Measurement and Uncertainty Analysis Guide

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

Table of Contents

last updated Aug 03, 2022

if your copy dates from the SP22 semester or before, it is out of date!

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:

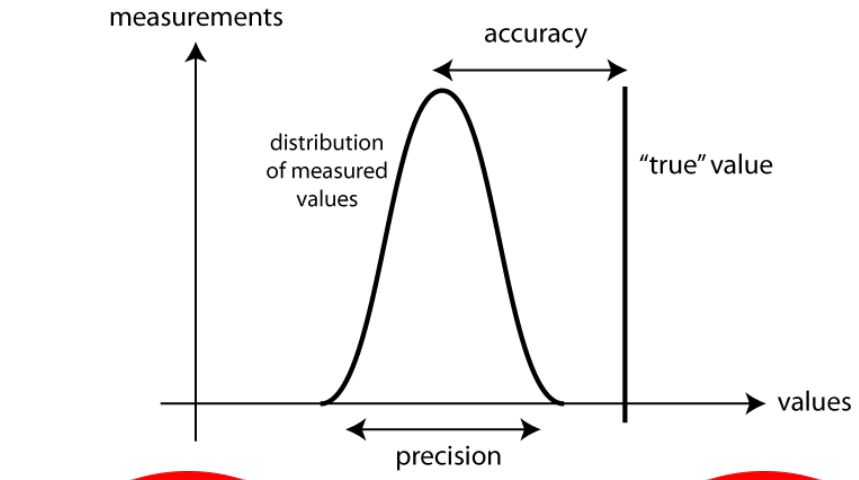
$$\text{measurement} = (\text{best estimate} \pm \text{uncertainty}) \text{ units}$$

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 ± 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:



high precision, low accuracy

low precision, high accuracy

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.

To determine the *accuracy* of a particular measurement, we must 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:

$$\text{Relative Uncertainty} = \left| \frac{\text{uncertainty}}{\text{measured quantity}} \right| \qquad (1)$$

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 ± 0.5 g has a relative uncertainty of

$$\frac{0.5g}{75.5g} = 0.00\bar{6} = 0.7\%$$

Relative Error

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

$$\text{Relative Error} = \frac{\text{measured value - expected value}}{\text{expected value}} \qquad (2)$$

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 many 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; 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 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 several 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.

Observation	Width (cm)
#1	31.33
#2	31.15
#3	31.26
#4	31.02
#5	31.20

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

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

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.

$$\text{Average Deviation, } \bar{d} = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + ... + |x_N - \bar{x}|}{N}$$

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**, aka **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:

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

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

Observation	Width (cm)	Deviation (cm)	
#1	31.33	+0.14	= 31.33 - 31.19
#2	31.15	-0.04	= 31.15 - 31.19
#3	31.26	+0.07	= 31.26 - 31.19
#4	31.02	-0.17	= 31.02 - 31.19
#5	31.20	+0.01	= 31.20 - 31.19

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

The *average* deviation for this example is $\bar{d}$ = 0.09 cm, whereas the *standard* deviation is:

$$s = \sqrt{\frac{(0.14)^2 + (0.04)^2 + (0.07)^2 + (0.17)^2 + (0.01)^2}{5-1}} = 0.12 \text{ cm}$$

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. 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.

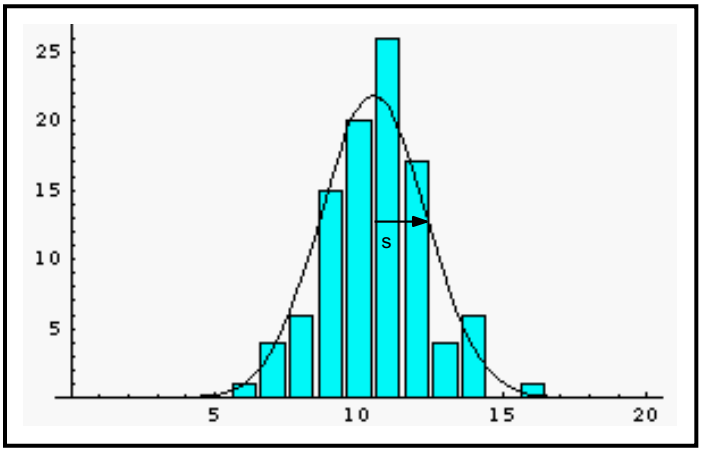


Figure 2. A Normal Distribution (Bell Curve) Based on 100 Measurements

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).

$$\text{Standard Deviation of the Mean, or Standard Error (SE)}, \quad \sigma_{\bar{x}} = \frac{s}{\sqrt{N}} \qquad (3)$$

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 ≈ 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 $\sqrt{N}$.

