Physics 53

Electromagnetism & Optics

Laboratory Course Manual

Fall 2010

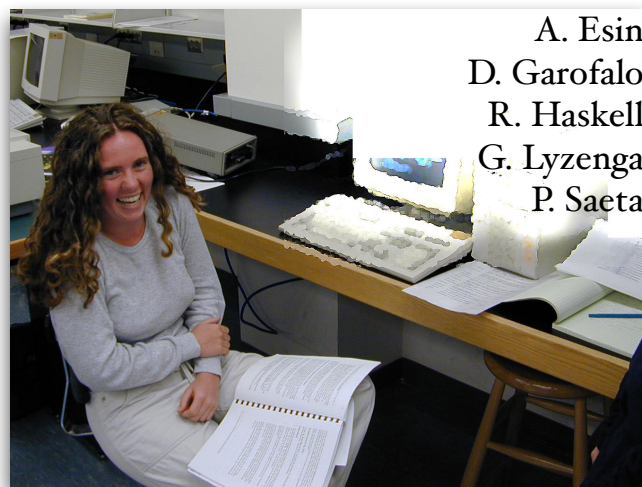### Instructors

E. Connolly
J. Eckert
A. Edwards
A. Esin
D. Garofalo
R. Haskell
G. Lyzenga
P. Saeta

"Nothing in the world can take the place of persistence. Talent will not; nothing is more common than unsuccessful men with talent. Genius will not; unrewarded genius is almost a proverb. Education will not; the world is full of educated derelicts. Persistence and determination alone are omnipotent."

Calvin Coolidge

Contents

Part I

Introductory Material

Quick Start Guide for First Meetings

The following is a brief summary of the essential information you'll need to prepare for the first meetings of E&M Lab. Further details will be found in the Overview section and in other resources provided to you by the instructors.

First Meeting

The first meetings of each lab section will be during the week of Aug. 30 (except for Monday sections, which will first meet the following week). The sections will then meet on succeeding weeks for two hours throughout the semester, following the tentative schedule of meetings and breaks on the following page. Although it is not heavy, there is some preparation for the first meeting. Prior to coming to lab, **you must read the lab manual introduction and description for Experiment 1 (pp. i–xiii, 2–5)**. Also come to lab prepared with a brown cover laboratory notebook of the approved type: either **Ampad #22-157 Computation Book** (preferred) or National Brand #43-648 Computation Notebook. One or both are available from Huntley bookstore.

Subsequent Meetings

After the first two weeks, you will perform four of the five available experiments according to a schedule to be worked out with your instructor. With your instructor's consultation, you will sign up for the experiments and lab partners you will work with during the semester.

The experiments in this course are challenging, and require advance reading and preparation. It is very important that you come to the lab prepared. Before the first laboratory meeting for a particular experiment, **watch the short movie** introducing the experiment you will be doing and read the appropriate section of this lab manual. The movies are available in both Quicktime and video podcast format at the course web site: http://www.physics.hmc.edu/courses/p053/. We will not require pre-lab homework or exercises; however, part of your grade will be based on your instructor's assessment of how prepared you are to conduct each experiment.

Lab periods run for 2 hours every meeting, and each experiment will take two lab periods. The first week (meeting "a") will typically be useful for familiarizing yourself with and calibrating the apparatus and making a first set of measurements; the second week (meeting "b") will normally be used for completion of data taking, recovery from any problems or mistakes in the first week, and consultation with your instructor about analysis and write-up.

Write-up of the Experiment

A typical experiment summary is 2 typed pages, turned in along with your lab notebook for documentary support, at a time and place to be announced by each instructor. All write-ups should be typed/word processed, with all graphics neatly and professionally incorporated. An example is available on the course web site. The experiment summary will be what is primarily read and marked by your instructor. The lab book may not be graded in detail, however the degree to which it serves as a valid reproducible archival record of your work in lab will form a part of your grade. In the write-up, you may assume your reader is familiar with your experiment, the apparatus, and the basic procedures, so re-hashing of these is not needed. However, you should detail innovations, important observations, or unique insights that you have made during the experiment period. *The write-up should be entirely derived from the content of your laboratory notebook, and not from memory or other undocumented sources.* See the rest of this manual for more detailed discussion of write-ups and lab books.

Schedule

Following is a *tentative* schedule for the semester. Your instructor may choose to shift meetings around; please note any changes that your instructor makes.

Key: Mtg. 3b is the second lab session ('b') of the third experiment a student does (not to be confused with Experiment 3 in the manual).

	Week of	Section 1 M 12:40 Lyzenga	Section 3 M 18:00 Eckert	Section 5 Tu 15:15 Connolly	Section 7 W 15:15 Garofalo	Section 9 Th 12:40 Esin	Section 11 F 12:40 Edwards
1	8/30			Expt. 1a	Expt. 1a	Expt. 1a	Expt. 1a
2	9/6	Expt. 1	Expt. 1a	Expt. 1b	Expt. 1b	Expt. 1b	Expt. 1b
3	9/13	Expt. 1	Expt. 1b	Mtg. 2a	Mtg. 2a	Mtg. 2a	Mtg. 2a
4	9/20	Mtg. 2a	Mtg. 2a	Mtg. 2b	Mtg. 2b	Mtg. 2b	Mtg. 2b
5	9/27	Mtg. 2b	Mtg. 2b	Mtg. 3a	Mtg. 3a	Mtg. 3a	Mtg. 3a
6	10/4	Mtg. 3a	Mtg. 3a	Mtg. 3b	Mtg. 3b	Mtg. 3b	Mtg. 3b
7	10/11	Mtg. 3b	Mtg. 3b	Mtg. 4a	Mtg. 4a	Mtg. 4a	*Open*
8	10/18				Mtg. 4b	*Open*	Mtg. 4a
9	10/25	Mtg. 4a	Mtg. 4a	Mtg. 4b	*Orals*	Mtg. 4b	Mtg. 4b
10	11/1	Mtg. 4b	Mtg. 4b	Mtg. 5a	Mtg. 5a	Mtg. 5a	Mtg. 5a
11	11/8	Mtg. 5a	Mtg. 5a	Mtg. 5b	Mtg. 5b	Mtg. 5b	Mtg. 5b
12	11/15	Mtg. 5b	Mtg. 5b	*Report Prep*	*Report Prep*	*Report Prep*	*Report Prep*
13	11/22	*Report Prep*	*Report Prep*	*Open*	*Open*		
14	11/29	*Open*	*Open*	*Open*	*Open*	*Open*	*Open*
15	12/6	*Makeup Week*					

	Week of	Section 2 M 15:15 Eckert	Section 4 Tu 12:40 Connolly	Section 6 W 12:40 Haskell	Section 8 W 18:00 Garofalo	Section 10 Th 15:15 Esin	Section 12 F 15:15 Saeta
1	8/30		Expt. 1a	Expt. 1a	Expt. 1a	Expt. 1	Expt. 1
2	9/6	Expt. 1a	Expt. 1b	Expt. 1b	Expt. 1b	Expt. 1	Expt. 1
3	9/13	Expt. 1b	Mtg. 2a	Mtg. 2a	Mtg. 2a	Mtg. 2a	Mtg. 2a
4	9/20	Mtg. 2a	Mtg. 2b	Mtg. 2b	Mtg. 2b	Mtg. 2b	Mtg. 2b
5	9/27	Mtg. 2b	Mtg. 3a	Mtg. 3a	Mtg. 3a	Mtg. 3a	Mtg. 3a
6	10/4	Mtg. 3a	Mtg. 3b	Mtg. 3b	Mtg. 3b	Mtg. 3b	Mtg. 3b
7	10/11	Mtg. 3b	Mtg. 4a	Mtg. 4a	Mtg. 4a	Mtg. 4a	*Open*
8	10/18			*Open*	*Open*	*Open*	Mtg. 4a
9	10/25	Mtg. 4a	Mtg. 4b	Mtg. 4b	Mtg. 4b	Mtg. 4b	Mtg. 4b
10	11/1	Mtg. 4b	Mtg. 5a	Mtg. 5a	Mtg. 5a	Mtg. 5a	Mtg. 5a
11	11/8	Mtg. 5a	Mtg. 5b	Mtg. 5b	Mtg. 5b	Mtg. 5b	Mtg. 5b
12	11/15	Mtg. 5b	*Report Prep*	*Report Prep*	*Report Prep*	*Report Prep*	*Report Prep*
13	11/22	*Report Prep*	*Open*	*Open*	*Open*		
14	11/29	*Open*	*Open*	*Open*	*Open*	*Open*	*Open*
15	12/6	*Makeup Week*					

Overview
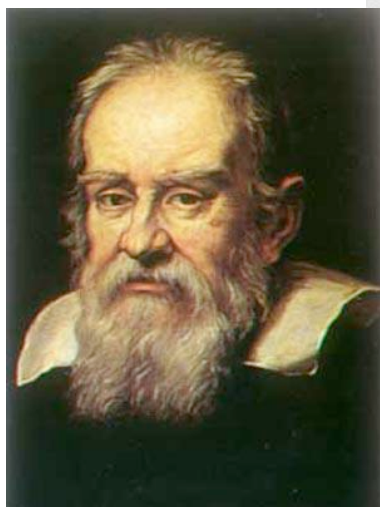
Eppur si muove (but it does move...)

Galileo Galilei, 1564–1642

Our understanding of the physical universe derives from comparing honest, careful observations to theoretical models. The principal goal of this course is for you to learn how to conduct experiments:

1. to figure out just what you need to do to obtain the necessary high-quality data

2. to make careful observations

3. to record those observations accurately

4. to compare them to a theory

5. to communicate the observations, results, and conclusions clearly and succinctly

None of these is as easy as it may appear at first glance. Your success in the course will be determined by the care and thought you demonstrate in these five aspects of conducting experiments.

1. What do I need to do?

A great chef uses a recipe as a source of ideas, as a suggestion about how to accomplish a goal; she or he doesn't necessarily follow the recipe meticulously, the way a lesser cook might. It would be possible for us to spell out in excruciating detail each step you should follow in conducting the experiments of this course, but we choose not to. We are uninterested in educating lesser cooks who are bound merely to follow the recipes developed by others. Such an approach gives an impoverished view of experimental science. Rather, we want you to understand how to conduct original research, which involves experimental design in addition to careful observations and analysis. We won't be answering for you questions such as

- "How many data points do I need?"

- "Do I need to measure both doublet lines, or is one enough?"

- "Should I measure the resistance of the inductor?"

- "Can you set up this circuit for me?"

even though this may frustrate you at times. Learning how to answer such questions is an important part of understanding what it means to conduct an original experiment.

2. Fiddling

Never trust your equipment. Rather, assume that Murphy's law applies: *Anything that can go wrong will go wrong.* Check out your equipment! Make sure you understand how it works, and how well it works. When you use a multimeter to measure resistance, for example, what value does it report when you simply connect the leads together? Does the scale read zero when nothing is on it? What does it read when a 50-g test mass

sits in the pan? Does the detector signal change when you block the laser beam?

Before you begin to take data for any experiment, play with the equipment to make sure you understand how it behaves.

3. Document

The second principal task is to record what you do. Not what you should have done, or thought you might do, or what you think you did yesterday. What you *actually did*, and as contemporaneously as practicable. Every scientist can tell you stories about efforts wasted because he or she failed to keep adequate notes on what went on in the lab, from hours, to days, to months or even more. Our memories fade annoyingly quickly, but it is critical to know whether *these* data were taken before or after *that* setting was changed. **Document your work as you go**.

- **Sketches and figures are good**. Use them to clarify (and sometimes avoid) discussion, define symbols, explain equations, etc. A picture speaks 10^3 words.

- **Tables are good**. Almost all measurements should be repeated, and the results collected in a table. Figure out what columns you'll need, head them and indicate units and uncertainties (or put uncertainties in their own column).

- **Lists are good**. If you have a multi-part procedure, a bulleted or numbered list is easier to figure out than such glowing phrases as "Then we turned the *xxx* on and ...", "Then we fiddled with the alignment of the *xxx* until it was good ...", "Then ..." A list saves time and is easier to read.

- **Algebra is good**. Using numbers without associated algebraic explanation is **evil**.

- **Units are good**. Failing to put units on a value is **evil**.

- Set off key results and formulas with boxes.

4. Analyze

An experiment is about **understanding**. The more understanding you can develop *while* you are conducting the experiment, the better your experiment will be. Although it is tempting to lapse into "data-taking mode" — mindlessly writing down measurements — don't! Instead, ask yourself if your data are making sense as you take them.

For this, a spreadsheet and plot are invaluable. Use one of the lab computers or your own notebook computer to allow you to monitor data quality as you go, or plot by hand in your notebook. If you use a computer to record your data, print the data table out and tape it into your notebook. If you finish a section of the experiment and have produced a graph, print it out and tape it into the notebook, as well. Annotate the graph before moving on: what conclusion(s) can you draw from the graph?

Uncertainties

Most measurements are not infinitely precise, either because you use instruments of imperfect accuracy to make the measurement or because the precise position or configuration you are attempting to quantify is difficult to determine. Coming up with a highly accurate measurement of your height, for example, may pose a number of difficulties: Standing up straight or slouching? With or without accounting for hair? Looking straight ahead, or slightly down?

One way to estimate the uncertainty of a measurement of a quantity y is to measure y several times and compute the mean ($\bar{y}$), (sample) standard deviation (σ), and standard deviation of the mean ($\sigma_{\bar{y}}$) of the measurements:

$$\bar{y} = \frac{1}{N} \sum_{n=1}^{N} y_n$$

$$\sigma_y = \left[\frac{1}{N-1} \sum_{n=1}^{N} (y_n - \bar{y})^2 \right]^{1/2}$$

$$\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{N}}$$

Scientific calculators typically employ the definition given here for σ_y; you can check yours by using it to compute the standard deviation of the values $y_n = \{1, 2, 3, 4, 5\}$. If you get $\sigma_y \approx 1.581$, it does. Having computed the mean and standard deviation of the mean, $\sigma_{\bar{y}}$, (also known as the **standard error**), you will then typically assume that your determined value is $\bar{y} \pm \sigma_{\bar{y}}$. That is, the uncertainty in the measurement is the standard deviation of the mean, $\delta y = \sigma_{\bar{y}}$. Roughly speaking, the standard deviation of the mean represents the 68% confidence interval: 68% of the time, the *true* value of y will lie in the range $\bar{y} - \sigma_{\bar{y}} \leq y_{\text{true}} \leq \bar{y} + \sigma_{\bar{y}}$. *These expressions are justified when the errors are random and normally distributed.* Fortunately, random errors are frequently normally distributed, which means that they have a Gaussian (bell-shaped) distribution. See Chapter 4 of Taylor for further discussion.

Sometimes you will estimate the uncertainty based on the stated accuracy of a measuring device. Usually you compare the device accuracy and the measurement standard error and take the larger, or add them in quadrature, which is a fancy way of saying that you take the square root of the sum of the squares. **It is not good enough to report a measured value; you must**

quantify its uncertainty, as well. *Please record in your notebook how you arrived at your uncertainty estimate.*

Error propagation

Suppose that you are measuring the speed of an object by measuring the time t it takes to travel a small distance d. The speed is then $v = d/t$, but what is the uncertainty in speed, δv? In this case, the desired quantity v is a function of two measured quantities, $d \pm \delta d$ and $t \pm \delta t$, namely $v(d, t) = d/t$.

Let's say that you measured the distance d with a ruler and you performed the time measurement with a stopwatch. First assume that the time measurement is much less certain than the distance measurement, so that we may neglect δd compared to δt. As illustrated in Fig. 1, the magnitude of the uncertainty in the deduced value of v for a given uncertainty δt depends on the value of t. For the comparatively small value t_1, δt produces a large uncertainty in the value of v_1, whereas for the larger value t_2, it produces a much smaller uncertainty in v_2.

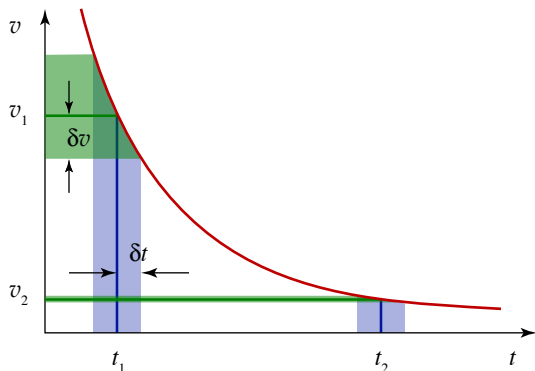Figure 1 Effect of uncertainty in a time measurement, δt, on the uncertainty in a deduced speed, δv.

In general, you may approximate the uncertainty in a derived quantity f from an uncertainty in the measured quantity x using a straight-line approximation to the function $f(x)$:

$$\delta f \approx \left| \frac{\partial f}{\partial x} \right| \delta x \tag{1}$$

Returning to our example of measuring speed, let us now assume that the uncertainty in d is significant and that the deduced value of v is uncertain both because of δt and δd. There is no reason to suspect that if you measured the distance to be higher than its true value, you are more likely to have measured a time that is higher than its true value. The errors in these measurements are **uncorrelated**. When this is the case, add the two errors in quadrature using the formula

$$\delta v = \left[\left(\frac{\partial v}{\partial d} \delta d \right)^2 + \left(\frac{\partial v}{\partial t} \delta t \right)^2 \right]^{1/2} \tag{2}$$

The extension to more uncorrelated (independent) variables is straightforward. Equations *(1) and (2) are really all you need to know to propagate errors.*

We can specialize this formula to a common case. If $f = x^m y^n$, then

$$\frac{\delta f}{f} = \sqrt{ \left(m \frac{\delta x}{x} \right)^2 + \left(n \frac{\delta y}{y} \right)^2 } \tag{3}$$

This expresses the **fractional error** in the function $f(x, y)$ in terms of the fractional error of its arguments x and y.

Example: As an example, suppose you have measured the speed of an object to an error of 3% and its mass to within 1%. What is the error in its kinetic energy?

Solution: Since $K = \frac{1}{2} m v^2$, the relative error is

$$\frac{\delta K}{K} = \sqrt{ \left(\frac{\delta m}{m} \right)^2 + \left(2 \frac{\delta v}{v} \right)^2 }$$
$$= \sqrt{(0.01)^2 + [2(0.03)]^2} = 0.0608$$

which is 6%. Thus, the uncertainty in the speed produces essentially all the uncertainty in the kinetic energy, and this relative uncertainty is double the relative uncertainty in the speed, because kinetic energy depends on the second power of speed.

For complicated expressions, taking the derivatives algebraically can be a real pain and isn't worth the effort. Using a spreadsheet or calculator, you can get a better answer with much less work by evaluating the discrete changes numerically.

Example: Suppose that you have measured the distance r between the center of a magnetic compass and the geometric center of a bar magnet, as well as the deflection angle ϕ of the compass needle and the distance b from the center of the magnet to either (fictitious) magnetic pole of the bar magnet. According to

$$\frac{M}{B_h} = \frac{2\pi}{\mu_0} r^3 \left(1 - \frac{b^2}{r^2} \right)^2 \tan \phi \tag{2.5}$$

the propagated error in M/B_h would be a rather nasty function of δr, $\delta \phi$, and δb. However, by computing the change in M/B_h caused by using $r + \delta r$ instead of r—and keeping the other parameters fixed—you can quantify the uncertainty caused by δr. Similarly, using $\phi + \delta \phi$ instead of ϕ you can quantify the change induced by an error in measuring the compass needle angle, and

similarly for δb. If you make one column in your spreadsheet for each of these shifts, you can compare them and see which one(s) cause the dominant shifts, adding them in quadrature if more than one value produces a significant shift.

You should work out error expressions *in the lab* and evaluate them numerically for a sample data point. This will allow you to determine which terms in the quadrature sum are negligible. A spreadsheet can be a great time-saver for error propagation, showing you quickly which terms dominate and allowing you to replicate the sample calculation for all your data. If you have questions, please ask your instructor.

5. Report

Following the second week of each experiment, you will type a roughly 2-page summary of the experiment containing a **concise** discussion of important points: what worked, what didn't, what you can conclude, and what you would do differently next time. Include procedure only where it departs from the directions in the lab manual. You may refer to entries and figures in your lab notebook by page, as appropriate. **Focus on your work and insights, not on a regurgitation of material in this lab manual.**

This summary should be turned in with your lab notebook write-up; it is the primary end product of your experiment. The lab notebook, while secondary, must provide documentary support for the results and claims of the summary.

All work that is handed in for credit in this course, including technical reports, is regulated by the HMC Honor Code, which is described in general terms in the student handbook. In application, this Code means simply that *all* work submitted for credit shall be your own. You should not hesitate to consult texts, the instructor, or other students for general aid in the preparation of laboratory reports. However, you must *not* transcribe another student's work without direct credit to him or her, and you must *give proper credit* for any substantial aid from outside your partnership. To reiterate, while you may discuss the experiment with your lab partner, *your analysis of the experiment should be done individually.*

Schedule

The first two weeks consist of an introduction to the laboratory, basic DC circuits and Ohm's law. After that, four experiments will be done, each scheduled for two meetings (consult your section's schedule for details). The first week (meeting "a") of each experiment will typically be devoted to understanding and calibrating the apparatus and making a first round of measurements; the second week (meeting "b") will normally be for completion of data taking and analysis. Makeup work for excused absences must be arranged with your laboratory instructor, and is generally done during the last week of the semester.

Early in the semester, your instructor will work out with you the selection of four experiments and lab partners you will work with during the semester. Although every effort will be made to accommodate individual preferences as to experiments, scheduling restrictions may prevent some students from getting all their first choices; we will appreciate your flexibility here. The usual means of evaluating your lab work will be the experiment summary, as described in the Quick Start section of this manual. This, together with the documentary support of your lab notebook, will form the bulk of your instructor's basis for grades. These lab writeups are to be prepared and turned in individually by each student. Collaboration between lab partners follows rules similar to those for homework in classes; partners may share ideas and general methods (and data, obviously), but the final product turned in by each must be his/her own work.

About once during the semester (more or less; see your professor), the written experiment summary will be replaced with an oral report presented by the lab partners as a team. The purpose is essentially the same as that of the written summaries, but presented in an oral presentation format, usually lasting about 15 minutes with graphical exhibits and a question-and-answer period. Again, your professor will provide you with more detailed information as to the scheduling and format of these presentations.

What to bring to the laboratory

1. A good (non-erasable) pen

2. A calculator (and a knowledge of such things as degree/radian mode, etc.)

3. A brown-cover laboratory notebook of the approved type: **Ampad #22-157 Computation Book** or National Brand #43-648 Computation Notebook

4. This laboratory experiment manual (also available on the course web site)

5. The error analysis textbook, *An Introduction to Error Analysis*, by John R. Taylor. Part I of this text is assumed knowledge and must be mastered for satisfactory completion of the experiments.

If you have a notebook computer, you may find it very useful to bring to the laboratory to help in data recording and analysis.

Laboratory notebook

Your lab notebook will be an essential part of your laboratory work in this course, and it should contain a running account of the work you do. Entries should be made while the experiment is in progress, and you should use a standard format. Your notebook should:

1. provide the reader with a table of contents at the beginning, page 1, listing the number and title of the experiment, the date or dates when it was done, the page numbers in the notebook, and the name of your partner (see Table 1);

2. contain all pertinent information, schematic diagrams, observations, data, rough calculations, results, and conclusions. Think of your entries as being those in an informal diary or journal relating daily experiences.

In the laboratory, each experiment will be performed by a team of two investigators. *Respect for and cooperation with your partner is essential; neither partner should seek to dominate the partnership.* Each partner should take a turn working with the apparatus and making measurements; each should be involved in running the computer for data acquisition or analysis, as appropriate.

Each person is responsible for the complete documentation of the work performed and its analysis. That is, while you may discuss the experimental results with your lab partner, **your analysis of the experiment should be done individually**. Remember, you will write a semester-end technical report based on one of the experiments, and therefore a complete record of your observations and conclusions is essential when trying to recall the pertinent facts of an experiment that you may have completed weeks previously. Besides its use for the technical report, your lab notebook, along with the written summary and/or oral presentation, will be part of your grade on all four major experiments.

The exact form for any particular day's record will depend on the type of experiment and may vary considerably from one experiment to the next. Please observe the precautions emphasized in the laboratory instructions and appendices, and accord the research-type equipment the respect it deserves. Report any damaged equipment to your instructor immediately. In two of the experiments, computers are used to speed data gathering and analysis.

> **Warning:** **Do not alter the data acquisition computers or their programs in any way.**

Do not install desktop backgrounds, utilities, or software on any laboratory computer. Tampering with these computers, even if intended to be harmless, *violates the academic Honor Code.*

Head each experiment with the same material provided in the Table of Contents. Begin with a *brief* statement of the purpose of the experiment and a brief outline of the procedure or intended procedure. Two or three sentences should be enough. There is no need to copy the procedure contained in the laboratory manual. Please note that you do not always have at the beginning all the information you need to prepare a full description of purpose or procedure. The objectives may be poorly defined at the start and become crystallized only in the final stages of the experiment.

Whenever practical and appropriate, include a large, clear drawing, sketch, block diagram, or photograph of your experimental arrangement, to scale when necessary. Indicate clearly on the sketch critical quantities such as dimensions, volumes, masses, etc. Avoid excessive detail; *include only essential features.* In some cases you will want to record the manufacturer, model name or number, and HMC identification number of apparatus you use, as it may be essential that you get the same apparatus later, or someone else may wish to compare his results with yours. In any case you will need to know and record the relevant accuracies and limitations of the instruments you use.

After the data are recorded, there will generally be some calculations to be made. **Make necessary calculations in the lab notebook as you go along** to verify that your data is giving reasonable results; do not postpone all calculations after the experiment. If the data are all treated by some standard procedure, describe the procedure *briefly* for each calculation, giving any formulas that are to be used. Computer spreadsheets, such as Excel, are extremely powerful for handling data reduction and error propagation. Make sure, however, that the relevant equations are shown *algebraically* in the notebook.

(Define *any* quantities appearing in the formula that have not previously been defined.) *Always give a sample calculation*, starting with the formula, substituting experimental numbers, and carry the numerical work down to the result.

Here are some general rules for making entries in your **laboratory notebook**:

1. Use permanent ink, not pencil or erasable ink.

2. Do *not* use scratch paper—*all* records must be made directly in the notebook. Cultivate the skill

Table 1: Table of contents of a laboratory notebook.

Expt. #	Title	Partner	Grade	Date	Page #s
Expt. 1	Sampling theory	Marge Inovera		10/03/06	3–11

of committing your thoughts to paper as you are working; this will prove valuable to you later. (Left-hand pages may be used for scratch work.)

3. Do not erase or use "white out"—draw a single line through an incorrect entry and write the correct value nearby. Apparent errors sometimes later prove to be important.

4. Record data in tabular form whenever possible *with uncertainties* and *give units* in the heading of each column. If you are taking data on a computer, make sure that you print out a copy of the spreadsheet holding the raw (and analyzed) data, and that it is appropriately labeled. Trim the paper as appropriate and tape the sheet into the lab notebook.

5. Define all symbols used in diagrams, graphs, and equations whenever necessary to make the description or discussion understandable by another reader.

6. You may produce graphs using a data analysis program (Igor, Kaleidagraph, Excel) or by plotting by hand. In either ease, it is critical to plot data as you take it and to consider whether the data and plot make sense. If you use a computer to prepare the graph, be sure to annotate the graph either before printing or after it has been taped into your notebook. Discuss the significance of the graph in the notebook narrative.

7. Do not staple or tape a stack of figures or tables in your notebook. Tape each one in separately and provide appropriate explanation and discussion in a caption.

8. Determine the uncertainties in your data and results *as you go*, and let the calculations determine the number of measurements needed.

9. Record qualitative observations as well as numbers and diagrams. This is often very important to give meaning to otherwise unintelligible pages of numbers.

10. Do not fall into the habit of recording only your uncommented data in lab, leaving blank pages or spaces for description and calculation to be finished later. Entries should be made in order corresponding to the work you are doing, much like

a diary report, although complicated computations and analyses are usually undertaken after the data taking procedures have been completed.

11. Not all students, or all professors, can produce a showcase-type notebook, but your work should be as neat and orderly as possible. Sloppiness and carelessness cannot be overlooked even when the results are good.

12. Your notebook will be a success if you *or a colleague* could use it as a guide in repeating or expanding upon the particular experiment at a much later date.

Grading policies

Your laboratory grade will be based upon:

75% Experiment summaries and notebook, oral report(s), pre-lab preparation and in-lab performance and progress.

25% Technical report on one of the experiments.

Semester-end technical report

Each student is required to write one technical report for this course. This represents a significant fraction of your grade, so it is important that it be turned in on time and be well written. It will be based on an in-depth treatment of one of the four experiments you performed during the semester. After consultation with your professor, you may decide it is desirable to take supplementary data for your report. For this purpose, and for consultation with your professor on the content of your report, a dedicated meeting date is provided in the semester schedule for technical report preparation. Check with your laboratory professor for the due dates of the first draft (for comments) and of the final tech report. Allow sufficient time for several revisions after you receive comments on your draft. **Appendix D** of this manual contains some discussion of writing style guidelines for your report, which should be prepared in the manner of a more-or-less formal publication. Your professor can provide you with further guidance and examples of proper technical report style.

Philosophical Musings

Some Notes on Scientific Method

from *Zen and the Art of Motorcycle Maintenance* by Robert M. Pirsig

Actually I've never seen a cycle-maintenance problem complex enough really to require full-scale formal scientific method. Repair problems are not that hard. When I think of formal scientific method, an image sometimes comes to mind of an enormous juggernaut, a huge bulldozer–slow, tedious, lumbering, laborious, but invincible. It takes twice as long, five times as long, maybe a dozen times as long as informal mechanic's techniques, but you know in the end you're going to get it. There's no fault isolation problem in motorcycle maintenance that can stand up to it. When you've hit a really tough one, tried everything, racked your brain and nothing works, and you know that this time Nature has really decided to be difficult, you say, "Okay, Nature, that's the end of the nice guy," and you crank up the formal scientific method.

For this you keep a lab notebook. Everything gets written down formally so that you know at all times where you are, where you've been, where you're going, and where you want to get. In scientific work and electronics technology, this is necessary, because otherwise the problems get so complex you get lost in them and confused and forget what you know and what you don't know and have to give up. In cycle maintenance, things are not that involved; but when confusion starts, it's a good idea to hold it down by making everything formal and exact. Sometimes just the act of writing down the problems straightens out your head as to what they really are.

The logical statements entered into the notebook are broken down into six categories: (1) statement of the problem, (2) hypotheses as to the cause of the problem, (3) experiments designed to test each hypothesis, (4) predicted results of the experiments, (5) observed results of the experiments, and (6) conclusions from the results of the experiments. This is not different from the formal arrangement of many college and high-school lab notebooks, but the purpose here is no longer just busy-work. The purpose now is precise guidance of thoughts that will fail if they are not accurate.

The real purpose of scientific method is to make sure Nature hasn't misled you into thinking you know something you don't actually know. There's not a mechanic or scientist or technician alive who hasn't suffered from that one so much that he's not instinctively on guard.

That's the main reason why so much scientific and mechanical information sounds so dull and so cautious. If you get careless or go romanticizing scientific information, giving it a flourish here and there, Nature will soon make a complete fool out of you. It does it often enough anyway, even when you don't give it opportunities. One must be extremely careful and rigidly logical when dealing with Nature: one logical slip and an entire scientific edifice comes tumbling down. One false deduction about the machine and you can get hung up indefinitely.

In Part One of formal scientific method, which is the statement of the problem, the main skill is in stating absolutely no more than you are positive you know. It is much better to enter a statement "Solve Problem: Why doesn't cycle work?" which sounds dumb but is correct, than it is to enter a statement "Solve Problem: What is wrong with the electrical system?" when you don't absolutely know the trouble is in the electrical system. What you should state is "Solve Problem: What is wrong with cycle?" and then state as the first entry of Part Two:

"Hypothesis Number One: The trouble is in the electrical system." You think of as many hypotheses as you can; then you design experiments to test them to see which are true and which are false.

This careful approach to the beginning questions keeps you from taking a major wrong turn which might cause you weeks of extra work or can even hang you up completely. Scientific questions often have a surface appearance of dumbness for this reason. They are asked in order to prevent dumb mistakes later on.

