Student Laboratory Manual

Introduction

Laboratory classes differ from lecture classes in that you are expected to do an experiment and report your results. Good performance on an experiment requires an understanding of three main areas of knowledge: experimental technique, analysis of data, and presentation of that data. Your grade depends on each of these three areas. This Student Laboratory Manual is designed to help you complete your *physics* laboratory projects, although it may be helpful when performing experiments in the other sciences (chemistry, biology, etc.).

Philosophy of Laboratory Experiments

The experiments you will perform are designed with two goals in mind: First and foremost, they are an introduction to the experimental technique which applies to *all* sciences—not just physics. As such, the study of measurement and uncertainties are of prime importance.

Secondly, these experiments are an extension of your lecture and will help you to understand the concepts presented there. Without a basic understanding of the concepts presented during lecture, you will have a difficult time completing the projects. Sometimes you will be asked to do an experiment *before* you have had the lecture in class. This cannot always be avoided. Therefore, it pays to read the experiment well ahead of time and be prepared to do some reading in your textbook before you come to lab.

Here are a few words of advice to help you get the most out of your laboratory course:

- **NEVER FABRICATE DATA**
  Fabricating data in the sciences is equivalent to plagiarism in the humanities and will be treated in a similar fashion. Experimental science is all about discovering how the world works. Data that doesn't seem right still tells us something useful—perhaps that a piece of equipment is broken or a particular methodology is flawed. Fabricated data wastes everyone's time: it wastes time to fabricate them, and it leads your colleagues to believe things that aren't true. These are cardinal sins in science. They will not be tolerated in your place of employment and they will not be tolerated here.
- **Report broken equipment**
  If you find that your equipment is not functioning properly, please let your instructor know **immediately**. The Department is more concerned with equipment that is in proper working order than it is with blaming anyone for its malfunctioning.
- **Print out and read the laboratory hand-out well in advance of your class.**
  By printing out the hand-out as soon as possible, you can read it at your leisure. Not only will you have a chance to understand the theory behind the experiment, but you'll have some time to think through the process of taking your data. Fifteen minutes spent reading the hand-out before-hand will save you an hour or more performing the experiment.
  In addition, most laboratory instructors will *not* allow you to print out your hand-out in class. Doing so early will give you time to handle broken printers (or, worse, a broken computer).
- **Copying your lab partners' work.**
  Beware purely copying your partner's work—you'll get the same grade.
- **Penmanship**
  If your experiment requires you to turn in a write-up, **be neat.** Fill-in tables/blanks are for **final** answers. Illegible reports often earn lower grades than clearly-written ones. Use pencil if necessary.
- **Clean up your station**
  Be responsible and help keep the labs operating smoothly. Once you are through taking data, be certain to return the station to a condition where the next student can begin taking data. Return borrowed hardware to its rightful place. Not only is this being polite to your fellow students, but it shows you to be a mature scientist.

Presenting Your Results

A proper laboratory write-up is clear and concise and, when it is read, it is easy to determine what data was gathered and what it means. There are certain guidelines that you must follow in order to adhere to this standard.

Formatting Numbers

Leading zeros

If the number you are quoting is less than a whole unit, you **must** place a leading zero before the decimal point:

CORRECT: 0.876 cm
INCORRECT: .876 cm

Because the leading zero makes your data much more legible, no journal will publish your data without it.

Scientific Notation

Your calculator may report numbers such as "1.708E-19," but this is not proper format for your data. To be accepted, all numbers in scientific notation must be in "×10" notation. So the above number would be reported as

$$1.708 \times 10^{-19} \text{ kg}$$

Uncertainties

As we will discuss later, no measurement of a continuous variable (*e.g.* length, voltage, speed, etc.) is ever exact. But with care, we can estimate how uncertain we are about a measured value. There are two ways to present this data.

If we take the example of a voltage which has an average value of 3.597 V and an uncertainty of ±0.545 V, we should write this in one of two ways:

$$V = (3.6 \pm 0.6) \text{ V} \quad \text{or} \quad V = 3.6(6) \text{ V}$$

The first form is probably more familiar to you. But notice that we have rounded the average value so that it matches the number of decimal places of the uncertainty. This is important. If our uncertainty is 0.6 V, then it makes little sense to discuss a voltage difference of 0.003 V.

The second form has been gaining ground for a long time given its brevity. The n numbers in parenthases describe the uncertainty in the last n digits. This form is preferred for your labs as long as there are only one or two digits in the parenthases. However, if the equivalent "±" form is used, then this, too, is acceptable.

Units

In any experiment, it is not simply numbers that are collected, but *measurements.* We ask how large an object is, or how long it took a procedure to complete, or what the charge of an object is. But each of these measurements could have different base units. Stating that the time that it took was "3.7" is useless. Was it 3.7 seconds? 3.7 days? 3.7 clock-ticks? Therefore, you **must put units beside every number you report**. An important detail is that there should be a single *unbroken* space between the number and its unit: 15.4 m/s.

Formatting Tables

When each data point has a few attributes associated with it, a table is often the correct way to present this information. Perhaps you took some pictures of galaxies and need to present their image sizes, angular size, distance, redshift, and recessional velocity. The table below displays this information clearly.

Table 2. Galaxy Data for Determining the Hubble Constant

Galaxy	**Image Size** (cm)	**Angular Size** (mas)	**Distance** (Mpc)	**Redshift**	**Velocity** (km/s)
NGC 2903	21.4	3.2	6.9	0.002	511.9
NGC 2276	4.3	0.6	36.6	0.008	2386.1
NGC 3245	4.4	0.6	36.6	0.007	2249.0
NGC 3516	3.3	0.5	48.8	0.008	2386.1

A properly organized table lists each data point and its associated attributes on a single row. (There is only one reason to organize a table by *columns* instead of rows: if you have just a few data points and many attributes. This, however, is a very rare circumstance and—when it does occur—there are better ways to present the data.)

Title

Every table requires a title. The title tells the reader how all the entries are related. Without a proper title, your data are meaningless numbers arranged in a pleasing way.

Column Organization

The **first** column will list the names of the data points. The remaining columns each list *one* attribute for those data. The ordering of the remaining columns is somewhat fluid. However, there are two general guidelines to help you decide in what order to put them.

First, attributes which are *derived* from other attributes should *follow* (*i.e.* be placed to the right of) the attributes from which they are derived. For example, if your table lists the length, width, and area of certain rectangular objects, then length and width should come *before* area.

Second, if you are going to create a plot of two attributes (*e.g.* Distance vs. Velocity from Table 1), then the attribute plotted on the abscissa (x-axis) should come first. Notice how the table was arranged so that the smaller distances come first.

Finally, notice how **all the entries are lined up at the decimal point.** This is very helpful for legibility and is good practice.

Header

The header row is extremely important. It tells us what each column contains. Often the header row is in **bold**.

Units

Unlike the rule for stand-alone measurements, you **should not** put units beside every entry in a table. However, you **must put the units in the table header.** In Table 1, neither the "Galaxy" column nor the "Redshift" column require units. For the former, "units" make no sense. For the latter, redshift is a pure number (a ratio) and has no units.

Formatting Equations

A proper equation has three characteristics: It is centered on its own line, it is referenced by a number at the edge of the page, and each variable is explained in the following text. For example, Newton's Law of Gravitation is given by

$$F = \frac{GMm}{R^2} \tag{7}$$

where F is the force of attraction, G is the universal gravitational constant with a 2010 value of $6.67384(80) \times 10^{-11}$ N m^2/kg^2, M and m are the masses of the two objects, and R is the center-to-center distance between the two objects.

Formatting Graphs

The purpose of a graph is to display data in a visual form. Often the graph is there to help emphasize the existence of a trend or detail the nature of a trend.

To do its job, a graph must contain certain elements. A graph missing any of the following will fail to clearly convey the information it contains. Look at Figure 1 as you read this section.

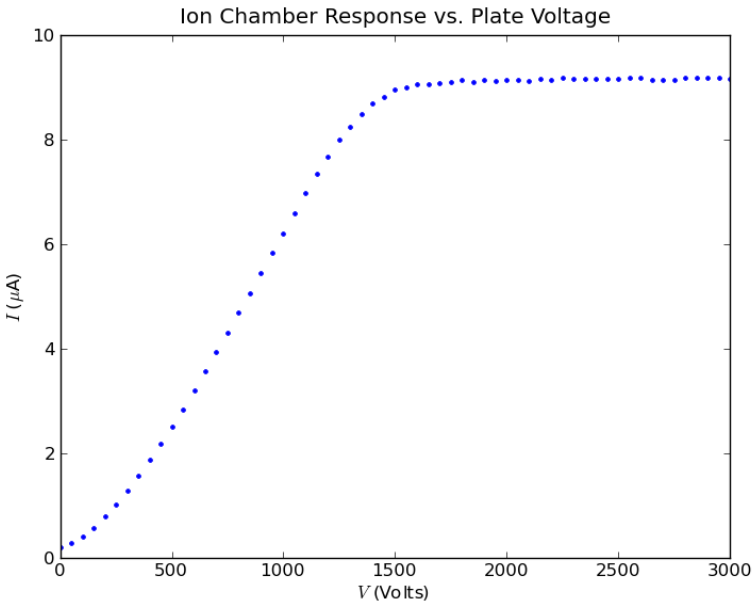


Figure 1. A Well-Formatted Graph

Title

As with tables, *every* graph must contain a title. The title conveys the setting in which the data were taken. Titles for graphs often have the structure of "Dependent Variable vs. Independent Variable." or "Dependent Variable as a function of Independent Variable." Beware: do **not** use the axes alone as your title. "Current vs. Voltage" doesn't say anything about the experiment; so the reader has no way to interpret the graph. Something just a bit more descriptive can make it quite clear: "Current Through a Resistor vs. Applied Voltage." If your graph is demonstrating a particular effect, the title might simply be the name of that effect: "Demonstration of Ohm's Law for a 570 Ω Resistor." There is no need to get too wordy: you will be explaining this graph in detail in your report.