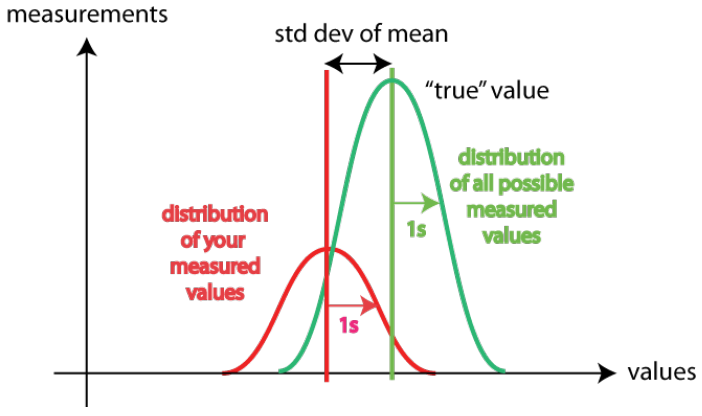


Figure 3. Standard Deviation vs Standard Error

In PHYS118 and 119, we often limit the number of measurements to save time. For a typical set of 3 measurements (instead of 5 or 10, for example), the normalization factor is $\sqrt{3} = 1.73$. Therefore, the difference between citing the uncertainty as the standard deviation or the standard deviation of the mean is less than a factor of 2. At this level of experimentation, this discrepancy is not considered significant – students may choose whichever calculation they are using to calculate the uncertainty as long as the method is cited.

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 several 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., 1.20×10^3 indicates 3 sig figs).

A calculator's 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 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 = 3×10^2 m^2.

Based on this example, the number of significant figures reported for a value can *imply* a degree of precision and can *suggest* a rough estimate of the relative uncertainty:

1 significant figure may suggest a relative uncertainty of about 10% to 100%
2 significant figures may suggest a relative uncertainty of about 1% to 10%
3 significant figures may 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 might suggest an uncertainty of ±1, or a fractional uncertainty of ±1/99 = ±1% (one could 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 isn't much point debating the factor of two). The smallest 2-significant-figure number, 10, also might suggest 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 better way to determine uncertainty. An explicit warning to this effect should accompany the use of this method in an exam or submitted lab work.

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	(2 significant figures)
	× 7328.7	(5 significant figures)
	$48369.42 = 48 \times 10^3$	(2 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	5560.5
	+54	+0.008
	278	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:

measured density = 8.93 ± 0.475328 g/cm^3 **INCORRECT!**

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 for a standard deviation to be reported to a precision of 3 significant figures, more than 10,000 readings would be required!*

N	Relative Uncertainty*	Significant Figures Valid	Implied Uncertainty
2	71%	1	±10% to 100%
3	50%	1	± 10% to 100%
4	41%	1	± 10% to 100%
5	35%	1	± 10% to 100%
10	24%	1	± 10% to 100%
20	16%	1	± 10% to 100%
30	13%	1	± 10% to 100%
50	10%	2	± 1% to 10%
100	7%	2	± 1% to 10%
10000	0.7%	3	±0.1% to 1%

Table 2. Valid Significant Figures in Uncertainties

*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:

measured density = 8.9 ± 0.5 g/cm^3 **CORRECT!**

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 many 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 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 presentation. 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

In this document and in the literature, uncertainties can be mathematically expressed in different styles – common choices include σ_x, σ_y, σ_z, σ_t, etc, or δ_x, δ_y, δ_z, δ_t, etc. You may encounter either style in this document. The symbol "Δ" can often be seen as well; however, we will reserve "Δ" for its tradition meaning of "difference between" as opposed to "small change in" or "uncertainty in."

Imagine that we have a result *f* that depends on other variables such as *x, y*, *z*, *t,* etc. (for example, *f* equals *x* divided by *y,* or *f* = some constant multiplied by *t* and added to *z*). We want to find the uncertainty in *f* if each of the measured quantities *x, y*, *z*, *t* has its own uncertainty. That is, we want to find out how the uncertainty in one set of variables (usually the independent variables *x, y*, *z*, *t*) propagates to the uncertainty in another set of variables (usually the dependent variable *f*). 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 but provides a better statistical estimate of the uncertainty. The Physics 118/119 courses focus on the upper-lower bound (ULB) estimate. The Physics 281L course focuses on quadrature; however, students in 118/119 are allowed if not encouraged to learn and use quadrature. In addition, quadrature forms the basis for *combining* uncertainties, which will be covered in an upcoming section of the guide.