Part Three, that part of formal scientific method called experimentation, is sometimes thought of by romantics as all of science itself, because that's the only part with much visual surface. They see lots of test tubes and bizarre equipment and people running around making discoveries. They do not see the experiment as part of a larger intellectual process, and so they often confuse experiments with demonstrations, which look the same. A man conducting a gee-whiz science show with fifty thousand dollars' worth of Frankenstein equipment is not doing anything scientific if he knows beforehand what the results of his efforts are gong to be. A motorcycle mechanic, on the other hand, who

honks the horn to see if the battery works is informally conducting a true scientific experiment. He is testing a hypothesis by putting the question to nature. The TV scientist who mutters sadly, "The experiment is a failure; we have failed to achieve what we had hoped for," is suffering mainly from a bad script-writer. An experiment is never a failure solely because it fails to achieve predicted results. An experiment is a failure only when it also fails adequately to test the hypothesis in question, when the data it produces don't prove anything one way or another.

Skill at this point consists of using experiments that test only the hypothesis in question, nothing less, nothing more. If the horn honks and the mechanic concludes that the whole electrical system is working, he is in deep trouble. He has reached an illogical conclusion. The honking horn only tells him that the battery and horn are working. To design an experiment properly, he has to think very rigidly in terms of what directly causes what. This you know from the hierarchy. The horn doesn't make the cycle go. Neither does the battery, except in a very indirect way. The point at which the electrical system directly causes the engine to fire is at the spark plugs; and if you don't test here at the output of the electrical system, you will never really know whether the failure is electrical or not.

To test properly, the mechanic removes the plug and lays it against the engine so that the base around the plug is electrically grounded, kicks the starter lever, and watches the sparkplug gap for a blue spark. If there isn't any, he can conclude one of two things: (a) there is an electrical failure or (b) his experiment is sloppy. If he is experienced, he will try it a few more times, checking connections, trying every way he can think of to get that plug to fire. Then if he can't get it to fire, he finally concludes that a is correct, there's an electrical failure, and the experiment is over. He has proved that his hypothesis is correct.

In the final category, conclusions, skill comes in stating no more than the experiment has proved. It hasn't proved that when he fixes the electrical system, the motorcycle will start. There may be other things wrong. But he does know that the motorcycle isn't going to run until the electrical system is working and he sets up the next formal question: "Solve problem: what is wrong with the electrical system?"

He then sets up hypotheses for these and tests them. By asking the right questions and choosing the right tests and drawing the right conclusions, the mechanic works his way down the echelons of the motorcycle hierarchy until he has found the exact specific cause or causes of the engine failure and then he changes them so that they no longer cause the failure.
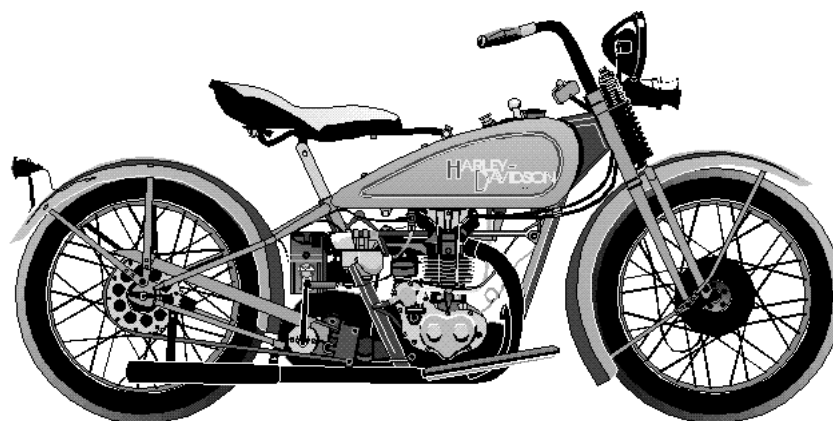
Part II

Experiments

Experiment 1

Ohm's Law and DC Circuits

Summary

At the commencement of a well-known technical institution a few years ago, graduates were handed a battery, some wire, and a light bulb and asked to light the bulb. Their fumblings and confusions were videotaped and presented in the documentary *Minds of Our Own*. By the end of this experiment you should be well-equipped to shine at graduation time, bringing fame and glory to Harvey Mudd College.

1.1 Overview

Metals conduct electricity, which means they have mobile charges (electrons) that can flow when "encouraged." A battery or a power supply can provide the encouragement by raising the potential energy of charges on one side of the metal with respect to the other. Charges then flow from the high-energy side to the low-energy side, constituting an **electric current**, which is defined as the amount of charge (coulombs) passing a cross section of the conductor per second. The unit of current is the **ampere** (1 A = 1 C/s) and the symbol for current is I.[1]

How much current should flow through a given piece of metal or conductor? That's a complicated question. The current could depend on the energy difference between the two ends, the shape of the conductor, the material of the conductor, its temperature... lots of things. To simplify a bit, we define the **voltage difference** V across the conductor as the difference in potential energy that a unit charge has at opposite ends of the conductor. Then for a given conductor, Ohm's law claims that the current flowing *through* the conductor is proportional to the voltage difference *across* the conductor: $I \propto V$. That is, if you double the energy incentive for charges to move from one side to the other, then you double the rate at which they move from one side to the other. If Ohm's law holds for a given object, then the **resistance** of the object is given by $R = V/I$.

Ohm's law is not a law of the same class as Newton's laws of motion. Some objects obey Ohm's law quite well; others don't at all. In this experiment you will test whether an object that is labeled "resistor" actually obeys Ohm's law. You will then study the behavior of an incandescent light bulb.

1.2 Theory

Aristotle (384–322 BCE) held that objects on Earth fall at steady speed, and heavier objects fall faster than light objects. Think about dropping a stone and a piece of paper at the same time from the same height and you may be willing to muster some charity for the seemingly naive Aristotle, who was said to have performed his experiments with small stones in water. Galileo Galilei (1564 – 1642) argued persuasively against this doctrine, holding that insofar as air resistance could be neglected, all objects fall the same way: at a steadily increasing rate. That is, they accelerate. You know all about that.

Electrons in a wire behave a whole lot more like Aristotelian pebbles than Galilean cannonballs. In response to an applied push (the electric field associated with the applied voltage difference), the free electrons in the metal drift with an average speed that is proportional to the strength of the push. In consequence, the electric current that flows in the conductor is proportional to the applied voltage; double the voltage and you double the amount of current. (If the current gets too large, this law breaks down. The resistor may, too!)

If an object obeys Ohm's law, its **resistance** is given by dividing the voltage drop across the object by the current flowing through it,

$$R = \frac{V}{I} \tag{1.1}$$

Resistances are measured in **ohms**: 1 Ω = 1 V/A. The resistance of a conductor depends on the material of the conductor, as well as the conductor's size and shape. Georg Simon Ohm (1789 – 1854) described the proportionality implied by (1.1) in 1827, and it is known as **Ohm's law** (usually written in the form $V = IR$).

Why don't the electrons accelerate under the influence of an applied force? They do, but they are continually colliding with vibrating metal atoms in the wire. The collisions transfer momentum and energy from the electrons to the wire, so that the (average) electron drift

---

[1] The I stands for the *intensity* of the current.

velocity reaches a steady value, much like the terminal velocity of an object in free fall. These collisions heat the atoms of the wire, an effect called **Joule heating**. The rate at which electrical energy is converted to heat (dissipated) is

$$P = IV = I^2 R = \frac{V^2}{R} \qquad (1.2)$$

where in the final two forms we have assumed ohmic behavior and used (1.1) to eliminate either V or I. If I is in amperes, V in volts, and R in ohms, then the power P is in watts (W).

Two conservation laws govern the flow of charges in conducting circuits and are known as **Kirchhoff's laws**:

(a) Because electric charge is conserved, the algebraic sum of currents from any point in a circuit is zero. That is, the sum of the currents arriving at a point in the circuit is equal to the sum of the currents leaving the same point.

(b) Because the electrostatic force between any pair of charges is conservative, we can define a unique value of electrostatic potential to each point in a

circuit (up to an overall constant). Therefore, the sum of voltage changes across all elements forming a closed loop in a circuit is zero.
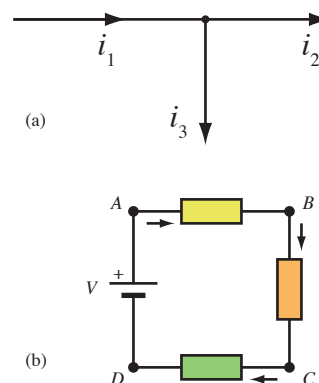
These laws are illustrated in Fig. 1.1.

Figure 1.1 (a) The sum of currents entering the junction is equal to the sum of currents leaving it: $i_1 = i_2 + i_3$. (b) The algebraic sum of voltage drops around a closed circuit vanishes: $V_{AB} + V_{BC} + V_{CD} + V_{DA} = 0$.

1.3 Precautions

Measuring current

To measure the current flowing through a component, insert an ammeter *in series* so that all the current that flows out of the wire *must* flow through the ammeter (see Fig. 1.2). By design, an ammeter has very low resistance to allow the current to traverse it unimpeded. If the ammeter is connected in parallel, so current can choose whether to pass through the ammeter or the resistor you are investigating, nearly all the current will choose to pass through the ammeter, **thereby destroying it!**
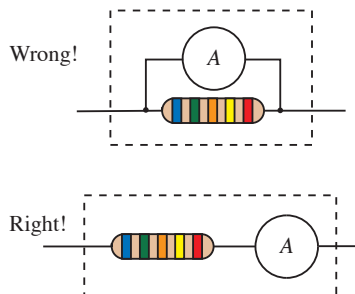
Figure 1.2 To measure current passing through a circuit element you must insert the ammeter *in series* with the element, as shown in the lower figure.

Meters have a maximum current or voltage they can handle. Carefully respect this limit or you will damage

the equipment.
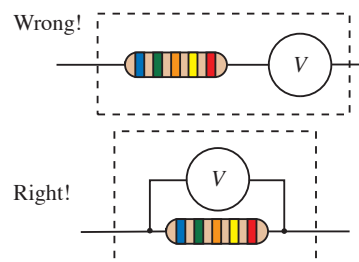
Figure 1.3 To measure the voltage drop across a circuit element, you must wire the voltmeter *in parallel with* the element, as shown in the lower figure.

Measuring voltage

To use a voltmeter to measure the voltage drop across a component, wire the voltmeter *in parallel* with the component, as shown in Fig. 1.3. This is *opposite* to the case of the ammeter. A voltmeter ideally has a very high resistance, preventing appreciable current from flowing through it, which would change the voltage drop it is trying to measure. If you hook it up in series with the component you are investigating, you probably won't damage the meter, but you will disrupt the flow of current through your circuit and get nonsensical

readings.

Resistors have a maximum allowable power dissipa-

1.4 Procedure

Team up with a partner of comparable experience with electrical circuits and proceed through the various parts of the experiment at a pace appropriate to your background. Each partner should work with the equipment and contribute ideas; neither partner should monopolize the equipment or be relegated to the role of recorder. Take turns! If you are saddled with an uncooperative partner, please raise the issue with your instructor.

Your goal for the first day should be to get through the first part: Testing Ohm's Law. Consult your instructor if you have questions or if you need to verify that a circuit is properly wired.

You must document your work carefully in your lab notebook. We recommend that you put the table of contents in the back of the notebook, and that you start the diary of your work in the front. That way, you'll have as much room for the table of contents as you need. If your notebook from Physics 28 has enough room left, you may use it for Physics 53.

Be sure to reproduce schematics of the circuits you construct and describe your observations and measurements in sufficient detail that they could be reproduced by someone else at a later time. Neatness and organization are worth cultivating!

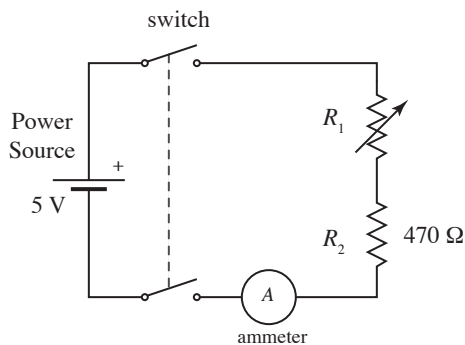### 1.4.1 Testing Ohm's Law

Figure 1.4 A circuit for testing Ohm's law. (The arrow through R_1 means that its value is variable.)

In this first experiment you will seek to confirm that the current and voltage are indeed proportional, as predicted by Ohm's law (1.1). By varying the value of R_1 in the circuit shown in Fig. 1.4, you can change the current flowing through the circuit. By measuring this

tion. You can use (1.2) to compute the power that will be produced as heat for a given current I.

current at the same time as you measure the voltage *across* resistor R_2, you can determine whether R_2 obeys Ohm's law and you can compare the value of resistance you obtain from your measurements with R_2's nominal resistance of 470 Ω.[2] See Fig. 1.2 and Fig. 1.3 if you have questions about how to wire the meters, and ask your instructor if you still have questions.

Things to consider:

- What is a good way to confirm a linear relationship? How will you "convict object R_2 of ohmic behavior" most persuasively?

- For purposes of testing whether the object labeled R_2 obeys Ohm's law, does it matter whether the object labeled R_1 obeys Ohm's law? Why or why not?

- Assuming that the object labeled R_1 behaves as a resistor, what value of R_1 will produce the largest current through the circuit?

- How can you quantify the errors in the current and voltage measurements? — Actually, I'll give you a heads up on this one. On the under side of each multimeter you will find a fairly cryptic description of the procedure. In essence, the uncertainty is the sum of two terms: one term that is proportional to the value displayed and a constant term, which is typically 1 or 2 units of the least significant digit on the display. For example, if the meter says "0.03% + 1 digit," and you read 3.245 V on the display, you would compute an uncertainty of

$$\delta V = (3.245 \text{ V})(0.0003) + 0.001 \text{ V}$$
$$= 0.0019735 \text{ V} \rightarrow 0.002 \text{ V}$$

Before collecting data, you should write carefully thought-out answers to the previous questions in your notebook — in addition, of course, to the appropriate circuit diagram.

Finish your analysis of this section, print out a graph, tape it into your notebook, annotate it, and summarize your results before moving on to the next section.

1.4.2 Lighting a bulb

An incandescent light bulb consists of a thin metal filament (typically tungsten) in a glass bulb filled with

---

[2]*Nominal* means, roughly, "in name alone." You should treat nominal values with some suspicion until you have confirmed their accuracy.

an inert gas at low pressure. In this part of the experiment, you will investigate whether a CM 130 light bulb behaves like a resistor.

The dimensions of the filament are chosen so that the filament emits white light when its rated voltage is applied. Exceeding the rated voltage (or current) shortens the life of the filament or burns it out suddenly.

> **Warning:** **The bulb you will study is rated for a maximum current of 0.15 A.**

First the easy way, using the multimeter as an ohmmeter. Set the meter to read Ω (or $k\Omega$) and use a combination of clip leads and cables to wire the light bulb across the red port labeled "V-Ω" and the black port labeled "Common." The meter will then give you a rough idea of the bulb's resistance. You might consider what happens when you clip the leads together.

Now the careful way. To power the bulb, you will use an adjustable-voltage power supply like the one shown in Fig. 1.5. Turning the knob labeled "6 V" allows you to adjust the voltage between the 6-V output port and the common from zero up to slightly more than 6 volts. With nothing connected to the supply, power it on and make sure that the meter is set to display the voltage of the 6-V port. Adjust the knob to observe how many turns you must make to cover the full range from 0 V to slightly more than 6 V. (Each partner should do this.) Then rotate the knob so the voltage reads zero and turn off the supply.

Figure 1.5 A DC power supply. Use the 6-V output port and the common.

You already know how to include a voltmeter and an ammeter in your circuit. Note that you face a choice when wiring some ammeters as to which port to use. Aim to use the low-current port when you are measuring currents below the rated capacity of that port. That is, if you are measuring a current less than 100 mA, use the port labeled 100 mA. Before switching ports, however, you should stop the current by turning off the power supply.

To include the bulb in the circuit, attach clip leads to the two contacts on the socket. Once you have wired the circuit *with the power supply off*, turn it on and ramp the voltage up slowly. The ammeter should register current flow before the voltage departs much from zero. If not, reduce the voltage to zero and check the wiring of the circuit. When your circuit is wired properly, study the behavior of the lamp. Don't forget to plot the data as you take it. If you don't know how to set up Kaleidagraph, Igor, or Excel to update the plot as you add data, ask your instructor.

1.5 Questions

We conduct experiments to investigate phenomena and answer questions. At the conclusion of this experiment you should have printed out appropriate plots, taped them into your lab notebook, and annotated the plots with a discussion of their salient features. Then, you should answer the following questions in your notebook:

1. Does the resistor box set at "470 Ω" function as a resistor and obey Ohm's law? If so, what is its resistance? How confident can you be of your conclusions?

2. Does the light bulb you studied obey Ohm's law? If so, how closely? If not, why not? How do you think the multimeter measures resistance?

3. Do you still agree with your original thoughts on how best to conduct this experiment, and why? (Refer to the questions you answered in the section 1.4.1.) If your thoughts have changed, when did they change and what caused the change?

4. If asked to work on this experiment for an extra hour or two, how would you productively spend that time? Avoid answers like "take more data;" instead, point to specific types of data you'd try to take, or changes you'd make to the procedure to clear up remaining issues. If you view your results as conclusive, well-understood, and without any loose ends, perhaps you have *no* future work to suggest.

Experiment 2

Magnetic Fields and an Absolute Determination of Current

Abstract

Electric currents produce magnetic fields, which in turn exert torques on permanent magnets. By using a simple geometry to permit the field to be readily calculated from the current, and by carefully balancing the torques from the current and from the magnet, it is possible to measure the current (an *electrical* quantity) by making measurements solely of the *mechanical* quantities of mass, length, and time. Your mission: to measure a current *as accurately as possible* using a ruler, some compasses, a bar magnet, a stopwatch, and a scale.

2.1 Overview

1. Current in a coil produces a magnetic field $\mathbf{B}_{\text{coil}}$ along the axis at the coil's center. The field is proportional to the current i flowing in the coil and given by (2.2) below. Measuring $\mathbf{B}_{\text{coil}}$ will allow you to measure i.

2. To measure $\mathbf{B}_{\text{coil}}$, you will produce an equal and opposite magnetic field with a bar magnet, as determined using a magnetic compass needle. To quantify the field of the bar magnet, and therefore $\mathbf{B}_{\text{coil}}$, you will measure the strength of the bar magnet's dipole moment M. The magnetic field of the bar magnet is proportional to M and falls off with distance according to (2.3) below.

   The calibration requires two measurements. First, you will use a compass needle to compare the strength of the bar magnet's field to that of the Earth. Since the compass needle aligns with the horizontal component of the total magnetic field at its center, by orienting the bar magnet's field perpendicular to the Earth's, you can deduce their relative strength from the angle of rotation from local magnetic north. Measurements at several magnet positions yield a value for M/B_h, where B_h is the local value of the horizontal component of Earth's magnetic field.

3. Second, you will suspend the bar magnet in a sling and make it oscillate about the Earth's field. The rate of oscillation is proportional to both M and B_h, allowing you to deduce their product, MB_h. Since you already know M/B_h, you may calculate both the magnet strength, M, and the local value of Earth's magnetic field, B_h.

4. Using the compass needle to make $B_{\text{bar}}(x)$ equal to $B_{\text{coil}}(i)$ for different currents i, you can deduce the current i from the magnet position x.

As the above list makes plain, this experiment involves many individual measurements which are combined to determine the current flowing in the coil. They may be done in any order. Each measurement contributes to the uncertainty of that determination, but some contribute more than others; *your challenge is to make the combined uncertainty as small as possible.*

2.2 Background

The great mathematician, astronomer, and physicist Johann Carl Friedrich Gauss first realized in 1833 the possibility of determining electromagnetic quantities, such as electric current, by measuring only mechanical quantities, such as mass, length, and time. To honor this work a common set of electromagnetic units are called **gaussian units**, and the gaussian unit for magnetic field strength is called the **gauss**.[1] This experiment is a variant of Gauss's original design.

Later work on electric currents focused on resistance standards as the practical means of quantifying current, and the common unit of current, the **ampere**, was defined to have a convenient magnitude. The Italian physicist Giovanni Giorgi showed in 1901 that it was possible to combine the mechanical units of the metric system (meters, kilograms, and seconds) with an electrical unit (ohms or amperes) to produce a coherent system of units. In 1954 the CGPM[2] officially added the fundamental unit of the ampere to the *Système international* (SI), using the definition

---

[1] 1 gauss $= 10^{-4}$ tesla, the SI unit of magnetic field. Gaussian units use the gram, centimeter, and second as base units for mechanical quantities, although Gauss himself used the millimeter. See http://www.bipm.org/en/si/history-si/.

[2] General Conference on Weights and Measures, *Conférence générale des poids et mesures*, in French.
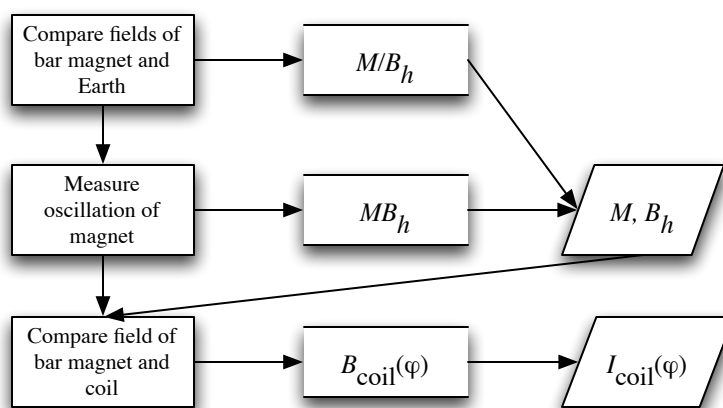
Figure 2.1: Logical flow of the experiment.

The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 m apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} N/m of length.

The reconciliation of mechanical and electrical units was accomplished through the definition of the constant μ_0, the **magnetic permeability of free space**; it is *defined* to have the exact value

$$\mu_0 \equiv 4\pi \times 10^{-7} \text{ kg m A}^{-2} \text{ s}^{-2}$$

and is used in expressions such as

$$d\mathbf{B} = \frac{\mu_0}{4\pi} \frac{i d\mathbf{l} \times \hat{\mathbf{r}}}{r^2} \qquad (2.1)$$

which relate currents and geometric quantities to magnetic fields. Notice that the strange factor of 10^{-7} in the definition of μ_0 is chosen to give the ampere a "reasonable" size.

Recent work on circuits that permit one to transfer one electron at a time across a conducting bridge may lead in the future to a more refined definition of the SI unit of charge, the coulomb, in terms of the electron charge. In the meantime, the coulomb is defined as $1 \text{ C} = 1 \text{ A} \times 1 \text{ s}$, and the accepted number of electrons in a coulomb is $6.241\,506 \times 10^{18}$ (1996 value).

Background information on the theory of magnetic dipoles in external fields can be found in HRK Chapter 32, sections 5 and 6. Review on the theory of magnetic fields produced by current loops will be found in HRK Chapter 33, sections 1 and 2.

Earth's magnetic field

The strength of Earth's magnetic field is several tens of microteslas, but the actual magnitude and direction depend upon position on the surface.

Figure 2.2 illustrates the overall shape of Earth's field. Note that 1 tesla = 1 weber/meter^2 = 1 newton/(ampere meter) = 10^4 gauss.
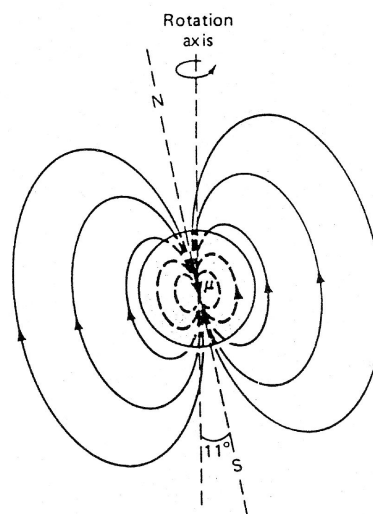
Figure 2.2 Geometry of Earth's magnetic field.

Recent evidence suggests that we are in a period of rapid change in Earth's field, perhaps heading to a reversal of the field's direction. For the present, the values in the following table describe Earth's field.

Place	$B_{\text{horizontal}}$	B_{vertical}
magnetic equator	35 μT	≈ 0
magnetic poles	≈ 0	70 μT
So. California	25 μT	45 μT

Fluctuations at a given position are on the order of 0.01 μT; the rate of gradual yearly change is on the order of 0.1 μT, and magnetic storm fluctuations due to solar activity are on the order of 0.1 μT/min.

2.3 Theory

The magnetic field produced at the center of a circular coil of N turns, radius R, and current i comes directly from applying the Biot-Savart law, (2.1), and results in

$$B_{\text{coil}} = \frac{\mu_0 N i}{2R} \qquad (2.2)$$

The magnetic field from a bar magnet on its symmetry axis is also fairly simple. A magnetized object like a bar magnet is well approximated as a magnetic **dipole** at distances large compared with the dimensions of the magnet. (Its field looks very much like the Earth's field shown in Fig. 2.2.) The magnetic field of a dipole drops off in magnitude proportional to $1/r^3$. However, the behavior is a bit more complicated at distances comparable to the size of the magnet, as is the case in this experiment. An adequate description is to think of the magnet as consisting of two point-like poles of equal and opposite strength, separated by a distance $2b$. [Note: Such magnetic "monopoles" do not actually exist, as far as anybody knows. They are however, a convenient fiction to use here for the purpose of describing the magnet's close-in field, at least approximately.] You can readily show that the resulting approximate magnetic field at a distance r from the center of the magnet is

$$B_{\text{bar}} = \frac{\mu_0}{4\pi} \frac{2Mr}{(r^2 - b^2)^2} \qquad (2.3)$$

In the above expression, $M = 2bp$ is the **magnetic dipole moment**, and p is the monopole strength of each (fictitious) pole. In SI units, dipole moment has the units of A m^2. The geometry is illustrated in Fig. 2.3.
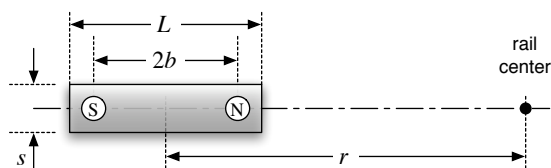

Figure 2.3 Bar magnet dimensions and locations.

When the magnetometer rail is oriented as shown in Fig. 2.4, perpendicular to the Earth's horizontal magnetic field B_h, the compass magnet points in the direction of the net horizontal vector sum of the Earth's field and that of the bar magnet. Accordingly, the angle of deflection ϕ is given by

$$\tan \phi = B_{\text{bar}}/B_h \qquad (2.4)$$

Combining equations (2.3) and (2.4), we obtain an expression for the *ratio* M/B_h :

$$\frac{M}{B_h} = \frac{2\pi}{\mu_0} r^3 \left(1 - \frac{b^2}{r^2}\right)^2 \tan \phi \qquad (2.5)$$

By measuring the angle ϕ for a variety of distances r (on both sides) we obtain a best estimate of the ratio M/B_h. If we knew the value of B_h, we would be almost finished, since the dipole moment—and therefore the magnetic field B_{bar}—would be known. However, the Earth's magnetic field is highly spatially variable, especially in a building environment filled with ferromagnetic materials, so we only know its value roughly in advance. Measuring b and the oscillation frequency of the bar magnet in Earth's field gets us around this problem.

When the bar magnet is immersed in the Earth's magnetic field, it experiences a **torque**, given by $\boldsymbol{\tau} = \mathbf{M} \times \mathbf{B}_h$. Note that the magnitude of the torque is proportional to the *product* MB_h. Once we know both the product and the ratio of M and B_h, we know their separate values, since $M = \sqrt{(M/B_h)(MB_h)}$.

If suspended by a thread of negligible torsion, the magnet will oscillate back and forth under the restoring torque of the Earth's magnetic field. The magnet's equation of motion is

$$I \frac{d^2\theta}{dt^2} = -MB_h \sin\theta \qquad (2.6)$$

where I is the rotational moment of inertia of the bar magnet about its center. This is the same equation that governs a pendulum. For oscillations of small amplitude, the oscillation frequency f is related to MB_h by the expression

$$MB_h = 4\pi^2 f^2 I \qquad (2.7)$$

Once M and B_h have been determined, the final result of the experiment can be obtained. Knowing M, we can use (2.3) to calculate B_{bar} at any distance r. For a variety of current values i (both positive and negative), we have measured the distance r at which the compass deflection is zero, at which point $|\mathbf{B}_{\text{bar}}| = |\mathbf{B}_{\text{coil}}|$. Then from (2.2) we obtain i, a measurement of the current which we have obtained without reference to external electrical standards.
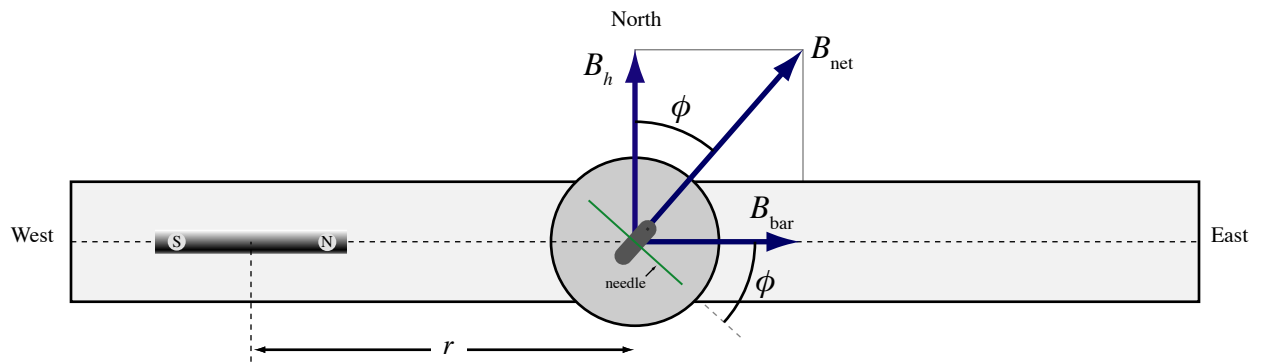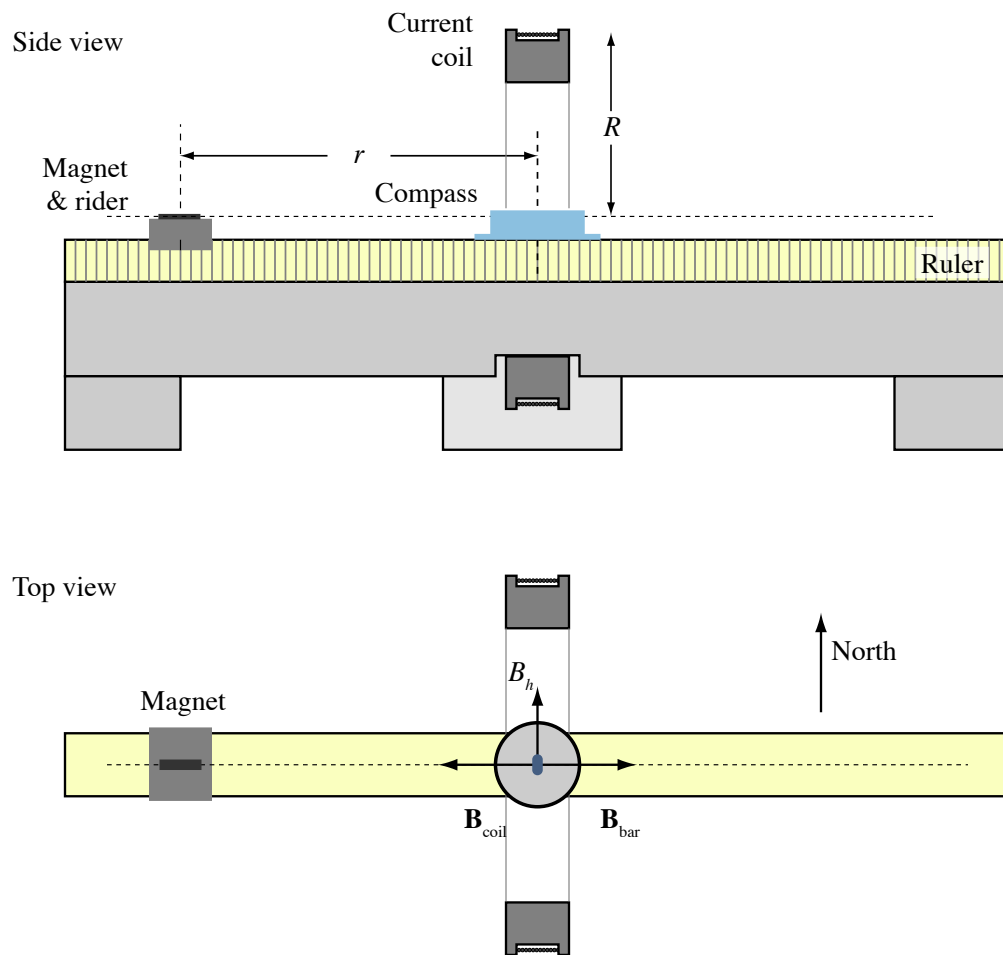
Figure 2.4: Top view of the magnetometer rail assembly.

