Measurement and Uncertainty Analysis Guide

“It is better to be roughly right than precisely wrong.” – Alan Greenspan

Table of Contents

THE UNCERTAINTY OF MEASUREMENTS .......................................................... 2
RELATIVE (FRACTIONAL) UNCERTAINTY ................................................ 4
RELATIVE ERROR ......................................................................................... 5
TYPES OF UNCERTAINTY ........................................................................ 6
ESTIMATING EXPERIMENTAL UNCERTAINTY FOR A SINGLE MEASUREMENT .................................................. 9
ESTIMATING UNCERTAINTY IN REPEATED MEASUREMENTS ....................... 9
STANDARD DEVIATION ............................................................................ 12
STANDARD DEVIATION OF THE MEAN (STANDARD ERROR) ....................... 14
WHEN TO USE STANDARD DEVIATION VS STANDARD ERROR .......... 14
ANOMALOUS DATA .................................................................................. 15
BIASES AND THE FACTOR OF N–1 ............................................................. 15
SIGNIFICANT FIGURES ............................................................................ 16
UNCERTAINTY, SIGNIFICANT FIGURES, AND ROUNING ......................... 17
PROPAGATION OF UNCERTAINTY ............................................................ 20
THE UPPER-LOWER BOUND METHOD OF UNCERTAINTY PROPAGATION .... 20
QUADRATURE ............................................................................................. 21
COMBINING AND REPORTING UNCERTAINTIES ................................... 26
CONCLUSION: “WHEN DO MEASUREMENTS AGREE WITH EACH OTHER?” .......... 27
MAKING GRAPHS ..................................................................................... 29
USING EXCEL FOR DATA ANALYSIS IN PHYSICS LABS ......................... 31
GETTING STARTED ................................................................................... 31
CREATING AND EDITING A GRAPH .......................................................... 32
ADDING ERROR BARS .............................................................................. 33
ADDING A TRENDLINE ............................................................................ 33
DETERMINING THE UNCERTAINTY IN SLOPE AND Y-INTERCEPT ............... 34
INTERPRETING THE RESULTS .................................................................. 34
FINAL STEP – COPYING DATA AND GRAPHS INTO A WORD DOCUMENT ....... 34
USING LINEST IN EXCEL ......................................................................... 35
APPENDIX I: PROPAGATION OF UNCERTAINTY BY QUADRATURE ............. 39
APPENDIX II: PRACTICAL TIPS AND EXAMPLES FOR MEASURING AND CITING UNCERTAINTY .......... 41

last updated July 25, 2019

if your copy dates from the SP19 semester, it is not the latest version!
The Uncertainty of Measurements

Some numerical statements are exact: Mary has 3 brothers, and $2 + 2 = 4$. However, all measurements have some degree of uncertainty that may come from a variety of sources. The process of evaluating the uncertainty associated with a measurement result is often called uncertainty analysis or sometimes error analysis.

The complete statement of a measured value should include an estimate of the level of confidence associated with the value. Properly reporting an experimental result along with its uncertainty allows other people to make judgments about the quality of the experiment, and it facilitates meaningful comparisons with other similar values or a theoretical prediction. Without an uncertainty estimate, it is impossible to answer the basic scientific question: "Does my result agree with a theoretical prediction or results from other experiments?" This question is fundamental for deciding if a scientific hypothesis is confirmed or refuted.

When making a measurement, we generally assume that some exact or true value exists based on how we define what is being measured. While we may never know this true value exactly, we attempt to find this ideal quantity to the best of our ability with the time and resources available. As we make measurements by different methods, or even when making multiple measurements using the same method, we may obtain slightly different results. So how do we report our findings for our best estimate of this elusive true value? The most common way to show the range of values that we believe includes the true value is:

\[ \text{measurement} = (\text{best estimate} \pm \text{uncertainty}) \text{ units} \]

As an example, suppose you want to find the mass of a gold ring that you would like to sell to a friend. You do not want to jeopardize your friendship, so you want to get an accurate mass of the ring in order to charge a fair market price. You estimate the mass to be between 10 and 20 grams from how heavy it feels in your hand, but this is not a very precise estimate. After some searching, you find an electronic balance that gives a mass reading of 17.43 grams. While this measurement is much more precise than the original estimate, how do you know that it is accurate, and how confident are you that this measurement represents the true value of the ring's mass? Since the digital display of the balance is limited to 2 decimal places, you could report the mass as $m = 17.43 \pm 0.01$ g. Suppose you use the same electronic balance and obtain several more readings: 17.46 g, 17.42 g, 17.44 g, so that the average mass appears to be in the range of $17.44 \pm 0.02$ g. By now you may feel confident that you know the mass of this ring to the nearest hundredth of a gram, but how do you know that the true value definitely lies between 17.43 g and 17.45 g? Since you want to be honest, you decide to use another balance that gives a reading of 17.22 g. This value is clearly below the range of values found on the first balance, and under normal circumstances, you might not care, but you want to be
fair to your friend. So what do you do now? The answer lies in knowing something about the accuracy of each instrument.

To help answer these questions, we first define the terms *accuracy* and *precision*:

**Accuracy** is the closeness of agreement between a measured value and a true or accepted value. Measurement error is the amount of inaccuracy.

**Precision** is a measure of how well a result can be determined (without reference to a theoretical or true value). It is the degree of consistency and agreement among independent measurements of the same quantity; also the reliability or reproducibility of the result.

The accuracy and precision can be pictured as follows:

![Image of accuracy vs precision](image)

*Figure 1. Accuracy vs Precision*
The uncertainty estimate associated with a measurement should account for both the accuracy and precision of the measurement. Precision indicates the quality of the measurement, without any guarantee that the measurement is "correct." Accuracy, on the other hand, assumes that there is an ideal "true" value, and expresses how far your answer is from that "correct" answer. These concepts are directly related to random and systematic measurement uncertainties (next section).

Note: Unfortunately the terms error and uncertainty are often used interchangeably to describe both imprecision and inaccuracy. This usage is so common that it is impossible to avoid entirely. Whenever you encounter these terms, make sure you understand whether they refer to accuracy or precision, or both. In this document, we will emphasize the term “uncertainty” but will use the term “error,” as necessary, to avoid confusion with commonly found examples and standard usage of the term.

In order to determine the accuracy of a particular measurement, we have to know the ideal, true value, sometimes referred to as the “gold standard.” Sometimes we have a “textbook” measured value, which is well known, and we assume that this is our “ideal” value, and use it to estimate the accuracy of our result. Other times we know a theoretical value, which is calculated from basic principles, and this also may be taken as an “ideal” value. But physics is an empirical science, which means that the theory must be validated by experiment, and not the other way around. We can escape these difficulties and retain a useful definition of accuracy by assuming that, even when we do not know the true value, we can rely on the best available accepted value with which to compare our experimental value.

For the gold ring example, there is no accepted value with which to compare, and both measured values have the same precision, so there is no reason to believe one more than the other. We could look up the accuracy specifications for each balance as provided by the manufacturer, but the best way to assess the accuracy of a measurement is to compare it with a known standard. For this situation, it may be possible to calibrate the balances with a standard mass that is accurate within a narrow tolerance and is traceable to a primary mass standard at the National Institute of Standards and Technology (NIST). Calibrating the balances should eliminate the discrepancy between the readings and provide a more accurate mass measurement.

Relative (Fractional) Uncertainty

Precision is often reported quantitatively by using relative or fractional uncertainty, given by the ratio of the uncertainty divided by the average value:

\[
\text{Relative Uncertainty} = \left| \frac{\text{uncertainty}}{\text{measured quantity}} \right| \quad (1)
\]
The relative uncertainty is dimensionless but is often reported as a percentage or in parts per million (ppm) to emphasize the fractional nature of the value. A scientist might also make the statement that this measurement “is good to about 1 part in 500” or “precise to about 0.2%”. The relative uncertainty is important because it is used in *propagating* uncertainty in calculations using the result of a measurement, as discussed in a later section. For example, \( m = 75.5 \pm 0.5 \) g has a relative uncertainty of

\[
\frac{0.5}{75.5} = 0.006 = 0.7\%
\]

**Relative Error**

Accuracy is often reported quantitatively by using **relative error**:

\[
\text{Relative Error} = \frac{\text{measured value} - \text{expected value}}{\text{expected value}} \tag{2}
\]

If the expected value for \( m \) is 80.0 g, then the relative error is:

\[
\frac{75.5 - 80.0}{80.0} = -0.056 = -5.6\%
\]

**Critical Notes:**

- The minus sign indicates that the measured value is *less* than the expected value – *unless explicitly stated, the term “relative error” is signed and does not in and of itself refer to a magnitude*.
- The denominator is neither the measured value nor the average of the measured and expected value – *the relative error can only be cited when there is a known expected value or gold standard*.
- A relative error of 100% means that the upper bound could be as much as *twice* the value. A relative error of 200% means that the upper bound could be as much as *triple* the value, and so on. Do not misinterpret a 200% error as twice the value.
Types of Uncertainty

Measurement uncertainties may be classified as either random or systematic, depending on how the measurement was obtained (an instrument could cause a random uncertainty in one situation and a systematic uncertainty in another).

Random uncertainties are statistical fluctuations (in either direction) in the measured data. These uncertainties may have their origin in the measuring device, or in the fundamental physics underlying the experiment. The random uncertainties may be masked by the precision or accuracy of the measurement device. Random uncertainties can be evaluated through statistical analysis and can be reduced by averaging over a large number of observations (see “standard error” later in this document).

Systematic uncertainties are reproducible inaccuracies that are consistently in the “same direction,” and could be caused by an artifact in the measuring instrument, or a flaw in the experimental design (because of these possibilities, it is not uncommon to see the term “systematic error”). These uncertainties may be difficult to detect and cannot be analyzed statistically. If a systematic uncertainty or error is identified when calibrating against a standard, applying a correction or correction factor to compensate for the effect can reduce the bias. Unlike random uncertainties, systematic uncertainties cannot be detected or reduced by increasing the number of observations.

When making careful measurements, the goal is to reduce as many sources of uncertainty as possible and to keep track of those that cannot be eliminated. It is useful to know the types of uncertainties that may occur, so that we may recognize them when they arise. Common sources of uncertainty in physics laboratory experiments include:

Incomplete definition (may be systematic or random) - One reason that it is impossible to make exact measurements is that the measurement is not always clearly defined. For example, if two different people measure the length of the same string, they would probably get different results because each person may stretch the string with a different tension. The best way to minimize definition uncertainty is to carefully consider and specify the conditions that could affect the measurement.

Failure to account for a factor (usually systematic) – The most challenging part of designing an experiment is trying to control or account for all possible factors except the one independent variable that is being analyzed. For instance, you may inadvertently ignore air resistance when measuring free-fall acceleration, or you may fail to account for the effect of the Earth’s magnetic field when measuring the field near a small magnet. The best way to account for these sources of uncertainty is to brainstorm with your peers about all the factors that could possibly affect your
result. This brainstorm should be done before beginning the experiment in order to plan and account for the confounding factors before taking data. Sometimes a correction can be applied to a result after taking data to account for an uncertainty that was not detected earlier.

**Environmental factors** (systematic or random) - Be aware of uncertainty introduced by the immediate working environment. You may need to take account of or protect your experiment from vibrations, drafts, changes in temperature, and electronic noise or other effects from nearby apparatus.