The Upper-Lower Bound Method of Uncertainty Propagation

This method uses the uncertainties of each variable *x, y*, *z*, *t* to calculate the maximum and minimum values of the quantity *f*. You can also think of this procedure as examining the worst-case scenario. As a first example, consider the division of two variables – a common example is the calculation of average speed:

$$v_{avg} = \Delta x / \Delta t$$

Let's say an experiment is done repeatedly and measures a distance traveled of $\Delta x = 30 \pm 0.5$ m during a time of $\Delta t = 2 \pm 0.1$ sec. To find the upper and lower bound of v_{avg}, the uncertainties must be chosen to create the worst-case scenario for the uncertainty in v_{avg}; note that this choice requires making the numerator as large as possible and the denominator as small as possible for the upper bound (max), and *vice versa* for the lower bound (min):

$$v_{avg-\max} = \frac{30 + 0.5 \text{ m}}{2 - 0.1 \text{ sec}} = 16.05 \text{ m/s} \qquad v_{avg-\min} = \frac{30 - 0.5 \text{ m}}{2 + 0.1 \text{ sec}} = 14.05 \text{ m/s}$$

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 asymmetry is a typical outcome when using the upper-lower bound method). The uncertainty should be expressed as the *most conservative* value. Thus:

v_{avg} = 15.00 ± 1.05 m/s **PREFERRED!**

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

v_{avg} = 15.05 ± 1.00 m/s **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 the "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 ± 1.1 m/s, and then round again to 15 ± 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 $\theta = 25° \pm 1°$ and you need to find $f = \cos\theta$:

$f_{\min} = \cos(26°) = 0.8988$ $f = \cos(25°) = 0.9063$ $f_{\max} = \cos(24°) = 0.9135$

The differences are $f - f_{min} = 0.0075$ and $f_{max} - f = 0.0072$. 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 θ is associated with the minimum f, and vice versa
- guard digits are retained to make informed choices about the final answer
- although θ was only measured to 2 sig figs, f is known to 3 sig figs

Next consider the more general case where we measure something about a *functional* relationship between two variables such as speed and time. As a simple example, consider the kinematic function $v = v_0 + at$, where v is the speed of an object at time t when the object undergoes constant acceleration a and starts with an initial speed of v_0.

If we record multiple values of v and t, we can create a plot of those data and determine both the slope (a) and its uncertainty and the intercept (v_0) and its

uncertainty. Later in this document we will see how to find the slope and intercept, but here let's assume that this procedure finds a value of a = 30±2 cm/s^2 and v_0 = 2±0.5 cm/s. If we then ask what the expected speed *and its uncertainty* is for a value of t = 3 seconds, we can predict that

$$v_{upper} = (32 \text{ cm/s}^2)(3 \text{ s}) + 2.5 \text{ m/s} = 98.5 \text{ cm/s}$$
$$v = (30 \text{ cm/s}^2)(3 \text{ s}) + 2.0 \text{ m/s} = 92 \text{ cm/s}$$
$$v_{lower} = (28 \text{ cm/s}^2)(3 \text{ s}) + 1.5 \text{ m/s} = 85.5 \text{ cm/s}$$

Our final answer for the speed at 3 seconds (after rounding) would therefore be v = 92 ± 7 cm/s.

Note that we did not include an uncertainty for the 3-second mark – this is because the functional relationship maps inputs to outputs where the *model* (characterized by a slope and intercept) is uncertain. Our premise is that the speed can be modeled as $v = v_0 + at$, and our experimental procedure discovered the uncertainties in the model (namely, the uncertainties in v_0 and a). With those uncertainties being known, we have used the model to predict speeds at specific times; the specific times are theoretical and have no uncertainties per se. That said, if the question does involve any uncertainty in the time, this uncertainty can straightforwardly be included in the same ULB fashion to further determine the impact on the values of speed.

We can extend this idea to any linear relationship, which can be expressed as $y = mx + b$, where any of the variables in the equation could have some uncertainty. We wish to figure out what the uncertainty is in y when we know the uncertainties in m ($\pm\delta m$), x ($\pm\delta x$), and b ($\pm\delta b$). Applying the general ULB approach, we get a straightforward result:

$$y_{upper} = (m + \delta m)(x + \delta x) + (b + \delta b)$$
$$y_{lower} = (m - \delta m)(x - \delta x) + (b - \delta b)$$

Later you will learn how to find m, δm, b, and δb from linear regression techniques when given a set of data $\{x, y\}$. Once the values of m and b are known, they can be used to straightforwardly predict a value y_i and its uncertainty when given a value x_i and its uncertainty δx_i.

The ULB 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 (no uncertainty), while others may be uncertain; the range of these uncertain terms could be used to predict the upper and lower bounds on the total expense.