Figure 2.5: Schematic views of magnetometer rail assembly and current coils. Compass orientation shown is for $\mathbf{B}_{coil}$ exactly cancelling $\mathbf{B}_{bar}$.
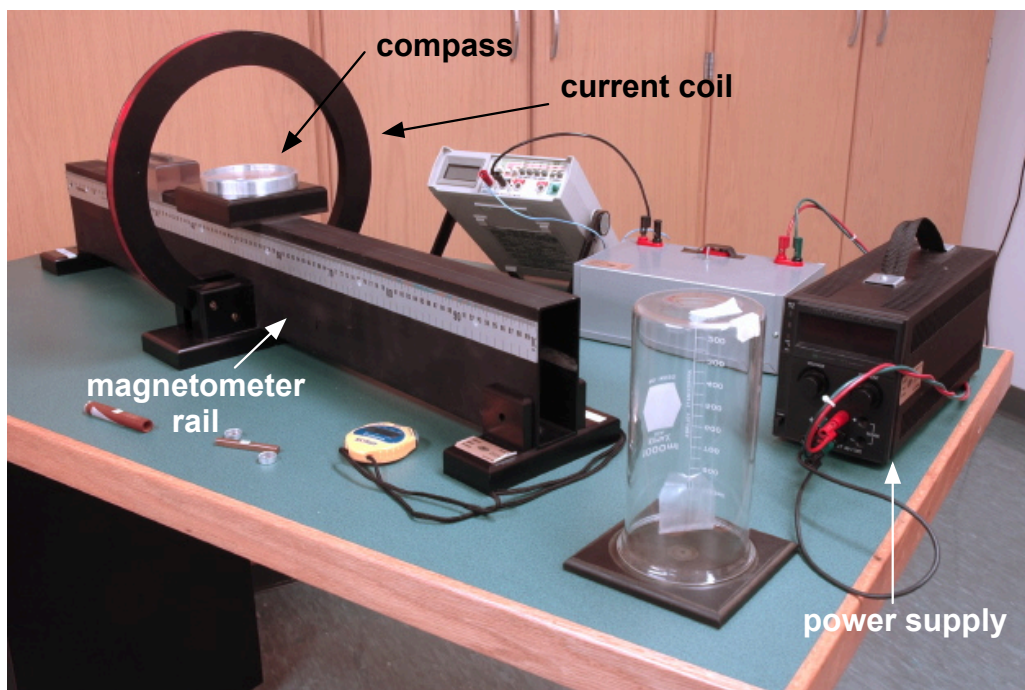
Figure 2.6: Photograph of the magnetometer apparatus.

2.4 Experimental Equipment and Procedures

Guidance on how to carry out the three kinds of measurements needed to complete this experiment (two to calibrate the bar magnet and one to compare measured and calculated currents) are given below. As you (re)read these procedures, keep in mind your goal: to determine the current with the greatest possible accuracy. Each measurement contributes to the overall uncertainty in your determination of the current, although some will likely add much more than others. Keep track of these uncertainties as you go and think carefully about which measurements are key to reducing the final uncertainty.

2.4.1   Preliminary setup

Figure 2.5 shows another layout view of the magnetometer rail. As seen in the photo of Fig. 2.6, the scale on the side of the rail and the scribe mark on the side of the plastic magnet holder allow the distance r between the magnet and rail centers to be measured. Throughout the experiment, all superfluous magnets and magnetic materials should be removed from the vicinity of the apparatus. Carefully orient the rail perpendicular to the ambient magnetic field, so that the indicator needle is parallel to the rail axis. Rotate the circular housing of the compass so that the zero reading of the scale is aligned with the needle.

It may be helpful to run the same current through the coil in opposite senses to confirm that the rail has been properly aligned, since equal currents should produce needle deflections of equal magnitude. It would be a good time to re-read Appendix A about the hazards of electric currents at this point. The circuit used to supply the current to the magnet coils is shown in Fig. 2.7. The reversing switch allows you to change the direction of current, so that you can perform trials with the magnetic field pointing in both directions. **To protect this switch from arcing damage, please reduce the current to zero before reversing direction.** The load resistor is located inside the switch box, and its purpose is to limit the current flowing in the circuit. Make sure that the meter (labeled "A" in the diagram) is set to ammeter mode. To minimize the effects of stray magnetic fields, the switch box, the multimeter and power supply should be kept several feet from the compass, and twisted pair wire leads are used to connect the coil to the supply and switch box.
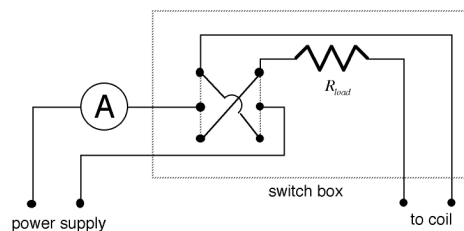

Figure 2.7   Circuit diagram of coil current supply with reversing switch

Adjust the current using the knobs on the power supply, one of which limits the voltage, the other the current. You can tell which one is controlling at any point by the light above the knob. To use the same current in both directions, use the current knob to set the current (which means you need to have the voltage knob turned high). Then turn the voltage knob to zero, throw the reversing switch, and ramp the voltage up until the current knob again limits the current.

> **Warning:** **Do not exceed a current of 0.7 A.**

Once you are satisfied with the orientation of the rail, its position on the laboratory bench can be marked with tape so it can be returned to the same position when it has to be moved later in the experiment. Please remove the tape before leaving the laboratory.

2.4.2   Compass deflection measurements

After selecting a bar magnet to use, mark its poles so that you can keep track of its orientation throughout the experiment. (*You should also identify the magnet you are using in your notebook, since you will need to use it again the following week. As a courtesy to the other students, please try not to drop or damage the magnets in any way that would require them to retake their data from the previous week!*)

To determine the true magnetic center of your magnet, compare the distances r measured on the right and left sides of the compass where the compass deflection is $45°$. If these distances differ slightly (as they probably will), the magnet's effective center is offset from its geometric center. You should carefully record this offset distance and use it during the rest of the experiment to correct all your measured values of r.

When the magnet is placed on the rail, the compass deflects to a new equilibrium position at angle ϕ. Although in theory a single measurement of r and ϕ would suffice to determine M/B_h from (2.5), in practice a larger number of measurements is needed to determine the ratio with good precision. Measurements should be taken at a generous number of locations on both sides of the center, at distances that probe both the near-dipole field and the far-field $1/r^3$ regions. In this way, it is possible to capture and account for any non-idealities in the bar magnet (more on this subject below).

> **Warning:** **When handling the bar magnet in this and subsequent parts of the experiment, take care not to subject it to mechanical shocks or magnetic influences that could change its magnetization.**

Note that in (2.5), r is assumed to be a positive number. However, it will be convenient for us to define r as negative on one side of the rail in order to separate our north pole and south pole measurements from one another; we just need to remember to take the absolute value of r when using (2.5). In selecting positions to take measurements, bear in mind that the least relative error in the quantity $\tan\phi$ is obtained for angles near $45°$. So while it is desirable to obtain measurements over a wide range of angles, those obtained near $45°$ will contribute with the most weight to determining M/B_h.

2.4.3   The Bar magnet and torsional oscillations

As shown in Fig. 2.8, in the experiment's second part the magnet is suspended from a slender thread beneath an inverted glass beaker. In the presence of the external field $\mathbf{B}_h$, the magnet oscillates about its equilibrium position. The measurement of the frequency of small-amplitude oscillation is accomplished with a stopwatch, and can be performed with quite high precision by timing the duration of several consecutive periods.
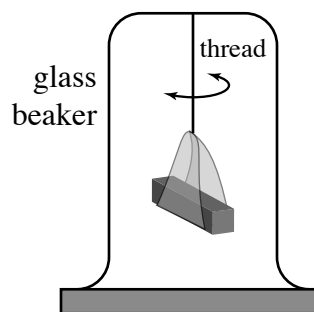
Figure 2.8   Torsional magnet suspension.

For the measurement to be meaningful, it must be carried out in the *same location* that the M/B_h measurement was made. To do this, the coil and compass are temporarily removed from the rail assembly, allowing the beaker and magnet to be placed on the rail in their place. (*Note: The aluminum mounting base for the compass can be replaced with a non-metallic one, which eliminates the eddy currents that otherwise tend to damp out the magnet oscillations rather rapidly.*)

It is also necessary to make accurate physical measurements of the bar magnet itself. The moment of inertia of the uniform bar of length L and square cross section s^2 is,

$$I = \frac{m(L^2 + s^2)}{12} \qquad (2.8)$$

The mass m is measured with a lab balance and the dimensions are measured with calipers (along with accompanying uncertainties).

Consideration of equations (2.3) and (2.5) also shows that it is necessary to determine the distance $2b$ between the effective poles of the magnet. There are different possible approaches to getting this quantity.

A simple preliminary approach is to estimate the positions of the poles visually using small compasses, and measure the distance $2b$ between them with a ruler. This approach will give a fairly good "ballpark" value. Measure and record this value for b by this method. An alternative approach, which may yield a more accurate estimate of b, is to fit to (2.5) rewritten as (2.9).

At this point, you have all the measurements you need to calibrate the bar magnet. See the analysis section below for suggestions on how to reduce your data to determine the bar magnet's dipole moment M. Even if you don't have time for a thorough analysis of the data from the deflection measurements at this point, you should at least compute a value of M/B_h from a single data point and combine this with the MB_h value from this section to obtain an estimate of M to carry forward to the following section.

2.4.4 Measuring the current

Now that you have calibrated the magnet, you can use its field to measure the current in the coil. If you have not already done so, you should determine the number N and radius R of the coil's turns, for use in (2.2). As described above, you will use the two knobs on the power supply, as well as the switch box, to adjust the current flowing in the coil. **Remember to ramp the current to zero before reversing the current direction.** For each current value indicated on the multimeter adjust the position of the bar magnet until the compass needle returns to the zero position, at which point the field of the magnet and field of the coil are equal and opposite.

> **Warning:** Do not exceed a current of 0.7 A.

Explore several current settings, both positive and negative, in the range of 0.4 to 0.7 A, as read by the multimeter. The more measurements of this kind that are made, the better you'll be able to characterize any systematic errors or asymmetries in the measurements. Note that for each current setting, there are *two* magnet positions that can be found to cancel the field of the current: on the right and left of the compass (*i.e.*, at positive and negative r).

2.5 Experimental Analysis

The principal goal of this experiment is to calculate the current i in the coil as accurately as possible, and to compare these calculations to the "accepted" value read from the multimeter. Think carefully about how best to demonstrate the quality of your measurements and to exhibit any systematic discrepancies.

Perhaps the most significant challenge here is in distilling the best value of M from all of the rail measurements. You should first carry out a quick "ballpark" calculation to verify that your results make sense, before getting too deeply into the detailed analysis. It is

at this stage that it is usually easiest to catch a simple mistake, such as incorrect unit conversion or a missing factor of 2, etc. *Set up an organized spreadsheet to help with the analysis.* You will need to have the spreadsheet ready to use by the second day in the lab, so be sure to set it up before coming to lab.

Pick a single representative measurement of r vs. ϕ near $\sim 45°$ and using (2.5), get a provisional value of M/B_h. Then using (2.7), get a provisional value for $M B_h$; you can defer any error propagation until later, since your purpose here is to perform a quick check. If everything is correct up to this point, the dipole moment M should be on the order of $0.5 - 1.0 \,\mathrm{A\,m^2}$ and the horizontal magnetic field B_h should be roughly 2×10^{-5} T. If your results are not in this range, go back and find the mistake before proceeding further. Finally, set equations (2.2) and (2.3) equal to one another using your provisional values for M and b to obtain the current i for one of your trials in the third part of the experiment. If all is well, it should agree with the multimeter reading within about 10% or so. Next we'll discuss how the final analysis will hopefully refine this result to 1% or better.

The simplest approach would be to obtain separate estimates of M/B_h from (2.5) for every measurement, and then take their average. However, this approach has some disadvantages. Some of the important sources of error, such as errors in the pole spacing b for example, tend to push the M/B_h estimate off in one direction *systematically*, rather than randomly. As a result, averaging over a lot of measurements yields a result that retains a bias from the "true" value, which cannot be averaged out.

A more sophisticated and potentially better approach is to *fit* the expression of (2.5) to the unknown parameters M/B_h and b, given the experimental measurements of ϕ vs. r. Solving the expression for ϕ, we want to fit the experimental data to the function,

$$\phi = \arctan\left\{\frac{\mu_0}{2\pi} \frac{(M/B_h)}{|r|^3 \left(1 - \frac{b^2}{r^2}\right)^2}\right\} \qquad (2.9)$$

Using a plotting/fitting computer utility such as Igor, Kaleidagraph, or Origin and proper error weighting, you can solve for the unknowns M/B_h and b, as well as their uncertainties. Fig. 2.9 shows schematically how we might expect this fit to look when plotted. Sometimes, however, the fit yields unreasonable values for b, so be careful. You may decide that it is preferable to fix the value of b during the fitting operation to extract a best-fit value for M/B_h.
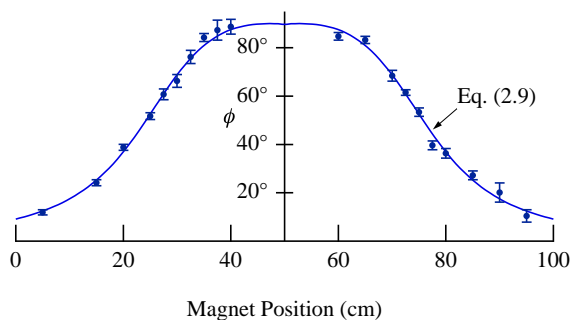
Figure 2.9   Compass rotation ϕ vs. magnet position data fitted to (2.9). Note that the independent variable plotted on the x axis here is not the quantity r used in the discussion and formulae. Rather, r is the magnet position minus the rail center position, converted to meters.

2.5.1   Fitting

The use of a plotting program to perform such a parameter fit is a great labor-saving convenience, and pretty easy. However, it pays to be familiar with the software and how it works before you need to use it. Here are some hints and reminders for using **Igor**, **Kaleidagraph**, and **Origin** to analyze this experiment:

Igor

Igor allows you to name the fit parameters anything you like. For simplicity here, we will use the names **b** and **MoverB** and call the independent variable **r**. Igor works only in radians, so we must convert into degrees. Assuming that distances are measured in meters, the fitting function definition for Igor would look like:

```
f(r) = (180/pi)*atan(MoverB*2e-7/(abs(r^3)*
  (1-(b/r)^2)^2))
```

Igor also requires you to provide starting guess values for the fitting parameters, but these are specified in the 'Coefficients' pane of the fitting dialog. Igor reports the value of χ^2, so if you would like the reduced χ^2 value, divide by the number of degrees of freedom.

Kaleidagraph

When defining your fitting function, Kaleidagraph requires the independent variable (i.e., r) to be named

m0. The most recent version allows user-defined fit parameter names, but earlier versions require the parameters (i.e b and M/B_h) to be named `m1` and `m2` (and so on). So with the program set to compute in degrees (rather than radians), the definition of your fitting function might look something like this:

```
invtan(m2*2e-7/(abs(m0^3)*(1-(m1/m0)^2)^2));
m1=0.03;
m2=32000
```

Note that it is necessary to provide the initial guess values following the function definition, or the program won't be able to get a solution. Also remember that you *must* specify a **weighted** fit, or else the χ^2 goodness-of-fit figure you obtain will be meaningless. In Kaleidagraph, you need to divide this number by the number of degrees of freedom (# of data pts. – # of parameters) in order to get χ^2 per degree of freedom.

Origin

Like Igor, when defining your fitting function, Origin allows you to name the parameters whatever you like. For simplicity here, we will use the names **b** and **M** and call the independent variable **R**. Origin works only in radians, so we must convert into degrees. The fitting function definition for Origin would look like:

```
(180/pi)*atan(M*2e-7/(abs(R^3)*
  (1-((b^2)/(R^2)))^2))
```

Origin also requires you to provide starting guess values for the fitting parameters, but these are specified in the 'Start Fitting Session' dialog, rather than in the function definition. In Origin, the reported χ^2 value is already normalized to χ^2 per degree of freedom.

All

In all cases the fitting programs do your error analysis for you automatically. The fitted values of the parameters are reported along with their propagated uncertainties. However, before you use these values uncritically, remember that these uncertainties are only valid to the extent that the residuals of the fit are randomly distributed *and* χ^2 per degree of freedom for the fit is approximately unity. See http://www.physics.hmc.edu/analysis/fitting.php for more information.

Experiment 3

RLC Resonance

Abstract

Current in a series circuit comprising a resistor, capacitor, and inductor tends to oscillate at a frequency that depends on the inductance L of the inductor and the capacitance C of the capacitor, although resistance damps out the oscillations. You will study how the oscillating current generated by an applied AC voltage depends on frequency, observing the maximum response (**resonance**) at the natural frequency set by L and C.

3.1 Overview

While the current flowing in a purely resistive electrical circuit is simply proportional to the applied voltage at any given instant, the responses of other electrical components can depend on how the applied voltage varies in time. In this experiment, you will explore the behavior of a circuit composed of two such components: an inductor (of inductance L) and a capacitor (of capacitance C), along with a familiar resistor (of resistance R). You will study the current generated in your circuit in response to an applied voltage which is sinusoidal in time; in particular, you will study how the response depends on the frequency of the applied voltage.

A familiar mechanical analogy can be helpful in

thinking of this oscillating circuit. When you push a child on a swing, the amplitude of the child's motion builds up swing by swing; with each push you pump energy into the child-swing system. Unless you are warped, you tend to push the child at the **natural frequency**, which is the frequency at which the child swings by herself when nobody is pushing. You (probably) match your driving frequency to the natural frequency because that strategy yields the maximum amplitude of motion for your effort.

In this experiment, you get to be warped and push at all sorts of frequency, from the pathetically low to the ridiculously high. Between these extremes lies the
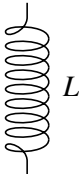
Element	Relation	Properties
R	$V_R = IR$	A **resistor** is a piece of poorly conducting wire, a pipe through which electric current flows. The magnitude of the current depends on the voltage across the resistor: double the voltage and you double the electric current. The mechanical analog of a resistor is a viscous liquid, like honey, through which you are trying to push an object. Resistors turn good, clean electrical energy into heat, a process called **Joule heating**.
C	$V_C = \dfrac{Q}{C}$	A **capacitor** is a pair of conducting plates facing one another across a gap. Electric charges flowing onto one plate produce an electric field in the gap between the plates, causing the voltage of one plate to differ from that of the other. The mechanical analog of a capacitor is a spring: the more charge on the capacitor, the greater the force required to keep it there. Note, however, that as the capacitance C increases, it becomes easier to pile up charge: a bigger capacitor behaves like a softer spring. As charge builds up, capacitors store energy in the electric field between their plates.
L	$V_L = L\dfrac{dI}{dt}$	An **inductor** is a coil of wire. As current flows in the coil, it generates a magnetic field inside the coil, which opposes a further change in the current. The mechanical analog of inductance is mass, which opposes a change in the state of motion. The greater the inductance L, the more the inductor opposes a change in the current that flows through it. Inductors store energy in the magnetic field generated by the current flowing through them.

Figure 3.1: Basic properties of resistors, capacitors, and inductors.

natural frequency, the one the system likes the best, as judged by the amplitude of the motion it experiences. In the case of the swing, it's the child that moves. In the *RLC* circuit of this experiment, it's the electric charge that 'swings' back and forth. To help you gain an intuition for the three circuit elements involved, they are introduced in Fig. 3.1.

1. Before lab — Read this chapter of the lab manual. The theoretical introduction is significant, so devote adequate time to understanding it. It will make your experience in lab much easier and more productive. If needed, you may read HRK Chapter 37 for more details.

2. Choose a specific experiment option in consultation with your instructor. If you choose the standard option, you will study a series RLC circuit, observing the current generated by an oscillat-

ing applied voltage. You will observe the phenomenon of **resonance**, in which the maximum response occurs at the natural frequency set by L and C. You will measure the quality factor of the resonance to observe how it depends on the resistance R. You will determine the true values of L and C from observing the behavior of the circuit, and you will demonstrate your understanding of series RLC circuits by identifying several properties of a hypothetical 'mystery graph' generated for that purpose.

3. You must analyze the results and present a rigorous discussion of your findings. Often, the experience gained (and mistakes made) during the first week's meeting will allow you to be that much more efficient and effective in completing the experiment the following week.

3.2 Theory

Resonance is a ubiquitous and very important phenomenon in nature; it is not restricted to electronic circuits. "Pumping" a swing, blowing on a flute, and tuning a radio are all examples you are familiar with, although you may not have realized you were "resonating" at the time! The key feature of all resonance phenomena is that the size of the effect (for example, the height the swing goes or the volume of the sound) depends on the frequency of the excitation. Several Nobel Prizes in physics have been awarded over the years for studies of resonance — for example, nuclear magnetic resonance, which is the basis of magnetic resonance imaging, or MRI, an important medical tool of recent vintage. In this experiment you will study the phenomenon of resonance in a simple AC (alternating current) electrical circuit.

3.2.1   Why does it oscillate?

The electrical circuit you will study comprises a resistor, a capacitor, and an inductor connected in series. To understand why it likes to oscillate, imagine first a simplified circuit in which the two ends of an inductor are wired to the two leads of a capacitor, as shown in Fig. 3.2. Now imagine transferring some charge from one plate of the capacitor to the other and "letting go." What will happen? Well, initially there are a bunch of negative charges on one of the capacitor plates, all repelling each other, and an excess of positive charge on the other plate, doing the same. These charges create an electric field in the region between the capacitor plates, and this field stores energy U_E, as shown in Fig. 3.2 (a). The mechanical analog of this system is a mass attached to the spring, and the starting point corresponds to when the spring is maximally compressed.
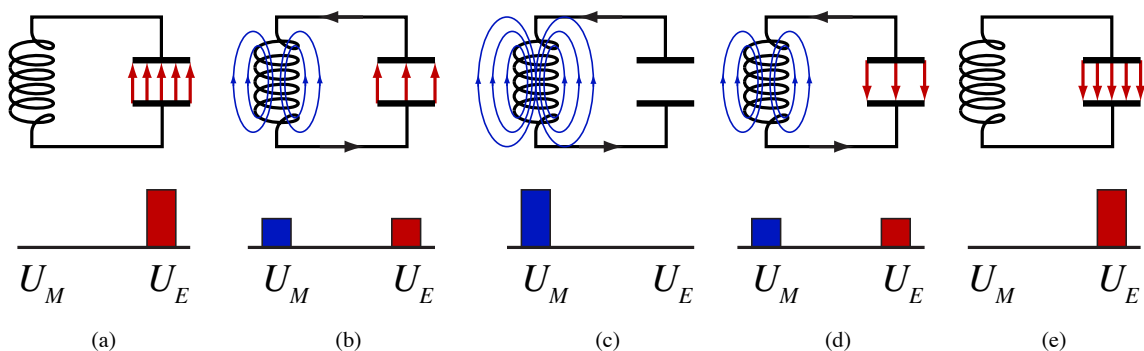

Figure 3.2: An inductor and capacitor circuit. The potential energy oscillates between the electric field of the capacitor (U_E) and the magnetic field of the inductor (U_B) as the charges oscillate between the top and bottom plates of the capacitor.

The electrons want to get away from one another by leaving the plate, but they can't jump across the gap to arrive at the positive plate. However, the voltage difference between the plates provides the incentive for the charges to seek the other plate by flowing out the lead and through the inductor, which is made of wire, after all. The wire of the inductor is wound so that as current flows through it, an appreciable magnetic field is established. That field stores energy, and this comes from the potential difference that drives the current. So, as the charge starts to leave the capacitor plate, two things happen. First, it reduces the electric field between the plates, which lowers the energy stored in the capacitor. At the same time it builds up a current, which stores energy in the inductor (Fig. 3.2 (b)). If there is no resistance, then the energy lost by the (electric field of the) capacitor is gained by the (magnetic field of the) inductor. This is analogous to the transfer of elastic energy stored in the spring to the kinetic energy of the mass.

After awhile, all the excess charge has moved from one plate of the capacitor to the other, so the voltage across the capacitor vanishes. Now all of the initial energy is stored in the magnetic field of the inductor and the current is at its maximum (Fig. 3.2 (c)). The current can't suddenly drop to zero, since then both the electric and magnetic energies would vanish and we would have lost the initial energy stored in the capacitor. So, it must keep flowing, just as the mass keeps moving as it passes the equilibrium point of the spring, where it has maximum kinetic energy. As the current continues to flow it piles up charge on the opposite capacitor plate, which produces an electric field that increasingly opposes the current, just as the spring now tends to slow the mass down (Fig. 3.2 (d)). Eventually, the current comes to zero, and we find that all the excess charge has been transferred from one plate of the capacitor to the other, as shown in Fig. 3.2 (e). (The mass has come to rest when the spring is maximally stretched.)

We are now half way through the cycle, and it is not too hard to imagine what will follow. The charge on the capacitor will begin to flow back through the inductor, trading stored electrical energy for stored magnetic energy until the current is maximum (and in the opposite sense from before), at which point the current will begin charging the plates of the capacitor to return them to their original condition. The energy of the system oscillates back and forth between the capacitor and the inductor as the charge oscillates between plates of the capacitor. And just as the natural frequency of oscillation of a mass-spring system depends on both the stiffness of the spring and the magnitude of the mass, the natural oscillation frequency of the LC circuit depends on the inductance L and the capacitance C.

What happens when we include a resistor in the circuit? Resistors convert electrical energy into heat, taking electrical energy out of the circuit and causing damping. Just as friction causes the mass to come to rest, resistance causes the current oscillations to die down. To produce a steady-state oscillation it is necessary to supply energy to the circuit to replace that dissipated in the resistor each cycle. This energy will be supplied by a function generator, which produces a sinusoidally oscillating voltage at a frequency you can adjust.

To understand the behavior of the current and voltage in an RLC circuit in mathematical detail, we begin with the basics of elements in an AC circuit, and how their frequency-dependent behavior is described mathematically.

Briefly, the three topics we should understand for the purposes of this experiment are:

- how we can represent AC voltages as complex numbers (what are amplitude and phase?)

- the relationship between voltage and current for each circuit element, which is called *impedance*

- how to analyze a circuit with more than one element (how to apply the Kirchhoff laws)

3.2.2 AC voltages as complex numbers

A steady AC voltage is represented as a sinusoidal function of time, with a fixed amplitude, frequency and phase shift. We will call the the amplitude V_0, the phase ϕ, and the angular frequency ω (*ordinary* frequency ν is $1/T = \omega/2\pi$). Therefore, we write the time-dependent voltage as

$$V = V_0 \cos(\omega t + \phi) = V_0 \, \mathrm{Re}\left[e^{i(\omega t + \phi)}\right] \qquad (3.1)$$

In the latter half of the above equation, we make use of Euler's identity to express the sinusoid in terms of complex exponentials. In this notation the imaginary unit is i and the real part of the expression gives us the cosine part.[1] At first this seems an unfriendly complication, but rest assured that the exponential expressions are much easier to manipulate algebraically than are sines and cosines. In practice, *we typically leave the taking of the real part as tacitly understood,* and just leave the expression in its complex form: $V_0 e^{i(\omega t + \phi)}$.

If we assume that the only time dependence in the system is in the sinusoidal voltage at ω (that is, the system is in steady state), our expressions simplify even further. Then without any loss of generality we can simply factor out the time dependence in the expressions, and focus our attention on the amplitude and phase of the voltage. Schematically we have replaced real voltages with a complex number as in,

$$V = V_0 e^{i\phi} \qquad (3.2)$$

---

[1] Note that engjneers sometjmes use $j = \sqrt{-1}$ as the jmagjnary unjt, presumably so jt wjll not be confused wjth electrjcal current.

We can perform all needed calculations using this complex representation. When working with a quantity like (3.1) that has explicit time dependence, the observable instantaneous voltage is always obtained by taking the real part. When we leave out the time dependence as in (3.2), the absolute value of the complex number tells us its amplitude, while the argument or phase tells how much the sinusoid lags or leads in its cycle relative to the cosine function.[2]

This algebraic representation of the voltage has a geometrical analog called a phasor diagram. A phasor is a vector in the complex plane whose magnitude is the amplitude of the voltage and whose direction gives the phase.[3] For example, we show the phasor diagram for an AC voltage in Fig. 3.3. In the illustrated case, the angle the phasor makes with the real axis is just ωt; if the waveform were delayed in phase by an amount ϕ (as it is in (3.1)), then this angle would be $\omega t - \phi$.
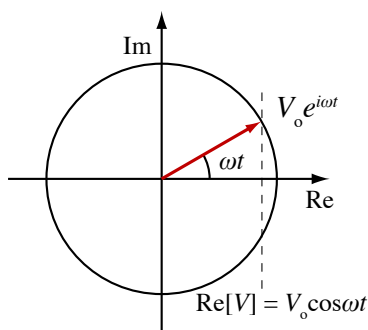

Figure 3.3 Phasor representation of a sinusoidally varying voltage at time t. The actual measured voltage is the projection of the phasor on the real axis.

The locus of points traced out by the phasor as a function of time is a circle as shown in Fig. 3.4. The true voltage oscillates between $\pm V_0$.
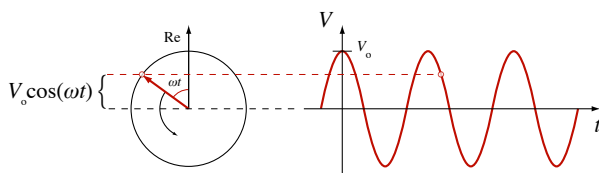

Figure 3.4 Time dependence of relationship between phasor representation and real observed sinusoidal voltage. Notice that the real axis has been rotated to the vertical to permit a more direct comparison with the $V(t)$ graph.

Now consider two AC voltages, both at the same frequency, as shown in Fig. 3.5. The voltages have different phases. In Fig. 3.5 (a), the phase difference is shown as the angular separation of the phasor vectors.

In Fig. 3.5 (b) this phase difference is appreciable in the different instantaneous magnitudes of the two voltages, even though they have the same amplitude.
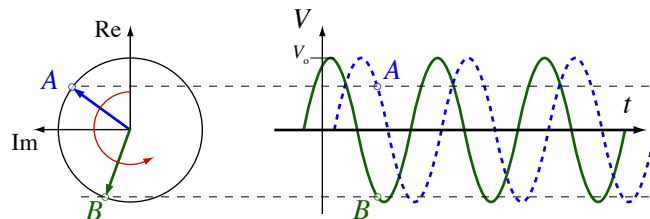
Figure 3.5 Relationship between phasors of different phase. Which phasor/voltage leads the other?

3.2.3 Complex impedance

Impedance is just a slight generalization of the concept of resistance, with which you are already familiar in the context of Ohm's law, (1.1). Just as in the case of the resistance of a resistor, impedance is the ratio of voltage to current for a given element, be it a resistor, capacitor, or inductor. In general, this differs from resistance because the ratio of voltage to current may depend on frequency, and because there may be a phase difference (lag or lead) between the sinusoidal voltage and current. For this reason, the impedance of a circuit element is in general a number with both magnitude and phase (a complex number) and depends on frequency (rather than being a fixed constant). In general we define impedance Z by the relation, $V = IZ$, where all the quantities may be complex. If for example V and I have different phases, then

$$V = V_0 e^{i\phi_1} \quad \text{and} \quad I = I_0 e^{i\phi_2} \tag{3.3}$$

imply that

$$Z = \frac{V_0}{I_0} e^{i(\phi_1 - \phi_2)} \tag{3.4}$$

Impedance of a resistor

This is the simple and familiar case of Ohm's law. As we discussed above, for a pure resistor there is no phase lag or frequency dependence, so the impedance reduces to the simple (real) result,

$$V_R = IR \quad \text{implies} \quad \boxed{Z_R = R} \tag{3.5}$$

Thus the impedance for a resistor is just the same as its DC resistance.

Impedance of an inductor

---

[2]Recall that if a complex number $z = x + iy$ is not already expressed in polar form, its absolute value and argument are given by $A = |z| = \sqrt{x^2 + y^2}$ and $\phi = \arg(z) = \arctan(y/x)$. So $x + iy = Ae^{i\phi}$ in this example.

[3]The "arrows" Feynman uses in QED to discuss photons are phasors that represent the complex probability amplitude.