**Instrument resolution** (random) - All instruments have finite precision that limits the ability to resolve small measurement differences. For instance, a meter stick cannot be used to distinguish distances to a precision much better than about half of its smallest scale division (typically 0.5 mm). One of the best ways to obtain more precise measurements is to use a null difference method instead of measuring a quantity directly. Null or balance methods involve using instrumentation to measure the difference between two similar quantities, one of which is known very accurately and is adjustable. The adjustable reference quantity is varied until the difference is reduced to zero. The two quantities are then balanced, and the magnitude of the unknown quantity can be found by comparison with a measurement standard. With this method, problems of source instability are eliminated, and the measuring instrument can be very sensitive and does not even need a scale. This type of measurement is more sophisticated and will typically not be used in the introductory physics courses.

**Calibration** (systematic) – Whenever possible, the calibration of an instrument should be checked before taking data. If a calibration standard is not available, the accuracy of the instrument should be checked by comparing with another instrument that is at least as precise, or by consulting the technical data provided by the manufacturer. Calibration errors are usually linear (measured as a fraction of the full scale reading), so that larger values result in greater absolute errors.

**Zero offset** (systematic) - When making a measurement with a micrometer caliper, electronic balance, or electrical meter, always check the zero reading first. Re-zero the instrument if possible, or at least measure and record the zero offset so that readings can be corrected later. It is also a good idea to check the zero reading throughout the experiment. Failure to zero a device will result in a constant offset that is more significant for smaller measured values than for larger ones.

**Physical variations** (random) - It is always wise to obtain multiple measurements over the widest range possible. Doing so often reveals variations that might otherwise go undetected. These variations may call for closer examination, or they may be combined to find an average value.
Parallax (systematic or random) - This error can occur whenever there is some distance between the measuring scale and the indicator used to obtain a measurement. If the observer’s eye is not squarely aligned with the pointer and scale, the reading may be too high or low (some analog meters have mirrors to help with this alignment).

Instrument drift (systematic) - Most electronic instruments have readings that drift over time. The amount of drift is generally not a concern, but occasionally this source of uncertainty can be significant.

Lag time and hysteresis (systematic) - Some measuring devices require time to reach equilibrium, and taking a measurement before the instrument is stable will result in a measurement that is too high or low. A common example is taking temperature readings with a thermometer that has not reached thermal equilibrium with its environment. A similar effect is hysteresis, wherein the instrument readings lag behind and appear to have a “memory” effect, as data are taken sequentially moving up or down through a range of values. Hysteresis is most commonly associated with materials that become magnetized when a changing magnetic field is applied.

Last but not least, some uncertainties are the result of carelessness, poor technique, or bias on the part of the experimenter. The experimenter may use a measuring device incorrectly, or may use poor technique in taking a measurement, or may introduce a bias into measurements by expecting (and inadvertently forcing) the results to agree with the expected outcome. Gross uncertainties of this nature can be referred to as mistakes or blunders, and should be avoided and corrected if discovered. As a rule, these uncertainties are excluded from any uncertainty analysis discussion because it is generally assumed that the experimental result was obtained by following correct and well-intentioned procedures – there is no point to performing an experiment and then reporting that it was known to be incorrect. The term human error should be avoided in uncertainty analysis discussions because it is too general to be useful.
Estimating Experimental Uncertainty for a Single Measurement

Any measurement will have some uncertainty associated with it, no matter the precision of the measuring tool. How is this uncertainty determined and reported? The uncertainty of a single measurement is limited by the precision and accuracy of the measuring instrument, along with any other factors that might affect the ability of the experimenter to make the measurement.

For example, if you are trying to use a meter stick to measure the diameter of a tennis ball, the uncertainty might be ±5 mm, but if you use a Vernier caliper, the uncertainty could be reduced to maybe ±2 mm. The limiting factor with the meter stick is parallax, while the second case is limited by ambiguity in the definition of the tennis ball’s diameter (it’s fuzzy!). In both of these cases, the uncertainty is greater than the smallest divisions marked on the measuring tool (likely 1 mm and 0.05 mm respectively). Unfortunately, there is no general rule for determining the uncertainty in all measurements. The experimenter is the one who can best evaluate and quantify the uncertainty of a measurement based on all the possible factors that affect the result. Therefore, the person making the measurement has the obligation to make the best judgment possible and report the uncertainty in a way that clearly explains what the uncertainty represents:

Measurement = (measured value ± standard uncertainty) (unit of measurement)

where “± standard uncertainty” indicates approximately a 68% confidence interval (see sections on Standard Deviation and Reporting Uncertainties).

Example: Diameter of tennis ball = 6.7 ± 0.2 cm

Estimating Uncertainty in Repeated Measurements

Suppose you time the period of oscillation of a pendulum using a digital instrument (that you assume is measuring accurately) and find that $T = 0.44$ seconds. This single measurement of the period suggests a precision of ±0.005 s, but this instrument precision may not give a complete sense of the uncertainty, and you should avoid reporting the uncertainty in this fashion if possible. If you repeat the measurement several times and examine the variation among the measured values, you can get a better idea of the uncertainty in the period. For example, here are the results of 5 measurements, in seconds: 0.46, 0.44, 0.45, 0.44, 0.41. For this situation, the best estimate of the period is the average, or mean:

$$\text{Average (mean)} = \frac{x_1 + x_2 + \ldots + x_N}{N}$$
Whenever possible, repeat a measurement several times and average the results. This average is generally the best estimate of the “true” value (unless the data set is skewed by one or more outliers which should be examined to determine if they are bad data points that should be omitted from the average or valid measurements that require further investigation). Generally, the more repetitions you make of a measurement, the better this estimate will be, but be careful to avoid wasting time taking more measurements than is necessary for the precision required.

Consider, as another example, the measurement of the width of a piece of paper using a meter stick. Being careful to keep the meter stick parallel to the edge of the paper (to avoid a systematic error which would cause the measured value to be consistently higher than the correct value), the width of the paper is measured at a number of points on the sheet, and the values obtained are entered in a data table. Note that the last digit is only a rough estimate, since it is difficult to read a meter stick to the nearest tenth of a millimeter (0.01 cm) – we retain the last digit for now to make a point later.

<table>
<thead>
<tr>
<th>Observation</th>
<th>Width (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>31.33</td>
</tr>
<tr>
<td>#2</td>
<td>31.15</td>
</tr>
<tr>
<td>#3</td>
<td>31.26</td>
</tr>
<tr>
<td>#4</td>
<td>31.02</td>
</tr>
<tr>
<td>#5</td>
<td>31.20</td>
</tr>
</tbody>
</table>

Table 1. Five Measurements of the Width of a Piece of Paper

Average = \( \frac{\text{sum of observed widths}}{\text{number of observations}} \) = \( \frac{155.96 \text{ cm}}{5} \) = 31.19 cm

This average is the best available estimate of the width of the piece of paper, but it is not exact. We would have to average an infinite number of measurements to approach the true mean value, and even then, we are not guaranteed that the mean value is accurate because there is still likely some systematic uncertainty from the measuring tool, which is difficult to calibrate perfectly unless it is the gold standard. So how do we express the uncertainty in our average value?

One way to express the variation among the measurements is to use the average deviation. This statistic tells us on average (with 50% confidence) how much the individual measurements vary from the mean.

Average Deviation, \( \bar{d} = \frac{\sum_{i=1}^{N} |x_i - \bar{x}|}{N} \)

The average deviation is a sufficient measure of uncertainty; however, it is important to understand the distribution of measurements. The Central Limit
Theorem proves that as the number of independent measurements increases, and assuming that the variations in these measurements are random (i.e., there are no systematic uncertainties), the distribution of measurements will approach the normal distribution, more commonly known as a bell curve. In this course, we will assume that our measurements, performed in sufficient number, will produce a bell curve (normal) distribution. In this case, the standard deviation is an alternate way to characterize the spread of the data. The standard deviation is always slightly greater than the average deviation, and is used because of its mathematical association with the normal distribution.
**Standard Deviation**

To calculate the standard deviation for a sample of $N$ measurements:

1. Sum all the measurements and divide by $N$ to get the **average**, or **mean**.
2. Subtract this **average** from each of the $N$ measurements to obtain $N$ "deviations."
3. **Square** each of the $N$ deviations and add them together.
4. Divide this result by $(N-1)$ and take the square root.

To convert this into a formula, let the $N$ measurements be called $x_1, x_2, ..., x_N$. Let the average of the $N$ values be called $\bar{x}$. Then each deviation is given by

$$\delta x_i = x_i - \bar{x}, \text{ for } i = 1, 2, ..., N$$

The standard deviation is then:

$$s = \sqrt{\frac{(\delta x_1^2 + \delta x_2^2 + ... + \delta x_N^2)}{(N-1)}} = \sqrt{\frac{\sum \delta x_i^2}{(N-1)}}$$

In the meter stick and paper example, the average paper width $\bar{x}$ is 31.19 cm. The deviations are:

<table>
<thead>
<tr>
<th>Observation</th>
<th>Width (cm)</th>
<th>Deviation (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>31.33</td>
<td>+0.14</td>
</tr>
<tr>
<td>#2</td>
<td>31.15</td>
<td>-0.04</td>
</tr>
<tr>
<td>#3</td>
<td>31.26</td>
<td>+0.07</td>
</tr>
<tr>
<td>#4</td>
<td>31.02</td>
<td>-0.17</td>
</tr>
<tr>
<td>#5</td>
<td>31.20</td>
<td>+0.01</td>
</tr>
</tbody>
</table>

**Table 1 (completed). Five Measurements of the Width of a Sheet of Paper**

The **average** deviation is: $\bar{d} = 0.09$ cm

The **standard** deviation is: $s = \sqrt{\frac{(0.14)^2 + (0.04)^2 + (0.07)^2 + (0.17)^2 + (0.01)^2}{5-1}} = 0.12$ cm

The significance of the standard deviation is this: if you now make one more measurement using the same meter stick, you can reasonably expect (with about 68% confidence) that the new measurement will be within 0.12 cm of the estimated average of 31.19 cm. In fact, it is reasonable to use the standard deviation as the uncertainty associated with this **single** new measurement. However, the uncertainty
of the average value is the standard deviation of the mean, which is always less than the standard deviation (see next section).

Consider an example of 100 measurements of a quantity, for which the average or mean value is 10.50 and the standard deviation is $s = 1.83$. Figure 2 below is a histogram of the 100 measurements, which shows how often a certain range of values was measured. For example, in 20 of the measurements, the value was in the range 9.50 to 10.50, and most of the readings were close to the mean value of 10.50. The standard deviation $s$ for this set of measurements is roughly how far from the average value most of the readings fell. For a large enough sample, approximately 68% of the readings will be within one standard deviation (“1-sigma”) of the mean value, 95% of the readings will be in the interval $\bar{x} \pm 2s$ (“2-sigma”), and nearly all (99.7%) of the readings will lie within 3 standard deviations (“3-sigma”) of the mean. The smooth curve superimposed on the histogram is the normal distribution predicted by theory for measurements involving random errors. As more and more measurements are made, the histogram will better approximate a bell-shaped curve, but the standard deviation of the distribution will remain approximately the same.

