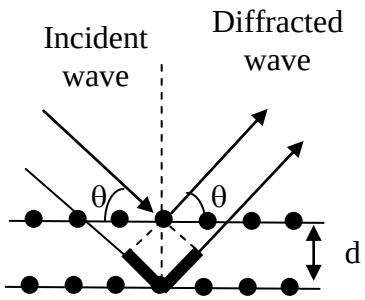


Figure 4: The geometry required to calculate the extra path length travelled by a wave scattered from the 2nd plane.

To produce a constructive interference, the scattering from separate planes separated by a distance d must reinforce each other. The condition for constructive interference is that the path difference from adjacent planes is an integer number of wavelengths $n\lambda$. The geometry required to calculate the path difference between adjacent planes is sketched in Fig. 4. From this figure, the length of the heavy line segments located between the 1st and 2nd planes represents the extra path length traveled and is given by 2dsin θ.

This argument, proposed by W.L. Bragg in 1913 to explain X-ray diffraction, is far from rigorous but it surprisingly is reproduced by more complete arguments based on electron scattering. The Bragg equation for intensity maxima (i.e. efficient specular reflection) is then

$$n\lambda = 2\, d \sin\theta \qquad (12)$$

This result is a simple consequence of the interference condition applied to this geometry. Note that since $\sin\theta \leq 1$, $\lambda \leq 2d$. This condition sets a range for the acceptable wavelengths to observe Bragg diffraction.

The Geometry of Periodic Crystals: Miller Indices

As can be inferred from Fig. 3, the value of *d* for a three dimensional lattice depends on the orientation of the lattice with respect to the incident beam. A systematic way to address this very general problem is required. For a three-dimensional lattice, the inter-planar spacing *d* for a particular set of planes may be calculated from the unit cell of the crystal and the Miller indices of the plane.

To understand Miller indices, consider the simple case of a cubic lattice of atoms sketched in Fig. 5. All space can be filled by translating this unit cell of atoms throughout all space. The lattice constant "*a*" is the length of the edge of the cube. The unit cell may be described by three lattice vectors: $\vec{a} = a\ \mathbf{i}$, $\vec{b} = a\ \mathbf{j}$, $\vec{c} = a\ \mathbf{k}$. The symbols **i**, **j** and **k** are orthogonal unit vectors which are used to define a right handed orthogonal coordinate system.

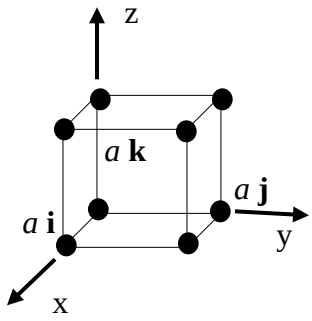


Figure 5: A simple cubic lattice illustrating the position of atoms on each corner of the cube.

Now consider a plane that cuts through the cube as illustrated in Fig. 6. Suppose the plane drawn cuts the lattice vectors at ½ $\vec{a}$, 2/3 $\vec{b}$ and 1 $\vec{c}$. The reciprocal of these intercepts are 2, 3/2, and 1. The smallest set of integers having the same ratios is 4, 3, and 2. This set of three integer numbers is known as the Miller indices and are written in the form (432).

For the general case, the Miller indices for an arbitrary plane is a set of integer numbers designated by (*hkℓ*). In summary, the general rule for finding the Miller index of a plane is:

1. Find the intersection of the plane with the lattice vectors.
2. Take the reciprocals of the intercepts.
3. Find the smallest set of integers with the same ratios.

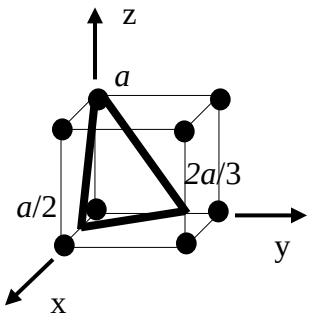


Figure 6: An arbitray plane passing through a simple cubic lattice.The Miller indices for the plane shown are (432).

Some common examples of low index planes are sketched in Figure 7.