The relationship between the voltage and current for an (ideal) inductor is

$$V_L = L \frac{dI}{dt} \qquad (3.6)$$

Applying the complex representation of the AC voltage and current, this becomes,

$$V_0 e^{i\omega t} = i\omega L I_0 e^{i\omega t} \quad \text{or} \quad V_0 = i\omega L I_0 \qquad (3.7)$$

Thus the impedance for an inductor is

$$\boxed{Z_L = i\omega L} \qquad (3.8)$$

If the current is assumed real across an inductor (cosine-like) then the corresponding voltage is purely imaginary (sine-like), or different in phase by 90°. Also, we see that for low frequencies an inductor presents little obstruction to the flow of current, but at large ω it takes a large voltage to cause current to flow; hence, the inductor will then act somewhat like a large resistor, neglecting phase effects. We say that an (ideal) inductor is a "short circuit" at low frequency and an "open circuit" at high frequency.[4]

Impedance of a capacitor

The relationship between voltage and current for a capacitor is

$$V_C = \frac{Q}{C} = \frac{\int I \, dt}{C} \qquad (3.9)$$

Again, going to complex representation we find,

$$V_0 e^{i\omega t} = \frac{I_0 e^{i\omega t}}{i\omega C} \quad \text{or} \quad V_0 = \frac{I_0}{i\omega C}$$

Thus, the impedance for the capacitor is

$$\boxed{Z_C = \frac{1}{i\omega C} = -\frac{i}{\omega C}} \qquad (3.10)$$

We see that the voltage and current of the capacitor also differ in phase by 90°, but in the sense opposite to that of the inductor. The capacitor is also opposite to the inductor in that it presents the greatest obstacle to current flow at low frequencies, and little obstruction as ω grows large. A capacitor is an "open circuit" at low frequency and a "short circuit" at high frequency. Much insight into asymptotic behavior can be achieved by considering these components to be large or small resistors in the appropriate frequency limit.

These are the basic concepts to understand; the detailed use of these impedances to derive the behavior of an RLC circuit is found in Sec. 3.4.

3.3 Experimental Procedure

3.3.1 Equipment and software

Having covered the theory behind this experiment, we now turn our attention to the equipment you will use to carry it out. Although in principle you could measure all the needed voltage amplitudes and phases manually with an oscilloscope, this would be very tedious and limit the amount of data you could take in a reasonable time. Instead, you have an Igor-based computer data acquisition system which greatly simplifies the procedure.[5]

The computer acquisition system has four analog signal input channels; one for the driving voltage and the remaining three for the voltages across the three circuit components (R, L, and C). The application interface provides a virtual oscilloscope window, in which you can view up to all four waveforms simultaneously in real time. In addition to displaying the waveforms, the program displays the frequency of the applied sinusoidal voltage signal, as well as the amplitude and phase of the voltages across the resistor, inductor, and capacitor (measured with respect to the driving voltage). A consistent color scheme is used throughout: the drive signal is black, and the resistor, inductor, and capacitor are represented by red, green, and blue, respectively ($RLC \rightarrow$ RGB).

The application also produces plots of the amplitude/phase results, and exports the data for use in other graphing and analysis programs. Within the Igor application, you can view or print graphs of amplitude and phase vs. frequency on either linear or logarithmic axes. The application can superimpose theoretical curves for the RLC circuit [as described in equations

---

[4]The inductor(s) you will use differ from an ideal inductor in two important respects. First, because it takes many turns of wire to produce a significant inductance, the wire of the inductor is fairly fine. This means that it has nontrivial resistance, as shown schematically in Fig. 3.6. Second, the wires of the inductor are wound around a ferromagnetic core to increase the inductance. Because of hysteresis in the iron of the core, the inductor in this experiment responds somewhat nonlinearly to the current. The significance of this effect depends on both the amplitude and frequency of the drive voltage. Exploring this nonlinearity can be a fruitful avenue for a technical report. Talk to your instructor about this if you are interested.

[5] Igor is a powerful data acquisition, analysis, and graphics program that runs on both Macintosh and Windows. Unlike Kaleidagraph, Igor is entirely scriptable: virtually everything that can be accomplished using the graphical user interface can be programmed within Igor using a C-like programming language. The programs that acquire and analyze data for the RLC and Fraunhofer diffraction experiments produce Igor experiment files holding your data, graphs, and tables, as well as a history of the operations you conduct. You may use them to export data in text format for further analysis in Kaleidagraph, Origin, or other program, or you may analyze your data further within Igor. See http://www.physics.hmc.edu/analysis/Igor.php for further information about using Igor.

(3.16), (3.18), and (3.20)]. There is also an option to plot data and theory in phasor form on the complex plane. The plots can also be readily copied and pasted into a Word document to keep a running description of your work. More detailed step-by-step instructions on the use of the program are provided at the end of this section.

In addition to the computer equipment, you are provided with a variety of components and hook-up wires with which to construct your circuit and connect it to the computer. Figure 3.6 shows a schematic diagram of the arrangement of components for the standard RLC circuit measurement. Attention should be paid to observing the correct polarity in the connection of the computer input channels; if reversed, the resulting voltage will be negative and therefore 180° out of phase with theory. Also note that you are provided with four digital multimeters, which when hooked up in parallel with each channel provide valuable additional diagnostic information.
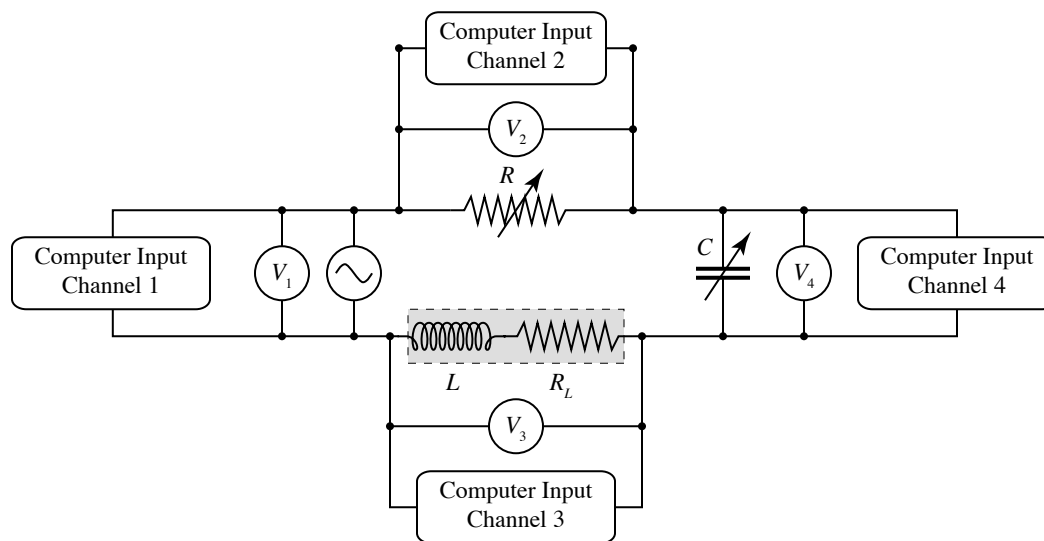

Figure 3.6: A series RLC circuit driven by an oscillating voltage. The voltmeters shown as circles are the digital multimeters. The boxes showing channels are the connections to the analog-to-digital converter in the computer allowing the computer to function as an oscilloscope. The (unavoidable) internal resistance of the inductor is shown explicitly.

Choosing an Experiment

This experiment provides you with the choice of several options to pursue. Depending on your particular interests and previous electronics experience, you may find one of the options more attractive than the others. The following description details the procedure for the standard experiment option, followed by descriptions of other possible alternative experiments. You will probably have time to do more than one.

(a) Standard option: series RLC circuit

Construct a series RLC circuit like the one shown in Fig. 3.6. The usual nominal values to use for the capacitance and inductance are $C = 5$ μF and $L = 500$ mH. With these values fixed, you will repeat the measurements for two different values of the resistance, $R = 50$ Ω and 500 Ω. Note that you can readily use one of your digital multimeters to directly measure the actual values of the resistances R and R_L. It is a good idea to do this before assembling the series circuit. However, the actual values of L and C may differ from their nominal manufacturer's values, so one of your objectives in this experiment will be to extract from your results refined estimates of these component values. **Pay attention to the polarity of the banana-to-BNC connectors: the outside conductor on the BNC cable goes to the marked (GND) side of the banana jacks.** What might happen if you reverse the polarity?

Start the oscilloscope in the Igor application and observe the four traces. Adjust the time per division so that you have two to five periods of oscillation for each curve displayed on the graph. Then click the **Save Traces** button and print a copy of the graph. Cut it out, tape it in your notebook, and then do the following:

1. From the plot, estimate the frequency and com-

pare with the value displayed on the function generator.

2. From the plot, estimate the phase of each of the colored traces and compare them with the values displayed on the Igor oscilloscope panel.

3. Sketch a phasor plot that represents each voltage (V_{app}, V_R, V_L, and V_C).

Now click the violet **¿?** button on the Igor oscilloscope. Two "manufactured" oscilloscope trace plots will be printed, one for each partner. These are voltages as a function of time, just like the experimental traces, but are calculated using "random" values for R, L, and C. It is also assumed that the inductor is ideal (*i.e.*, has no resistance). Eventually, each partner will need to analyze the traces of her or his plot, answering the questions listed below. However, it may be helpful to save this exercise until after you have taken data for the RLC circuit you have assembled. When you are ready, answer these questions about your mystery graph:

1. Determine the frequency of the curves.

2. Figure out whether the curve labeled α corresponds to V_R, V_L, or V_C, and explain your reasoning clearly. Similarly identify curves β and γ.

3. Determine whether the frequency is greater than, less than, or equal to the resonant frequency, and explain your reasoning clearly.

Using the Igor application, collect amplitude and phase data for each of the components. Vary the input frequency from the function generator over the range of roughly 20 to 900 Hz. You can monitor the data as you collect it on either a linear or logarithmic frequency scale in the data window (see Fig. 3.9). You can also superimpose theoretical curves. Unless you have made a mistake somewhere, your amplitude and phase results should appear quite close to the theoretical curves using the nominal values of R, L, C, and R_L.

Compare and contrast the results you get for the two different choices of the damping resistor, R. In particular, try to estimate the sharpness of the resonance peak with the "**quality factor**" defined as $Q = \omega_0/\Delta\omega$, where $\omega_0 = 1/\sqrt{LC}$ is the resonant (peak) frequency and $\Delta\omega$ is the peak's full width at half-maximum. A little algebra (see Sec. 3.4) shows that the quality factor of the voltage resonance should be related to L, C, and total resistance $(R + R_L)$ through

$$Q = \frac{\omega_0 L}{\sqrt{3}(R + R_L)} = \sqrt{\frac{L}{3C}}\frac{1}{R + R_L} \qquad (3.11)$$

Check the consistency of your experimentally determined values of Q and ω_0.

The culmination of your analysis will be in determining the best values of L and C consistent with your data. Although you could try a series of trial-and-error guesses for the values in question, this approach can be tedious and unwieldy. The preferred approach is to use "inverse-model" or fitting techniques. Igor is a very powerful data analysis environment with a very flexible nonlinear least squares fitting procedure built in. For the functions of the series RLC circuit, the necessary fitting functions are already defined for you. Using the **Fit** pop-up menu in the data window, you can fit each of the six data sets (amplitude and phase for the resistor, inductor, and capacitor) to equations (3.16), (3.18), and (3.20). Note that for many of these fits, the values of R and R_L are redundant: only their sum matters. You will need to hold fixed at least one of the resistances to obtain reasonable results.

The validity of the fitting procedure is an open question, for reasons that are discussed in Sec. 3.6. A good way to estimate the confidence with which a circuit parameter is determined by fitting is to compare the value obtained by fitting to different data. For example, do fits to the inductor phase and the inductor amplitude yield similar values for L? Are the residuals shown at the top of the fit window randomly distributed, or do they exhibit systematic errors? Explore your data and discuss your observations and conclusions in your notebook. Note that it is unnecessary to *print* each and every fit graph. Pick one or two representative graphs to print. Cut the two copies on the dashed line, tape a copy into your notebook, and write a brief discussion of its salient features and your conclusions about the extent of agreement or disagreement between your data and the linear theory of a series RLC circuit.

(b) Parallel RLC analysis option

This option may appeal to students who enjoy the mathematical or theoretical side of things. Construct a circuit as in the standard option, but instead of a series circuit, connect two of the three elements in parallel with one another and the third element in series with that pair. Which of the three elements, R, L, or C, to make the series element is up to you; making the resistor the series element $[R(L||C)]$ is a choice that leads to a particularly interesting resonance behavior. The theory of the parallel-series circuit is not derived in this lab manual, although the method to do so is essentially the same as in Sec. 3.4. The challenge of this option is to derive the appropriate equations describing your chosen circuit, and then compare them with experimental data you take for a single choice of R, L, and C. To keep the amount of work to a reasonable level, only derive the voltage amplitude and phase for the single series element.

After gathering data for your circuit, use the **Plot** button of the data window (Fig. 3.9) to extract and

plot either the amplitude or phase data from one of the channels. A new plot window with these data is created, which you can use for fitting. To begin the fitting procedure, with the plot window frontmost select **Analysis | Curve Fitting...** and be sure to check the **From Target** box. See the Igor web page at http://www. physics.hmc.edu/analysis/igor.php for more information on using Igor for fitting data. You may also export the data in text form for fitting in another program, such as Kaleidagraph. Click the **Save** button in the data window to do this.

One word of caution in constructing fitting functions for phases. Inverse trigonometric functions are multiple valued. You may need to add or subtract $180°$ in some cases.

(c) Unknown "black box" determination option

For those who find the standard experiment option a little too cut-and-dried, this option provides a little "mystery." Your instructor has a number of sealed boxes that contain a resistor, inductor and capacitor connected in series inside. In addition to the two end terminals (to which you connect the driving voltage), there is a pair of test terminals that gives you access to the voltage across one of the unknown components (you're not told which one of the three). In addition

to measuring the DC resistance across any pair of the four terminals, you gain information about this system by driving it with voltages of various frequencies, and using the computer system to measure and record the amplitudes and phases across the test terminals. Do not drive the system with voltages greater than 0.1 V or the inductors will not act as linear devices.

Using the methods and theory described above, you should be able to deduce the values of L, C, R and R_L inside the box, and which of the three components the test terminals are connected to. When using the computer to assist you in fits of your data, it is best to fix the value of R since you have measured it directly and the fit is not tightly constrained otherwise. You may have time to analyze more than one unknown box.
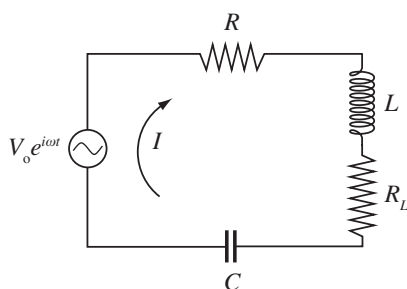
(d) Design your own experiment option

Using your own ideas and/or ideas from the above three options, you may want to devise your own custom experiment. It should use the basic methods and apparatus of the experiment as described above, but may emphasize some different aspect or circuit design that you find particularly interesting. Consult with your instructor for his or her approval of your experimental plan.

3.4 Analyzing a circuit with more than one element

The payoff for our development of the ideas of complex impedance is realized now in the ease with which we can analyze circuits consisting of combinations of these three elements. Series and parallel impedances add in a manner exactly analogous to DC resistances. As a result, analyzing a circuit composed of inductors, capacitors and resistors is no harder conceptually than dealing with the same circuit composed of all resistors.

As an illustration, we'll look at the series RLC circuit and compare it with a series circuit of three resistors, as illustrated in Fig. 3.7.

The analogous resistor circuit is quite easy to analyze. The current through each resistor must be the same, and follows from the series resistance of the three,

$$I = \frac{V_0}{R_1 + R_2 + R_3} \qquad (3.12)$$

Since the three resistors form a simple **voltage divider**, we can also easily write the voltage drop across each
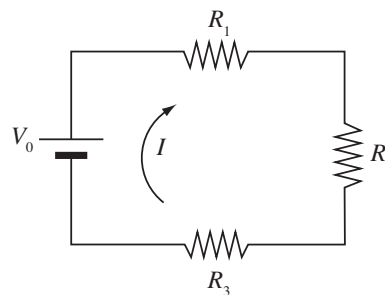
(a) AC circuit

(b) Resistor circuit

Figure 3.7: (a) An RLC circuit shown driven by an oscillating voltage. (b) An analogous DC circuit composed of three resistors.

one,

$$V_1 = V_0 \frac{R_1}{R_1 + R_2 + R_3}$$

$$V_2 = V_0 \frac{R_2}{R_1 + R_2 + R_3}$$

$$V_3 = V_0 \frac{R_3}{R_1 + R_2 + R_3}$$

Now we apply the same reasoning to the series RLC circuit. In this case the current through each element of the circuit is given by,

$$I = \frac{V_0}{Z_R + Z_L + Z_C} \tag{3.13}$$

Now the impedances are not all real, so we expect that the resulting current I will be some amount out of phase with the driving voltage. The voltages across each of the elements are now simply given by,

$$V_R = V_0 \frac{Z_R}{Z_R + Z_L + Z_C}$$

$$V_L = V_0 \frac{Z_L}{Z_R + Z_L + Z_C}$$

$$V_C = V_0 \frac{Z_C}{Z_R + Z_L + Z_C}$$

We must take note of the fact that by virtue of their construction as coils of wire, inductors generally have a non-negligible resistance, in addition to their designed inductive impedance. As a result, we have to replace the expression in (3.8) with one that takes the finite resistance of the inductor into account:

$$Z_L = i\omega L + R_L \tag{3.14}$$

In this expression, R_L is the resistance in ohms of the inductor that would be measured at DC ($\omega = 0$). Armed with these expressions, we can now predict all of the voltages and phases across every element of our circuit. For the **resistor**, we obtain,

$$V_R = V_0 \frac{R}{R + i\omega L + R_L + \frac{-i}{\omega C}} \tag{3.15}$$

which yields the amplitude and phase (with respect to the applied voltage),

$$|V_R| = V_0 \frac{R}{\sqrt{(R + R_L)^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}$$

$$\phi_R = \arctan\left(\frac{\frac{1}{\omega C} - \omega L}{R + R_L}\right) \tag{3.16}$$

For the **inductor**, we obtain,

$$V_L = V_0 \frac{i\omega L + R_L}{R + i\omega L + R_L + \frac{-i}{\omega C}} \tag{3.17}$$

which yields the amplitude and phase,

$$|V_L| = V_0 \left[\frac{R_L^2 + (\omega L)^2}{(R + R_L)^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}\right]^{1/2}$$

$$\phi_L = \arctan\left(\frac{\omega L}{R_L}\right) + \arctan\left(\frac{\frac{1}{\omega C} - \omega L}{R + R_L}\right) \tag{3.18}$$

And finally, for the **capacitor**, we obtain,

$$V_C = V_0 \frac{-i/\omega C}{R + i\omega L + R_L + \frac{-i}{\omega C}} \tag{3.19}$$

which yields the amplitude and phase,

$$|V_C| = V_0 \frac{1/\omega C}{\sqrt{(R + R_L)^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}$$

$$\phi_C = \arctan\left(\frac{\frac{1}{\omega C} - \omega L}{R + R_L}\right) - \frac{\pi}{2} \tag{3.20}$$

These derived expressions for the amplitude and the phase lag depend on frequency, and contain the mathematical essence of the resonance phenomena you will be observing. Although it will be up to you in the experiment and analysis to explore these functions, we present in Fig. 3.8 some typical plots to show you what you might expect.

In pondering these equations and their graphs, you should ask yourself the following questions:

- What is significant about the system's response at the **resonant frequency**, $\omega_0 = \sqrt{1/LC}$?

- Why do V_L and V_C approach the limits of zero and V_0 at opposite extremes?

- What influences the width of the peak in V_R, and the "overshoot" in V_L and V_C?

- Why do the phases always differ from each other by 90°, *except* for ϕ_L as $\omega \to 0$?

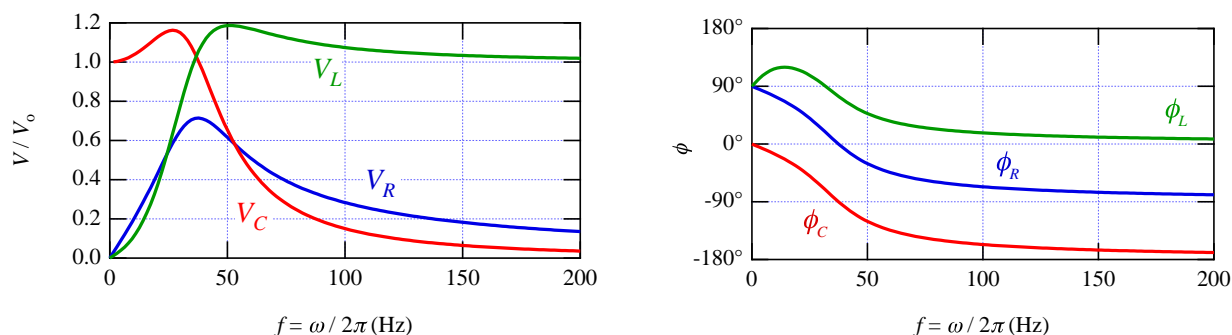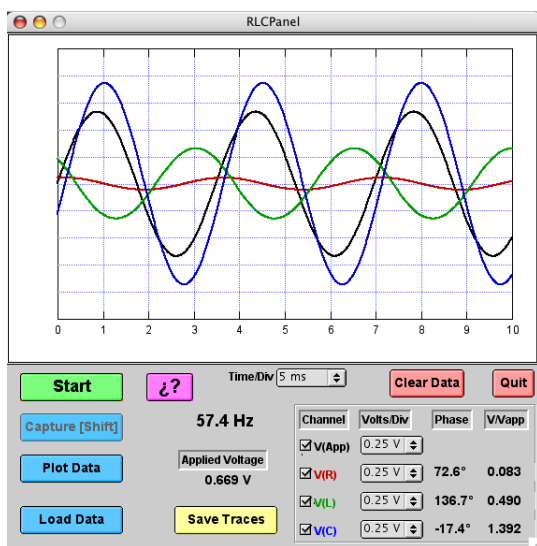- How is this system analogous to resonance of a mechanical spring-mass-damper system?

Figure 3.8: Example *RLC* voltage amplitudes as a function of frequency. For the illustrated case, $R = 100\ \Omega$, $L = 0.6$ H, $C = 30\ \mu$F, and $R_L = 40\ \Omega$.

(a) Main oscilloscope window.
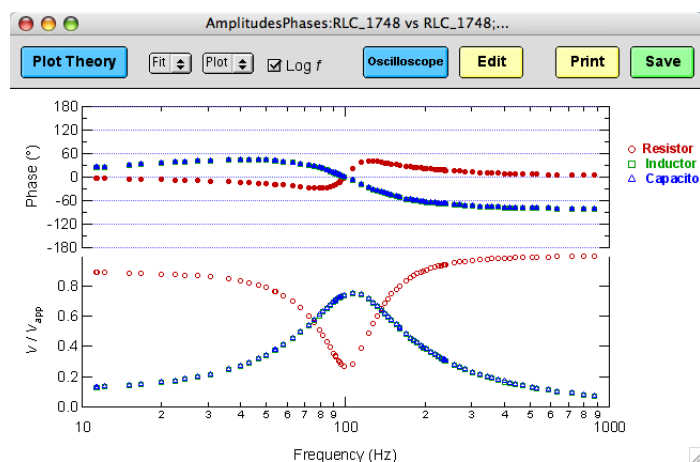
(b) Data window.

Figure 3.9: (a) Oscilloscope window of the *RLC* program. (b) Sample data taken with the *RLC* program.

3.5 Instructions for Using the Igor RLC Application

In the not-so-distant past, students took data on *RLC* circuits "by hand," using an oscilloscope and AC voltmeters. The Igor *RLC* application (written by HMC student Andy Fischer ('06) and Peter N. Saeta) is designed to acquire the data for voltages as a function of time for V_{applied}, V_R, V_L, and V_C. It will also calculate the normalized voltages and phase relative to V_{applied}, and plot them on two types of graph: V or ϕ versus frequency and a phasor plot. It takes care of a lot of measurements, allowing you to focus more on the circuit behavior. **However, you must pay careful attention to the circuit behavior as shown on the computer-based oscilloscope display as you take data, or you risk recording spurious data.** Remember: you do the experiment, not the computer.

Warning: Do not tamper with the data acquisition computer.

Warning: The program that runs this experiment sometimes loses mouse clicks while it is processing the data. Hold down the Escape key to stop the oscilloscope if your mouse clicks are getting lost.

The basic operation of the program is as follows:

1. Launch Igor (if it is not already running). Click on the Igor icon in the Dock at the right of the screen. It looks like this: .
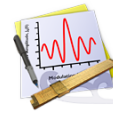
2. Click the button to run the *RLC* program.

3. The main window features an oscilloscope window, several buttons and popup menus to operate the oscilloscope, and indicators that report

frequency, amplitudes, and phases, as shown in Fig. 3.9. Before starting the oscilloscope, make sure that you have finished wiring the circuit, connected the signals to the inputs of the data acquisition box, and established a sinusoidal input voltage with a frequency in the range 50 Hz $< f < 200$ Hz and amplitude less than 1 V. Then click the **Start** button.

> **Warning:** **There is a delay of about 15 seconds the first time you start the oscilloscope as the National Instruments driver software is loaded into memory and initialized. Please be patient; it only happens the first time.**

4. The oscilloscope trace updates only a few times per second, much more slowly than a real oscilloscope. However, the Igor virtual oscilloscope also reports the frequency, amplitudes, and phases of the signals at the R, L, and C inputs to the ADC (analog-to-digital converter). *Note that the amplitudes are measured with respect to the drive voltage.*

5. For the program to determine the frequency accurately, aim to have 3 to 5 cycles on the screen. You can change the time base with the popup menu below the time axis. As you tune the frequency, the voltage across one or more components will decrease. Use the channel's gain setting to boost or decrease the signal level to allow the program to determine the signal's phase accurately.

6. Compare the frequency reading on the function generator to that on the computer oscilloscope; they should agree. Also compare the amplitude of the applied voltage to that shown on the multimeter. Note that the multimeter reports the root-mean-square voltage, which is smaller than the peak voltage by a factor of $\sqrt{2}$.[6]

7. To acquire data from the oscilloscope while it is running, either click on the **Capture** button or hold down the **Shift** key until the point appears on the amplitude/phase plot. This plot, shown as (b) in Fig. 3.9, appears automatically when you first capture a data point. If it is closed or hidden inadvertently, you can bring it forward or regenerate it by clicking on the **Plot Data** button of the main oscilloscope window.

    Mouse clicks are sometimes lost while the oscilloscope is running; you can use the Shift key for

taking data points and the **Escape** key to stop the oscilloscope. Note that you can take data continuously by holding the shift key down as you *gradually* change the frequency. If you go too fast, your data will suffer because the frequency changes during the measurement. You may also overload one of the channels.

8. To clear all data and start over, click the **Clear Data** button. Note that this does not save the data first; all unsaved data will be lost.

9. To remove points from the data, or to examine the raw numerical data, click the **Edit** button in the Amplitudes and Phases window. A spreadsheet window opens with one row for each data point. To delete a point, click in the "Row" column at the left to select the entire row. Then use the **Data | Delete Points. . .** command to delete the point(s). Note that there is no "undelete" function. Any graphs that depend on these data will be updated automatically.

10. To add theoretical curves to your data, click the **Plot Theory** button in the amplitude/phase window. Enter the circuit parameters. The expected curves (taking into account the inductor resistance) are shown superimposed on the data. (Note that this assumes a series RLC circuit. If you have constructed a different circuit, you can use the **Plot** menu to prepare a graph of a particular data set, on which you can superimpose a curve of your own specification. See below.)

11. To fit your data to the expected theoretical expressions, select the data set from the **Fit** pop-up menu. Enter initial guesses for the circuit parameters, and check any boxes for parameters you wish held fixed during the fitting procedure. Note that for many fitting operations, the value of R and R_L are redundant: only the sum of these resistances matters to the fit function. In such cases, you should make sure to hold at least one resistance value fixed.

    The fit results are shown on the plot window. Beside each parameter is either an open square (meaning the parameter was varied during the fit) or a filled square (meaning that it was held fixed). Note that the curves of the main data window are updated to reflect the fitted values of the parameters.

    A more complete summary of the fit results is printed to the **History** window at the bottom of

---

[6]The root-mean-square (rms) voltage is defined by

$$V_{\mathrm{rms}} = \sqrt{\langle V^2 \rangle} = \left[\frac{1}{T}\int_0^T V^2(t)\,dt\right]^{1/2} = \left[\frac{1}{T}\int_0^T V_0^2\cos^2(\omega t + \phi)\,dt\right]^{1/2} = \left[\frac{V_0^2}{T}\int_0^T \cos^2(\omega t + \phi)\,dt\right]^{1/2} = \frac{V_0}{\sqrt{2}} \qquad (3.21)$$

where T is the period.

the screen. The reported value of χ^2 (`V_chisq`) is not reduced. The fits use the "measured uncertainties" (see below), but rarely produce a value of χ^2 consistent with a good fit. Inspect the panel of residuals at the top of the plot window to investigate the validity of the fitting procedure. You should also study the variation in fitted parameters when you hold different parameters fixed and when you fit a different curve of the data set.[7]

12. When you have finished acquiring data for a given circuit, click the **Save** button of the data window. This saves (1) the complete Igor experiment, (2) the current batch of data, setting its name based on the current time, (3) a tab-delimited text version of the data with the same time stamp, and (4) an Igor binary version of the data with the same time stamp. The text version can be read by any data analysis program on any platform. Note that an Igor experiment file can hold several data runs; you should only need to save a single experiment file.

The text file format is tab-delimited text, which can be opened by any data analysis package. The first row holds column labels, and subsequent rows hold the data. There is a potential gotcha with text files: different operating systems use different characters to indicate the end of a line. UNIX computers use the linefeed character (LF = ASCII 10), Macintosh computers use the carriage return character (CR = ASCII 13), and DOS/Windows computers use both (CR+LF)! If your data analysis program has trouble reading the file, you can use Excel to open it first and handle the conversion of line terminators.

There are many ways to transfer your data from the lab computer to your own. Perhaps the easiest is to use a web-mail service (such as gmail or HMC's web mail interface at http://www.hmc.edu/webmail/). You can then attach your data file(s) to the message and mail it to both partners. Alternatively, you can mount your Charlie account. Near the bottom of the Dock (at the right of the screen) is an icon of a sphere inside a cube. Click that icon and enter your login information on Charlie. If someone has thoughtfully removed that icon, click on the Desktop to activate the Finder and select the **Go** | **Connect to Server. . .** command. Enter **afp://charlie.ac.hmc.edu** and click **Connect**. Once your class's Charlie volume mounts, select your personal folder and copy the files you wish to your Charlie account.

When you have finished, unmount Charlie by selecting the icon for your class's volume on the Desktop (not the one in the Dock you used to mount Charlie in the first place!) and using the **File** | **Eject** command.

13. If you would like to try your own fitting function (which is necessary if you have set up a circuit other than a series RLC circuit), use the **Plot** pop-up menu to copy the appropriate data and produce a plot window of it. Then select **Analysis** | **Curve Fitting. . .** and be sure to select the **From Target** check box. This limits the data shown in dialogs to that already in the figure. Click the **New Fit Function. . .** button and enter your fit function. Proceed sequentially through the tabs at the top of the dialog window, entering information as necessary. In particular, make sure you enter reasonable parameter values. You may click on the **Graph Now** button to see how close the data are to the curve computed with the values you have entered. When you are satisfied that it is close, click the **Do It** button to complete the fit. Results will be shown in the **History** area at the bottom of the screen.

14. The **Load Data** button allows you to restore a previously saved data set (it must have been saved within the same Igor experiment, however). Select the data set you wish from the pop-up menu and it will be restored to the main data window.

3.5.1   Troubleshooting

The RLC program is not perfect and needs you to make sure that you feed it appropriate signals to work properly. Here is a brief list of things to pay attention to.

- Make sure to have 3 to 5 full cycles displayed on the oscilloscope; the program needs these to determine the frequency and phase accurately. Change the time per division using the popup menu below the time axis.

- Make sure that the sensitivity of each channel is appropriate so that the voltage for that channel is neither too small nor too large.

- Monitor the shape of the sinusoidal curves at all times. Under certain circumstances you may drive the voltage across a circuit element beyond the range of the digitizer. You can correct this by reducing the drive voltage amplitude. In some circuits and frequency ranges, it may not be possible to obtain sinusoidal curves. You can document the curves you observe with the **Save Traces** button.

---

[7]If you leave multiple fit windows open at once, beware: the curves and residuals in all windows will update to reflect the fit parameter values obtained in the most recently performed fit. This behavior can actually be quite useful if you know it is happening, but misleading if you do not.

- The program does not gracefully handle frequencies above 1000 Hz or below 10 Hz.