![Figure 2. A Normal Distribution (Bell Curve) Based on 100 Measurements](image-url)


**Standard Deviation of the Mean (Standard Error)**

When reporting the average value of \( N \) measurements, the uncertainty associated with this average value is the *standard deviation of the mean*, often called the *standard error* (SE).

\[
\text{Standard Deviation of the Mean, or Standard Error (SE), } \sigma_x = \frac{s}{\sqrt{N}} \tag{3}
\]

The standard error is smaller than the standard deviation by a factor of \( 1/\sqrt{N} \). This reflects the fact that we expect the uncertainty of the average value to get smaller when we use a larger number of measurements. In the previous example, we divided the standard deviation of 0.12 by \( \sqrt{5} \) to get the standard error of 0.05 cm. The final result should then be reported as “average paper width = 31.19 ± 0.05 cm.” For this example, the relative uncertainty would be \( (0.05/31.19) \), or \( \approx 0.2\% \).

**When to Use Standard Deviation vs Standard Error**

For repeated measurements, the significance of the standard deviation \( s \) is that you can reasonably expect (with about 68% confidence) that the next measurement will be within \( s \) of the estimated average. It may be reasonable to use the standard deviation as the uncertainty associated with this measurement; however, as more measurements are made, the value of the standard deviation will be refined but it will not significantly decrease as the number of measurements is increased; therefore, it will not be the *best* estimate of the uncertainty of a set of measurements. In contrast, if you are confident that the systematic uncertainty in your measurement is very small, then it is reasonable to assume that your finite sample of all possible measurements is not biased away from the “true” value. In this case, the uncertainty of the average value can be expressed as the standard deviation of the mean, which is always less than the standard deviation by a factor of \( \sqrt{N} \).

![Figure 3. Standard Deviation vs Standard Error](image-url)
If you are not confident that the systematic uncertainty in your measurement is very small, the uncertainty that should be reported is the standard combined uncertainty \( (U_c) \) that includes all known uncertainty estimates (see section on Combining and Reporting Uncertainties later in this document).

**Anomalous Data**

The first step you should take in analyzing data (and even while taking data) is to examine the data set as a whole to look for patterns and outliers. Anomalous data points that lie outside the general trend of the data may suggest an interesting phenomenon that could lead to a new discovery, or they may simply be the result of a mistake or random fluctuations. In any case, an outlier requires closer examination to determine the cause of the unexpected result. Extreme data should never be “thrown out” without clear justification and explanation, because you may be discarding the most significant part of the investigation! However, if you can clearly justify omitting an inconsistent datum, then you may exclude the outlier from your analysis so that the average value is not skewed from the “true” mean. There are a number of statistical measures that help quantify the decision to discard outliers, but they are beyond the scope of this document. Be aware of the possibility of anomalous data, and address the topic as needed in the discussion included with a lab report or lab notebook.

**Biases and the Factor of N−1**

You may find it surprising that the best value (average) is calculated by normalizing (dividing) by \( N \), whereas the standard deviation is calculated by normalizing to \( N−1 \). The reason is because normalizing to \( N \) is known to underestimate the correct value of the width of a normal distribution, unless \( N \) is large. This underestimate is referred to as a bias and is the result of incomplete sampling (that is, the population of measurements, or sample, falls short of the entire population of measurements that could be taken), also known as “sampling error.” If the number of samples is less than or about 10, even the \( N−1 \) term (known as Bessel’s correction) is known to induce a bias. Determining the exact correction to minimize or eliminate bias depends on the distribution of the data, and there is no simple exact equation that can be applied; however, for small sample sizes that are quite common in introductory physics classes, a correction of \( N−1.5 \) may be more appropriate. If you use a correction factor of 1.5 in your lab reports, you must make this clear in your analysis and cite this Guide as a reference.
**Significant Figures**

The number of significant figures (sig figs) in a value can be defined as all the digits between and including the first non-zero digit from the left, through the last digit. For instance, 0.44 has two sig figs, and the number 66.770 has 5 sig figs. Zeroes are significant except when used to locate the decimal point, as in the number 0.00030, which has 2 sig figs. Zeroes may or may not be significant for numbers like 1200, where it is not clear whether two, three, or four sig figs are indicated. To avoid this ambiguity, such numbers should be expressed in scientific notation (e.g., $1.20 \times 10^3$ indicates 3 sig figs).

When using a calculator, the display will often show many digits, only some of which are meaningful (significant in a different sense). For example, if you want to estimate the area of a circular playing field, you might pace off the radius to be 9 meters and use the formula $A = \pi r^2$. When you compute this area, the calculator will report a value of 254.4690049 m$^2$. It would be extremely misleading to report this number as the area of the field, because it would suggest that you know the area to an absurd degree of precision – to within a fraction of a square millimeter! Since the radius is only known to one significant figure, it is considered best practice to also express the final answer to only one significant figure: Area = $3 \times 10^2$ m$^2$.

From this example, we can see that the number of significant figures reported for a value implies a certain degree of precision and can suggest a rough estimate of the relative uncertainty:

- 1 significant figure suggests a relative uncertainty of about 10% to 100%
- 2 significant figures suggest a relative uncertainty of about 1% to 10%
- 3 significant figures suggest a relative uncertainty of about 0.1% to 1%

To understand this connection more clearly, consider a value with 2 significant figures, like 99, which suggests an uncertainty of ±1, or a fractional uncertainty of ±1/99 = ±1% (one might argue that the implied uncertainty in 99 is ± 0.5 since the range of values that would round to 99 are 98.5 to 99.4; however, since the uncertainty here is only a rough estimate, there is not much point debating the factor of two.) The smallest 2-significant-figure number, 10, also suggests an uncertainty of ±1, which in this case is a fractional uncertainty of ±1/10 = ±10%. The ranges for other numbers of significant figures can be reasoned in a similar manner.

**Warning:** this procedure is open to a wide range of interpretation; therefore, one should use caution when using significant figures to imply uncertainty, and the method should only be used if there is no other better way to determine uncertainty. An explicit warning to this effect should accompany the use of this method.
Measurements & Uncertainty Analysis

Subject to the above warning, significant figures can be used to find a possibly appropriate precision for a calculated result for the four most basic math functions.

- For multiplication and division, the number of significant figures that are reliably known in a product or quotient is the same as the smallest number of significant figures in any of the original numbers.

  **Example:**
  
  \[
  6.6 \quad (2 \text{ significant figures}) \\
  \times 7328.7 \quad (5 \text{ significant figures}) \\
  48369.42 = 48 \times 10^3 \quad (2 \text{ significant figures})
  \]

- For addition and subtraction, the result should be rounded off to the last decimal place reported for the least precise number.

  **Examples:**
  
  \[
  223.64 \\
  +54 \\
  278
  \]

  \[
  5560.5 \\
  +0.008 \\
  5560.5
  \]

**Critical Note:** if a calculated number is to be used in further calculations, it is mandatory to keep guard digits to reduce rounding errors that may accumulate. The final answer can then be rounded according to the above guidelines. The number of guard digits required to maintain the integrity of a calculation depends on the type of calculation. For example, the number of guard digits must be larger when performing power law calculations than when adding.

**Uncertainty, Significant Figures, and Rounding**

For the same reason that it is dishonest to report a result with more significant figures than are reliably known, the uncertainty value should also not be reported with excessive precision. For example, it would be unreasonable to report a result in the following way:

\[
\text{measured density} = 8.93 \pm 0.475328 \text{ g/cm}^3 \quad \text{WRONG!}
\]

The uncertainty in the measurement cannot possibly be known so precisely! In most experimental work, the confidence in the uncertainty estimate is not much better than about ±50% because of all the various sources of error, none of which can be known exactly.

Therefore, unless explicitly justified, uncertainty values should be stated (rounded) to one significant figure.
To help give a sense of the amount of confidence that can be placed in the standard deviation as a measure of uncertainty, the following table indicates the relative uncertainty associated with the standard deviation for various sample sizes. Note that in order for a standard deviation to be reported to a precision of 3 significant figures, more than 10,000 readings would be required!

<table>
<thead>
<tr>
<th>N</th>
<th>Relative Uncertainty*</th>
<th>Significant Figures Valid</th>
<th>Implied Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>71%</td>
<td>1</td>
<td>±10% to 100%</td>
</tr>
<tr>
<td>3</td>
<td>50%</td>
<td>1</td>
<td>±10% to 100%</td>
</tr>
<tr>
<td>4</td>
<td>41%</td>
<td>1</td>
<td>±10% to 100%</td>
</tr>
<tr>
<td>5</td>
<td>35%</td>
<td>1</td>
<td>±10% to 100%</td>
</tr>
<tr>
<td>10</td>
<td>24%</td>
<td>1</td>
<td>±10% to 100%</td>
</tr>
<tr>
<td>20</td>
<td>16%</td>
<td>1</td>
<td>±10% to 100%</td>
</tr>
<tr>
<td>30</td>
<td>13%</td>
<td>1</td>
<td>±10% to 100%</td>
</tr>
<tr>
<td>50</td>
<td>10%</td>
<td>2</td>
<td>±1% to 10%</td>
</tr>
<tr>
<td>100</td>
<td>7%</td>
<td>2</td>
<td>±1% to 10%</td>
</tr>
<tr>
<td>10000</td>
<td>0.7%</td>
<td>3</td>
<td>±0.1% to 1%</td>
</tr>
</tbody>
</table>

*The relative uncertainty in the standard deviation is given by the approximate formula: $\frac{\sigma_{\sigma}}{\sigma} = \frac{1}{\sqrt{2(N-1)}}$

When an explicit uncertainty estimate is made using the standard deviation, the uncertainty term indicates how many significant figures should be reported in the measured value (not the other way around!). For example, the uncertainty in the density measurement above is about 0.5 g/cm$^3$, which suggests that the digit in the tenths place is uncertain, and should be the last one reported. The other digits in the hundredths place and beyond are insignificant, and should not be reported:

$\text{measured density} = 8.9 \pm 0.5 \text{ g/cm}^3 \quad \text{RIGHT!}$

An experimental value should be rounded to be consistent with the magnitude of its uncertainty. This generally means that the last significant figure in any reported value should be in the same decimal place as the uncertainty, and unless a large number of measurements have been taken, the uncertainty should be reported to only one sig fig.

In most instances, this practice of rounding an experimental result to be consistent with the uncertainty estimate gives the same number of significant figures as the rules discussed earlier for simple propagation of uncertainties for adding, subtracting, multiplying, and dividing.
When conducting an experiment, it is important to keep in mind that precision is expensive (both in terms of time and material resources). Do not waste your time trying to obtain a precise result when only a rough estimate is required. The cost increases exponentially with the amount of precision required, so the potential benefit of this precision must be weighed against the extra cost.

Important Notes about sig figs when interpreting both this guide and course materials:

a) On exams, homeworks, or solutions, you may see a quantity cited imprecisely (for example, “1200 m” instead of “1200.0 m”). This is typically done for readability, and does not imply anything about uncertainty unless the quantity is tied explicitly to an experiment. You may consider such values to be exact.

b) In contrast, you may see examples or solutions that appear to retain an excessive number of sig figs (including in this document). Such examples may have skipped the final rounding step for either clarity or readability. Often, these examples retain higher precision to help you ascertain that you have duplicated a calculation correctly, or to distinguish them from incorrect solutions that happen to be close to the correct answer. Again, unless such examples are tied explicitly to an uncertainty, you may ignore such inconsistencies; however, when you present an answer on an assignment or exam, you should minimize your uncertainty precision to one sig fig unless otherwise directed or justified.

c) Similarly, if an experiment described on an exam or homework claims to have more than one sig fig in the uncertainty, and you are not asked to address this aspect of the experiment, take the uncertainty precision at face value as correct.

d) Normally, when reviewing experimental questions (such as are found on the practicum or lab reports), graders carefully scrutinize your use of more than 1 sig fig when citing uncertainty. One exception has already been discussed, and is embodied in Table 2: if enough samples have been collected to justify 2 sig figs, then this must be explicitly discussed in your answers. A second exception could be when the leading significant figure is a “1” – it is possible to induce a bias when rounding to one sig fig in this case. Consider the example of rounding ±1.4 as opposed to ±9.4, to one sig fig. The value of ±1.0 represents a 40% discrepancy from ±1.4 (very large), whereas the value of ±9.0 represents a 4% discrepancy from ±9.4 (much smaller, albeit nonzero). The best approach here, as for virtually all exam and assignment examples, is to cite and discuss the high precision answer, which indicates that you know the details of the calculation, and then cite and discuss your final answer with suitable rounding and justification.