Axes

The axes of the graph give the data points meaning. Therefore, it is of great importance that some serious thought be given to them.

• Abscissa and Ordinate

Because the "x-" and "y-" axes rarely represent positions in the $\hat{x}$ and $\hat{y}$ directions, many people use the terms **abscissa** and **ordinate**, instead. The word *abscissa* comes from the verb *"to cut"* (think "scissors") and *ordinate* comes from the verb *"to (put in) order."* Just to be perfectly clear: the abscissa is the *horizontal* axis and the ordinate is the *vertical* one.

When choosing which axis to be the abscissa and which to be the ordinate, it is almost always correct to put the *independent* variable (*i.e.* the one you control directly) on the abscissa and the *dependent* variable (*i.e.* the one that is affected when you change the independent variable) on the ordinate. **However, if the slope of the graph is meaningful, then the axes should be set so that the slope has that meaning.**

For example, say that you had measured the current through a resistor as you varied the voltage across it. At first, you might consider putting the voltage values—the *independent variable*—on the horizontal axis. But because Ohm's law says that $R = V/I$, it would make more sense to put the voltage on the *vertical* axis so that the slope of the graph would be the resistance.

• Good Use of Space

If you are creating a graph by hand, utilize the space on the graph paper so that the graph is as large as possible. In general, divide the largest data value you have on an axis by the number of squares available on that axis. Then increase your maximum data value until your answer becomes divisible by 1, 2, 2.5, or 5. That will be your maximum value on that axis. The lower your factor, the better the choice. Check both ways of plotting your data to get the best use of your space.

• Tick Marks

Tick marks should start at zero and occur at regular intervals (typically of multiples of 1, 2, 2.5, or 5 units) and their value should be labeled clearly *below* the tick marks.

A frequent mistake is to start each axis at the minimum value for the data—not at zero. At first, this mistake seems to be economical as it uses as much of the graph as possible for data points. However, this method of creating axes is called "Suppression of the Zero." A graph which has its zero suppressed can be misleading.

In general, you should *never* suppress the zero on an axis. However, it isn't a terrible choice *if:* (1) You are only interested in *relative* changes in your data **and** (2) suppression of the zero is the *only* way to make good use of the space available.

For example, say you graph two data points: a voltage of 1,000,000 volts at zero seconds and a voltage of 1,001,000 volts at 10 seconds. If you suppress your zero on the voltage axis so that it starts at 1,000,000 volts, the change of 1,000 volts looks *huge*. But it is only a change of 0.1%. This is a misleading graph.

If, on the other hand, you are only interested in an accurate measurement of the *rate of change* (100 volts per second), then it pays to zoom into the graph and the zero-suppression isn't so damaging.

• Axis Labels

In addition, the axes require labels. The labels tell us what is being measured and with what units they were measured.

• Linear Axes

The standard graph is one where both axes are linear (so that each tick mark represents the same change in value). This format is useful for most graphs. There are two other widely used formats.

• Log-Linear or "Semi-log"

When the values on the one axis span an order of magnitude (*i.e.* a factor of ten) or more, but the values on the other do not, we may plot the *logarithm* of the values on that axis rather than the values themselves. This method of plotting data should be used if details in the data can be seen which would otherwise be lost.

In Figure 2, notice that the low-frequency shoulder is completely invisible if we use a linear axis for the abscissa, which spans more than three orders of magnitude. However, in Figure 3, this shoulder is clearly visible when plotted against the logarithm of the value on the abscissa.

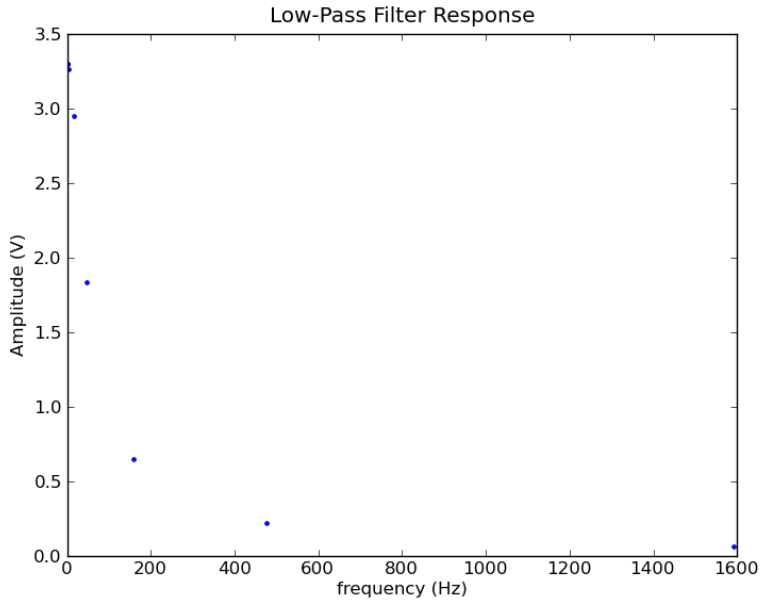


Figure 2. A Wide Range on Linear Axes

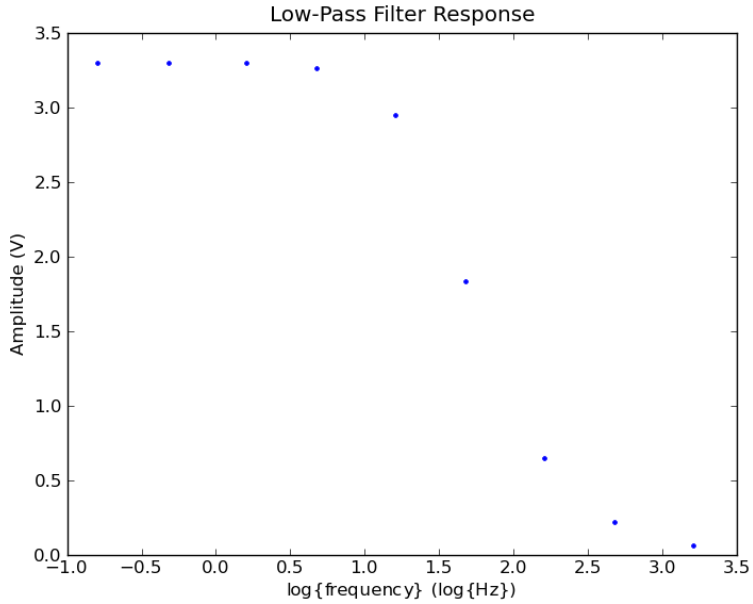


Figure 3. A Wide Range on a Logarithmic Scale

Instead of plotting the logarithm of a value on the axis, we can re-scale the axis itself. Although difficult to do by hand, most computer plotting programs offer this option. In Figure 4, the axis is no longer linear (the same distance in two different locations represents different step-sizes), but it is easier to read.

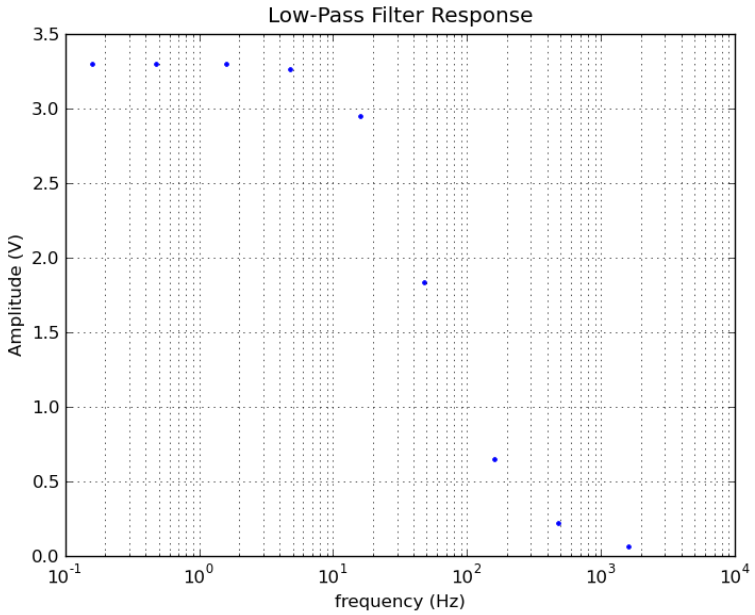


Figure 4. A Logarithmic Axis

When the ordinate has the logarithmic axis, a straight line on a semi-log plot will represent a relationship of

$$y = BM^x \tag{1}$$

Even if the values don't span a large range, we may use this set of axes if we expect this kind of relationship. We'll return to this in the section on Log-Linear or "Semi-log" (p. 19) plots, below.

• Log-Log

When *both* axes span an order of magnitude or more, we often plot the logarithm of both values rather than the values themselves. In this case, a straight line will represent a relationship of

$$y = Bx^m \tag{2}$$

Even if the values don't span a large range, we may use this set of axes if we expect this kind of relationship. Again, we'll return to this in the section on Log-Log (p. 19) plots, below.

Keep in mind that the goal of graphing your data is to *clearly illustrate* them. If a rule tells you to use one method of plotting, but that method would obscure the trend in your data, then that rule should be ignored. Here are two examples where we *should* bend the rules.

Notice that Figure 1 was *not* plotted in semi-log fashion, even though the range of voltages exceeded three orders of magnitude. This is because the data are fairly evenly spaced and there is no feature to the data which would be brought out using a semi-log plot. The odd scale would only confuse the reader.

Another case where we would bend the rules is if we did an experiment where we want to show the resistance of two devices. The experiment itself might have us vary the voltage through the device and record the current. Normally, we'd plot the voltage on the abscissa because that is the independent variable. However, resistance is given by $R = V/I$. If we plot the voltage on the *vertical* axis, then the slope of the graph would be the resistance—which is what we would want to emphasize.

