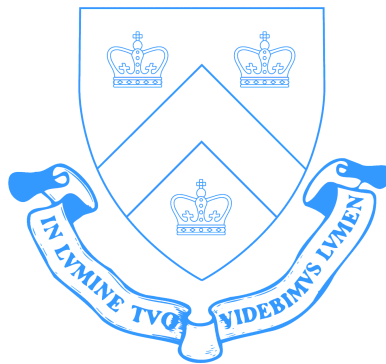


Experiments in Physics

Physics 1494

Columbia University



Department of Physics

Fall 2022 Edition

Contents

Introduction	v
1 Velocity, Acceleration, and g	1
2 Projectile Motion and Conservation of Energy	13
3 Magnetic Fields	21
4 e/m of the Electron	29
5 Polarization and Interference	37
6 Interferometer	49
7 The Spectrum of the Hydrogen Atom	61
8 Capacitance and the Oscilloscope	71
9 AC Circuits	85
10 Absorption of Beta and Gamma Rays	97
A Measurement and Error Analysis	109

Introduction

Purpose of the Course

The purpose of performing physics experiments in the laboratory is not only to reinforce the material you learn in physics lecture, but also to enhance your understanding of how scientific knowledge actually progresses. As you “perform” the experiments in this manual, you will encounter a messy reality quite unlike the tidy exercises of your textbooks — a reality of inconclusive results, subtle equipment problems, and experimental biases. It is our hope that this course will give you a taste of the difficulties and thrills of real experimentation; and that you come away with a greater appreciation of how the laws of nature are teased out of “real world” measurements.

In addition to engaging with the practicalities of the scientific method, another important skill that we hope to instill is the “art” of scientific writing. At the conclusion of each experiment, you will be expected to submit a lab report condensing the essential details for other readers to both understand, and be able to recreate your experiment for verification if they wished. Being able to convey your message in a simple, straightforward and efficient manner is a crucial skill regardless of the field you choose to pursue in the future.

Course Outline and Policies

This course consists of two parts: a required fifty-minute lecture component at the beginning of each experiment cycle, and a weekly three-hour session with a lab instructor. Many of the experiments in this manual require a large number of measurements and prior work. The lecture is meant to prepare you for each week’s lab works, thereby minimizing your time (and difficulties) in the lab session. For this reason, half of the lecture is typically devoted to a review of the physics relevant to the experiment, while the second half will describe the experiment itself. Particular emphasis is given to common practical issues and how to solve them, as well as tips on how to perform data analysis.

You will attend the same lab section each week and, under the guidance of a lab instructor, you will work on each experiment in a group. You will have a full week to write up a brief report detailing your measurements and conclusions. You will submit completed reports online prior to the next lab session.

The grade for the course is dominated by your weekly reports. In the laboratory, you may share data with your lab partner and discuss all aspects of the experiment. However, you are not to share anything beyond what is given to you, e.g. data points taken from a given figure, with other lab groups in your own section or any other section. Moreover, the data analysis and interpretation that goes into the lab report must be performed independent of your partner, and written in your own words. Failure to comply with this rule is considered plagiarism.

You will also be graded on your preparation for and participation in the experiments. Read this manual before coming to the labs, and even though we are not stepping in actual laboratories, please follow laboratory etiquette: no eating, no drinking, and no cellphones use during the lab sessions. Should you *really* need to do any of the above, like taking an urgent phone call, please do not disturb others and kindly mute yourself.

Finally, if you miss an experiment, you are not allowed to make up the work at another time; we do not accommodate make-up labs at another time. During the semester, you are allowed up to three absences. Greater than this amount means automatic failure of the course. Among the three allowed absences, at most two of them can be excused, which will not be counted towards the final grade. You are responsible for obtaining a proper excuse and submitting it to your lab instructor prior to your lab session. A third absence will automatically be a zero and will be dropped as the lowest grade.

Lab Report Format

General Remarks

Writing a lab report is the primary way your TA will know what you have done during the lab, how well you understood the experiment, and whether or not you know how to process the results. This is designed to serve as your first introduction to the world of scientific writing.

The formatting of your report should follow the "**IEEE Transactions on Magnetism**" format (http://ieeauthorcenter.ieee.org/wp-content/uploads/Trans_Magnetism_instructions.pdf), with some simplifications that will be detailed below.

Your lab report should include, but not necessarily be limited to, the following features:

- Concision. There is no need to include huge excerpts from a textbook, dozens of irrelevant plots, or a lengthy derivation (unless the lab manual specifically asks for it). You should consider 2-6 typed pages, including plots, to be the typical requirement for a lab report. There is a hard limit of 6 pages for each lab report.
- Graphical representation of data: for example, a histogram or an xy plot. Plots must include error bars on the data points, be clearly labeled, and be large and easy to read. Captions should follow figures as a description of the plot.

- Proper presentation of results. When you present your data, all measurements should contain an estimate of the uncertainty and the units of measurement.
- A description of the statistical and systematic errors that affected your measurement. Again, this description should be concise but well-developed.
- A discussion containing all relevant information and reasoning, allowing the reader to validate your conclusion.
- Past, Present, Future Tense. You should use past tense to describe results you have done, results from published papers should be described in present tense and only experiments you plan to do in the future should be described in future tense.

Using a Logical Format

To provide a disclaimer, if you actually go on further in research and read other scientific papers, you will notice that there is no “one” correct way to write a paper, and the style often depends on journal requirements, field of research, author styles, etc. As you progress on after the course, you will develop your own flavor for scientific writing, but as this is still an introductory course, the expectations will be a little firmer to build a stronger foundation for you that you can build on in your future careers.

You may want to structure your report using the following framework:

1. **Abstract (Optional):** The abstract is usually the last section one writes, and acts as a short paragraph summary of the contents of the entire report. A reader should be able to glean the essentials of what the rest of the report is going to cover, from reading the abstract.
2. **Introduction:** The introduction should provide the background for a reader unfamiliar with the field of the experiment, to jump into it. It should describe the scientific question you are looking to probe in the experiment, and motivate why you (and the reader) should be interested in it. Depending on the scenario, it can also be a good idea to very briefly mention the conclusion of the report as well.

Any theoretical equations or derivations should be developed in the Introduction and Method section. Anything that relies on a description of the particular setup should be presented in the method section, while more general concepts may be described prior in the introduction.

3. **Method:** The method section should contain a summary of the experiment *in your own words*. You are NOT to copy large chunks of the lab manual. A simple sketch of your apparatus accompanied by a one or two-sentence description is often a good place to start.

A good method section will also contain a description of the important parameters measured in the experiment and expectations for their relationships to one another (ie. what relationship, formula

or model describes the system you are studying). Finally, you can mention some of the details of your analysis in this section; e.g., “We perform a fit to the data of the form $y = mx + b...$ ”

4. **Results:** You should present the results of your measurements and analysis (i.e., error propagation) here. You don't need to include lengthy tables; it is often preferable to include plots and final values, with uncertainties. The plots, which must be clearly labeled, should speak for themselves. However, you should always contextualize your figures and fit analysis. This section should continue the narrative of the rest of the report.

Moreover, never forget that, whenever you are reporting an error associated with a certain quantity, it must be clearly explained. If the uncertainty is associated with a raw data point, you should justify why you assigned this uncertainty, often based on your specific measuring procedure. On the other hand, if the uncertainty is associated with a derived quantity (e.g. from the slope of a linear fit), then you should explain how you propagate it or derive it from a collection of measurements.

5. **Discussion:** In this section, you should demonstrate your grasp of the experiment by discussing the statistical significance of your measurements. Comment on whether or not your observations differ significantly from your expectations. If so, you should consider whether or not this could be a reasonable physical effect, or if in fact your experiment was affected by a systematic bias. When discussing possible sources of bias, you should be as specific and quantitative as possible. Simply writing that your observations were affected by vague factors such as “human error” or “bad equipment” is not sufficient.

In many physics experiments, there are many sources of error, but a dominant error (the source that contributes the most to the uncertainty) is often identified, and measures to address this particular error source will often be suggested.

You should also discuss the *precision* of your observations — i.e., the size of the statistical fluctuations in your answer — and think about how to improve it. (Do this even if your answer does not differ significantly from expectations.)

6. **Conclusion:** In the conclusion, you quickly summarize for the reader your results, the precision and accuracy of these results relative to expectations, and the possibility of improving your measurements.

This framework can help you organize your work, but remember, it is up to you to make the report readable and transparent. In writing your report, you should still aim to present your information in an effective manner, and you are welcome to explore tweaking this structure to do so.

Experiment 1

Velocity, Acceleration, and g

1.1 Purpose

The purpose of this experiment is to study the motion of an object undergoing uniform acceleration. Using a “frictionless” air track and a computer, you can observe the one-dimensional uniform motion of a metal rider. You will measure the constant velocity motion of the rider before and after it collides with an elastic bumper, and then incline the air track to accurately measure g , the local acceleration due to gravity.

1.2 Introduction

The motion of an object is typically described in terms of four quantities: time (t); position (\vec{r}); velocity (\vec{v}); and acceleration (\vec{a}). Velocity is the rate of change of position with respect to time, and acceleration is the rate of change of velocity.

$$\begin{aligned}\vec{v} &= \frac{d\vec{r}}{dt} \\ \vec{a} &= \frac{d\vec{v}}{dt} = \frac{d^2\vec{r}}{dt^2}\end{aligned}$$

The quantities \vec{r} , \vec{v} , and \vec{a} are vectors because motion can occur in more than one dimension. Moreover, position and velocity are generally complicated functions of time. Even along one dimension such as $x(t)$, the motion of a body can be difficult to analyze, since $a(t)$ may have some nontrivial time dependence. However, it turns out that the time dependence of position and velocity simplify greatly when an object undergoes constant acceleration $a(t) = a$:

$$\begin{aligned}v(t) &= \int a \, dt = v_0 + at \\ x(t) &= \int v(t) \, dt = x_0 + v_0t + \frac{1}{2}at^2\end{aligned}$$

These kinematic expressions were first deduced in the sixteenth century by Galileo, who systematically observed objects rolling down inclined planes. Starting the objects from rest, he found that the displacement of a body down the plane is proportional to the square of the time it is in motion.

In this experiment, you will perform a related study of motion with constant velocity and constant acceleration. Ordinarily, it is difficult to examine these kinds of motion with precision, since objects in free fall tend to move too rapidly, and frictional forces tend to arise in most everyday situations. These factors hinder a direct observation of the underlying physical principles of motion, and in fact this is one of the reasons why these principles were poorly understood until Galileo's careful experiments.

1.3 Experiment

In this lab, it will be possible to study motion in the absence of almost any friction by using a rider on a “frictionless” air track. The air track has rows of small air jets running down its side, which support the rider on a thin film of air and allow it to float just above the track. When the track is level and the rider is given a slight push, it will move with constant velocity; when the track is slightly inclined, the rider will experience a small acceleration due to the component of gravity that is parallel to the track.

The rider and air track are pictured in Fig. 1.1. To observe the motion of the rider, we need to make accurate measurements of its position at regular intervals of time. For this we employ a sonar device called a Sonic Ranger.

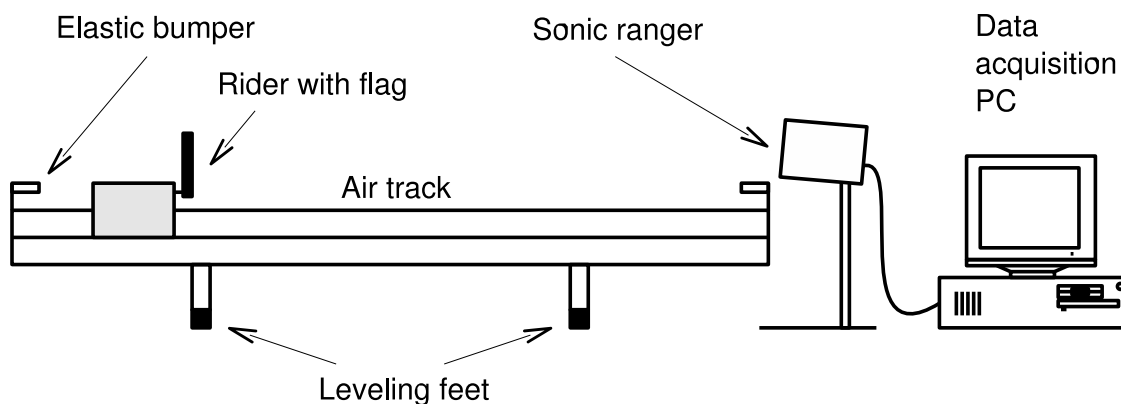


Figure 1.1: Setup of air track and PC.

The Sonic Ranger sends out discrete pulses of sound waves, which travel at some velocity v_s to an object on the track and are reflected back to the source. As reflected waves are detected by the Ranger, their arrival times are read out to a PC, which calculates the distance x to the object based on the time τ it takes the signal to make a round trip to and from the object:

$$x = \frac{v_s \tau}{2}$$

If a series of such measurements is made in rapid succession, then the computer can reconstruct the position of the rider over some time interval, using this information to calculate quantities such as the “instantaneous” velocity or acceleration of the rider as a function of time. The Sonic Ranger driver software calculates the velocity v at time t using the formula

$$v(t) = \frac{\Delta x}{\Delta t} = \frac{x(t + \Delta t) - x(t)}{\Delta t},$$

where Δt is the time interval between each emitted pulse. We operate the Sonic Ranger at a frequency $f = 20$ Hz, so that $\Delta t = 1/f = 0.05$ s. As long the rate of change of $v(t)$ is not too extreme, this Δt is small enough to give us a fairly accurate approximation to the instantaneous velocity.

1.4 Procedure

The procedure in this experiment is split into two parts: a setup segment in which you will familiarize yourself with the equipment; and a measurement segment during which you will observe motion with constant velocity and constant acceleration. While you should make yourself comfortable using the air track, Sonic Ranger, and PC, try to spend the bulk of your time in the laboratory working on the measurements.

Getting Started

When you begin the experiment, you will want to load the DataStudio file (**ENG-VELOCITYLAB**) for this lab. To do so, you may have to log on to the computer with the user name **student** and the password **student**.

You should now see the screen shown in Fig. 1.2, with two empty graphs displayed: position x (m) and velocity v (m/s) versus time t (s). To the left of the plots is a summary window, which you can use to browse through your data runs. The computer is now ready to take data.

To collect data, single click on the **Start** button. The plots will update and rescale automatically as data arrive. Clicking the **Stop** button ends data acquisition. To begin collecting new data, simply click the **Start** button again; the old trace will disappear (don't worry; it is still accessible via the summary window - click on Summary) and new data will plot in the two graphs.

NOTE: Please do NOT save your session in order that no change is made in the initial settings.

Testing the Sonic Ranger

The Sonic Ranger has two modes of operation, which are controlled by a switch on top of the Ranger module:

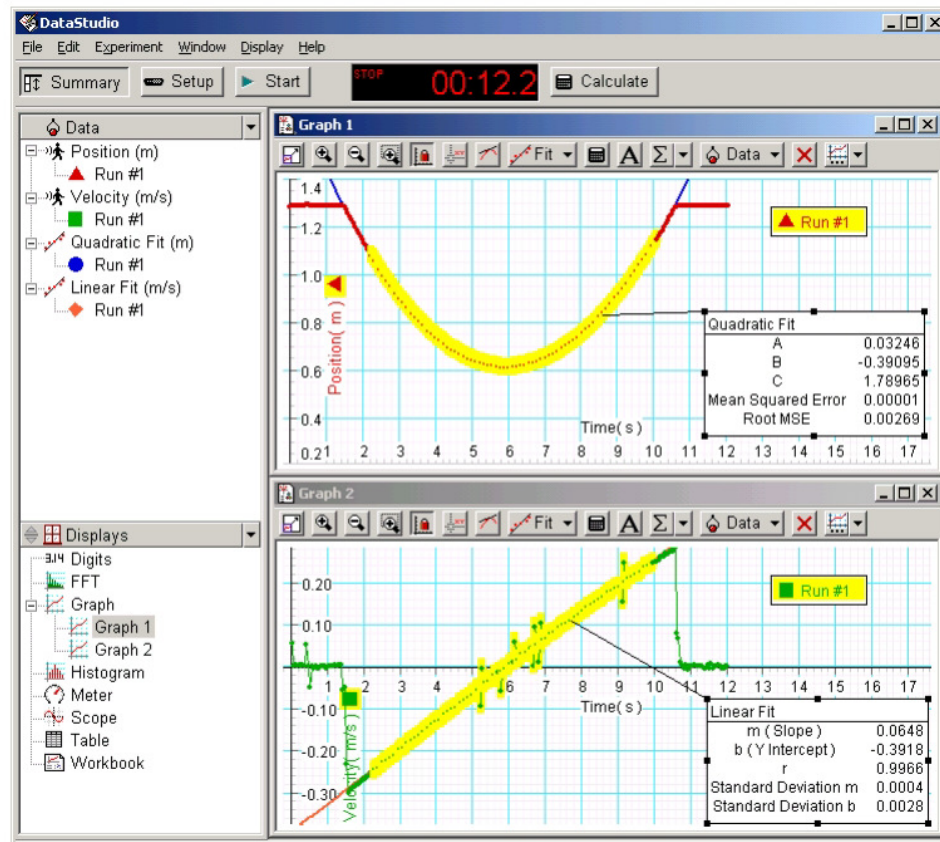


Figure 1.2: Main window of sonic ranger VELOCITYLAB software.

1. Person Mode: The Ranger emits a wide conic sonar beam for observing extended objects.
2. Cart Mode: The Ranger emits a narrow collimated beam, useful for observing a small object.

Before you start actual taking data, it is useful to familiarize yourself with the Sonic Ranger and the DataStudio software by trying the following simple experiment:

- Switch the Ranger to *Person Mode* using the switch.
- Stand about two meters from the Ranger and have your partner point it at you.
- As your partner clicks **Start**, walk slowly toward the Ranger with constant speed.
- End the data collection after about ten seconds.

Inspect the graphs and see whether or not you were able to walk with constant velocity. You may have to rescale the position and velocity axes to properly view the traces. In order to do this, sweep the mouse over the y -axis until the icon changes from an arrow to a small curly line. Then hold down the

left button and drag the mouse up or down to zoom in and out of the plot.

Q: Look at the position plot. Does x versus t look like a straight line, as you would expect for motion with constant velocity? Does the velocity plot appear as you expect?

During the next five minutes or so, try out some of the other features of the software, such as curve fitting. To perform an unweighted least squares fit to the data, drag the mouse over a portion of the trace while holding down the left button. Some of the data points should now be highlighted in yellow. Click on the **Fit** button at the top of the plot. (If the toolbar does not show the **Fit** button, you may need to modify the **Main Toolbar Settings** under **File**→**Options**.) A listbox will open up displaying a large set of functions: choose *Linear*, which fits the selected data to a function of the form $y = mx + b$. When you perform the fit, a textbox will pop up in the plot window, displaying values for the slope m , intercept b , and uncertainties in both.

Note that you can view old data traces by left-clicking the **Data** button in the plot window and selecting the name of the data run. All fits are listed under the Summary window.

Preparing the Sonic Ranger for Use with the Air Track

Care of the Rider and Air Track: *Please take care when using the rider and track. DO NOT let the rider sit on the track when the air is off, and DO NOT let the flag-side of the rider collide with the elastic bumper, since both of these can cause the rider and air track to scrape against each other, permanently damaging the equipment.*

To conduct the rest of the experiment, you will need to set up the Sonic Ranger over the air track. Simply follow these steps:

- Set the Ranger at the right end of the track, pointed horizontally and centered along the track.
- Switch the Ranger to Cart Mode.
- Turn on the air using the pump under the table.
- Gently place the rider onto the track.
- Start taking data. Check the Ranger alignment by giving the rider a small push up the track.
- The PC should plot the rider's position smoothly throughout the length of the track. If you see any jumps in the plot, you will need to align the Ranger more carefully.

Leveling the Air Track

To study constant velocity motion, it is necessary to level the air track as much as possible. You will find that when you first put the rider on the track, it will tend to drift in one direction or another. What you want to do is level the track such that when you place the rider at any position along its length, the rider stays more or less stationary.

To level the air track, use the two adjustable feet under the left side. The right foot is not adjustable, but you can raise or lower it using sheets of paper (do not use the shims!). It may not be possible to completely level the track, but try as best as you can to eliminate irregularities in the motion of the rider. Sometimes the track might have bumps on its surface. One thing you can do to make sure that it is aligned in the best possible way is to place the cart on different locations along the track. If its motion is slow and in random directions, your air track is fairly horizontal.

A Note About Distances

The Sonic Ranger is unable to detect objects that are too close, because it requires the round trip travel time τ of sonar pulses to have some minimum size. If τ is too small, the software cannot correctly interpret the data. Therefore, always work with distances greater than 20 cm from the Sonic Ranger.

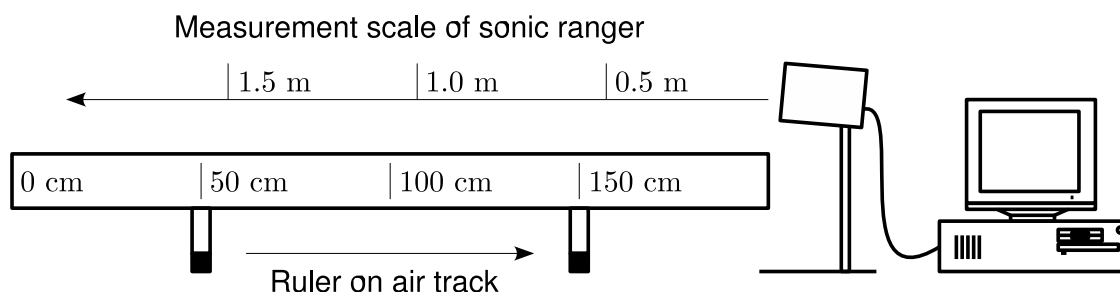


Figure 1.3: The measurement scale of the Sonic Ranger starts at its faceplate and increases toward the left, *opposite* to the ruler on the air track.

Note also that the scale on the air track increases as you read toward the right. Since the Sonic Ranger is aimed left, its scale increases toward the left, with the origin at the faceplate of the module, not the right end of the track (see Fig 1.3). Keep these two scales separate in your work. Use the readings from the Sonic Ranger for all experimental calculations, and use the scale on the air track only for positioning the rider in the same place when you are repeating an experiment over several trials.

Measurements

The data-taking segment of this experiment consists of two parts: motion with constant velocity, and motion with constant acceleration.

Motion with Constant Velocity

Once you are satisfied that the track is sufficiently level, set the rider at the 150 cm mark. You will now test the setup by quickly observing the constant velocity motion of the rider.

Observing the Rider: Begin collecting data, and then give the rider a gentle push toward the left. Make sure that you are able to take data over a substantial portion of the return trip after it has bounced off the elastic bumper at the left end, and also make sure that the position data is smooth and without jumps.

Q: Is the velocity graph what you expected? The time scales of the two graphs are always the same, so you should be able to see how the rate of change in the displacement graph corresponds to the velocity. How does motion with constant velocity on the air track compare to trying to walk with constant velocity?

The Coefficient of Restitution: When two objects collide and bounce away from each other, they tend to lose some of their energy in the collision, and the rebound velocity between them is therefore less than the initial velocity between them. This is why objects that are dropped will sooner or later stop bouncing. The elasticity of the collision can be described by e , the **coefficient of restitution**, which is defined as the speed after the collision divided by the speed before the collision:

$$e = \frac{|\vec{v}_f|}{|\vec{v}_i|}$$

A perfectly elastic collision would have a coefficient of restitution equal to one; an elastic “super” ball is a good example of an object whose coefficient of restitution is close to one. You can calculate e for the case of the rider colliding with the elastic bumper using the data you collect in this experiment.

- To measure e , place the rider at the 150 cm mark, click **Start**, and gently push the rider to the left.
- Make sure the software cleanly plots the motion before and after the rider impacts the bumper.
- Using the DataStudio fitting routines, perform a linear fit to the x versus t trace using the data before the collision. Record the slope v_i and the uncertainty σ_{v_i} on the slope obtained from the fit.

- Perform a linear fit to the x versus t trace using the data after the collision. Record the slope v_f and the uncertainty σ_{v_f} obtained from the fit.
- Repeat this measurement at least ten times, calculating e for each trial. Try to impart roughly the same gentle initial velocity v_i to the cart each time.

Gravitational Acceleration

A small constant force can be applied to the rider by inclining the track slightly. The component of gravity which acts on the rider parallel to the air track is equal to $g \sin \theta$, as indicated in Fig. 1.4. When analyzing this motion, use a coordinate system whose x -axis is parallel to the surface of the track. In this case, the constant acceleration of the rider a_x is simply the component of \vec{g} along this axis. By measuring this acceleration, you can infer the local acceleration due to gravity g .

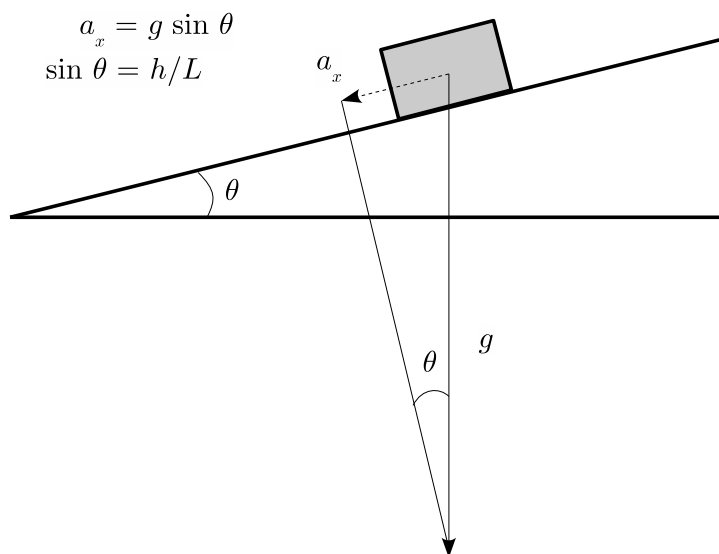


Figure 1.4: A convenient coordinate system for the measurement of g .

To incline the track, use a few of the metal shims provided to elevate the right side as shown in Fig. 1.5. The shims can be used to elevate the right support by some height h . Since the distance between the right and left supports is $L = 1.0$ m, the inclination angle of the track is given by $\sin \theta = h/L$. Therefore, the acceleration of the rider can be written as linear function of h :

$$a_x = \left(\frac{g}{L}\right) h$$

NOTE: Even though some of the shims might indicate some thickness on them, it is always a good habit to use a caliper to measure the total height yourself.

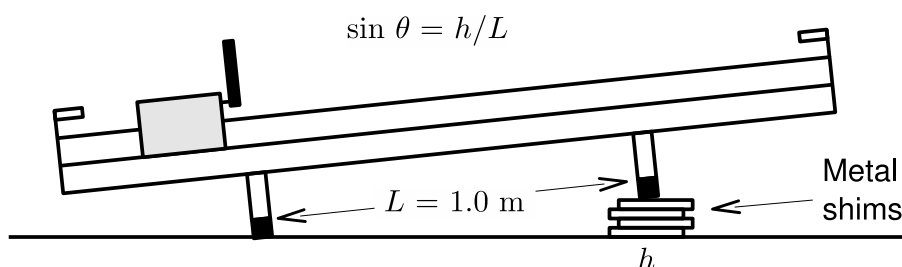


Figure 1.5: Elevate the air track using the metal shims.

You will measure the acceleration of the rider a_x as a function of 5 different heights h , using one, two, three, four, and five stacked shims. A convenient and accurate way to measure a_x for a given height is described below:

- Place the rider at the **150** cm mark and release it.
- Just as the rider reaches the elastic bumper on the left, click **Start**.
- Take data as the rider returns to the right, slows, stops, and reverses direction. Click **Stop** just as the rider reaches the left bumper a second time. Your position plot should look like a nice concave-up parabola, and your velocity plot should be a long straight line (more or less).
- Note how far the rider travels up the slope after the first collision (l_2). This is just the distance between the bottom and the top of your parabola, and can be measured using Smart Tool. Also measure the velocity right after the collision (v_1). This can be found easily from the v vs. t plot.
- Perform a linear least squares fit on the velocity data, recording the slope of the plot (a_x) and its uncertainty.
- Also make a quadratic fit of the form $x = At^2 + Bt + C$ to the position data. How does the quadratic coefficient A compare to a_x ?
- Repeat this measurement at least ten times for each height h . You don't have to perform the quadratic fit each time; just for a few trials. Don't take too long for each measurement, or you will run out of time! Record your measurements of l_2 and v_1 for each run, and also note the position of the elastic bumper on the track.

Summary of data:

- Constant velocity section:
 - At least ten values for v_i and v_f , and corresponding coefficient of restitution, e .
- Gravitational acceleration section:
 - At least ten values for a_x , l_2 and v_1 for each of 5 shim thicknesses, h .

1.5 Analysis

The analysis section of this experiment is fairly straightforward; using a spreadsheet program like Excel, you should complete it with little difficulty.

Coefficient of Restitution

During the constant velocity portion of the experiment, you collided the rider with an elastic bumper and recorded at least ten values of v_i and v_f (plus uncertainties). You will now estimate coefficients of restitution e_j for each pair of (v_i, v_f) measurements. This is largely an exercise in error propagation and statistics, so you will find the average \bar{e} of the observations in two separate ways.

The Unweighted Mean

Proceed as follows:

- Calculate 10 values e_j for each pair of measurements v_i and v_f using the equation

$$e_j = \left| \frac{v_f}{v_i} \right|$$

- For each of the 10 values of e_j , calculate the uncertainty σ_{e_j} by *propagating the errors* σ_{v_i} and σ_{v_f} . Assuming the errors in v_i and v_f are independent, σ_{e_j} is

$$\sigma_{e_j} = \sqrt{\left(\frac{\partial e}{\partial v_i}\right)^2 \sigma_{v_i}^2 + \left(\frac{\partial e}{\partial v_f}\right)^2 \sigma_{v_f}^2}$$

When you evaluate the partial derivatives $\partial e / \partial v_i$ and $\partial e / \partial v_f$, you will find that σ_{e_j} is just a function of v_i , v_f , σ_{v_i} , and σ_{v_f} . Once you have the formula, you can use it directly in a spreadsheet to find the 10 values of σ_{e_j} .

- Calculate the unweighted mean \bar{e} and standard error on the mean $\sigma_{\bar{e}}$. These expressions are given by the formulae

$$\begin{aligned}\bar{e} &= \frac{\sum_j^N e_j}{N} \\ \sigma &= \sqrt{\frac{\sum_j^N (e_j - \bar{e})^2}{N - 1}} \\ \sigma_{\bar{e}} &= \frac{\sigma}{\sqrt{N}}\end{aligned}$$

Note how these expressions *do not* account for the errors σ_{e_j} in each of the individual data points.

The Weighted Mean

Now, calculate the weighted mean \bar{e}_w and its standard error $\sigma_{\bar{e}_w}$ using the expressions

$$\begin{aligned}\bar{e}_w &= \frac{\sum_j e_j / \sigma_{e_j}^2}{\sum_j 1 / \sigma_{e_j}^2} \\ \sigma_{\bar{e}_w} &= \left(\sum_j \frac{1}{\sigma_{e_j}^2} \right)^{-1/2}\end{aligned}$$

These equations *do* account for the individual errors in each measurement, making data points with large σ_{e_j} less important in the sum. Comment on your results. For example, consider the following questions:

- How do \bar{e}_w and its error $\sigma_{\bar{e}_w}$ compare to \bar{e} and $\sigma_{\bar{e}}$?
- Do you see any evidence that the coefficient of restitution depends on the initial velocity of the glider? (Hint: try plotting e against v_i .) Does this result agree with your expectations? Why?
- When calculating σ_e from σ_{v_i} and σ_{v_f} , you should treat the errors as though they are independent. However, are the uncertainties in v_i and v_f truly independent? Or do you think small errors in v_i affect v_f ? Justify your answer.
- Do your values for e appear to be randomly distributed about \bar{e} , or \bar{e}_w ?