Practical tips for measuring and citing uncertainty, and uncertainty calculation examples, can be found in Appendix II.
Propagation of Uncertainty

Suppose we want to determine a quantity \( f \), which depends on \( x \) and maybe several other variables \( y, z \), etc. We want to know the uncertainty in \( f \) if we measure \( x, y \), etc, with uncertainties \( \sigma_x, \sigma_y \), etc. That is, we want to find out how the uncertainty in one set of variables (usually the independent variables) propagates to the uncertainty in another set of variables (usually the dependent variables). There are two primary methods of performing this propagation procedure:

- **upper-lower bound**
- **quadrature**

The upper-lower bound method is simpler in concept, but tends to overestimate the uncertainty, while the quadrature method is more sophisticated (and complicated) but provides a better statistical estimate of the uncertainty.

The Upper-Lower Bound Method of Uncertainty Propagation

This method uses the uncertainty ranges of each variable to calculate the maximum and minimum values of the function. You can also think of this procedure as examining the best and worst case scenarios. As an example, consider the division of two variables – a common example is the calculation of average speed:

\[
v_{\text{avg}} = \frac{\Delta x}{\Delta t}
\]

Let’s say an experiment is done repeatedly and measures a distance travelled of 30 ± 0.5 m during a time of 2 ± 0.1 sec. To find the upper and lower bound of \( v_{\text{avg}} \), the uncertainties must be set to create the “worst case scenario” for the uncertainty in \( v_{\text{avg}} \):

\[
v_{\text{avg}-\text{max}} = \frac{30 + 0.5 \text{ m}}{2 - 0.1 \text{ sec}} = 16.05 \text{ m/s}
\]

\[
v_{\text{avg}-\text{min}} = \frac{30 - 0.5 \text{ m}}{2 + 0.1 \text{ sec}} = 14.05 \text{ m/s}
\]

The best (expected) value for the average speed is 30/2 = 15.00 m/s. The upper bound is 1.05 m/s higher but the lower bound is 0.95 m/s lower (different from 1.05). This is a typical outcome when using the upper-lower bound method. The uncertainty should be expressed as the most conservative value. Thus:

\[
v_{\text{avg}} = 15.00 \pm 1.05 \text{ m/s} \quad \text{PREFERRED!}
\]

Note that it is not correct to take the difference and divide by two:

\[
v_{\text{avg}} = 15.05 \pm 1.00 \text{ m/s} \quad \text{NOT PREFERRED!}
\]
Although the last result satisfies symmetry between the bounds, it explicitly calculates an **incorrect** value of the best-known expected value of the average speed.

Many times, the difference between the so-called “preferred” and “not preferred” approaches is not significant enough to be an issue. Citing the uncertainty in this example to 3 sig figs would require a strong justification. The accepted approach (and the one that should be followed in the absence of such justification) is to round to $15.0 \pm 1.1$ m/s, and then round again to $15 \pm 1$ m/s. With this approach, arguments about whether the “actual answer” is 15.00 or 15.05 are irrelevant, because the experiment does not justify such precision.

This asymmetry arises in the case of non-linear functions as well. For example, suppose you measure an angle to be $\theta = 25^\circ \pm 1^\circ$ and you need to find $f = \cos \theta$:  

\[
\begin{align*}
 f_{\text{min}} &= \cos(26^\circ) = 0.8988 \\
 f &= \cos(25^\circ) = 0.9063 \\
 f_{\text{max}} &= \cos(24^\circ) = 0.9135
\end{align*}
\]

The differences are $f - f_{\text{min}} = 0.0075$ and $f_{\text{max}} - f = 0.0072$. When rounding the uncertainty to 1 sig fig, the most conservative value should be chosen (0.008). Thus the answer would be cited as $f = 0.906 \pm 0.008$. Note the following about this example:

- the maximum $\theta$ is associated with the minimum $f$, and vice versa
- guard digits are retained to make informed choices about the final answer
- although $\theta$ was only measured to 2 sig figs, $f$ is known to 3 sig figs

The upper-lower bound method is especially useful when the functional relationship is not clear or is incomplete. One practical application is forecasting the expected range in an expense budget. In this case, some expenses may be fixed, while others may be uncertain, and the range of these uncertain terms could be used to predict the upper and lower bounds on the total expense.

**Quadrature**

The quadrature method yields a **standard uncertainty** estimate (with a 68% confidence interval) and is especially useful and effective in the case of several variables that weight the uncertainty non-uniformly. The method is derived with several examples shown below.

For a single-variable function $f(x)$, the deviation in $f$ can be related to the deviation in $x$ using calculus:

\[
\delta f = \left( \frac{df}{dx} \right) \delta x
\]
Taking the square and the average yields:

$$\overline{\sigma_f^2} = \left( \frac{df}{dx} \right)^2 \overline{\sigma_x^2}$$

Using the definition of $\sigma$ yields:

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

Examples for power laws:

- $f = \sqrt{x}$
  - $f = x^2$
  - $\frac{df}{dx} = \frac{1}{2\sqrt{x}}$
  - $\frac{df}{dx} = 2x$
  - $\sigma_f = \frac{\sigma_x}{2\sqrt{x}}$ or $\frac{\sigma_f}{f} = \frac{1}{2} \frac{\sigma_x}{x}$
  - $\frac{\sigma_f}{f} = \frac{2}{x} \frac{\sigma_x}{x}$

Note that by judiciously normalizing, it is easy to express the relative (fractional) uncertainty in one variable with respect to the relative (fractional) uncertainty in another. Note also that the weighting is directly related to the power exponent of the function. Now reconsider the trig example from the upper-lower bound section:

- $f = \cos \theta$
  - $\frac{df}{d\theta} = -\sin \theta$
  - $\sigma_f = |\sin \theta| \sigma_\theta$

Note that in this situation, $\sigma_\theta$ must be in radians. For $\theta = 25^\circ \pm 1^\circ (0.727 \pm 0.017)$

$$\sigma_f = |\sin \theta| \sigma_\theta = (0.423)(\pi/180) = 0.0074$$

This is essentially the same result as upper-lower bound method. The fractional uncertainty follows immediately as:

$$\frac{\sigma_f}{f} = |\tan \theta| \sigma_\theta$$
The deeper power of the quadrature method is evident in the case where \( f \) depends on two or more variables; the derivation above can be repeated with minor modification. For two variables, \( f(x, y) \):

\[
\delta f = \left( \frac{\partial f}{\partial x} \right) \delta x + \left( \frac{\partial f}{\partial y} \right) \delta y
\]

The partial derivative \( \frac{\partial f}{\partial x} \) means differentiating \( f \) with respect to \( x \), while holding the other variables fixed. Taking the square and the average yields the generalized law of propagation of uncertainty by quadrature:

\[
(\delta f)^2 = \left( \frac{\partial f}{\partial x} \right)^2 (\delta x)^2 + \left( \frac{\partial f}{\partial y} \right)^2 (\delta y)^2 + 2 \left( \frac{\partial f}{\partial x} \right) \left( \frac{\partial f}{\partial y} \right) \delta x \delta y 
\]

(4)

If the measurements of \( x \) and \( y \) are uncorrelated, then \( \delta x \delta y = 0 \), and this reduces to its most common form:

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

Addition and Subtraction Example:

\[
f = x \pm y
\]

\[
\frac{\partial f}{\partial x} = 1, \quad \frac{\partial f}{\partial y} = \pm 1 \quad \rightarrow \quad \sigma_f = \sqrt{\sigma_x^2 + \sigma_y^2}
\]

When adding (or subtracting) independent measurements, the absolute uncertainty of the sum (or difference) is the root sum of squares (RSS) of the individual absolute uncertainties. When adding correlated measurements, the uncertainty in the result is simply the sum of the absolute uncertainties, which is always a larger uncertainty estimate than adding in quadrature (RSS). Adding or subtracting a constant does not change the absolute uncertainty of the calculated value as long as the constant is an exact value.
Measurements & Uncertainty Analysis

Multiplication example: \( f = xy \)

\[
\frac{\partial f}{\partial x} = y, \quad \frac{\partial f}{\partial y} = x \quad \rightarrow \quad \sigma_f = \sqrt{\sigma_x^2 + \sigma_y^2} \]

Dividing the above equation by \( f = xy \) yields:

\[
\frac{\sigma_f}{f} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2} \]

Division example: \( f = x/y \)

\[
\frac{\partial f}{\partial x} = \frac{1}{y}, \quad \frac{\partial f}{\partial y} = -\frac{x}{y^2} \quad \rightarrow \quad \sigma_f = \sqrt{\left(\frac{1}{y}\right)^2 \sigma_x^2 + \left(\frac{x}{y^2}\right)^2 \sigma_y^2} \]

Dividing the previous equation by \( f = x/y \) yields:

\[
\frac{\sigma_f}{f} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2} \]

When multiplying (or dividing) independent measurements, the relative uncertainty of the product (quotient) is the RSS of the individual relative uncertainties. When multiplying correlated measurements, the uncertainty in the result is just the sum of the relative uncertainties, which is always a larger uncertainty estimate than adding in quadrature (RSS). Multiplying or dividing by a constant does not change the relative uncertainty of the calculated value.

Note that the relative (fractional) uncertainty in \( f \) has the same form for multiplication and division: the relative uncertainty in a product or quotient depends on the relative uncertainty of each individual term.

As another example, consider propagating the uncertainty in the speed \( v = at \), where the acceleration is \( a = 9.8 \pm 0.1 \) m/s\(^2\) and the time is \( t = 1.2 \pm 0.1 \) s.

\[
\frac{\sigma_v}{v} = \sqrt{\left(\frac{\sigma_a}{a}\right)^2 + \left(\frac{\sigma_t}{t}\right)^2} = \sqrt{\left(\frac{0.1}{9.8}\right)^2 + \left(\frac{0.1}{1.2}\right)^2} = \sqrt{(0.010)^2 + (0.029)^2} = 0.031 \text{ or } 3.1\% \]
Notice that the relative uncertainty in $t$ (2.9%) is significantly greater than the relative uncertainty for $a$ (1.0%), and therefore the relative uncertainty in $v$ is essentially the same as for $t$ (about 3%).

Graphically, the RSS is like the Pythagorean theorem: The total uncertainty is the length of the hypotenuse of a right triangle with legs the length of each uncertainty component.

**Timesaving approximation:** “A chain is only as strong as its weakest link.”
If one of the uncertainty terms is more than 3 times greater than the other terms, the root-squares formula can be skipped, and the combined uncertainty is simply the largest uncertainty. This shortcut can save a lot of time without losing any accuracy in the estimate of the overall uncertainty.