The Appendix includes a variety of example calculations of ULB uncertainties.

The Quadrature Method of Uncertainty Propagation

The ULB method described above has one significant shortcoming – it invariably overestimates the propagation of uncertainty in measurements by assuming a "worst-case scenario" where any single measurement can be at the extremes of expectations. But measurements are inherently "statistical" in that they are sometimes closer to expectations and sometimes further from expectations – in other words, they aren't *always* extremes, and therefore describing the propagation of uncertainty as an extreme isn't appropriate – we can do better. A variety of techniques can provide better estimates of true propagation of uncertainty – the quadrature method is one such statistically valid way to estimate the propagation of uncertainty. The quadrature method has several additional advantages over the ULB method:

- Quadrature is significantly easier to apply to complex models that obscure an easy application of the ULB method.
- Quadrature easily generalizes to multiple variables.
- Quadrature provides an approach to combining unrelated uncertainties. *This feature is the most important one for our purposes*.

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. Our approach uses the fundamental principles of calculus; the method is derived with several examples shown below.

For a single-variable function f (x), a small change in f (δf, known as the *differential*) can be related to the deviation in x as follows:

$$\delta f = \left(\frac{df}{dx} \right) \delta x$$

Taking the square and the average yields:

$$\overline{\delta f^2} = \left(\frac{df}{dx} \right)^2 \overline{\delta x^2}$$

We now define the uncertainty in f as the magnitude of this differential, and refer to it symbolically as σ:

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

Two examples where f is a power law are shown below (square root and power of two). We start with the target power-law function f (top), consider how small

changes in x affect small changes in f (middle), and finally end up with a relationship between the relative uncertainty in x and the relative uncertainty in f (bottom):

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

Note that the final normalization (division by f) makes it easy to express the *relative* (fractional) uncertainty in one variable (x) as a *relative* (fractional) uncertainty in another (f) – this is what we mean by "propagation." Note that the weighting is directly related to the power exponent of the function:

- For the square root, the relative uncertainty in f is half that in x.
- For the power of 2, the relative uncertainty in f is twice that in x.

These power law expressions are the most important ones to know for most modules of PHYS118/119. Now let us reconsider the trig example from the ULB section of the document above. Our goal is to propagate the uncertainty in θ to the uncertainty in $\cos\theta$. We start with the target function f (left), consider how small changes in θ affect small changes in f (middle), and finally end up with a relationship between the relative uncertainty in theta and the relative uncertainty in $\cos\theta$ (right).

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

Note that in this situation, σ_θ *must* be in radians. For $\theta = 25° \pm 1°$ (0.727 ± 0.017 rad)

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

This is essentially the same result as ULB method. The *fractional* uncertainty follows immediately as:

$$\frac{\sigma_f}{f} = \lvert\tan\theta\rvert\sigma_\theta$$

These examples are straightforward and are little different than applying the ULB method. The deeper power of the quadrature method becomes evident in the case where f depends on two or more variables; the derivation above can be repeated with minor modifications. For two variables, $f(x, y)$:

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

Here we have introduced the "partial derivative." The partial derivative $\partial f/\partial x$ 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 (for two variables):

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

If the measurements of x and y are *uncorrelated*, then $\overline{\delta x \delta y} = 0$, and this reduces to its most commonly cited form:

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

This form is known either as the "root sum of squares" (RSS) or "adding in quadrature" and is the basis for comprehensive approaches to combined uncertainty propagation.

Addition and Subtraction Example: $f = x \pm y$

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

The independence of the operation (doesn't matter whether the function is addition or subtraction) is a significant simplification compared to ULB.

Multiplication example: $f = xy$

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

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:

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

Note that the relative uncertainty in f has the same form for multiplication and division: the relative uncertainty in a product or quotient depends only on the *relative* uncertainty of each individual term. *Just as for addition and subtraction, this independence of the operation (doesn't matter whether the function is division or multiplication) is a significant simplification compared to ULB.*

Note also that adding or subtracting a *fixed* constant does not change the absolute uncertainty of the calculated value because, by definition, a constant value has no uncertainty. Likewise, multiplying or dividing by a constant does not change the relative uncertainty of the calculated value.

The results for the four operations of addition, subtraction, multiplication, and division are summarized in the following table, when the measurements are *uncorrelated* (independent):

Operation	Final Propagated Uncertainty
Addition/Subtraction	Absolute propagated uncertainty is quadrature addition of individual absolute uncertainties
Multiplication/Division	Relative propagated uncertainty is quadrature addition of individual relative uncertainties

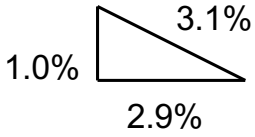
If the measurements are *correlated*, the absolute or relative uncertainties are straightforwardly added (not in quadrature). The result is always larger than adding in quadrature. Correlated measurements are considered in higher level physics courses and will not be addressed in PHYS118/119. *The quadrature propagation equation for arbitrary power laws is derived in Appendix I.*

As a more concrete 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.