Data

If you are plotting more than one set of data on the same graph, use different symbol for each set of data, and you *must* include a key or "legend" stating which symbol represents which data set.

Data Analysis

Taking data is only half the battle. Without an analysis of that data, the experiment provides nothing useful. The first concept we'll tackle is how to determine the *uncertainty* of a result. In other words, we need to come up with a measure of how much we trust a result.

Analysis of Uncertainties

Investigating the natural world requires making measurements. Even in principle, it is not possible to measure a quantity with perfect accuracy. We must learn to evaluate the uncertainty inherent in our data and report it clearly so that others can draw informed conclusions from that data. If you are unable to give at least a rough estimate of the uncertainty with which you know a certain quantity, you have done nothing useful: no journal will publish your results and no employer will pay you for them. For this reason, an important part of your laboratory work will be making simple estimates of the uncertainties in your measurements and your subsequent results.

When an experimentalist speaks of an "error," they are talking specifically about *measurement uncertainty*. They do *not* mean "mistake," such as using incorrect units or miscalculating a value. A good experimentalist carefully checks their work before reporting their results. So when we ask for you to analyze your errors, mistakes should *not* be included in your discussion.

There are three ways in which we must understand uncertainties: First, we must be able to determine the accuracy and precision of our primary measurements. Second, we must be able to determine what kinds of uncertainties might be present. Third, we must be able to deduce the uncertainty in the results derived from our measurements.

Measurement Uncertainty—Accuracy and Precision

We often take the two words "accuracy" and "precision" to mean the same thing. But they are two very different terms and it is important to understand the difference between them.

• Accuracy

The term "accuracy" describes how close our measurement is to the "true" value. It is very much a function of the method by which a measurement is made. For example, if I measure the distance between two points using my hand (the distance between the tip of my thumb and the tip of my pinkie finger is about 20 cm), then I might say that the distance between two objects is about 45 cm, but I wouldn't be certain that it wasn't 44 cm or 46 cm. If, however, I were to use a ruler, I would feel more certain about the measurement and might say that the distance was 46.7 cm. But the ruler only has tick-marks every millimeter, so I'd be hard-pressed to tell you if the real distance were 46.77 cm or 46.79 cm. Finally, if I use a micrometer which has tick-marks every 0.01 mm, I could state with more certainty that the distance was 46.782 cm (although micrometers typically don't measure distances this large).

• Precision

"Precision" is the word we use to describe how close *multiple* measurements are to each other. Any *individual* measurement will have a certain degree of accuracy depending on the method by which that measurement was taken, but a group of measurements of the same thing will be "precise" if they all lie near each other.

How accuracy and precision differ can be envisioned by imagining you are shooting many arrows into a target. The **accuracy** is how close you got to the bull's eye. The **precision** is how close together the arrows all sit.

An example of having high accuracy but low precision would be taking measurements of voltage with a noisy meter. The meter specifies voltages to three decimal places, but ten measurements might look like this:

Table 3. Multiple Measurements of the Same Circuit

Data point	Voltage (V)
1	3.454
2	3.879
3	3.123
4	5.009
5	3.501
6	3.726
7	2.891
8	3.403
9	3.310
10	3.672

• Measuring Accuracy: Significant Digits

As I became more certain of my measurement in the "Accuracy" subsection, I could add more **significant digits** to my result. The number of significant digits tells us how accurate we believe our measurement is. Let us recall the six rules for counting significant digits (in parentheses, I will give an example of a number with four significant digits):

(1) Every non-zero number is significant (*e.g.* 1234, 65.23)
(2) All zeros between significant digits are significant (*e.g.* 1024, 77.01, 120.0)
(3) Any zeros to the **left** of the left-most non-zero number are **not** significant (*e.g.* 00512.3, 0.002341)
(4) Any zeros to the **right** of a non-zero digit **and** to the right of the decimal point **are** significant (*e.g.* 0.1750, 12.00)
(5) Any zeros to the **right** of a non-zero number **that has no decimal point** are **indeterminate**
(6) All defined numbers have **infinite** significant digits

An example of rule (5) would be the number 4200. Does this number have only two significant digits (I know there are about 70 apples in a box and I have 60 boxes, so I have about 4200 apples, but I could have 4212 or 4193)? Does the number have three significant digits? Are all four significant (I hand-counted all the apples and I know there are exactly 4200 of them)?

The only way to specify the number of significant digits in this case is to use scientific notation:

$$4.2 \times 10^3 \text{ only has two significant digits}$$

$$4.20 \times 10^3 \text{ has three significant digits}$$

$$4.200 \times 10^3 \text{ has four significant digits}$$

Using scientific notation will also clear up questions with rule (3).

For rule (6), we have numbers such as $\pi \equiv 3.1415926\ldots$ and $e \equiv 2.71828\ldots$ and the speed of light is *defined* to be *exactly* 299,798,452 m/s. Also, if I have 4 oranges, that number also has infinite significant digits (I know that I **don't** have 3.999999876 or 4.0000003 oranges).

In any lab report, you will be critiqued on your use of significant figures. If you report that a driver measured their speed to be 21.3432 MPH using their speedometer, that number will come under severe scrutiny because typical speedometers aren't that accurate. A measurement of "5 m" would also be criticized if that datum was taken with a meter stick that had tick-marks every millimeter—the measurement was actually *more accurate* than reported. Always use the correct number of significant figures for your measurements.

• Measuring Precision: Standard Deviation

Sometimes it is desirable to state explicitly the limits between which we believe our value. If you were to measure your vehicle's speed with a standard speedometer, you might state it as 34.2 MPH just using significant digits and guessing between the tick-marks, or you could say (34.2 ± 0.1) MPH. Often the number, σ, following the "$\pm$" sign is called the **uncertainty, precision,** or, more colloquially, the **error bar**.

What we're saying by placing a precision value on our measurement is that we believe that, if we were to take multiple measurements of the same thing, all of our values have a high chance of falling in between $x - \sigma$ and $x + \sigma$. So in the speedometer example, we would say that the speed was certainly somewhere between 34.1 and 34.3 MPH. It *could* be between 34.0 and 34.4 MPH, but it's fairly unlikely to be as low as 33.9 MPH or as high as 34.5 MPH.

If we stick to simply using significant digits, there is an *implied* precision of one-half of the last reported digit. So a measurement of 2.34 V could be written as (2.34 ± 0.005) V and we would be fairly certain that the voltage was between 2.335 V and 2.345 V. Can we find a good way of estimating our uncertainty?

Consider a sample of data such as the measurements in Table 2. It would not be unreasonable to state that the value we found was the average of all our measurements:

$$\bar{x} \equiv \frac{1}{N}\sum_{i=1}^{N} x_i = 3.597 \text{ V}$$

What we now want is a measure of our deviation. We could start with averaging the difference between each measurement and our mean. But if you think about it, a typical set of measurements will have just as many values higher than the mean as they will lower than the mean. So if we average these differences, we expect to get near zero. Look carefully at such an average:

$$\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x}) = \frac{1}{N}\sum_{i=1}^{N} x_i - \frac{1}{N}\sum_{i=1}^{N}\bar{x} = \bar{x} - \frac{1}{N}N\bar{x} = 0$$

So this is not useful. However, if we *square* the differences, none of these values can be negative. Their average will represent the average squared distance from the mean (called the **variance** if you wanted to know):

$$\sigma^2 \equiv \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 \tag{3}$$

If we then take the square root of the answer, we arrive at the **standard deviation**:

$$\sigma \equiv \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{4}$$

For our table of data, we find:

$$\sigma = 0.545 \text{ V}$$

So we can report that our voltage was

$$(3.6 \pm 0.6) \text{ V} \quad \text{or, in more modern notation,} \quad 3.6(6) \text{ V}$$

Note carefully that it would make no sense to report a voltage of (3.597±0.545) V because our uncertainty of 0.545 V makes it clear that we really aren't sure of anything beyond the first decimal place. And maybe not even that.

Classification of Errors

We may break down errors into two broad categories: Random and Systematic.

• Random Errors

A random error is one which, every time a measurement is taken, introduces a random deviation in our result. Often the term "noise" is used to describe the agent causing the deviation. Various examples of random error are: uncontrolled vibrations in a mechanical system, a photometer which records ambient light

as well as that from the source of interest, a timer whose frequency drifts uncontrollably, a measurement which can be affected by cosmic rays—which strike relatively randomly, etc. Random errors will degrade your *precision.*

We account for random errors by taking more data. Because we expect random errors to be, well, *random,* we use more data and average them all together. Although the mean value might be representative of the true value, random errors will increase the standard deviation of a set of measurements.

• Systematic Errors

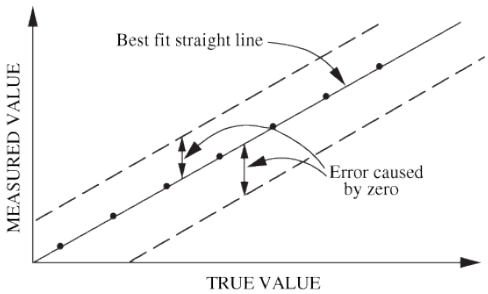
A systematic error is one which always shifts the result in a particular direction by a particular amount. It is caused by some agency that we either are unaware of or cannot remove before the measurement is made. Various examples of systematic error are: biological reaction times (ask a person to start a stopwatch and they will always be slow by about the same amount), thermal expansion of an object at a known temperature, measuring the radioactivity of an object (background radioactivity and the "dead time" of the detector will contribute a set amount to the measurement), etc. A systematic error will cause a measurement to lose *accuracy.* There is no mathematical way to account for a systematic error.