The quadrature method can be generalized to all power laws in the following way:

$$f = x^n y^m$$

$$\frac{\sigma_f}{f} = \sqrt{n^2 \left(\frac{\sigma_x}{x}\right)^2 + m^2 \left(\frac{\sigma_y}{y}\right)^2} \quad (5)$$

The proof of this is shown in Appendix I.

The uncertainty estimate from the upper-lower bound method is generally larger than the standard uncertainty estimate found from the quadrature method, but both methods will give a reasonable estimate of the uncertainty in a calculated value.

**Note:** Once you have an understanding of the quadrature method, it is not required to perform the partial derivative every time you are presented with a propagation of uncertainty problem in any of the above forms! Instead, simply apply the correct formula for the relative uncertainties.
Combining and Reporting Uncertainties

In 1993, the International Standards Organization (ISO) published the first official worldwide Guide to the Expression of Uncertainty in Measurement. Before this time, uncertainty estimates were evaluated and reported according to different conventions depending on the context of the measurement or the scientific discipline. Here are a few key points from this 100-page guide, which can be found in modified form on the NIST website (see References).

When reporting a measurement, the measured value should be reported along with an estimate of the total combined standard uncertainty $U_c$ of the value. The total uncertainty is found by combining the uncertainty components based on the two types of uncertainty analysis:

**Type A evaluation of standard uncertainty** – method of evaluation of uncertainty by the statistical analysis of a series of observations. This method is primarily associated with uncertainties that are *randomly* distributed.

**Type B evaluation of standard uncertainty** – method of evaluation of uncertainty by means other than the statistical analysis of a series of observations. This method is primarily associated with uncertainties that are *systematic*, including the precision of the measuring instrument and any other factors that the experimenter believes are important.

The individual uncertainty components $u_i$ should be combined using the *law of propagation of uncertainties*, commonly called the “root-sum-of-squares” or “RSS” method (very similar to quadrature). This combined standard uncertainty should be equivalent to the standard deviation of the result, making this uncertainty value correspond with a 68% confidence interval. If a wider confidence interval is desired, the uncertainty can be multiplied by a *coverage factor* (usually $k = 2$ or 3) to provide an uncertainty range that is believed to include the true value with a confidence of 95% (for $k = 2$) or 99.7% (for $k = 3$). If a coverage factor is used, there should be a clear explanation of its meaning so there is no confusion for readers interpreting the significance of the uncertainty value.

Be aware that the $\pm$ uncertainty notation might be used to indicate different confidence intervals, depending on the scientific discipline or context. For example, a public opinion poll may report that the results have a *margin of error* of $\pm 3\%$, which means that readers can be 95% confident (not 68% confident) that the reported results are accurate within 3 percentage points. Similarly, a manufacturer’s *tolerance* rating generally assumes a 95% or 99% level of confidence.
**Conclusion: “When do measurements agree with each other?”**

We now have the resources to answer the fundamental scientific question that was asked at the beginning of this error analysis discussion: “Does my result agree with a theoretical prediction or results from other experiments?”

Generally speaking, a measured result agrees with a theoretical prediction if the prediction lies within the range of experimental uncertainty. Similarly, if two measured values have *standard uncertainty* ranges that overlap, then the measurements are said to be *consistent* (they agree). If the uncertainty ranges do not overlap, then the measurements are said to be *discrepant* (they do not agree). However, you should recognize that these overlap criteria can give two opposite answers depending on the evaluation and confidence level of the uncertainty. It would be unethical to arbitrarily inflate the uncertainty range just to make a measurement agree with an expected value. A better procedure would be to discuss the size of the difference between the measured and expected values within the context of the uncertainty, and try to discover the source of the discrepancy if the difference is truly significant. Example:

\[
A = 1.2 \pm 0.4 \\
B = 1.8 \pm 0.4
\]

These measurements agree within their uncertainties, despite the fact that the percent difference between their central values is 40%. In contrast, if the uncertainty is halved (±0.2), these same measurements do not agree since their uncertainties do not overlap.

If two measurements (with similar uncertainties, each represented by ±1 sigma) barely overlap, then they differ by approximately 2-sigma, which means there is about a 5% chance that the measurements agree. When two values differ by 3-sigma or more, it is highly unlikely (less than 1% chance) that they agree, and we
would conclude with confidence that there is a discrepancy. Further investigation would be needed to determine the cause for the discrepancy. Perhaps the uncertainties were underestimated, there may have been a systematic error that was not considered, or there may be a true difference between these values.

An alternative method for determining agreement between values is to calculate the difference between the values divided by their combined standard uncertainty. This ratio gives the number of standard deviations separating the two values. If this ratio is approximately 1 or smaller, then it is reasonable to conclude that the values agree. If the ratio is on the order of 2, or more, then it is unlikely (less than about 5% probability) that the values are the same.

Example from above with \( u = 0.4 \): \[ \frac{|1.2-1.8|}{0.57} = 1.1 \] A and B likely agree

Example from above with \( u = 0.2 \): \[ \frac{|1.2-1.8|}{0.28} = 2.1 \] A and B likely do not agree

Note the following example of overlap:

These two measurements agree, despite their large difference in error bars and the fact that the “red” measurement is outside of the “green” error bars. There is no requirement that either measurement be included in the other measurement’s error bars – only that the error bars of the two measurements overlap.
Making Graphs

When graphs are required in laboratory exercises, you will be instructed to “plot A vs. B” (where A and B are variables). By convention, A (the dependent variable) should be plotted along the vertical axis (ordinate), and B (the independent variable) should be plotted along the horizontal axis (abscissa). Graphs that are intended to provide numerical information should be drawn on ruled graph paper. Use a sharp pencil (not a pen) to draw graphs, so that mistakes can be corrected easily. It is acceptable to use a computer (see the Excel tutorial below) to produce graphs. The following graph is a typical example in which distance vs. time is plotted for a freely falling object. Examine this graph and note the following important rules for graphing:

**Figure 3. Plot of Distance vs Time**

**Title.** Every graph should have a title that clearly states which variables appear on the plot. If the graph is not attached to another identifying report, write your name and the date on the plot for convenient reference.

**Axis labels.** Each coordinate axis of a graph should be labeled with the word or symbol for the variable plotted along that axis and the units (in parentheses) in which the variable is plotted.

**Choice of Scale.** Scales should be chosen in such a way that data are easy to plot and easy to read. On coordinate paper, every 5th and/or 10th line should be selected as
major division lines that represent a decimal multiple of 1, 2, or 5 (e.g., 0.1, 2, 0.05, 20, 500, etc.) – other choices (e.g., 0.3) make it difficult to plot and also read data. Scales should be made no finer than the smallest increment on the measuring instrument from which data were obtained. For example, data from a meter stick (which has 1 mm graduations) should be plotted on a scale no finer than 1 division = 1 mm, because a scale finer than 1 div/mm would provide no additional plotting accuracy, since the data from the meter stick are only accurate to about 0.5 mm. Frequently the scale must be considerably coarser than this limit, in order to fit the entire plot onto a single sheet of graph paper. In the illustration above, scales have been chosen to give the graph a roughly square boundary; avoid choices of scale that make the axes very different in length. Note that it is not always necessary to include the origin (‘zero’) on a graph axis; in many cases, only the portion of the scale that covers the data need be plotted.

**Data Points.** Enter data points on a graph by placing a suitable symbol (e.g., a small dot with a small circle around the dot) at the coordinates of the point. If more than one set of data is to be shown on a single graph, use other symbols (e.g., Δ) to distinguish the data sets. If drawing by hand, a drafting template is useful for this purpose.

**Curves.** Draw a simple smooth curve through the data points. The curve will not necessarily pass through all the points, but should pass as close as possible to each point, with about half the points on each side of the curve; this curve is intended to guide the eye along the data points and to indicate the trend of the data. A French curve is useful for drawing curved line segments. *Do not connect the data points by straight-line segments in a dot-to-dot fashion.* This curve now indicates the average trend of the data, and any predicted (interpolated or extrapolated) values should be read from this curve rather than reverting back to the original data points.

**Straight-line Graphs.** In many of the exercises in this course, you will be asked to linearize your experimental results (plot the data in such a way that there is a linear, or straight-line relationship between graphed quantities). In these situations, you will be asked to fit a straight line to the data points and to determine the slope and y-intercept from the graph. In the example given above, it is expected that the falling object’s distance varies with time according to \( d = \frac{1}{2}gt^2 \). It is difficult to tell whether the data plotted in the first graph above agrees with this prediction; however, if \( d \) vs. \( t^2 \) is plotted, a straight line should be obtained with slope = \( \frac{1}{2}g \) and y-intercept = 0.
Using Excel for Data Analysis in Physics Labs

Students have a number of software options for analyzing lab data and generating graphs with the help of a computer. It is the student’s responsibility to ensure that the computational results are correct and consistent with the requirements stated in this lab manual. Any suitable software can be used to perform these analyses and generate tables and plots for lab reports and assignments; however, since Microsoft Excel is widely available on all CCI laptops and in university computer labs, students are encouraged to use this spreadsheet program. In addition, there may be assignments during the semester that specifically require an Excel (or platform-equivalent) spreadsheet to be submitted.

Getting Started

This tutorial will lead you through the steps to create a graph and perform linear regression analysis using an Excel spreadsheet. The techniques presented here can be used to analyze virtually any set of data you will encounter in your physics studio. The experiment is a measurement of times and positions of a cart moving at constant speed.

To begin, open Excel. A blank worksheet should appear. Enter the sample data from the experiment (listed in the table below) and column headings shown below into cells A1 through D6. Save the file to a disk or to your personal file space on the campus network.

<table>
<thead>
<tr>
<th>Time (sec)</th>
<th>Position (m)</th>
<th>Time ±</th>
<th>Position ±</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.64</td>
<td>1.15</td>
<td>0.05</td>
<td>0.20</td>
</tr>
<tr>
<td>1.10</td>
<td>2.35</td>
<td>0.07</td>
<td>0.30</td>
</tr>
<tr>
<td>1.95</td>
<td>3.35</td>
<td>0.05</td>
<td>0.20</td>
</tr>
<tr>
<td>2.45</td>
<td>4.46</td>
<td>0.06</td>
<td>0.40</td>
</tr>
<tr>
<td>2.85</td>
<td>5.65</td>
<td>0.10</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Note that the uncertainties for the time and position (denoted ±) have been included. These are not necessary for a basic plot, but the studio lab reports and assignments require an uncertainty analysis, so you should get into the habit of including them.
Creating and Editing a Graph

Note: these instructions may not be precisely correct for different versions of Excel or on different platforms; however, these instructions should be roughly correct for all modern versions of Excel on all platforms. Other spreadsheet programs should have similar features. We suggest you use on-line resources or consult with your classmates or your TA for specific questions or issues.

You will be creating a graph of these data whose finished form looks like this:

Follow these steps to accomplish this:

1: Use your mouse to select all the cells that contain the data that you want to graph (in this example, columns A and B). To graph these data, select Chart on the toolbar.

2: Click on Scatter plots and choose XY (Scatter) or Marked Scatter with no lines. A default plot should appear in the spreadsheet, and it should be both moveable and resizable.

3: Using the Chart Layout tool, experiment with setting the title, axes, axis titles, gridlines and legends. At a minimum, we require that the plot be titled and that the x- and y-axes are descriptively labeled with units. We strongly suggest that all gridlines and the legend be removed for clarity.
Most graph features can be modified by double-clicking on the feature you want to change. You can also right-click on a feature to get a menu. Try changing the color of the plot area, the numbers on the axes, and the appearance of the data points. It is recommended that you always format the background area to white using the “Automatic” option.