Combining and Reporting Uncertainties

In general, uncertainties come in two different forms:

1. **Type A Uncertainties** are ***randomly distributed***, due to normal statistical fluctuations in a *series of observations*. These uncertainties are quantified using statistical techniques such as the standard deviation of the mean. Examples include:
    a. Electronic noise in the sensors used for measurement.
    b. Random variability in measurement (e.g., human reaction time when multiple observers use multiple stopwatches).
2. **Type B Uncertainties** are ***systematic***, due to some non-statistical bias that may or may not change as observations are made. These uncertainties cannot be quantified using statistical techniques and often depend on "available knowledge." Examples include:
    a. Offsets in sensors used for measurement (e.g., parallax when reading a meter stick).
    b. Human reaction time (e.g., multiple observers with multiple stopwatches *all of which run slow*)
    c. Unknown instrumental effects (e.g., uncalibrated electronic equipment properly)
    d. Instrument precision limitations.
    e. Other effects due to the environment, error, or poor experimental design.

Note*: do not confuse "Type A or B Uncertainties" with "Type I or Type II Errors." The latter are the false-positive and false-negative rates that occur in hypothesis testing.*

The above descriptions were first detailed in 1993, when 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. The key point from this 100-page guide, which can be found in modified form on the NIST website (see References), is the following:

The uncertainty of a measurement should be reported as the total **combined standard uncertainty** U_c, which is found by adding all Type A and Type B uncertainty components *in 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.

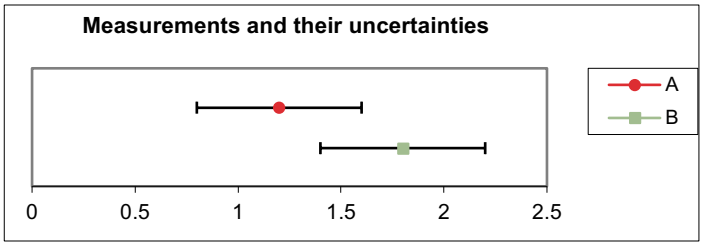
Be aware that the ± uncertainty notation might be used to indicate different confidence intervals, depending on the scientific discipline or context. For example, a public opinion poll may report that the results have a **margin of error** of ±3%, which means that readers can be 95% confident (not 68% confident) that the reported results are accurate within 3 percentage points. Similarly, an equipment manufacturer's **tolerance** rating generally assumes a 95% or 99% level of confidence.

Measurements and Their Agreement

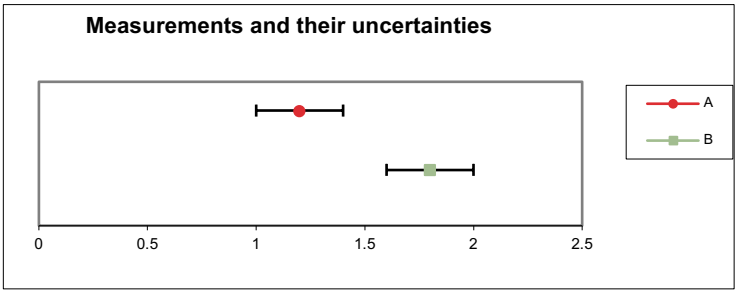
We now have the resources to answer a fundamental scientific question at the heart of scientific experimentation: "Does my result agree with a theoretical prediction or results from other experiments?"

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 ± 0.4
B = 1.8 ± 0.4



These measurements **agree** within their uncertainties, even though 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 (figure on next page):



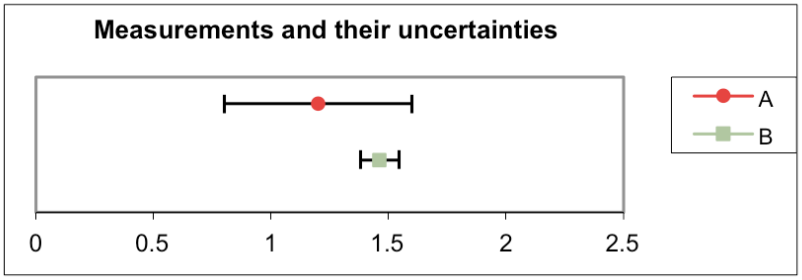
If two measurements (with similar uncertainties, each represented by ±1 sigma) barely overlap, then they differ by approximately 2-sigma, which means there is about a 5% chance that the measurements agree. When two values differ by 3-sigma or more, it is highly unlikely (less than 1% chance) that they 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: $\dfrac{|1.2-1.8|}{0.57} = 1.1$ A and B likely ***agree***

Example from above with u = 0.2: $\dfrac{|1.2-1.8|}{0.28} = 2.1$ 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 or Plots of Data

The construction of graphs (plots) is an important technique in experimental physics. They provide a compact and efficient way of displaying the functional relationship between two experimental parameters and of summarizing experimental results.