Gravitational Acceleration

For a set of air track heights h , you have a collection of acceleration values a_x obtained from linear fits of v versus t . Using the measurements of a_x as a function of h , you will estimate g . Proceed as follows:

- For each height h , you have 10 measurements of a_x . Find the unweighted mean \bar{a}_x of these values and the unweighted standard error on this mean $\sigma_{\bar{a}_x}$.

- You should now have 5 values of h , 5 values of \bar{a}_x , and 5 uncertainties on these means $\sigma_{\bar{a}_x}$. Plot the averages \bar{a}_x as a function of h , and include error bars on the points (given by $\sigma_{\bar{a}_x}$).
- Draw a best fit line of the form $\bar{a}_x = mh + b$ in the plot.
- Using Mathematica or Python, calculate the slope m , intercept b , and the standard errors in both.
- Use the slope m to estimate g ; remember that we expect $a_x = (g/L)h$. Use the standard error in the slope (from the output of your linear fit) to find σ_g .

Once you have completed the calculation of g , consider the following questions:

- Does your measurement of g agree with the nominal value $g = 9.80 \text{ m/s}^2$? If not, how significant is the discrepancy? That is, how many σ 's away from the nominal value is your result? Can this be explained by natural variation in g ?
- Is the intercept $b = 0$ within experimental error, or is there a statistically significant difference? What might cause this?
- Do you expect steeper slopes to be more accurate than shallower slopes for extracting g ? Why? Are your results consistent with this?
- Can you think of any systematic effects that may have been at work in this experiment? Do you see any evidence for their impact?
- What limits the precision of this way of measuring g ? Why is a pendulum measurement more precise? How could you increase the precision of your result — that is, decrease the size of σ_g ?

Estimating Friction Losses

In this section you will investigate the effects of friction, using your measurements of l_2 and v_1 from the gravity section of the experiment. Conservation of energy without friction implies

$$\frac{1}{2}mv_1^2 = mgl_2 \sin \theta$$

For some of the trials compute the following quantity¹:

$$\Delta = \frac{v_1^2}{2a_x l_2} - 1$$

Statistical deviations of this from zero will indicate that there is energy loss due to friction. You do not have to compute it for every trial, just for a small sample (two or three for each height).

¹This quantity is closely related to the coefficient of kinetic friction μ_f .

Experiment 2

Projectile Motion and Conservation of Energy

2.1 Purpose

In this experiment, you will use the trajectory equations of a body in two-dimensional free fall to predict where a projectile lands. The initial velocity of the projectile is determined by applying the conservation of energy for the projectile as it drops through a long inclined tube. You will test your prediction with repeated trials and a statistical analysis of the spread in the landing spots.

Preparation for this Experiment

NOTE: You must prepare some derivations at home; otherwise, you may have trouble finishing the lab in the time given.

2.2 Introduction

This lab should demonstrate the predictive power of applying physical principles correctly, show that predictions correspond to something in the “real world,” and provide insight about deciding what is important in making a measurement.

In order to understand the procedure used in this experiment, it is necessary to review some basics of energy, energy conservation, and projectile motion. Make sure you read and understand the following sections before coming to the laboratory.

Energy Conservation and Frictional Losses

One of the most fundamental principles of physics requires that total energy be conserved in all physical processes. Of course, there are many different kinds of energy, including mechanical energy, heat, chemical energy, and mass¹. Energy conservation implies that one kind of energy can transform into another — for example, an engine converts the heat energy released by burning fuel into mechanical energy. So long as all types of energy are accounted for, the total energy inside a closed system is a fixed constant.

Mechanical Energy

Mechanical energy is a general term that refers to the various types of kinetic energy (translational and rotational) and gravitational potential energy that an object can have. The kinetic energy of a point particle of mass m moving with velocity v is given by the well-known equation $K = 1/2 mv^2$.

In this experiment, we deal with a rolling ball, an extended object that not only translates but also rotates about its own center of mass. A proper estimate of its kinetic energy requires us to account for both types of motion:

$$\begin{aligned} K &= K_{\text{trans}} + K_{\text{rot}} \\ &= \frac{1}{2}mv^2 + \frac{1}{2}I\omega^2 \end{aligned}$$

In this expression, m is the total mass of the object, v is now the translational velocity of the center of mass, I is the moment of inertia, and ω is the angular velocity of the rotation. For a solid sphere of radius R rotating about its center of mass, $I = 2/5mR^2$. If the sphere rolls without slipping, then $\omega = v/R$. Therefore, the total kinetic energy of a sphere that rolls without slipping is

$$\begin{aligned} K &= \frac{1}{2}mv^2 + \frac{1}{2}\left(\frac{2}{5}mR^2\right)\left(\frac{v}{R}\right)^2 \\ &= \frac{1}{2}mv^2 + \frac{1}{5}mv^2 \\ &= \frac{7}{10}mv^2 \end{aligned}$$

In contrast to K , the gravitational potential energy U of an extended object is still quite simple. Near the Earth's surface, if a body's center of mass is some height h above an arbitrary level chosen such that $U = 0$, its gravitational potential energy is $U = mgh$. This means that when an object drops by a vertical height $\Delta h = h_2 - h_1$, the change in potential energy is $\Delta U = mg\Delta h < 0$ (since $h_2 < h_1$).

¹This is the content of Einstein's famous formula $E = mc^2$.

Estimating Losses Due to Friction

In an ideal world, every mechanical process would only involve exchanges of kinetic and potential energy. Conservation of energy ($\Delta E = \Delta K + \Delta U = 0$) in such a system would simply mean that losses in potential energy are matched by gains in kinetic energy:

$$\Delta K = -\Delta U$$

Unfortunately, all mechanical processes are affected by friction, which dissipates mechanical energy as heat. A proper expression of energy conservation should include the work done by friction W_f , so that we obtain $\Delta E = \Delta K + \Delta U - W_f = 0$, or

$$\Delta K = W_f - \Delta U$$

Estimation of W_f from the frictional force alone is not possible, so in this experiment, you will use a simple technique to estimate the energy lost to frictional heating. Your task will be to drop spheres through an inclined launch tube; refer to Fig. 2.1 as you read this procedure.

To find the energy lost to friction, one determines by trial and error the vertical distance $\Delta h'$ that the ball traverses through the track such that it just comes to rest at the far end of the track when released from the top. You should recognize that the loss in potential energy $mg\Delta h'$ must equal the energy lost to friction (since $\Delta K = 0$). If you assume that the same amount of energy is lost to friction when the track is tilted more steeply, you can use $W_f = mg\Delta h'$ for all tilting angles. It follows that

$$\begin{aligned}\Delta K &= W_f - \Delta U \\ &= mg\Delta h' - mg\Delta h \\ &= mg(\Delta h' - \Delta h)\end{aligned}$$

NOTE: The force of friction depends on the material composing the ball, so you have to measure the frictional loss separately for each sphere you use!

Parabolic Trajectory

The motion of a mass launched into free fall with initial velocity v_0 at an angle θ relative to the horizontal can be treated easiest by evaluating the horizontal (x) and vertical (y) position in terms of the time (t) as two independent motions:

$$\begin{aligned}x(t) &= x_0 + v_{0,x}t \\ y(t) &= y_0 + v_{0,y}t - \frac{1}{2}gt^2\end{aligned}$$

The initial velocity v_0 is estimated using energy conservation. You can break the velocity vector down into its x and y components using the geometry of the setup. Note that the ball only has one force, gravity, which acts on it during free fall; we neglect the effect of air resistance in this derivation.

2.3 Experiment

Fig. 2.1 shows the apparatus used in this experiment. A ball is released into a tube at the release point, rolls through the tube and emerges into space at the launch point. You measure the parameters shown explicitly in the figure, and the additional parameter, $\Delta h'$, which is used to estimate friction losses.

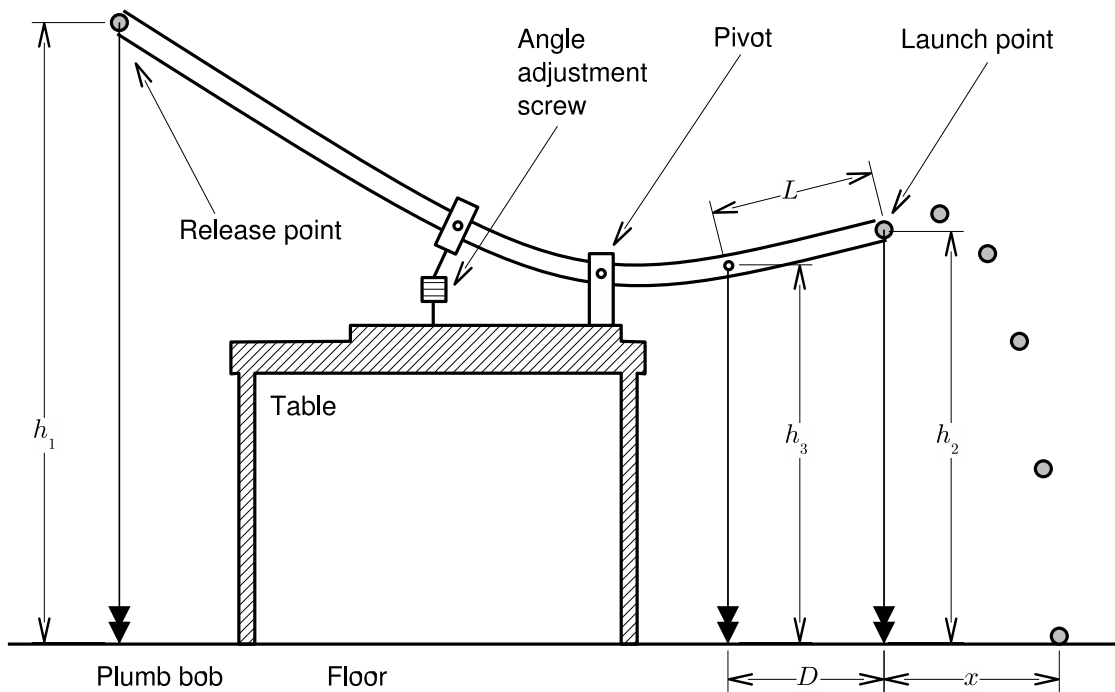


Figure 2.1: Launch tube used in this experiment, with relevant lengths labeled.

Prediction of Position

Before you do the experiment, you should derive at home a set of equations which provide a prediction for the x -position where the ball hits the ground ($y = 0$). This expression should depend only on measured parameters shown in the figure: h_1 , h_2 , h_3 , D , L , as well as the value of $\Delta h'$; your final equation should not depend on parameters that you do not directly measure, such as v_0 and θ .

Rather than derive a single complicated formula for x in terms of symbols for all the preliminary measurements, it will be more convenient to calculate, in sequence, several intermediate quantities and then combine these to find x . In other words, it is advisable to break up the entire problem into several smaller ones:

1. Find v_0 using the conservation of total mechanical energy and the energy lost to friction.

2. Find $v_{0,x}$ and $v_{0,y}$, the horizontal and vertical components of v_0 , by referring to the geometry of the final section of the track.
3. Find t , the time the ball is in the air, by considering the vertical motion involving $v_{0,y}$ and h_2 alone.
4. Finally, use t to find the expected range x .

If you want to check that your derived formula for the position x is correct, try plugging in the following data:

$$\begin{array}{ll}
 h_1 = 124.3 \text{ cm} & h_3 = 110.0 \text{ cm} \\
 h'_1 = 122.7 \text{ cm} & D = 27.7 \text{ cm} \\
 h_2 = 119.3 \text{ cm} & L = 29.2 \text{ cm} \\
 h'_2 = 120.5 \text{ cm} &
 \end{array}$$

The projectile should hit the ground at $x = 30.5$ cm.

Comparing Predictions with Observation

Using your derivation, for any configuration of the tube, you should be able to predict the x position where a ball hits the floor. The goal of the lab is to compare the measured position with the prediction by performing repeated measurements. This data will permit a measure of the spread, or uncertainty, and the reproducibility of the results.

2.4 Procedure

If you want to clean the tube before you start, there should be a swab on a string available. The procedure you should follow is described in detail below.

Estimating Friction Predicting Landing Position

- Choose your first ball as the heavy metal ball.
- Adjust the screw such that the ball, when released at the release point, just makes it to the launch point before reversing direction. Record h'_1 and h'_2 . (You only have to do this once for each ball.)
- Increase h_1 with the adjustment screw so that the ball will become airborne. Make sure that $\Delta h = h_2 - h_1$ is at least twice as big as $\Delta h' = h'_2 - h'_1$.
- Measure all the required quantities and predict where the ball will hit the floor. Use the plumb bobs to observe the alignment of the launch tube. Place a coin to test the predicted position.

Testing Prediction

At this stage, you can start to quantitatively compare the predicted landing position with actual trials.

- Do not change the launch tube settings that you just used.
- Place a sheet of white paper on the floor centered at the predicted location and place a piece of carbon paper on it. Tape them to the floor.
- On the sheet, draw a crosshairs (an x and z axis) where you expect the ball to land.
- Roll the metal ball at least 20 times; you should obtain 20 points marked on the white paper.
- Clearly label this sheet with the ball you used, the geometry of the apparatus (h_1, h_2, h_3, D, L), and the number of trials.

When you are finished with this first trial, alter the geometry of the setup and start the procedure again from the beginning. Do at least one other set of trials with the metal ball, remembering to keep $\Delta h \geq 2\Delta h'$. Then repeat the procedure using the plastic ball, with at least two separate settings for the setup. Don't forget that $\Delta h'$ is different for the metal and plastic spheres, so you need to remeasure this quantity when you start using the plastic sphere!

Before you begin your repeated trials, qualitatively test your prediction of the landing position.

After completing at least two sets of trials for at least two spheres, you may leave the laboratory and analyze the data at home.

Summary of data:

- Friction losses:
 - h'_1 and h'_2 for metal and plastic spheres
- Prediction:
 - h_1, h_2, h_3, D and L for two sets of trials, for each sphere
 - Quantitative measurement of spread in landing positions for each sphere.

2.5 Analysis: Measurement of Spread in Landing Positions

The analysis of the projectile data requires two steps. First you must estimate the *expected* horizontal landing position x of each sphere for each setup you used with the launch tube. The resulting number needs to include an uncertainty, so you must propagate the uncertainties in your measurements of h_1, h_2, h_3, L , and D . Call this prediction and its uncertainty $x_p \pm \sigma_p$.

Second, you need to calculate the mean positions *observed* along the direction of motion (\bar{x}) and perpendicular to the direction of motion (\bar{z}) observed during each set of trials for each sphere. These numbers should include experimental uncertainties, which you can easily calculate using your favorite spreadsheet program.

- Calculate the means using the usual formulae:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad \bar{z} = \frac{1}{N} \sum_{i=1}^N z_i$$

- Calculate the standard deviations of the data:

$$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N - 1}}, \quad \sigma_z = \sqrt{\frac{\sum (z_i - \bar{z})^2}{N - 1}}$$

- Calculate the standard deviations *of the means*:

$$\sigma_{\bar{x}} = \sigma_x / \sqrt{N}, \quad \sigma_{\bar{z}} = \sigma_z / \sqrt{N}$$

Report these results as $\bar{x} \pm \sigma_{\bar{x}}$ and $\bar{z} \pm \sigma_{\bar{z}}$.

To better visualize your data, calculate the differences between your actual data points (the spots from the carbon sheet) and the expected position (the crosshairs) for each trial. Plot these differences in histograms, separately for each ball. Create separate histograms for the x and z directions of your measurements.

As you write your report and comment on the results, consider the following questions:

- Do the predicted and observed mean positions x_p and \bar{x} agree within uncertainties?
- Do your data appear to be randomly distributed about the expected value? That is, do the histograms “look like” Gaussian curves?
- Do you observe any evidence for systematic errors in the data, i.e., an overall shift in the points from the expected landing site? If so, what could have caused these errors?
- How does the spread in along the x direction compare to the spread along z for each set of trials? Explain these results.
- Is the spread in the data about the same for the metal and plastic spheres? Should it be?
- What alterations can you make to the experiment to decrease the spread in the data?
- How do your results for the plastic sphere change if you assume it is hollow? (Hint: the moment of inertia of a hollow sphere is not $I = 2/5mR^2$.)

- How accurate is your measurement of friction? What factors limit its accuracy?
- Estimate the size of the frictional force in the tube, assuming (incorrectly) that it is constant over the length of the tube, i.e. that $W = f_f l$. Explain how you estimated the length of the tube l and how accurate you think your estimate is. Based on this, do you expect friction to be significant? (Hint: compare to the magnitude of other forces acting.)
- How do you expect your estimate to change if you consider friction to vary over the tube? What determines the magnitude of the frictional force? Where do you expect it to be maximal? Why?
- What are some implicit assumptions you are making when you use the theoretically derived equation?

Experiment 3

Magnetic Fields

3.1 Purpose

You will determine the strength of the magnetic field in the gap of an electromagnet by measuring the force on a current-carrying rod, and by observing the induced EMF in a closed loop.

- **CAUTION:** Always reduce the current through the electromagnet to zero before opening the circuit of the magnet coils.
- **CAUTION:** Remove wrist watches before placing hands near the magnet gaps.
- **CAUTION:** Do NOT place your laptop near the magnets!

3.2 Introduction

There are several techniques we can use to measure the strength of a magnetic field \vec{B} in some region. One technique exploits the fact that a current-carrying conductor will feel a deflection force in the presence of a magnetic field; another uses Faraday's Law and observes the EMF induced by a changing magnetic field in a closed loop. In this experiment you will employ both of them.

Force on a Current-Carrying Conductor

Consider a wire or conducting rod of length \vec{L} carrying a current i . The direction of the vector \vec{L} indicates the flow of the current. If this wire is placed into a magnetic field \vec{B} , it will experience a deflecting force

$$\vec{F} = i\vec{L} \times \vec{B} \quad (3.1)$$

If the direction of the current i and the field \vec{B} are perpendicular, as in Fig. 3.1, then the conductor will experience a vertical force $F = iLB$.

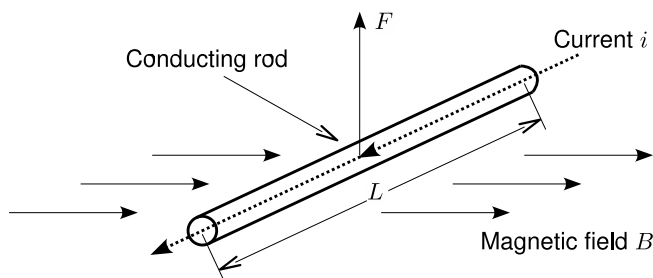


Figure 3.1: Force on a current-carrying wire.

If one knows the current i in the rod and the magnitude of the deflection force F , it is possible to infer the strength of the magnetic field B by observing this deflection.

EMF Induced in a Moving Coil

Suppose you place a closed conducting loop of area A in a region with some magnetic field \vec{B} . The flux of field lines through the loop is given by

$$\Phi = \int \vec{B} \cdot \hat{n} dA$$

If the plane of the loop and \vec{B} are perpendicular, the flux reduces to the simple expression $\Phi = BA$. Faraday's Law tells us that if the flux changes as a function of time — e.g., \vec{B} is a function of time, or the loop area moves with respect to the field — then an EMF

$$\varepsilon = -\frac{d\Phi}{dt} = -\frac{d}{dt}(BA)$$

will be induced across the loop. It is useful to think of the EMF as a voltage, rather than a battery, that drives a current i through the loop. This current can be measured, and used to infer ε . Note that if we replace the single loop by a coil consisting of N tightly wound loops of equal area, then the EMF for the coil becomes

$$\varepsilon = -N \frac{d\Phi}{dt} \tag{3.2}$$

A measurement of ε can, under the right circumstances, yield the strength of the magnetic field.

3.3 Experiment

There are two parts to this experiment; first, you will use balanced forces to find the magnetic field of an electromagnet, and second, you will use an induced emf to measure the magnetic field produced by a magnet.

Measurement of B using the Current Balance

In the first part of the experiment, you will use the magnetic force on a horizontal rod carrying current i to estimate the strength of the electromagnet. The basic experimental setup is shown in Fig. 3.2. The magnetic field you will be measuring is provided by an electromagnet, which uses the fact that a current produces a magnetic field. The strength of the magnetic field is proportional to the amount of current flowing, with a constant of proportionality determined by its geometry. In this experiment, a large, horizontal magnetic field B is produced in the air gap of a C-shaped electromagnet, which has many coils of wire wrapped around an iron core. The current, I , is supplied by an adjustable low voltage power supply.

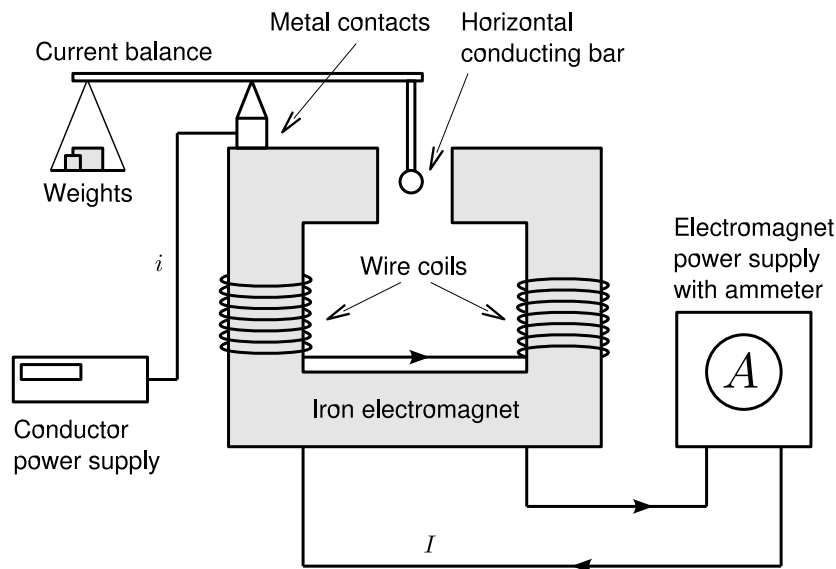


Figure 3.2: Setup of the wire, magnet, and balance for force measurement.

Given a B field in the electromagnet, one can increase or decrease the force on the rod by adjusting the current i in the rod. The size of the force is estimated using a balance: a set of weights can be used to bring the balance to equilibrium. Once in equilibrium, the force of gravity on the weights, $F = mg$, must equal the magnetic force on the conducting rod.

Measurement of B using Induction

An alternative method of measuring B is to use Faraday's law. Figure 3.3 shows the initial setup for the experiment. A rotary motion sensor is attached to the end of the cross-rod. It also acts as a pivot for the induction wand. The induction wand is a rigid pendulum with coil at its end that will swing

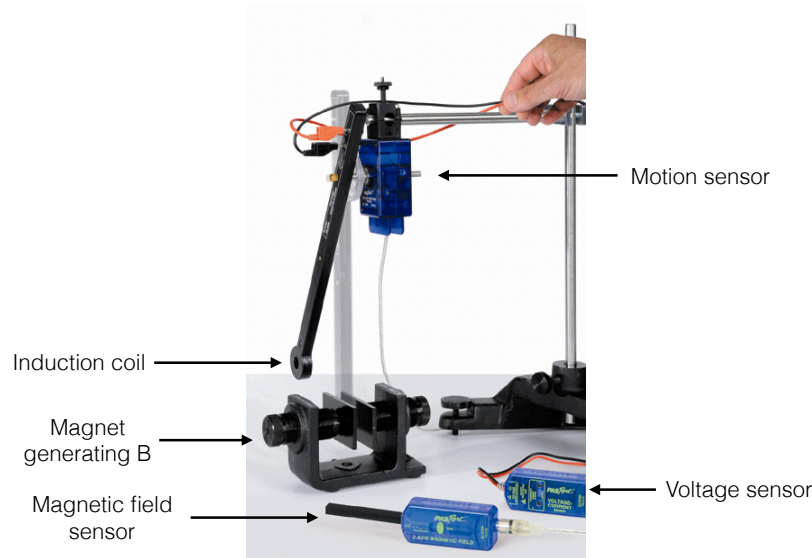


Figure 3.3: Experimental apparatus.

through a variable gap magnet. When the coil is suddenly inserted into (or removed from) the field, the flux through the coil rapidly changes, inducing an EMF ε . The EMF will drive a current i through the coil while $d\Phi/dt$ is nonzero. This can be easily measured. If we are interested in the net average result over an interval of time Δt , then Faraday's formula tells us that

$$\varepsilon = -NA \frac{\Delta B}{\Delta t} \quad (3.3)$$

where N is the number of turns in the coil, A its area and ΔB the net change in magnetic field.

In this particular setup, the coil has $N = 200$ turns and its outer diameter is $d = 3.1$ cm. When the coil passes the gap in between the magnet, it goes from a region of zero magnetic field (outside the gap) to a region of non-zero field (center of the gap). An EMF is therefore induced on the coil and it also follows that the net change in magnetic field ΔB (equation (3.3)) is exactly the field generated by the magnet.

3.4 Procedure

Current Balance

You will begin the experiment by measuring the force exerted on the conducting rod by the electromagnet. First, connect the electromagnet to the large power supply, using the banana plugs provided. (Be

sure to follow any instructions that are posted on the power supply.) This will provide the current I generating the magnetic field. The current in the rod, i , is supplied by an HP E3610A power supply, which you will use in constant current mode.

- Measure L , the length of the conducting rod, using the calipers provided.
- Turn the adjustment knob on the left of the balance to lower the balance arm onto its knife edge. Your balance may need adjusting, by turning the screw on the threaded rod, so that it is horizontal. If there is not already a mark indicating the horizontal position, you may draw one on the mirror.
- To use the power supply of the current rod in constant current mode, begin with the current dial turned all the way down (counterclockwise) and the voltage dial turned all the way up (clockwise).
- Set the range to 3 A (button out).
- Connect leads wires from the balance arm to the $+$ and $-$ terminals.

You can now set the current to the desired level, but note that the digital meters on the power supply show the actual voltage and current being supplied, so you will not normally see any current unless you have a complete circuit, i.e. the leads are connected and the balance is resting on its knife edge. If you want to set the current level without closing the circuit, you can hold in the **CC Set** button while you turn the current dial. Once you close the circuit, the voltage adjusts automatically to maintain the constant current level, and the **CC** (Constant Current) indicator light should be on.

At this point, you are ready to begin the actual measurements:

- Set the current I through the electromagnet at 4 A.
- When the current is set, the rod in between the electromagnet's poles should be pulled downward. If the rod moves upward instead, reduce the current on the large power supply to zero, flip the direction of the current through the rod, then turn the current back to 4 A.
- For at least five values of the balance current i , determine the force $F = mg$ needed to bring the balance to equilibrium. **NOTE:** it is easier to make the final adjustment on i once weights have already been selected for the approximate current value. Change i slowly, to avoid oscillations of the balance.
- Use the mirror behind the balance to line up the pointer and its image as you balance the rod.
- Repeat this measurement for other values of I , e.g., 3.5 A, 3 A, 2.5 A, and 2 A.
- Be sure to record uncertainties for every measurement.

Induced EMF

CAUTION: Do not place the magnet very close to the computer.

After you set up the apparatus, this part of the experiment will consist of two steps: measurement of the magnetic field ΔB generated by the magnet and measurement of the EMF ε induced on the swinging coil. This will be done with the help of DataStudio.

Most likely, the apparatus will be ready for data taking and you will not have to do anything. If that is not the case, set it up following these steps:

- Adjust the height of the stand so that the coil is in the middle of the magnet. Adjust the gap between the magnet poles to about 1 inch so the coil will be close enough to the magnet poles and will be able to pass through the magnet without hitting the pole plates.
- Plug the voltage sensor banana plugs into the banana jacks of the induction wand. Connect the voltage sensor to the USB link that is plugged into the computer. Make sure that the voltage sensor cables will not exert any torque on the induction wand as it swings. If it does, it helps to hold the wires up while recording data.
- Attach the rotary motion sensor and the magnetic field sensor to the computer. Notice that the magnetic field sensor's probe has 2 white dots: one for sensing perpendicular magnetic field lines and the other for axial field lines. You will use the perpendicular sensing element of the sensor. Note the direction of the arrow. **NOTE:** The sensor is set to measure the field in units of Gauss (G). You will have to convert it into Tesla (T). Remember that $1 \text{ G} = 10^{-4} \text{ T}$.
- Open the "Induced EMF" DataStudio file on the computer. Make sure that the sampling rate setting (which can be found in the "Setup" menu) shows the following:

Magnetic field: 1 Hz; Angle: 10 Hz; Voltage: 500 Hz or 1000 Hz.

Once this is done you can proceed and measure the magnetic field ΔB as follows:

- Pull the magnet out of the setup so that you can place the magnetic field sensor in between the pole plates. Hold the magnetic field sensor such that the perpendicular sensing element is along the magnet's field lines in between the magnet and is at the center of the pole plates. See figure 3.4.
- Click "Start" to measure the maximum magnetic field strength between the poles. Hold the sensor in the center for around 6 seconds. Click "Stop". You will see a straight line in the magnetic field panel. Record the mean value of the measured magnetic field. From the sign of the recorded value determine the direction of the magnetic field between the pole plates. Note which pole of the magnet is the North pole.



Figure 3.4: Measuring magnetic field in between the pole plates.

- Move the probe 1 cm away from the center of the pole plates and repeat the measurement of the magnetic field at this new spot. Now move it 2 cm away from the center and repeat the measurement. Compare all three measurements. Note that the value of ΔB that you will have to use during data analysis is the one obtained from the center of the magnet since that is where the coil goes through.

Now you know what value of ΔB to use in equations (3.3) you can go to the next step and measure ε and Δt . In particular, you will have to do the following:

- Put the magnet back in the setup as shown in figure 3.3. Make sure that the coil can pass through the gap without hitting the pole plates.
- Pull the wand back and click “Start”. Release the wand allowing it to swing through the magnet *once* and then click “Stop”.
- Click and drag the mouse all over the first peak to highlight the first peak and find the average voltage by getting the Mean of the highlighted part of the graph. This will be your average ε . Note that before the wand passes through the pole plates there may be very small voltage readings. Make sure to exclude those data points as you select the peak, and try to select the same amount of data points on both sides of the peak.
- Use the Smart Tool to determine the difference in time (Δt) from the beginning to the end (width) of the first peak.

3.5 Analysis

Current Balance

- Using the current balance data, plot iL , the product of the rod current and rod length, against the force values $F = mg$.
- Using a linear least squares analysis, determine the slope of the plot and its corresponding uncertainty. From there find $B \pm \sigma_B$. This is the magnitude of the electromagnet's field.
- Repeat these steps for all the values of I that you set in the electromagnet.
- Plot all of the data on the same set of axes, and extract the slope B for each setting of I .
- In another graph, plot B against I , and draw a line or curve of best fit, commenting on the shape. Is it what you expect?
- You have considered only the horizontal part of the current-carrying rod in your analysis. What effect does the vertical part (which is also within the magnetic field) have?

Induced EMF

- Given the experimental values of ΔB and Δt , compute the EMF ε . Compare the computed value with the one measured experimentally. Are they compatible within errors?
- Identify on the graph where the coil is entering the magnet and where the coil is leaving the magnet.
- Why is the sign of the EMF of the second peak opposite to the sign of the first peak?
- Why is the EMF zero when the coil is passing through the exact center of the magnet?

Experiment 4

e/m of the Electron

4.1 Purpose

In this experiment, you will measure the charge to mass ratio of the electron, a quantity known as e/m . The procedure demonstrates how we can use fairly basic equipment to observe one of the fundamental constituents of matter.

4.2 Introduction

The “discovery” of the electron by J.J. Thomson in 1897 refers to the experiment in which it was shown that “cathode rays” behave as beams of particles, all of which have the same ratio of charge to mass, e/m . At the time, Thomson correctly recognized that he had isolated a fundamental particle of nature; for his work, he won the Nobel Prize in 1906.