- The function generator may require some settling time following a change in the frequency to

achieve steady state behavior. If you encounter this issue, adjust your data acquisition strategy accordingly.

3.6 Program Details

This section provides details about the internal working of the *RLC* program. Reading through this section will help you understand how the program does what it does, and how the data it takes might be less than ideal.

The voltages from the four input ports are sampled at a rate that depends on the setting of the **Time/division** pop-up menu. Because the channels are multiplexed and sampled sequentially, the nth data point from each channel does not correspond to the exact same instant of time. However, the phase delay from channel to channel is predictable, and the data have been corrected for this offset.

The maximum sampling rate of the National Instruments USB-6009 when recording 4 channels of data is 11 kS/s for each channel, although the program uses a maximum rate of 10 kS/s. This is insufficient to produce a smooth sinusoidal curve at frequencies greater than about 200 Hz. For display purposes, spline interpolation is used to produce a smooth curve through the data points, and to permit the temporal origin to be shifted (to remove the interchannel delay and to ensure that the phase of the drive voltage is 0).

3.6.1 Uncertainties

To determine the frequency, phase, and amplitude of each signal, the applied voltage is fitted first to determine the signal frequency and to establish the origin of phase. For this and subsequent fits, the *raw* data (not the spline-adjusted data) are used. This fit removes any DC offset in the data, uses as uncertainty for each data point half the voltage corresponding to a change in the least-significant bit of the analog-to-digital converter, and produces frequency, amplitude, and phase values for the drive voltage.

The data for each other channel is then fitted to a sine wave with a possible DC offset. The frequency is held fixed at the frequency determined from the drive voltage. The fitted amplitude and its uncertainty are divided by the amplitude of the drive voltage, and the phase value is shifted by the phase of the drive voltage. The recorded phase uncertainty is the unmodified value (that is, any uncertainty in the phase of the drive voltage is ignored).

This procedure does not account for potential non-linearities in the circuit, since it forces each signal to be a pure sine wave. The recorded uncertainties can thus be much smaller than warranted by the agreement between the fitted curve and the data, since these values presume agreement and a reasonable value of χ^2.

Abstract

Light bends on passing from one medium to another. By analyzing this refraction you will measure the factor by which light is slowed on entering acrylic or water. Curved surfaces cause the magnitude of the bending to depend on surface position, producing lenses which can be used to form images. You will study image formation using both converging and diverging lens systems.

4.1 Overview

The bending of light as it passes from one medium to another is known as **refraction**. This experiment consists of two independent investigations, both involving the refraction of light. You may choose one or the other. In the first investigation, titled 'Refraction,' you will measure various properties of light transmitted and reflected at a flat interface between two materials. In the second investigation, 'Lenses,' you will see how refraction through a curved surface causes light at different positions to bend differently, leading to the image-forming capabilities of converging and diverging lenses.
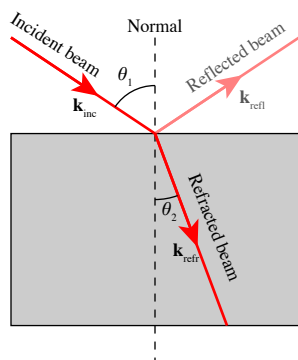

Figure 4.1   Refraction at the interface between two transparent media.

If you do the first investigation, you will study the relationship between the angle at which light arrives at an interface between two dissimilar media and the angle at which it leaves the interface on the opposite side. This relationship is known as **Snel's law**.[1] Light bends at the interface because its speed of propagation is different in the two media. The ratio of the speed of light in the vacuum to its speed in a medium is called the **index of refraction** n of the medium. For isotropic media,

this single number is all that is required to characterize the bending, although the index of refraction typically depends on the frequency or wavelength of light.[2]

Snel's law may be derived from Fermat's principle, which holds that the time taken by a light ray to pass between two points should be an extremum. As discussed in Feynman's *QED*, the probability amplitude that a photon will travel from point A to point B may be computed by summing its amplitude for traveling along *every* path from A to B. The "arrows" (probability amplitudes) corresponding to paths near the extremal path predicted by Fermat's principle all point in the same direction (add in phase), and so reinforce one another. Arrows from other paths tend to point equally in all directions, thus canceling each other out. This behavior leads to Snel's law of refraction. The index of refraction of each medium generally depends on the wavelength λ of the light, although this dependence may be weak. You may compare the refraction of different materials or the refraction by the same material at different wavelengths. You may also investigate the phenomena of **total internal reflection**, in which there is no beam transmitted through the interface and the reflection is perfect, and **polarization on reflection**, which occurs at the **Brewster angle**. You may also investigate how light refracts through a prism.

The second investigation focuses on lenses, both converging and diverging. Converging lenses can form **real images** of **real objects**; that is, the image of a tangible object can be viewed by placing a paper or screen "downstream" of the lens. Converging lenses are thicker in the middle than at the edges, which causes light rays to bend towards the axis of cylindrical symmetry (called the optic axis). **Diverging lenses** are thinner in the center and cause light to bend away from the axis of symmetry. If you look at something through

---

[1] Before you reach for your red pen, this is how Willebrord Snel spelled his name. See C. F. Bohren, *Clouds in a Glass of Beer: Simple Experiments in Atmospheric Physics* (Dover, New York, 1987/2001) 100. See also http://www.snellius.tudelft.nl/snelvanroyen.html.

[2] For anisotropic media, such as crystals, the index of refraction may be a more complicated beast. Not only does it depend on wavelength, but also on the direction of propagation and of polarization.

a diverging lens, it appears smaller. A diverging lens produces a **virtual image** of a real object, which means that the object is located on the "upstream" side of the lens. Therefore, you cannot view the image on a screen, since the screen would block the light on the way to the lens! However, using a combination of con-

verging and diverging lenses, you *can* produce a real image from the diverging lens, which you can view on a screen. You will use various lenses both alone and in combination to study image formation and to measure the focal lengths of the lenses.

4.2 Background and Theory — Refraction

The angles of incidence θ_1 and refraction θ_2 are illustrated in Fig. 4.1, and are related to each other by Snel's law

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \qquad (4.1)$$

where n_1 and n_2 are the **indices of refraction** of the two media. By definition, the index of refraction is the ratio of the speed of light in the vacuum to the speed of light in the medium. The refractive index of air may be taken equal to one to the accuracy of this experiment.

The law of refraction has an interesting history. The ancient Greeks knew a version of this law that worked only near normal incidence. Credit for the discovery of (4.1) is customarily given to the Dutch mathematician and scientist, Willebrord Snel van Royen (1580–1626), who was the son of a University of Leiden mathematics professor. He followed in his father's footsteps at the University of Leiden, focusing his research on mathematics, including trigonometry, cartography, and optics. He discovered the law of refraction in 1621,[3] but did not publish his work. The first published account was given by the French philosopher, mathematician, and scientist René Descartes (1596–1650) in *La Dioptrique* (1637). Descartes did not acknowledge Snel's work, although many believe that he had seen it. Notwithstanding this bit of awkwardness, the law of refraction is known as Descartes' law in French-speaking countries.

Something very interesting happens when the angle of incidence, θ_1, is large enough so that $n_1 \sin \theta_1 \geq n_2$. For Snel's law to hold, the sine of the refracted angle must exceed unity! This means that the angle of refraction is imaginary, and no light crosses the interface in this case: the beam undergoes **total internal reflection**. Knowing n_2, which for air may be taken to be unity, you can deduce the refractive index n_1 by finding the critical angle, which is the smallest angle θ_1 at which you observe total internal reflection.

Nearly two centuries after Snel's and Descartes' work, Etienne Louis Malus (1775–1812) discovered in 1808 that light reflected from a glass plate at a certain angle is polarized. Shortly thereafter in 1815, David

Brewster (1781–1868) worked out that when the angle between the reflected and refracted beams is 90°, the reflected beam is polarized with its electric vector perpendicular to the plane of incidence. The angle of incidence that achieves this condition is called **Brewster's angle** and is illustrated in Fig. 4.2.
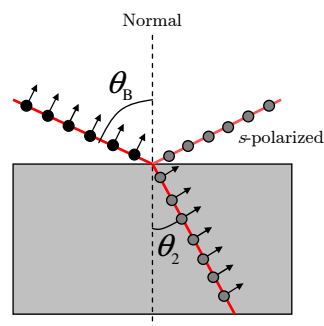

Figure 4.2 Here an unpolarized beam of light (shown with polarization arrows pointing in the plane of the page *and* out of the page) is incident on a transparent medium at Brewster's angle. The reflected beam is polarized with its electric vector perpendicular to the plane of incidence, which is the plane of the figure. This polarization is called *s*-polarization, where the *s* stands for the German *senkrecht*, which means "perpendicular."

There is a simple way to understand Brewster's condition. The reflected beam is generated by the oscillating dipoles inside the medium that are induced by the electric field of the refracted wave. However, no radiation is generated along the axis of a dipole. So when the refracted electric vector points in the direction of the reflected beam, no reflected wave is generated. In this case, all the light with an electric vector lying *in* the plane of incidence (*p*-polarized light) that is incident at Brewster's angle is transmitted through the interface. By combining Brewster's condition with Snel's law, you can show that

$$\tan \theta_B = \frac{n_2}{n_1} \qquad (4.2)$$

4.3 Experimental Procedures — Refraction

---

[3]M. Born and E. Wolf, *Principles of Optics*, 7th edition (Cambridge, 1999) xxvi.

In this part of the experiment you are provided with a helium-neon laser of wavelength 632.8 nm, attached to a circular fixture, the center of which is occupied by a semicircular piece of transparent optical material. The optical sample can be rotated so that the entering laser beam is presented at any desired angle of incidence to the flat face of the sample (away from the laser). A measuring tape has been affixed to the circular outside screen to permit you to measure the position of the reflected and refracted laser spots, from which you can infer the angles θ_1 and θ_2.

Before beginning to take data, verify the alignment of the beam through the center point of the sample piece, and re-check this centering periodically throughout the session. Think carefully about how you will convert distance measurements around the circumference of the measuring screen into the angles of reflection and refraction. As suggested by Fig. 4.3, knowing the angular position of the zero point and the reflected and refracted rays allows you to determine the angles θ_2 and θ_1. Prepare a spreadsheet to hold the raw position data, as well as to compute the angles and their uncertainties. An example is shown in Fig. 4.4.

One convenient way to check the alignment of the incident beam is to look for the faint reflection of the beam from the curved first surface of the sample; if the beam is entering the sample normal to its surface, this reflection will be directed straight back toward the source.

Warning: Do not look directly into the laser beam.

Before starting the experiment, describe the beam you see and give approximate dimensions. Use these observations to estimate the uncertainty in your angular measurements.

4.3.1 Measurements for Snel's law

Using the optical acrylic (looks like glass) sample, measure the angles θ_1 and θ_2 at several positions. Make enough well-spaced measurements that you produce a smooth graph in your spreadsheet and have a good basis for later performing a fit to determine the index of refraction, n, of optical acrylic. (If you are efficient, you should be able to perform that fit while you are still in the laboratory. As always, ask your professor for help with this, if needed.)

4.3.2 Analysis

There are two principal ways to analyze your measurements. Pick one or the other.

1. Plot θ_2 vs. θ_1 and fit directly to Snel's law, (4.1). Use your uncertainties $\delta\theta_1$ (and $\delta\theta_2$) on the plot and take advantage of a nonlinear fitting program. In this case, the program takes care of transforming the uncertainties for you.
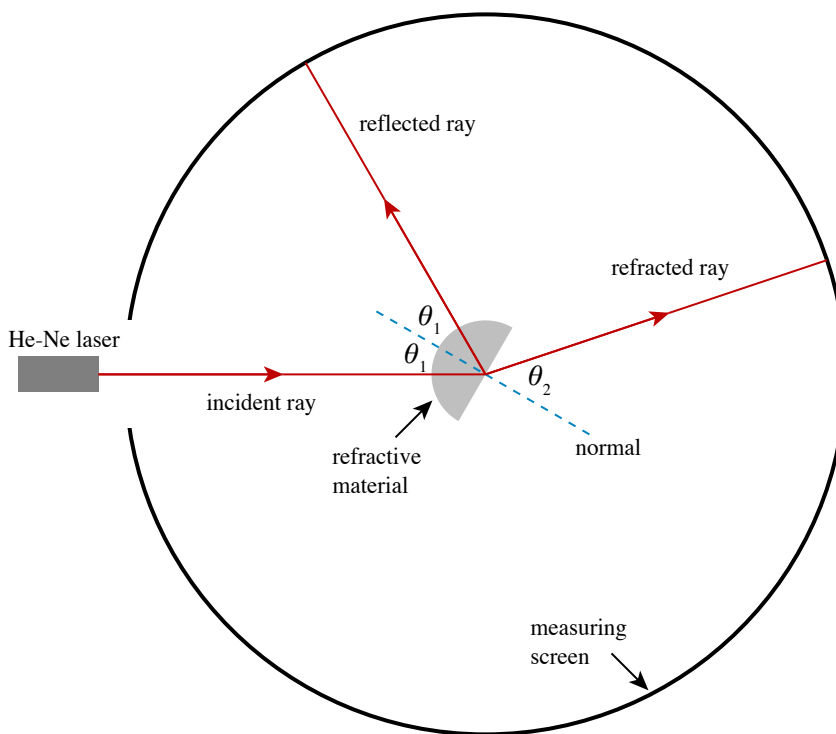

Figure 4.3: Index of refraction experimental setup.

Refl (mm)	Unc (mm)	Refr (mm)	Unc (mm)		θ_1	$\delta\theta_1$	θ_2	$\delta\theta_2$		n	δn
2738.5	1	987	2		0.071	0.003	0.048	0.002		1.463	0.060
2450	1	936	2		0.320	0.003	0.209	0.002		1.513	0.013
2503.5	1	945	2		0.274	0.003	0.179	0.002		1.520	0.016

Figure 4.4: A portion of a spreadsheet to record raw angular position data and to compute angles of reflection and refraction. A labeled diagram or other explanation in the lab notebook should define the measured quantities **Refl** and **Refr**, while sample calculations or symbolic formulas should show how the later columns of the spreadsheet are obtained.

2. Prepare a plot of $\sin\theta_2$ vs. $\sin\theta_1$. According to (4.1), you should obtain a straight line with slope equal to the index of refraction of the semicircular prism (either acrylic or water). Note that if you choose this option, you will need to compute $\delta(\sin\theta_2)$ from $\delta\theta_2$ via

$$\delta(\sin\theta_2) \approx \left| \frac{d\sin\theta_2}{d\theta_2} \right| \delta\theta_2 = |\cos\theta_2|\delta\theta_2 \qquad (4.3)$$

This expression is valid only when $\delta\theta_2$ is measured in radians. You should be able to explain why.

In both cases, note that fitting routines only account for the errors in the *dependent* variable (plotted on the y axis). If one variable has notably larger uncertainties than the other, put that one on the y axis and invert the fitting function, if necessary. For the curious, see http://www.physics.hmc.edu/analysis/fitting.php#ErrorsOnBoth for a method of accounting for *both* x and y uncertainties.

4.3.3 Further investigations

Once you have completed these measurements for the acrylic sample, there are many opportunities for further investigation. If time permits on the first day, select one of the following options:

1. **Index dependence on material**: Replace the sample with the liquid sample container filled with distilled water, or another liquid provided by the instructor. This will allow you to contrast the refractive indices of different materials.

2. **Index dependence on wavelength**: Replace the 633-nm (red) He-Ne laser with a 532-nm (green) diode laser. Repeat the measurements on the acrylic sample. This will allow you to observe the phenomenon of dispersion, which is the subtle dependence of the refractive index on wavelength. Note: this is a small effect and requires very careful measurements!

On the second day, you should explore total internal reflection, Brewster's angle and polarized reflection, and the refraction of an equilateral prism. You should describe carefully in your lab notebook (using both words and sketches) how you conduct each of these investigations.

1. Use the semicircular prism to investigate total internal reflection. When you have found the minimum angle the produces total internal reflection (and its uncertainty, of course!), use it to compute the index of refraction of the prism and compare the value you obtain to the one you got from analyzing your Snel's law data.

2. There are several pieces of sheet polarizer at the experimental station. Get a sense for what these do by viewing the room lights through them in various combinations and orientations, and by consulting your instructor. Then use the polarizers to investigate the polarization of the laser. Then see if you can identify Brewster's angle for a semicircular prism that you have already characterized. Does the angle you measure agree with (4.2)?

3. Replace the semicircular prism with an equilateral prism and prepare a diagram in your lab notebook of the path taken by the laser beam as it refracts through the prism. How does the angle at which the beam emerges from the prism depend on the rotation angle of the prism? From the observed refraction, determine the refractive index of the material of the prism.

4.4 Background and Theory — Lenses

Both the ancient Greeks and Romans knew that a curved glass surface could focus light. Aristophanes' play *The Clouds* (424 BCE) mentions a "burning glass," which was used to focus the sun's rays to produce fire. The behavior of such a convex lens is illustrated in Fig. 4.6. The Arabian mathematician Alhazen (965–
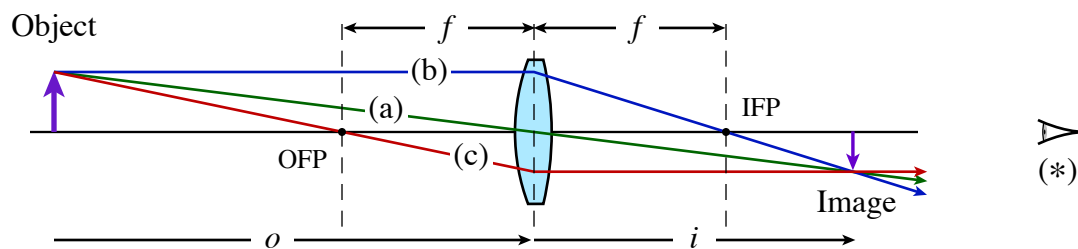
Figure 4.5: Formation of a real image by a converging lens. Light diverges from a point on the object at the left, which is a distance o from the lens. It passes through the lens and converges to form an inverted image of the point a distance i to the right of the lens. The three principal rays are shown: ray (a) passes through the center of the lens undeviated; ray (b) arrives at the lens parallel to the optic axis and leaves passing through the image focal point (IFP), a distance f from the lens; ray (c) passes through the object focal point (OFP) on the way to the lens, and emerges from the lens parallel to the optic axis.

1038) described how the lens in the human eye forms images on the retina, and the use of lenses in spectacles dates from the high Middle Ages in Italy.

What is an image? Briefly, it is a position at which light rays converge and from which they diverge. Looking at Fig. 4.5 we see that the image to the right of the lens is formed where the rays that diverge from a point on the object converge to a point again. If you were to place an opaque screen at this position and look at the screen from the left, you would see an upside-down version of the object. What if instead you put a thin tissue paper at this position and look at the tissue from the right? Again you would see the upside-down replica of the original object. Lastly, imagine removing the tissue and looking at the same position from the right. Your eye would still see the upside-down replica at the same position, except it would now be brighter and clearer. [*The above is critical to your understanding images; reread it until you understand!*]
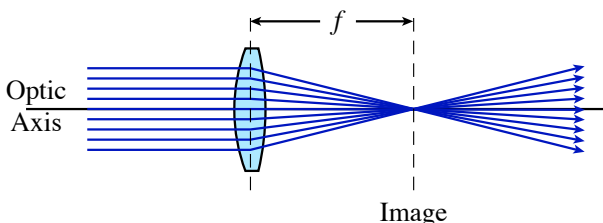

Figure 4.6   A parallel beam of light converges to a point at a distance f from a convex (converging) lens.

The general theory of lenses is quite complicated; lens design is both art and science.[4] However, in many circumstances a simple approximate theory suffices. We assume that the rays of light make only small angles θ

with the **optic axis**, which is the symmetry axis of the lens(es), so that we may neglect the difference between $\sin\theta$ and θ. This is called the **paraxial approximation**. We also assume that we may neglect the separation of the two surfaces of the lens (the thin-lens condition). Under these assumptions, the distance from the center of the lens to the object (o) and the distance from the center of the lens to the image (i) are related by the **thin lens equation**,

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f} \qquad (4.4)$$

where f is the **focal length** of the lens. These distances are illustrated for a converging lens in Fig. 4.5. If you'd like to play with an interactive simulation of image formation by a converging lens, try http://www.physics.uoguelph.ca/applets/Intro_physics/kisalev/java/clens/. (There's also a simulation for diverging lenses at this site.)

The focal length of a lens depends upon the index of refraction n of the lens material, and on the radius of curvature of the two surfaces of the lens, according to the **lens maker's formula**[5]

$$\frac{1}{f} = (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \qquad (4.5)$$

You can readily derive this formula by using some elementary trigonometry and Snel's law in the small-angle limit: $n_1\theta_1 = n_2\theta_2$. Because the index of refraction depends on the wavelength of light, a singlet lens focuses different colors slightly differently. You will have an opportunity to observe this effect by interposing colored glass filters in the beam from the source and looking for a subtle difference in the focal position of green and red light.

---

[4]Please note that the singular of *lenses* is *lens*, not *lense*.

[5]This formula was first obtained by the Italian mathematician Bonaventura Cavalieri (1598–1647) in the year of his death, 1647. For a beam of light propagating from left to right, the sign of a radius of curvature, r_n, is positive if the surface bulges outward to the left, and negative if it bulges outward to the right. Hence, for a biconvex lens, which has two curved surfaces and bulges in the center, $r_1 > 0$ and $r_2 < 0$.

Ray diagrams such as Fig. 4.5 and Fig. 4.8 are extremely helpful in understanding the behavior of lenses, and they are easy to construct. Here are the rules:

(a) A ray that passes through the center of the lens is undeviated.

(b) A ray that propagates parallel to the optic axis (the symmetry axis of the lens) emerges from the lens heading for the image focal point, a distance f downstream of the lens.

(c) A ray that passes through the object focal point a distance f before the lens emerges from the lens parallel to the optic axis.

The location of the image can be described in two ways. First, if a screen is placed at the location of the image, a focused picture of the object appears. Second, if you were to look at the system of the object and lens by putting your eye at the location marked with a star at the right of Fig. 4.5, you would see light rays appearing to originate at the location of the image. That is, the object would "appear" to be at the location of the image. As shown in Fig. 4.5, it would be inverted and it may be either magnified or reduced in size. You can readily show from principal ray (a) that the magnification is

$$\gamma = -\frac{i}{o} \tag{4.6}$$

The negative sign signifies that the image is inverted.

4.4.1 Diverging lens

Diverging lenses are thinner in the center and thicker at the edges; they divert incoming parallel rays away from the optic axis, as illustrated in Fig. 4.7. On the downstream side of the lens the rays appear to diverge from a point a distance $|f|$ behind the lens, which

is the location of the **virtual image**. It is virtual, not *real*, since you can't image it by placing a paper or screen there, the way you can with a real image produced by a converging lens. Because the focus happens on the upstream side of diverging lenses, their focal lengths are negative.

The same set of three principal rays may be drawn for a diverging lens to locate the virtual image of a real object, as shown in Fig. 4.8. The trick is to remember that the focal points have exchanged places. Therefore, a ray that arrives at the lens parallel to the optic axis emerges from the lens *as though it had come* from the image focal point, which is $|f|$ *in front of* the lens.
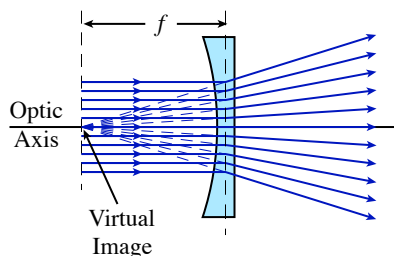
Figure 4.7   Formation of a virtual image from a collimated beam by a diverging lens. The image is virtual because it is on the "wrong" side of the lens; you can't put a screen there and see it without blocking the light on its way to the lens.

Similarly, a ray that is heading towards the object focal point a distance $|f|$ *beyond the lens* emerges from the lens parallel to the optic axis.

Just as diverging lenses have negative focal lengths, virtual images have negative image distances i. In the case illustrated in Fig. 4.8, the image distance i is equal to $-\frac{2}{3}o$.
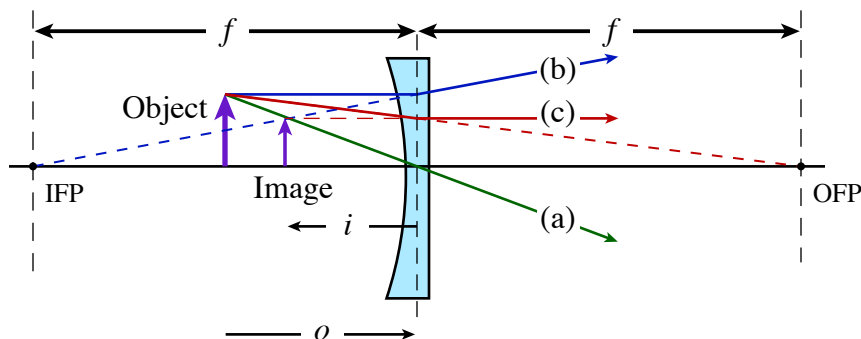
Figure 4.8: Formation of a virtual image by a diverging lens. The image is not inverted, is closer to the lens than the object, and is reduced in size. As in Fig. 4.5, the principal rays are shown: ray (a) passes through the center of the lens undeviated; ray (b) arrives at the lens parallel to the optic axis and leaves as though it came from the image focal point (IFP), which is $|f|$ from the lens on the upstream side; ray (c) is heading towards the object focal point (OFP), which is a distance $|f|$ downstream of the lens, and emerges from the lens parallel to the optic axis.

Since there are virtual images, could there be virtual objects? What would they look like? Well, light diverges from a real object, as shown at the left of Fig. 4.5. Light also diverges from the real image formed by the converging lens in that figure. What if you were to stick another lens downstream of that first image? The rays would diverge from the first image and propagate to the second lens; the image of the first lens serves as the object of the second lens. On the other hand, if you were to slide that second lens upstream of the real image, as illustrated in Fig. 4.9, the rays would encounter the second lens *before* they had finished the job of converging to produce the first lens's image (which is the second lens's object). Thus, they constitute a **virtual object** for the second lens. Not surprisingly, virtual objects have a negative object distance o.
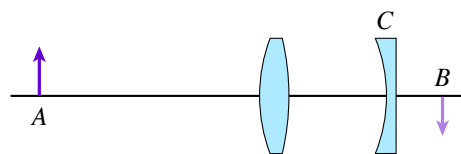

Figure 4.9   A converging lens forms an image of an object A at position B, which is downstream of lens C. The image at B is a virtual object for lens C.

Amazingly, the thin lens equation, (4.4), handles all cases for both kinds of lenses, provided you respect the sign conventions outlined above. For easy reference, here's a table summarizing them:

Qty	Positive	Negative
f	converging lens	diverging lens
o	upstream of lens	downstream of lens
i	downstream of lens	upstream of lens

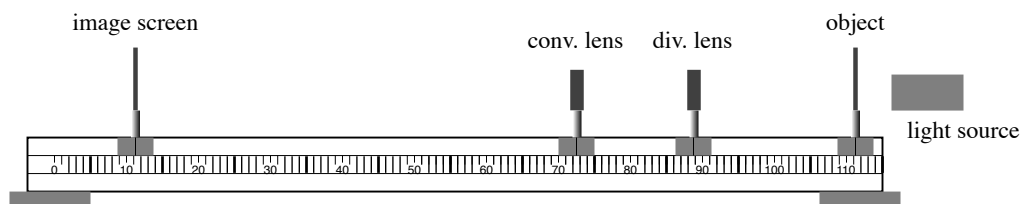## 4.5 Experimental Procedure — Image Formation

Figure 4.10: Layout of the optical bench and components.

4.5.1   Preliminaries

> **Warning:** **Do not touch the surfaces of optical components; handle all optics with care.**

Your lab station should have several converging lenses and at least one diverging lens, mounted in interchangeable protective holders. The lens mounts allow you to exchange lenses without having to realign. Select the converging lenses and make a quick estimate of their focal length by projecting an image of an overhead fluorescent light on the floor. You'll probably need to put the lens closer to the floor than about half a meter to see an image of the glowing tube. In this case, $i \ll o$, and if we neglect $1/o$ compared to $1/i$ in (4.4), then the distance from the lens to the floor is approximately equal to the focal length.

Take a look through each lens to get a feel for what it does to light coming from the opposite side, then select one of the converging lenses with a focal length less than 25 cm. You will work with it first. Throughout your work with the optical bench it is critical for you to record carefully exactly what you are doing. You will need to make many sketches, explain tersely how you made your measurements, and record uncertainties. Since you will be forming images on screens and measuring their positions, it will require some judgment on your part to determine where the position of best focus is. Typically the uncertainty you assign to your measurements will be influenced both by your ability to read the bench scale, and by the uncertainty inherent in estimating the position of best focus. Use your scientific judgment here and explain the rationale clearly in your lab notebook.[6]

---

[6]Beware the NFL absurdity: the referee carefully peels off gargantuan bodies as they wrestle for the ball, lovingly extracting the pigskin and setting it gently upon the field somewhere in the vicinity of where it might have been at the end of forward progress (but before the post-progress tug-of-war has moved the ball around). The chain gang then trots vigorously out, having carefully marked its

Sometimes your ability to find the position of best focus can be improved by adjusting the brightness of the object. For example, if the images formed exhibit flaring that makes it hard to judge sharpest focus, a piece of diffusing tissue paper between the illuminator and the object can dim the images enough to improve their visibility. You may also find it helpful and/or interesting to insert a colored filter to reduce chromatic aberration.

Record raw values first, then compute derived quantities. That is, if you are really interested in o, but you measure the position x_{obj} of the object and x_{lens} of the lens so you can compute $o = x_{\text{lens}} - x_{\text{obj}}$, record both x_{obj} and x_{lens} in your data table, as well as their difference.

First, set up the object, white image screen, and illuminator on the optical bench. Place the object at the end of the optical bench next to the illuminator and adjust it for some convenient height. The illuminator should be carefully adjusted to obtain bright images:

1. Adjust the height of the illuminator to the same height as the object and position it at the end of the bench.
2. Remove both the image screen and the object from the bench, and focus the filament of the lamp on the far wall (which is effectively at infinity) by sliding the part that contains the socket in and out.
3. Mount the image screen at the opposite end of the bench and adjust the tilt of the lamp so the beam is centered on the screen, at the same height as the illuminator.

4.5.2 Converging lens

Now replace the object and place the image screen at the far end of the optical bench. Place the converging lens with the shortest focal length between the object and screen, and align the system. Adjust the position of the lens to produce as clear an image as possible on the screen. How reproducible is the position of the lens? That is, if you slide the lens back and then reposition it to produce the clearest image, how much variation in the optimal position of the lens do you get from iteration to iteration? Do you and your partner agree? Does the sensitivity depend on where the screen is located?

Pick a particularly good data point and compute the focal length of the lens, as well as its uncertainty. You may find the following formulas helpful:

$$\delta\left(\frac{1}{x}\right) = \left|\frac{\partial(1/x)}{\partial x}\right|\delta x = \frac{\delta x}{x^2} \tag{4.7}$$

$$\delta\left(\frac{1}{f}\right) = \sqrt{\left[\delta\left(\frac{1}{o}\right)\right]^2 + \left[\delta\left(\frac{1}{i}\right)\right]^2}$$

$$\delta f = f^2\sqrt{\left(\frac{\delta i}{i^2}\right)^2 + \left(\frac{\delta o}{o^2}\right)^2} \tag{4.8}$$

Note that (4.8) assumes that the error in o is uncorrelated with the error in i, which may or may not be the case, depending on how you measure.

Repeat this procedure with the other converging lenses.

Even better measurements

Suppose that we plot a series of image and object distance measurements in the form of a graph with $1/o$ on the x-axis and $1/i$ on the y-axis. Rewriting the lens equation,

$$\frac{1}{i} = \frac{1}{f} - \frac{1}{o} \tag{4.9}$$

we see that if the thin lens formula is obeyed, our graphed data should describe a straight line with a slope of -1 and a y-intercept of $1/f$, as shown in Fig. 4.11. Choose one converging lens and determine the focal length in this way. Plan your experiment sensibly; which converging lens you use and what data points you take can significantly affect the quality of your result.

Notice that your data points up to this point all reside in the first quadrant of this plot. Your challenge, now, is to add *at least one point* in quadrants II and IV. After you have added these, you will be able to make a very fine fit for the focal length f of the lens. Note that you will need to use another converging lens somehow to obtain points in quadrants II and IV. If you are clever, you can introduce this second converging lens in such a way that you don't need to measure (or know) its focal length to obtain the desired points in quadrants II and IV. Be sure to document your work very carefully. If you are stuck, you might skip to the following section on the diverging lens, and then come back.
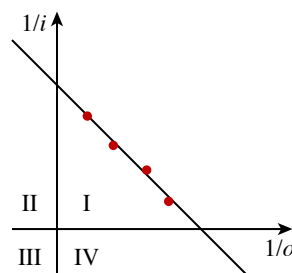
Figure 4.11   A plot of $1/i$ vs. $1/o$ should yield a straight line with slope -1 whose x and y intercepts are $1/f$.