When graphs or plots 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 can be drawn on ruled graph paper – use a sharp pencil (not a pen) to draw graphs, so that mistakes can be corrected easily. More commonly, it is recommended to use a computer to produce graphs (typically by using spreadsheet software such as Excel). An example is shown below:

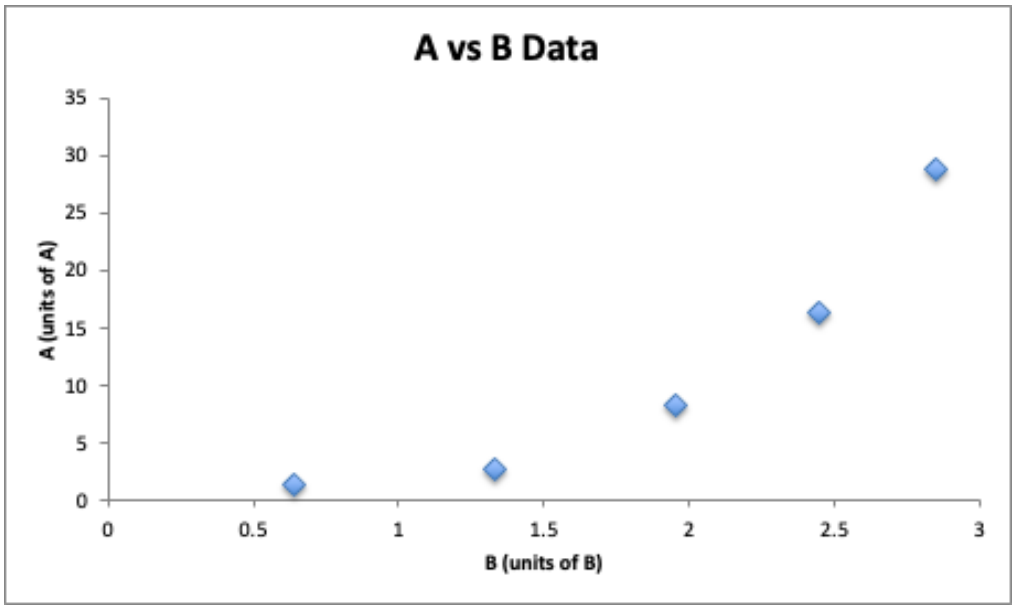


Fig 4. Plot of A versus B

Note the following important rules for graphing:

Title. Every graph should have a title that clearly states the goal or outcome of the plot; it is allowable to substitute the names of the variables that 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, l, 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 l mm graduations) should be plotted on a scale no finer than l division = l 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. Commonly, scales are 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 solid 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., ■ or ◆ or ▲) to distinguish the data sets. If drawing by hand, a drafting template is useful for this purpose. If using a spreadsheet, these symbols are commonly chosen automatically but can be changed.

Curves. As a rule, you should ***not*** draw curves through the data point – because of the random aspect of data acquisition, any curve will typically not pass through all the points, although the agreement may be close. Instead, the plot should be a *scatter plot* with unconnected dots, as shown in Fig. 4. A *best fit* curve to the data can then be automatically generated by a *regression* process (not shown in this example). The regression curve indicates the average trend of the data, and any *predicted* (interpolated or extrapolated) values can and should be read from the regression curve rather than reverting to the original data points.

Linear (straight-line) Graphs. In virtually every exercise 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 best slope and best y-intercept from the graph. This is termed a *linear* regression. In the example shown in Fig. 4, if the A data were distance an object falls in the air from rest and the B data were time, we might expect that the falling object's distance varies with time according to $d = \frac{1}{2}gt^2$. It is difficult to tell whether the plot in Fig. 4 agrees with this prediction; however, if d vs. t^2 is plotted instead, a straight line should be obtained if the data actually matches the model of a falling object. If so, the linearized slope would be $\frac{1}{2}g$ and the y-intercept would be close to zero.

Using Excel for Data Analysis in Physics Labs

Students have several 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 document. 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 generally installed on all CCI laptops and in university computer labs, students are encouraged to use this spreadsheet program or an equivalent (e.g., Apple Numbers or Google Sheets). 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. Several instructional videos have been created for the 118/119 courses and can be found on Sakai. Check these videos for details on using the Excel software so that you can duplicate this example.