Thomson’s experiment, and the one we will conduct today, is a simple application of the motion of a charged particle in a magnetic field. Recall that if a particle of charge e moves with velocity \vec{v} into a region with some magnetic field \vec{B} , it will feel a magnetic deflection force

$$\vec{F}_{mag} = e\vec{v} \times \vec{B} \quad (4.1)$$

Note the similarities between this expression and the magnetic force on a current-carrying wire: $\vec{F}_{wire} = i\vec{L} \times \vec{B}$. This equation is in fact just a generalization of the force on a point charge, since a current is nothing more than a flow of charge.

Consider the special case where \vec{v} and \vec{B} are perpendicular, and the magnitude of \vec{B} is uniform. As shown in Fig. 4.1, the force \vec{F}_{mag} is directed in such a way that the particle moves in a circle of radius r , with the plane of the circle perpendicular to \vec{B} .

Since the particle moves in uniform circular motion, the magnetic force is a centripetal force,

$$\vec{F}_c = m \frac{v^2}{r} \quad (4.2)$$

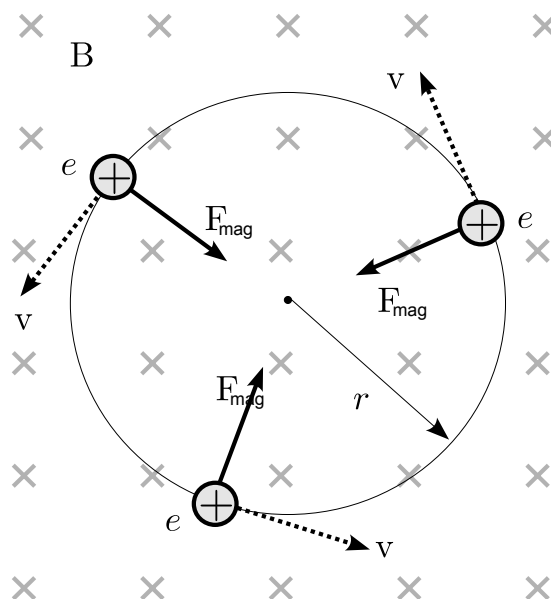


Figure 4.1: Deflection of a charged particle in a uniform field \vec{B} directed into the page.

We can consider equations (4.1) and (4.2) for the special case where \vec{v} and \vec{B} are perpendicular, and the magnitude of \vec{B} is uniform. They suggest that, in principle, if we could measure the incoming velocity v of the particle, we could use this result to directly measure the charge to mass ratio e/m .

In practice, direct measurements of v are not feasible. However, if we use some known potential difference V to accelerate the particle from rest to a speed v , we could rewrite the speed in terms of V using energy conservation:

$$U_{kin} = \frac{1}{2}mv^2 \quad \text{and} \quad U_{pot} = eV \quad (4.3)$$

By substituting v in equations (4.2) and (4.3), e/m can be expressed directly in terms of V , B , and the easily observed radius of curvature r :

$$\frac{e}{m} = \frac{2V}{B^2 r^2} \quad (4.4)$$

4.3 Experiment

In the setup you will use, electrons are emitted at a very low velocity from a heated filament, accelerated through an electrical potential V to a final velocity v , and finally bent in a circular path of radius r in a magnetic field B . The entire process takes place in a sealed glass tube in which the path of the electrons can be directly observed. During its manufacture, the tube was evacuated and backfilled with a small

trace of helium gas. When electrons in the beam have sufficiently high kinetic energies (≥ 10.4 eV), a small fraction of them will ionize helium atoms. Recombination of the helium ions, accompanied by the emission of a characteristic blue light, occurs very near the point where the ionization took place. As a result, the path of the electron beam is visible to the naked eye as a thin blue beam of light.

Electron Gun in Vacuum Tube

Figure 4.2 shows the indirectly heated cathode and the anode plate used to accelerate the electrons. The cathode is heated by passing a current directly through the heater. A variable positive potential difference of up to 500 V is then applied between the anode and the cathode in order to accelerate the electrons emitted from the cathode. Some of the accelerated electrons come out as a narrow beam through a small aperture on the grid.

The tube is set up so that the beam of electrons travels perpendicular to a uniform magnetic field B , and is initially vertical. The B field is produced by the current I running through a pair of large diameter coils (so-called “Helmholtz coils”) designed to produce optimum field uniformity near the center.

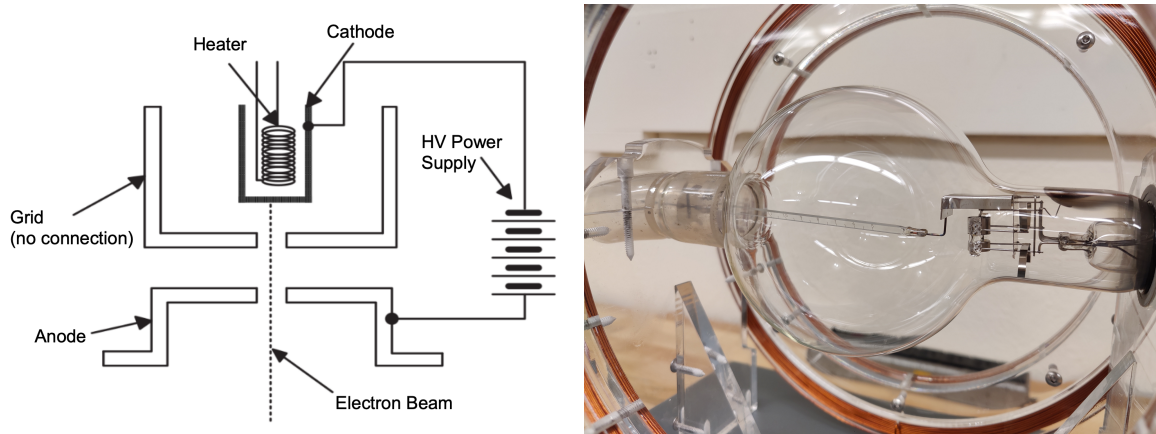


Figure 4.2: Diagram of the electron gun in the sealed glass tube.

The Helmholtz Coils and the Uniform Magnetic Field

A current I flowing in a single loop of wire of radius R produces a magnetic field on the symmetry axis given by:

$$B_{loop} = \frac{\mu_0 R^2 I}{2(R^2 + x^2)^{3/2}}$$

x is the distance from the plane of the loop. The electromagnet used in this experiment, shown in Fig. 4.3, consists of two loops of wire with N turns each, separated by a distance R (the same R as

the coil radius). The coils contribute equally to the field at the center ($x = R/2$), so at that point

$$B_I = \frac{\mu_0 R^2 N I}{(R^2 + (R/2)^2)^{3/2}} = \left(\frac{4\pi \times 10^{-7} N}{R(1 + 1/4)^{3/2}} \right) I = C \cdot I \quad (4.5)$$

The setup in the laboratory has $N = 132$ turns per coil, where each coil has a radius $R = 0.1475$ m; before you come to the lab, use these numbers to calculate the constant C (units of T A^{-1}). This arrangement, called a pair of **Helmholtz coils**, yields a highly uniform field in the region at the center.

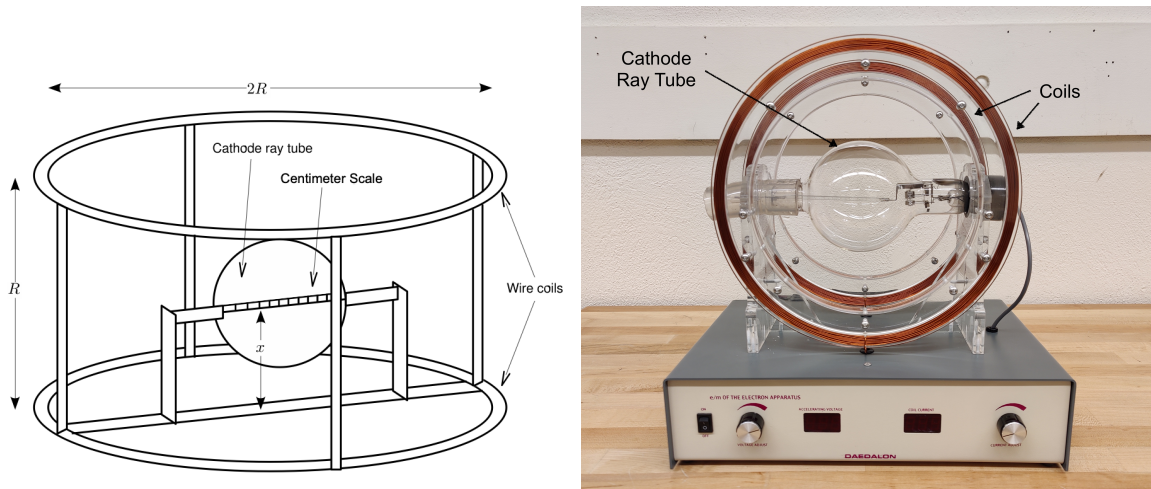


Figure 4.3: Helmholtz coils used to produce a uniform magnetic field.

Estimating the Charge to Mass Ratio

The electrons are therefore emitted into a region where a uniform magnetic field acts perpendicular to the motion of the electrons. The magnitude of the magnetic field can be adjusted until the resultant circular path of the electron beam just reaches the far end of the centimeter scale. The scale extends from the electron gun in a direction perpendicular to that in which the electron beam was emitted — i.e., along a diameter of the circular orbits. The scale numbers on the scale fluoresce when struck by the electron beam. Then for given values of V , B and r it would be possible to determine e/m from eq. (4.4).

However, the net field \vec{B} in which the electrons move is not only due to the Helmholtz coils, but also to the magnetic field in the ambient environment \vec{B}_E . A part of the ambient magnetic field is due to the Earth's magnetic field, but there may also be contributions from nearby ferromagnetic materials in the lab. Hence, the total field inside is the vector sum of \vec{B}_I and \vec{B}_E . We can minimize the effect of \vec{B}_E by aligning the \vec{B}_I field with the direction of the needle of a compass close to the setup. The

horizontal component of \vec{B}_E is now aligned with \vec{B}_I . In fact, the value of \vec{B}_E is so small that we can safely ignore it for the purpose of this lab. Incorporating eq. (4.5) into eq. (4.4) and rearranging terms, we obtain

$$I \simeq \left(\frac{1}{C} \sqrt{\frac{2V}{e/m}} \right) \frac{1}{r} \quad (4.6)$$

Now, instead of performing a single measurement of r , we can measure the variation of r with B (or I) at fixed values of V , and use the fact that the coil current I is a linear function of the **curvature** $1/r$.

Note that eqs. (4.2), (4.3), and (4.6) apply only to electrons with trajectories on the outside edge of the beam — i.e., the most energetic electrons. This is because some electrons in the beam will lose energy through collisions with helium atoms.

4.4 Procedure

Derivations

Before you begin taking data, derive equations (4.4) and (4.6) using the concepts and basic assumptions described in the *Introduction*.

Check your work with your TA. When you write your lab report, you should include relevant equations from your derivation in your “Introduction” and “Methods” sections. Recall that your lab report should be **clear** and **concise**, so you should **not** include the full derivation in your report.

Orientation of the Coil and Tube Setup

For reasons already explained, we would like to orient the Helmholtz coils such that their axes are parallel to the horizontal direction of the ambient magnetic field. To do this,

- Make sure that the setup is turned off before doing anything.
- Put a compass on the desk at a place close to the setup. Wait for the needle to stabilize.
- Rotate the setup such that the axes of the coils are parallel to the needle of the compass.
- The coil axis should now be aligned with the horizontal component of the ambient magnetic field. An example is shown in Fig. 4.4

Please exercise caution as you align the Helmholtz coils: the cathode ray tube is very delicate and may break if a large force is applied to it. It is recommended that you never touch the tube or the coils, and only maneuver the setup by touching the base.



Figure 4.4: Example of the Helmholtz coils aligned to the ambient field.

Measurement of the Circular Orbits

Once you are done with the alignment, you can turn on the setup and start taking data. Make sure only the dim incandescent ceiling lights in the room are on.

- Turn on the power switch. The unit will perform a 30-second self-test, indicated by the digital display changing values rapidly. During the self-test, the controls are locked out, allowing the cathode to heat to the proper operating temperature. When the self-test is complete, the display will stabilize and show “000”. Although the unit is now ready for operation, a 5-10-minute warm-up time is recommended before taking careful measurements.
- Turn the Voltage Adjust control up to 200 V and observe the bottom of the electron gun. The bluish beam will be travelling straight down to the envelope of the tube.
- Turn the Current Adjust control up and observe the circular deflection of the beam. When the current is high enough, the beam will form a complete circle within the envelope. The diameter of the beam can be measured using the internal centimeter scale inside of the tube. The scale numbers fluoresce when struck by the electron beam.

- For an accelerating voltage of 200 V, measure the beam diameter for a series of coil current settings. Alternatively, determine the coil current necessary to bring the beam to each prescribed distance on the tube arm. If the mark falls between steps on the ammeter, interpolate the current.
- Repeat the measurements for additional accelerating voltages: 100 V, 300 V, 400 V, and 500 V. Note that both the voltage and current outputs are controlled by an on-board microprocessor, which locks out the controls at both the minimum and maximum settings. The range for the voltage V is 100 – 500 V, while the range for the current I is 1 – 3.0 A.
- When all of the data has been collected, switch off the apparatus.

Summary of data:

- I and r for four or five values of V
- Derivations for equations (4.4) and (4.6)

4.5 Analysis

When you finish the lab, you should have a data table for each accelerating voltage V you used. The tables will contain values of I as a function of r . To determine the charge to mass ratio e/m , you will plot I versus $1/r$ and perform the usual linear least squares analysis on the data.

- For each voltage V , plot I against $1/r$. Plot all of the curves in a single chart, and remember to convert r to meters.
- As already discussed, we expect that I should be a linear function of $1/r$

$$I = A \cdot \frac{1}{r} + D$$

Perform a weighted linear least squares fit to the data to find the slope A , the intercept D , and the standard errors σ_A and σ_D for each of the five curves.

- Comparing the equation above to eq. (4.6) you will find that e/m can be expressed in terms of A . Use this fact to get five estimates of e/m from the five curves. Remember to propagate errors in A , and V to get the uncertainty $\sigma_{e/m}$ in each of the estimates.
- Average your results to find $\overline{e/m}$. When finding the average, weight the data points by the errors $\sigma_{e/m}$ — i.e., find a **weighted average** and a **weighted standard error**. Report these weighted averages as your final result.

When you have completed the analysis, comment on the results. Consider the usual set of questions:

- Do your results for e/m disagree significantly (i.e., accounting for statistical uncertainties) with the accepted value

$$e/m = 1.758 \times 10^{11} \text{ C kg}^{-1}$$

- What possible systematic errors could be affecting your results?
- Which measurements of I are more precise, those for smaller r or larger r ? Why?
- Does the value for D agree with the expected value, within uncertainties? Can you explain any significant discrepancies, if they exist?

Note that accuracy can be difficult to achieve in this laboratory. Do not become frustrated if your results exhibit significant discrepancies with respect to the accepted value of e/m . Discuss possible sources of the discrepancies and suggest several means to improve the experiment.

Experiment 5

Polarization and Interference

5.1 Purpose

The purpose of this experiment is to investigate two basic concepts of wave mechanics: polarization and interference. You will conduct the experiment with light, but the general concepts apply to many other types of waves.

5.2 Introduction

Polarization

Within the enormous range of wave phenomena observed in nature, we can identify two basic types of waves:

1. **Longitudinal:** waves whose vibrations are parallel to the direction of wave propagation. When a longitudinal wave moves through a medium, one observes density variations — regions of compression and rarefaction — in the material. Sound is an example of a longitudinal wave.
2. **Transverse:** waves whose vibrations (or displacements) are perpendicular to the direction of propagation. Light is an example of a transverse wave.

Polarization occurs only in transverse waves. It refers to the axis of the waves' oscillation, which is always perpendicular to the wave direction. For a wave traveling in three spatial dimensions, the direction of polarization can occur anywhere within the plane perpendicular to the direction of propagation. If a wave oscillates along only one axis within this plane, it is said to be *polarized*.

Light is electromagnetic radiation; it consists of oscillating electric and magnetic fields that vibrate perpendicular to the direction of the wave. When we speak of the polarization state of light, we refer specifically to the vibration axis of the electric field.

Typically, light produced by an incandescent bulb or a candle is *unpolarized*; that is, the electric field associated with the emitted light oscillates in random directions within the plane perpendicular to the direction of propagation. However, unpolarized light can become polarized under several circumstances. One of these occurs when we pass unpolarized light through a polarizing filter, as shown in Fig. 5.1. For visible light, such filters are typically made of polymer films, plastics whose molecules form long chains oriented along one axis.

When an unpolarized light wave encounters the filter, the component of the electric field oscillating parallel to the molecular chains is absorbed in the material. The component of the field perpendicular to the molecules' long axes is transmitted. Therefore, the direction perpendicular to the molecular chains is called the *transmission axis*.

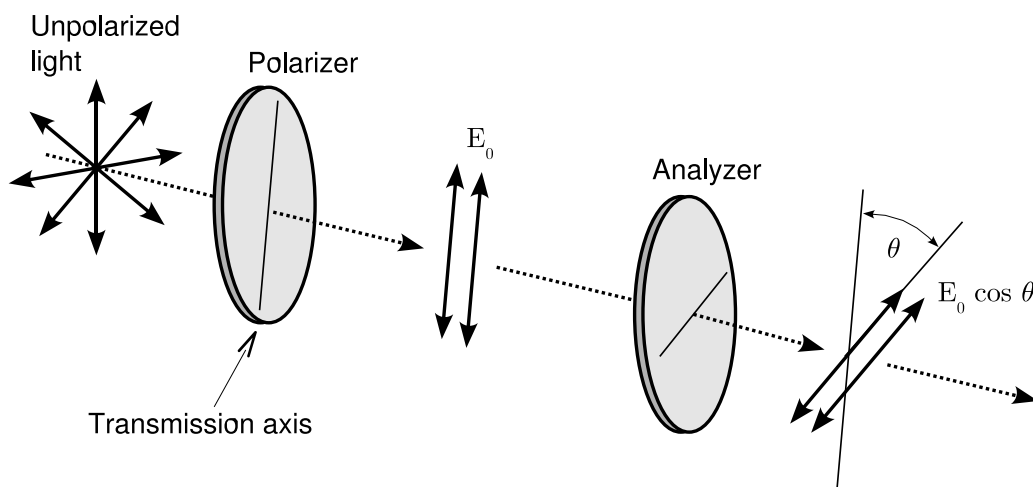


Figure 5.1: Creation and detection of polarized light.

Suppose a wave has been polarized by a filter such that its polarization is characterized by an electric field \vec{E}_0 . We can analyze the polarization state by passing the light through a second filter (called an analyzer) whose transmission axis is oriented at an angle θ with respect to the first filter. The effect of the analyzer is to pick off the component of \vec{E}_0 parallel to its transmission axis, $E_0 \cos \theta$. The remaining light is absorbed. In this experiment, we use a laser, which is a polarized light source, and shine the light through a polarizer (the analyzer), recreating the physics of two successive polarizers.

Since the intensity of the light wave is proportional to $|\vec{E}|^2$, the intensity of light after it has passed through the analyzer is

$$I = I_0 \cos^2 \theta, \quad (5.1)$$

where $I_0 = |\vec{E}_0|^2$. This shift in the intensity, known as Malus' Law, can be detected by a human eye or a photometer.

Interference: Young's Double Slit

Another key property of waves is their ability to superimpose. That is, when two waves encounter each other in a medium, the resulting wave is simply the algebraic sum of the two individual waves. The combination of two or more waves into a third is called *interference*. Interference can occur constructively or destructively:

1. **Constructive Interference:** the displacements of the two waves occur in the same direction, so the sum is an even larger wave (see Fig. 5.2, left).
2. **Destructive Interference:** the displacements of the two waves occur in the opposite direction, so the sum is smaller than either component (see Fig. 5.2, right).

Interference can be totally constructive, totally destructive, or some combination of both. The amount depends on the relative phases of the component waves — that is, the relative location of each wave in its oscillatory cycle.

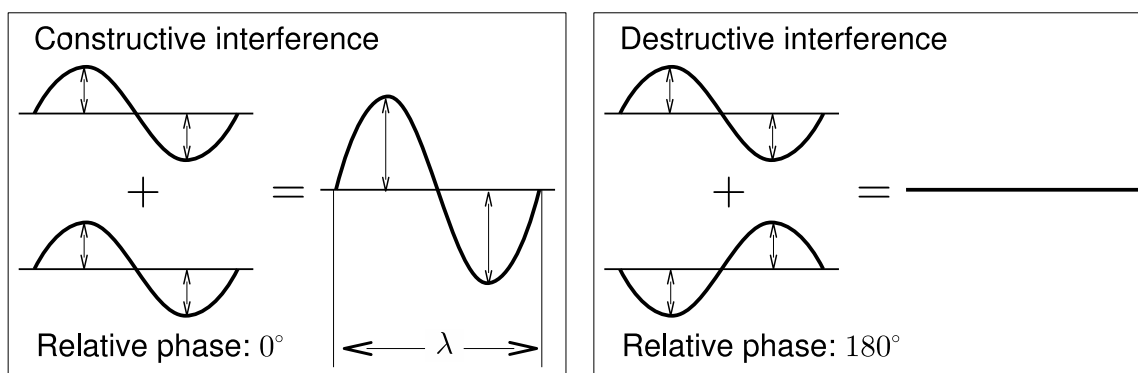


Figure 5.2: Totally constructive and totally destructive interference of two waves.

When visible light constructively interferes, the result is an increase in the intensity of the light. When it destructively interferes, the intensity decreases. The first clear demonstration that this actually occurs was carried out in 1801 by Thomas Young. After passing a collimated light beam through two narrow slits, Young observed the wavelike interference of the beam, which formed a pattern of bright and dark spots.

Young's double slit experiment, illustrated in Fig. 5.3, permits what remains of an incoming wave (from the left) to travel to a distant screen (on the right) along two different paths l_1 and l_2 . The light waves from the two slits interfere, resulting in an interference pattern of bright (constructive) and dim (destructive) patches as viewed on the screen.

Using the geometry of Fig. 5.3 and Fig. 5.4, we can quantitatively predict where bright and dark patches will appear on the screen. The two slits are separated by a distance d , and located a distance

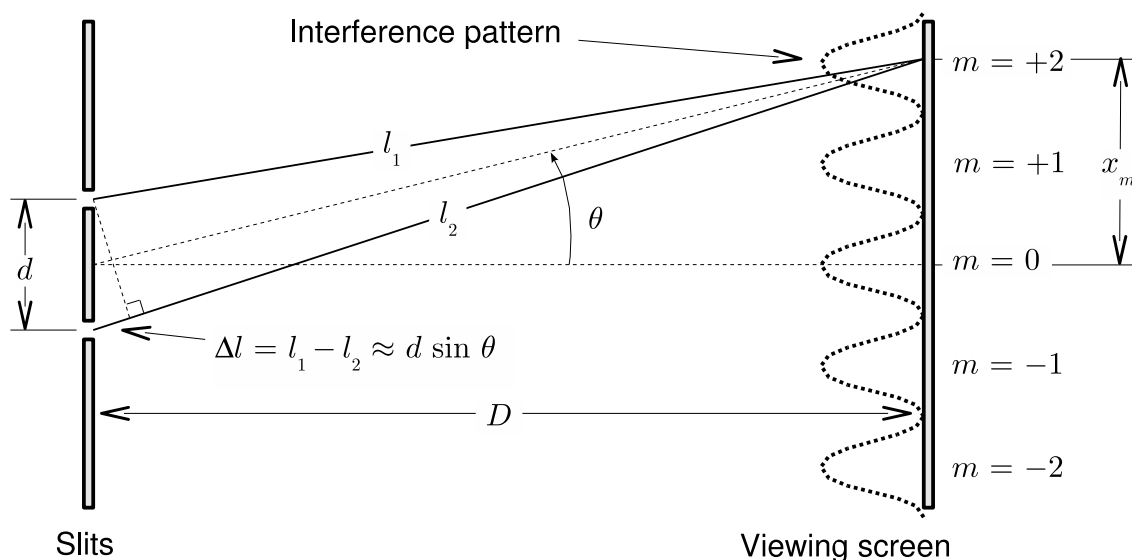


Figure 5.3: Interference of light emanating from two small slits. Note that the horizontal distance D has been shortened and distorted in this drawing.

D from the screen. We shine monochromatic (single-wavelength) light of wavelength λ from the left onto the double slit, which allows two light waves to propagate from the two slits. Straight ahead, they will always be in phase because they travel the same distance to the screen. But when the two waves propagate at an angle θ , they cover different distances to reach a specific point on the screen¹.

From Fig. 5.4, it is clear that the difference in path length Δl between the two slits is

$$\Delta l = d \sin \theta$$

It is this Δl that determines the location of intensity maxima and minima in the interference pattern on the screen. For an intensity minimum to occur, we must have destructive interference between the two waves. This happens when the length difference between the waves' paths is a half-integer multiple of the wavelength of the light:

$$\Delta l = \left(m + \frac{1}{2}\right)\lambda, \quad m = 0, \pm 1, \pm 2 \dots$$

When Δl takes on these values, the relative phase between the two waves will be 180° (consult Fig. 5.2 to convince yourself). For an intensity maximum, the relative phase must be 0° , and this occurs when Δl is an integer multiple of the wavelength:

$$\Delta l = m\lambda, \quad m = 0, \pm 1, \pm 2 \dots$$

¹We assume that the two rays are parallel ($\theta_1 \approx \theta_2$), a good approximation since the distance to the screen D is effectively infinite compared to the distance d between the two slits.

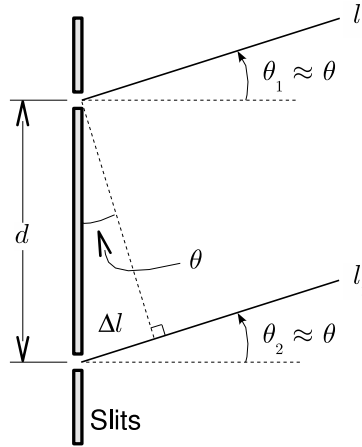


Figure 5.4: Geometry of the double slit, assuming that the emitted rays are effectively parallel.

Therefore, there exists a set of angles θ_m where the intensity maxima will occur, satisfying

$$d \sin \theta_m = m\lambda$$

For small θ_m (in radians), we can use the approximation $\sin \theta \approx \tan \theta$. If x is the distance on the screen from the central maximum, then the position of the m^{th} maximum is given by

$$\sin \theta_m \approx \tan \theta_m \approx \frac{x_m}{D}$$

Combining all of our results, we find that the position of the m^{th} maximum is a linear function of m :

$$x_m = \left(\frac{\lambda D}{d} \right) m \quad (5.2)$$

The Interference Pattern

The cartoon in Fig. 5.3 suggests that the intensity pattern produced by the double slit should look roughly like a sine wave, with distance between the peaks depending on the slit separation d and the screen distance D . While this is true, it is not the entire story.

In fact, the true intensity pattern should appear as depicted in Fig. 5.5. The double slit diffraction pattern (solid line) is periodic, but it also contains some larger features. This results from the fact that a combination of interference effects are present in the double slit experiment: first, there is interference of the two slits, as we have already discussed, and, secondly, there is a diffraction effect that occurs from each individual slit. The larger pattern in the figure is the intensity profile you would observe if you just passed the light through one slit. The intensity of the single slit pattern is given by

$$I = I_0 \left(\frac{\sin(\pi a \sin \theta / \lambda)}{\pi a \sin \theta / \lambda} \right)^2,$$

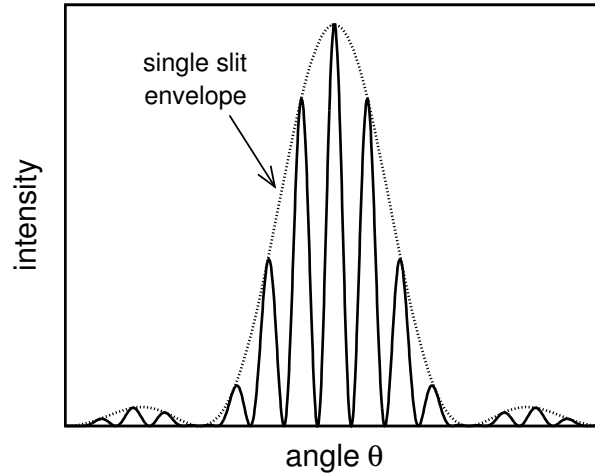


Figure 5.5: Intensity pattern of the double slit: a double slit pattern modulated by a single slit envelope.

where a is the width of the slit. From this result, you can show that the single slit minima occur when

$$\frac{\pi a \sin \theta}{\lambda} = n\pi, \quad n = \pm 1, \pm 2, \dots$$

$$\sin \theta_n = n \frac{\lambda}{a}$$

Using the small angle approximation we applied earlier, we can express this result in terms of the positions x_n of the single slit minima on the viewing screen:

$$x_n = \left(\frac{\lambda D}{a} \right) n \quad (5.3)$$

5.3 Experiment

To conduct the experiment, you will use the components shown in Fig. 5.6. The primary pieces of equipment that you will need to use are:

Light Sensor with Rotary Motion Sensor (RMS): Light sensor is a light-sensitive device that measures the total power of light incident on it, and RMS measures and reports the current transverse position. It is mounted on a linear translator, along which it can move on the direction transverse to the optical bench.

Double/Single Slit Disk: Holds various kinds of slits that will be used in this experiment. Slits can be switched by rotating the disk.

Polarizer (with RMS): Polarizes unpolarized light or act as analyzer. The angle it rotates through can be measured and reported with a RMS.

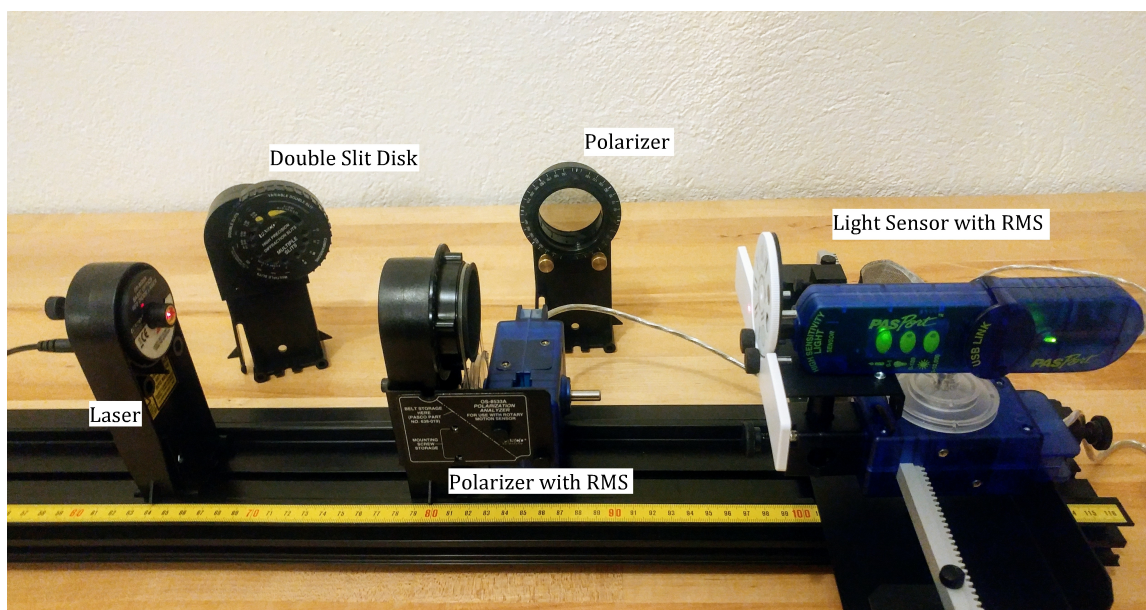


Figure 5.6: Equipment components used in the experiment.