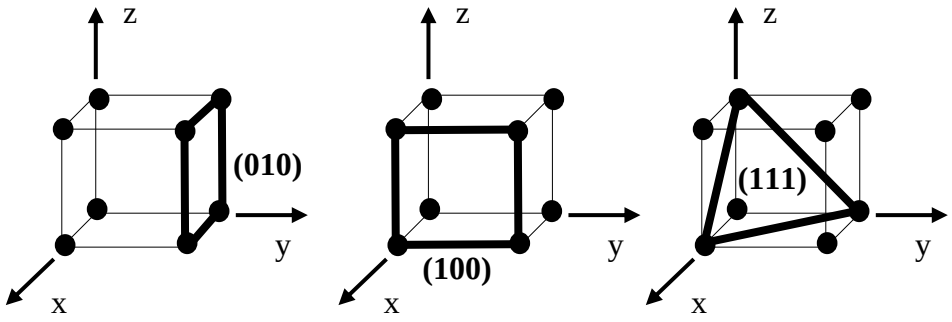


Figure 7: A few low index planes indicated by the solid lines and their respective Miller index designation.

The General Condition for Diffraction in Cubic Crystals

For a cubic lattice, you can show that the distance between adjacent parallel planes having Miller index ($hk\ell$) is given by

$$d_{hk\ell} = \frac{a}{\sqrt{h^2 + k^2 + l^2}} \qquad (13)$$

Substitution of this result into the Bragg condition (Eq. 12) gives

$$n\lambda = 2d_{hk\ell} \sin\theta = \frac{2a \sin\theta}{\sqrt{h^2 + k^2 + l^2}} \qquad \text{or,}$$

$$2a \sin\theta = n\lambda \sqrt{h^2 + k^2 + l^2} \quad . \qquad (14)$$

This equation forms the basis for analyzing diffraction data acquired below.

Crystal Structures

Fourteen distinct crystal lattices are allowed by symmetry to fill all space in a periodic way. The simplest lattices are the three cubic lattices. In all three classes the cell is cubic and all edges have equal lengths.

The simple cubic (sc) cell has one atom at each of the cube's eight corners (see Fig. 5). The body-centered cubic (bcc) cell has one atom at each corner, plus one atom in the center of the cubical volume. The face-centered cubic (fcc) cell has one atom at each corner plus one atom in the center of each of the six faces of the cube (See Fig. 8). There are

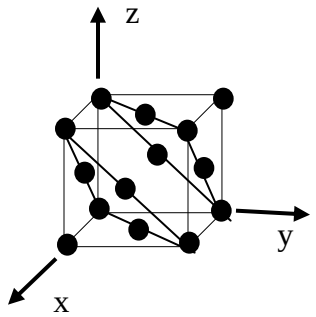


Figure 8: Placement of atoms in an fcc lattice

many elements with fcc unit cell structure. Examples are: Al, Ca, Cu, Ag and Au. These elements show similar diffraction patterns.

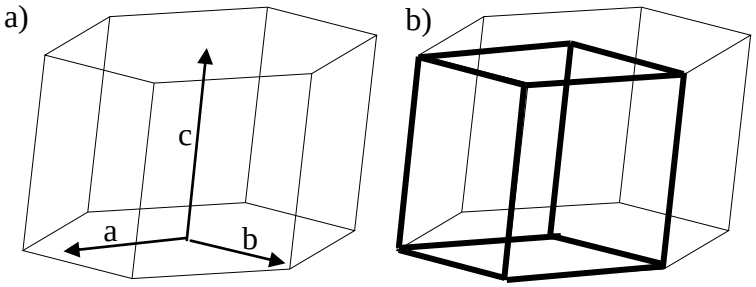


Figure 9: In a), a sketch of the hexagonal system with the lattice vectors defined. In b), the primitive hexagonal unit cell is indicated by the heavy lines.

Another common lattice type has hexagonal symmetry. The simple hexagonal cell has atoms at each of 12 corners and at the center of each hexagon. The length of the sides of the two hexagons are equal, but not equal to the height, denoted by "c", which is the distance between the hexagons (see Fig 9(a)). The smallest unit cell that can be translated to fill all space is a part of the hexagonal cell as shown in Fig. 9(b). There are a variety of different classes of unit cells having hexagonal symmetry. A form of carbon (graphite), Mg, Zn and Cd are examples of materials possessing hexagonal symmetry.

The fcc Selection Rule for Diffraction

The intensity of reflection from a particular set of crystal planes depends on a detailed theory of the diffraction process. In general, planes with low density of atoms (large ($hk\ell$)) will have a small scattering amplitude. Thus planes with low ($hk\ell$) will diffract more strongly. One advantage of defining Miller indices is that a rigorous diffraction theory can be cast in such a form that the intensity scattered by a particular set of lattice planes is related to the Miller indices of the plane. For an fcc crystal like aluminum, you can show that the reflected intensity is proportional to a geometrical structure factor S, which is defined as:

$$S=1+e^{i\pi(h+k)}+e^{i\pi(h+\ell)}+e^{i\pi(k+\ell)} \qquad (15)$$

This factor is **non zero** only if all three indexes are either even or all three indexes are odd. Zero is considered to be an even number.