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.

Time (sec)	Position (m)	Time ±	Position ±
0.64	1.15	0.05	0.20
1.10	2.35	0.07	0.30
1.95	3.35	0.05	0.20
2.45	4.46	0.06	0.40
2.85	5.65	0.10	0.40

Table 3. Position versus Time Data

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 Plot

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

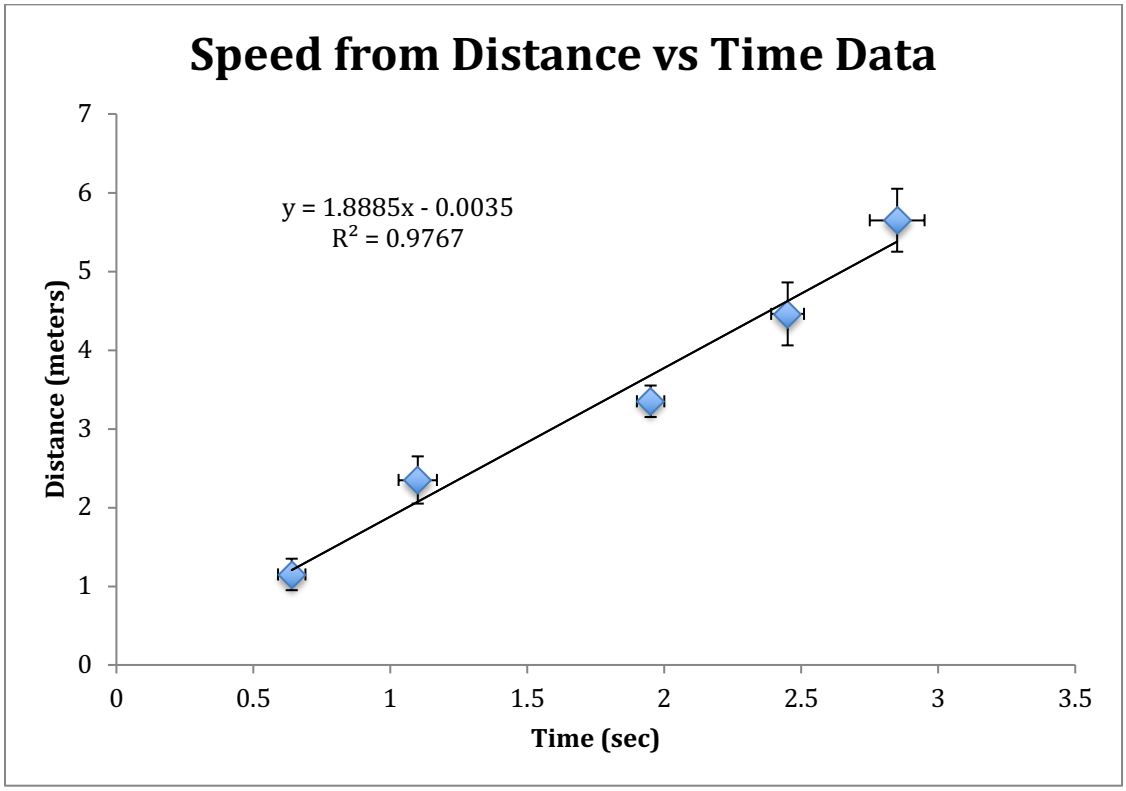


Figure 5. Trendline Analysis of Data from Table 3.

Follow these steps to accomplish this (refer to the videos for details):

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. From the plotting options, choose any scatter plot (e.g., **XY (Scatter)** or **Marked Scatter**) with no lines. A simple plot of the data should appear in the spreadsheet, and the plot should be both moveable and resizable.
3. Using the **Chart Layout/Format/Design** tool (the name varies with version), 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

The methods for adding error bars vary significantly with the different versions of Excel through the years and can also differ slightly for different platforms. In general, you will format your ***data series*** and access the ***error bar options***. 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*** options to display the values on the plot. Select the positive error value field and then click and drag in the corresponding Excel column of uncertainties. Repeat for the negative values. 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.

Consult the supplied Sakai videos for details on how to add error bars to your 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. In statistical circles this is referred to as "regression." The goal is to find a graph shape that best fits the data and is consistent with your theoretical prediction. Since we are trying to linearize the data to get a straight line characterized by a slope and an intercept, we are looking to perform a "linear regression." For our two variables x and y, we expect to be connected by a linear relationship: $y = mx + b$. A graph of y vs x should be a straight line which has a slope of m and intersects the y-axis at the value $y = b$.