During the first part of the experiment, you will use the polarizer with RMS to test Malus' Law, with the laser being the light source. In the second part of the experiment, you will remove the polarizer from the bench, install the slits and observe double and single-slit diffraction of laser light. The laser is extremely useful because it emits intense light of a single wavelength, and because this light is coherent (all light waves are in phase).

It is **very** important that lasers, polarizers and disks should stand upright at all times. Laying them on their sides exposes them to damage from dust and to getting scratched/damaged by another object placed on top of it. This also exposes you to damage to your grade!

Safety Note

Although the laser is of relatively low intensity, it can be dangerous in certain circumstances unless used carefully. In particular, do not use the laser in such a way that it can shine into any person's eye. (The warning label on the laser states "Do NOT stare into beam!"). When you are not using the laser, remember to turn it off.

5.4 Procedure

Polarization

- Make sure only the dim incandescent ceiling lights in the room are on.
- Place the Polarization Analyzer with the Rotary Motion Sensor (RMS) between the laser and the Light Sensor.
- Rotate the aperture disk in front of the Light Sensor so that the open aperture is in front of the light sensor. Make sure that the light from the laser is aligned properly with the aperture. Push the 0-10,000 sensitivity button on the side of the Light Sensor.
- Push all components on the Optics Track as close together as possible.
- Open “Polarization” in DataStudio. Make sure the Rotary Motion Sensor will measure the Angular Position and use Large Groove in the setup.
- Click Start and slowly rotate the polarizer through 360 degrees (one revolution), then click Stop. Try to move slowly and steadily through the turning points.
- Using the Smart Tool record the coordinates of at least 20 data points from the Intensity vs. Angular position graph. The associated measurement uncertainties are ± 1 on 0-10,000 sensitivity, ± 0.01 on 0-100 sensitivity for the relative intensity and $\pm 0.1^\circ$ for the angle.
- Make sure to record the maximum intensity, I_0 , and the respective angular position, θ_0 . Maximum intensity occurs when the polarizer is aligned with the direction of the polarized laser light, so you will need to subtract θ_0 from your measured angles to make sure you are recording the angle *difference* between the polarization directions of the laser and the polarizer.

Double Slit Diffraction Pattern

- Replace the USB connector of the RMS on the Polarizer with that of the RMS on the Light Sensor.
- Set the Double Slit disk to the ($a = 0.04\text{mm}$, $d = 0.25\text{mm}$) position. Here a refers to the slits' width and d to the separation between the two slits.
- Place the laser at the far end of the Optics Track. Record the wavelength of the laser as printed on its back. Mount the High Precision Double Slit Disk to the optics bench next to the laser, with the printed side toward the laser.

- Place the cardboard in between the laser and the Double Slit Disk. Push the Double Slit Disk close to the laser so that the cardboard can stand against the components. Make sure that laser light goes through the slit on the cardboard and into the double slit. See Fig. 5.7.

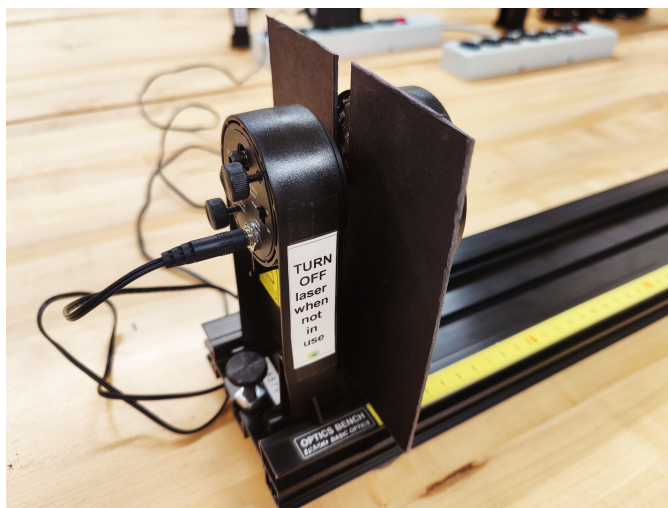


Figure 5.7: Setup for double slit experiment.

- Measure and record the distance D from the position of the double slit to the tip of the Light Sensor. You can use either the scale on the track or use a meter stick.
- Move the Light Sensor slightly left or right along the linear translator until you can see the beam somewhere on the white screen. Use the adjustment screws on the back of the laser to adjust the position of the laser beam from left-to-right and up-and-down to make the pattern on the white screen as bright as possible and aligned with the height of the slits on the Light Sensor.
- Use slit #3 on the Light Sensor and push the 0-100 sensitivity button.
- Move the Light Sensor to the left end of the linear translator arm as viewed from the laser.
- Open “Interference” in DataStudio. Click Start.
- **Slowly** move the Light Sensor across the translator arm and scan the full pattern. Hold the rear of the RMS down against the linear translator bracket so it does not wobble up and down as it moves. Try to move the sensor as smoothly as possible.
- Click Stop when you have finished scanning the light pattern and reached the right end of the linear translator.
- Use the Smart Tool to record the positions of the maxima on the central interference pattern.

Single Slit Envelope

- Please keep the data from the double-slit part. You can choose to hide the previous data in DataStudio so that the new data look clear.
- Move the Light Sensor to the center of the linear translator arm so that you can see the interference pattern on the white screen.
- Adjust the position of the cardboard carefully, until it covers one of the double slits, and you can see a clear single slit pattern on the white screen. See Fig. 5.8.



Figure 5.8: Single slit pattern on the white screen.

- To make better measurements, you will probably want to use a different slit on the light sensor. Try slit #5 for example. Make sure to take note of any changes made.
- Move the Light Sensor to the left end of the linear translator arm as viewed from the laser.
- Click start in DataStudio. You can now record the positions of the minima of the single slit pattern at the same distance D . Remember to click stop. If you show the data of double-slit part on top of that of single-slit, there should be a nice overlap resembling Fig. 5.5.
- Take a screenshot of the overlapping curves.

Summary of data:

- Polarization section:
 - Maximum intensity I_0 and the corresponding position θ_0
 - Variation in the intensity with polarizer angle, 20 pairs
- Double slit diffraction section:
 - Wavelength of the laser
 - Distance D from slits to linear translator
 - Slit separation d
 - Positions of maxima, x_m
- Single slit envelope
 - Distance D from slit to linear translator
 - Slit width, a
 - Positions of minima

5.5 Analysis

Polarization

To test Malus' Law, plot I/I_0 against $\cos^2 \theta$, where θ is the angle difference between the polarizer's angular position and θ_0 . Does this look like a straight line, as you would expect from Malus' Law? Perform a regression analysis and record the slope and intercept of the line, with their uncertainties. You can also choose to use some non-linear fitting tools to fit your original data with the \cos^2 function, and check the quality of the fit. In addition, consider the following questions:

- What reading do you obtain when the polarizer is at an angle of 90° relative to where maximum intensity is recorded? This is called the noise of your measuring device. What reading would you expect if there was no noise?
- What is the signal to noise ratio? The signal is the reading at θ_0 . What does this tell you?
- If you had a lot of background light, how could you reduce the influence it would have on your results: by changing the setting, or by changing your data analysis?
- How does the noise arise? Is it possible to eliminate it completely?
- What is the physical meaning of the y -intercept of your plot? If it is different from zero, is the difference statistically significant?
- What results would you get if you performed the same experiment using the incandescent bulb as a light source, instead of the laser?

Double Slit Diffraction Pattern

- Begin analyzing your double slit data by plotting the positions of the double slit diffraction maxima x_m against the order number m .
- Perform a least squares analysis to find the slope of the data (with uncertainties), and use the slope to estimate the wavelength λ of the laser light.
- Compare your result to the wavelength of the laser recorded before.
- How many fringes can you see? Why can't you see more? How could you improve the experiment to see more fringes?
- What limits the precision of this measurement?
- Does it matter how close the slits are to the laser? What effect does changing the slit-to-light sensor spacing have?

Single Slit Envelope

- Create a plot of x_n versus the order number n (which counts up in either direction from the central maximum).
- Determine the slope of this plot, and use it and the value of λ you found above to estimate the slit widths a . Do the widths agree with the nominal width written on the component?
- How big is the double slit envelope compared to the single slit envelope? What does this mean for the relative sizes of the slit width and slit separation? What happens if we change a of the double-slit?

Experiment 6

Interferometer

6.1 Purpose

In this experiment, you will use an interferometer base designed to be used in Michelson or Fabry-Perot mode. In the first part of the experiment, you will set up the device to obtain two separate measurements of the wavelength of a He-Ne laser. Once you have determined the wavelength of the laser, you will measure the indices of refraction of air.

6.2 Introduction

An interferometer is a device that splits a beam of light into several parts and then recombines them to form an interference pattern. The pattern can be used to measure the wavelength of the light (or other lengths) with tremendous accuracy. By interference, we naturally mean the behavior of waves when they superimpose. When two waves encounter each other, the resulting wave is the sum of the two individual waves. Interference can be totally constructive, totally destructive, or a combination of both (see Fig. 6.1). The amount depends on the relative phases of the component waves.

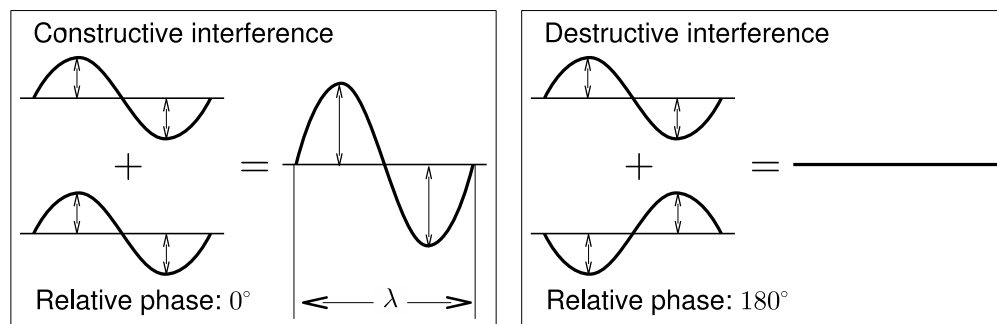


Figure 6.1: Totally constructive and totally destructive interference of two waves.

6.3 Fabry-Perot Interferometer

In a Fabry-Perot Interferometer, two partial mirrors are aligned parallel to one another, forming a reflective cavity. Figure 6.2 shows one ray of light entering such a cavity and reflecting back and forth inside. At each reflection, part of the beam is transmitted, splitting each incident ray into a series of rays. Since the transmitted rays are all split from a single incident ray, they have a constant phase relationship (assuming a sufficiently coherent light source is used).

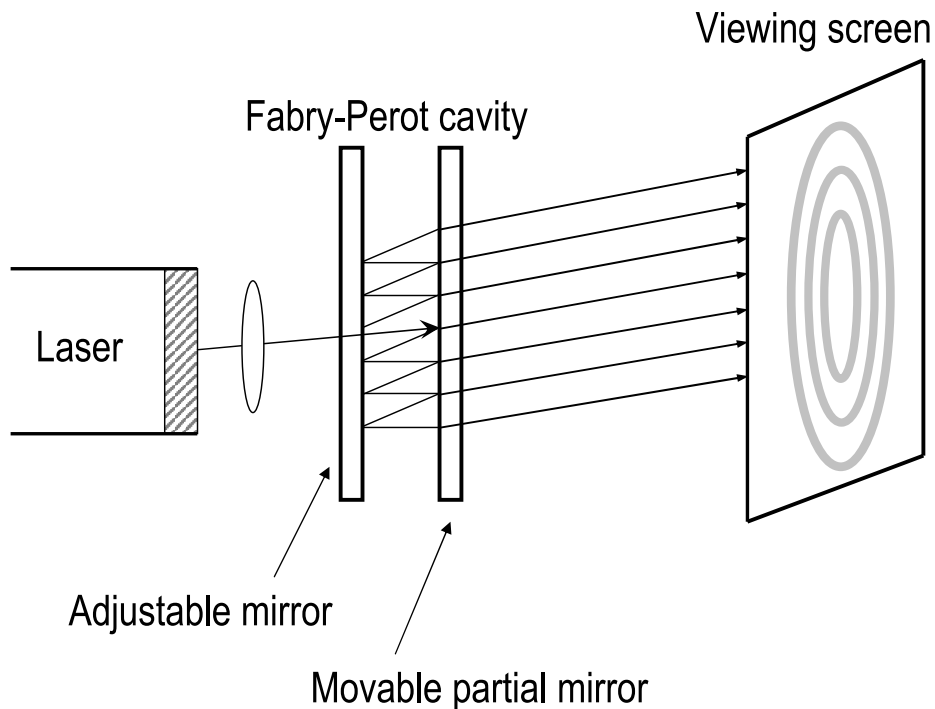


Figure 6.2: Diagram of the Fabry-Perot interferometer. This device uses a cavity to split each ray into a series of beams. (Note that the reflections in the cavity are offset for clarity.)

The phase relationship between the transmitted rays depends on the angle at which each ray enters the cavity and on the distance between the two mirrors. The result is a circular fringe pattern, with fringes that are thin, bright, and more widely spaced. An example of this pattern is depicted in Fig. 6.3. The sharpness of the Fabry-Perot fringes makes it a valuable tool in high resolution spectrometry. Moving one of the mirrors will change the distance (and thus phase relationship) between rays and will cause the fringe pattern to shift. When the movable mirror translates by a distance $\lambda/2$, the new fringe pattern is identical to the original:

$$\lambda = \frac{2d_m}{m}$$

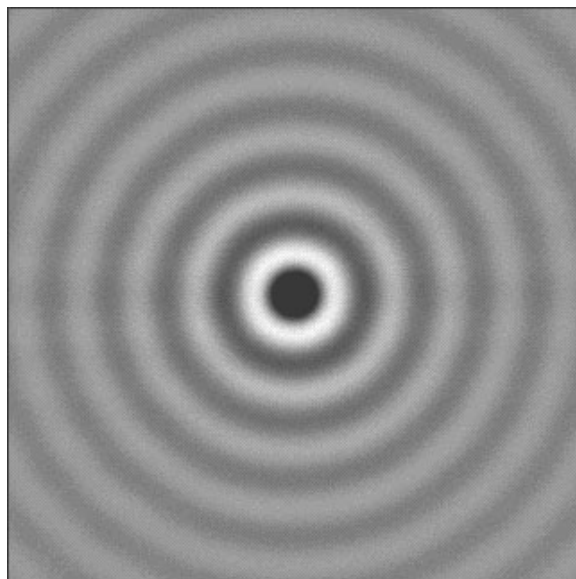


Figure 6.3: Example of the fringe-like interference pattern generated by an interferometer.

6.4 Michelson Interferometer

A.A. Michelson designed and built the interferometer in 1881 to test for the existence of the “ether,” a hypothesized medium through which light was believed to propagate, like waves in water¹. Today, the Michelson interferometer is used for a number of tasks, including measurements of the wavelength of light, measurements of extremely small distances, and for investigating optical media. The most well-known version is a 3 mile-long version called the Laser Interferometer Gravitational-Wave Observatory (LIGO), used to detect gravitational waves (the co-founders won the 2016 Nobel Prize in Physics!).

Figure 6.4 shows a diagram of a Michelson interferometer. A ray of light from the laser strikes a beam splitter, which reflects 50% of the incident light and transmits the other 50%. The incident beam is therefore split into two beams: one transmitted toward a movable mirror (M_1), and another reflected toward a fixed-position, adjustable-angle mirror (M_2). Both mirrors reflect the light directly back toward the beam splitter. Half the light from M_1 is reflected from the beam splitter to the viewing screen, and half the light from M_2 is transmitted through the beam splitter to the viewing screen, where it interferes with the light from M_1 .

Since the two interfering beams of light were split from the same initial beam, they were initially in phase. Their relative phase when they meet at any point on the viewing screen therefore depends only on the difference Δl in the lengths of their optical paths.

¹Michelson’s famous experiments discredited the ether hypothesis, hastening the development of Special Relativity.

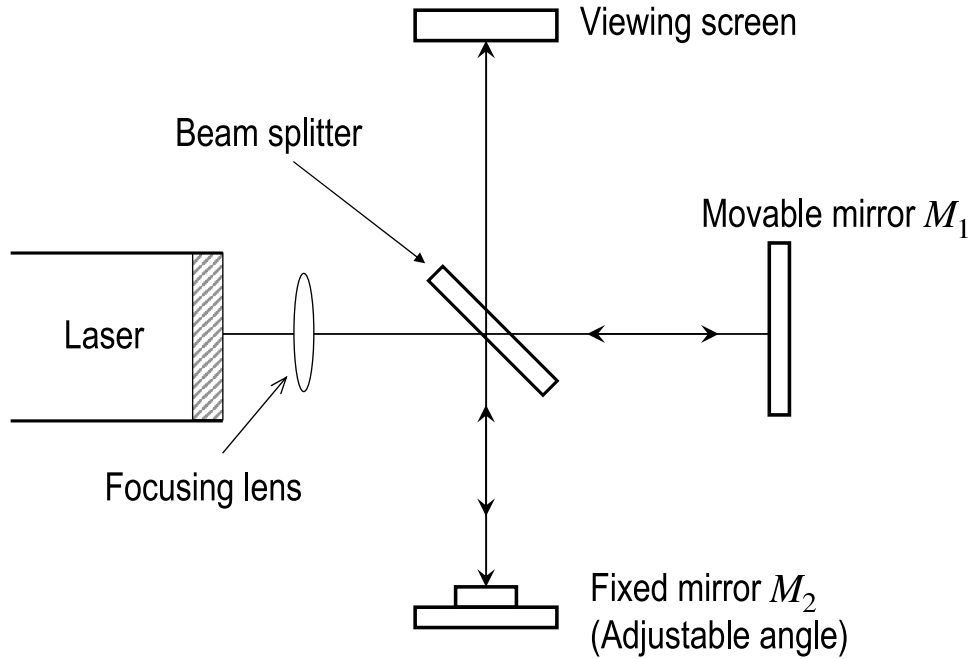


Figure 6.4: Diagram of the Michelson interferometer. A single beam is split into two rays by a half-silvered mirror. The path difference between the rays is varied with a movable mirror.

Convince yourself of the following:

1. By moving M_1 , the path length of one of the beams can be varied.
2. Moving M_1 by $\lambda/4$ (a quarter wavelength) closer to the beam splitter will reduce the overall optical path of that beam by $\lambda/2$.
3. Moving M_1 will cause the interference pattern on the screen to change

By slowly moving the mirror a measured distance d_m and counting m , the number of times the fringe pattern is restored to its original state, the wavelength λ of the light can be calculated as

$$\lambda = \frac{2d_m}{m} \quad (6.1)$$

Conversely, if the wavelength λ is already known, one can reverse the procedure and estimate the change in position d_m . In this manner, the interferometer can be used to measure extremely small distance scales.

6.5 Index of Refraction of Air

Introduction

In the Michelson interferometer, the characteristics of the fringe pattern depend on the phase relationship between the two interfering beams. There are two ways to change this phase relationship: the first is to change the distance traveled by one or both beams (by moving the movable mirror, for example); and the second is to change the medium through which one or both of the beams pass. This second method can be used to measure the index of refraction of air.

When light travels through some medium, its wavelength λ varies according to the formula

$$\lambda = \frac{\lambda_0}{n}$$

where λ_0 is the wavelength of the light in vacuum and n is the index of refraction of the medium. For reasonably low pressures, the index of refraction for a gas varies linearly with the gas pressure. As the index of refraction varies, so does the wavelength λ in the medium.

Pressure and Changing Index of Refraction of Air

You will find n_{air} by placing a vacuum cell along one leg of the Michelson interferometer. As you pump air out of the cell, the wavelength of the light inside will increase relative to that outside. This causes a shift in the interference pattern.

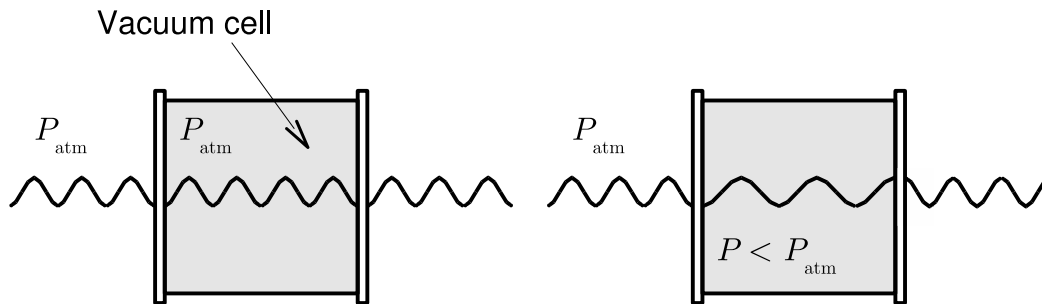


Figure 6.5: Shift in the wavelength of light in the vacuum cell when air is pumped out.

Originally there are $N_i = 2d/\lambda_i$ wavelengths of light within the cell (considering that the laser beam passes through the cell twice, on the forward and return trip). At the final pressure there are $N_f = 2d/\lambda_f$ wavelengths within the cell. The difference between these values is equal to the number of fringes you might observe. Therefore, $m = N_f - N_i = 2d(n_f - n_i)/\lambda_0$, or

$$n_f - n_i = \frac{m\lambda_0}{2d} \quad (6.2)$$

6.6 Experiment

Laser Setup

- Set the interferometer base on a lab table with the micrometer knob pointing toward you.
- Position the laser alignment bench to the left of the base, approximately perpendicular to the interferometer base, and place the laser on the bench.
- Secure one of the component holders (52-54) on the left side of the interferometer base, and another on the right side, both facing in.
- Turn the laser on. Using the leveling screws on the laser bench, adjust its height until the laser beam passes through the center of both component holders. (Use a piece of paper to check the beam path.) You may need to shift the laser bench slightly in order to adjust the horizontal path of the laser as well. This is most easily done by gently sliding the rear end of the laser transverse to the axis of the alignment bench.
- Place the movable mirror (58) in the recessed hole on the interferometer base.
- Check that the beam is reflected back from the movable mirror close to the laser aperture, and secure the mirror. (If it isn't, you may need to make small adjustments to the position of the movable mirror, and/or the alignment of the laser, until it is.) You are now ready to set up the interferometer in any of its modes of operation.

Measurement of Laser Wavelength (in Fabry-Perot Mode)

The Fabry-Perot interferometer is only used once, at the start of the lab, to measure the wavelength of the laser light.

- Align the laser and interferometer base as described in the Laser subsection above.
- Mount the adjustable mirror (51) on the interferometer base in Fabry-Perot position, facing the movable mirror.
- Place the viewing screen (62) on the component holder behind the movable mirror using its magnetic backing. You should see several images of the laser beam on the viewing screen.
- Using the thumbscrews, adjust the tilt of the adjustable mirror until there is only one bright dot on the screen.

- Attach the 18 mm FL lens (63,64) to the magnetic backing of the component holder in front of the laser, so that the lens is on the side of the component holder furthest from the laser. Adjust the position of the lens until the diverging beam is centered on the adjustable mirror. A clear sharp interference pattern should be visible on the viewing screen.

Now you are ready to measure the wavelength of the laser.

- Adjust the micrometer knob to a middle reading (approximately $500\ \mu\text{m}$). In this position, the relationship between the micrometer reading and the mirror movement is most nearly linear.
- Turn the micrometer knob one full turn counterclockwise. Continue turning counterclockwise until the zero on the knob is aligned with the index mark. Record the micrometer reading.
- Adjust the position of the viewing screen so that one of the marks on the millimeter scale is aligned with one of the fringes in your interference pattern. You may find it easier to count the fringes if the reference mark is one or two fringes out from the center of the pattern, or you may find that counting “bullseyes” is easier.
- Rotate the micrometer knob slowly counterclockwise. Count the fringes as they pass your reference mark. Continue until some predetermined number of fringes (at least 20) have passed your mark. As you finish your count, the fringes should be in the same position with respect to your reference mark as they were when you started to count. Record the final reading of the micrometer dial.
- Record d_m , the distance that the movable mirror moved toward the beam splitter according to your readings of the micrometer knob. Remember, each small division on the micrometer knob corresponds to $1\ \mu\text{m}$ ($10^{-6}\ \text{m}$) of mirror movement.
- Record m , the number of fringe transitions that you counted.
- Repeat the above steps 10 times, recording your results each time.

Measurement of Laser Wavelength (in Michelson Mode)

- Remove the 18mm FL lens (63,64).
- Mount the adjustable mirror on the interferometer base in Michelson position. Place the second component holder opposite the adjustable mirror and attach the viewing screen to its magnetic backing.
- Position the beam splitter (55) at a 45° angle to the laser beam, within the crop marks, so that the beam is reflected to the adjustable mirror. Adjust the angle of the beam splitter as needed so that the reflected beam hits the adjustable mirror near its center.

- There should now be two sets of bright dots on the viewing screen; one set comes from the fixed mirror and the other comes from the movable mirror. Each set of dots should include a bright dot with two more dots of lesser brightness (due to multiple reflections). Adjust the angle of the beam splitter again until the two sets of dots are as close together as possible, then tighten the thumbscrew to secure the beam splitter.
- Using the thumbscrews on the back of the adjustable mirror, adjust the mirror's tilt until the two sets of dots on the viewing screen coincide.
- Attach the 18 mm FL (63,64) lens to the magnetic backing of the component holder in front of the laser such that the lens is on the side of the component holder furthest from the laser. Adjust the position of the lens until the diverging beam is centered on the beam splitter. You should now see circular fringes on the viewing screen. If not, carefully adjust the tilt of the adjustable mirror until the fringes appear.

You will now repeat the measurement of the laser wavelength that you performed with the Fabry-Perot setup, which you can refer to on page 55.

- Record m the number of fringe transitions that you counted, and the final position on the micrometer knob, for 10 trials.

Measuring the Index of Refraction of Air

To measure n_i , the index of refraction of air, we will pump air out of a vacuum cell and measure the change in the number of fringes, m , for some pressure change $P_f - P_i$. Once you know how changes in pressure affect the index of refraction, you can determine n_i , the index of refraction of air, knowing P_f , P_i , and n_f . As we pump air out of the chamber, two things will happen:

- Pressure inside the cell drops $\Delta P = P_f - P_i$
- The refractive index, n_f , reduces. Recall that the refractive index of absolute vacuum is defined to be $n_{\text{vacuum}} = 1$.

Dividing both sides of equation (6.2) by ΔP , we get

$$\frac{n_f - n_i}{P_f - P_i} = \frac{m\lambda_0}{2d} \frac{1}{P_f - P_i} \quad (6.3)$$

The quantity on the left side is the slope of an n vs. P graph. Since we assume that the index of refraction varies linearly with pressure (at low pressures), this ratio is constant for any corresponding difference between n and P . Thus, we can rewrite the formula as a linear equation.

$$n_f - n_i = C(P_f - P_i) \quad (6.4)$$

where C is a constant that will be experimentally determined. Since the left-hand side is constant, the quantity on the right-hand side of the equation is also a constant.

$$C = \frac{m\lambda_0}{2d} \frac{1}{P_f - P_i} \quad (6.5)$$

- Align the laser and interferometer in Michelson mode.
- Place the rotational pointer between the movable mirror and the beam splitter.
- Attach the vacuum cell to its magnetic backing and push the air hose of the vacuum pump over the air outlet hole of the cell.
- Adjust the alignment of the fixed mirror as needed so the center of the interference patterns clearly visible on the viewing screen. The fringe pattern will be somewhat distorted by irregularities in the glass end-plates of the vacuum cell, but this is not a major concern.
- For accurate measurements, the end-plates of the vacuum cell must be perpendicular to the laser beam. Rotate the cell and observe the fringes. Based on your observations, how can you be sure that the vacuum cell is properly aligned?

Now you are ready to measure the wavelength of the laser.

- To collect data for fringe number count as the pressure in the cell changes, begin with a cell at atmospheric pressure.
- **Slowly** pump on the handle which will evacuate air from the cell and record the number of fringes, m , that have passed during the interval, and the pressure P at the end of the interval.
- Note that the pressure you measure is actually a *gauge pressure* $P - P_{\text{atm}}$, since the gauge starts at zero when the cell is at atmospheric pressure. Convert your ΔN values to $m = N - N_{\text{atm}}$, the number of fringes at a given pressure minus the number of fringes corresponding to atmospheric pressure. (You don't know N_{atm} , the number of fringes corresponding to the cell at atmospheric pressure, but you can determine the number of fringes that have passed since you started evacuating the cell, and this is $m = N - N_{\text{atm}}$). Obtain at least 7 data points for m and various ΔP .
- Convert your values for m to $n - n_{\text{atm}}$ using eq. (6.2).

NOTE: the following quantities will come in handy:

$$\begin{aligned} P_{\text{atm}} &= 76 \text{ cm Hg} \\ P_{\text{vacuum}} &= 0 \text{ cm Hg} \\ d &= 3.0 \text{ cm} \end{aligned}$$

Summary of data:

- Fabry-Perot section:
 - m, d_m for 10 runs
- Michelson section:
 - m, d_m for 10 runs
 - $m, \Delta P$ for several runs

6.7 Advice for Accurate Fringe Counting

- It's not necessary that your interference pattern be perfectly symmetrical or sharp. As long as you can clearly distinguish the maxima and minima, you can make accurate measurements.
- It's easy to lose track when counting fringes. It can help to focus on a fringe a few away from the center, or to use the "disappearing bullseye" technique, in which you observe the central minimum, to count the fringes.
- When turning the micrometer dial to count fringes, always turn it one complete revolution before you start counting, then continue turning it in the same direction while counting. This will almost entirely eliminate errors due to backlash in the micrometer movement. (Backlash is a slight slippage that always occurs when you reverse the direction of motion in a mechanical system.) Turning the micrometer dial clockwise moves the movable mirror toward the right.
- Always take several readings and average them for greater accuracy or have multiple members count the same run and compare results.
- Warm up the laser to reduce intensity and polarization variations.

6.8 Analysis

Laser Wavelength

- For the measurement of the laser wavelength, take your results for the number of fringes moved m and the distance d_m traveled by the movable mirror and convert them to a wavelength.
- Average these wavelengths, and find the standard deviation of the mean: $\bar{\lambda} \pm \sigma_{\bar{\lambda}}$.
- Compare the results from the two separate methods. Do they agree within uncertainties? Does one method appear to be more accurate than the other?

Index of Refraction of Air

- Make a plot of $n - n_{\text{atm}}$ vs. $P - P_{\text{atm}}$ for your data.
- Draw a line of best fit and determine the slope, C . This slope is slightly different from the ratio described by eq. (6.3). However, you should be able to convince yourself that changing the x and y values by a constant, n_{atm} and P_{atm} , shouldn't change the slope at all (these constant values will subtract out when calculating the slope). Therefore, we can use our graph of $n - n_{\text{atm}}$ vs. $P - P_{\text{atm}}$ to determine our constant C .
- Once you know the slope of your line, you know the constant value of the right-hand side of eq.(6.4). Now you can plug in values on the left-hand side. Take n_{atm} , the index of refraction of air at atmospheric pressure, to be the unknown value — this is the original pressure in the cell before you pumped the air out.
- For the other three values on the left-hand side you can plug in the atmospheric pressure and the pressure in a vacuum, both in cm Hg, and the index of refraction of a vacuum. This will allow you to calculate the index of refraction in air.
- Compare your value of n_{air} at atmospheric pressure and standard temperature to that given in a recognized source (eg. *CRC Handbook of Chemistry and Physics*). Comment on the results.

General Questions

- The Michelson interferometer can be used to measure extremely small distance scales. What is the smallest distance scale that can be measured? What uncertainty is associated with this? How could the precision be increased?
- What is the smallest distance scale that can be measured by the Fabry-Perot interferometer? Which is better, the Michelson or the Fabry-Perot? Why?
- How could you change either of these interferometers to measure smaller distance scales? What if you wanted to use the same equipment as you had in this experiment?