To find systematic errors, we have to be smart and look carefully at our data to determine if the results as a whole are high or low of our expected value. Another way to find systematic errors is to make the measurement by a different method and compare the results with the first method. After careful analysis, we might be able to pinpoint a systematic error and remove or correct for it.

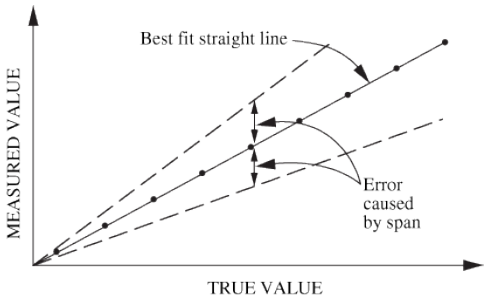
• Instrument Errors

Because we are dependent on tools to help us measure things, one very important type of systematic error is the "Instrument" or "Calibration error." Imagine stepping on a bathroom scale and finding you weigh 374 lb. Clearly, there is some sort of instrument error (we hope!) There are three main ways an instrument can fail to report an accurate value for a measurement:

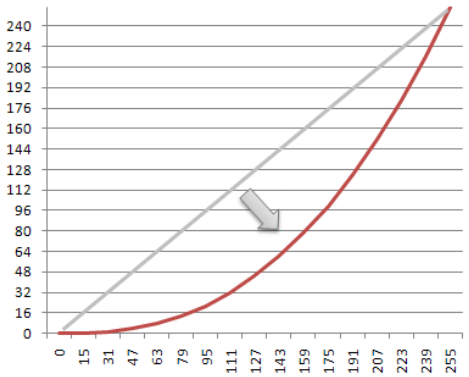
- **The Zero Error**
  One of the first errors that we can check for is called the "Zero Error." This error occurs when instrument reports a non-zero value for something which we know should be zero. For example, if nothing is on the scale, it should read 0 lb. So if the bathroom scale had a zero error of 200 lb, then it would read 200 lb with nothing on it and your true weight would be 174 lb (assuming the scale suffers from no other errors).



- **The Span (or Slope) Error**
  The Span Error is more insidious. Even if the scale reads zero when no weight is on it, it can still read too high or too low if there is a span error. In this case, increasing the input by some amount increases the output by some factor more (or less) than that amount. If your scale had a span error of two, then it's output is twice the input and your true weight is only 187 lb (again, assuming that your scale suffers from no other errors).



- **The Linearity Error**
  Finally, our instrument may suffer from Linearity Error. In this case, the output is no longer a linear function of the input. This error is actually a very common error: many sensors are non-linear. However, when they are put into measurement equipment, the device is equipped with a computer which corrects for this (and the other two) errors before reporting a measurement.



Sometimes it is hard to determine if an error is random or systematic. Consider measuring the speed of sound every day of the year. Because the speed of sound is dependent on temperature, on a hot day, your measurement will be higher than average. On a cold day, it will be lower. Each deviation can be looked at on a case-by-case basis and seen to be systematic. However, as a set of data, the error will appear random if one didn't know to take temperature into account.

• NON-Errors

There are two "errors" typically made in introductory laboratory courses: reporting something vague called "human error" and mindlessly blaming the equipment for errors. It is important to realize that the term "Human Error" does *not* represent a valid classification of a **measurement** error. This term is used by safety agencies solely to describe a mistake made by a person that causes an accident. Although it is understood that humans are capable of only so much accuracy and precision, this term will *not* be accepted in a lab write up as a valid source of error. Classify the error you are considering as either random or systematic.

Finally, it is not considered appropriate to blame uncertainties on equipment without justifying such a claim. If you think the equipment is at fault, was it a zero, span, or non-linearity error? How would you prove that this error existed? Consider these points carefully before citing the equipment as the source of your errors.

Comparing Values

One of the more common mistakes made in laboratory courses is to say something vague like "...and these two values are close..." or "...they don't differ by much..." **These statements have no meaning.** They are entirely subjective. Somehow we must find an *objective* measure for the "closeness" between two values. Depending on what we know, we have a few to choose from: percent error, ercent difference, and percent change.

• Percent Error

When we have a predicted value, often called the "theoretical" value, we can use percent error to compare our experimental value to this prediction:

$$\%\ \text{Error} = \frac{\text{Experimental value} - \text{Theoretical value}}{\text{Theoretical value}} \times 100 \tag{5}$$

Notice that there are no absolute value bars. A negative percent-error has meaning: it tells us that the experimental value was *lower* than the theoretical value.

• Percent Difference

When you **don't** know the correct theoretical value (or that value is zero) and you merely wish to compare two answers, we don't really know which answer to take as "true." So we compare the difference to the *average* of the two values:

$$\%\ \text{Difference} = \left|\frac{\text{value 1} - \text{value 2}}{\text{mean value}}\right| \times 100 \quad \text{where} \quad \text{"mean value"} = \frac{1}{2}(\text{value 1} + \text{value 2}) \tag{6}$$

Notice that, in this case, there *are* absolute value bars. This is because a negative value would only mean that value 2 was higher than value 1, but this is irrelavant because we don't know which—if either—of the values is "correct."

Pay attention to large percent errors and differences. They are a clue that you have done something incorrectly. Ignoring these clues can cost you points in your lab.

• Percent Change

Sometimes we are comparing a measurement at one time to the same measurement at some other time (or under different conditions). The best way to do this is to use the following formula:

$$\% \text{ Change} = \frac{\text{Final Value} - \text{Initial value}}{\text{Initial value}} \times 100 \tag{7}$$

Notice that if the final value is *lower* than the initial value, you will get a negative percent change. That is a useful thing to know.

Propagation of Uncertainties

Say I make measurements of the length, width, and height of a box. These measurements contain uncertainties. What, then, is the uncertainty in the volume I compute from these three measurements? For the following three variants, let's assume that I measure $x = (100 \pm 0.5)$ cm, $y = (50 \pm 0.5)$ cm, and $z = (20 \pm 0.5)$ cm.

• The Naïve Approach

In the naïve approach, we consider the case where all the numbers are at their upper limit, and then consider the case where all the numbers are at their lower limit:

$$V_{\text{high}} = (100.5 \text{ cm})(50.5 \text{ cm})(20.5 \text{ cm}) = 104042.625 \text{ cm}^3$$
$$V_{\text{low}} = (99.5 \text{ cm})(49.5 \text{ cm})(19.5 \text{ cm}) = 96042.375 \text{ cm}^3$$

We could take the volume as the expected result:

$$V = (100 \text{ cm})(50 \text{ cm})(20 \text{ cm}) = 100000 \text{ cm}^3$$

And then half the range would be

$$\frac{\delta V}{2} = \frac{V_{\text{high}} - V_{\text{low}}}{2} = 4000.125 \text{ cm}^3$$

And we then might report

$$V = (100000 \pm 4000.125) \text{ cm}^3$$

It really doesn't make sense to quote a precision with more than one or two significant figures (if you're off by 4000, does it really matter if you're off an additional 0.125?). So let's revise our answer to

$$V = (1.00 \pm 0.04) \times 10^5 \text{ cm}^3 \quad \text{or, in more modern notation,} \quad V = 1.00(4) \times 10^5 \text{ cm}^3$$

Notice that I only kept three significant digits in the result. This is because I have an uncertainty in the third digit: ± 4. Because I'm uncertain of this digit, digits to the right of this one can't be significant.

To make things prettier, I might convert to cubic meters:

$$V = (0.100 \pm 0.004) \text{ m}^3 \quad \text{or, in more modern notation,} \quad V = 0.100(4) \text{ m}^3$$

In general, there's nothing wrong with this method of estimating uncertainties in computed results. However, the method is cumbersome and tends to overestimate the uncertainty. To see how this is cumbersome, say the first value was the mass of some paint: (100 ± 0.5) g and the second two numbers were the

length and width of the part I painted. If I then wanted to know what the mass per unit area, ρ, was, the high value and low value would need to be computed as

$$\rho_{\text{high}} = \frac{100.5 \text{ g}}{(49.5 \text{ cm})(19.5 \text{ cm})} = 0.104 \text{ g/cm}^2$$
$$\rho_{\text{low}} = \frac{99.5 \text{ g}}{(50.5 \text{ cm})(20.5 \text{ cm})} = 0.0961 \text{ g/cm}^2$$

Notice how I had to use the *low* values for length for the *high* mass per unit area and vice-versa. Using the naïve method requires paying attention to this. To see how it overestimates the error, we need to look at our other two methods.

• The Significant Figures Approach

Consider the two numbers 15.87 and 12.1. The first number is uncertain in the third decimal place. The second number is uncertain in the second decimal place. We would like to add those two numbers.

Adding these two numbers requires lining up the decimal points. Let's put a question mark in the place of the uncertain digits and add them:

$$\begin{array}{r} 15.87? \\ +\ 12.1?? \\ \hline 27.9?? \end{array}$$

Notice that we *lost* information about the second decimal place. If you try this method with subtraction, you will find the same result. So we come upon our first rule for combining significant figures:

(1) **When adding or subtracting, only keep the left-most decimal place common to both numbers.**

Now say we'd like to *multiply* these two numbers. Although you may have learned a slower, more error-prone method of multiplication, we'll use the old-fashioned method of stacking. Let's put the three-significant-digit number underneath:

$$\begin{array}{r} 15.87? \\ \times\ 12.1?? \\ \hline ????? \\ ????? \\ 1587? \\ 3174? \\ 1587? \\ \hline 192?????? \end{array}$$

Our result is 192.??????. This means we only have *three* significant digits. But this is the number of significant digits of our *least*-well-known number. Try this with any two numbers and the result will be the same:

(2) **When multiplying or dividing, only keep the number of significant digits equal to the lesser of the two numbers.**

Let's now try our volume calculation from before.