---

location with respect to a yard line, and stretches the chain to see whether the nose of the ball has exceeded the requisite 10 yards. This final measurement is made with great precision as television cameras zoom in for exacting scrutiny. The moral: if you can't tell better than a foot where the lens should go, don't bother measuring its position to the nearest thousandth of an inch!
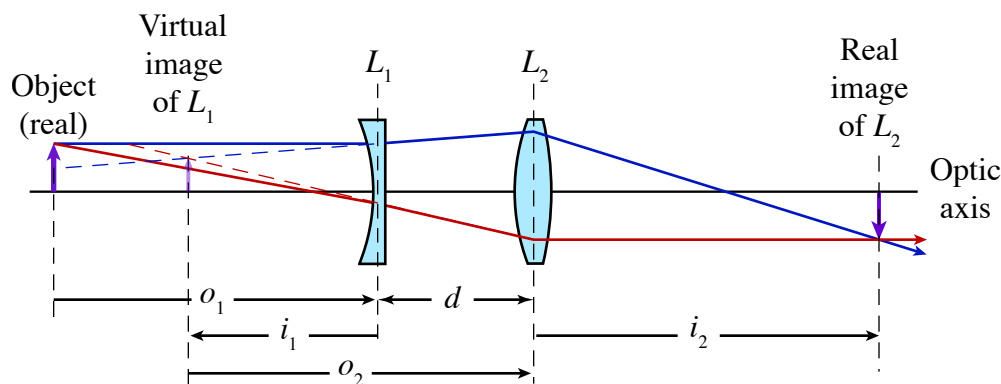
Figure 4.12: A method for forming a real image from the virtual image of a diverging lens. The diverging lens L_1 forms a virtual image of the real object at the left. It is a virtual image because it is upstream of the lens, so you couldn't visualize the image with a screen. Nonetheless, this virtual image then serves as the input object to converging lens L_2, which forms a real (inverted) image at the right. It is real because you could (and will!) place a screen at that position to observe the image.

4.5.3 Diverging lens

In this part of the experiment, you will investigate the properties of a diverging lens, again by forming images on a screen, this time using the diverging lens in combination with one or more converging lenses. A possible geometry is shown in Fig. 4.12. The virtual image of the diverging lens L_1 is located a distance $|i_1|$ upstream of the lens, and a distance $o_2 = |i_1| + d$ upstream of lens L_2.

How could you determine the position of the image of L_1? With both lenses and the screen in place to produce a sharp image of L_2, remove L_1. The image on the screen will disappear (or grow diffuse). Now slide the object forward until the image on the screen is sharp again. At this point, the object must be at the position where the virtual image of L_1 was located before L_1 was removed. Be sure you can explain why this procedure works.

4.5.4 Lens maker's formula

Shared between the two experimental stations is a diopter gauge (see Fig. 4.13), which provides an way of inferring focal length based on the lens maker's formula (4.5). The gauge places three round-ended rods against the lens, providing a measurement of the curvature of the lens surface. It then reads the refractive power of the lens, based on the measured curvature and an assumed index of refraction of 1.5. Refractive power is measured in diopters, with units of inverse meters (m^{-1}); the refractive power in $(\mathrm{m}^{(}-1))$ is equal to $1/f$, where f is the focal length in meters.

Figure 4.13 The diopter gauge, which allows you to measure the refractive power of a glass lens surface. It works by placing three round-ended rods against the lens surface. The two outer rods are fixed; the central one is spring loaded and operates the dial of the gauge. The black object is a cap to protect the rods. Please keep the gauge capped when you are not using it to measure a lens surface.

The diopter gauge assumes you are using it on the curved surface of a lens whose opposite surface is planar. All the lenses provided have a planar side and a curved side. The curved side is accessible to the diopter gauge, and the planar side is inset into the mount. Measure the curvature of the surfaces of the lenses you have investigated and compare the focal lengths you determine this way to the focal lengths you obtained previously. Be sure to use the positive (black) scale for converg-

ing lenses (which have convex surfaces and positive focal lengths), and the negative (red) scale for diverging lenses (which have concave surfaces and negative focal lengths).

> **Warning: Please handle the diopter gauge and lenses with care. Remember, you are testing theories of optics, not gravitation!**

4.5.5 Chromatic aberration

If time permits, consider investigating chromatic aberration by forming images with a single converging lens in combination with a colored-glass filter. Because the refractive index of glass decreases with increasing wavelength, the focal length of a singlet lens is shorter in the green than in the red. See (4.5). The change in focal length is fairly small—of order 1%—but if you're careful, you can observe it.

4.5.6 Depth of field

Another possible phenomenon you can investigate is **depth of field**. For a camera lens, this is the range of distances from the lens that can be imaged (sufficiently) clearly on the film (or sensor) plane. By adjusting the aperture of the lens using an iris diaphragm, which is an adjustable circular opening that limits the effective diameter of the lens, the photographer can blur the background (by using a large open aperture) or make everything sharp (by using a small aperture). The ratio of the focal length f of the lens to the open diameter ϕ is called the "f/number." A young Ansel Adams founded Group f/64 in 1932 as a statement of his æsthetic favoring razor-sharp "straight" photography.

A camera lens typically operates at high conjugate ratio, which means that the lens is much closer to the film plane than to the object being imaged. Select a converging lens of intermediate focal length and set up an arrangement with the screen near the far end of the rail, to allow you a large range of object positions. Form an image with a conjugate ratio of about 4. Mount an iris diaphragm on the object side of the lens and as close to the lens as possible and investigate how the open aperture affects the range of object positions that produce a well-focused image.

An alternative way to investigate depth of field, for those with access to a digital camera permitting manual control of the exposure settings, is by taking a sequence of exposures using a suitable target that allows you to figure out the range of distances that are in focus. Such an investigation would be appropriate for a technical report. Starting from Fig. 4.5 and the thin lens equation (4.4), you can develop a theory for depth of field to compare with your experimental observations.

Experiment 5

Fraunhofer Diffraction

Abstract

Diffraction refers to the fanning out of a wave upon encountering a constriction or obstacle. You will study Fraunhofer diffraction from single and multiple slits illuminated with a laser beam. Each point in the opening serves as a source of secondary waves, whose interference produces the diffraction pattern. Comparing the observed patterns to theoretical curves permits you to make very accurate measurements of the slit widths and slit separations.

5.1 Overview

If we take the quantum-mehcanical view of light as composed of photons, diffraction phenomena can again be derived, this time from the laws of quantum mechanics expressed for probability amplitudes. This derivation may be more familiar to you, although our experiment operates far from the single-photon regime. As discussed by Feynman in *QED*, the probability for a photon to pass from the source to the detector may be calculated by adding up the "arrows" (probability amplitudes) for each path from the source to the detector, and then taking the square of the resulting arrow (amplitude). For a plane wave normally incident on one or more slits, the arrows at the plane of the slits all point in the same direction (the amplitudes are in phase with one another), and we may consider this the starting point for the interference effects. (The alignment of the arrows everywhere in the slit plane is part of the simplification that constitutes *Fraunhofer* diffraction.) For a finite slit width, the distance from a point within the slit to the detector depends upon position within the slit.

Therefore, arrows arriving at the detector from different positions within the slit point in different directions and interfere with one another.

You will investigate the nature and origin of the diffraction patterns from a single slit and from various arrays of multiple slits. You should observe the patterns both qualitatively on a piece of paper and quantitatively using a photodetector, which is mounted behind a narrow window to yield a well-defined diffraction angle. By translating the window/detector with a micrometer, and by knowing the distance between the slit and the window, you can record the diffracted intensity as a function of diffraction angle. A computer data logging program will assist you in this effort, but you will have to supply the distances it needs to convert a voltage proportional to micrometer rotation into a diffraction angle. From this calibration, you will then be able to make a very precise determination of the slit width (and separation, in the case of multiple slits).

5.2 Background and Theory

The first mention of diffraction phenomena appears in the work of Leonardo da Vinci (1452–1519). Over a century later, a professor of mathematics and physics from Bologna, Francesco Maria Grimaldi (1613–1663), made more careful observations of the phenomenon by illuminating a thin rod by a "pencil" of sunlight admitted into a darkened room. He was surprised to see a shadow behind the rod that exceeded in width the geometric shadow, and that it was fringed with colored bands. Grimaldi coined the term *diffraction*, which

means "breaking up," to describe the phenomenon and he developed a theory that light was a fluid that exhibited wave-like motion. His observations were published soon after his death, and this work stimulated Isaac Newton (1642–1727) to study optics.

Newton took up Grimaldi's observations in the third book of his *Opticks* (1704). However, Newton's insistence on a corpuscular theory of light prevented him from developing a satisfactory theory of diffraction. A century later Thomas Young (1773–1829) reported a

---

[1]See Walter Scheider, *The Physics Teacher* **24** (1986) 217. Thomas Young's interests were many and varied; the fact that his contemporaries weren't particularly interested in his diffraction experiments was perhaps not terribly troubling to him. He took up other pursuits. A child prodigy, he was reading fluently at age two and by age 16 was proficient in Greek, Latin, and eight other languages. He put to good use this facility for languages by helping decipher the Rosetta stone, which had been discovered in 1799 by French troops preparing the foundation for a fort in the Nile Delta. Young provided the key breakthrough for deciphering ancient Egyptian hieroglyphs: cartouches on the stone contained a phonetic representation of the royal name Ptolemy. Jean-François Champollion took up where Young left off, applying his linguistic skills and knowledge of Coptic to decipher the rest of the stone. Besides working as a medical doctor,

simpler and more convincing demonstration of "double-slit" diffraction at the 24 November 1803 meeting of the Royal Society of London,[1] in which he split a narrow pencil of light into two parts using a card held edge-on into the beam. However, British science was unprepared to admit that the lion, Newton, could have erred, and further development of the wave theory of light came on the Continent with the work of François Jean Arago (1786–1853) and especially Augustin Fresnel (1788–1827). Fresnel produced a mathematical theory of diffraction, for which he was awarded the *Grand Prix* of the French Academy of Sciences in 1819.[2]

5.2.1 Single slit

To compute the diffraction pattern for a single slit, we must add the amplitudes arising from each point in the slit. We assume that the slit is uniformly illuminated at normal incidence. Let x be the position in the slit, which has width a, as indicated in Fig. 5.1.
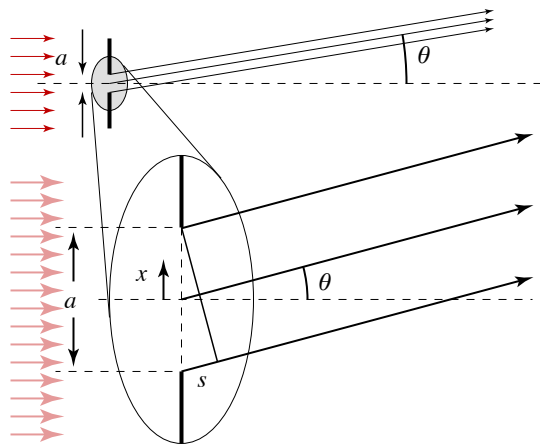

Figure 5.1 Geometry of Fraunhofer diffraction. A uniform wave illuminates a slit of width a at normal incidence. The diffracted light is observed on a screen far from the slit, so that we may approximate as parallel the rays coming from each point in the slit. Compared to the ray coming from the center of the aperture, a ray at x must travel an extra distance of $-x \sin \theta$. Hence, its phase differs from the central ray's by $-kx \sin \theta$.

If we center our coordinates on the slit, $-\frac{a}{2} \leq x \leq \frac{a}{2}$, then the extra distance traveled by a wave originating at point x in the slit, compared to the wave originating at $x = 0$, is $-x \sin \theta$. Its amplitude is therefore proportional to $e^{-ikx \sin \theta}$, where $k = 2\pi/\lambda$. We can thus add amplitudes (arrows) across the slit with the integral

$$\psi(\theta) \propto \int_{-a/2}^{a/2} e^{-ixk \sin \theta} \, dx$$

$$= \frac{e^{-ika/2 \sin \theta} - e^{ika/2 \sin \theta}}{-ik \sin \theta}$$

so

$$\psi(\theta) \propto a \frac{\sin(\pi a \sin \theta / \lambda)}{\pi a \sin \theta / \lambda} \tag{5.1}$$

Squaring and normalizing, we get the intensity

$$\boxed{I(\theta) = I_0 \operatorname{sinc}^2 \left(\frac{\pi a \sin \theta}{\lambda} \right)} \tag{5.2}$$

where

$$\operatorname{sinc} \phi \equiv \frac{\sin \phi}{\phi} \tag{5.3}$$

The single-slit intensity distribution of (5.2) is shown in Fig. 5.2. Of course, when the angle θ is small, it is possible to approximate $\sin \theta \approx \theta$, which simplifies this expression somewhat.

♣ Is the small-angle approximation appropriate in this experiment? You should answer this in the lab with the equipment before you.
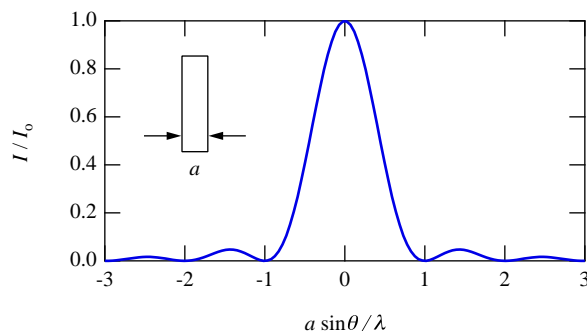

Figure 5.2 Diffracted intensity $I(\theta)/I_0$ far from a single slit, as given by (5.2).

5.2.2 Double slit

It is actually slightly easier to compute the diffraction pattern from a double slit, at least crudely. If we can neglect the width of the slits, then there are only two "arrows" to sum, giving

$$\psi(\theta) \propto e^{-i\frac{1}{2}kd \sin \theta} + e^{i\frac{1}{2}kd \sin \theta}$$

$$= 2 \cos \left(\frac{\pi d \sin \theta}{\lambda} \right) \tag{5.4}$$

Young also investigated the elastic properties of materials—Young's modulus relates the stretching of a solid (the strain) to the applied force per unit area (the stress).

[2]There's a great story about this prize. The prize committee included such notables as Arago, Siméon Poisson, Jean-Baptiste Biot, and Pierre-Simon Laplace. Poisson, in particular, was sure that Fresnel's theory was hogwash and set about a calculation of the shadow behind an opaque disk to see if he could identify a fatal flaw. Indeed, Poisson deduced that along the axis of the disk, where the shadow "should" be darkest, Fresnel's theory predicted a bright spot, which was obviously absurd. Arago requested that Poisson's prediction be tested, and to the general astonishment of the jury, the spot duly appeared! Fresnel won the prize, but credit for the spot is conventionally awarded to the skeptical Poisson.

where d is the distance between the slits, and we have used $e^{i\phi} + e^{-i\phi} = 2\cos\phi$.

However, a more realistic calculation takes into account the finite width a of each slit. We have already computed the sum of all the arrows from one slit, in (5.1), assuming the slit is centered on the origin. All we need to do is slide one slit up $d/2$ (which multiplies the integral by $e^{-i\frac{1}{2}kd\sin\theta}$) and slide the other down $d/2$ (which multiplies the integral by $e^{i\frac{1}{2}kd\sin\theta}$), to get

$$\psi(\theta) \propto e^{-i\frac{1}{2}kd\sin\theta}\,\mathrm{sinc}\,\Phi + e^{i\frac{1}{2}kd\sin\theta}\,\mathrm{sinc}\,\Phi$$

$$= 2\cos\left(\frac{\pi d\sin\theta}{\lambda}\right)\mathrm{sinc}\,\Phi \qquad (5.5)$$

where

$$\Phi = \frac{\pi a\sin\theta}{\lambda} \qquad (5.6)$$

Squaring and normalizing gives the observed intensity pattern,

$$\boxed{I(\theta) = I_0\cos^2\left(\frac{\pi d\sin\theta}{\lambda}\right)\mathrm{sinc}^2\left(\frac{\pi a\sin\theta}{\lambda}\right)} \qquad (5.7)$$
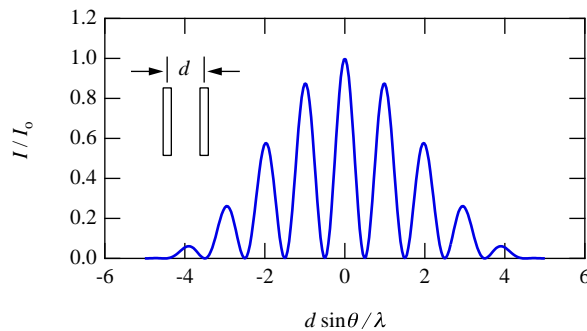
which is plotted in Fig. 5.3.

Figure 5.3 Diffracted intensity from two identical slits of width a whose centers are separated by distance $d = 5a$.

5.2.3 Multiple slits

You can work out the theory for N slits by generalizing the double-slit case. There is one more mathematical trick that helps tidy up the answer. With N equally spaced slits the phase difference between adjacent slits is $r = e^{-ikd}$ and we must sum

$$S_N = 1 + r + r^2 + \cdots + r^{N-1} \qquad (5.8)$$

which is a **geometric series**. Multiply (5.8) by r and subtract from (5.8) to get

$$S_N = 1 + r + r^2 + \cdots + r^{N-1}$$
$$rS_N = \phantom{1 + {}} r + r^2 + \cdots + r^{N-1} + r^N$$
$$(1-r)S_N = 1 - r^N \qquad (5.9)$$

Substituting the definition of r and symmetrizing gives the result

$$\psi = \frac{e^{i\frac{1}{2}Nkd\sin\theta} - e^{-i\frac{1}{2}Nkd\sin\theta}}{e^{i\frac{1}{2}kd\sin\theta} - e^{-i\frac{1}{2}kd\sin\theta}}$$

$$\psi = \frac{\sin\left(\dfrac{N\pi d\sin\theta}{\lambda}\right)}{\sin\left(\dfrac{\pi d\sin\theta}{\lambda}\right)} \qquad (5.10)$$

An example of the application of this expression to calculate a 5-slit diffraction pattern is shown in Fig. 5.4.
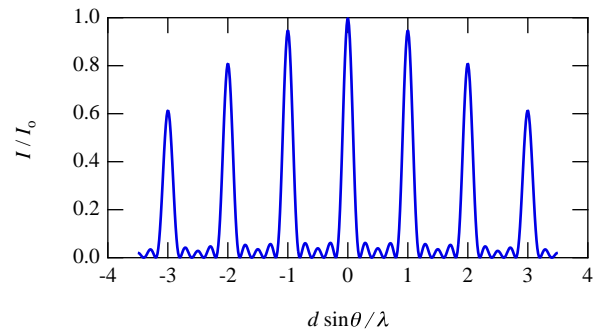
Figure 5.4 Diffracted intensity from five identical slits of width a separated by $d = 8a$.

5.3 Experimental Procedures

Make sure the HeNe laser is turned on so that is warmed up by the time you are ready to make measurements. Also check that the shared 5-volt power supply is on. It is located on a small wooden platform between the two Fraunhofer diffraction setups. *Do not turn on the photocell power supply yet.*

5.3.1 Alignment procedure and preliminary observations

Warning: **The laser mount and optical rail overall do not handle vibrations and bumping very gracefully. Take care not to bump the setup and to keep vibrations of the photometer to a minimum.**
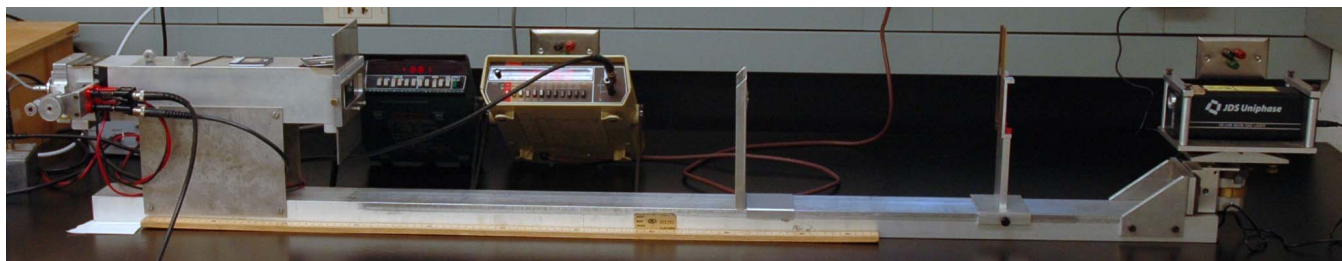
Figure 5.5: Fraunhofer photometer. The beam from the laser at the right passes through a partially silvered glass slide (optional) before encountering a slide with vertical slits at the entry to a light-tight box. The photodetector is at the extreme left of the photograph, mounted to a translation stage that allows it to be swept across the diffraction pattern of the slit(s).

It is likely that the system is still reasonably well aligned from the previous lab meeting and will need only minor adjustments. To align the system, first use the alignment target. Put the target close to the laser and translate the laser perpendicular to the optic axis until the beam falls on the target spot. Do this by adjusting the control screws on the laser mount. The control screws are on the far side of the laser mount, so look/feel for them carefully! To align the laser beam along the optic axis of the apparatus, place the target just in front of the photocell box and rotate the laser about vertical and horizontal axes to center the laser

spot again. Repeat these two steps until no further adjustments are necessary to center the laser on the target spot at both locations.

Before continuing with the final alignment, you should observe a few examples of the single- and double-slit diffraction/interference patterns you will be measuring, and determine the effects of varying the slit width and separation. To do this, remove the photocell mount at the rear of the photometer box and project the pattern on a piece of paper after mounting a slide of single slits in the holder on the front of the box.

5.3.2 Removing the photocell mount

The photocell mount is secured to the photometer by means of a set screw. The photocell mount is located at the rear of the light-tight box and the removal of the photocell mount is a simple matter of loosening a set screw and then sliding the photocell mount from the backside of the light-tight box. (See Fig. 5.6.)
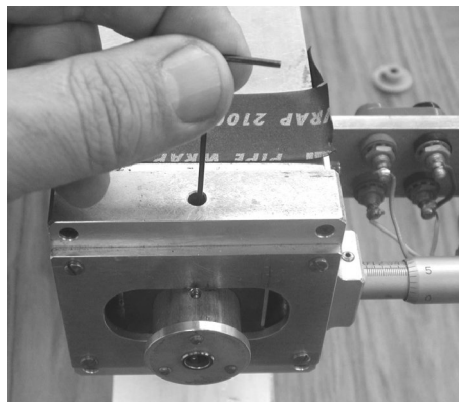
Figure 5.6 Removing the detector. Be gentle with the set screw; a modest torque suffices to hold it in place and make proper electrical contact.

The set screw securing the photocell mount to the light-tight box is accessed via a hole drilled through the top of the aluminum plate that is attached to the rear of the light-tight box. Move the photocell mount to the midpoint of its travel. Look into the hole in the above-mentioned aluminum plate. You should be able to see the set screw. Use the hex wrench provided to loosen the screw.

> **Warning:** Do not remove the set screw. Simply back it out a single turn. Do not remove the cable from the photodetector; handle it gently to avoid compromising the wire connections inside.

If you're unable to see the set screw when the photocell mount position is set to the midpoint of its travel, try moving the photocell mount back and forth a short distance while looking into the access hole. See your instructor for help.

Center the broadest slit in the laser beam. (Note: The clamping screw holding the slide in place should be tightened only very lightly by hand. Over-tightening can damage the slide.) Sketch the diffraction pattern seen when projected onto a piece of paper.

Warning: Do not look directly into the laser beam!

Try some of the other single slits; then change to the double slits. From your sketches write down how varying the distances a and d affects the patterns. Next remove the slide with the slits from the front of the photocell box. You should now determine carefully for later use the distance from the slits on the slide to the photocell window when it is remounted in its box.

Here are some things to worry about while the photocell is still out:

♣ Note that the width of the window in front of the photodiode is 130 μm. How would a narrower or wider window affect your data?

♣ How should the window be oriented?

♣ What is the distance between the target slits and the window in front of the detector?

When you have sufficiently pondered these questions, replace the photocell in the mounting bracket on the back of the photometer box and tighten the set screw that holds the photocell in place. This screw needs to make good contact with the photocell to ground the cell properly.

The photocell is positioned by turning a micrometer drive attached to the back of the photometer box. The micrometer drive is connected by a gear arrangement to a 10-turn potentiometer. With 5 V applied across the entire potentiometer, the relative position of the photocell is given by the voltage $(0 - 5$ V$)$ between the center tap and one end of the potentiometer, as read by a voltmeter.

5.3.3 Potentiometer

A potentiometer (as we are using the term) is nothing more than a resistor that allows one to make connections across the ends of the resistor and between one end of the resistor and the movable center tap (Fig. 5.7). If 5 V is set up across the ends of the resistor, then the voltage between the center tap and one end of the resistor will vary from 0 to 5 V linearly with the position of the tap.

If you haven't already done so, now would be a good time to sketch the experimental apparatus and explain how it works; *i.e.*, how you can use it to measure a diffraction pattern.
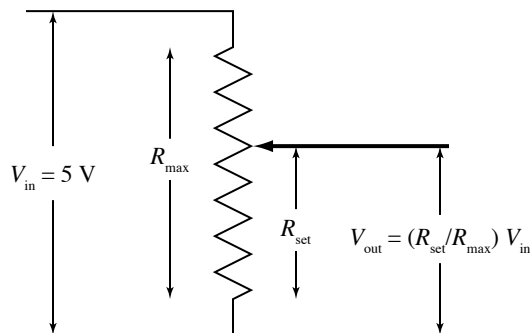
Figure 5.7   Potentiometer position circuit.

For the final alignment, you will use two voltmeters to monitor the photocell position and the photo-voltage. To avoid saturating the photocell keep the photo-voltage below 5 V; you need to reduce the laser beam intensity by adjusting the partially silvered mirror ("beam attenuator") mounted on the rail near the laser (Fig. 5.8). Position the mirror so that the laser beam passes through it about one-fourth of the way from the top.

Warning: If the beam that reflects from the slide heads back into the laser, it can destabilize the laser. Rotate the slide slightly to prevent this.

Figure 5.8   The attenuator is the partially silvered glass slide at the left. Note the small red dot of the helium-neon laser beam, which has a wavelength of 632.8 nm.

Finishing the alignment

Turn on the photocell power. Using the same hex wrench that you used to remove the photocell mount, remove the gear from the end of the potentiometer. Rotate the micrometer handle while watching the reading on the voltmeter connected to the photocell output. Adjust the photocell position so that the voltage obtained is a maximum. Then adjust the beam attenuator so that the photocell has an output just less than 5 V, so

it is not saturated by the laser beam. Next, adjust the potentiometer until the voltmeter connected across the center tap of the potentiometer reads 2.5 volts. Without rotating either shaft, reinstall the gear onto the end of the potentiometer so that the two gears line up and are flush with each other. This should center the maximum intensity peak in the range of travel.

At this point you should make fine adjustments of the vertical position of the laser beam spot by adjusting the tipping screw on the laser's seesaw bracket. Adjust for maximum photo-voltage. Once again adjust the beam attenuator to give a non-saturating peak photo-voltage close to, but not exceeding, 5.0 V.

Change the micrometer position and note the intensity profile of the laser beam. If you did everything correctly, you should see a steep but smoothly varying profile with a peak photo-voltage of a bit less than 5.0 V at a micrometer position voltage near 2.5 V.

5.3.4 Beam profile

You will be using a computer program to aid you in taking data and in comparing your results with the theoretical expressions of section 5.2. The voltages from the potentiometer and photodetector are converted to digital values by an analog-to-digital converter (ADC), which is connected to a USB port on the computer. During an acquisition, these voltages are sampled several times per second and plotted in real time as photo-voltage vs. position voltage, allowing you to monitor the data as you slowly rotate the micrometer or potentiometer gear to translate the detector window.

To get a feel for how the acquisition program works, you should first take a scan of the laser beam itself, with no target slit mounted at the front of the photometer box. If Igor is not already running, launch the program by clicking on the Igor icon in the Dock at the right of the screen. It is shown in Fig. 5.9. Then select the **Fraunhofer Diffraction** program.[3]
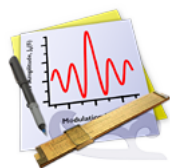
Figure 5.9   Igor icon.

Rotate the micrometer to one edge of its range, note the position of the micrometer (see Fig. 5.10), then click

the **Take Data** button.

> **Warning:** There is a long delay (up to 20 seconds) the first time you take data, as the National Instruments driver software is loaded into memory. Be patient and do not touch the computer during this period.

You will be asked to block the beam with your hand so the program can measure the "zero level" of the photocell, which may be negative.[4] The zero-level reading will be subtracted from subsequent readings.

Smoothly turn the micrometer shaft or gear to advance the detector. Watch the computer screen as you go to make sure that the data are smooth and quiet. *Do not reverse the direction of the photocell motion while collecting data.* If you do, you will see the effects of "screw backlash" in your data. Scan through the complete 5-V range of the potentiometer, then hold down the **Shift** key to finish the acquisition. **Note that the vertical axis rescales automatically during the acquisition, so that the noise in early points may appear exaggerated at first.**
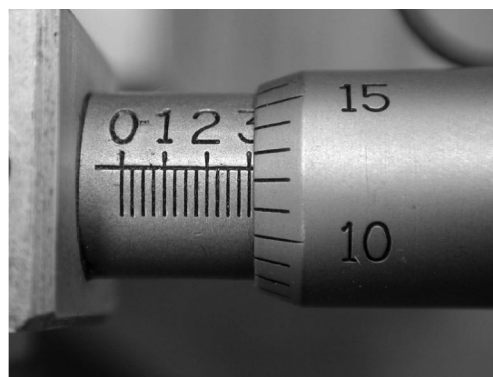
Figure 5.10   Reading the micrometer: it takes four revolutions to advance by one "major" unit. What do you suppose those units are? How could you check?

You may save and print these data, if you wish, but you should at least note the width of the beam, which you can readily convert from volts to distance. Note that the procedure described above measures the width of the laser beam at the detector plane, not the slit plane where it is apt to be slightly narrower.

♣ Would you say that the laser profile you have measured justifies the assumption of a uniform plane wave incident on the slit plane?

---

[3] Igor is a powerful data acquisition, analysis, and graphics program that runs on both Macintosh and Windows. Unlike Kaleidagraph, Igor is entirely scriptable: virtually everything that can be accomplished using the graphical user interface can be programmed within Igor using a C-like programming language. The programs that acquire and analyze data for the *RLC* and Fraunhofer diffraction experiments produce Igor experiment files holding your data, graphs, and tables, as well as a history of the operations you conduct. You may use them to export data in text format for further analysis in Kaleidagraph, Origin, or other program, or you may analyze your data further within Igor. See http://www.physics.hmc.edu/analysis/Igor.php for further information about using Igor.

[4]A negative reading from an amplified photodiode does not mean that you are detecting negative photons! Rather, it means that the amplifier has a slight negative voltage offset.

Determine the distance the micrometer has advanced and use this information to calculate the width of the laser beam.

♣ Is your value reasonable?

5.3.5 Taking data

At this point you are ready to take diffraction data. Select the slide of single slits and note the labeling of the slide and the slits (it will be handy later to have identified each slit before installing the slide). Place the slide in the holder at the front of the photocell box and center the slit you have selected under the laser spot. Advance the micrometer until the potentiometer voltage is about 2.5 V, and scan a short way in either direction to look for the maximum intensity. At the maximum, adjust the attenuator slide to produce a peak intensity slightly lower than 5 V. Then set the micrometer at the edge of its travel.

Start the data acquisition and carefully scan the micrometer across the full range of its motion, watching the computer screen to monitor the data acquisition as you go. You should also confirm that the program is in agreement with the voltage levels reported on the multimeters. Hold down the **Shift** key when you have finished.

When you are satisfied with your data, click the print button (in the **Save & Print** tab) and print a graph of the voltage from the photodiode as a function of the voltage across the position potentiometer. You will be asked to enter the first name of each lab partner and your section, so the program knows where to save your data. Cut the two copies of the graph apart; partners should tape a copy in their notebook and analyze the results. Note that the wavelength of the HeNe laser is 632.8 nm.

♣ Is the overall shape of the pattern consistent with theoretical expectations?

♣ Does the spacing of the minima or maxima agree with theory?

♣ Are there details of the pattern that differ from theoretical expectations?

Do not proceed until this work has been completed in your lab notebook.