EXPERIMENTAL APPARATUS AND PROCEDURE

Apparatus

For high energy electron diffraction, a beam of electrons strikes a target located inside a vacuum tube or vacuum chamber. The electrons diffracted by the target are observed when they strike a phosphor screen located on the opposite side of the target (see Fig. 10).

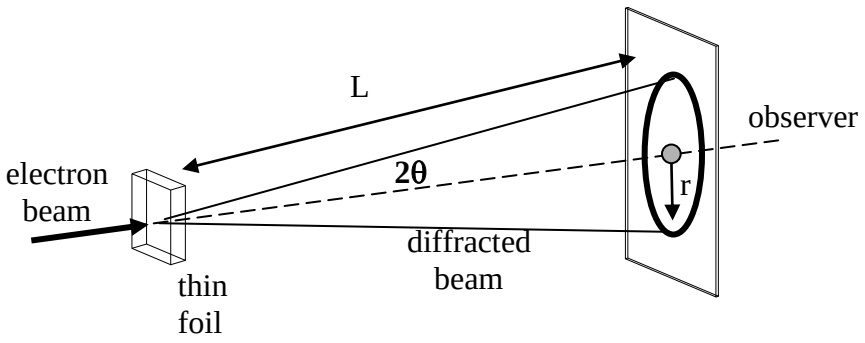


Figure 10: The geometry relevant to a transmission electron diffraction experiment.

For the apparatus you will use, the target contains both polycrystalline aluminum and graphite samples. The desired material can be selected by moving the electron beam with the controls provided for this purpose. A microammeter should be attached to the terminals at the back of the Sargent-Welch apparatus. This meter measures the electron beam current. Look on the left hand side of the tube's housing to find the value of *L*.

It is convenient to analyze the diffraction pattern by measuring the radius formed when the diffracted beam strikes the fluorescent screen. The radius *r* of the diffraction pattern can be measured with a caliper. Since θ will be a small angle, we can replace sinθ with θ in Eq. 14. Also note that the angle between the incident and diffracted beam is 2θ as shown in Fig. 3. This fact is also indicated in Fig. 10. For small θ, tan2θ ≅ 2θ = r/L. This means that Eq. 14 can be rewritten as

$$\frac{a\, r_{hk\ell}}{L} = \lambda\sqrt{h^2+k^2+\ell^2} \qquad (16)$$

where $r_{hk\ell}$ is the radius produced by scattering from the ($hk\ell$) plane. In our case we observe only the first order of diffraction, so n=1.

General Experimental Procedure

Set the accelerating voltage of tube at 6 kV, keep the beam current as low as possible. Never let the beam current exceed 10 μ*A*. Focus the beam and move it using the horizontal and vertical controls until you obtain a pattern of rings without bright spots. Defocus the beam to see the shadow of the target. Look through the window at the side of the box to see the target inside the tube. Warning: Do not let the undiffracted beam spot sit still on the screen or you will burn the screen!

Refocus the beam. Measure the diameters of the rings with a clear plastic ruler or a caliper and record the results. The rings may be slightly distorted by the tube, so you should measure and record at three different orientations the diameter of each ring and use the average in your calculations.

Repeat your measurements for four other voltages at 7, 8, 9 and 10 kV. The electron beam will shift when you change the voltage, so you will have to bring it back to its original location or at least to another place where aluminum rings may be found.

Characterizing Crystalline Materials

It is important to have a general understanding of the crystalline make-up of materials before taking diffraction data. In general terms, a particular piece of material can be either amorphous or crystalline. Amorphous materials are characterized by the