First we notice that x has three significant digits and y and z both have only two. By using our second rule, we find that the answer must have only two significant digits:

$$V = (1.0 \pm 0.5) \times 10^5 \text{ cm}^3 \quad \text{or} \quad V = (0.1 \pm 0.05) \text{ m}^3$$

or, in more modern notation,

$$V = 1.0(5) \times 10^5 \text{ cm}^3 \quad \text{or} \quad 0.10(5) \text{ m}^3$$

where the precision comes from the implied precision of only having two significant digits. This method is simpler, but notice how our precision got *much* worse. Again, if we overestimate a precision, this is not a terrible thing, but it means we might not be able to say what we would like to.

• The Sophisticated Approach

We can use calculus to tell us how much a function varies if we change one of its parameters. For the following, we'll assume that we are trying to compute some value, Y, using the inputs A, B, and C. This could be the perimeter of a rectangle, in which case the perimeter, Y, is the sum of four variables:

$$Y(A, B, C, D) = A + B + C + D$$

In the case of a volume measurement, the volume, Y, would be the product of length, width, and height:

$$Y(A, B, C) = ABC$$

The uncertainties in our measured values will be given by δA, δB, and δC. The uncertainty in the computed value is, then, δY.

The details are in Appendix A, but the results are the following three rules:

If we have a function which is the sum or difference of a number of variables,

$$Y = A \pm B \pm C$$

Then we can compute the uncertainty in Y with the following formula:

$$(\delta Y)^2 = (\delta A)^2 + (\delta B)^2 + (\delta C)^2 \qquad (R1)$$

If the function is some kind of product of our three variables,

$$Y = ABC \quad \text{or} \quad Y = \frac{AB}{C} \quad \text{or} \quad Y = \frac{A}{BC} \quad \text{or} \quad Y = \frac{1}{ABC}$$

Then we can compute the uncertainty in Y via:

$$\left(\frac{\delta Y}{Y}\right)^2 = \left(\frac{\delta A}{A}\right)^2 + \left(\frac{\delta B}{B}\right)^2 + \left(\frac{\delta C}{C}\right)^2 \qquad (R2)$$

Finally, if our function is a power of one of our variables (for example, the area of a circle),

$$Y = A^\alpha$$

Then we can compute the uncertainty by

$$\frac{\delta Y}{Y} = |\alpha| \frac{\delta A}{A} \qquad (R3)$$

Let's try our volume example one last time. Because $V = xyz$, we use (R2),

$$\begin{aligned} \frac{\delta V}{V} &= \sqrt{\left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta y}{y}\right)^2 + \left(\frac{\delta z}{z}\right)^2} \\ &= \sqrt{\left(\frac{0.5}{100}\right)^2 + \left(\frac{0.5}{50}\right)^2 + \left(\frac{0.5}{20}\right)^2} \\ \frac{\delta V}{1 \times 10^5} &= 0.027386 \\ \delta V &= 2738.6 \end{aligned}$$

Therefore we may write

$$V = (1.000 \pm 0.027) \times 10^5 \text{ cm}^3 \quad \text{or} \quad V = (0.1000 \pm 0.0027) \text{ m}^3$$

Or, in more modern notation,

$$V = 1.000(27) \times 10^5 \text{ cm}^3 \quad \text{or} \quad V = 0.1000(27) \text{ m}^3$$

Although this method seems complicated, it gives the best estimate of our uncertainties. After you've used it a few times, it will become second nature.

• Combining Formulæ

What happens if we have something like the volume of a cylinder:

$$V = \pi r^2 h$$

Here we must work with both (R3) and (R2). The key is to look at the formula as the product of r^2 and h (π is exact and has no uncertainty). This requires us to use (R2). But to get δr^2, we'll need (R3).

Let's start with (R2):

$$\left(\frac{\delta V}{V}\right)^2 = \left(\frac{\delta r^2}{r^2}\right)^2 + \left(\frac{\delta h}{h}\right)^2$$

Now, (R3) tells us that

$$\frac{\delta r^2}{r^2} = 2\frac{\delta r}{r}$$

So we can make the substitution:

$$\left(\frac{\delta V}{V}\right)^2 = \left(2\frac{\delta r}{r}\right)^2 + \left(\frac{\delta h}{h}\right)^2$$

Although it is tempting to multiply through by $V^2 = (\pi r^2 h)^2$ on both sides before we compute, it's usually easier if we don't. The fractional errors ($\delta x/x$) are used in so many places that it makes sense to use those more often than not.

• Constant Factors

Say we measure the radius of a circle. If we have an uncertainty of δr in the radius measurement, what's the uncertainty in the circumference? To answer this, we use (R2):

$$\left(\frac{\delta C}{C}\right)^2 = \left(\frac{\delta(2\pi)}{2\pi}\right)^2 + \left(\frac{\delta r}{r}\right)^2$$

Now, of course, 2π has no uncertainty, so that term is zero and we have

$$\left(\frac{\delta C}{2\pi r}\right)^2 = \left(\frac{\delta r}{r}\right)^2$$
$$\frac{\delta C}{2\pi r} = \frac{\delta r}{r}$$
$$\delta C = 2\pi\, \delta r$$

So because the circumference is 2π times the radius, the uncertainty in the circumference is also 2π times the uncertainty in the radius. This is true in general. Notice the uncertainty in the average in the next example.

• The Average of Independent Measurements

Recall the voltage data in Table 2. Because we didn't know what the precision of the instrument really was (and we suspected it was much greater than the ±0.0005 V implied by the resolution of the device), we used the standard deviation of a bunch of measurements of the same circuit to determine that the *measured* uncertainty was closer to $\sigma = 0.30$ V.

But what happens when each measurement comes *with its own uncertainty?* For example, say we used different methods to measure the density of an unknown object and we had the following table:

Table 4. Multiple Measurements of the Density of an Unknown

Data point	Density (g/cm^3)	Uncertainty (g/cm^3)
1	3.5	0.01
2	3.6	0.006
3	3.5	0.004
4	3.7	0.07
5	3.5	0.06
6	3.7	0.03
7	3.4	0.04
8	3.4	0.02
9	3.5	0.015

We still can state that the average value is $\bar{\rho} = 3.53$ g/cm^3, but what is our uncertainty in this measurement? We can use (R1) and (R2) to determine what the uncertainty in the average is in the following way:

First, the sum of the values is $S = 31.8$ with an uncertainty of $\delta S = 0.11$ (calculated using (R1)). The average, of course, is the sum divided by nine:

$$A = S/9$$

So we can use (R2) to find

$$\left(\frac{\delta A}{A}\right)^2 = \left(\frac{\delta S}{S}\right)^2$$

$$\delta A = \delta S \left(\frac{A}{S}\right)$$

$$\delta A = \delta S \left(\frac{S/9}{S}\right) = \frac{\delta S}{9}$$

and our average becomes

$$\bar{\rho} = (3.530 \pm 0.012) \text{ g/cm}^3 \quad \text{or} \quad 3.530(12) \text{ g/cm}^3$$

Before we leave this topic, it's important to emphasize that it is *always* better to take as many measurements as your time constraints allow. To see this, let's say we take N measurements of something. Each measurement has *the same* uncertainty: δx. We calculate the average value in the normal way:

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

where S is the sum of the measurements. The uncertainty, δS, is given by (R1):

$$\delta \bar{x} = \frac{1}{N}\delta S = \frac{1}{N}\sqrt{\sum_{i=1}^{N} (\delta x)^2} = \frac{1}{N}\sqrt{N(\delta x)^2} = \frac{1}{\sqrt{N}}\delta x$$

So the uncertainty of four measurements is one-half the uncertainty of a single measurment. And the uncertainty of one hundred measurements is 10% that of a single measurement.

Graphical Analysis of Data

The old cliché states that the picture is worth a thousand words. And often enough, if we plot our data, we may elucidate trends that were not obvious in any other way.

The easiest trend to notice on a graph is when the data fall on a straight line. We then have two tasks before us: we must be able to determine the slope and y-intercept of that line.

Plotting Straight Lines

Even if you have a computer with data analysis software, it is important to know how to draw the best-fit line to your data by hand. This way, you can make a first-order guess at what the answer *should* be so that you can verify that the computer gave you the correct answer.

The equation for a straight line is exactly the equation your high-school algebra teacher taught you:

$$y = mx + b \tag{8}$$

The *best* straight line for your data points is the line for which the slope and intercept give the best prediction at every x_i. In other words, for your choice of m and b, the prediction

$$y(x_i) = mx_i + b$$

is as close to each actual data point, (x_i, y_i) as possible.

Most likely, your best-fit line will go through *none* of your data points. You will have to be judicious in drawing this line. It will get close to some data points and be further from others. To know what the *best* line is will require some calculus. But you should be able to draw a fairly good line by eye—unless the data points are very scattered. Appendix C describes how to compute the line of best fit to your data.

Slope and Intercept

• The Slope

The slope of a straight line is the variable, m, in equation (8). The value of that slope was also given by your high-school teacher. If the line starts at point (x_1, y_1) and ends at point (x_2, y_2)

$$m = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} \tag{9}$$

Therefore, after you've drawn your best-fit line, pick two points on the line and compute the slope using equation (9). But which two points should you choose?

First and foremost: **do not use any data points**. Even if your line appears to go through a data marker, that does *not* mean it goes through that exact data point and you have no justification to use those numbers as a point on your line. The points you pick *must* be points clearly found on your line.

How far apart should your points be? To answer this question, let's look at the uncertainty in the calculated value for your slope. Using (R2),

$$\left(\frac{\delta m}{m}\right)^2 = \left(\frac{\delta(\Delta y)}{\Delta y}\right)^2 + \left(\frac{\delta(\Delta x)}{\Delta x}\right)^2$$

We can make the left-hand side smallest by making Δx and Δy, the distance between our measurement points, as *large* as possible. Therefore, the two points you use should be as far apart as possible if we want the smallest uncertainty in our slope calculation.