Experiment 7

The Spectrum of the Hydrogen Atom

7.1 Purpose

In this experiment, you will observe the discrete light spectrum from a gas discharge lamp. You will find that the spectrum consists of a collection of sharp monochromatic lines. Using a diffraction grating spectrometer, you will be able to measure the wavelength of the emitted light to better than one part in a thousand. Therefore, it is crucial to make all calculations to five significant figures.

7.2 Introduction

Spectrum of the Hydrogen Atom

When gases are subjected to large applied voltages, they tend to undergo dielectric breakdown and emit bright light. If one examines the light from such a gas discharge with a spectrometer¹, one finds that the light consists mainly of a few bright lines of pure color on a generally dark background. For example, excited hydrogen gas will emit four visible lines during breakdown: red, green, blue, and violet. This phenomenon contrasts sharply with the continuous spectrum of colors observed in light from the sun or an incandescent bulb.

The property of light that we observe as color is actually related to its wavelength λ . In the late nineteenth century, J.J. Balmer discovered an equation that correctly predicts the wavelengths of the visible lines in the hydrogen spectrum:

$$\frac{1}{\lambda} = R \left(\frac{1}{2^2} - \frac{1}{n_i^2} \right) \quad (7.1)$$

In this expression, $n_i = 3, 4, 5, \dots$, and R is the so-called Rydberg constant,

$$R = 1.0974 \times 10^7 \text{ m}^{-1}$$

¹A device that decomposes light into its constituent wavelengths.

Other spectral lines beyond the visible wavelengths can be observed in hydrogen and other gases; in hydrogen, these wavelengths are given by the general formula

$$\frac{1}{\lambda} = R \left(\frac{1}{n_f^2} - \frac{1}{n_i^2} \right),$$

where n_f and n_i are integers.

Balmer derived his formula for the hydrogen spectrum empirically; at the time, his result did not have a fundamental explanation grounded in classical physics. In fact, classical electromagnetism predicts that hydrogen atoms should radiate *continuously*, and, even worse, that they should be highly unstable². Neither prediction is observed, suggesting deep flaws in the classical description.

The origin of the line spectrum became a major problem that was not resolved until 1913, when Niels Bohr suggested an alternative theory for atoms. He proposed that the valence electron could only exist in certain energy states, and could only “jump” between these discrete states discontinuously. When a “jump” occurred, the atom would emit light to conserve energy. Since the energies were discrete, the emitted light should always have the same fixed set of colors (wavelengths).

Bohr was able to derive a formula for the energy of the hydrogen atom’s quantum energy levels in terms of the mass m and charge e of the electron, the permittivity of free space ϵ_0 , Planck’s constant h , and an integer n :

$$E_n = - \left(\frac{me^4}{8\epsilon_0^2 h^2} \right) \frac{1}{n^2}$$

Therefore the energy emitted by the atom during transition from an initial level n_i to a final level n_f is

$$\Delta E = E_i - E_f$$

The relationship between the energy E and the wavelength λ of the light is due to Planck: $E = hc/\lambda$. Hence, one can derive an expression for the wavelength of light emitted during an atomic transition:

$$\frac{1}{\lambda} = \frac{\Delta E}{hc} = \frac{me^4}{8\epsilon_0^2 h^3 c} \left(\frac{1}{n_f^2} - \frac{1}{n_i^2} \right) = R \left(\frac{1}{n_f^2} - \frac{1}{n_i^2} \right)$$

Resolving a Spectrum with a Diffraction Grating

In order to decompose a spectrum, one can use a so-called transmission grating. The grating is nothing more than a slab of material with a large number of tiny parallel slits. Transmission gratings are often made of finely machined glass or even crystals. The spacing d between the slits is called the “lattice constant” of the grating.

Consider a collimated (parallel) light beam incident on a grating from the left, as shown in Fig. 7.1. Each slit will diffract the beam, and act in turn as a new source of waves. The waves all begin in phase

²The electrons should fall into the nuclei within a short timespan, collapsing the atom.

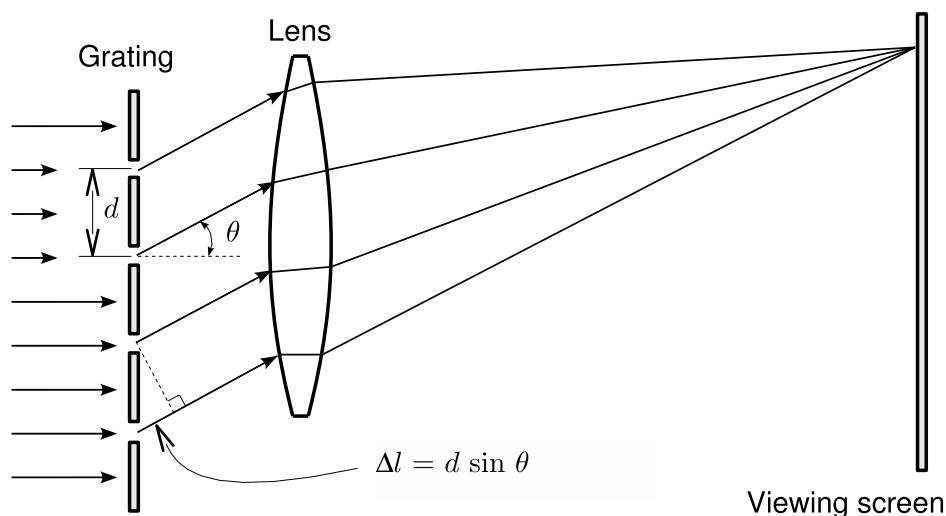


Figure 7.1: A diffraction grating with slit separation d . The location of the diffraction maximum on the viewing screen is dependent on the wavelength λ of the incident light. Note that the horizontal scale in this diagram has been highly distorted.

at the slits, but depending on the angle with which they leave the grating (called the diffraction angle θ), they travel different paths to the viewing screen and may be out of phase by that point. From the figure, the difference in the path length for two adjacent slits is

$$\Delta l = d \sin \theta$$

In order for the grating to form a maximum at some point on the viewing screen, the waves must be in phase there. This will occur if the path length difference is an integral multiple of the wavelength:

$$\Delta l = m\lambda, \quad m = 0, 1, 2, \dots$$

Therefore, we find that maxima will occur whenever

$$d \sin \theta = m\lambda, \tag{7.2}$$

where m is called the order of the maxima. Equation (7.2) tells us that the maxima for a given λ will occur at different angles with respect to the laser beam direction. This is how the transmission grating decomposes a spectrum into its individual wavelengths. When light shines through the grating, particular colors will appear at several lower and higher angles relative to the forward direction $\theta = 0^\circ$. The order number m refers to the relative position of a maximum with respect to $\theta = 0^\circ$.

7.3 Experiment

The equipment used in this experiment, called a diffraction grating spectrometer, is depicted in Fig. 7.2. The spectrometer contains three major components: a collimator tube, a rotating table, and a telescope.

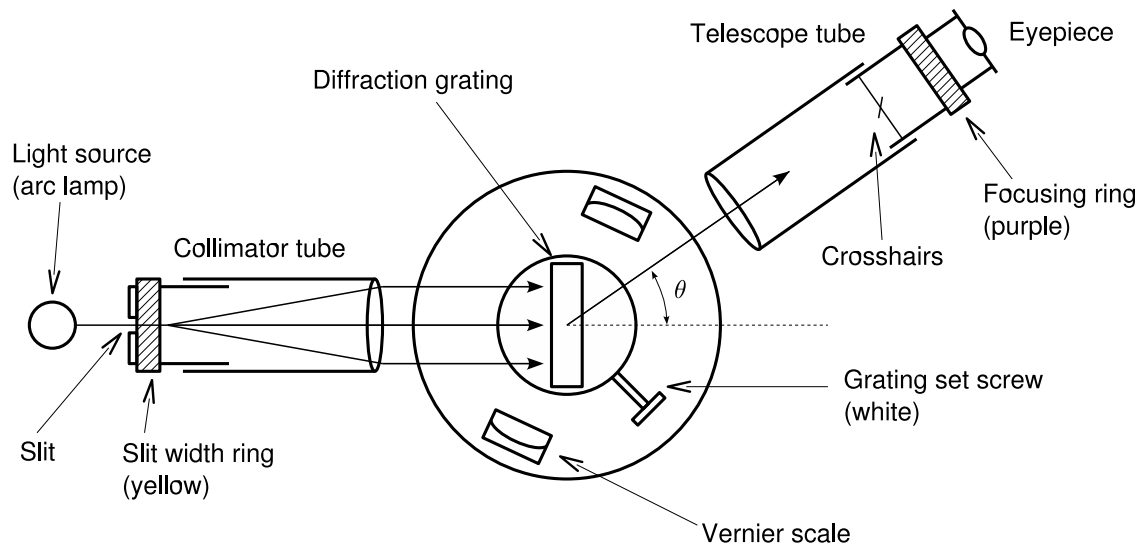


Figure 7.2: Schematic of the spectrometer you will use to observe the hydrogen spectrum.

The collimator tube takes light from excited hydrogen gas, provided here by an arc lamp, and uses a lens to collimate the beam — that is, make the light rays from the source parallel. When the parallel rays exit the tube, they travel to the transmission grating, which is mounted in the center of the rotating spectrometer base. The light diffracted by the grating may be viewed through an eyepiece at the end of the telescope tube.

The telescope is able to swivel with respect to the grating, allowing you to sweep through a set of angles θ and observe the angle dependence of the various spectral lines. Using eq. (7.2), you can then use this angle to determine the wavelength of each line you observe. To determine the angles, the spectrometer base contains a graduated circle attached to the telescope. As you turn the telescope, you can read off the angle θ using the angular scale scored into the circle.

How to Read the Angular Scale

The angular scale in the base is not a standard ruler, but a Vernier ruler. With a standard angular scale, you would probably be able to resolve angles down to the nearest degree or half degree. In this experiment, we would like considerably more precision in our measurements. Therefore, the spectrometer contains a scale that allows users to measure angles with great accuracy, to the nearest arcminute

(1 arcmin = $1' = 1/60^\circ$).

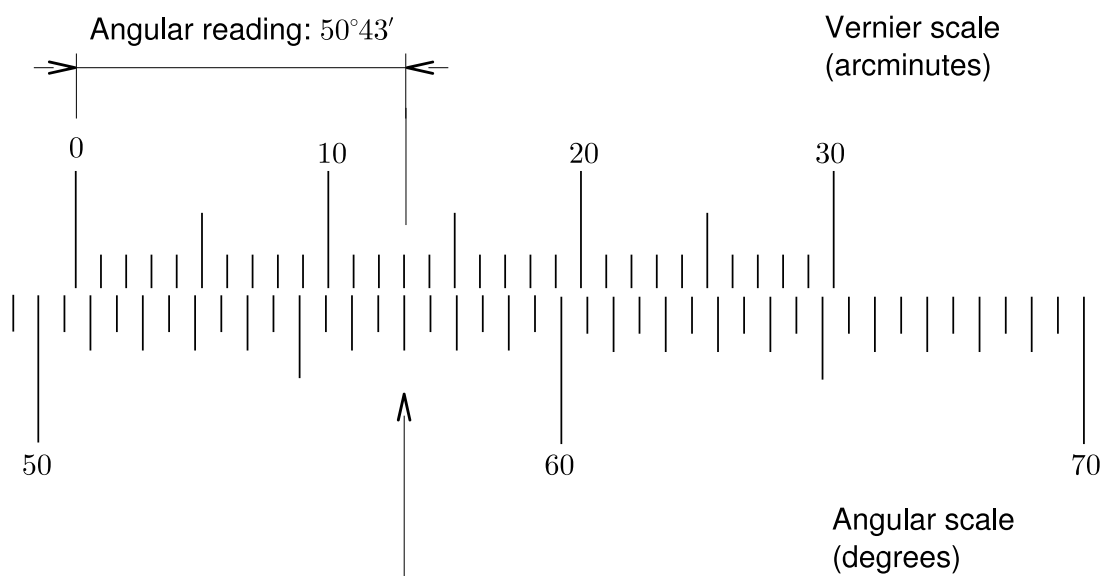


Figure 7.3: Sample reading from the degree and Vernier scales in the spectrometer base.

The device can achieve this precision by having two scales rather than one. The first is a standard degree scale running from 0° to 360° , and the second is a Vernier scale running from $0'$ to $30'$ (0° to 0.5°). To understand how these two scales work together, consult Fig. 7.3 as you read the following procedure.

1. Begin by finding the zero marker on the Vernier scale.
2. Scan down from the Vernier scale to the next line on the degree scale, as read from the left. This line is the angle θ , accurate to the nearest half degree ($30'$).
3. Reading from left to right, find the line on the Vernier scale that best lines up with a line on the degree scale. This value marks, in arcminutes, your position between two ticks on the degree scale.
4. Add the first reading from the degree scale to the second reading from the Vernier scale. You have now measured θ to the nearest arcminute.

Again, refer to Fig. 7.3 as you read this procedure. In the figure, the 0 mark on the Vernier scale is between 50.5° ($50^\circ 30'$) and 51.0° ($51^\circ 0'$). Hence, the base measurement is $50^\circ 30'$, since we are reading

from the left. On the Vernier scale, the mark 13 best matches a mark on the degree scale. Therefore, within the half degree interval, we add an additional $13'$. The resulting measurement is

$$\theta = 50^\circ 30' + 13' = 50^\circ 43' = 50.717^\circ$$

Measurement Error

For the spectrometer, you will be using the Vernier scale to measure the angular separation of the various slit images (or the angular separation of the various colored lines).

A simplistic view might be to take the error of each measurement to be the maximum resolution of the Vernier ruler, but what would be a better approach for this? *Hint: It has to do with the finite width of the lines you are measuring.*

7.4 Procedure

The first part of the experiment will basically include the procedure to set up the equipment. This should be done with as much care as possible. Only then will you be able to measure the wavelength on the limit of our apparatus. If you don't set up the spectrometer correctly you will get systematic errors, skewing your results.

Adjusting the Spectrometer

- Take the grating out of the holder and close the green knob.
- Rotate the yellow knob such that the slit is about half open.
- In the straight-through position (180°), look through the eyepiece and turn the purple focusing ring until you see a sharp image of the slit.
- Loosen the red knob and move the telescope tube until the crosshairs are in the middle of the slit. Tighten red knob.
- Open the green knob and turn the tabletop such that the zero mark from the Vernier scale with the magnifying glass is lined with either 180° or 360° from the outer scale. Always use only this Vernier scale and don't switch to the other one in between. Close the green knob and don't open it again for the rest of the experiment!
- Now you can fine adjust the relative position of the inner and outer scale by turning the blue knob. Line up the zero on the Vernier scale and $180^\circ/360^\circ$ and on the degree scale as carefully as possible. **NOTE:** For some of the spectrometers, there is a small mark to the left of the zero

mark on the Vernier scale. Make sure that you line up the zero mark, and not the extra mark, with $180^\circ/360^\circ$.

- Put the grating in the holder such that it is perpendicular to the telescope tube-collimator tube line. Close the white screw to lock the grating.

Obtaining the Grating Lattice Constant

After adjusting the spectrometer you will measure the yellow line of a helium discharge lamp. Since we know that the wavelength of this light is $\lambda = 5.8756 \times 10^{-7}$ m, we can determine the lattice constant d of the grating quite accurately. Even though the grating has $600 \text{ lines mm}^{-1}$ written on it, this is only an approximation. We want to know the lattice constant to five significant digits and not just three, and so you have to measure it.

- Switch on the helium lamp and line the spectrometer up such that you can see the slit well illuminated by the lamp as you look through the telescope.
- Put the black cardboard over the front end of your collimator tube and cover the spectrometer with a black piece of cloth to block light from your surroundings (but be careful not to block the telescope with the cloth).
- **NOTE:** This step and the next should be performed in the dark; therefore, switch on the incandescent light and turn off the fluorescent light in the room.
- Open the red knob and move the telescope tube to the left until the crosshairs are in the center of the yellow line. You should first see a few blue and green lines, then the isolated yellow line, and then red lines. The yellow line should be somewhere around 20° degrees.
- Note down the angle in degrees and minutes where you see the first order of the yellow line. Do the same on the right side and average these two numbers.
- Use the average and plug it into the grating equation ($m = 1$) to determine the lattice constant d to at least 5 significant figures.
- How many lines mm^{-1} does this lattice constant correspond to? Can you also see the second and third order yellow lines on either side of $\theta = 0^\circ$?

Measuring the Spectrum of Hydrogen Atoms

At this point, you are ready to use your calibrated equipment to observe the hydrogen spectrum. The procedure is similar to the steps you already followed during the calibration.

- Switch off the He lamp and set up the spectrometer for the hydrogen lamp.
- The light purple line you see in the middle is the 0th order line.
- There are four visible lines in the spectrum: one red (furthest out), one greenish-blue, one purple-blue and one dark purple. The dark purple line is very faint and you may not be able to see it. Look on both sides of the straight-through position since the dark purple line may be clearer on one side than the other. **Note:** You may also see “background” lines between the red and greenish-blue lines - be careful!
- Measure the angles of the four lines for at least two orders on both sides (the first and second order). Average the angles from both sides for each order.
- How many orders do you see on either side? E.g., look for the red line and count how often it appears as you go further out.
- Do the higher orders overlap? That is, does a new order start before the old one ends? What do you expect? Why?

Summary of data:

- Grating lattice constant section:
 - θ_L and θ_R for yellow line of He
- Hydrogen spectrum:
 - θ_L and θ_R for first and second order red, green, blue and purple lines.

7.5 Analysis

The analysis of the data should be fairly easy. There are not too many data points, but remember that you should propagate errors!

- Put together formulas (7.1) and (7.2) and solve the resulting expression for n_i .
- Use your data, the value for R , and $n_f = 2$ to determine n_i for both orders independently.
- Remember that n_i should be an integer number labeling the initial atomic shell occupied by the excited valence electron in each hydrogen atom.
- Are the results integers, or close to them?

- Are the integer numbers consistent with the predictions of the Balmer Formula and the Bohr Model?
- Which order gives the best results?
- For which color did the electrons jump from the lowest shell? Explain why you could have predicted that anyway.
- What energy difference do these shells correspond to? How does the He/H get excited into these states?
- For which energy level did you have the most precise measurement? Can you explain why?
- Why can you only see four lines in the Hydrogen spectrum?
- What is the 0th order line? Why does it appear?
- To what precision (in percent) were you able to measure the data? Compare that to other experiments you have performed in this course. Discuss.
- What is the precision of measurement on the spectrometer? What precision does this correspond to on the wavelength? How does this compare to measurements of the laser wavelength in the Polarization/Interference and Interferometer labs?
- Discuss the factors that determine the uncertainty in your measurements. Which of these are random, and which are systematic?

Experiment 8

Capacitance and the Oscilloscope

8.1 Purpose

The purpose of this experiment is to investigate a capacitor as it charges and discharges in DC circuits, and observe the behavior of combinations of capacitors in a circuit. You will also learn to use one of the most useful pieces of equipment in the entire course: the oscilloscope.

8.2 Introduction

A capacitor is a device for storing electric charge and energy. The simplest capacitor one can think of is just a pair of parallel metal plates, as shown in Fig. 8.1. When a charge $+Q$ is placed on the upper plate and $-Q$ on the lower plate, a potential difference V is established between them.

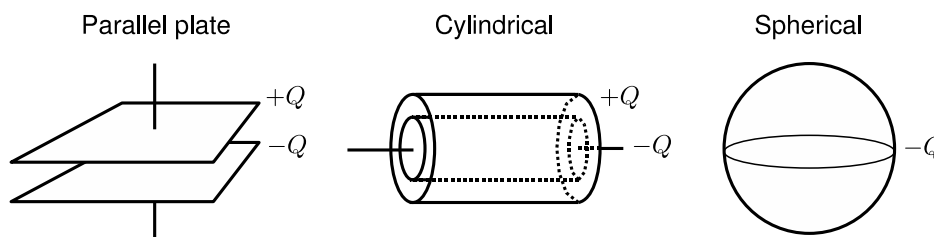


Figure 8.1: Some symmetrical geometries that can be used to make a capacitor.

The amount of charge that can be stored on the plates is directly proportional to the voltage V . The constant of proportionality is called the capacitance C :

$$Q = CV$$

As you might expect, the dimensions and configuration of the plates affect how much charge they can store. This geometrical dependence is reflected in the value of C . For any plate geometry — planar,

cylindrical, spherical, etc. — the capacitance always takes on the general form

$$C = \left(\frac{\text{Geometry factor}}{\text{units of length}} \right) \cdot \epsilon_0$$

Charging a Capacitor

A capacitor can be charged by attaching conducting wires to its plates and using a battery to establish a voltage ϵ across the device. However, the capacitor plates will not immediately gain a charge $Q = C\epsilon$; it gradually fills with charge, with the charging rate smoothly decreasing over time.

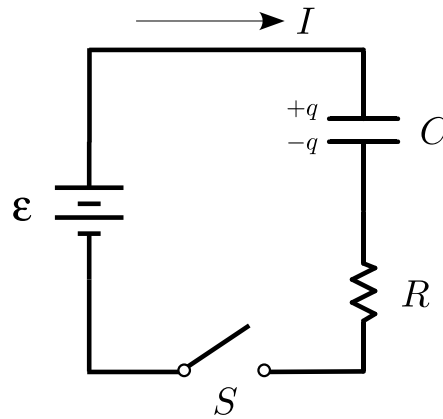


Figure 8.2: Circuit used to charge a capacitor C .

To visualize this process, consider a battery of EMF ϵ connected in series with a capacitor C , a resistance R , and a switch S (Fig. 8.2). We start with the switch open, so that no current flows in the circuit and C is uncharged. Then we flip S shut, allowing a current I to flow through the circuit. By Kirchhoff's Loop Rule, the sum of the potential drops around the circuit is zero:

$$\begin{aligned} \Delta V_{\text{battery}} + \Delta V_{\text{capacitor}} + \Delta V_{\text{resistor}} &= 0 \\ \epsilon - \frac{q}{C} - IR &= 0 \end{aligned}$$

We have written q as the instantaneous value of the charge on C . The instantaneous current and charge are related by $I = dq/dt$, so that

$$\epsilon - R \frac{dq}{dt} - \frac{q}{C} = 0$$

This differential equation can be solved for q (try it yourself), yielding

$$\begin{aligned} q(t) &= Q \left(1 - e^{-t/RC} \right) \\ I(t) &= \frac{\epsilon}{R} e^{-t/RC} \end{aligned}$$

The quantity RC , called the time constant of the circuit, sets the time scale for how long it will take to fill the capacitor up to a given level. For a given resistor R , a larger capacitor will take longer to charge.

Discharging a Capacitor

A capacitor initially filled with charge $Q = C\varepsilon$ can be discharged if a wire is used to “short” the two plates, allowing charge to flow freely between them. Of course, as in the previous case, the discharge will not be instantaneous, but will occur at a rate that depends on the size of the capacitor and the internal resistance of the wire.

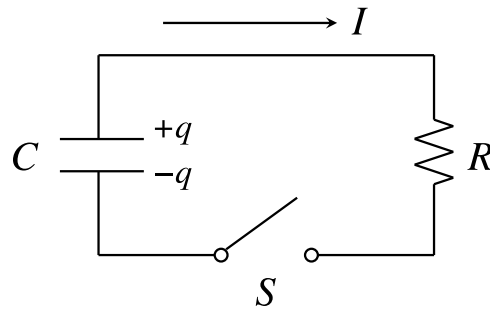


Figure 8.3: Circuit used to discharge a capacitor C .

For an example of discharge, consider the circuit depicted in Fig. 8.3. When the switch S closes, the resistor R begins to draw current I , draining the capacitor. Again, using Kirchhoff's Loop Rule,

$$\begin{aligned} IR + \frac{q}{C} &= 0 \\ R \frac{dq}{dt} &= -\frac{q}{C} \end{aligned}$$

This differential equation is easily solved for q :

$$\begin{aligned} q(t) &= Q e^{-t/RC} = C\varepsilon e^{-t/RC} \\ I(t) &= -\frac{\varepsilon}{R} e^{-t/RC} \end{aligned}$$

Comparing this with our earlier result, we see that the current that flows when discharging a capacitor from a given potential ε decreases in time just like the current in a charging capacitor. The difference in the sign of I between the two cases indicates that the currents flow in opposite directions. This makes sense — the capacitor is losing charge, so $I(t)$, the rate of charge moving through the capacitor, should be negative.

8.3 Experiment

Measuring Large RC Charging with an Ammeter

Figure 8.4 shows the circuit to be wired for measuring the current $I(t)$ when charging a capacitor C through a resistance R up to 15 V. For this part of the lab, you will use a bank of three capacitors attached to a piece of wood with three switches and labeled $30\ \mu\text{F}$ (microfarads).

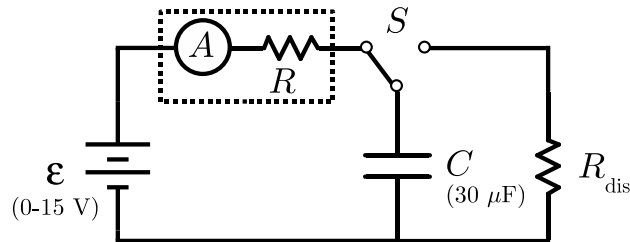


Figure 8.4: Switch circuit used to analyze charging of a capacitor C .

The capacitors in the bank are connected in parallel, such that their combined capacitances add:

$$C_{\text{total}} = C_1 + C_2 + C_3$$

By closing the switches one by one, you can increase the value of C in increments of $10\ \mu\text{F}$. The resistor R_{dis} in the circuit is used solely to discharge the capacitor bank between measurements. You will use an ammeter to measure the current I flowing through the capacitor as a function of time. The internal resistance of the ammeter, R , determines the time constant $\tau = RC$ of the charging cycle.

A Warning About Electric Shocks

A capacitance of $10\ \mu\text{F}$ sounds like a tiny amount, but don't let the prefix “micro” fool you: the bank of capacitors can store a significant amount of charge. If the fully charged bank were to discharge through your body, you would get quite a nasty shock.

While there is no need to fear the equipment, please take care when touching the bank and switches. Do **not** touch two terminals on the battery or on the capacitor bank simultaneously. When handling electrical equipment, a good rule of self-preservation is to put one hand in your pocket, which decreases the likelihood that a conducting path will close through your body.

Large RC Discharging

Use the circuit shown in Fig. 8.5 to measure the discharge current through R with the same three capacitances C used earlier. You will use the same equipment discussed above.

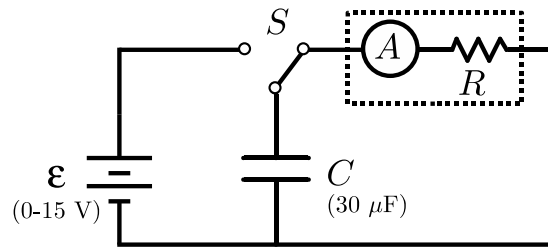


Figure 8.5: Switch circuit used to analyze discharging of a capacitor C .

You will use the switch to connect the battery across the capacitor, charging it up. Then you can flip the switch to discharge the capacitor through the ammeter, whose internal resistance R determines the time constant $\tau = RC$. By measuring the current I through the ammeter as a function of time, you can estimate the time constant and R .

Measuring Small RC Charge and Discharge with an Oscilloscope

In the first segment of the lab, you will use a combination of resistance and capacitance chosen such that $\tau = RC$ is long enough to allow observation of $I(t)$ with an ammeter. However, even for reasonable values of R and C , such as 15 $\text{k}\Omega$ and 0.01 μF , τ is too small to be measured in this manner ($\tau = 0.15$ ms).

To observe signals that vary rapidly in time (particularly periodic signals), we can use a digital oscilloscope, which is essentially a computer tailor-made to graphically display voltages as a function of time. A digital oscilloscope works by taking an input voltage $V(t)$ and sampling the waveform at regular time intervals using an **analog-to-digital converter** (ADC). The device has an internal clock, just like any home PC, that determines the sampling rate. Most digital oscilloscopes use sampling rates on the order of 1 GHz (10^9 samples per second).

The oscilloscope stores an analog input $V(t)$ as discrete set of sample points $\{V_i\}$. It then displays the waveform on a screen (usually an LCD) as a function of time. To display the function, the scope uses an interpolation scheme to calculate the value of $V(t)$ in between the sample points $\{V_i\}$. Effectively, the display rebuilds the analog function from the sampled waveform¹. The scope will properly reconstruct those waveforms whose maximum frequency is less than half the sampling rate; otherwise, the reconstructed signal becomes highly distorted by the digitization process (why?).

Because the digital scope is essentially a computer, it not only displays fast signals as a function of time, but also performs a number of real-time calculations on those signals. Familiarizing yourself with the scope and learning to use these functions is a primary goal of this experiment.

¹This is analogous to how a stereo plays sound using the digitally recorded audio information stored on a CD.

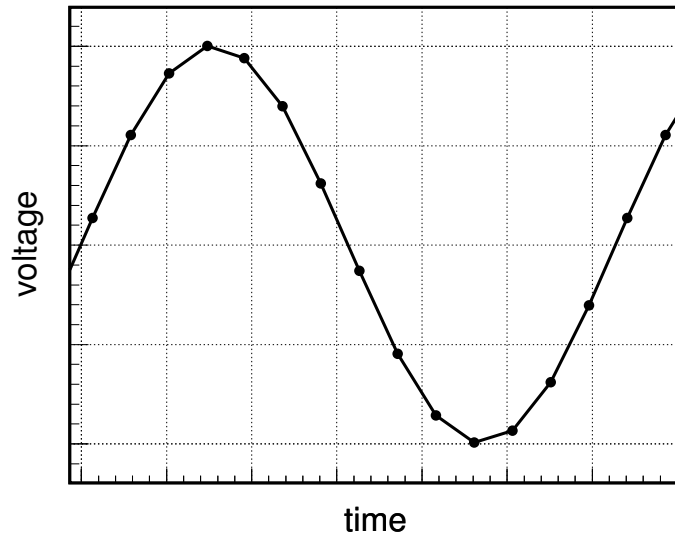


Figure 8.6: Analog-to-digital conversion of a sinusoidal voltage $V(t) \rightarrow V_i$. An oscilloscope displays the voltage on an LCD screen by interpolating between the sampled points V_i .

The Tektronix TBS-1052B-EDU Digital Oscilloscope

The digital scope used in this laboratory is the Tektronix TBS-1052, depicted in Fig. 8.7. Do not be intimidated by the number of buttons on the faceplate; there are only several functions that you will need to learn in order to use this device.

For your reference, the controls and functions you will need to use are:

VERTICAL Controls: The vertical controls for CH 1 and CH 2 are identical. These controls primarily affect the display of the vertical axis, or voltage, of the input. Of these controls, you will be using:

1. Position (1): this knob positions a displayed waveform vertically on the LCD screen.
2. Scale (2): this adjusts the scale of the vertical (voltage) axis of the LCD display, allowing you to view smaller or larger signals.

CH1/CH2 Input (3): These inputs, compatible with BNC-type transmission cables, are used to carry signals to the scope for display. There are two channels, allowing you to display two inputs simultaneously.

HORIZONTAL Controls: The horizontal controls affect the display of the horizontal time axis. The controls you will use are:

1. Position (4): this knob will move a displayed waveform horizontally on the LCD screen.



Figure 8.7: Faceplate of the oscilloscope used in this experiment.

2. Scale (5): this knob selects the horizontal time scale of the LCD trace. Quite useful for viewing different time features of a signal.

Ext Trig (6): This is where you plug in the trigger signal, that is, some periodic wave that you know is at the same frequency as the signal you are trying to analyze.