**Adding Error Bars**

Right-click on a datum point and choose **Format Data Series...** and select **Error Bars**. Click on the appropriate **Error Bar tab** (X or Y) and choose **Both** under **Display** and **Cap** for the **End style**. Fixed values or percentages can be set, for example, but if you have separate columns of uncertainty values for each datum, as shown above, then select **Custom** to specify the values. In the subsequent custom error bar window, select the positive error value field and then click and drag in the corresponding Excel column of uncertainties. Repeat for the negative value and click OK. Your custom error bars will then be applied. Repeat for the other axis. Note that if you create separate columns for the positive and negative error bars, they can be set independently. Also note that error bars may not be visible if they are smaller than the size of the datum point on the plot.

**Adding a Trendline**

The primary reason for graphing data is to examine the mathematical relationship between the two variables plotted on the x- and y-axes. To add a trendline and display its corresponding equation, right-click on any datum point and **Add Trendline**. Choose the graph shape that best fits your data and is consistent with your theoretical prediction (usually Linear). Click on the “Options” tab and check the boxes for "Display equation on chart" and "Display R-squared value on chart.” A good fit is indicated by an $R^2$ value close to 1.

**Caution:** When searching for a mathematical model that explains your data, it is very easy to use the trendline tool to produce nonsense. This tool should be used to find the simplest mathematical model that explains the relationship between the two variables you are graphing. Look at the equation and shape of the trendline critically:

- Does it make sense in terms of the physical principle you are investigating?
- Is this the best possible explanation for the relationship between the two variables?

Use the simplest equation that passes through most of the error bars on your graph. You may need to try a couple of trendlines before you get the most appropriate one. To clear a trendline, right-click on its regression line and select **Clear**.
Determining the Uncertainty in Slope and Y-intercept

The $R^2$ value indicates the quality of the least-squares fit, but this value does not give the error in the slope directly. Given the best fit line $y = mx + b$, with $n$ data points, the standard error (uncertainty) in the slope $m$ can be determined from the $R^2$ value by using the following formula:

$$\sigma_m = m \sqrt{\frac{(1/R^2) - 1}{n-2}}$$  \hspace{1cm} (6)

Likewise, the uncertainty in the y-intercept $b$ is:

$$\sigma_b = \sigma_m \sqrt{\frac{\sum x^2}{n}}$$  \hspace{1cm} (7)

These values can be computed directly in Excel or by using a calculator. For this sample set of data, $\sigma_m = 0.1684$ m/s, and $\sigma_b = 0.33$ m. Note that a value of $R^2$ of exactly 1 leads to slope and intercept uncertainties of zero. Carefully exam the Excel $R^2$ value -- although it may display as exactly 1, it likely is not exactly 1. If your value is indeed exactly 1, it indicates an error in how you have plotted your data.

The uncertainty in the slope and y-intercept can also be found by using the LINEST function in Excel. Using this function is somewhat tedious and is best understood from the Help feature in Excel.

Interpreting the Results

Once a regression line has been found, the equation must be interpreted in terms of the context of the situation being analyzed. This sample data set came from a cart moving along a track. We can see that the cart was moving at nearly a constant speed since the data points tend to lie in a straight line and do not curve up or down. The speed of the cart is simply the slope of the regression line, and its uncertainty is found from the equation above: $v = 1.9 \pm 0.2$ m/s. (Note: If we had plotted a graph of time versus distance, then the speed would be the inverse of the slope: $v = 1/m$) The y-intercept gives us the initial position of the cart: $x_0 = -0.0035 \pm 0.33$ m, which is essentially zero.

Final Step – Copying Data and Graphs into a Word Document

Copy your plots and data table from Excel to Word. Just select the graph (or cells) and use the Edit menu or keyboard shortcuts to copy and paste.
Using LINEST in Excel

LINEST is an alternative linear least square fitting function in Excel. The results of the LINEST analysis are virtually identical to the linear trendline analysis described above; however, LINEST provides a single-step calculation of both the slope and intercept uncertainties, instead of the multi-step procedure described above.

- Start with a table for time and velocity (right).
- The LINEST function returns several outputs; to prepare, select a 2 by 5 array below the data, as shown.
- Note, velocity was mistakenly labeled as having units of m/sec in the table to the right.
- Under the Insert menu, selection Function, then Statistical, and finally LINEST as shown below (and hit OK).
Select the y-values and x-values from the table and write ‘TRUE’ for the last two options (below). Again, press OK.
The LINEST function is an array function; therefore, you must tell Excel you are done with the array. Highlight the formula in the formula bar as shown below. Press Ctrl+Shift simultaneously with Enter (Mac users, press Command-Enter or Shift-Control-Enter depending on your version of Excel)
The result is that the array of blank cells that was selected previously is now filled with data (although the new data isn’t labeled). See the image below for the labels that can be added after the fact (e.g., “Slope,” “R^2 Value,” etc):

![Table 1: Time and Velocity Data]

<table>
<thead>
<tr>
<th>time (m)</th>
<th>uncertainty in time</th>
<th>velocity (m/sec)</th>
<th>uncertainty in velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>0.1</td>
<td>5.5</td>
<td>0.3</td>
</tr>
<tr>
<td>3.7</td>
<td>0.3</td>
<td>7.7</td>
<td>0.2</td>
</tr>
<tr>
<td>4.9</td>
<td>0.2</td>
<td>9.8</td>
<td>0.7</td>
</tr>
<tr>
<td>6.9</td>
<td>0.4</td>
<td>12.3</td>
<td>0.9</td>
</tr>
<tr>
<td>8.3</td>
<td>0.6</td>
<td>15.7</td>
<td>0.2</td>
</tr>
<tr>
<td>11.5</td>
<td>0.9</td>
<td>16.2</td>
<td>0.3</td>
</tr>
<tr>
<td>13.3</td>
<td>0.8</td>
<td>17.2</td>
<td>0.6</td>
</tr>
<tr>
<td>17.9</td>
<td>0.3</td>
<td>18.9</td>
<td>0.6</td>
</tr>
<tr>
<td>20.3</td>
<td>0.6</td>
<td>25.3</td>
<td>0.9</td>
</tr>
</tbody>
</table>

![Using Linest function:]

- Slope: 0.937876188
- Slope Uncertainty: 0.102847501
- R^2 Value: 0.922358443
- Y-intercept: 4.983072931
- Y-intercept uncertainty: 1.188614381

**References:**


Appendix I: Propagation of Uncertainty by Quadrature

Consider a quantity \( f \) to be calculated by multiplying two measured quantities \( x \) and \( y \) whose uncertainties are \( \sigma_x \) and \( \sigma_y \), respectively. The question is how to propagate those uncertainties to the calculated quantity \( f \). From the chain rule of calculus, the change in \( f \) due to changes in \( x \) and \( y \) is:

\[
\delta f = \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y
\]

Squaring and averaging yields:

\[
(\delta f)^2 = \left( \frac{\partial f}{\partial x} \right)^2 (\delta x)^2 + \left( \frac{\partial f}{\partial y} \right)^2 (\delta y)^2 + 2 \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \delta x \delta y
\]

For uncorrelated measurements, \( \delta x \delta y \) is zero. Consider the average square change in quantities to be the uncertainty in each of \( x, y, \) and \( f \); that is, \( (\delta f)^2 = \sigma_f^2 \), etc. Then:

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

To generalize it to arbitrary powers of \( x \) and \( y \), consider the function \( f = x^n y^m \); substituting this into the last equation and dividing by \( f \) yields the relative uncertainty:

\[
\frac{\sigma_f}{f} = \sqrt{\left( \frac{\partial f}{\partial x} \right)^2 \frac{\sigma_x^2}{x^n y^m} + \left( \frac{\partial f}{\partial y} \right)^2 \frac{\sigma_y^2}{x^n y^m}}
\]

The partial derivatives are \( \frac{\partial f}{\partial x} = nx^{n-1} y^m \) and \( \frac{\partial f}{\partial y} = mx^n y^{m-1} \). Substituting these yields:

\[
\frac{\sigma_f}{f} = \sqrt{n^2 \left( \frac{\sigma_x}{x} \right)^2 + m^2 \left( \frac{\sigma_y}{y} \right)^2}
\]

This expression looks complicated but it simplifies to something rather simple:

\[
\frac{\sigma_f}{f} = \sqrt{n^2 \left( \frac{\sigma_x}{x} \right)^2 + m^2 \left( \frac{\sigma_y}{y} \right)^2}
\]
The result is that the relative (fractional) uncertainty in $f$ is the root of the squared sum (RSS) of individual uncertainties in $x$ and $y$. Examples in the body of this document include:

$$f = xy \quad \frac{\sigma_f}{f} = \sqrt{(1)^2 \left( \frac{\sigma_x}{x} \right)^2 + (1)^2 \left( \frac{\sigma_y}{y} \right)^2}$$

$$f = x/y \quad \frac{\sigma_f}{f} = \sqrt{(1)^2 \left( \frac{\sigma_x}{x} \right)^2 + (-1)^2 \left( \frac{\sigma_y}{y} \right)^2}$$

$$f = xy^2 \quad \frac{\sigma_f}{f} = \sqrt{\left( \frac{\sigma_x}{x} \right)^2 + 4 \left( \frac{\sigma_y}{y} \right)^2}$$

Note that the result for multiplication and division is the same (division is just a power law with a negative exponent). Also note that variables that appear with a higher power are weighted more heavily in the propagation.

For some functions, especially non-linear trig functions, you may have to perform the derivatives to find how the uncertainty propagates; however, for many functions, performing the derivatives each time is not required – merely apply the equation highlighted in yellow above.
Appendix II. Practical Tips and Examples for Measuring and Citing Uncertainty

I. Instrument Precision and Accuracy

Instrument precision is the smallest reading an instrument can provide. For some instruments, this precision is unambiguous; for example, a digital balance whose readout is displayed as __. __ g has an instrument precision of 0.1 g. In contrast, a meter stick’s instrument precision is open to interpretation. If the finest mark on the meter stick is 1 mm, it could be claimed that the instrument precision is 0.001 m, or 1 mm; however, a different observer might claim that it is possible to interpolate between the finest marks, and therefore might claim that the instrument precision is 0.0005 m, or 0.5 mm. In all cases, the choice of instrument precision and the reasoning should be included in any report that incorporates that instrument’s use.

In contrast to the instrument precision, the instrument accuracy refers to the ability to correctly measure the value of a known standard. If a digital balance with a measurement precision of 0.1 g is used to measure a weight independently known to have a mass of exactly 500.0 g, and the result is 495.3 g, the accuracy is no better than about 5 g, even though the precision is 0.1 g. Several strategies for addressing this discrepancy are listed below:

a) Cite the value as 495.3 ± 4.7 g. Unacceptable: citing the uncertainty to 2 sig figs is not warranted for less than 50 measurements (Table 2).

b) Cite the value as 495.3 ± 5 g. Unacceptable: citing the measurement to a higher precision than the uncertainty is not warranted without significant justification.

c) Cite the value as 495 ± 5 g. Acceptable but not optimal: the measurement is clearly incorrect because it is known that the weight is exactly 500 g. If only a single measurement is possible, citing as shown is allowed, but should be accompanied by a discussion about the known discrepancy and the introduction of a systematic error.

d) Cite the value as 495.3 ± 0.1 g. Allowed but only if a known calibration standard or other information about the accuracy is unavailable: the calibration status of the instrument should be consulted. If this is not possible, it is not unreasonable to assume that the instrument has an implied accuracy that matches its precision, as shown in this example.