• Linear Axes

The standard graph is one where both axes are linear (so that each tick mark represents the same change in value). This format is useful for most graphs. If your data fall on (or near) a line, then you can discuss the closeness with which your data follow the trend

$$y = mx + b$$

If the data already fall on a straight line, then our plotting job is done. If not, there are two common plotting tricks to make non-linear data fall along a straight line.

• Log-Linear or "Semi-log"

If we plot the *logarithm* of the dependent variable (so, $\log y$) against the independent variable, x, it is possible to get a straight line. Let's consider the equation of such a line and see what it tells us about the relationship between the values, themselves. Using the equation for a line:

$$\begin{aligned} \log y &= mx + b \\ y &= 10^{mx+b} \\ y &= (10^b)(10^m)^x \\ y &= BM^x \end{aligned}$$

Therefore, if we expect that the y values depend on x as an *exponent*, we might prefer to plot these data in semi-log fashion. A good example of this would be the current, I, entering a capacitor, C, charging through a resistor, R:

$$I(t) = \frac{\varepsilon}{R} e^{-t/RC}$$

In this case, we'd expect the intercept, b, to be the logarithm of ε/R. The slope, m, would take a bit of work but we find that:

$$\begin{aligned} M^t &= e^{-t/RC} \\ t \log M &= -\frac{t}{RC} \log e \\ \log M &= -\frac{\log e}{RC} \\ m &= -\frac{\log e}{RC} \end{aligned}$$

• Log-Log

If, instead, we plot the logarithms of *both* variables against each other, we find the following relationship *if* we then get a straight line when we do so:

$$\begin{aligned} \log y &= m(\log x) + b \\ y &= 10^{m \log x + b} \\ y &= 10^b 10^{m \log x} \\ y &= B \left(10^{\log x}\right)^m \\ y &= Bx^m \end{aligned}$$

So, if we expect that the y values depend on some *power* of the x values, we might choose to plot our values in log-log format. A perfect example of this might be the force between two charged objects as a function of separation, x:

$$F = \frac{q_1 q_2}{4\pi\epsilon_0} \frac{1}{x^2}$$

In this case, we would expect the slope of the line to be $m = -2$ and the intercept, b, to be the logarithm of $q_1 q_2/(4\pi\epsilon_0)$.

Laboratory Reports

The purpose of a laboratory report is to communicate the results of an experiment you performed to your colleagues. Because you might be reporting to people who are unfamiliar with the experiment (perhaps they are reading your report many years later), the report should be self-contained: it must state clearly what the experiment was about and how it was conducted. You cannot expect the reader to have read the laboratory hand-out (indeed, in a real experiment, there is no laboratory hand-out).

The laboratory report is to scientists what an essay is to those in the fine arts. Therefore, good English and clear thinking are just as important in your report as they are in an essay. However, a laboratory report is *not* to be written like a story.

No portion of your laboratory report is to be copied verbatim from the hand-out. Use your own words and experiences from having performed the experiment, yourself.

The Structure of a Laboratory Report

Below is the outline for a laboratory report and what each section should contain. Please use this template when you write your laboratory reports.

I. **Introduction**

The introduction, as the name suggests, introduces the idea of the experiment. The introduction is a few paragraphs which should contain three pieces of information:

A. **Rationale**

In the real world, you would explain why this experiment was necessary. Clearly, in this course, the reason is that it was assigned for you to perform. You can probably skip this subsection.

B. **Goal**

What, *specifically*, are you attempting to discover via this experiment?

C. **Methodology**

In *general* (you'll describe the specifics later), how was the experiment configured so that you could determine the values you need to prove your goal? Why is this the proper method for doing so?

II. **Theory**

The theory section provides the details of how the method you chose will yield the results you are interested in discovering. In this section, you should provide your logic—including graphs and equations where necessary. This theory then provides a prediction for the results you expect to obtain. Derive this theoretical prediction here.

In addition, you must provide the mathematical formulæ that you will use to compute your uncertanties. Most if not all formulæ belong in this section. Please read about how to format formulæ above.

III. **Experimental Procedure**

The procedure section provides the details of the experiment itself. What you briefly described in the introduction must now be made clear enough so that your reader could, if desired, repeat your experiment. This section should also include details about any modifications you had to make to get the experiment working. You may *reference* sections of your laboratory manual for brevity.

IV. **Data/Results**

In this section, you present your *final* results. **This section is not for raw data.** Place any raw data (such as asked for in your Raw Data Sheet) in the Appendix.

A. **Main Data Table or Graph**

Display the results of your experiment as clearly as possible including error bars.

B. **Main Result**

Your main result for these experiments is typically one or two numbers. These numbers are derived from the data you presented in your main data table above. Clearly derive these results using the theory you described in the Theory section and report the mean values *plus* their standard deviations.

C. **Error**

Your main result will be different from your theoretical prediction. Report that error here.

V. **Discussion and Conclusions**

In this section, you sum up the experiment focusing on your results. Did you expect the results you obtained? Why was your error as high (or low) as it was? Why was your standard deviation as high (or low) as it was? What systematic and random errors do you think affected your results the most? What would have made the experiment more precise or easier to perform? Do your results connect to other results or theories? Does this experiment suggest a follow-up experiment you could perform?

VI. **Appendix**

This section contains the ancillary material which may be of interest but does not directly support the conclusions of the report.

A. **Raw Data Sheet**

This is where you put your signed raw data sheet.

B. **Sample Calculations**

Because these lab reports are as much practice using new algorithms as they are practice in communicating results, you need to show a sample computation for *each* type of value you compute.

How to Write A Laboratory Report

Writing a laboratory report takes some effort, but if you approach it correctly, you will know what needs to be written in each section. Here is one method for writing a report that will cover all the points you need to discuss for an experiment.

(1) **Write the Discussion and Conclusions section first.**
The Discussion and Conclusions section is *the most important* section of your report. In it, you discuss your results and compare those results to your theoretical predictions. Once you are satisfied with this section, you need to support it with the data you took and the theoretical predictions you made.

(2) **Write the Theory section next.**
Now that you know what calculations you needed to make to state your conclusion, report the theory that supports all the computations you made. Do not add in anything else unless needed to support something else in this section.

(3) **Insert any tables, graphs, or final values into your Data section.**
The conclusions you made in your Discussion and Conclusions section requires the data you took. Report all the data that you mentioned in your Discussion and Conclusions section in a clear format so that you can simply refer to that data instead of repeating it.

(4) **Write your Experimental Procedure section.**
Now that you know what the important experimental values are, write your Experimental Procedure section so it is clear how you were able to measure those values.

(5) **Write your Introduction last.**
Only after you have written the rest of your report and have the whole experiment in your mind should you write the Introduction section. This section should be a one- or two-paragraph overview of the experiment you performed—possibly stating your main conclusion briefly.

Common Laboratory Report Mistakes

The purpose of a laboratory report is to communicate the results of an experiment you performed to your colleagues. Because you might be reporting to people who are unfamiliar with the experiment (perhaps they are reading your report many years later), the report should be self-contained: it must state clearly what the experiment was about and how it was conducted. You cannot expect the reader to have read the laboratory hand-out (indeed, in a real experiment, there is no laboratory hand-out).

The laboratory report is to scientists what an essay is to those in the humanities. Therefore, good English and clear thinking are just as important in your report as they are in an essay. However, a laboratory report is *not* to be written like a story.

Below is a list of common mistakes made in physics lab reports. The more of these errors that you can avoid, the higher your score will be. The errors in the first two sections will be graded more strictly than those lower on the list.

Missing Information

Your Laboratory Manual details the structure and reasoning behind each section in your lab report. Although your first few lab reports will be partial reports, a full lab report must have *all* the sections listed above. Each of these five sections is necessary to fully describe your experiment because they explain: what your experiment was about, how you collected your data, how you interpreted it, what your final results were, and how those results match your expectations. Missing any of this information or putting it in an unexpected location makes your report less useful to the reader. Common errors in these five sections include:

M0. **No Introduction**
When asked for, you need to give a *brief* overview of your experiment. You cannot expect the reader to know what your experiment was about. Do *not* put experimental or theoretical details in this section. It is meant to be short and give an overview of your experiment.

M1. **Missing Equations**
When asked for, the Theory section is for explaining how you will take your raw data and compute your final data from it. Often an equation is missing. The most common omission is the method for computation of uncertainties.

M2. **Missing Methods**
A particular method for calibrating a device or obtaining data is missing.

M3. **Missing Data**
Your data section is for your final data. You must present *all* of the results of your experiment.

M4. **Missing Comparisons**
The Discussion and Conclusion section must compare your data to your expected results. Those comparisons must be numerical.

Unscientific Arguments

U1. **Unsubstantiated Claims**
If you make a claim in your report, you *must* back it up with evidence. This typically happens when one is trying to explain their differences between a theoretical result and the results of their experiment. A common claim is that the equipment is broken in some unspecified way. If you think the equipment is not functioning properly, you must provide evidence for this claim. What tests did you perform to verify that the equipment is faulty?

A very closely related claim is one that tries to blame a large difference on a very small effect. "People walking down the hallway," or "Likely due to the air conditioning" have been mentioned with no supporting evidence. These explanations will not be accepted without a valid rationale for the size of the purported effect.

U2. **Non-Numerical Comparisons**
When comparing the values of two measurements, one must use *numerical* comparisons such as percent error or percent difference. Without this numerical comparison, "weasel words" such as "about," "almost," "close to," "kind of," "roughly," or "sort of" have no meaning. These weasel words are subjective and do not constitute a scientific argument.

U3. **Lack of Uncertainty**
No measurement is ever exact. Every result carries an uncertainty. If you cannot estimate the uncertainty of a value you determined, it is not a useful value and no journal will publish your results. You must state an uncertainty with every value you report.