Now click the **Normalize** tab. To convert your position data (in volts) to angles, you must enter the *complete angular range* of motion of the detector window (corresponding to a 5-V change in the potentiometer voltage). You should already have a sketch of the relevant geometry in your notebook, as part of the analysis you just finished. Use the distance from the target slit to the detector window to compute the complete range of angles traveled by the detector, in radians.

(Is the small-angle approximation appropriate?) Enter this value in the box, then click the **Normalize** button. The Normalize routine both scales and shifts the data to produce a maximum peak height of 1 centered at $\theta = 0$. For best results you should make sure to go smoothly and rather slowly over the central maximum so the program can make an accurate determination of the peak height and position.

Comparing to theory

To compare your data to (5.2), click the **Slit Dimensions** tab and enter the slit width. Nominal values for the slit widths (and separations) are printed on the slides. *More accurate values are listed in a table at the end of this section of the manual.* Then select **1 slit** from the popup menu. You should see the computed curve superimposed on your data, as well as a panel of residuals at the top of the graph. The residual of a data point is its vertical distance from the computed curve. When a curve accurately describes experimental data, the residuals should be randomly distributed with respect to the zero line. When there are significant disagreements between the curve and the data, the residuals will exhibit a pattern. Note that the program assumes that $\lambda = 632.8$ nm, which is the wavelength of the HeNe laser.

If the calculated curve is either fatter or skinnier than the central peak in your data, click on the **Hypothetical Dimensions** tab and enter a different value for the slit width, adjusting that value until you get the best agreement. Note that when a hypothetical curve is plotted, the residuals panel is adjusted to show residuals for the hypothetical curve.

♣ To narrow the diffraction peak do you have to increase or decrease the value of the slit width? Why?

If you find that you must vary the hypothetical slit width appreciably from the expected value, stop and take stock of your procedure. The slit widths were measured carefully with a high-quality optical microscope (these accurate values are listed in a table at the end of this section). Is it likely that the table of slit widths is significantly off? What else could account for the discrepancy? Jot some ideas in your notebook as you test various hypotheses for the discrepancy.

Table 5.1 Key Experimental Parameters	
HeNe laser wavelength	$\lambda = 632.8$ nm
Detector window width	$w = 130\ \mu$m

Saving and printing

When you have found the best value for the hypothetical slit width, save and print your data. Click the **Save & Print** tab and click the **Save** button. *Do not*

use the **Save** *command in the* **File** *menu.* You will be asked for each partner's name and your section number so your data can be stored in your own local directory (you will be able to copy the files to your own computer or your Charlie account later). A dialog will tell you the name of the data file, which you should record in your lab notebook. This information is also copied to the **History** window at the bottom of the screen.

Then click **Print**. You will be asked for a (brief) comment. You may leave this blank, or enter a short description (e.g., "Marge Inovera turned the micrometer"). After this dialog is closed, you will be pestered with one more to check the positioning of the graph annotations. Sometimes the text will overlap your data. Using the mouse, drag any annotations that are poorly positioned into appropriate places, then click the **Print** button.

Two copies of the graph will be printed, one for each partner. Cut the page in half, tape the graph into your lab notebook, and comment briefly on the extent of agreement between your data and the calculated curve(s) on the same notebook page. To test repeatability, you should realign the apparatus and record a second run.

At this point you should be up and running with the acquisition program. If all has gone well in the first week, you should be well into computer data taking by your second lab period. Repeat the experiment using another single slit, and two different double slits. If time permits, check out the three-, four-, and five-slit configurations, as well. However, the focus of your work should not be on producing the greatest number of graphs for the notebook, but in making careful measurements and in analyzing the trends and discrepancies between your data and theoretical predictions.

Readjust the laser optics to eliminate any saturation of the photocell response at intensity maxima and for a maximum as large as possible but not to exceed 5 V. Carefully re-center each set of slits in the laser spot. There is a second set of gears you can use to change the range of the photocell positions. These gears should be used for some of the slits to record a useful portion of the pattern. **Note that you will need to recompute the total angular range for the new gear set.** There is also one $3\times$ beam expander to be shared between the two setups, should you wish to change the size of the beam. It consists of a diverging lens followed by a converging lens, and the distance between them can be adjusted by screwing in or out the diverging lens. Note that the expander needs to be put far from the laser to work properly.

In analyzing your results, there are a number of opportunities for quantitative comparison with theory. Your experimental curves should have the correct general qualitative form, but they will unavoidably differ from the ideal theory in some particulars. **The positions (and uncertainties!) of the experimentally ob-** served maxima and minima should be used together with different trials for the hypothetical parameters to derive refined estimates of the slit dimensions and their uncertainties for one particular single slit and one particular pair of double slits.** (These should be compared with direct microscopic measurements listed in the table at the end of this section.)

Discrepancies in the heights, depths and symmetries of the peaks should give you food for thought and further analysis. The theory section makes certain simplifying assumptions about the illumination and geometry that may not describe your experiment accurately. Many of these "non-idealities" can be profitably simulated and studied at length in the context of a laboratory report. Finally, **you should use the trends of your plots for the 3-, 4-, and 5-slit patterns to describe the effect of increasing the number of slits, and explain the cause of this effect**.

5.3.6 Program details

The program has two adjustable parameters to tailor the acquisition to your needs: the number of individual analog-to-digital values to average to produce each data point, and the number of data points per second to record. While you are acquiring data, the data acquisition board samples the two input channels at a steady rate of 10^4 samples per second on each channel. To reduce electronic and digitization noise, many samples are averaged together to produce each data point, and their standard error is recorded as the uncertainty of the data point.

You may set the number of samples to average for each data point. The default value is 500, which means that each data point is acquired during a time of 50 ms. You may also set the number of data points you wish recorded each second. Note that this value is only approximate. The default value is 12, but you may wish to sample either more or less rapidly.

Note that when you normalize your data, the uncertainties are scaled by the same factor as the data, so that their relative value remains unchanged.

Working offline

As you perform an experiment using Igor, the various data runs you take are stored within an experiment file (extension **.pxp**), along with a host of other settings and code to recreate graphs and other windows. In addition, the code that actually runs the experiment and data acquistion is included "by reference," which means that it lives in files on the lab computer separate from the experiment file. This means that if you wanted to head to the computer lab to continue analyzing your data using one of the keyed copies of Igor available in both the Mac and PC labs, you would be stymied by the missing procedure files when you tried to open the experiment file. There are a couple of ways around this

problem.

First, you could use the text-only versions of the data that are saved along with the experiment file (described below), loading them into a new Igor experiment file one by one. Alternatively, you can have all the procedure files copied into your experiment file, so that they will be available when you open the experiment on another computer. This would allow you to use the fitting functions available in the lab, among other things, at the expense of a few more hundred kilobytes in the experiment file. Igor calls this copying business *adoption*, and you can adopt all the necessary procedure files by holding down the shift key and selecting **File | Adopt All...** after you have saved your experiment file at least once. If you want to continue to analyze your data *using Igor* after you leave the sophomore lab, you should Adopt All. If you would prefer to use a different software package, read on.

5.3.7 Saved data

The program saves your data in two different places. First, it is copied into a data folder within the Igor experiment itself. This allows you to store many different runs inside a single Igor file. Igor's data folders are like directories or folders on a hard drive. You can access these data by clicking on the **Load Data** button in the **Acquire** tab. Second, it is exported to a tab-delimited text file in a subfolder of the **Data** folder on the Desktop of the computer. Open this folder, then open your section's folder, and then open the folder named for you and your partner. Each file is named **Fraun_xxxx**, where **xxxx** represents the time at which the data were saved.

The text file format is tab-delimited text, which can be opened by any data analysis package. The first row holds column labels, and subsequent rows hold the data. There is a potential gotcha with text files: different

operating systems use different characters to indicate the end of a line. UNIX computers use the linefeed character (LF = ASCII 10), Macintosh computers use the carriage return character (CR = ASCII 13), and DOS/Windows computers use both (CR+LF)! If your data analysis program has trouble reading the file, you can use Excel to open it first and handle the conversion of line terminators.

There are many ways to transfer your data from the lab computer to your own. Perhaps the easiest is to use a web-mail service (such as gmail or HMC's web mail interface at <http://www.hmc.edu/webmail/>). You can then attach your data file(s) to the message and mail it to both partners. Alternatively, you can mount your Charlie account. Near the bottom of the Dock (at the right of the screen) is an icon of a sphere inside a cube. Click that icon and enter your login information on Charlie. If someone has thoughtfully removed that icon, click on the Desktop to activate the Finder and select the **Go | Connect to Server...** command. Enter **afp://charlie.ac.hmc.edu** and click **Connect**. Once your class's Charlie volume mounts, select your personal folder and copy the files you wish to your Charlie account.

When you have finished, unmount Charlie by selecting the icon for your class's volume on the Desktop (not the one in the Dock you used to mount Charlie in the first place!) and using the **File | Eject** command.

5.3.8 Further reading

A laser beam is an almost ideal monochromatic light source, but it has both a finite spatial width and a finite spectral width. In this experiment it is safe to neglect the variation in wavelength. If you would like to learn a bit more about the spatial profile of laser beams, you can read a quick summary at <http://en.wikipedia.org/wiki/Gaussian_beam>.

Measured dimensions of slits, in microns. Separations are measured from center to center.

Slide	Slit	Slit Width(s)	Separation		Slide	Slit	Slit Width(s)	Separation
1	A	20.5			4	A	23.5	
	B	41.0				B	45.0	
	C	85.0				C	83.0	
	D	166.0				D	166.0	
2	A	42.0, 41.5	252.8		5	A	39.0, 40.0	248.3
	B	41.0, 41.5	511.3			B	42.5, 40.0	516.3
	C	83.0, 81.0	253.0			C	81.0, 81.0	251.5
	D	83.0, 82.0	513.5			D	82.0, 85.0	508.5

Slide	Slit	Slit Widths	Separation(s)
3	A	43.0, 45.0	129.0
	B	42.0, 44.0, 40.0	127.0, 126.0
	C	44.0, 42.0, 44.0, 44.0	130.0, 132.0, 134.5
	D	43.0, 43.5, 41.0, 44.0, 45.0	125.8, 128.8, 128.5, 128.5
6	A	40.0, 41.0	131.0
	B	38.0, 40.0, 38.0	127.0, 128.0
	C	40.5, 40.0, 40.5, 40.0	132.3, 133.8, 134.3
	D	40.0, 39.5, 38.0, 40.0, 40.5	125.8, 129.3, 128.0, 128.3

Experiment 6

The Grating Spectrometer and the Balmer Series

Abstract

The colors emitted by many sources of light provide a spectral fingerprint that uniquely identifies them. You will use a **diffraction grating** to study the discrete line spectrum of two elements having a single electron in their outermost shell: sodium and hydrogen. Using literature values for the wavelengths of the yellow sodium doublet, you will calibrate the period of the diffraction grating. You will then measure the wavelengths of four visible hydrogen lines and compare them to the Balmer equation, deducing a value for the Rydberg constant.

6.1 Overview

A **diffraction grating** is a transparent or opaque substrate on which a series of fine parallel, equally spaced grooves are patterned. When a light wave illuminates a grating, each groove scatters light in all directions. However, the scattering will be appreciable only in directions for which the light from neighboring grooves arrives in phase. The behavior of diffraction gratings can thus be readily explained using the wave theory of light.

Diffraction gratings may be used to disperse the colors of a polychromatic source and to identify its component wavelengths. To make a quantitative identification of the wavelengths, you will need to calibrate the grating, which means to determine the distance d between rulings. The grating you will use has a *nominal* period[1] of 600 lines/mm, but you will determine d much more accurately than the 1 significant figure implied by this value by observing the yellow light of a sodium lamp. You have probably seen the bright yellow light of sodium lamps in parking lots, where they are used because of their energy efficiency and power. By taking as given the literature values for the wavelengths of the two adjacent sodium lines,[2] and by measuring precisely the angles at which the yellow light is diffracted by the grating, you can determine d.

Having calibrated the grating, you will then study the spectrum of a Balmer hydrogen discharge tube, which allows you to investigate the energy levels of a hydrogen atom. When an atom in an excited electronic state relaxes to a lower-energy state, it emits the energy difference between the states in the form of a photon of light. In the case of hydrogen, most of these transitions produce photons either in the ultraviolet or the infrared. For four transitions ending in the $N = 2$ level, the emitted photons are visible and you can study them by eye using the grating you have calibrated with the sodium lines.
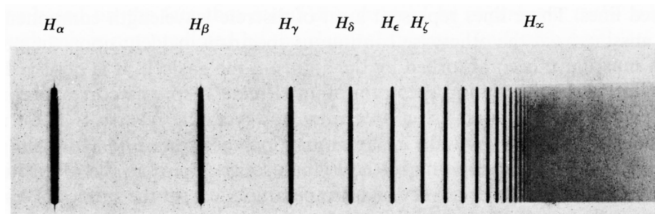
Figure 6.1   The line spectrum of hydrogen.

6.2 Theory

In 1666 while Isaac Newton (1642–1727) was home in Lincolnshire because the plague had closed down Cambridge University, he discovered that whereas a single prism of glass would cause sunlight to fan out into a rainbow of colors, a subsequent prism caused no further color changes. Newton reasoned that sunlight was a mixture of all colors, and that a prism caused the various colors to bend (or refract) differently, thereby producing a smooth, continuous band of gradually varying appearance.

---

[1] "Nominal" here means "named"; that is, the manufacturer's stated value. How many significant figures is the manufacturer promising?

[2] The discrete wavelengths at which a particular atom or molecule emits light are called "lines" because this is how they appeared on photographic plates in the early days of spectroscopy (see Fig. 6.1). A diffraction grating fans the light out in one direction, but doesn't affect the extent of the beam in the perpendicular direction. Hence, the "signal" on the photographic plate looks like a line segment.

A century and a half later in 1802, English chemist and physicist William Wollaston noted dark lines in the smooth solar spectrum. Twelve years later an orphaned glassmaker, Joseph von Fraunhofer (1787–1826), discovered that the smooth solar spectrum was actually interrupted by a series of dark lines. [3]

The positions of the spectral lines of different substances made a challenging puzzle to nineteenth century scientists—and to modern quantum theory, for all but the simplest atoms! In 1885, J. J. Balmer (1825–1898) showed that the wavelengths of the visible spectral lines of hydrogen were given by the expression

$$\frac{1}{\lambda} = R\left(\frac{1}{N^2} - \frac{1}{n^2}\right) \tag{6.1}$$

where n is an integer and $N = 2$ for the visible lines of the Balmer series. The hunt was on to find a theory that could explain Balmer's relationship.

At the end of the century, Max Planck (1858–1947) found he could explain the smooth (dominant) part of the solar spectrum with the help of the relation

$$\Delta E = h\nu = \frac{hc}{\lambda} \tag{6.2}$$

between the frequency and the quantum of energy emitted or absorbed by an atom. Niels Bohr (1885–1962) combined the Planck relation with his own idea of angular momentum quantization to produce a simple theory for the hydrogen atom that reproduced Balmer's relation, (6.1). Bohr's model described unphysical microscopic solar systems, but the principal idea of angular momentum quantization survived the birth of modern quantum mechanics in 1925–26 and the hydrogen atom calculations of Heisenberg (1901–1976) and Schrödinger (1887–1961).
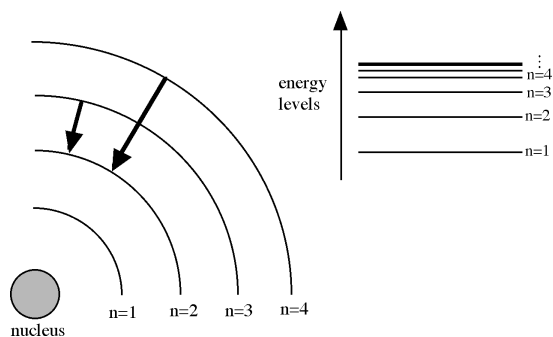

Figure 6.2   A schematic representation of quantized atomic energy levels.

Quantum mechanics also tells us that electrons in orbit around a nucleus cannot have just any energy (unlike planets orbiting the Sun, which can), but are restricted to certain quantized energy levels (Fig. 6.2). This prediction of quantum mechanics explained the previously mysterious fact that elements emitted light only at discrete wavelengths. Quantum mechanics was able to explain and quantitatively predict the wavelengths of these spectral lines for hydrogen, the simplest element. This confirmed the relation that had been empirically discovered by Balmer for the wavelength of the visible hydrogen lines,

In (6.1), N and n are integers corresponding to the lower and upper energy levels, respectively. For the visible Balmer series, the lower level always corresponds to $N = 2$. However, n is different for each spectral line in the series, each corresponding to a different upper energy level. The constant R is known as the Rydberg constant, with an accepted value (1998) of $R = 1.097\,373\,157 \times 10^7$ m^{-1}. In this experiment you will first observe and carefully measure the wavelengths of these spectral lines, and then determine best values for n and R.

The technique used in this experiment to measure the wavelengths makes use of diffraction gratings. When *parallel* light is incident upon an array of many equally spaced rulings on a transparent substrate, light emerges on the other side at certain specific angles where constructive interference occurs among the simultaneously illuminated rulings. These intensity maxima occur at angles θ for which

$$d\sin\theta = m\lambda \qquad m = 0, \pm1, \pm2, \dots \tag{6.3}$$

Here, d is the spacing between rulings on the grating, and the integer m is referred to as the order of the observed interference maximum. The theory of diffraction gratings is a straightforward extension of the Fraunhofer diffraction theory presented in Sec. 5.2. In particular, the behavior of the grating is roughly as indicated in Fig. 5.4, which shows the diffracted intensity from 5 identical slits, except that the transmission peaks for the grating are much narrower and the small wiggles at the bottom are much more rapid and imperceptibly tiny.

As indicated in Fig. 6.3, the purpose of the grating spectrometer is to measure the angles θ at which lines are observed, allowing their wavelengths to be determined. The light from some source enters the instrument at a slit opening on the collimator tube. We need

---

[3]Fraunhofer's childhood augured anything but a scientific career. Orphaned at age 11, he was not strong enough to become a wood turner and was apprenticed to a glassmaker, Philipp Anton Weichselberger. Weichselberger did his level best to frustrate the intellectual curiosity of the boy for the next two years, but Fraunhofer's life took a sudden turn when Weichselberger's house/workshop collapsed in 1801. After being buried in the rubble for several hours, Fraunhofer was rescued and came to the attention of Prince Elector Max IV Joseph, who directed the rescue operation. The prince took an interest in Fraunhofer and made sure that he had books and time to study while he learned the craft of lens grinding.

At age 22 he was appointed head of the glass factory in Benediktbeuern, where he worked to develop new types of glass and on methods to avoid streaking and inhomogeneities. He expanded the product line to include telescopes, binoculars, microscopes, magnifying glasses, and the best and largest telescopes at the time. He invented the spectroscope in 1814 and the diffraction grating in 1821.

the light striking the grating to be parallel rather than diverging, so the lens in the collimator tube serves this purpose.

After striking the grating and emerging at some angle θ, the light is collected for viewing by the telescope assembly. The lens on this assembly focuses the paral-

lel light into an image which is viewed by the observer through the eyepiece. The principal observational challenge in conducting this experiment is learning to use the instrument to measure θ most effectively with a minimum of systematic error and uncertainty.

$$d \sin \theta = m\lambda, \qquad m = 0, 1, 2, \cdots$$
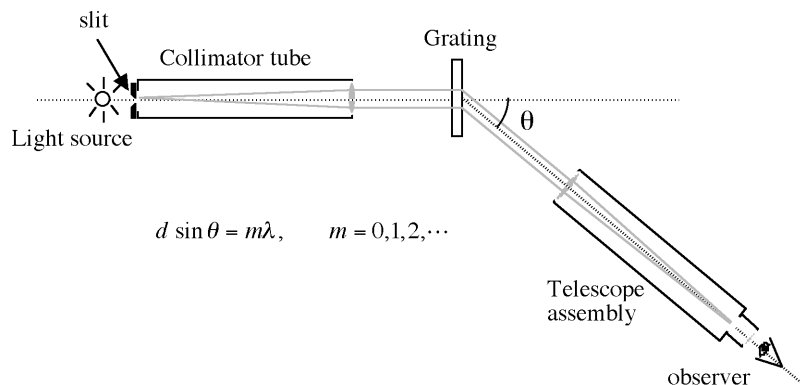
Figure 6.3: A schematic view of the operation of a grating spectrometer.
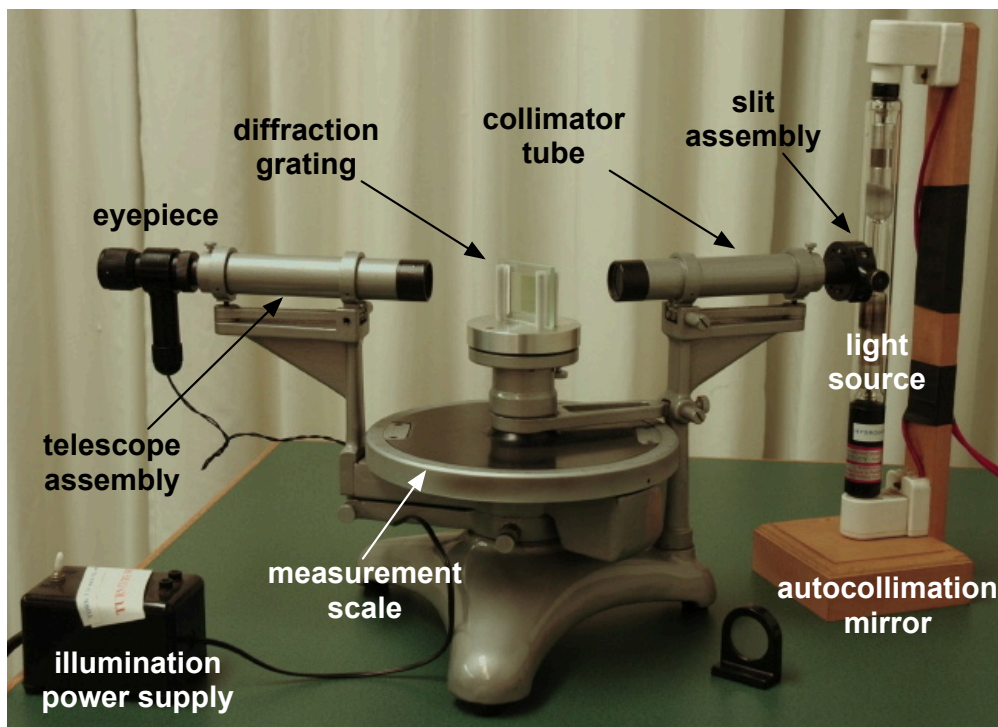
Figure 6.4: Photo of grating spectrometer and supporting equipment.

6.3 Using the Spectrometer

The grating spectrometer is shown in Fig. 6.4, with its important components and adjustments labeled. You should begin by familiarizing yourself with the lo-

cation and function of all the parts, so that you can use them comfortably while working in the dark. The important moving parts are the telescope assembly, the

rotating grating table, and the circular measurement scale. Each of these can be moved by hand, and locked in place by tightening a lock screw. Avoid damaging the instrument by forcing the components to move when they are locked. Only moderate "finger tight" pressure should be required to hold adjustments in place, once they are made.

> **Warning: Never touch a grating! Gratings cannot be cleaned, and are significantly degraded by fingerprints and dirt.**

In addition to the rotational adjustments, you will adjust the focus of the telescope assembly, the collimator tube, and the width of the slit opening. Take time now to find the following adjustment locations and see how they work:

- Telescope arm lock screw — *under the circular scale on the right side of the center post*

- Telescope arm fine adjustment — *under the rim of the scale on the front of the telescope*

- Circular scale lock screw — *under the circular scale on the left side of the center post*

- Grating table lock screw — *above the circular scale on the center post*

- Reticle (cross hairs) focus adjustment — *outer black ring of telescope eyepiece*

- Telescope image focus lock — *knob on top of telescope near eyepiece*

- Collimator focus lock — *knob on top of collimator near slit*

- Slit width adjustment — *black knob on side of slit assembly*

6.3.1 Focusing and alignment procedure

For the spectrometer to work properly, the light that diverges from the input slit must be collimated (made into a non-diverging beam) by the collimator tube, the grating must be perpendicular to the collimated input light, the telescope assembly must be focused to image collimated light correctly, and the angular scale must be configured to read out the deflection angle of the light being viewed. The following procedure will allow you to achieve these conditions.

1. First you must make sure the last stage of the eyepiece is adjusted to give focused viewing for your eyes. Look through the eyepiece and rotate the reticle focus adjustment until the black crosshairs appear clear and sharp.

2. Next you must adjust the telescope assembly to image collimated (parallel) light properly. Turn on the illumination power supply, which controls a light source near the eyepiece, shining backwards through the telescope assembly towards the grating table. Place the autocollimation mirror on the grating table, aimed to reflect light back into the telescope. Rotate the mirror slowly while looking through the eyepiece until you see a green cross (or at least a greenish blob of fuzz). If the grating table is tilted away from horizontal, it may be necessary to adjust this tilt to make the green cross appear in the telescope field of view.

   Slide the eyepiece in its sleeve until the green cross appears sharp and clear. Now the green light is being collimated by the telescope lens on its way to the mirror, and refocused for clear viewing on its way back to the eyepiece. By sliding the eyepiece in its sleeve, adjust the focus of the green cross until it is sharpest. Lock the eyepiece in place, keeping the reticle line parallel to the slit image. You should now see *both* the black crosshairs and the green cross sharply focused.

3. At this point it would be helpful to make sure that both partners can see both the green cross and the crosshairs simultaneously in focus. Either partner is allowed to adjust the reticle focus *only*. This accommodates for differences between eyeballs; no other adjustment should be necessary. If it is, go back to the previous step.

4. Next you must make sure the slit assembly is adjusted so that light diverging from the slit is properly collimated. Turn on the sodium vapor lamp and place it at the slit opening of the spectrometer. Line up the telescope and collimator tube to form a straight line path. Looking through the telescope eyepiece, you should see a bright image of the slit. Loosen the collimator focus lock and slide the slit assembly in and out until the slit image appears in sharp focus. Also take this opportunity to rotate the slit assembly so the slit image is vertical. Then lock the assembly in place with the collimator focus lock.

5. Check that the black crosshairs are properly oriented. You should see a single vertical crosshair, aligned with the slit image which you just adjusted. If the crosshair is not vertical, loosen the eyepiece and rotate it to bring the crosshairs into alignment. Verify that the green cross is still in focus before locking the eyepiece back in place.

6. You must now make sure that grating is perpendicular to the path of light through the spectrometer. Remove the autocollimation mirror and rotate the grating table so that the ruled face of the

grating faces the telescope. Make sure the grating is centered in the holder. *Do not touch the grating surface!* Now look through the eyepiece and rotate the grating table until the grating reflects some of the the green cross light back down the telescope. Locate the brightest image of the green cross, rotate the grating to center the image in the eyepiece, and adjust the tilt of the grating table to put the strongest green cross at the horizontal crosshair in the top half of the field of view. [Since the source of the green cross is at the bottom of the field of view, its image should appear symmetrically in the top part.] Lock the grating table to prevent further rotation, and turn off the green illumination source.

7. Finally, you must set the circular angle scale. With the slit image carefully centered on the crosshairs, adjust the circular scale until it is zeroed, and lock it in place. Note that the scale is zeroed when the zero marking on the outer ring aligns with the zero marking on the inner scale. The inner scale need not be located at the telescope, and in fact should be left out to the side for easier viewing.

6.4 Experimental Measurements

6.4.1 Grating calibration

From (6.3), it is evident that the observed position θ of a spectral line allows its wavelength λ to be determined provided that the grating spacing d is known. The grating in this experiment is known to have a nominal spacing of about 600 lines/mm, but a more precise determination of d is required. For this purpose, we will take the wavelengths of the yellow sodium doublet (588.995 nm and 589.592 nm) as accepted standards.

Turn on the sodium vapor lamp, if it is not already on, and place it at the slit opening of the spectrometer. It will take several minutes for the light to change from a pinkish mixture of contaminant wavelengths to the pure yellow color characteristic of Na. While it is warming up, rotate the telescope arm of the spectrometer and look for a series of red lines, as well as the yellow lines of the sodium doublet. [Until you have calibrated the grating, you are not in a position to measure the wavelength of some of the red lines; however, a subsequent investigation of these lines to determine the contamination gas could be an interesting topic for a technical report.] Using the nominal value for d, estimate the angle(s) at which you expect to find the sodium doublet, and use the spectrometer to look for them.

During this step, you'll be learning how to use the spectrometer, which will help you find the more difficult lines of the hydrogen spectrum. One of the adjustments available to you is the slit width. The wider the slit opening, the brighter is the light seen, which may be helpful in trying to view particularly faint lines. (You can also see fainter lines more easily by darkening the room, using a black cloth to shield outside light, and allowing 10 minutes or so for your eyes to become dark adapted.) A wider slit opening, however, has the undesirable effect of making the line image wider, and therefore making the uncertainty of its position greater. You'll note that if you make the slit wide enough, the images of the two doublet lines will merge together into one wide line.

Make adjustments until you achieve a good compromise between maximum sharpness and adequate visibility. Now is a good time to make and record estimates of the uncertainties in θ and λ associated with your choice(s) of slit width. Also note that in addition to the uncertainty caused by finite slit width, the spectral lines will have a finite width due to the resolving power of the grating. The smallest resolved wavelength difference $\Delta\lambda$ is

$$\frac{\lambda}{\Delta\lambda} = Nm \qquad (6.4)$$

where m is the order and N is the number of grating rulings illuminated. If you use a piece of paper to cover part of the width of the grating, you should see the apparent line width increase accordingly. Using standard error propagation, this effect should be incorporated into your uncertainty budget.
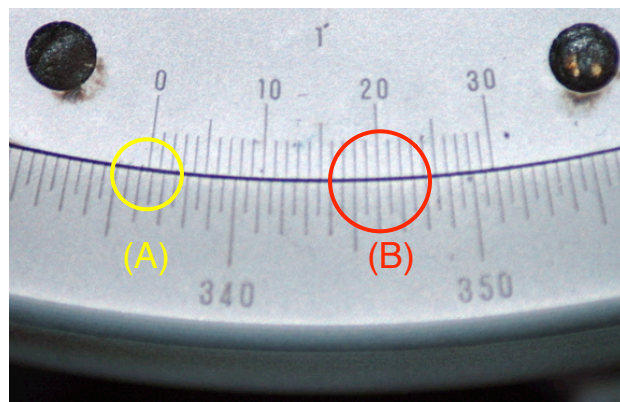

Figure 6.5 The vernier allows you to read angles to the nearest minute of arc. Start at (A) to determine $336° < \theta < 336°30'$. Then look for the lines that match up best at (B) to obtain the final value $336°20'$.

Make careful measurements of θ for both the sodium doublet lines on both sides of zero, for as many orders

as you can. Figure 6.5 shows how to read the vernier to about 1 minute of arc. At this point you may wish to thank the ancient Sumerians for the sexagesimal system we use for telling time and measuring angles. If you have been careful in the alignment of the spectrometer and grating, the right and left angles of deflection should agree with one another to better than $10'$; if they do not, now is the time to tweak the grating position to get symmetric readings, in preparation for the hydrogen measurements.

From your Na measurements, determine the best value for d to be used in the study of the hydrogen spectrum in the next part. You will probably note that some orders appear brighter than others, and that the lines may appear brighter or clearer on one side than the other. This is because the grating is blazed or optimized for maximum output in a particular direction. Make note of this information, as some of the fainter hydrogen lines may be hard to observe in any but the brightest and best conditions. Consider carefully the uncertainty of your data in deciding whether to use the first- and/or second-order measurements in obtaining a calibration for d and also in determining the hydrogen wavelengths.

6.4.2 Hydrogen observations

Replace the Na lamp with the Balmer hydrogen discharge tube lamp.

> **Warning: Be sure to turn on the cooling fan and direct it toward the base of the Balmer tube, or overheating may result.**

Observe the hydrogen spectrum. You will be looking for four visible wavelength lines, corresponding to colors of deep violet, violet, blue-green, and red. (The shortest-wavelength line can be particularly hard to see, and may require some patience and observation skill to locate and measure.)

Record the color and angular position of all the spectral lines in all of the orders in which they can be adequately observed. Include appropriate uncertainties for each measurement, so that the correct relative weighting can be assigned to observations of differing quality in coming up with a final set of "best results" for the hydrogen wavelengths. Calculate the final values and propagated uncertainties for the wavelengths of the hydrogen Balmer lines.