Item d) is the most common scenario in Physics 118 activities. Any such assumption about substituting precision for accuracy should be included in a discussion, with an acknowledgment that a systematic error may have been introduced. The recommended procedure to address this issue is to acquire multiple “identical” samples of the “known” weight, and compute an average value and standard deviation. This would resolve the question of whether citing the precision as an uncertainty is warranted; however, it would not resolve the accuracy question.
best approach to the accuracy question is a calculation of the standard error, but it is extremely important to remove any systematic uncertainties—any measurements performed without known calibration standards available are susceptible to systematic effects, and should be cited with the combined (Type A and Type B) standard uncertainty.

II. Stacking Example 1

Assume you are asked to measure the mass of a typical penny (according to the US Mint, currently made pennies have a nominal mass of 2.5 g) with a scale whose accuracy is known to be ±0.2 g but reads to a precision of 0.1 g. One can immediately discard the choice of 0.1 g as the accuracy. Measuring one penny might yield a measurement of 2.4 ± 0.2 g, and this would be the only measurement possible for that one penny. Likewise, another penny might yield a measurement of 2.5 ± 0.2 g.

Is there a way to get a more precise measurement? In this case, yes, because you are asked to find the mass of a typical penny. By stacking pennies and measuring more than one of them at the same time, dividing by the number of pennies measured can provide a more precise answer. For example, assume that you measure the mass value of five pennies separately, with these results (all with an accuracy of ±0.2 g): 2.4, 2.4, 2.5, 2.4, 2.6. The relative uncertainty of each measurement is about 8%.

Further assume that when you measure all five at the same time, the value is 12.3 ± 0.2 g, yielding a relative uncertainty of about 2% for the stack. The mean value for a typical penny is therefore (retaining guard digits) 2.460 g. But what do we assign as the uncertainty? One might argue that the uncertainty is still 0.2 g; however, we can plausibly rewrite the sum as:

\[
\text{sum} = (2.4 \pm 0.04) + (2.4 \pm 0.04) + (2.5 \pm 0.04) + (2.4 \pm 0.04) + (2.6 \pm 0.04) = 12.3 \pm 0.2 \text{ g}
\]

This is demonstrably the same as the stacked value of 12.3 ± 0.2 g, and suggests that, based on the upper-lower bound method, the uncertainty can also be divided by five. Therefore, as a general rule, we can reasonably divide both the stacked value and its uncertainty by \( N \), and assert the value for a typical penny as 2.46 ± 0.04 g. Note that as per Table 2, the uncertainty is cited to only one sig fig, because only five measurements were made.

Assume instead that the accuracy of the scale is unknown; a working proposition is that the scale’s mechanism will correctly report that an object of 0.1 g is somewhere between 0.05 g and 0.15 g. The reading of 12.3 g for all 5 pennies together could therefore have values falling between 12.25 and 12.35 g. Dividing this by 5 yields a nominal value of 2.460 g with a high and low of 2.470 g and 2.450 g, respectively. The value of 2.460 ± 0.010 g could then be cited as 2.46 ± 0.01 g to one sig fig; however, although the results seems “better,” there is no confidence that this value is correct. In measurement cases were the accuracy of the measuring device is not known, the result should explicitly cite this discrepancy.
III. Stacking Example 2

As another example, assume that a student is asked to measure the period of a pendulum and discuss any systematic effects that could cause the measurement to be inaccurate. The student responds: “I held the pendulum bob still at an angle of 45°, and started a stopwatch at the moment I let go. I waited for the pendulum to swing back and then stopped the stopwatch at the moment the bob reached the starting position. The reading on the stopwatch is the best measurement of the period of the pendulum of this length; however, friction is a systematic effect that could cause the measurement to be inaccurate.” Critique this response.

This measurement can be improved greatly by stacking. In this example, “stacking” means to allow the pendulum to swing for multiple periods, and divide by the number of periods. Despite this improvement, multiple issues remain with such a measurement. In particular, the use of the term “friction” is too vague to be instructive. In the introductory physics course, “friction” is reserved to mean either a) two surfaces in contact and sliding against each other (kinematic friction) or b) the resistance between two surfaces in contact to begin sliding (static friction). Neither of these apply to the pendulum bob. Instead, the bob may be slowed down by drag, which is a more appropriate term than “friction” when fluids (e.g. air) are present – the physics of drag and the physics of kinematic sliding friction are quite different; therefore, making the distinction is a superior answer that will be rewarded with credit. In the case of drag, stacking may result in a worse answer for the period, because as more swings are counted, drag effects multiply. Finally, we note that the simple pendulum is a non-linear device – the harmonic oscillator period for a simple pendulum applies only for small angles (good to about 1% at angles of less than 10°). For angles such as 45°, there may be poor agreement between the measured and theoretical periods unless non-linearity is taken into account.
IV. Approach to Uncertainty When Accuracy is Cited as a Percentage.

Instrument accuracy may be described in terms of a percentage, such as “± 0.1%.” For high quality instruments, this usually means 0.1% of the full-scale (FS) reading, across the entire measurement range of the instrument. For example, if a digital balance can measure a maximum value of 199.9 g, the accuracy would be 0.1% of FS (200 g), or ± 2 g (worse than the instrument precision is 0.1 g). In this case, follow the guidelines in Example I.

However, the origin of the percentage uncertainty may be ambiguous, and could imply a percentage of the current reading. If the interpretation is ambiguous, it can be evaluated by comparing to the precision of the instrument. As an example, assume it is claimed that the accuracy of the balance is 0.1% for any mass, and that the measurement of one object reads 53.4 g. The percent accuracy of 0.1% implies and accuracy of ±0.05 g, which is better than the instrument precision of 0.1 g! This inconsistency suggests that a “stack of one” item is inappropriate. Stacking should be used, but it cannot be applied arbitrarily; a stack of two would not unambiguously resolve the question, but a stack of three would.

The general approach in this case is to stack enough objects so that the percentage accuracy exceeds the precision of the measurement. The approach in Example I can then be used; however, under all circumstances, any assertion of an uncertainty beyond 1 sig fig is only justified for N=50 or above (Table 2).

V. Use of Containers

Assume you are given a quantity of water and you are asked to measure its weight using a scale. You need a container, but the container itself has a weight. The procedure is to weigh the container by itself, and then the container and the water together, and subtract to get the weight of the water. However, in this case, the uncertainties of the two measurements must be added. This should be self-evident from an upper-lower bound perspective; the quadrature analysis is a more formal statement of the same result (see Quadrature section).
VI. Explaining Discrepancies

A simple pendulum is known to have a period of oscillation $T = 1.55$ s. Student A uses a digital stopwatch to measure the total time for 5 oscillations and calculates an average period $T = 1.25$ s. Student B uses an analog wristwatch and the same procedure to calculate an average period for the 5 oscillations and finds $T = 1.6$ s.

- Which student made the more accurate measurement? *The measurement made by Student B is closer to the known value and is therefore more accurate.*
- Which measurement is more precise? *The measurement made by Student A is reported with more digits and is therefore more precise.*
- What is the most likely source of the discrepancy between the results? Although the difference in period measurements is only 0.3 s, the original timing measurements may have differed by 5 times this amount since we are told that the average period was calculated from the total time for 5 oscillations. So even though reaction time (typically ~0.2 s) is a likely source of systematic error, this would not explain the discrepancy of approximately 1.5 s, or about one period. Therefore, the most likely source of the discrepancy is that Student A mistakenly measured only 4 oscillations instead of 5. This example shows that a measurement with greater precision is not always more accurate.

VII. Propagating Uncertainty and Reporting Values

A student uses a protractor to measure an angle of $\theta = 85^\circ \pm 1^\circ$. What should she report for $\sin \theta$?

$$\sin(\theta) = 0.996 \pm 0.002.$$ *The uncertainty can be determined from either the upper/lower bound method or propagation of uncertainty using partial derivatives. Both methods yield the same result when rounded to one significant figure. Note that it would be inappropriate to round this further to $1.00 \pm 0.00$, because it implies an uncertainty of zero.*
VIII. Outliers

A group of students are told to use a meter stick to find the length of a hallway. They take 6 independent measurements as follows: 3.314 m, 3.225 m, 3.332 m, 3.875 m, 3.374 m, 3.285 m. How should they report their findings, and why?

No data should be excluded when presenting the raw numbers. Therefore, all values should appear in any data tables, and all values should appear on plots. An example is shown below for these data:

The measurement of 3.875 m is an outlier (10 standard deviations from the other values) and is most likely a mis-measurement (see 2nd bullet point below). In this case, the best estimate would be the average and standard error for the 5 “good” values (3.31 ± 0.02 m). Note two aspects of this approach:

• If a systematic error is suspected, then using the standard error may also be suspect. The standard deviation is 0.06 m; therefore, a more conservative answer would be 3.3 ± 0.1 m.

• Be careful before throwing out data – Nobel Prizes have been won because of outliers (although probably not in 118)! Also think about the experiment carefully: in this example, there is likely to be a misused meter stick, but in the case of gold prospecting, only the higher outliers are kept, and everything ELSE is thrown out!

IX. Limiting Factors

A Vernier caliper (an instrument that can routinely measure lengths to 0.001 inch, or less than 0.05 mm), is used to measure the radius of a tennis ball. The value is presented as \( R = 3.2 \pm 0.1 \) cm (half of \( D = 6.4 \pm 0.2 \) cm). Why is this result not more precise?

The uncertainty of this measurement is determined not by the resolution of the caliper, but by the imprecise definition of the ball’s diameter (it’s fuzzy!).
X. Uncertainty Propagation I

Equipment is used to measure the acceleration of a glider on an inclined air track as accurately as possible. The distance between two photogates is measured very accurately as $d = 1.500 \pm 0.005$ m. The experiment is run 5 times with the glider starting from rest at the first photogate, and the photogate times are recorded as $t$(in sec) = 1.12, 1.06, 1.19, 1.15, 1.08. What is the best estimate of the acceleration of the glider?

The average time is 1.12 s with a standard error of 0.02 s. The relationship between the distance and acceleration is $d = \frac{1}{2}at^2$; the best estimate for the acceleration is therefore $a = 2.392$ m/s$^2$ (retaining guard digits). The quadrature result for the uncertainty in the acceleration is:

$$\frac{\sigma_a}{a} = \sqrt{\left(\frac{\sigma_d}{d}\right)^2 + 4\left(\frac{\sigma_t}{t}\right)^2}$$

The relative uncertainty in the time is more than 5x the relative uncertainty in the position. In addition, the quadratic dependence on time creates an additional factor of 4, so the uncertainty in time dominates, and the uncertainty in the distance can be neglected. Therefore:

$$\sigma_a = 2a\left(\frac{\sigma_t}{t}\right) = 2(2.392 \text{ m/s}^2)\frac{0.02}{1.12} = 0.085 \text{ m/s}^2$$

The final result is $a = 2.39 \pm 0.09$ m/s$^2$, or $2.4 \pm 0.1$ m/s$^2$. 
XI. Uncertainty Propagation II

A student performs a simple experiment to find the average acceleration of a falling object. He drops a baseball from a building and uses a string and tape measure to measure the height from which the ball was dropped. He uses a stopwatch to find an average time of fall for 3 trials from the same height and reports the following data: $h = 6.75 \pm 0.33 \text{ m}$, $t = 1.14 \pm 0.04 \text{ s}$. Determine the average acceleration, and comment on the accuracy of the result (do you think any errors were made?). What suggestions would you make to improve the result?