U4. **"Human Error"**
As has been mentioned in many other places, the term "Human Error" is allowed in only one place: It is allowed when you are making a report on the cause of an accident. The term "Human Error" is

too vague to be used in a scientific report when one is discussing sources of uncertainty and will not be allowed in your report.

U5. **Classifying Mistakes as Uncertainties**
Often, students will assert their uncertainties stem from not reading the lab, making mistakes in calculations, not using the tools properly, and similar errors in sceintific practice. These are *not* sources of uncertainty. They are *mistakes and blunders.* No good scientist uses them as excuses for a poorly performed experiment. Good scientists are knowledgeble, careful, fix their mistakes, and report good data.

Formatting Errors

F0. **Formatting of Data**
Most of the formatting issues are covered in your Laboratory Manual. Please read it carefully and follow the instructions for formatting numbers, units, graphs, and tables. You will be graded according to how well you follow those formatting guidelines.

F1. **Procedure**
The procedure should be a brief description of how *you* performed the experiment. **Do not write instructions to the reader.** Write clearly enough that your experiment could be performed by the reader. Each paragraph should focus on what needed to be done to perform your experiment and precisely **how** you did it. You should be thourough, but do not dive into minute details unless they are critical to the experiment. It is also not necessary to describe precisely how any data collection software was run (*e.g.* "And then I clicked on the 'stop' button.").

F2. **Results**
This section is *only* for presentation of those results which support your conclusions. Typically you need only report one or two numbers, a table, and/or a graph **which are clearly labeled**.

Do not report raw data. Any explanations of the data belong in the appropriate section of the lab report (usually the "Conclusions" section).

F3. **Conclusion**
This is possibly the most important part of your lab report. Many students do not appropriately compare their results to their theory section and try to explain their discrepancies.

If the details of a calculation are important, show the equation (formatted as outlined in your Lab Manual) and then simply quote the numerical result. Do *not* write any calculations in-line with your text. **Sample calculations belong in the Appendix.**

English Errors

E1. **Use Paragraphs**
Paragraphs break your exposition into manageable "chunks." Without them, any points you may try to make will be lost in the "word salad" and you won't get credit for them. Every paragraph should start with an introductory sentence that states the main point of the paragraph. The following sentences support the introductory sentence, and the last two sentences sum up the paragraph and transition to (introduce) the next paragraph.

E2. **Be Succinct, but not Terse**
Lab reports summarize your experiment. They should *not* be long-winded expositions. However, you should be detailed enough to clearly explain your methods and conclusions. Do not "pad" your report with unnecessary or redundant statements. Always ask the question "does this sentence have information that is not already covered in the rest of the report?" If not, omit it.

E3. **Past Tense, First Person, and Passive Voice**
When you make a report of your experiment, it has already happened. Every sentence in your report *must* be in the **past** tense. The only exceptions are when you back up claims with on-going events or make future predictions in your conclusions.

Although this is often debated, in this lab it is OK to use the first person. When you say "I measured the volume using vernier calipers," this is a (hopefully) truthful statement. More correctly, you are working with a lab partner (or two) and you might prefer the word "We."

Many people will avoid using the first person by using passive voice. Instead of writing "I measured the volume using vernier calipers," the subject of the sentence becomes what was measured: "The volume of the cylinder was measured using vernier calipers." This is typically the preferred sentence structure for a lab report, but try not to over-use it as it makes the report sound mechanical.

E4. **Commas, Run-on Sentences, and Homonyms**

First and foremost, learn the correct placement of commas. The informal rule is to place a comma wherever you might pause when speaking—but this is not always correct usage. Commas belong immediately after introductory clauses: "Because of this, we..." or "Next, we..." Commas also set off independent clauses: "When we measured the volume, using the calipers supplied to us, we made measurements at three locations." Finally, commas *always* precede the word "and" in lists: "We measured three cylinders: the blue one, the green one, and the red one."

A run-on sentence is a sentence that is built of smaller sentences "pasted together" with commas and the word "and": "We placed the cylinder in the loop of string and tied it tightly to keep it from touching the bottom of the beaker, and then we lowered it slowly into the water while watching the value displayed on the scale, trying to pay attention to the significant digits." Can you see how this one sentence could be broken into three to make it more readable?

Although even your web-browser can check your spelling, it cannot tell you if you've mis-used a word that is spelled correctly. Learn the difference between "its" and "it's" (hint: the first is a possessive, the second is a contraction). "There," "they're," and "their" are also commonly confused.

E5. **Unprofessional Writing**

Your report is a formal document. Most of us use informal wording when we speak, but you are *not* allowed to use it in your report. Write your lab report in a professional manner. This means avoiding "filler" words such as "Well," or "Obviously," or "Seriously." Also avoid all slang, colloquialisms, and clichés.

Appendix A: Derivation of the Sophisticated Approach

Calculus tells us that the change in the value of a function $f(x_1, x_2, \ldots, x_n)$ when we change its parameters, x_j, is given by the relation

$$\Delta y = \frac{\partial f}{\partial x_1}\Delta x_1 + \frac{\partial f}{\partial x_2}\Delta x_2 + \cdots + \frac{\partial f}{\partial x_n}\Delta x_n \tag{10}$$

Most books will simply state that the Δx_j are our uncertainties, δx_j, in our measurements and that we should consider two cases: The worst case is if we got all of our measurements wrong and in the same direction. Then all the Δx_j are at their maximum values and the uncertainty in our function, Δy, is precisely the sum of each of the terms in equation (10). The best case is that each of our uncertainties cancels with all the others so that $\Delta y = 0$. Because we expect to be somewhere in between, we choose to add all the terms "in quadrature." This will give us a good estimate:

$$\delta y = \sqrt{\left(\frac{\partial f}{\partial x_1}\delta x_1\right)^2 + \left(\frac{\partial f}{\partial x_2}\delta x_2\right)^2 + \cdots + \left(\frac{\partial f}{\partial x_n}\delta x_n\right)^2} \tag{11}$$

However, we can make this argument much stronger.

The standard deviation of y, σ_y, is given by the square root of the variation of y:

$$\sigma_y^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2$$

We're interested in computing this variation. We, therefore, need to construct this sum.

The mean value, $\bar{y}$ of a function of n variables, $y = f(x_1, x_2, \ldots, x_n)$, is given by the value of the function at the mean value of each of its variables: $\bar{y} = f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$.

For a particular set of measurements, $x_j^{(i)}$, the value of the function changes to $y_i = f(x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)})$ for the ith set of values.

Equation (10) then tells us that

$$y_i - \bar{y} = \frac{\partial f}{\partial x_1}(x_1^{(i)} - \bar{x}_1) + \frac{\partial f}{\partial x_2}(x_2^{(i)} - \bar{x}_2) + \cdots + \frac{\partial f}{\partial x_n}(x_n^{(i)} - \bar{x}_n)$$

We can now calculate what the variation is:

$$\begin{aligned}
\sigma_y^2 &= \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left[\frac{\partial f}{\partial x_1}(x_1^{(i)} - \bar{x}_1) + \frac{\partial f}{\partial x_2}(x_2^{(i)} - \bar{x}_2) + \cdots + \frac{\partial f}{\partial x_n}(x_n^{(i)} - \bar{x}_n)\right]^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left\{\left[\frac{\partial f}{\partial x_1}(x_1^{(i)} - \bar{x}_1)\right]^2 + \left[\frac{\partial f}{\partial x_2}(x_2^{(i)} - \bar{x}_2)\right]^2 + \cdots + \left[\frac{\partial f}{\partial x_n}(x_n^{(i)} - \bar{x}_n)\right]^2\right. \\
&\qquad \left. + 2\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}(x_1^{(i)} - \bar{x}_1)(x_2^{(i)} - \bar{x}_2) + 2\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_3}(x_1^{(i)} - \bar{x}_1)(x_3^{(i)} - \bar{x}_2) + \cdots\right\} \\
&= \left(\frac{\partial f}{\partial x_1}\right)^2 \frac{1}{N}\sum_{i=1}^{N}(x_1^{(i)} - \bar{x}_1)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 \frac{1}{N}\sum_{i=1}^{N}(x_2^{(i)} - \bar{x}_2)^2 + \cdots + \left(\frac{\partial f}{\partial x_n}\right)^2 \frac{1}{N}\sum_{i=1}^{N}(x_n^{(i)} - \bar{x}_n)^2 \\
&\qquad + 2\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}\frac{1}{N}\sum_{i=1}^{N}(x_1^{(i)} - \bar{x}_1)(x_2^{(i)} - \bar{x}_2) + 2\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_3}\frac{1}{N}\sum_{i=1}^{N}(x_1^{(i)} - \bar{x}_1)(x_3^{(i)} - \bar{x}_2) + \cdots \\
&= \left(\frac{\partial f}{\partial x_1}\right)^2 \sigma_{x_1}^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 \sigma_{x_2}^2 + \cdots + \left(\frac{\partial f}{\partial x_n}\right)^2 \sigma_{x_n}^2 + 2\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}\sigma_{x_1x_2}^2 + 2\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_3}\sigma_{x_1x_3}^2 + \cdots
\end{aligned}$$

The $\sigma^2_{x_i}$ are simply the squares of the standard deviations of the x_i. The terms containing $\sigma^2_{x_i x_j}$ are correlation coefficients and are equal to zero if the variables are independent. Assuming that they *are* independent,

$$\sigma_y = \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 \sigma^2_{x_1} + \left(\frac{\partial f}{\partial x_2}\right)^2 \sigma^2_{x_2} + \cdots + \left(\frac{\partial f}{\partial x_n}\right)^2 \sigma^2_{x_n}} \tag{12}$$

which is equivalent to equation (11).