TRIGGER Controls: This set of controls adjusts where the oscilloscope searches for an input signal and how it displays it. When you are comfortable displaying signals, you may explore the trigger menu. Novices should instead use the oscilloscope's Autoset feature. However if you do want to play with it, the two useful controls are:

1. Trigger Menu (7): Press once to display Trigger Menu on the screen. Press for more than 1.5 seconds to display the trigger waveform instead of the channel waveform.
2. Trigger Level (8): If you are having trouble setting the trigger level correctly, push this knob to set trigger level to 50%. This will automatically set it halfway between the max and min of the trigger signal. From here you can adjust easily.

RUN/STOP (9): Freezes the current image on the screen.

MEASURE Button (10): The scope can perform a number of automated measurements for you. To access these, hit the Measure button at the top of the faceplate, and then use the menu box buttons to pick the settings you want.

AUTOSET Button (11): An extremely useful button, located at the top right of the faceplate. Auto-set obtains a stable waveform display for you, automatically adjusts the vertical and horizontal scales, and sets up the trigger.

Multipurpose Knob (12): You may turn or push this knob after activating one of the button functions at its right. Press Measure to select the type of automated measurements in CH1 or CH2. Turn Multipurpose knob to scroll and highlight, then push knob to select automatic measurement type. If you wish to use the Cursor function, press the Cursor button, then turn the Multipurpose knob to position the selected cursor. Automatic measurements are more accurate than cursor measurements.

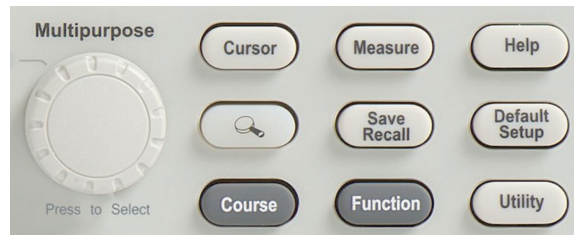


Figure 8.8: The multipurpose knob and the functions at its right.

MENU BOX Buttons: These are five blank buttons located immediately to the right of the LCD screen. When you enable any menu function on the scope (TRIGGER, CH 1, MEASURE, etc.), these buttons will help you to control the functions.

8.4 Procedure

Large RC Charging

Set up up the capacitor bank, switch, power supply, ammeter, and $10\text{ k}\Omega$ resistor to measure large RC charging, as shown in Fig. 8.4. Before closing the switch to the power supply and starting each new measurement, momentarily flip the switch such that the capacitor bank discharges across the resistor.

Set up the bank such that the total capacitance is $10\text{ }\mu\text{F}$; then flip the switch connecting the capacitor in series with the battery. As the capacitor bank charges, measure I at regular time intervals and record your results with uncertainties. If necessary, discharge the capacitor and repeat the measurement to acquire more data points. When you have finished, set up the bank for $20\text{ }\mu\text{F}$ and $30\text{ }\mu\text{F}$ and repeat these measurements.

Make sure to also record the voltage setting you use on the adjustable power supply.

Large RC Discharging

Set up the circuit drawn in Fig. 8.5. Again, you will measure the discharge rate for three values of the capacitance C .

Before closing the switch to begin the measurement, momentarily flip the switch to connect the power supply across the capacitor. DO NOT under any circumstances take leads from the battery and touch them to the capacitor terminals by hand! Use the switch.

When the capacitor is charged, flip the switch to allow it to discharge through the ammmeter. Record your values of $I(t)$. Make sure to also record the voltage setting you use on the adjustable power supply.

Display Periodic Signals on the Oscilloscope

Use the signal generator to input sine waves and square waves of various frequencies. You should connect the leads from the function generator to the oscilloscope CH 1 or CH 2. Be sure that the ground on the function generator is connected to the ground on the oscilloscope (the ground banana plug male connector is black). Press AUTOSET to acquire the signal.

Vary the shape (square, sine or sawtooth), the amplitude, and the frequency of the input signals produced by the function generator and observe the effect on the oscilloscope display. You will need to change the Vertical scale setting using the knob on the appropriate channel to make the signal appear larger or smaller on the screen; the Horizontal Scale knob will control the horizontal scale. One full division refers to the length of the sides of the square boxes.

Practice with Automatic Measurement

Set the function generator to output a sine wave and display the result using Channel 1. You will now use the oscilloscope to automatically output basic information about the signal.

- Push the **MEASURE** button to display the Measure menu.
- Push the second menu box to select **CH1**.
- Turn the **Multipurpose** knob to highlight **Period** then push the **Multipurpose** knob to select it.
- Turn the **Multipurpose** knob again to highlight **Frequency** then push to select it.
- Turn the **Multipurpose** knob again to highlight **Peak-Peak** then push to select it.
- Push the round **Menu On/Off** button below the last menu box to turn off the menu.

The bottom of the LCD screen should now display the frequency, period, and peak-to-peak amplitude of the input. Vary the amplitude and frequency of the input waveform using the function generator

controls, and observe what happens on the LCD screen. Try out the measurement for square and triangular waves, if the function generator allows it.

Note that you must display at least one full cycle of the waveform for the period and frequency calculations to work properly. You can alter how much of the wave you display by turning the Horizontal Scale knob.

Observation of a Short RC Time

Set up your circuit as shown in Fig. 8.9 Use the smallest R and C available. For this part you should use a square wave with a frequency of about 100 – 150 Hz; adjust the time scale (the x -axis scale) so that half of the charge/discharge cycle fits in the LCD window, as in Fig. 8.10. Also, make sure that the ground (black banana plug jack) of your function generator is connected to the ground of your oscilloscope.

Note that you may need to tune the frequency of the function generator a bit to get your display to look like Fig. 8.10. As you change the frequency, it will also be necessary to alter the time axis on the scope. Be patient: the measurement cannot be done well unless you have a nice exponential curve.

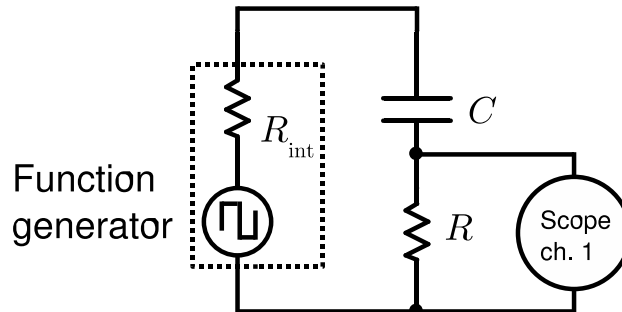


Figure 8.9: Observing the time behavior of a small RC combination.

You can get a fairly exact value for the time constant, RC , from the oscilloscope display by using the cursor function as follows:

- Adjust the Vertical and Horizontal scale settings of the input so that one part of the charge/discharge cycle fills the LCD screen, as in Fig. 8.10.
- Push the **Cursor** button.
- Push the top menu box button to select the **Amplitude** mode using the **Multipurpose** knob.
- In **Amplitude** mode, turn the **Multipurpose** knob and place one cursor line at the top of your curve.

- Push the last menu box to select **Cursor 2** and place it at the bottom of your curve.
- The delta menu box on the right-hand side displays the value of the difference ΔV between the two cursor values. This is the total voltage difference across your resistor over the cycle.
- After a time $t = RC$, the voltage will drop to $1/e = 37\%$ of its initial value. To find this point, calculate 37% of ΔV and move the bottom cursor to that location.
- Once you have the voltage cursors appropriately placed, use the Cursor menu to switch to **Time** mode and align the vertical cursors with their positions in voltage mode. Read off the time difference Δt between these two points. This time difference is just the time constant τ .
- Repeat this procedure for several different values of the function generator frequency.

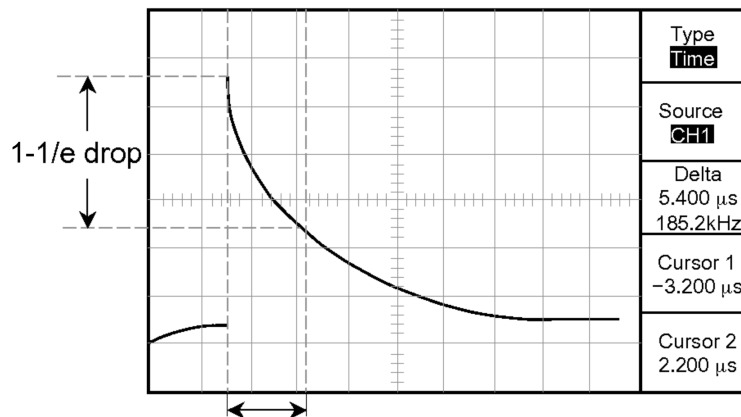


Figure 8.10: Estimate of the time constant using the oscilloscope Cursor function. (Note that this figure is not shown on the correct scale.)

Summary of data:

- Current I vs time: three capacitances, charging and discharging
- Resistance R , capacitance C for **all** circuit elements used
- ΔV and $t = RC$ from oscilloscope: several function generator frequencies f .

8.5 Analysis

Large RC Charging and Discharging

The analysis of the capacitor bank data proceeds in the same way for both the charging and discharging cycles.

- Plot your current data $I(t)$ vs time, and use a logarithmic scale on the y -axis. Because $I(t)$ is an exponential function in time, it will look like a straight line when plotted in this manner. This happens because the logarithm *linearizes* the exponential:

$$\begin{aligned}\ln I &= \ln \left(I_0 e^{-t/RC} \right) \\ &= \ln I_0 - \left(\frac{1}{RC} \right) t\end{aligned}$$

In other words, $\ln I$ is a linear function of t :

$$\ln I = mt + b$$

- Transform your current data $I(t)$ into linear form by taking the \ln of the values you recorded. Remember to transform the uncertainties on the data points as well. Perform a linear regression analysis of $\ln I$ versus t , and from the slope m and its standard error σ_m estimate the time constant $\tau = RC$ for this circuit. Do you find that the time constants obtained from charging and discharging the circuit agree within errors? What could cause any disagreement?
- Given that we use $10\mu\text{F}$, $20\mu\text{F}$ and $30\mu\text{F}$, we should see that $\tau_{30\mu\text{F}}/\tau_{20\mu\text{F}} = 3/2$, $\tau_{30\mu\text{F}}/\tau_{10\mu\text{F}} = 3$ and $\tau_{20\mu\text{F}}/\tau_{10\mu\text{F}} = 2$. Do your results agree within errors?

Short RC Measurement with the Oscilloscope

- Compare your measurements of τ from the digital scope with the estimate you expect using RC . Remember that your value for R should not only include the resistance of the circuit component you used, but also the $50\ \Omega$ output impedance of the function generator.
- Do the calculated time constants agree within uncertainties, and with RC ? If there is a significant difference, can you explain why?
- Consider that typical circuit elements may have tolerances as poor as 10%. Does this explain any discrepancies?
- Does the magnitude and shape of the voltage across the resistor agree with what you expect? Justify your results.

- What should the voltage difference across the capacitor look like? Draw the time-dependent behavior, and indicate the magnitude of the variation.
- What should the voltage across both elements, the resistor and capacitor, look like? Why would you expect that anyway?
- The scope has an approximate internal resistance of $1\text{ M}\Omega$. Estimate what effect this has on the time constant of the circuit. What if you used a $1\text{ M}\Omega$ resistor in the circuit instead of the $10\text{ k}\Omega$ resistor? By how much would this change the time constant?
- What would happen if you used a much higher frequency, say 1 kHz ? Can you estimate what range of frequencies will give you accurate measurements of τ ?

Experiment 9

AC Circuits

9.1 Purpose

You will be investigating the voltage and current response of an alternating circuit (AC) to various impedance changes. The experiment will make extensive use of the digital oscilloscope.

9.2 Introduction

An alternating current, or AC, circuit is a device containing the usual elements (resistors, capacitors, etc.), but driven by a voltage that varies in time. In this experiment, we will study the behavior of AC circuits using series combinations of a resistor R , inductor L , and capacitor C , as depicted in Fig. 9.1.

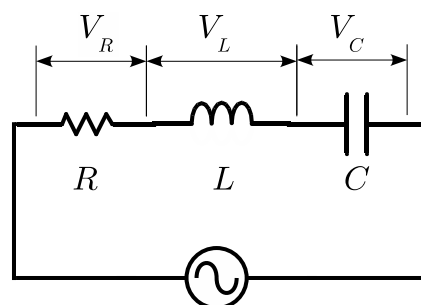


Figure 9.1: A series RLC circuit driven by a sinusoidal voltage, with components voltages indicated.

Suppose some time-dependent voltage $V(t)$ drives the circuit in Fig. 9.1. We can analyze the

circuit's behavior by writing down Kirchhoff's Loop Rule:

$$\begin{aligned}\sum \Delta V &= 0 = V(t) - V_R - V_L - V_C \\ V(t) &= IR + L \frac{dI}{dt} + \frac{q}{C} \\ &= IR + L \frac{dI}{dt} + \frac{1}{C} \int I dt\end{aligned}$$

We would like to solve this differential equation for $I(t)$ — and hence V_R , V_L , and V_C — given a driving voltage $V(t)$. For arbitrary $V(t)$, the equation is quite difficult to analyze. However, if $V(t)$ is sinusoidal, driving the circuit with an angular frequency $\omega = 2\pi f = 2\pi/T$ such that

$$\begin{aligned}V(t) &= V_{\max} \sin \omega t + \phi_V \\ I(t) &= I_{\max} \sin \omega t + \phi_I,\end{aligned}$$

then the current in each circuit element will also be sinusoidal, though perhaps out of phase with the voltage across that element.

Voltage and Current in the Resistor R

By Ohm's Law, the voltage across the resistor is $V_R = IR$. Because the resistor is in series with the voltage supply, the current running through it must have the same amplitude and phase as $I(t)$. Substituting $I(t)$ into Ohm's Law, we find

$$V_R = I(t) \cdot R = I_{\max} R \sin \omega t$$

That is, the voltage and current in the resistor have the same time dependence, and are *in phase* with each other.

Voltage and Current in the Inductor L

For a circuit element like the inductor, the voltage can be obtained from Faraday's Law,

$$V_L = L \frac{dI}{dt}$$

Again, because the inductor is in series with the voltage supply and the resistor, the current moving through it is the same $I(t)$ used above. Substituting and differentiating $I(t)$, we obtain

$$\begin{aligned}V_L &= \omega L I_{\max} \cos \omega t \\ &= \omega L I_{\max} \sin \left(\omega t + \frac{\pi}{2} \right) \\ &= I_{\max} X_L \sin \left(\omega t + \frac{\pi}{2} \right)\end{aligned}$$

The quantity $X_L = \omega L$ is called the **reactance** of the inductor. The comparison of X_L to the resistance R in Ohm's Law is actually a good way to think about how the inductor works. For an inductor, the current running through the device will be out of phase with the voltage across it.

There are two things you should note about this result:

1. In the inductor, the voltage and current are *out of phase*; the voltage "leads" the current by a phase shift $\phi = +90^\circ$.
2. The maximum voltage occurs when $\sin(\omega t + \pi/2) = 1$; at this location,

$$V_{\max} = \omega L I_{\max}$$

By replacing ωL by the variable X_L , we get an expression that looks rather like Ohm's Law:

$$V_{\max} = I_{\max} X_L$$

Voltage and Current in the Capacitor C

For the capacitor, the voltage V_C is obtained from the relationship

$$\begin{aligned} V_C &= \frac{q}{C} = \frac{1}{C} \int I(t) dt \\ V_L &= \frac{1}{C} \left(-\frac{I_{\max}}{\omega} \cos \omega t \right) \\ &= \frac{1}{\omega C} I_{\max} \sin \left(\omega t - \frac{\pi}{2} \right) \\ &= X_C I_{\max} \sin \left(\omega t - \frac{\pi}{2} \right) \end{aligned}$$

The expression $X_C = 1/\omega C$ is also called the reactance; but unlike X_L , it actually decreases as a function of ω . The capacitor is like a mirror image of the inductor, strongly resisting the flow of current when the driving frequency ω is small, but barely resisting the flow of high frequency signals.

As before, notice two things about this result:

1. The voltage and current are *out of phase*, but the voltage now "lags" behind the current by a phase shift $\phi = -90^\circ$.
2. At the maximum voltage, where $\sin(\omega t - \pi/2) = 1$,

$$V_{\max} = \frac{I_{\max}}{\omega C}$$

Again, replacing $1/\omega C$ by a new variable X_C yields an equation that looks like Ohm's Law:

$$V_{\max} = I_{\max} X_C$$

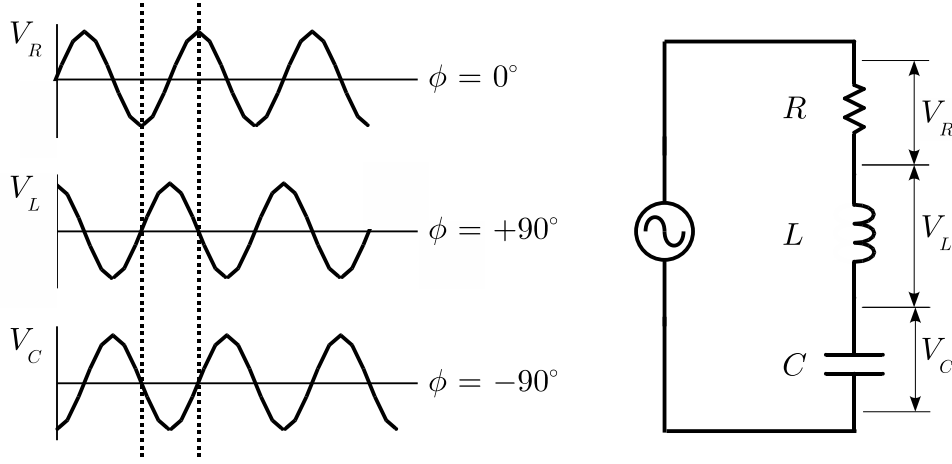


Figure 9.2: Summary of the various phase relationships for a series RLC circuit.

Impedance and Phase of the RLC Circuit

Now that we understand the voltages across each individual element V_R , V_C , and V_L , we can try to make sense of the overall RLC combination. The total voltage V_{RLC} across the series combination is just the algebraic sum of the voltage drops across each component:

$$\begin{aligned}
 V_{RLC} &= V_R + V_L + V_C \\
 &= I_{\max} \left[R \sin \omega t + X_L \sin \left(\omega t + \frac{\pi}{2} \right) + X_C \sin \left(\omega t - \frac{\pi}{2} \right) \right] \\
 &= I_{\max} [R \sin \omega t + (X_L - X_C) \cos \omega t]
 \end{aligned} \tag{9.1}$$

Now, we know that the solution V_{RLC} must be sinusoidal, and must have some phase ϕ :

$$\begin{aligned}
 V_{RLC} &= V_{\max} \sin(\omega t + \phi) \\
 &= V_{\max} [\sin(\omega t) \cos \phi + \cos(\omega t) \sin \phi]
 \end{aligned} \tag{9.2}$$

Since eq. (9.1) and eq. (9.2) must be equal at all times t , we can equate the coefficients of $\sin \omega t$ and $\cos \omega t$ and solve for ϕ :

$$\left. \begin{aligned} V_{\max} \cos \phi &= I_{\max} R \\ V_{\max} \sin \phi &= I_{\max} (X_L - X_C) \end{aligned} \right\} \Rightarrow \tan \phi = \frac{X_L - X_C}{R} = \frac{\omega L - 1/\omega C}{R}$$

The overall phase shift of the RLC combination depends on the reactive parts of the circuit X_L and X_C . Since these are functions of frequency, the phase will change if the driving frequency ω changes. As a final step, we can eliminate ϕ from these two equations by squaring and adding them:

$$\begin{aligned}
 V_{\max} &= I_{\max} \sqrt{R^2 + (X_L - X_C)^2} \\
 &= I_{\max} Z,
 \end{aligned}$$

where $Z = \sqrt{R^2 + (X_L - X_C)^2}$. Once again, this looks suggestively like Ohm's Law: voltage equals current times “resistance.” The quantity Z , called the **impedance** of the RLC circuit, is in fact just a generalized, frequency-dependent resistance. The **resistive** part of Z (just the resistor R) resists all signals in a frequency-independent way; it is also responsible for dissipating the power delivered by the voltage supply as heat. The **reactive** part of Z (the combination $X_L - X_C$) does not dissipate energy, but is responsible for determining the circuit's frequency response.

A way of visualizing this is using phasor diagrams. Notice how all the equations derived in this subsection can be obtained via Pythagoras Theorem from Fig. 9.3.

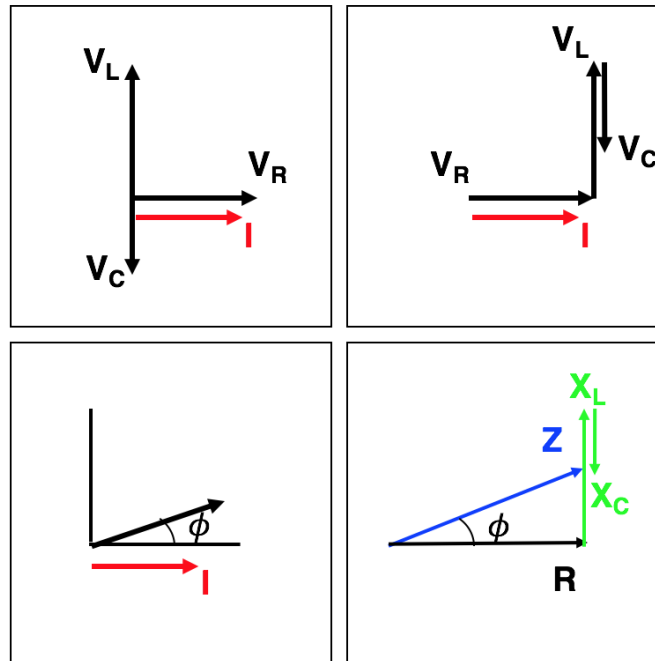


Figure 9.3: A series of phasor diagrams providing visual illustration on how the reactances influence the voltage and current of a RLC circuit. The voltage contributions of the components have to be treated as vector sums due to the phase differences between them.

Frequency Response of the RLC Circuit

Because the RLC circuit contains reactive elements, the voltage V across the resistor depends on the frequency ω of the AC source. The frequency dependence, depicted in Fig. 9.4, has a **resonance**; that is, there is a frequency ω_0 where the impedance Z takes on a minimum value, so the current takes on a maximum value for a given driving voltage.

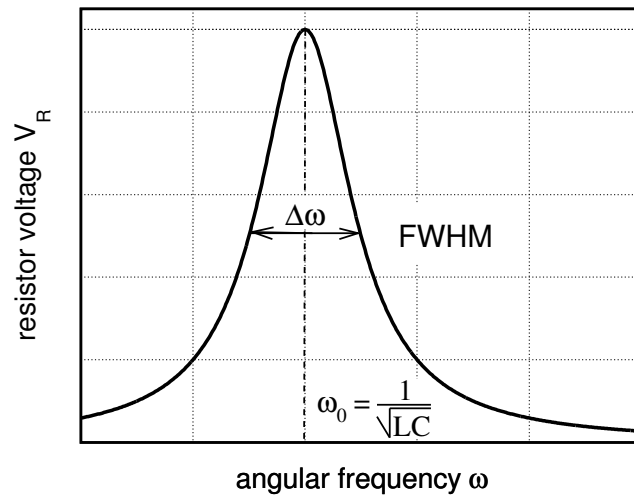


Figure 9.4: Frequency response of the RLC circuit, showing the resonance frequency ω_0 .

The impedance is clearly a minimum when $X_L = X_C$, so that the resonance frequency is

$$\begin{aligned}\omega_0 L &= \frac{1}{\omega_0 C} \\ \omega_0 &= \frac{1}{\sqrt{LC}}\end{aligned}$$

The resonant behavior of the RLC circuit actually has a pretty intuitive explanation. Remember that at low frequencies, the capacitor has a high “resistance” to current flow, but the inductor allows current to move freely. At high frequencies, it is the inductor that resists the signal, while the capacitor has a very low resistance. In between, there is a crossover region around ω_0 where the influences of both effects are minimal.

The width of the peak at half of the maximum voltage or current, denoted $\Delta\omega$ or FWHM (full width at half maximum) depends on the size of the resistor in the RLC combination. Smaller resistors will produce sharper peaks, making the circuit more sensitive to frequencies near ω_0 .

The frequency response of the RLC combination makes it a very useful filter for AC signals. Called a bandpass filter, circuits like this are employed when we want to pick out a frequency from a range of possible signals. A perfect example is the radio tuner in an automobile.

9.3 Experiment

You will record several observations of the behavior of the series RLC circuit shown in Fig. 9.5. The procedure makes heavy use of the digital oscilloscope used in the Capacitance experiment. If you still feel uncomfortable using the scope, refer back to that section of the lab manual to refresh your skills.

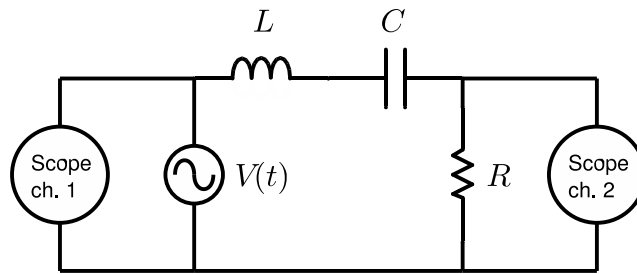


Figure 9.5: Setup of the RLC circuit, function generator, and scope.

Besides the oscilloscope, there are only a few pieces of equipment you will need to use:

1. A variable resistor, in the form of a decade resistor box.
2. A capacitor provided for you in the lab. Be sure to note the capacitance.
3. Two inductors: a small 150 mH circuit component, and a large induction ring of unknown L .
4. A function generator, which you will use to send sinusoidal signals through the circuit.

9.4 Procedure

Setting Up

The first thing you should do is look at your driving signal on the oscilloscope.

- Connect the output of the function generator to CH 1 on the oscilloscope using wire leads of a single color.
- Make sure the ground on the function generator is connected to the ground on the scope. Ground is indicated GND on the black female BNC-to-banana plug connector.
- Set up the function generator to produce a sinusoidal signal in the range of 1 kHz, with a peak-to-peak voltage of $V_{pp} = 20$ V.
- Set up the oscilloscope to trigger on the function generator. You can do this by pressing the Trigger button, and then on the menu that appears on the right-hand side of the display, press the button next to trigger source until CH 1 appears.

You should now be able to alter the VOLTS/DIV and SEC/DIV settings in order to view the sine wave produced by the function generator. If you want to take a shortcut, you can quickly obtain the

appropriate settings by pressing the AUTOSET button. Note also the following general hints for dealing with the equipment:

- If at any time the scope display is very noisy, it is possible that you have lost the trigger. You can fix this by adjusting the trigger level knob in the TRIGGER controls, or turn off CH 2 and press AUTOSET.
- The frequency recorded by the scope reads is f , in Hz, or cycles per second. The angular frequency of your RLC circuit, ω , given in rad sec^{-1} , is related to f by $\omega = 2\pi f$.

Resonance

- Use a multimeter to measure the resistance of the 150 mH inductor. Connect the inductor, the capacitor, and the decade resistor box in series with the function generator using cables of a single color (different from that used above), as depicted in Fig. 9.5.
- Connect CH 2 of the oscilloscope across the resistor, **making sure that the ground on the oscilloscope and the ground side of the resistor are connected to the ground on the function generator**; use wires of a color different than the two colors used in earlier procedure steps. Your oscilloscope and function generator now share a common ground.
- Set the decade resistor box to a resistance of $50\ \Omega$. Make sure you see a stable display of both CH 1 and CH 2.
- Using your circuit components, estimate the resonance frequency you expect to observe.
- Use the frequency knob on the function generator to try to find the circuit resonance. Sweep through a set of frequencies until you observe the signal on CH 2 peaking around some value. NOTE: observe when the peak-to-peak voltage of the sine wave on CH 2 is maximized.
- When you have found the resonance frequency ω_0 , slowly scan over frequencies and record the peak-to-peak voltage you observe in CH 2. You can use the scope's MEASURE function to automatically display the peak-to-peak signal amplitude.
- Make sure you scan through a wide enough frequency range that you can record the FWHM of the signal. Record the voltage amplitude at 20 points (at least) around the location of the resonance.

Repeat this procedure for a lower value of the resistance, for example $10\ \Omega$, and a much higher value, for example $500\ \Omega$. Make sure you record the value of R for each voltage table you record.

When you are finished, replace the 150 mH inductor with the large copper ring. Be careful: the ring is quite **heavy**! Sweep with the function generator through a range of frequencies until you find the

resonance frequency ω_0 . Using this measurement and the value of C , estimate (with uncertainties) the inductance L of the copper ring.

Phase of Driving Voltage and V_R

- Replace the copper ring in your setup with the small inductor you used earlier.
- Set the decade resistor box to a resistance of $30\ \Omega$.
- Make sure you can see the driving signal on CH 1 and the voltage across the resistor on CH 2. Note whether the voltage across the resistor is ahead of, behind, or the same as the driving voltage.
- Vary the frequency of the function generator to determine the relationship between the driving voltage and the current in the circuit at, above and below the resonant frequency ω_0 , and record your qualitative observations.
- For five values of the frequency at and around ω_0 , measure the phase difference between the driving voltage $V(t)$ and the current in the resistor. You can do this using the CURSOR function of the scope. Measure the total time for one full cycle of the driving signal, T_d , and then measure the time difference $t_R - t_d$ between the max (or min) of the driving signal and the max (or min) of the signal across the resistor.
- Be sure to record error and propagate it. The phase difference ϕ is equal to

$$\phi = 2\pi \frac{t_R - t_d}{T_d}$$

Note that this formula implies that a positive phase shift corresponds to the driving voltage peaking earlier, and therefore to the left of, the resistor voltage.

Phase of $V(t)$, V_L , and V_C

Above you found that the driving voltage and the voltage across the resistor (and thus the current in the circuit) are in phase at the resonant frequency. In this last part of the lab you will observe the phase differences between the driving signal $V(t)$, the inductor V_L , and the capacitor V_C .

Phase Shift of the Inductor

Keep the circuit in its original configuration, as in Fig. 9.5, but increase the frequency of the function generator to a value well above resonance — e.g., 10 kHz. At a sufficiently high frequency, $X_L = \omega L \rightarrow \text{big}$, and $X_C = 1/\omega C \rightarrow \text{small}$. That is, the inductor dominates the circuit, so its reactance will determine the overall phase shift ϕ that you observe in the voltage across the resistor.

To determine the phase shift, just measure the phase of the voltage across the resistor as before. It is only necessary to make one measurement in this case. The expression for the measured phase shift also remains the same as before.

Phase Shift of the Capacitor

Again, leave the configuration of the circuit alone, but decrease the frequency to well below resonance — e.g., 10 to 50 Hz.

At this low frequency, $X_L = \omega L \rightarrow \text{small}$, and $X_C = 1/\omega C \rightarrow \text{big}$. That is, the capacitor dominates the circuit, and its reactance determines the overall phase shift ϕ that you observe in the voltage across the resistor. To determine the phase shift, the procedure is the same as what you tried above.

Summary of data:

- Resonance section:
 - Capacitance C and inductance L of elements
 - V_{pp} vs frequency, for three values of resistance R
 - Resonance frequency ω_0 , for unknown L
- Phase shifts:
 - Phase difference between voltage and current: frequencies close to resonance
 - Phase shift for high frequency
 - Phase shift for low frequency

9.5 Analysis

Resonance of the RLC Circuit

1. For the three resistor values you used, plot the peak-to-peak voltages you recorded as a function of frequency. Put all three plots on the same set of axes, and normalize each to the maximum voltage from each dataset.
2. What is the resonant frequency of the circuit you observed? Report an angular frequency, ω . Be sure to record an error for your measurement.
3. Calculate what the resonant frequency should in theory be by reading the values off the circuit elements and using the formula for at the resonant frequency. Remember to propagate errors. Does your measured value agree with the theoretical value within error? What is your relative accuracy?
4. Using your plots, calculate the FWHM $\Delta\omega$ for each set of data. Be sure to estimate an uncertainty.