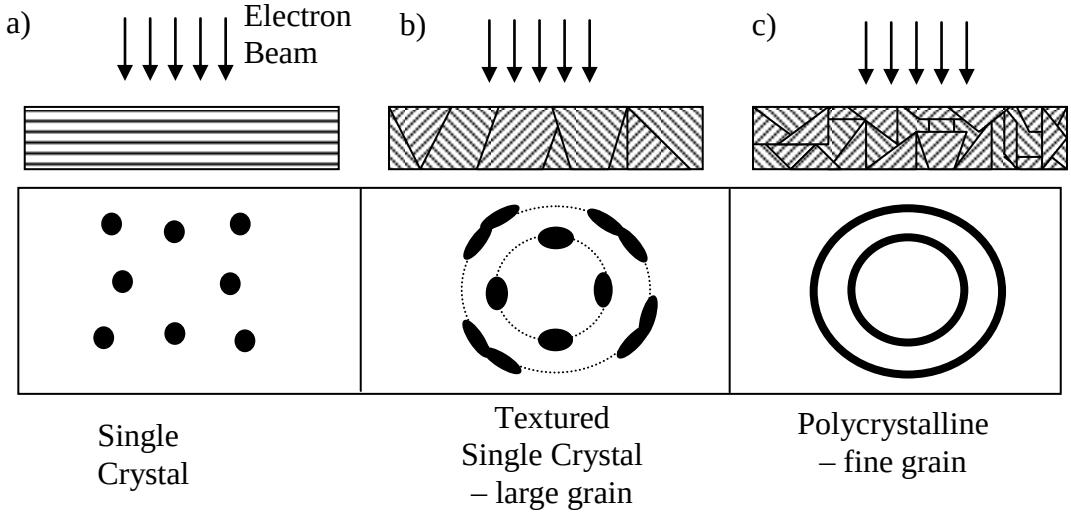


Figure 11: Examples of varying degreses of crystallinity and the resulting diffraction patterns. In a), a perfect single crystal. The diffraction pattern is characterized by well defined spots. In b), a large grain textured sngle crystalline material. The diffraction spotsa begin to broaden and streak due to different orientations of individual grains. In c), a fine grained polycrystalline material. The individual diffraction spots produced by one grain now merge into complete circles, reflecting the randomn orientation of one grain with respect to any other.

absence of long-range periodicity in atom position and are not of interest in this lab.

Crystalline materials can be categorized in terms of their long-range periodicity as indicated in Fig. 11. The most crystalline material is characterized by a perfectly periodic arrangement of atoms throughout. Such a material is said to be a single crystal. Single crystals require great care in growth and can be tens of centimeters in size. Single crystals are easy to damage and their crystalline state can be substantially degraded. Materials lacking long-range periodicity are classified as crystalline with a textured or multiple-domain structure. Such materials may have a high degree of periodicity extending over a characteristic distance – say a few millimeters. Beyond this, another domain is encountered which itself is highly crystalline, but with an arbitrary orientation with respect to adjacent domains. These domains (or grains) tend

to have a characteristic size that is replicated throughout the material. If the size of the single crystal domains is small enough, then the material is said to be polycrystalline. In a polycrystalline material, the single crystal domain or grain size may only be 1 to 10 microns in size.

In the experiments below, you will examine a textured single crystal of graphite and a polycrystalline sample of aluminum.

PYROLITIC GRAPHITE TARGET

Diffraction from a Textured Single Crystal

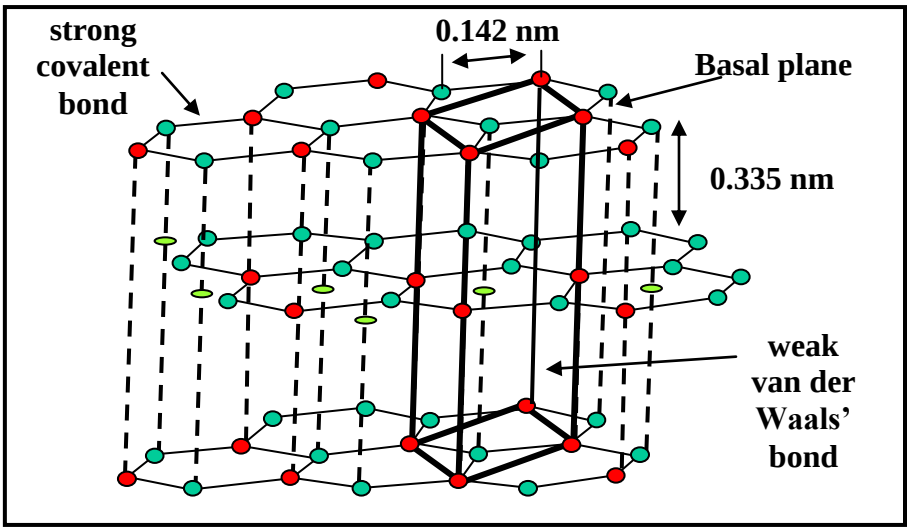


Figure 12: The crystal structure of pyrolitic graphite. The small dots represent the position of C atoms that are spaced by 0.142 nm from each other on an hexagonal grid. The inter-sheet spacing betweeen graphene planes is 0.335 nm. The heavy lines delineate the unit cell which has dimensions of a=b= 0.246 nm and c = 0.670 nm. The first and thrid graphene layers are aligned, but the second layer is laterally translated (dotted lines) to form what is known as an ABAB . . stacking sequence.