6.4.3 Forensic investigation

If time permits, return to the sodium lamp, which should now be cool again, and investigate the pink light it gives off before the discharge warms up. At room temperature, the vapor pressure of sodium in the lamp is too small to operate the electric discharge. Therefore, the lamp is filled with buffer gas to help the lamp get going. For the sodium lamp to work efficiently, the sodium vapor must be heated by the discharge to $260°C$. Different buffer gases may be used in sodium lamps to help launch the discharge and heat the sodium. Such gases must not react with sodium metal, or sodium vapor would be chemically removed from the cell.

You can figure out what is in the lamp if you use the spectral fingerprint of the buffer gas(es). Record as precisely as you can the positions of several prominant spectral lines, and convert them to wavelengths. You can then look them up at http://cfa-www.harvard.edu/amdata/ampdata/kurucz23/sekur.html.

To use the web-based database, you must enter ranges for several parameters. Here's how to find the sodium lines you have already observed:

1. Set the range from 500 nm to 700 nm.

2. Set the oscillator strength range from -2 to 4. Bright lines have large oscillator strength.

3. Set the lower level range from 0 to 10^5 cm^{-1} and the upper level range from 0 to 4×10^5 cm^{-1}.

4. Select all the Na items in the left column, then click **Search**.

You should see a listing of 8 spectral lines, including the two prominent lines of the sodium doublet. Notice that these are the only entries in this wavelength range that have the ground state as the lower energy level. Notice further that these lines are described by Na I, which is traditional spectroscopic notation for the neutral sodium atom. (Na II indicates singly ionized sodium, Na III is doubly ionized, etc.)

Another useful web site for spectral lines is maintained by the National Institute for Standards and Technology (NIST). It is available at http://physics.nist.gov/PhysRefData/Handbook/Tables/findinglist.htm. At this page, the lines are sorted by wavelength in 100-nm bands. However, you can click on the **Element Name** button to select a particular element, from which you can then display the strongest spectral lines, or focus only on the lines of the neutral atom.

6.5 Experimental Analysis

One simple way to evaluate your results is to compare them with accepted literature values, shown in the table below. Statistically characterizing any discrepancies as random or systematic, probable or improbable, is important to demonstrating your understanding of the experiment.

Color	Wavelength
red	656.27 nm
blue-green	486.13 nm
violet	434.05 nm
deep violet	410.17 nm

A somewhat more instructive analysis is obtained by comparing your results with the predictions of the Balmer relation of (6.1). There are several ways to manage this, of which one is given here:

- A plot of $1/\lambda$ as a function of $1/n^2$ should theoretically yield a straight line, whose slope and y-intercept are related to the Rydberg constant and the value for the constant $N = 2$. (Here, the consecutive integers, n, index the energy levels corresponding to each line; it will be up to you to determine convincingly what they are for the observed lines by comparing the linearity of the fit of the Balmer formula for several plausible choices of the sequence of n values.) You should prepare the plot giving the best straight line fit, along with appropriate error bars and a properly weighted linear fit determining R and N. Remember to describe how you performed your analysis in your lab notebook.

Discuss your results and their agreement (or lack thereof) with theoretical and accepted values. (How does your value of the Rydberg constant constrain other fundamental physical constants?) As quantitatively as possible, ascribe any discrepancies to sources of uncertainty in the experimental equipment and procedures. For example, if there is a lingering systematic bias between your wavelength estimates from the right and left sides, you should quantitatively investigate the origin of this error, and attempt to correct for it in coming up with your final results. Although the results of this experiment are actually well known to the physics world, you should approach the analysis task as if the results are original, and strive to be as precise and rigorous as possible in obtaining them.

Part III

Appendices

Naively, it would seem that a shock of 10 kV would be more deadly than 100 V. This is not necessarily so! Individuals have been electrocuted by appliances using ordinary house currents at 110 V and by electrical equipment in industry using as little as 42 V direct current. The real measure of a shock's intensity lies in the amount of current forced through the body, not in the voltage. Any electrical device used on a house wiring circuit can, under certain conditions, transmit a fatal current.

From Ohm's law we know

$$I = \frac{V}{R}$$

The resistance of the human body varies so greatly it is impossible to state that one voltage is "dangerous" and another is "safe." The actual resistance of the body varies with the wetness of the skin at the points of contact. Skin resistance may range from 1000 Ω (ohms) for wet skin to over $500,000$ Ω for dry skin. However, if the skin is broken through or burned away, the body presents no more than 500-Ω resistance to the current.

The path through the body has much to do with the shock danger. A current passing from finger to elbow through the arm may produce only a painful shock, but the same current passing from hand to hand or from hand to foot may well be fatal. The practice of using only one hand (keeping one hand in your pocket) while working on high-voltage circuits and of standing or sitting on an insulating material is a good safety habit.

The Physiological Effect of Electric Shock

Electric current damages the body in three ways:

1. it interferes with proper functioning of the nervous system and heart

2. it subjects the body to intense heat

3. it causes the muscles to contract.

Figure A.1 shows the physiological effect of various currents. Note that voltage is not a consideration. Although it takes a voltage to make the current flow, the amount of current will vary, depending on the body resistance between the points of contact.

As shown in the chart, shock is relatively more severe as the current rises. At values as low 20 mA, breathing becomes labored, finally ceasing completely even at values below 75 mA. As the current approaches 100 mA, ventricular fibrillation of the heart occurs —

an uncoordinated twitching of the walls of the heart's ventricles. Above 200 mA, the muscular contractions are so severe that the heart is forcibly clamped during the shock. This clamping protects the heart from going into ventricular fibrillation, and the victim's chances for survival are good.
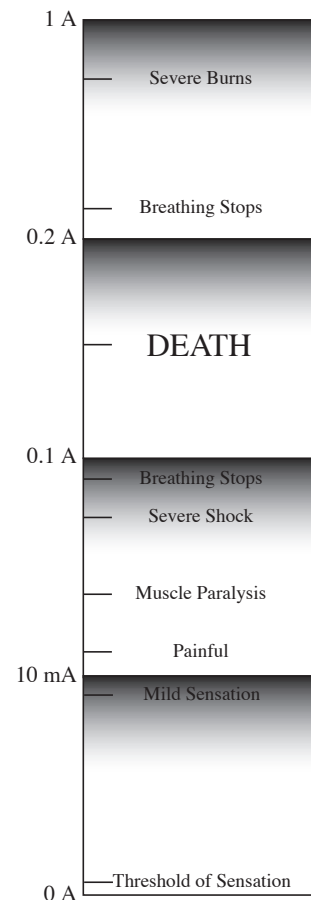
Figure A.1   Common injuries at different currents

AC is four to five times more dangerous than DC, because it stimulates sweating that lowers the skin's resistance. Along that line, it is important to note that the body's resistance goes down rapidly with continued contact. Induced sweating and the burning away of the skin oils and even the skin itself account for this. That's why it's extremely important to free the victim from contact with the current as quickly as possible before the climbing current reaches the fibrillation-inducing level. The frequency of the AC has a lot of influence over the severity of the shock. Unfortunately,

60 Hz (the frequency used in the lab) is in the most harmful range — as little as 25 V can kill. On the other hand, people have withstood 40 kV at a frequency of 1 MHz or so without fatal effects.

A very small current can produce a lethal electric shock. Any current over 10 mA will result in serious shock.

Summary

Voltage is not a reliable indication of danger, because the body's resistance varies so widely it's impossible to predict how much current will be made to flow through the body by a given voltage. The current range of $100 - 200$ mA is particularly dangerous, because it is almost certain to result in lethal ventricular fibrillation.

Victims of high-voltage shock usually respond better to artificial respiration than do victims of low-voltage shock, probably because the higher voltage and current clamps the heart and hence prevents fibrillation. AC is more dangerous than DC, and low frequencies are more dangerous than high frequencies. Skin resistance decreases when the skin is wet or when the skin area in contact with voltage source increases. It also decreases rapidly with continued exposure to electric current.

Prevention is the best medicine for electric shock. That means having a healthy respect for all voltage, always following safety procedures when working on electrical equipment.

In case a person does suffer a severe shock, it is important to free him from the current as quickly as can be done safely and to apply artificial respiration immediately. The difference of a few seconds in starting this may spell life or death to the victim. Keep up the artificial respiration until advised otherwise by a medical authority.

References:

1. Electronics World, p. 50, December 1965

2. Tektronix Service Scope, No. 35, December 1965, The Fatal Current

Resistors and Resistance Measurements

Types of Resistors

Resistance R is defined as the ratio of the potential difference V between the terminals of a conductor to the current I in the conductor, or

$$R \equiv \frac{V}{I}$$

If the resistance of a conductor is independent of the potential difference and the current; that is, if a plot of V vs I is linear, then the conductor obeys Ohm's law.

Almost all conductors offer some resistance to the flow of current and it is often desirable to keep such resistance in a circuit at a minimum. However, special devices called resistors are widely used in electrical circuits solely for their resistive value.

The most common resistor used in electronic circuits is made of carbon molded in the form of a small cylinder with wires attached at each end. The resistance value is generally designated by a set of three colored bands on the resistor. The band closest to the end of the resistor represents the first figure of the resistance; the next band gives the second figure; and the third band gives the number of zeros which must be added to get the total resistance. Figure B.1 below shows the values associated with each color.

For example, a resistor with bands of green-yellow-orange would be 54000 Ω, or 54 kΩ. For resistances ranging 0.1–9.9 Ω, gold is used for the third band. Thus, violet-white-gold is 7.9 W. This labeling system can describe resistances ranging from 0.01 Ω (black-brown-silver) to 99 MΩ (white-white-blue).

The fourth band indicates the tolerance; that is, the accuracy assured by the manufacturer in designating the resistance value. The tolerance is expressed in percent of the nominal resistance value, with silver representing 10 percent and gold 5 percent. The fifth band is a failure rate indicator.

Another commonly used resistor is constructed by depositing a thin metallic film on an insulating form. This type produces little noise in electronic circuits, and obeys Ohm's law over a wide range of currents and temperatures. A third type of resistor is made of a coil of resistance wire wound on an insulated form. These resistors can have very precise resistances, but it must be kept in mind that the inductive effect of the coil will decrease this accuracy when used in a high frequency circuit. Resistive values of these resistors do not follow the colored-band convention described in Fig. B.1, but are instead printed numerically on the resistor's body.

E.I.A. RESISTOR COLOR CODE

1st Significant Figure		2nd Significant Figure		Multiplier		Tolerance		Failure Rate Level (%/1000 Hours)	
Color	Value	Color	Value	Color	10^n	Color	Tolerance	Color	Level
Black	0	Black	0	Black	0	Silver	± 10%	Brown	M = 1.0%
Brown	1	Brown	1	Brown	1	Gold	± 5%	Red	P = 0.1%
Red	2	Red	2	Red	2			Orange	R = 0.01%
Orange	3	Orange	3	Orange	3			Yellow	S = 0.001%
Yellow	4	Yellow	4	Yellow	4				
Green	5	Green	5	Green	5				
Blue	6	Blue	6	Blue	6				
Violet	7	Violet	7	Silver	-2				
Grey	8	Grey	8	Gold	-1				
White	9	White	9						

Figure B.1: Resistor color band values

The accuracy of a resistor may be affected by a number of factors. For resistors of very low values (less than 10 Ω, the resistance of the connection to the resistor can become a significant part of the total resistance. For resistors of very high values (greater than 10 MΩ, leakage of current between the contacts on the surface of the resistor may significantly decrease the resistance, though this effect may be greatly reduced by the use of special insulating materials. The resistance of almost all resistors changes with temperature, and for this reason it is often necessary to hold the temperature of a resistor constant in order to maintain stability in an electronic circuit.

All resistors are limited in the amount of current which they can carry. An excessive current will burn out a resistor, but even a small overload will cause the resistor to become overheated and change its resistance. To determine the maximum current which a resistor can carry, use

$$P = I^2 R$$

The power rating of some resistors is marked on the resistor body; but for the common carbon type, only its physical size indicates the power it can handle. The smallest ones with a diameter of 1/8 inch are rated at 0.5 W, 1/4 inch at 1 W, and 3/8 inch at 2 W.

Resistors designed to operate "hot," such as tube filaments, heating coils, and light bulbs, have a much lower resistance when they are cold than when they are hot. Thus, a heavy surge of current occurs when the circuit is first closed, and at that instant the filament wire is most likely to burn out.

Study the various resistors available to you, determining their type, nominal resistance value, tolerance, and power rating. Some dissected resistors are on display in the laboratory and should be examined for an understanding of their internal construction.

Appendix C

Electrical Techniques and Diagrams

Precautions

Electrical experiments call for special precautions, since applying power to an incorrectly wired circuit can seriously damage or destroy expensive equipment. The key rules are:

1. When in doubt about a circuit, do not apply power. Recheck each branch and consider the function of each element.

2. If still in doubt, ask your instructor or lab assistant to check the circuit.

It is disconcerting to hear a student say, "I don't know whether this is right or not, but let's turn it on and see what happens." Always check the polarity of the circuit before wiring up a DC meter and understand your circuit before "trying" it. Attention should be paid to the power rating of every piece of electrical equipment — make certain never to exceed it. Include an appropriate fuse in your circuit whenever possible.

It will not be necessary for you to have your instructor check every circuit before it is put into operation, especially if it is one with which you are familiar. **However, in the case of a new circuit or one involving expensive equipment, the instructor or lab assistant should be asked to check the circuit before any power connections are made.** Any time that a circuit does not behave as it should, open all switches immediately and correct the problem before applying power again.

Elementary Electrical Techniques

Good results in electrical experiments require that the circuits be correctly and neatly wired and that you understand the reasons for each connection. Translating circuit diagrams to actual circuits quickly and accurately (and vice versa) takes practice. Some students may already be skilled in reading and wiring circuits, while for others it may be a new experience. The following notes are written primarily for beginners.

1. Circuits are usually specified by a schematic diagram using the symbols illustrated in Fig. C.1. Actual laboratory circuits translate these schematic diagrams into wires connected between binding posts or spring clips on the apparatus. Power supplies, batteries, and meters have polarities which must be watched carefully.

2. Every circuit has a main series part in which most of the current flows. Identify this main series circuit and connect it up first, then attach the branches which make up the rest of the circuit.

3. Although in a schematic diagram it is often convenient to indicate a junction point at the middle of a connecting wire, in practice it is necessary to make every junction at a terminal of some apparatus. The results are equivalent if you assume negligible resistance in the connecting wires. Since this assumption is not always admissible, care must be taken not to include a long wire in series with a small resistance. **Never use a meter terminal as a junction point.** Only one wire should be connected to any meter terminal so that it is possible to remove or replace any meter without otherwise disturbing the circuit.

4. A circuit should always be "dead" when it is being made up or when any changes are being made, and the circuit should be checked carefully thereafter before it is powered. A "dead" circuit is one in which at most one terminal of any battery is connected, and no power plugs are inserted into power sockets.

Figure C.1   Common circuit Symbols

One of your requirements for Physics 53 is writing a technical report. You can write the report on any experiment you have completed except for Exp. 1. Pick an experiment that will make an interesting paper and one for which your results are reasonable. There is no page limitation but you should restrict yourself to 1500 – 2500 words. This allows your report to be complete and informative, but also requires concision. Templates in Microsoft Word and LaTeX format are available for download at the Physics Kiosk.[1] Starting with one of these template files—which has a number of helpful suggestions and examples—will streamline the writing process somewhat. However, it is only fair to point out that real research papers go through many, many drafts before being submitted for publication— sometimes dozens. It is not easy to write clearly and concisely for most mortals, and so we write and rewrite and rewrite. . .

You are strongly urged to submit a complete draft of your paper to your professor for review. This draft should be a serious attempt at a complete and finished document — the more effort you put into this draft, the more your instructor will be able to help you improve it. Give your professor at least one week to read the draft, and leave yourself time to incorporate the feedback before submitting the final paper. The technical report counts for a significant portion of your course grade (see Introduction), so please take it seriously.

Learning to produce a good technical paper takes time, and for that reason you will be given several opportunities (groan) at HMC to improve your technical writing. We know from alumni surveys that the great majority of Harvey Mudd alumni have found that writing skills were much more important to their careers than they anticipated, and that they wish they had worked harder on writing while in college. Save your reports after they are returned to you to serve as models for future reports.

Goals

Your purposes are:

- to entice the reader to continue beyond your title and abstract to read the entire paper;

- to provide appropriate technical background[2] to motivate the experiment you conducted;

- to explain clearly the relevant procedural issues and results;

- to summarize key experimental findings in figures (and occasionally in tables);

- and to provide conclusions of your findings that are as quantitative as possible.

Strategy

Although it is tempting to begin writing before you have finished analyzing your results, *don't*. The linchpin of a physics paper is the graph (or graphs) that summarize your principal results. Frequently, this graph displays a comparison of experimental data and a theoretical curve, with a panel of residuals showing the discrepancy between the two. An example is shown in Fig. D.1. Until you have prepared this graph, you don't have a story to tell. So prepare the graph first.
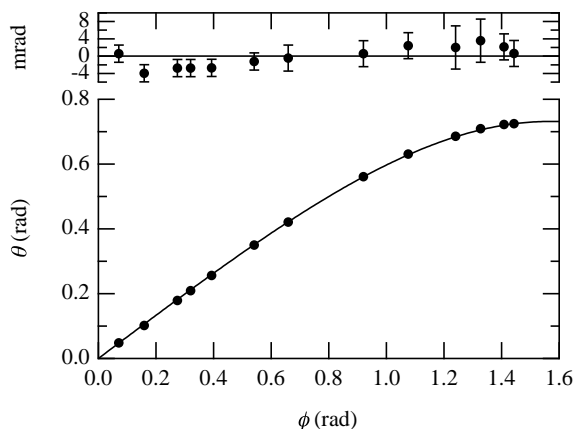

Figure D.1 Angle of incidence θ (in acrylic) vs. angle of refraction ϕ (in air) at $\lambda = 632.8$ nm. The smooth curve is a weighted least-squares fit to the equation $\theta = \arcsin(\frac{\sin\phi}{n})$, for which $\tilde{\chi}^2 = 1.0$. The fitted value of n is $1.xxx \pm 0.002$. [I get to hide my value, but you don't!]

If you have questions about your analysis, and how to prepare this graph, speak to your professor—early! When it's ready, bring him or her a copy of the graph to discuss it and to make sure that your analysis is sound.

With the revised figure in hand, you have the focus of your paper. Everything that appears in your paper before this figure leads up to the figure; everything that appears after it discusses its significance. Hold this notion firmly in mind as we review the sections of the paper.

---

[1] http://www.physics.hmc.edu/kiosk.php

[2] In a real technical paper you should also provide a synopsis of the technical history that underpins your experiment.

Sections of a Technical Paper

The organization of a technical report follows a formulaic pattern, which is illustrated in Table D.1. The elements of the paper are listed in the order in which they appear in the paper, *not* the order in which you should write them! Although the abstract appears third on the list, you should probably leave it to the very last. Abstracts must be extremely concise and they can be a challenge to write. It is better to build momentum by writing some of the easier sections first.

Theory

Assuming that you have completed the analysis of your data, and that you have prepared the figure(s) that summarize your results, I recommend beginning with the **theory** section. For a Physics 53 technical report, this is apt to be a concise presentation of the theory from the laboratory manual. Write for an educated person who understands mathematics and who has some science background, but who has not conducted the experiment you did. The theory section should acquaint or remind your audience of the crucial theoretical background to your experiment, without attempting to fill in all the steps. Make sure to define all symbols you employ, and to explain the mathematics in plain English. Omit steps in a derivation that are merely algebraic manipulations.

Number equations so that you can refer to them later in the text. Remember that variables are set in italics (x), but functions and numbers are set in Roman (upright) font:

$$f(x) = \sin(2x + \phi) \qquad (D.1)$$

If you are using Microsoft Word, be sure to use the Equation Editor to enter equations in proper format. You can find it at **Insert** | **Object**, and then selecting **Microsoft Equation 3.0** or something similar from the list of objects. If you get tired of running this gauntlet, you can assign this task to a tool in the toolbar and/or a key combination using the **Tools** | **Customize...** command.

Experiment

In writing this section you should have in mind a reader who has your background in physics and is familiar with typical laboratory apparatus, but has no knowledge of the specific experiment you are describing and probably has no intention of repeating it. Therefore, you should describe your experimental apparatus and the procedures you followed only insofar as it is necessary for him or her to understand how you made your measurements. A careful drawing of the apparatus is often useful here.

In this regard it is important for you to understand the distinction between a technical report and a labo-

Table D.1: Elements of a technical paper.

Title	
Author	author's name author's institution or company date of submission
Abstract	Summary of the principal facts and conclusions of the paper. Usually fewer than 100 words.
Introduction	Concise discussion of the subject, scope, and purpose of the experiment. It should be accessible to the educated non-expert.
Theory	Succinct development of the theory related to the experiment.
Experimental methods	Brief description of experimental apparatus and procedures.
Results	Summary of important data and results, using figures (graphs) wherever possible. Include error estimates for all relevant quantities. Note that data are typically presented graphically, not in tabular form, although there are exceptions. If you feel a table is more appropriate, discuss this with your instructor.
Discussion or Conclusions	Further analysis and discussion of results. Specific conclusions. Recommendations if needed. "Graceful termination."
References	The complete list of sources cited in the text.

ratory manual. The latter is designed primarily to give instructions on how an experiment is to be carried out, while the former describes to the reader what the author has done. Though it is sometimes necessary in a technical report to describe the procedures that were followed in carrying out the experiment, such material should be kept to a minimum, and *under no circumstances should it instruct the reader as though he or she were a student in a laboratory.*

Data and Results

State first (or show first) your data before providing an interpretation. If you averaged 10 measurements of the pendulum period, simply state that the average of 10 measurements was 1.87 s with a standard error of 0.08 s.[3]

Usually, you will present your data in graphical form, as in Fig. D.1. Occasionally, a table is more appropriate. However, do not display raw data in a table.

Figures and tables must be introduced and discussed in the text; they cannot be placed "inline" with the text, and they certainly cannot be used to begin a section. Examples of the proper way to introduce figures abound in this manual.

Some form of error analysis should be included with your results, and the uncertainty of the final result should be stated clearly. The analysis of errors must be thorough and convincing. Your goal should be to persuade a skeptical reader that your quantitative results are believable. This will often require a more thorough analysis of experimental errors than was done in your lab notebook.

Discussion and Conclusions

This section will often contain further analysis and discussion of your results in order to support and clarify your conclusions. Your final conclusions should be quantitative, precise, and specific, but they should not be simply a summary of the experimental results. Rather, they should be convictions arrived at on the basis of evidence previously presented and analyzed.

Until this section, the paper has grown progressively narrow, beginning with a broad introduction and tapering down to the nitty-gritty of your actual experiment. Once you have presented your most careful analysis, it is time to open up the discussion again, placing your work in a broader context. Make sure that you fulfill any promise you make to the reader in the Introduction as to what your paper would demonstrate or prove. You may want to include suggestions for further work.

Introduction

The Introduction needs to apprise the reader quickly of the topic and scope of the paper, to engage the reader, and to persuade her or him to continue reading. It may be the most challenging part of the paper to write. The first paragraph of the Introduction is particularly critical, since it plays a major role in determining the reader's attitude toward the paper as a whole. You will probably need to rewrite this section more than other sections.

- Make the precise subject of the paper clear relatively early in the Introduction. You should assume your reader is someone with essentially your background in physics but with no particular knowledge of the experiment you performed. Thus you should include background material only to the extent necessary for the reader to understand your statement of the subject of the paper and to appreciate the scientific reasons for your doing the experiment to be described.

- State the purpose of the paper clearly. This statement should orient the reader with respect to the point of view and emphasis of the report and what he should expect to learn from it. A well-done Introduction will also be of great help to you in providing a focus for your writing and in drawing your final conclusions (see below).

- Indicate the scope of the paper's coverage of the subject. State somewhere in the introductory paragraphs the limits within which you treat the subject. This definition of scope may include such topics as whether the work described was experimental or theoretical (In this particular case, the work is almost certainly experimental), the exact aspects of the general subject covered by the paper, the ranges of parameters explored, etc.

An example is given below in the section on passive voice.

Abstract

The abstract should be a concise and specific summary of the entire paper. It should be as quantitative as possible and include important results and conclusions. Including all this in fewer than 100 words takes careful thought (and probably considerable rewriting). An example:

The speed at which light propagates through air at room temperature and atmospheric pressure was studied using a time-of-flight technique. A value of $(2.995 \pm 0.003) \times 10^8$ m/s was obtained, with the dominant error arising from the bandwidth of the oscilloscope used to observe the light pulses. This value is consistent with literature values for c and the refractive index of air.

---

[3]You may also report this result as (1.87 ± 0.08) s. Note that "standard error" and "standard deviation of the mean" are synonyms.

References

Oy vey! Reference formatting is one of the dumbest things in technical writing. Your paper should have references to *published* literature, where appropriate, but the format of those references as they appear in standard physics journals exhibits a remarkable (and pointless) variety. Every journal seems to have its own preferred way of doing things. Physical Review uses one style, except for Physical Review B, which uses a different one. Neither includes article titles, although IEEE style does. Physical Review used to put references at the bottom of the page; now they appear as endnotes and all you see in the body of the text is a number in brackets (or a superscripted number, in *Phys. Rev. B*). Neither approach is as civilized and useful as an (author, date, page) in-text citation style, which is sadly uncommon in physics publishing.[4] Unless your professor tells you otherwise, you may use either numbers or (author, date, page) citations in the text of your report. No matter what the format, though, the complete reference must appear in the list of references at the end. If you use a numeric style, subsequent references to the same source should use the same number, and the list of references should appear in the order of first citation. For the (author, date, page) format, alphabetize the references by author.

- No item should appear in the list of references without specific citation within the body of the text.

- Cite published and commonly available sources (not the lab manual).

In the References section at the end of the paper, include the complete bibliographic information for each of the cited works. Following is a list of references styles that roughly follow the American Institute of Physics Style Manual. Note that the titles of majors works are *italicized*, and volume numbers are set in **boldface**.

1. (journal article) Gale Young and R .E. Funderlic, J. Appl. Phys. **44** (1973) 5151.

2. (translated journal article) V. I. Kozub, Fiz. Tekh. Poluprovodn. **9** (1975) 2284 [Sov. Phys. Semicond. **9** (1976) 1479].

3. (book) L. S. Birks, *Electron Probe Microanalysis*, 2nd ed. (Wiley, New York, 1971), p. 40.

4. (book section) D. K. Edwards, in *Proceedings of the 1972 Heat Transfer and Fluid Mechanics Institute*, Edited by Raymond B. Landis and Gary J. Hordemann (Stanford University, Stanford, CA, 1972), pp. 71-72.

5. R. C. Mikkelson (private communication).

6. James B. Danda, Ph.D. thesis, Harvard University, 1965.

Passive Voice

In *The Elements of Style*, Strunk and White suggest using the active voice as much as possible. This is a worthy goal in scientific writing, but often the role of the investigator is irrelevant to the topic at hand. Because the passive voice emphasizes the phenomenon not the investigator, it is frequently more appropriate in scientific writing. Compare the following introductory paragraphs, the first taken from a student's technical report, and the second a rewritten version:

*By running a current through a coil of wire, **we** can produce a magnetic field which will deflect the needle of a compass. When **we** have done this, **we** can take a bar magnet and position it such that the needle is no longer deflected. Once this has been done, **we** need only to measure the distance from the needle's center to the magnet. Using this and the properties **we** have already measured from the magnet and the coils, **we** can calculate the current through the loop. Because **we** only took measurements of mass, length, and time **we** have determined a value for current not based on any other electrical or magnetic quantities. [Emphasis added.]*

Hans Christian Oersted discovered in 1820 that a current flowing in a wire generates a magnetic field proportional to the current.[1] By comparing this field to that produced by a calibrated permanent magnet, and by using a geometry in which the constant of proportionality between field and current can be readily calculated, it is possible to measure the current—an electrical quantity—by measuring only the mechanical quantities mass, length, and time. C. F. Gauss first proposed a technique for making such an absolute determination of current in 1823.[2] We report here a modification of his technique which yields current measurements having an accuracy of 1%.

In my opinion, the use of the first person in the first paragraph does not add useful information; rather, it detracts from the message that these phenomena exist independent of the observer. The second paragraph uses some active voice and some passive voice, throws in some historical information, and keeps the discussion conceptual, as befits an introduction.

---

[4]Bibliographic database software, such as Endnote or ProCite, can simplify the task of reference formatting, but is apt to be overkill for the technical report. These programs understand the styles of a vast number of publications, and can format the bibliographic information in their databases to match the requirements of whichever publication you target. You can also store notes and other information associated with references in your bibliography, and reuse this information in multiple publications. For extended research projects, such software can be extremely convenient.

Make no mistake: good writing of any kind, whether technical or not, requires practice and a willingness to rewrite. Simply put, *it's hard!*

Do's and Don't's

Do's

1. Double space the report, except for the abstract.

2. Divide the report into clearly labeled sections (Abstract, Introduction, Theory, . . .).

3. Number all pages.

4. Place tables and figures as soon as possible after they are referred to in the text.

5. Give each table and figure a complete title and/or self-descriptive caption.

6. Give each table and figure a complete title and/or self-descriptive caption.

7. Define all symbols used.

8. Number all equations. (Place numbers near right-hand margin.) Each equation should be on a sepa-rate line. Use complete sentences and fit all equations into the sentences that introduce them.

9. Avoid imprecise, vague, or unclear language. Carefully define your terms and use them consistently. Avoid jargon and excessive abbreviations.

10. Tell a story. Organize your information so that the narrative proceeds smoothly and logically.

11. Use SI units unless there is a compelling reason not to. That is, use μm or nm, but not Å.

Don't's

1. Do not repeat every sordid detail from a lab diary; report the logically crucial ones.

2. Do not distort unpleasant realities; call it like it is. Don't attempt to cover up flaws in your data; do explain departures from theory when you can.

3. Do not scan someone else's figure.

Technical Typography

1. Algebraic symbols (variables and constants) should be italicized. $E = mc^2$, not E = mc². In Microsoft Word, modify the fonts yourself or use the Equation Editor. In LATEX type inline equations inside dollar signs: `$E = mc^{2}$`.

2. Subscripts are italicized when they refer to symbols, but Roman when they are words

$$B_{\rm bar} = B_x \tan\phi \qquad (D.2)$$

LATEX: `B_{\rm bar} = B_{x} \tan\phi`

3. Vectors are written in boldface (usually not italicized), but their magnitude is set in italics. Functions such as sine are not italicized.

$$\boldsymbol{\tau} = \mathbf{M} \times \mathbf{B} \quad \text{or} \quad \vec{\tau} = \vec{\mathbf{M}} \times \vec{\mathbf{B}}$$
$$\tau = MB\sin\theta$$

LATEX: `\boldsymbol{\tau} = \mathbf{M} \times \mathbf{B}`

4. Refer to the second figure as "Fig. 2". Write out "Figure 2" if this starts the sentence.

5. Figures should have informative captions. Most people "read" a paper by looking at the title, abstract, and figures. By making the figures clear and the captions informative, you increase your chances of getting the reader to read the text.

6. Number pages.

7. <u>Do not underline</u>. Book titles should be *italicized.* Volume numbers should appear in **bold**. Example: D. Quayle, J. Irrepro. Res. **21** (1996) 1234. [Note that "1234" is the page number of the article.]

8. Get your symbols for units right: "Hz" not "hz", "Ω" not "Ohm", "30°" not "30 degrees". Note: units are not pluralized: don't use "secs", "hrs", etc. Never ever screw up the capitalization of SI prefixes: there's a crucial difference between MHz and mHz.

9. Abbreviations (e.g., RLC, rms, etc.) must be defined when first used (except in a title, where they should usually be avoided).

10. A number and the unit that modifies it should be separated by a (nonbreaking) space (e.g., 50.1 Hz, not 50.1Hz). The space should be a hyphen when the number-unit phrase functions as an adjective before the verb (e.g., A 50-Hz signal was applied across points A and B). In Microsoft Word, head to the **Insert** | **Symbol** command and select the **Special Characters** tab to find how to insert a nonbreaking space. In LATEX it is typed with a tilde (`~`).

11. To type a multiplication symbol in Microsoft Word, use the **Insert** | **Symbol** and select Symbol font to get the palette of mathematical symbols. Then select the multiplication sign and click **Insert**. In LaTeX use the `\times` command in math mode.

12. If you use superscripts as endnote markers, they should appear after the punctuation mark (e.g., ". . . is given by the Biot-Savart law.[1]") If you use numbers in brackets, they go before (". . . is given by the Biot-Savart Law [1].").

13. Do not start a sentence with a symbol.

14. Doan furget two speel cheque. An prufreed. Swapping papers with a friend for proofreading is often very helpful.

Proofing marks

stet. Latin for "let it stand." Leave as it was (I screwed up and marked something and then decided to "unmark" it).

awk. awkward—rewrite the sentence or paragraph

it. italics