Uncertainty in Angle from Uncertainties in Rise and Run

Frequently, one measures an angle by measuring the "rise," Δy, and the "run," Δx and then finding the angle:

$$\theta = \mathrm{Tan}^{-1}\left\{\frac{\Delta y}{\Delta x}\right\}$$

The question is: what is the uncertainty in the angle, $\delta\theta$, given uncertainties δx and δy? Equation (11) states

$$\delta\theta = \sqrt{\left(\frac{\partial}{\partial x}\mathrm{Tan}^{-1}\left\{\frac{y}{x}\right\}\right)^2 (\delta x)^2 + \left(\frac{\partial}{\partial y}\mathrm{Tan}^{-1}\left\{\frac{y}{x}\right\}\right)^2 (\delta y)^2}$$

Once we know that

$$\frac{d}{dx}\mathrm{Tan}^{-1}\{x\} = \frac{1}{1+x^2}$$

then

$$\begin{aligned}
\delta\theta &= \sqrt{\left[\frac{1}{1+(y/x)^2}\left(-\frac{y}{x^2}\right)\right]^2 (\delta x)^2 + \left[\frac{1}{1+(y/x)^2}\left(\frac{1}{x}\right)\right]^2 (\delta y)^2} \\
&= \sqrt{\left[-\frac{y}{x^2+y^2}\right]^2 (\delta x)^2 + \left[\frac{x}{x^2+y^2}\right]^2 (\delta y)^2} \\
&= \frac{1}{r^2}\sqrt{(y\,\delta x)^2 + (x\,\delta y)^2}
\end{aligned}$$

or

$$= \frac{1}{r}\sqrt{(\sin\theta\,\delta x)^2 + (\cos\theta\,\delta y)^2}$$

where $r^2 = x^2 + y^2$. Notice that if the uncertainties are the same (say we measured x and y in the same way such that $\delta x = \delta y \equiv \delta z$), then the uncertainty in angle becomes particularly simple:

$$\delta\theta = \frac{\delta z}{r}$$

Beware: $\delta\theta$ is in *radians*. You'll have to convert if you want it in degrees.

Application to Speed Measurement

Some speed traps are set up by a law enforcement official flying overhead and measuring the time it takes a vehicle to cross a pair of lines some distance apart. We'd like to know the uncertainty in that measurement if we have some uncertainty in the time, σ_t and distance, σ_x. Given that

$$s = x/t$$

equation (12) says

$$\begin{aligned}
\sigma_s &= \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial t}\right)^2 \sigma_t^2} \\
\sigma_s^2 &= \frac{1}{t^2}\sigma_x^2 + \frac{x^2}{t^4}\sigma_t^2 \\
\frac{\sigma_s^2}{s^2} &= \left(\frac{t^2}{x^2}\right)\frac{1}{t^2}\sigma_x^2 + \left(\frac{t^2}{x^2}\right)\frac{x^2}{t^4}\sigma_t^2 \\
\left(\frac{\sigma_s}{s}\right)^2 &= \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_t}{t}\right)^2
\end{aligned}$$

just as required by (R2).

To get a feel for these uncertainties, let's re-write this as

$$\sigma_s = \frac{s}{x}\sqrt{\sigma_x^2 + s^2\sigma_t^2}$$

Now we have to fill in the blanks. First, notice that the uncertainty increases as the measured speed increases. It also decreases as the measured distance increases. All of the following are best guesses.

The measured distance is what would probably take a car doing the speed limit (55 MPH) about 5 seconds to traverse. That distance is about 1/13 mile.

The uncertainty in *measuring* that distance is the uncertainty in determining when the vehicle crosses the lines. We'll guess an uncertainty of 4 feet.

The uncertainty in measuring the time taken depends on the ability of the measurer to start and stop the stopwatch. We'll assume 0.1 second for each action.

This makes the equation for the standard error look like:

$$\begin{aligned}
\sigma_s &= \frac{s}{1/13}\sqrt{(4\text{ feet})^2 + s^2(0.1\text{ second})^2} \\
&= 13s\sqrt{(1/1320\text{ mile})^2 + s^2(1/36,000\text{ hour})^2} \\
&= 13s\sqrt{5.74\times10^{-7} + s^2(7.72\times10^{-10})}
\end{aligned}$$

So here's a table of speeds and uncertainties:

Measured Speed	Uncertainty (MPH)	Uncertainty (%)
30	0.44	1.47
40	0.70	1.75
50	1.03	2.06
60	1.43	2.38
70	1.90	2.71
80	2.44	3.05
90	3.06	3.40
100	3.74	3.74
150	8.26	5.50
200	14.6	7.30

Appendix B: Logarithms

Logarithms were invented when we wanted to answer the question, "What must x be to satisfy

$$y = a^x$$

where y and a are known values?" The value a must be greater than unity. We then define a function called "the logarithm with respect to base a" which answers our question:

$$\begin{aligned} y &= a^x \\ \log_a y &= x \end{aligned}$$

In plotting, we're often interested in the case where $a = 10$. Typically we leave off the subscript when using this base:

$$\log 10 = 1 \quad \text{or} \quad \log 1000 = 3 \quad \text{or} \quad \log 2 = 0.301$$

In physics, we often use base $e \equiv 2.718281828\ldots$. But instead of writing $\log_e$, we use the abbreviation "ln" which stands for "natural logarithm."

$$\ln\{10\} = 2.303 \quad \text{or} \quad \ln\{1000\} = 6.908 \quad \text{or} \quad \ln\{2\} = 0.693$$

Logarithms have some useful properties, and they were used extensively before the advent of cheap calculators. First and foremost, taking a logarithm is the inverse of exponentiation, so we have the following operations:

$$a^{\log_a x} = x \quad \text{and} \quad \log_a a^x = x$$

Secondly, because of the nature of exponents, the logarithm of a product is equal to the *sum* of the logarithms:

$$\begin{aligned} xy &= a^{\log_a x} a^{\log_a y} \\ xy &= a^{\log_a x + \log_a y} \\ \log_a\{xy\} &= \log_a \left\{a^{\log_a x + \log_a y}\right\} \\ \log_a\{xy\} &= \log_a x + \log_a y \end{aligned}$$

From this rule, we can derive the result of taking the logarithm of a power:

$$\log_a \{x^z\} = \log_a x + \log_a x + \cdots + \log_a x = z \log_a x$$

Finally, we might be interested in switching bases. So if we know how to compute $y = \log_a x$, how would we get $z = \log_b x$? From the definition of the logarithm, we have

$$\begin{aligned} x &= b^z \\ \log_a x &= \log_a \{b^z\} \\ \log_a x &= z \log_a \{b\} \\ \frac{\log_a x}{\log_a b} &= z \end{aligned}$$

Appendix C: Linear Least-Squares Fitting

Say we have a bunch of data points and would like to fit those points to a straight line. This straight line becomes our model. Whatever slope, m, and intercept, b, we choose for a model, this model will predict the value

$$y(x_i) = mx_i + b$$

for some x_i. We would like the values m and b which *minimizes* the difference between each prediction, $y(x_i)$, and the actual data point, y_i. So we're interested in the value of

$$\eta_i = y_i - y(x_i)$$

Minimizing η_i isn't useful. Why? We can make η_i as small as we like by letting y_i become as negative as we want (remember, minimizing means "making small." No number is smaller than $-\infty$). So let's minimize the *square* of this number:

$$\eta_i^2 = [y_i - y(x_i)]^2$$

This value can, at the very least, be zero. But we want the squared-errors for *all* our data points to be as close to zero as possible. So let's add up all the squared errors and minimize *that:*

$$\epsilon \equiv \sum_i \eta_i^2 = \sum_{i=1}^{N} [y_i - y(x_i)]^2 = \sum_{i=1}^{N} [y_i - mx_i - b]^2$$

Minimization, as every calculus student knows, is a requirement to take a derivative and set it equal to zero. We have two adjustable parameters, so let's find the values of m and b which minimize our squared error:

$$\frac{\partial \epsilon}{\partial m} = 0 = -2 \sum_{i=1}^{N} [y_i - mx_i - b] x_i \tag{F1}$$

$$\frac{\partial \epsilon}{\partial b} = 0 = -2 \sum_{i=1}^{N} [y_i - mx_i - b] \tag{F2}$$

Equations F1 and F2 are two equations with two unknowns. Algebra tells us that we can solve for these two unknowns. First, let's rearrange them:

$$m \sum_{i=1}^{N} x_i^2 + b \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x_i y_i \tag{F1}$$

$$m \sum_{i=1}^{N} x_i + Nb = \sum_{i=1}^{N} y_i \tag{F2}$$

We can now use Cramer's Rule to solve:

$$m = \frac{\begin{vmatrix} \sum_i x_i y_i & \sum_i x_i \\ \sum_i y_i & N \end{vmatrix}}{\begin{vmatrix} \sum_i x_i^2 & \sum_i x_i \\ \sum_i x_i & N \end{vmatrix}} = \frac{N \sum_i x_i y_i - \left(\sum_i x_i\right)\left(\sum_i y_i\right)}{N \sum_i x_i^2 - \left(\sum_i x_i\right)^2}$$

$$b = \frac{\begin{vmatrix} \sum_i x_i^2 & \sum_i x_i y_i \\ \sum_i x_i & \sum_i y_i \end{vmatrix}}{\begin{vmatrix} \sum_i x_i^2 & \sum_i x_i \\ \sum_i x_i & N \end{vmatrix}} = \frac{\left(\sum_i x_i^2\right)\left(\sum_i y_i\right) - \left(\sum_i x_i\right)\left(\sum_i x_i y_i\right)}{N \sum_i x_i^2 - \left(\sum_i x_i\right)^2}$$

An interesting side-light is when we want to draw the best-fit *horizontal* line to our data. In this case, we *require* that $m = 0$ and equation F2 tells us that

$$b = \frac{1}{N}\sum_{i=1}^{N} y_i$$

which is clearly the average value of the y_i.