Calculating the acceleration as for the previous example yields $a = 2h/t^2 = 10.387 \text{ m/s}^2$. In this case, neither relative uncertainty in time or distance dominates, so both should be included in a quadrature calculation:

$$\sigma_a = a \sqrt{\left(\frac{\sigma_h}{h}\right)^2 + 4\left(\frac{\sigma_t}{t}\right)^2} = \left(\frac{0.33}{6.75}\right)^2 + 4\left(\frac{0.04}{1.14}\right)^2 = 0.888 \text{ m/s}^2$$

The final answer would be written as $10.4 \pm 0.9 \text{ m/s}^2$. Note that this answer is higher than the value for $g$ in a vacuum. Although the uncertainty range includes the value $9.8 \text{ m/s}^2$, we would expect the value for acceleration to be at least somewhat or well below $g$ because of the effect of drag. Therefore, this result should be viewed with some suspicion. Suggestions include doing more trials and improving the accuracy of each measurement. Accuracy in the time is where the most gains can be made, because the relative uncertainty of the time is weighted twice that of the relative uncertainty of the distance.
XII. Uncertainty Propagation 3

A student uses a meter to measure the capacitance of two parallel plates that are 22 cm in diameter and separated by a piece of cardboard ($\kappa = 3$) that is 5 mm thick. What capacitance would you expect this student to measure? If the relative uncertainties in the diameter and plate separation and dielectric constant are 2%, 5%, and 10%, respectively, what uncertainty in the measurement would you cite? What suggestions would you give to reduce the effect of stray capacitance?

The nominal value would be:

$$C = \frac{Q}{V} = \kappa \varepsilon_0 \frac{A}{d} = \frac{3(8.85 \times 10^{-12} \text{ C}^2 \text{N}^{-1} \text{m}^{-2}) \pi (0.11 \text{ m})^2}{0.005 \text{ m}} = 202 \text{ pF}$$

$$\sigma_C = C \sqrt{4 \left(\frac{\sigma_\varepsilon}{\varepsilon_0}\right)^2 + \left(\frac{\sigma_d}{d}\right)^2 + \left(\frac{\sigma_\kappa}{\kappa}\right)^2} = (202 \text{ pF}) \sqrt{4 \left(0.02\right)^2 + \left(0.05\right)^2 + \left(0.10\right)^2} = 23.99 \text{ pF}$$

Note that since the area $A$ depends on the diameter squared, the uncertainty in the diameter has a higher weight than the other variables and cannot be neglected in the uncertainty calculation, even though it is significantly smaller. The answer should be cited as $202 \pm 24 \text{ pF}$.

To reduce stray capacitance, minimize the length of wire leads used in the measurement process. If possible, zero the meter with the leads in place before making the measurement.
XIII. Uncertainty from Linear Regression

In an investigation to empirically determine the value of $\pi$, a student measures the circumference and diameter of 5 circles of varying size, and uses Excel to make a linear plot of circumference versus diameter (both in units of meters). A linear regression fit yields the result of $y = 3.1527x - 0.0502$, with $R^2 = 0.9967$ for the 5 data points plotted. How should this student report the final result? Does the empirical ratio of $C/D$ agree with the accepted value of $\pi$?

The theoretical result is $C = \pi D + 0$. By inspection, the experimental value of $\pi$ is given by the slope of the regression, or 3.1527. The uncertainty in this value is:

$$\sigma_m = m \sqrt{\frac{1}{R^2} - 1} = (3.1527) \sqrt{\frac{1}{0.9967} - 1} = 0.1047$$

The final experimental value for $\pi$ is then $3.15 \pm 0.10$, or $3.2 \pm 0.1$. Note that the experimental intercept is not zero, but not enough information has been given to calculate the intercept’s uncertainty. The value for $\pi$ can therefore only be trusted to the extent that $-0.0502$ (the intercept) is experimentally equivalent to zero. Therefore, the uncertainty in the intercept must be larger than 0.0502; if not, there is more discussion to be written!
XIV. Analyzing the Effect of Multiple Variables

A student uses a photogate to determine the speed of a cart on a track. A “flag” on top of the cart has length \( d \) and runs in the direction of travel (parallel to the length of the cart). The flag is positioned to block the photogate beam as the cart crosses. When the beam is unblocked after the flag move on, the amount of time \( \Delta t \) that has elapsed is recorded. Since \( d \) and \( \Delta t \) are assumed to be accurate, an accurate calculation of the cart’s speed can be made.

Normally the flag is oriented exactly in line with the cart and parallel to the track as described, but instead suppose that the flag is tilted (as viewed from above). As viewed by the photogate, the flag will appear to be shorter than it really is. Suppose that an experimenter does not realize that the flag on the cart is angled away from the path of motion. Suppose further that the photogate timer is inaccurate. If the flag is actually tilted at an angle of 30° and the timer actually measures 10% less time than it should, what relative error in speed will result from these two systematic effects? Ignore friction, rolling resistance, and drag in this problem.

Calculating the relative error in the flag length is not the correct solution. As far as the student knows, the flag is correctly oriented. A flag error results in an incorrect time measurement, and the incorrectly working timer further compounds this error. The correct speed, using the variables that are measured, is simply \( v = \frac{d}{\Delta t} \), where \( d \) is the flag length and \( \Delta t \) is the time that the photogate is blocked. The incorrect speed arises entirely from erroneous time measurements; the measured time is shortened by a factor of \( \cos 30° \) and then is shortened again by a separate factor \( f \) by which the timer is defective:

\[
\text{incorrect speed} = \frac{d}{f \cdot \cos 30° \cdot \Delta t}
\]

The relative error is given by equation 2:

\[
\text{relative error} = \frac{\text{measured} - \text{expected}}{\text{expected}}
\]

An acceptable solution on a practicum, for \( \cos(30°) = 0.866 \) and \( f = 0.9 \), is:

\[
\text{Relative speed error} = \frac{\frac{d}{f \cdot \cos 30° \cdot \Delta t} - \frac{d}{\Delta t}}{\frac{d}{\Delta t}} = \frac{1}{f \cdot \cos 30°} - 1 = +28.3\% \text{ (too fast)}
\]
Discussion

The time measured by a photogate for a flag of length $d$ moving at speed $v$ is $t = \frac{(dcos\theta)}{v}$, where $\theta$ is the flag angle measured from the line of travel: for any flag length $d$, the time measured by the photogate approaches zero as $\theta$ approaches $90^\circ$. If the photogate itself measures incorrectly by a factor of $f$, then the time measured is $t = f\left(\frac{dcos\theta}{v}\right)$. If the photogate is measuring 10% less than it should, then $f = 0.9$. We now calculate the relative time error:

$$\text{Relative time error} = \frac{t_{\text{meas}} - t_{\text{expected}}}{t_{\text{expected}}} = \frac{f}{v} \frac{d}{v} \frac{cos30^\circ}{1} = 0.9 \cos30^\circ - 1 = -0.22058$$

The photogate measures a time that is 22.1% too low in total. We now ask, for the speed in question (whatever it is), what the relative error in that speed is for the actual length of the flag, which is both measured and expected to be $d$:

$$\frac{v_{\text{meas}} - v_{\text{expected}}}{v_{\text{expected}}} = \frac{d_{\text{meas}} - d_{\text{expected}}}{d_{\text{expected}}} = \frac{d}{0.77942 \frac{d_{\text{expected}}}{t_{\text{expected}}}} - \frac{d}{t_{\text{expected}}} = \frac{1}{0.77942} - 1 = +0.283$$

This is the same answer as before.

There are a variety of incorrect approaches to this question. We list below the most common errors that have been observed on the practicum:

a. Flag length relative error: $\frac{d_{\cos30^\circ} - d}{d} = \cos30^\circ - 1 = -13.4\% (-13\% \text{ to } 2 \text{ sig figs})$

   Note the wrong sign – the flag indeed appears shorter (negative) but the answer should result in higher speed.

b. Missing the time measurement error: $\frac{d}{\cos30^\circ} \frac{\Delta t}{\Delta t} = \frac{1}{\cos30^\circ} - 1 = +15.5\%$

c. Time measurement error: $\frac{d}{1.1 \cos30^\circ} \frac{\Delta t}{\Delta t} = \frac{1}{1.1 \cos30^\circ} - 1 = +4.97\%$

d. Time measurement error: $\frac{0.9 \Delta t}{\cos30^\circ} \frac{d}{\Delta t} = 1.1 \cos30^\circ - 1 = -4.74\%$

   Note the wrong sign – answer should result in higher speed if the timer reads low.

e. incorrect: $\pm13\%$ error on flag length $\pm10\%$ error on timer = $\pm23\%$ or $\pm3\%$

f. incorrect: quadrature calculation $\sqrt{(13\%)^2 + (10\%)^2} = \pm16.4\%$
Measurements & Uncertainty Analysis

(e) and (f) are incorrect for the same reason. Errors cannot be summed except in cases of simple addition – quadrature must be invoked; however, even quadrature is incorrect, because this is NOT an uncertainty propagation question! This is an estimate of a one-sided systematic effect.

**Challenge question:** suppose the timer actually incorrectly measures 10% more time than it should – what relative error in speed results?

Answer: for a value of \( f = 1.1 \), the result is identical to example (c) in the common errors list above:

\[
\text{Relative speed error} = \frac{\frac{d}{1.1 \cos 30^\circ}}{\frac{d}{\Delta t}} = \frac{1}{1.1 \cos 30^\circ} - 1 = +4.97\% \text{ (too fast)}
\]

The take-home message here is two-fold. First, it can be difficult to distinguish between a careless mistake (“oh when you said 10% I thought it was 1.1 instead of 0.9”) and a systematic effect. An even more important observation is that measurements in an introductory physics class tend not to be very accurate – a 5% uncertainty is quite good in an intro class and could easily be dismissed as “probably just due to sloppy measurement technique.” Be careful to point out the possibility of systematic effects in your lab reports, and cite specific potential sources of systematic error in your lab report discussion.
XV. Interpretation of Plotted Data

Students connect a solenoid to a variable power supply. They then use a magnetic field sensor to measure the magnetic field strength along the central axis at the center of the solenoid. A plot of the measured magnetic field strength $B$ as a function of the emf $\varepsilon$ of the power supply is shown to the right (error bars not shown). Theoretically, the value of $B$ should be $\mu_0 n I$, where $n$ is the number of turns per unit length, and $I$ is the current.

One of the students notes that the horizontal component of the magnetic field of the Earth is $2.5 \times 10^{-5} \, \text{T}$. Is there evidence from the graph that the Earth’s magnetic field systematically affected their measurement? What can you say about the orientation of the solenoid?

*The theoretical equation for the magnetic field is $B = \mu_0 n \frac{\varepsilon}{R}$*

where $\varepsilon$ is the voltage (emf) applied to the solenoid and $R$ is the resistance of the coil. Therefore, the plot of $B$ vs $\varepsilon$ is expected to be linear, but there should be no magnetic field when no current is flowing through the solenoid (that is, the y-intercept should be zero on this graph). However, it is not difficult to see that the plot has an intercept at approximately the value of the Earth’s natural magnetic field, as shown in red.

*Since the Earth’s magnetic field presumably is parallel to the surface of the Earth and appears as a systemic offset in the graph, it is not unreasonable to conclude that the solenoid was detecting the Earth’s magnetic field and was thus oriented horizontally during this set of measurements.*