5. What happens to the resonant frequency when R is increased? What happens to the full width at half max?
6. For the L you measured using the large copper ring, how does the estimate compare to the value marked on the inductor?

Phase of Driving Voltage and V_R

Compare your measured values of the phase difference with the theoretical value derived for ϕ :

$$\tan \phi = \frac{\omega L - 1/\omega C}{R}$$

Remember to use the value for f recorded from the oscilloscope and convert this to an angular frequency ω . Assume that the internal resistances of the circuit components, other than the decade resistor box, are negligible. Does your measured value agree with the theoretical value, within error? If there is a large discrepancy, does accounting for the $50\ \Omega$ output impedance of the function generator help?

Phase Difference of $V(t)$ and the Capacitor/Inductor

Does the inductor voltage V_L lead the driving voltage by 90° , as expected? Does the capacitor voltage V_C lag by 90° ? If there are any discrepancies, can they be explained by errors in your measurements?

General Questions

We mentioned that the inductor resists signals with high frequencies. What property of the inductor causes it to behave this way? HINT: Think of an inductor as a current loop or solenoid, and consider what happens when you try to vary the current through the loop.

Similarly, given what you know about how capacitors charge and discharge, why do you think it tends to block DC signals but likes high frequency AC signals?

You observed the behavior of the capacitor and inductor indirectly, by looking at the signal across the resistor. Why is this necessary?

Experiment 10

Absorption of Beta and Gamma Rays

10.1 Purpose

The objectives of this experiment are: to study the behavior of beta (β) and gamma (γ) rays passing through matter; to measure the range of β -particles from a given source and hence to determine the endpoint energy of decay; and to determine the absorption coefficient in lead of the gamma radiation from a given source.

Safety Note

The radioactive sources used in this experiment are extremely weak and therefore can be handled without any danger of overexposure to radiation. However, it is always prudent not to handle any radioactive source more than is necessary, and to clean off after handling the sources and lead sheets.

10.2 Introduction

All nuclei heavier than lead (and many isotopes of lighter nuclei) have a finite probability of decaying spontaneously into another nucleus plus one or more lighter particles. This phenomenon, known as radioactivity, was first investigated systematically at the end of the nineteenth century. At the time, physicists roughly classified the decay products of radioactive nuclei into three basic categories, based on the particles' mass, charge, and ability to penetrate matter. Since the fundamental nature of these particles was not understood, they were simply called α (alpha), β (beta), and γ (gamma).

Alpha particles: One of the decay products may be a bound state of two protons and two neutrons, the stable nucleus of a helium atom. This is the α particle.

Beta particles: Alternatively, a nucleus with more neutrons than it can stably maintain may decay by emission of an electron (β -decay), which corresponds to the conversion of a neutron to a proton.

These electrons may emerge with a kinetic energy of up to several million electron-Volts (MeV).

Gamma particles: After α or β decay, the residual nucleus may be left in an excited state. (The nucleus has discrete energy levels, just like an atom as a whole.) When the nucleus decays to a lower energy state, a photon is emitted to conserve energy, as in an atomic transition. However, the typical energies of these so-called gamma ray photons are on the order of MeV, almost 10^6 times as large as the energy released during atomic transitions.

NOTE: Due to the small energies acquired by individual particles, the typical energy unit used in particle physics is the electron-Volt, or eV. It is defined as the energy acquired by a single electron that has been accelerated from rest through a potential difference of one Volt.

Details of Beta Decay

β particles, or electrons, are emitted by the decay of neutrons in radioactive nuclei. It may surprise you to discover that neutrons are not stable particles; on average, a free neutron outside a nucleus will live for about 15 minutes before it decays into a proton, electron, and an antineutrino (an extremely light, neutral particle).

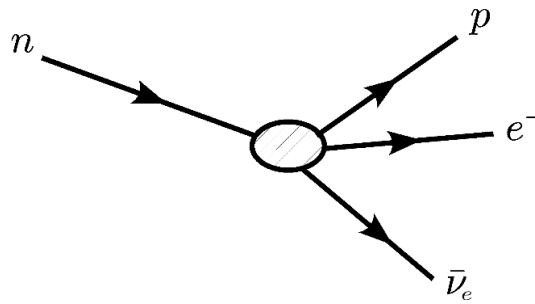


Figure 10.1: The β -decay of a neutron into a proton, electron (e^-), and electron antineutrino ($\bar{\nu}_e$).

Consider a single neutron of mass m_n . In the typical units of particle physics, $m_n = 939.5 \text{ MeV}/c^2$, where c is the speed of light. When the neutron decays, its *mass-energy* $E = m_n c^2 = 939.5 \text{ MeV}$ is used to create the decay products. Most of the energy immediately goes into creating the proton ($m_p = 938.2 \text{ MeV}/c^2$); a small bit creates the mass of the electron ($m_e = 0.511 \text{ MeV}/c^2$); and a negligible fraction creates the antineutrino ($m_{\bar{\nu}_e} \approx 0$). Whatever happens to be left over goes into the kinetic energy of the products.

In early experiments on β -decay, it was observed that the electrons do not simply emerge from the nucleus with the same energy every time (see Fig. 10.2). Rather, due to the presence of the antineutrino (which can use up the kinetic energy available in the decay), the β has a spectrum of possible energies.

These range from $m_e c^2$ — the electron is created, but has essentially no kinetic energy — to $(m_n - m_p) c^2$ — the electron soaks up nearly all of the extra kinetic energy available in the decay.

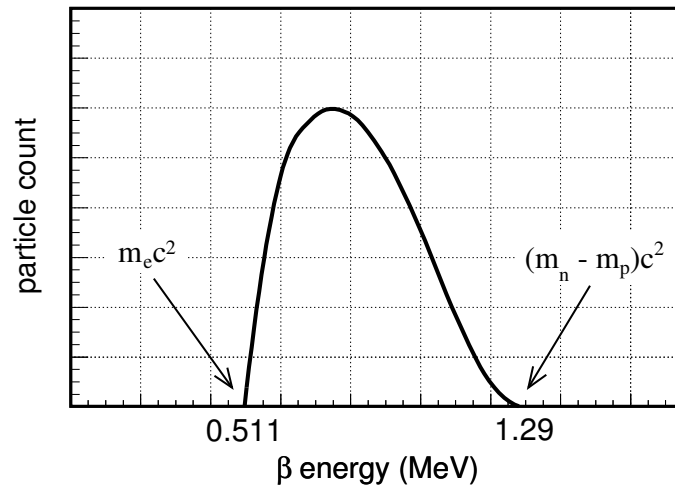


Figure 10.2: Total energy spectrum (i.e. rest plus kinetic energy) of β -particles emitted by a decaying neutron.

Interaction with Matter: Energy Loss and Range of Beta Particles

When an energetic charged particle traverses matter, it will give electrostatic impulses to nearby electrons and thus ionize atoms in the material. Because of its ionizing action, a charged, incident particle in matter will continuously lose kinetic energy, eventually coming to rest after traversing a path length called its **range**.

For a particle of known charge and mass, there will be a unique range associated with each incident energy. A formula can be theoretically deduced for the rate of energy loss — and hence the range — of a particle (of known mass, charge, and initial velocity) in a particular “stopping” material of known electron density and ionization potential.

In each interaction with atomic electrons, however, an incident electron may be scattered through a wide range of angles. As it traverses the material, it may follow a rather tortuous, winding path, especially at low energies. The actual path of the electron may be considerably longer than the observed distance that it penetrates into the material. For this reason, the incident electron range is not sharply determined, and the theoretical calculation is of limited usefulness for electrons of less than 1 MeV in energy. In this experiment, therefore, we will use an approximate empirical relationship¹ between range

¹L. Katz and A. S. Penfold, Rev. Mod. Phys. **24** (1952), 1.

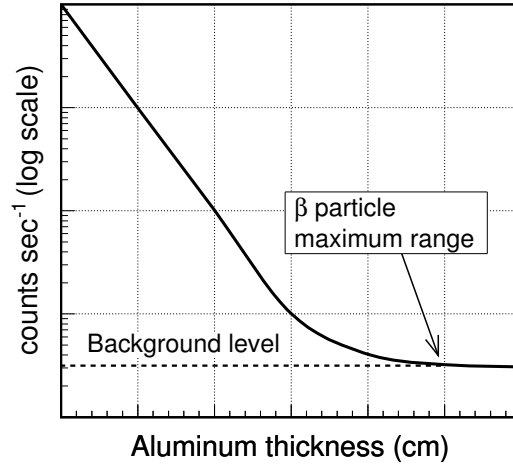


Figure 10.3: Typical absorption curve of β particles in a material.

and energy for low energy electrons:

$$r = \frac{0.412 \text{ g cm}^{-2}}{\rho} E^{1.29}, \quad (10.1)$$

where r is in cm, E is in MeV, and ρ is the density of the stopping material in g cm^{-3} .

Interaction with Matter: Absorption of Gamma Rays

Gamma rays are electromagnetic radiation, but their energies are so high — and hence their wavelengths are so short — that they effectively behave like particles. Gamma rays are electrically neutral, so they do not ionize matter in the same way that electrons do. Instead, they will interact with absorbers in one of three ways:

Compton Scattering: This refers to a photon-electron collision in which the energy lost by the scattered photon is given to the recoil electron in the material.

Photoelectric Effect: The photon is absorbed by an atom as a whole, releasing an electron with kinetic energy equal to $E_\gamma - E_b$, where E_γ is the photon energy and E_b is the relatively small binding energy of the electron in the shell from which it is released.

Pair Production: If the photon has energy greater than $1.02 \text{ MeV}/c^2$ (twice the electron rest mass), it can create an electron-positron pair in the neighborhood of a nucleus. This will not be an issue in today's experiment, since we do not use such an energetic source.

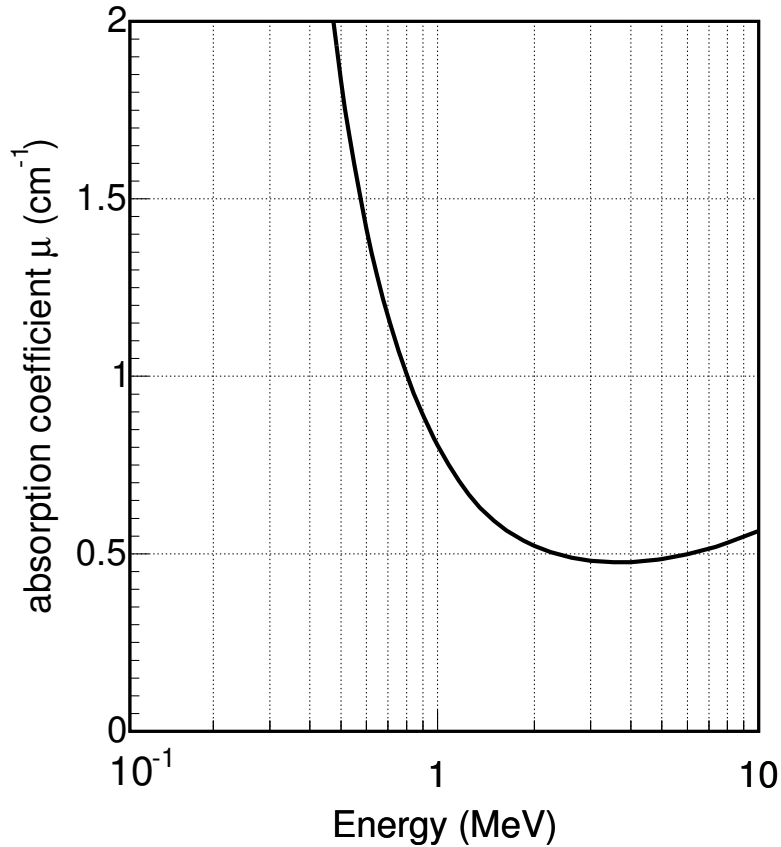


Figure 10.4: Linear absorption coefficient for γ rays in lead, as a function of energy.

The total probability that a γ particle will interact in a given block of material is the sum of the probabilities of these three processes. It also depends on the photon's energy and the atomic structure of the material. Typically, these numbers are folded into a single parameter μ , called the linear absorption coefficient of the material. This number, plotted for lead² in Fig. 10.4, physically represents the probability that a photon will be absorbed in a given unit length of the material it is moving through. The coefficient is defined by the equation

$$\frac{dN}{dx} = -\mu N, \quad \mu \text{ in cm}^{-1},$$

where N is the number of incident photons and dN is the number absorbed in an absorber of infinitesimal thickness dx (in cm). As in any process where the rate of decrease is proportional to the number present (such as the discharge of a capacitor), the solution of this differential equation is

$$N(x) = N_0 e^{-\mu x}$$

²J.H. Hubbell and S.M. Seltzer, Report NISTIR 5632 (1995).

where $N(x)$ is the number of photons passing through x cm of absorber and $N_0 = N(x)$ at $x = 0$.

A Note on Counting Particles: Accuracy and Counting Statistics

Particles decay randomly in time from a radioactive source (over a period shorter compared to the half-life). When counting particles in a detector, the probability distribution for measuring a given number of counts in a given time interval is an almost bell-shaped curve, called a **Poisson distribution**, centered around some most probable value $\mu = N_0$.

The Poisson distribution has standard deviation about the peak of $\sigma_\mu = \sqrt{N_0}$. If N counts are measured in an interval, the best estimate of the error is $\sigma_N = \sqrt{N}$. Note that the magnitude of the statistical error — your uncertainty in the measurement — increases significantly for trials involving a very small number of counts.

That is, the accuracy of a count N measured in such an experiment goes like

$$\frac{\sigma_N}{N} = \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}$$

So, for a high counting rate in a given interval — e.g., 900 counts in one minute — the estimated error will be $\sqrt{900} = 30$ counts per minute, for 3.3% error. For a much lower counting rate in the same interval — say 25 counts — an uncertainty of ± 5 counts per minute amounts to a 20% statistical error. To achieve the same precision as in the first case, you would have to collect 900 counts, which would be a 36-minute measurement at the present rate. While such a long measurement is impractical for this lab, you should aim for high accuracy by taking longer measurements when necessary.

10.3 Experiment

Detection of Charged Particles

The apparatus used here to observe β and γ particles emitted from radioactive nuclei is a Geiger counter. Like most other particle detectors, the Geiger counter uses the fact that high energy particles will ionize matter as they move through. Inside this particular device, the ionization produced by a charged particle causes a violent electrical discharge.

As shown in Fig. 10.5, the counter consists of a metal cylinder (cathode) with insulating ends supporting a fine axial wire (anode). When a charged particle enters the chamber, it ionizes the gas (e.g., argon) inside. Electrons freed during the ionization will drift toward the positively charged anode wire, accelerated more and more by the rapid increase in electric field near the wire. When an electron acquires kinetic energy greater than the ionization energy of the gas molecules, it can create by collision a new ion and electron, which in turn can accelerate and create another ion, and so on, thus initiating

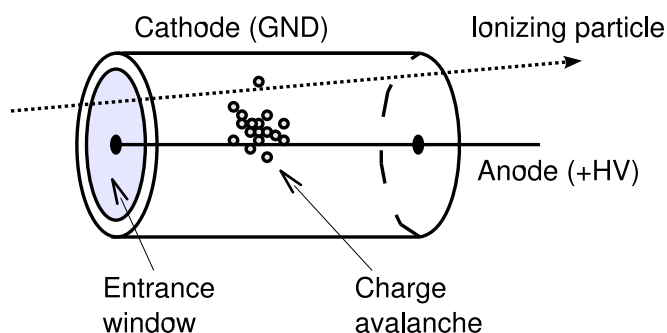


Figure 10.5: Basic workings of the Geiger counter used in this experiment.

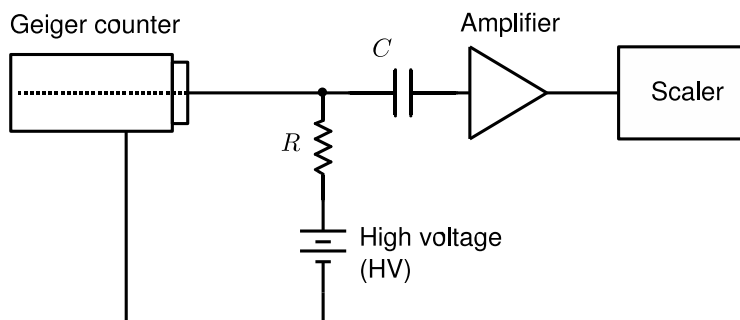


Figure 10.6: The Geiger counter circuit.

an avalanche of charge. The process, called a *Townsend avalanche*, is possible only if the voltage maintained between anode and cathode is sufficiently high.

The basic counter circuit, shown in Fig. 10.6, supplies a positive high voltage of up to 900 V to the center wire. When an avalanche occurs, current flows through R , the counter side of R drops in potential, and this negative pulse is fed through C to a stage of amplification and then to a scaling (counting) device.

10.4 Procedure

The procedure in this experiment is divided into four parts: setting up the Geiger counter; measuring the ambient radiation in the room, which you must correct for in the experiment; determining the range r of β particles; and measuring the linear absorption coefficient and energy of γ rays emitted from a radioactive source.

Setting up the Geiger Counter

Every Geiger tube that is in good working order has a plateau region in which its counting rate is relatively insensitive to changes in the high voltage (HV) supply. Before any measurements are taken, one must calibrate the Geiger tube such that the working voltage is around the plateau region. This region can be found following these steps:

- Place the source under the tube and increase the voltage from its lowest value until the tube just begins to count.
- Raise the voltage in 20 V steps while taking 15 s counts for each voltage setting. The curve of counting rate versus voltage will increase and then eventually start to level off.
- The plateau region is reached once the count rate rises by less than 10% for a 100-volt increase. Do not raise the voltage further as this may damage the tube due to the continuous breakdown.

Set the HV to a value on this plateau for the remainder of the experiment. If this procedure is followed correctly, HV variations may then be ignored as a source of error. Plot your calibration results in your report.

Background Measurement

In order to make accurate counting measurements of the sources, it is necessary to know the counting rate due to natural background radiation. This natural radiation will lead to excess counts in the Geiger counter that are not due to your sources. The background must be properly measured and subtracted out of the data if you are to obtain trustworthy results.

By and large, most of the background radiation affecting the experiment comes from cosmic rays entering the earth's atmosphere. In addition, there will be some excess counts due to the Cs-137 gamma sources nearby in the room, whose gamma rays can pass through the side of the detector. At a distance of 30 cm, for example, a Cs-137 source contributes roughly as many counts as the natural background radiation; doubling the distance reduces its contribution to one-fourth this level.

It is best to try to minimize these secondary effects by keeping your detector far from other sources, and by shielding your own Cs source with the lead sheets when not in use. Ideally, you want to keep these background effects constant during the course of the experiment. If all of your data are shifted by roughly the same constant amount, then it is possible to isolate the desired results by subtracting out this constant background.

Range of β Particles

You will use the unstable isotope Thallium-204 to generate the β particles in this experiment. The range of the most energetic of the decay electrons can be determined by placing aluminum foil absorbers

between the source and the Geiger counter.

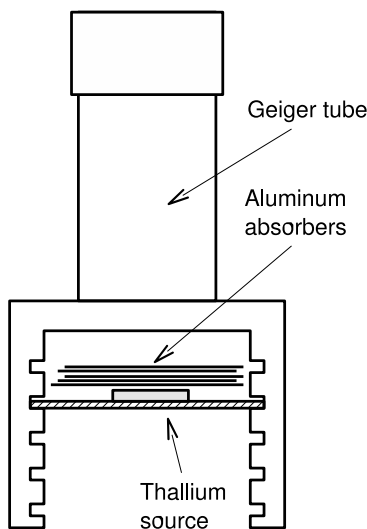


Figure 10.7: Setup of the Geiger counter, source, and absorbers for β range measurement.

Refer back to the typical absorption curve in shown in Fig. 10.3. The maximum range r is the point where the absorption curve meets the background. You should start by making a careful measurement of background, and you should repeat this measurement again after taking the absorption curve to check for constancy.

- Place the Thallium source on the second shelf below the detector, as shown in Figure 10.7, to maximize the number of counts while leaving enough room to stack aluminum absorbers.
- Begin taking measurements for the absorption curve, adding aluminum foils until the counting rate reaches the background level. Adjust the preset counting interval as needed to keep the relative uncertainties in the counts small.
- **Note:** The thickness of the aluminum absorbers is 2 mils, i.e. 0.002 inches, (not millimeters).
- Make a table of the results, including the background level, and include estimates of the statistical error in each measurement.
- Note that you should not subtract background from your data for this measurement.

Absorption of γ Rays

Your γ -ray source will be Cesium-137, which decays with a half-life of 30 years to Barium-137. As shown in Fig. 10.8, a small fraction of the time a Cs nucleus will β -decay directly to the ground state

of Barium-137. More than 90% of the time, however, the nucleus emits a 0.52 MeV β -particle and transforms into an excited state of the Ba nucleus. Almost immediately, the nucleus relaxes down to its ground state and emits a 0.662 MeV γ ray.

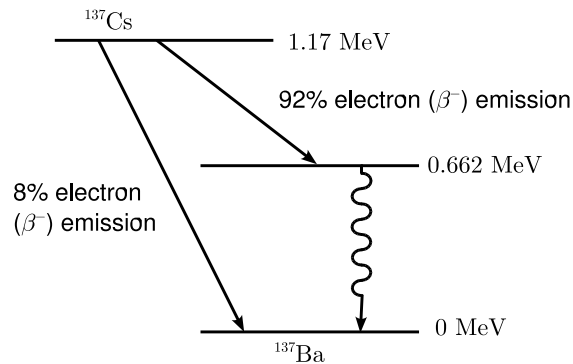


Figure 10.8: Decay modes of Cesium-137.

Lead absorbers are used for the gamma absorption study. They are thick enough (0.062 inches) so that one absorber will stop all the β particles emitted during Cs decay.

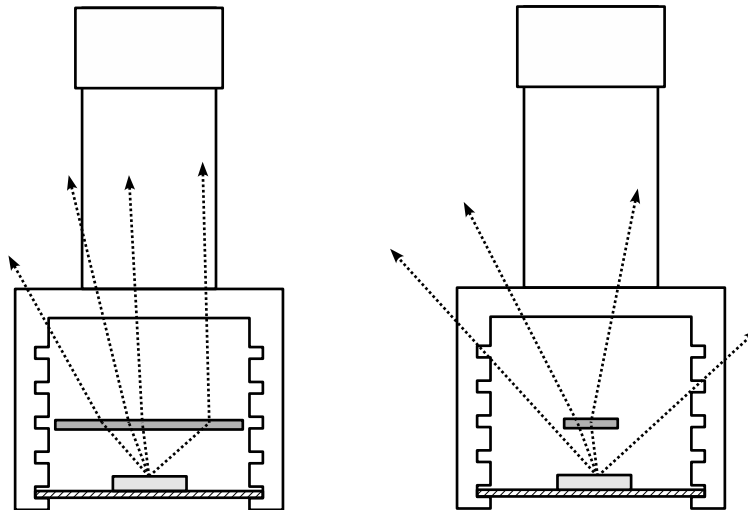


Figure 10.9: Setup of the Geiger counter for measurement of γ particles.

The gamma rays are detected by means of the same Geiger counter already used. Note that the efficiency of the Geiger counter for detecting photons is much less than for detecting the β particles, since it depends on a somewhat indirect detection scheme. For a photon to be detected inside the counter, it must collide with the gas or wall of the counter and liberate an electron to initiate the the

discharge cycle.

Figure 10.9 represents one of the effects of γ ray scattering in the lead sheets. Increasing the area of lead through which the γ rays pass tends to increase, rather than decrease, the number of counts one measures, since gamma rays which otherwise would not have entered the detector may now backscatter into it.

The arrangement shown in Fig. 10.10 is designed to reduce the backscattering effect. By keeping the lead sheets high above the source, one simultaneously reduces the excess area exposed to γ rays and the effective difference in area between the top and bottom sheets.

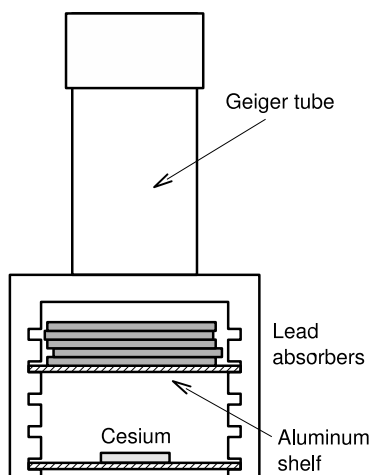


Figure 10.10: Setup of the Geiger counter for measurement of γ particles.

- Take measurements for an absorption curve. Note how this differs fundamentally from β absorption: there is no maximum range for γ rays passing through lead. Rather, one expects to lose a fixed fraction of the remaining γ rays passing through each successive layer of lead absorber. For this reason, the theoretical absorption curve never intersects the background curve. When the absorption curve is plotted on semi-log scale with background subtracted, the exponential decay curve appears as a straight line.
- Adjust the preset counting interval as needed to keep the relative uncertainties in the counts small.
- Make a table of the data and estimated errors.

You may quickly plot the background-subtracted data in the lab to check your work; otherwise, you are free to finish your work at home.

Summary of data:

- Background count rate
- Thickness of aluminum sheets, thickness of lead absorbers
- Count rate vs absorber thickness, for Tl source
- Count rate vs absorber thickness, for Cs source

10.5 Analysis

Estimate of the Range of Beta Particles

- Using your count table for the absorption of β particles by aluminum, make a semi-log plot of the counts vs. the absorber thickness. If your measurements are of different durations, you will need to scale them to a constant duration which is convenient for graphing.
- Determine the approximate value of the maximum range from the graph.
- Use eq. (10.1) to compare your result with the value of 0.765 MeV for the maximum beta energy for Thallium-204 as measured in a magnetic spectrometer. Note that the density of aluminum (Al) is $\rho = 2.702 \text{ g cm}^{-3}$.

Linear Absorption Coefficient μ

- Plot your γ ray rate counts versus lead thickness; first transform the count data by taking the ln.
- Perform a linear fit to the data, and find the absorption coefficient μ from the slope of the fit.
- Use this result and Fig. 10.4 to estimate the energy of the gamma rays. Compare with the accepted value. What factors limit the precision of this estimate?
- The efficiency of the Geiger counter is much lower for γ rays than for α or β particles. How else could you measure γ rays? What about α and β particles?

Appendix A

Measurement and Error Analysis

A.1 Measurement, True Values, and Errors¹

In a typical experiment, we are usually interested in determining the value of one or more physical quantities: the width of a block of glass, the period of a pendulum, the mass of a body, and so on. Such measurements are always subject to influences and uncertainties that will affect the observations. It is the job of the experimenter to try to minimize these effects, but it is never possible to completely eliminate them. Therefore, we need a method to quantitatively handle the errors that creep into every experiment; i.e., we need to perform a statistical analysis of our data.

Case Study in Randomness: Fairness of a Coin

Consider the following seemingly trivial experiment. A student wants to decide if a coin is fair (or “unbiased”). If the coin is fair, then when she flips it, the number of times it lands heads-up should roughly equal the number of times it lands heads-down.

The student flips the coin 50 times and counts the number of heads n . She expects that if the coin is fair, heads should come up half the time, giving her a count $n = 25$. In this trial, she actually counts 28. To check that she didn’t miscount, she decides to flip the coin another 50 times. This time, she counts $n = 19$ heads.

Intrigued, the student decides to repeat her experiment a large number of times N (e.g., $N = 100$). After finishing all 100 trials, she histograms the frequency of each head count (see Fig. A.1). The histogram indicates that over the course of 100 trials, the most common result² is 25 heads out of 50 tosses. However, there is significant spread in the data: six trials have a head count under 20; ten trials have a count greater than 30.

¹Adapted from “Introduction to Experimental Error,” Susan Cartwright, University of Sheffield, UK (2003).

²This is referred to as the *mode*.

The student also decides to calculate the average (or mean) head count \bar{n} for the 100 trials, knowing that this number is related to the “most likely” outcome of the experiment. She expects 25 heads on average, but calculates

$$\bar{n} = \frac{1}{N} \sum_{i=1}^N n_i = 24.79$$

This result is close to 25, but not exactly. The difference is small, but she must ask: is this evidence that the coin is slightly biased? Or are random fluctuations that occur from trial to trial responsible for the discrepancy? How should she quantify these fluctuations?

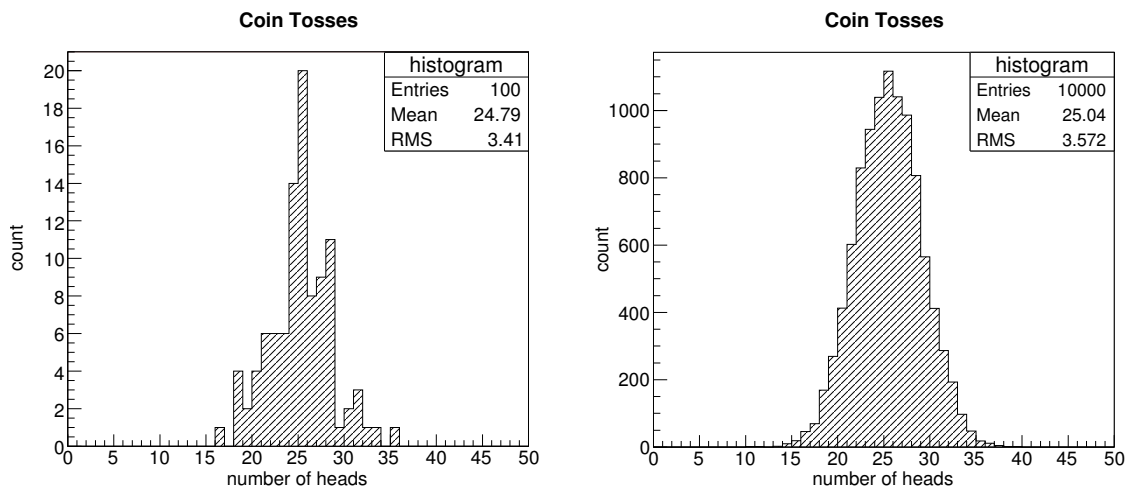


Figure A.1: LEFT: histogram of the number of head counts recorded in 100 trials when tossing a fair coin 50 times per trial. RIGHT: same histogram for 10,000 trials.

Case Study in Precision: Thickness of a Block of Glass

Random fluctuations are a source of uncertainty³ in all experimental quantities. However, these fluctuations may sometimes be too small to be observed. In such cases, we remember that physical measurements of infinite precision are impossible. Hence, we can use the *resolution* of our measurement devices to estimate the uncertainties in our observations.

For example, suppose a student is asked to measure the width of a block of glass. He uses a plastic ruler accurate to the nearest millimeter, and finds that the block appears to be 10.0 mm thick. Since he is careful, he tries the same measurement at various points along the length of the block. Each time he gets the same result: 10.0 mm.

³Uncertainty is often referred to as “experimental error.”

Does this set of observations suggest that the block is 10.0 mm thick, *exactly*? No; it only indicates that any fluctuations in the length of the block are smaller than what can be measured with a cheap ruler. In this situation, we quantify the uncertainty in the measurement as one half of the smallest division that appears on the scale of the instrument. (The rule also holds for instruments with digital readouts.) For a ruler accurate to the nearest mm, that means an uncertainty of ± 0.5 mm.

<i>Method</i>	<i>Typical Error</i>
cheap ruler	± 0.5 mm
draftsman's ruler	± 0.2 mm
calipers with vernier scale	± 0.05 mm
traveling microscope	± 0.005 mm
interferometer	± 0.00001 mm (if n known accurately)

Table A.1: Experimental error on the thickness of a glass slab.

The resolution lets us place an upper bound on the unobservable but present random fluctuations in the measured quantity. As Table A.1 indicates, we can improve the resolution by using more sensitive instruments.