The crystal structure of graphite belongs to the hexagonal class of symmetries (see Fig. 9). In graphite, carbon atoms form a hexagonal pattern in a basal plane that is known as a graphene sheet. The individual graphene sheets are then stacked one on top the other as shown in Fig. 12 to form a single crystal of graphite. Pyrolitic graphite is a textured form of single crystal graphite with a strongly preferred orientation. Its structure is comprised of domains or grains of single crystal graphite. Each domain is randomly rotated about the $\vec{c}$ axis and each grain is oriented with a strong preference for parallel $\vec{c}$ axes.

Since the penetration of 6 keV electrons is on the order of a few hundred nanometers, an extremely thin graphite target is required. If the single crystal graphite domains

are sufficiently large, it may be possible for the electron beam to strike a part of the target that is one single crystal. When this situation is met, you will obtain the best diffraction pattern. On the screen you will then see a pattern something like Fig. 13.

The diffraction pattern of pyrolitic graphite is hexagonal, but rotated 30° with respect to the original lattice. The inner most hexagon corresponds to n=1, the next hexagon to n=2, etc. Again, diffraction will occur when $n\lambda = 2d \sin\theta$.

For diffraction through small angles, $2\sin\theta \cong 2\theta \cong r/L$.

Thus the separation between the planes responsible for diffraction will be given by

$$d = \frac{n\lambda L}{r} \qquad (16)$$

where $r=D/2$ is the distance from the center of the pattern to a point on the hexagon. If you see several spotty rings instead of a pattern like Fig. 13, the electron beam is not striking a single crystal but rather several crystals. With a little patience, a single crystal can be found.

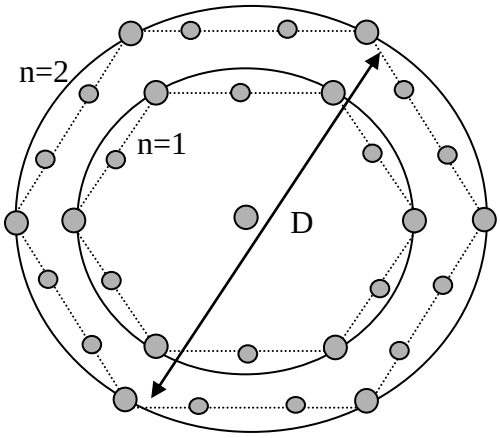


Figure 13: A schematic digram of the expected diffraction pattern from a single crystalline region of the pyrolitic graphite target. The spots on the corner of the hexagon form the dominant contribution to the diffracted beam.

To begin, set the tube voltage at 6 keV. Keeping the beam current below 10 μA, search for the graphite pattern. Record the diameter D (three times with about a 120^{o} degree between the measured lines) and the order n of each concentric hexagonal pattern. Repeat your measurements for four more voltages of 7, 8, 9 and 10 kV.

Using Eq. 16, calculate the value of the plane separation d for each order of diffraction which appear at all five voltages you used. Without a more thorough analysis, it is not obvious how the values of d are related to the lattice constants of graphite which are specified by a=0.24612 nm, c= 0.6707 nm. From the placement of atoms in the unit cell, it is easy to estimate a few inter-atomic distances. These can be compared to your data to learn which planes are responsible for the diffracted spots. Estimate your errors in d, collect your data and results in a table. Be prepared to discuss your findings.

ALUMINUM TARGET

Diffraction from a Polycrystalline Material

The crystal structure of single crystal aluminum is face-centered cubic. The unit cell is a cube with an atom at each corner and also at the center of each face, as in Fig. 8. In the case of polycrystalline aluminum, many individual single crystals (ranging in size from a few microns to a few 100 microns) are randomly oriented with respect to the electron beam. Because of the large number of crystals illuminated, many of the crystallites will be oriented so as to satisfy the Bragg diffraction equation. The diffraction pattern from any given crystallite will be a pattern of spots, but because all orientations are possible, the individual diffraction spots will merge into a cone of diffracted electrons. The result is that a pattern of concentric rings will be seen on the screen.