Importance of Experimental Errors

The scientific approach to understanding the world is based on a number of fundamental assumptions and techniques. Two of the most important are:

- Experiments are reproducible. If you say experiment X produces result Y , I should be able to do experiment X myself and expect to get result Y .
- Theories are tested by experiment. If I assert that theory A is an improvement on theory B , I should be able to point out experimental results that are explained by A and not B , and also predict the results of new experiments yet to be performed.

Clearly, uncertainties are vital to our interpretation of experimental results. If two individuals perform an experiment, it is likely that they will not report exactly the same result due to random fluctuations. However, the relevant question is not whether their results are exactly the same, but whether they agree within the range set by experimental errors. If so, the results are said to be consistent. If not, the discrepancy may be due to an interesting physical effect that should be investigated.

Reporting Results

Hopefully you are now convinced that whenever something is measured, an uncertainty is always involved in that measurement. The complete specification of a physical quantity always includes this uncertainty,

as well the units in which the quantity was measured. For example:

1. $m = 9.0$ kg: WRONG — no uncertainty.
2. $m = 9.0 \pm 0.3$: WRONG — no units.
3. $m = 9.0 \pm 0.3$ kg: RIGHT.

A.2 Precision and Accuracy

Precision

The uncertainty (or “experimental error”) reported above is perhaps more accurately described as the *precision* of the measurement. The uncertainty reflects the range of values in which we expect to measure a physical quantity, most of the time. In other words, it is the typical scatter that we see in data when we make repeat measurements, and it is related to the reproducibility of a measurement. A precise measurement is one in which the scatter of the data is small.

Accuracy

Independent of scatter, a value is said to be accurate if it is numerically close to some “true” value. We can define the accuracy (or “relative error”) of an experimental result as

$$\text{accuracy (in \%)} = 100 \times \frac{|\text{expt} - \text{true}|}{\text{true}}$$

Note how precision and accuracy are far from the same thing. Values are precise when the scatter in the values about some mean is small. However, this does not imply that the values are close to the true value. Ideally, we want to achieve both accuracy and precision; i.e., we want to obtain results close to the true value with little scatter.

A.3 Types of Errors: Statistical and Systematic

There are two fundamentally different types of experimental error. *Statistical* errors are random in nature: repeated measurements will differ from each other and from the true value by amounts which are not individually predictable, although the average behavior over many repetitions can be predicted. For example, the spread in the number of heads n in our coin-flipping experiment, or the resolution uncertainties of an instrument, are statistical errors. There is an extensive mathematical literature dealing with statistical errors, and most of the rest of this note will be concerned with them.

Systematic errors arise from problems in the design of an experiment. They are not random, and tend to affect all measurements in some well-defined way. For example, suppose you are asked to read

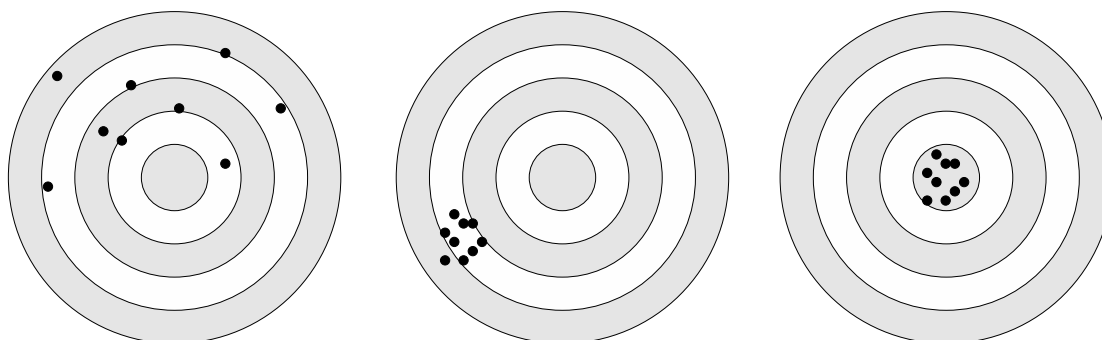


Figure A.2: Precision and accuracy visualized as a bullseye: imprecise and inaccurate results (left); precise but not accurate (center); precise and accurate (right).

the temperature from a mercury-in-glass thermometer. The position of your eye with respect to the scale on the glass introduces a parallax error — a reading error related to the relative position of the scale and the top of the mercury column.

If you consistently view the scale from below the top of the column, your temperature measurements will be consistently too low. If you consistently view the scale from above the top of the column, all of your temperature measurements will be too high. In either case, your data points will systematically shift away from the correct readings.

Systematic errors can be extremely subtle and difficult to diagnose. Repeating measurements usually won't help, but you can identify them by looking for some of the following common symptoms:

- Curved lines on a plot appear where straight lines were expected (especially on log plots);
- Nonzero values appear where zero was expected (e.g., nonzero intercepts in plots);
- Inability to reproduce results, even on the same equipment. This could indicate a dependence on ambient temperature or pressure, or on the running time of the apparatus (common with lasers).

In experiments in the undergraduate lab, systematic errors are often discovered by hindsight during the analysis phase of the experiment. Ideally, one would not want to do things this way; it is better to think about and record possible sources of systematic error *before* starting an experiment, and then attempt to eliminate these effects as measurements are conducted. In practice, this can be quite a challenge. The ability to identify systematics and neutralize their effects is a skill acquired through much practice.

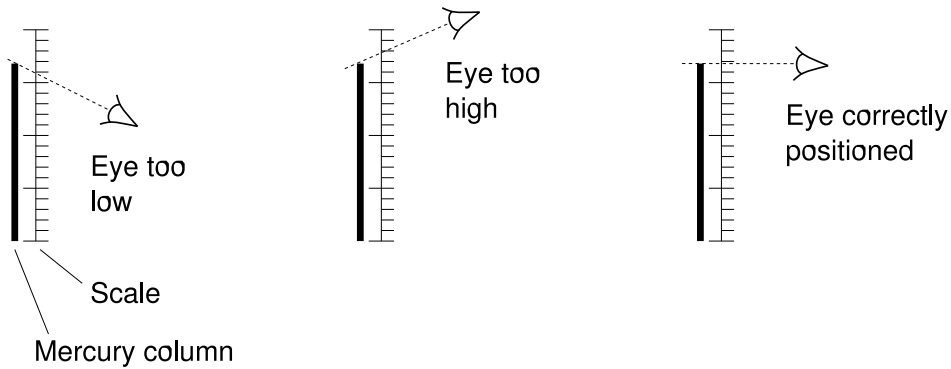


Figure A.3: Systematic errors introduced by parallax. From Les Kirkup, *Data Analysis with Excel: An Introduction for Physical Scientists*, Cambridge (2002).

A.4 Statistical Uncertainties in Measured Quantities

When we record data in the laboratory, our task is to provide a quantitative estimate of the uncertainties in our measurements. This task divides into two parts: first, we estimate the errors on directly measured quantities; second, we use these to calculate the resulting errors on derived quantities.

Basic Statistical Concepts

For single measurements, our only guide to evaluating the error is our knowledge of the experimental setup and our own capabilities. We already discussed using the resolution of an instrument to roughly estimate the uncertainty on a single measurement. This is unlikely to be very accurate, but it is still a reasonable guide to the precision of the final result.

Repeated measurements provide a more satisfactory estimate, because we can use the spread in the observed values to derive the characteristics of the underlying random error. Suppose we make N independent measurements of some quantity x . The average or mean of the N samples is the sum of the measured values divided by the total number N :

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

As mentioned earlier, the mean of a sample is an estimate of the “true” value of x . The uncertainty in this estimate is given by the standard deviation, also known as the root-mean-square (RMS) deviation s :

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}}$$

The quantity s is the best estimate of the error on an individual measurement x_i . However, our best estimate of the true value is the mean \bar{x} , not any individual measurement. The error on this is called the standard error of the mean:

$$\bar{x} \pm \frac{s}{\sqrt{N}}$$

That is, given a data sample $\{x_i\}$, we expect that the mean \bar{x} — our estimate of the true value of x — should usually fall within the range given by $\pm s/\sqrt{N}$. Note that this range decreases in size as the sample size N goes up. As we increase the number of measurements, our estimate of the mean and the precision of our result improves.

The Gaussian Distribution

It is often convenient to assume that the distribution of measurements about the true value is given by a bell-shaped curve called the Gaussian probability distribution:

$$P(x) dx = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

In this expression, μ is the true value (or mean, or “expectation value”) of the quantity we are trying to measure; σ , the standard deviation, is the typical spread expected for measurements of μ ; and $P(x) dx$ is the probability that a given measurement will fall into the interval between x and $x + dx$. In the language of statistics, we say that x is a Gaussian random variable.

For certain types of error, it can be proven mathematically that this is the appropriate formula. In such cases, μ and σ can be predicted before performing any measurements. As an example, consider again the experiment in which a student flips a coin to determine its fairness. A well-known result from elementary statistics tells us that if the coin has a probability p of landing heads up⁵, then for n tosses we expect

$$\begin{aligned}\mu &= np \\ \sigma &= \sqrt{np(1-p)}\end{aligned}$$

Look at the right panel of Fig. A.1. As the student increases the number of trials, the distribution of the head count starts to look more and more like a bell-shaped curve. In the limit as $N \rightarrow \infty$, this distribution will become a perfect Gaussian curve.

In practice, we do not measure μ and σ , because no one has time to let $N \rightarrow \infty$. Rather, we have to use a finite data sample to *estimate* their values. You probably already guessed that the best estimator for μ corresponds to the arithmetic mean \bar{x} , and the estimator for σ is the standard deviation s . Note that while s refers to the standard deviation of a finite data set and σ refers to the theoretical

⁵For a fair coin, $p = 0.5$.

spread in the data, the symbols σ and s tend to be used interchangeably. In a report, you would record your observation of μ as

$$\mu \approx \bar{x} \pm \sigma_{\bar{x}} = \bar{x} \pm \frac{\sigma}{\sqrt{N}} \approx \bar{x} \pm \frac{s}{\sqrt{N}}$$

For example, note how the estimate of μ improves in the coin toss experiment when the student increases N from 100 to 10,000:

$$\begin{aligned}\bar{x}_{100} &= 24.79 \pm \frac{3.41}{\sqrt{100}} = 24.79 \pm 0.34 \\ \bar{x}_{10000} &= 25.040 \pm \frac{3.572}{\sqrt{10000}} = 25.040 \pm 0.036\end{aligned}$$

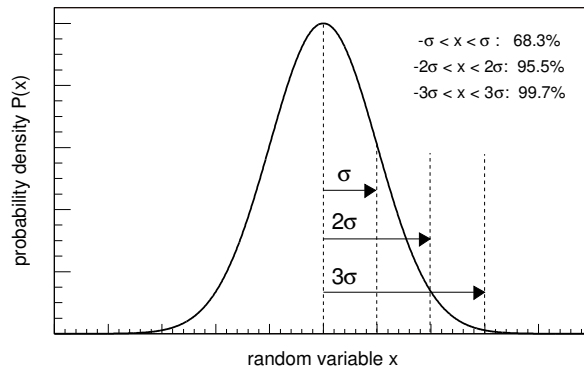


Figure A.4: Graphical representation of the Gaussian distribution. About 68.3% of the measurements of x should fall within one σ of μ ; 95.5% should fall within 2σ of μ ; and 99.7% should fall within 3σ of μ .

A.5 Errors on Derived Quantities

In most experiments, the desired quantity is not measured directly, but is calculated from other quantities which are measured. In this case we need to know how to deduce the error on the calculated result from the estimated errors on the measured values.

Error on $f(x)$ Due to an Error in x

Suppose we have measured x to be $X \pm \Delta X$. The quantity we want to determine is some function $f(x)$, and $f(X) = F$. What is the error on F , ΔF , corresponding to ΔX ?

Common sense does work; we can evaluate $f(X + \Delta X)$ and $f(X - \Delta X)$ to get $F \pm \Delta F$. However, it is often convenient to use the fact that

$$\frac{df}{dx} = \lim_{\Delta X \rightarrow 0} \frac{\Delta F}{\Delta X}$$

If ΔX is small, this gives

$$\Delta F \approx \left| \frac{df}{dx} \right|_{x=X} \Delta X,$$

or in a more standard notation,

$$\sigma_f = \left| \frac{df}{dx} \right| \sigma_x$$

We take the absolute value because (by convention) we choose ΔF and ΔX (σ_f and σ_x) to be positive.

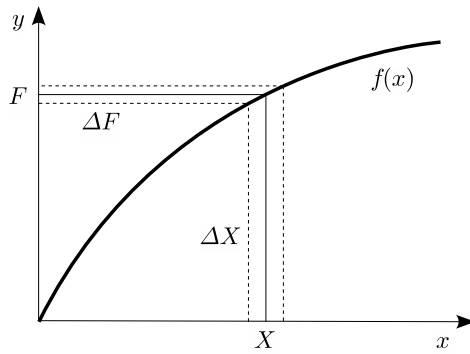


Figure A.5: Relationship between uncertainty of a function and slope.

Examples

Consider a measured quantity x with uncertainty σ_x . The resulting uncertainty in $f(x)$ is

$f(x)$	σ_f
x^2	$2x\sigma_x$
x^n	$nx^{n-1}\sigma_x$
$\sin x$	$(\cos x)\sigma_x$
$\ln x$	$\frac{1}{x}\sigma_x$

NOTE: when propagating uncertainties in angles, the angles must always be expressed in radians. Also note that the formula for the error on a logarithm applies only to natural logarithms — the derivative of $\log_{10} x$ is not $1/x$.

Functions of More than One Variable

Suppose we measure two quantities x and y , and the value that we need is a function of both, $f(x, y)$. What is the uncertainty in f given errors in x and y ? If x and y are independent, and the uncertainties are relatively small, then we add the errors *in quadrature*⁶:

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x} \sigma_x \right)^2 + \left(\frac{\partial f}{\partial y} \sigma_y \right)^2$$

As in the single variable case, all of the partial derivatives are evaluated at the measured values X and Y . Note that this equation is easily extended to the case where f is a function of N variables:

$$\sigma_f^2 = \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} \sigma_{x_i} \right)^2$$

Examples

The following formulas are derived from the expression above:

$f(x)$	σ_f^2
$x + y$	$\sigma_x^2 + \sigma_y^2$
xy	$(y\sigma_x)^2 + (x\sigma_y)^2$
x/y	$(\frac{\sigma_x}{y})^2 + (\frac{x\sigma_y}{y^2})^2$

A.6 Combining Data: Fits and Averages

We have already seen how repeated measurements of a quantity can be averaged to obtain an improved estimate of its true value. This is a simple case of combining data. We shall often meet more complex cases: for example, we want to combine two measurements 1.5 ± 0.2 and 1.70 ± 0.05 . It is intuitively obvious that the true value is likely to be between 1.5 and 1.70, but probably closer to the latter (the more precise value).

The Weighted Mean

Suppose you have measured the same quantity using several different techniques, which naturally give different errors. How do you combine the results to get the best possible estimate of the true value?

Clearly, you want to somehow weight your data so that the more precise values have more influence on the final answer. For Gaussian variables with independent errors, the appropriate weight is $1/\sigma_i^2$. In

⁶For a derivation of this expression, see Stuart L. Meyer, *Data Analysis for Scientists and Engineers*, Wiley (1975).

other words, we define the weighted mean by

$$\bar{x}_w = \frac{\sum_i x_i / \sigma_i^2}{\sum_i 1 / \sigma_i^2}$$

and the standard error of the weighted mean by

$$\sigma_{\bar{x}_w} = \left(\frac{1}{\sum_i \frac{1}{\sigma_i^2}} \right)^{1/2}$$

Both of these expressions reduce to the usual values when the $\sigma_i = \sigma$ (the errors are all the same):

$$\bar{x}_w = \frac{\sum_i x_i / \sigma^2}{\sum_i 1 / \sigma^2} = \frac{1 / \sigma^2 \sum_i x_i}{1 / \sigma^2 \sum_i (1)} = \frac{\sum_i x_i}{N} = \bar{x}$$

and

$$\sigma_{\bar{x}_w} = \left(\frac{1}{\sum_i \frac{1}{\sigma^2}} \right)^{1/2} = \left(\frac{1}{\frac{1}{\sigma^2} \sum_i (1)} \right)^{1/2} = \left(\frac{\sigma^2}{\sum_i (1)} \right)^{1/2} = \left(\frac{\sigma^2}{N} \right)^{1/2} = \frac{\sigma}{\sqrt{N}} = \sigma_{\bar{x}}$$

In the weighted average, data points with large uncertainties are guaranteed to contribute almost nothing to the overall mean. Provided that N is reasonably large, the weighted and unweighted means should be roughly the same. If they give drastically different values, it is likely that your error estimates are off.

The Unweighted Linear Least Squares Fit

A common situation is the case where two variables or suitable functions of two variables are linearly related:

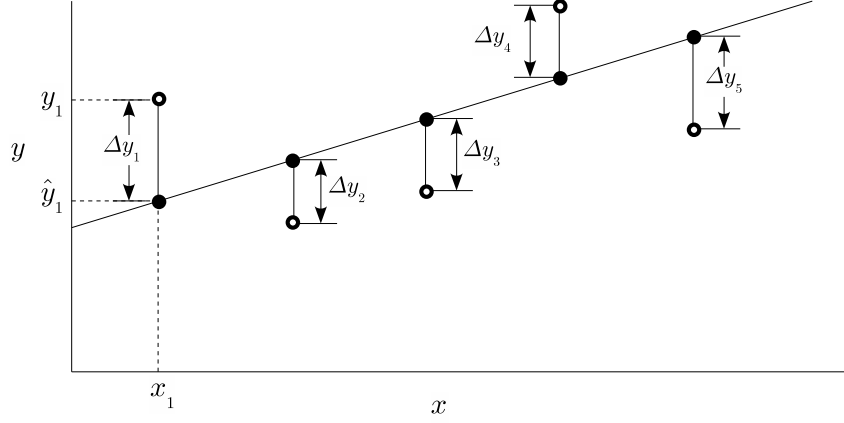
$$y = ax + b,$$

where a and b are unknown constants. Typically, the data consist of N pairs of observations (x_i, y_i) and the desired physical quantity is a or b or both. The problem is to derive the values of a and b which best fit the observations, and the corresponding uncertainties σ_a and σ_b .

To find the best line through x - y data, we need to decide on a measure of the goodness of fit of the line to the data. Fig. A.6 graphically depicts what we want to do. At each point (x_i, y_i) , there is a fitted value of y $\hat{y}_i = ax_i + b$. The best fit line should somehow minimize the differences $\Delta y_i = y_i - \hat{y}_i$ between the fitted values and the data points, known as the *residuals* of the fit.

The standard best-fit algorithm used by most statistics packages is the Method of Least Squares, which involves minimizing the quantity

$$\chi^2 = \sum_{i=1}^N \frac{(y_i - \hat{y}_i)^2}{\sigma_i^2},$$

Figure A.6: x - y plot showing residuals. From Kirkup (2002).

where σ_i is the uncertainty in the data point y_i . The expression χ^2 (chi-square) is the weighted sum of squares of the residuals. For simplicity, let's assume for now that $\sigma_i = \sigma$, i.e., all the errors are the same. To find the best fit line, we write χ^2 in terms of the fit parameters a and b and differentiate with respect to a and b :

$$\begin{aligned}\chi^2 &= \sum_i \left(\frac{y_i - ax_i - b}{\sigma} \right)^2 \\ \frac{\partial \chi^2}{\partial a} &= 0 = -\frac{2}{\sigma^2} \sum x_i (y_i - ax_i - b) \\ \frac{\partial \chi^2}{\partial b} &= 0 = -\frac{2}{\sigma^2} \sum (y_i - ax_i - b)\end{aligned}$$

Manipulating these equations to solve for a and b , we find

$$\begin{aligned}a &= \frac{N \sum x_i y_i - \sum x_i \sum y_i}{D} \\ b &= \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{D}\end{aligned}$$

where

$$D = N \sum x_i^2 - \left(\sum x_i \right)^2$$

The uncertainties in the fit parameters are tedious to derive; if all of the uncertainties are equal ($\sigma_i = \sigma$), you can show that

$$\begin{aligned}\sigma_a &= \sigma \left(\frac{N}{D} \right)^{1/2} \\ \sigma_b &= \sigma \left(\frac{\sum x_i^2}{D} \right)^{1/2}\end{aligned}$$

These expressions look intimidating, but take comfort in two things: they are actually rather easy to calculate using a spreadsheet; and most statistics packages contain functions that automatically estimate a , b , σ_a , and σ_b for you⁷.

The Weighted Linear Least Squares Fit

On a few rare occasions (to be discussed momentarily), you will want to perform a weighted least squares fit to your data that accounts for different σ_i . In this case, the expressions for the fit parameters and their standard errors change slightly:

$$a_w = \frac{\sum \frac{1}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2}}{E}$$

$$b_w = \frac{\sum \frac{x_i^2}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2}}{E}$$

and

$$\sigma_{a,w} = \left(\frac{\sum 1/\sigma_i^2}{E} \right)^{1/2}$$

$$\sigma_{b,w} = \left(\frac{\sum x_i^2/\sigma_i^2}{E} \right)^{1/2}$$

where the denominator E is given by

$$E = \sum \frac{1}{\sigma_i^2} \sum \frac{x_i^2}{\sigma_i^2} - \left(\sum \frac{x_i}{\sigma_i^2} \right)^2$$

The bad news is that few statistics packages automate these functions. The good news is that you will need to use them rarely, if ever.

Checking the Fit with the Residuals

Once you have completed the fit, a very revealing technique to test its goodness is to examine the distribution of residuals, Δy . Specifically, one usually plots the residual $\Delta y = y - \hat{y}$ on the vertical axis against x . If the fit is good and only random errors exist to obscure the true value of y at any x , then residuals should be scattered randomly about the $\Delta y = 0$ axis.

If we can discern a pattern in the residuals, then one or more factors may be affecting the data:

1. We have chosen the wrong equation to fit to the data, or a systematic error is present.
2. We should be using a weighted fit, because the uncertainties in the y values are growing or shrinking as a function of x .

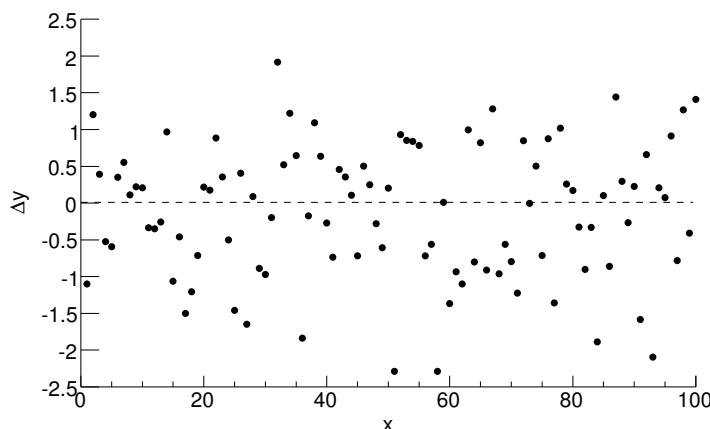


Figure A.7: Ideal distribution of residuals, with no discernable pattern. From Kirkup (2002).

Typical plots of residuals which illustrate patterns that can emerge are shown in Figs. A.7 to A.9.

Note that if the errors on y are Gaussian, you should expect to see about $2/3$ of the residuals (68%) lying on the line $\Delta y = 0$ within their 1σ error bars. If too many points lie off the line, you have underestimated your errors, or there is a systematic effect. If too few lie off the line, you have overestimated your errors, or the errors are not really all independent. If the errors are truly Gaussian, it is highly unlikely (1 in 2000 chance) that all the Δy_i should be within 1σ of zero. If this happens, there must be something wrong!

A.7 Quoting Errors

Once we have analyzed our data, we must present and correctly interpret our results. The aim here is to provide the information in a manner that is easily grasped by a reader who is not necessarily familiar with the details of the experiment.

Quoting the Numbers

Error estimates should not be quoted to unjustified precision. Usually one or two significant figures is enough: e.g., 14.2 ± 0.6 m or 14.18 ± 0.17 m are acceptable. If you are going to use these numbers in subsequent calculations, beware of round-off errors that result from throwing away too many significant figures too early. (This is typically not a problem if you do all your calculations using a spreadsheet.)

Once you have decided how many significant figures to use in your uncertainty, round the main value to the same number of decimal places. If using scientific notation, always use the same exponent for

⁷Example: the LINEST function in Excel and OpenOffice.

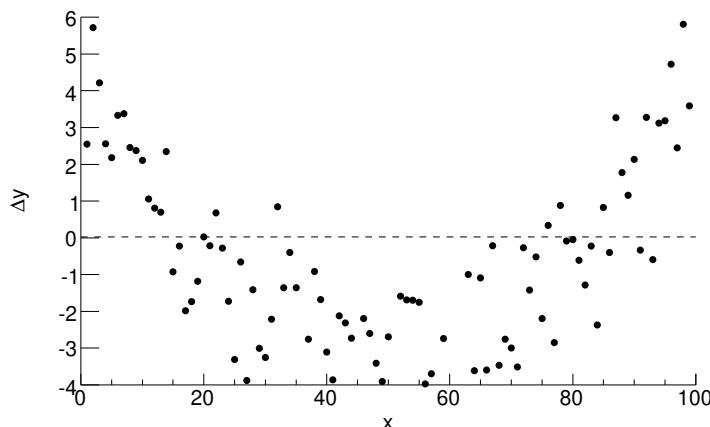


Figure A.8: Residuals revealing an incorrect equation fitted to the data. From Kirkup (2002).

the main figure and its uncertainty:

$$h = (6.62606876 \pm 0.00000052) \times 10^{-34} \text{ J s}, \quad \text{not} \quad h = 6.62606876 \times 10^{-34} \pm 5.2 \times 10^{-41} \text{ J s}$$

When reporting numbers, favor scientific notation or metric prefixes over presenting numbers with long strings of zeros. For example,

$$25 \pm 1 \mu\text{m} \quad \text{and} \quad 2.5 \pm 0.1 \times 10^{-5} \text{ m}$$

are much easier to read and understand than

$$0.000025 \pm 0.000001 \text{ m}$$

In summary:

- Unless instructed otherwise, do not quote errors to more than two significant figures.
- Quote values such that they have the same number of decimal places as the error.
- When using scientific notation, quote the value and error with the same exponent.
- When reporting a numerical value, never forget the error or the units (unless the value is dimensionless).

What to Say About Errors in Primary Quantities

As you measure quantities, try to work out errors as you go along. Do not wait until you have a final answer and then think about the uncertainties. This will often help improve your results. If, during the

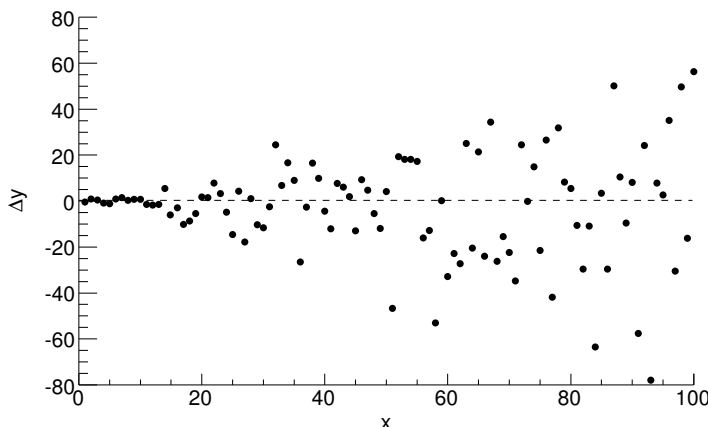


Figure A.9: Pattern in residuals suggesting weighted fit should be used. used. From Kirkup (2002).

course of the experiment, you find that some effect is contributing to the statistical errors in your data, you can take steps to mitigate the effect and improve your results.

When discussing error estimates in your lab notebook or final report, you should be as specific as possible. That is,

- Carefully state the source or sources of uncertainty, how you estimated the quoted values, and whether the resulting error is statistical or systematic.
- If the quoted value is a mean of N observations, state N , and make clear that the quoted error is the standard deviation (i.e., the error on an individual measurement) or the standard error of the mean.
- Combine different sources of statistical error by adding in quadrature: $\sigma_{\text{tot}}^2 = \sigma_1^2 + \sigma_2^2 + \dots$
- If you are able to quantify systematic errors, report them separately, e.g.,

$$1.57 \pm 0.09(\text{stat}) \pm 0.05(\text{sys}) \text{ kg}$$

What to Say About Errors in Derived Quantities

When reporting your results to the outside world, you do not need to derive or quote the basic error formulas used to propagate your uncertainties. Your intended readers will be trained scientists who expect you to do things right. While you should not go through lengthy derivations in your main report, in this course you should attach scratch work and derivations to the back of your write-up each week.

In the report itself, you need to:

1. Make sure that you have explained how the derived quantity is related to the measured quantities (i.e., give the equation).
2. Carefully explain any systematic errors and how you have dealt with them.

How to Compare Experimental and Predicted/Accepted Values

In most lab experiments, the result you obtain can be directly compared with a theoretical expectation or a “book value.” In making a comparison, it is essential to take the errors into account. The basic idea is to subtract the true value from your result, calculate the error on the difference, and observe how far the result is from zero. You can express this difference in terms of the error using the expression

$$\text{difference} = \frac{|\text{expt} - \text{true}|}{\sigma}$$

For example, suppose you measure the landing position of a projectile. You expect it to land at $x = 0$, but you find that it actually lands at $x = 1.3$ m. Deciding whether or not this is a reasonable result depends on the uncertainty of the measurement. Roughly speaking,

- If the difference between expectations and measurements is within 1σ of zero, e.g., 1.3 ± 1.4 m, the results are in *good agreement*. For Gaussian data, 68% (or two-thirds) of the random fluctuations about the true value should occur in this region.
- If the difference is between 1σ and 2σ , e.g., 1.3 ± 0.8 m, the results are *consistent*. There is about one chance in three ($\sim 100\% - 68\%$) that the discrepancy is due to a random fluctuation.
- If the difference is between 2σ and 3σ , e.g., 1.3 ± 0.5 m, there is about one chance in 20 that this is a random fluctuation away from the true value. A fluctuation this large probably (but not definitely) indicates a statistically significant fluctuation.
- If the difference is bigger than 3σ , e.g., 1.3 ± 0.3 m, then your discrepancy is almost certainly statistically significant. The chance probability of a random fluctuation this big is about 1%.

If you find that the difference between your result and the accepted value is statistically significant, do not ignore it. For example, it is not acceptable to say, “The difference between my result of 1094.2 ± 0.3 and the accepted value of 1095.5 is 1.3 ± 0.3 , but this is less than 0.1%, so my value is very close.” Remember, for a result to be truly accurate, your answer must agree with the accepted value within the precision of your measurements. In this particular case, the experimental result is 4.3σ away from the expected value, a large discrepancy. Either the error estimate is significantly off, or there is a systemic effect at work, or you have discovered some new physics!

Note that it is not acceptable to simply say, “The discrepancy is due to inaccurate measurements.” The inaccuracy of your measurements is already known from your error estimate; your job is to determine its source. Acceptable explanations for a statistically significant discrepancy might include:

- Different conditions. For example, the book value is quoted at 0° C, but you worked at room temperature.
- Dubious approximations. For instance, a theoretical prediction assumes a small angle approximation, but you carried out your measurements at 10° (0.175 rad).
- Possible systematic effects, such as a device that you could not calibrate accurately. **NOTE:** it is not sufficient for you to just say, “Our instrument was not calibrated.” You need to estimate how large the miscalibration would have to be to cause the observed discrepancy. In this way, you can determine if a suggested systematic error is reasonable.

Finally, when quoting values from a text, always specify the source. Do not write in your report, “The book value of Planck’s constant is...” Instead, you should include the source in the report, either as a footnote, as a cited reference at the end, or as an inline reference: “Planck’s constant is $(6.62606876 \pm 0.00000052) \times 10^{-34}$ J s (source: PDG, Phys. Rev. **D66** (2002) 010001).”