Set the voltage to 6 kV and examine the diffraction pattern from the aluminum target. The innermost ring (smallest $r_{hk\ell}$) will correspond to the orientation with the smallest value of $h^2+k^2+\ell^2$, which turns out to be (111). The sequence of rings continues with (200), (220), (311), (222), (400), (331), (420), (422), and so forth. Some rings may appear to be "double", but these are really separate rings due to two slightly different orientations.

For the fcc unit cell in Fig. 8, note the presence of planes parallel to the faces of the cube and halfway between them. The electron waves reflected from the opposite cube faces will be out of phase with the waves reflected from the intermediate plane between these faces, and therefore no diffraction will be seen from the (100), (010), or (001) directions. As discussed above, no diffraction will occur unless h,k, and ℓ are either all odd or all even.

Repeat your measurements for four more voltages of 7, 8, 9 and 10 kV. Calculate the lattice constant "a" for polycrystalline aluminum using Eq. 16 and the value of λ calculated from the accelerating voltage of the tube (Eq. 10).

State which crystal plane is associated with each ring. Calculate the value of "a" for each voltage and each ring and compare with the accepted value of a=0.405 nm obtained from X-ray diffraction measurements. Estimate the size of your errors. Collect your data and results in Table 1 and discuss your findings.

Rings with different sets of Miller Indices will show up with different intensities or brightness. There is no simple theory that predicts which ring associated with an allowed set of Miller Indices will be visible, or how many rings will be visible. Since the electron diffraction pattern obtained in this experiment is very similar to the pattern produced by X-ray power diffraction and by Laue single crystal diffraction, you should look up information on X-ray diffraction patterns on polycrystalline aluminum. A helpful example is given in the discussion below and in Table 1.

The information in Table 1 shows you that the first four lowest ($hk\ell$) combinations should give you four clearly visible rings. It is not clear whether you will be able to observe the reflections which correspond to the combinations (222) or (400). Since you will see at least two more faint rings and you do not know what $hk\ell$ combinations they correspond to, try several allowed combinations of ($hk\ell$) until you obtain values for "a" which are consistent with what you obtained from the four innermost rings.

The relative scattered X-ray intensity pattern given in Table 1 is from the X-ray Diffraction Information Handbook (compilations). The source was a λ=0.154056 nm wavelength $CuK\alpha_1$ X-ray with a Ni filter to absorb background radiation. The measured lattice constant "a" was *a*=0.404958 nm. A chemical analysis of the target revealed a 99.9% aluminum purity. Background metals were Si, Cu, Fe, Ti, Zr, Ga, S with relative concentrations less then 0.007% each. Iron contamination was the largest at 0.007%.

Table 1: Using λ=0.154056 nm wavelength X-ray radiation the following relative intensities were obtained

Miller Indices	Intensity of Ring	Lattice Constant
$hk\ell$	Int.	*a*(nm)
111	100	
200	47	
220	22	
311	24	
222	7	
400	2	
331	8	
420	8	
422	8	

SWITCHING OFF THE EQUIPMENT

When you are finished, it is important that you turn the High Voltage down before turning it off. This allows the H.V. capacitors in the power supply to discharge, thus avoiding damage to the tube and power supply.

ADDITIONAL QUESTIONS

The relativistic kinetic energy is defined as the difference between the total energy and the energy of the rest mass (m_0).

$$E_K = E_{total} - m_o c^2 = \frac{m_o c^2}{\sqrt{1-\gamma^2}} - m_o c^2 \quad with\ \gamma \equiv \left[\frac{v}{c}\right] .$$

A particle is moving with relativistic velocity if $E_K \geq m_o c^2$. Using this condition and the above definition of E_K, determine the value of v/c at the onset of the relativistic domain i.e when $E_k = m_0 c^2$. Is the velocity of a 10 keV electron, in your experiment, relativistic?
Hint: set $E_K = E_{total} - m_o c^2 = 10$ keV.

Why are the target, screen, and conducting coating on the tube connected to ground?

Can particles other than electrons be diffracted? If so, give examples.

State clearly how these experiments demonstrate that the wave picture of an electron is valid.

If a conduction electron, already inside a metal, travels through a periodic array of atoms would it also be forced to undergo diffraction? What consequences might this have?

In deriving Eq. 16, tan 2θ was set equal to r/L. Is this approximation valid for the geometry of the tube you are using? How do you justify the setting of the scattering angle in Fig. 10 equal to 2θ instead of θ?

NOTES