Suppose we make N measurements of x and y with values (x_1, y_1), (x_2, y_2), ..., (x_N, y_N). We assume that the measurements of x have negligible error and the measurements of y have standard errors σ_1, σ_2, ..., σ_N. Finding the best straight-line amounts to finding the best estimates for m and b in a least-squares sense. The best such estimates are those that minimize the weighted sum of squares (chi-squared):

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - (b + mx_i)}{\sigma_i} \right)^2$$

Note that the quantity $b + mx_i$ is the expected value of y when $x = x_i$, thus $y_i - (b + mx_i)$ is just the deviation of the measured value of y from the expected value. The least squares method finds values of m and b that minimize the sum of the squares of these deviations weighted by their respective uncertainty.

In this course, you will use data analysis software like Excel to automatically calculate the best values of *m* and *b* and their respective errors σ_a and σ_b. If your data points have widely varying uncertainties, those with large uncertainties should be given less weight and a weighted fit should be used. In this course, we will assume (possibly incorrectly) that all data points are to be given *equal* weight.

Always select the Options "display equation on chart" and "display R-squared value on chart." A very good linear fit is indicated by an R^2 value close to 1, but values as low as 0.7 are not unexpected.

Consult the supplied Sakai videos for details on how to add a trendline to your plot.

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 the result 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.

For linear plotting, the result of the regression analysis will be a value of the slope and *y*-intercept of the trendline, but you have not completed the analysis until you have found the uncertainties in both these quantities.

Determining the Uncertainty in Slope and Y-intercept

Given the best fit line $y = mx + b$, there are two methods for finding the uncertainties in the slope *m* and *y*-intercept *b*. The methods are mathematically equivalent; therefore, either method of computation is allowed.

Method 1: First principles calculation of least squares uncertainty.

These uncertainties are calculated in a two-step process. For *n* data points, the standard error (uncertainty) in the slope *m* can be determined from the R^2 value by using the following formula:

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

The uncertainty in the y-intercept *b* is then given by:

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

These values can be computed directly in Excel or by using a calculator. For our example sample set of data, σ_m =0.1684 m/s, and σ_b = 0.3330 m. Note that a value of R^2 of *exactly* 1 results in slope and intercept uncertainties of exactly *zero*, which is highly unlikely to be correct. 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 result for our Table 3 data are shown in Fig. 5. After a regression line has been found, the trendline equation must be interpreted in terms of the context of the situation being analyzed. This data set came from a cart moving along a track. We can see that the cart was moving at essentially a constant speed since the data points in Fig. 5 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 = 1.8885 ± 0.1684 m/s. We apply our sig fig rules for a ***final*** answer of v = 1.9 ± 0.2 m/s. (Note: If we had plotted a graph of time versus distance, 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 ± 0.3330 m, which is essentially zero.

Method 2: LINEST calculation of least squares uncertainty.

The uncertainty in the slope and y-intercept can also be found by using the LINEST 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 position (Table 3, shown right).
- Follow our video instructions for creating a LINEST table.
- In general, the 4 arguments that go into the LINEST routine are [x-values, y-values, TRUE, TRUE]
- The result is shown to the right. *Note: the LINEST routine only shows the numbers highlighted in blue.*

	A	B	C	D
1		Time (sec)	Position (m)	
2		0.64	1.15	
3		1.1	2.35	
4		1.95	3.35	
5		2.45	4.46	
6		2.85	5.65	
7				
8	LINEST Results:			
9	slope →	1.88849214	-0.0035089	← intercept
10	u-slope →	0.16841081	0.33298611	← u-intercept
11	R^2 →	0.97669811	0.30976054	
12		125.74492	3	
13		12.0654252	0.28785478	

References:

Baird, D.C. Experimentation: An Introduction to Measurement Theory and Experiment Design, 3rd. ed. Prentice Hall: Englewood Cliffs, 1995.

Bevington, Phillip and Robinson, D. Data Reduction and Error Analysis for the Physical Sciences, 2nd. ed. McGraw-Hill: New York, 1991.

ISO. *Guide to the Expression of Uncertainty in Measurement.* International Organization for Standardization (ISO) and the International Committee on Weights and Measures (CIPM): Switzerland, 1993.

Lichten, William. *Data and Error Analysis., 2nd. ed.* Prentice Hall: Upper Saddle River, NJ, 1999.

NIST. *Essentials of Expressing Measurement Uncertainty*. http://physics.nist.gov/cuu/Uncertainty/

Taylor, John. *An Introduction to Error Analysis, 2nd. ed.* University Science Books: Sausalito, 1997.

Appendix I. Extension of Quadrature to Arbitrary Power Laws

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

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

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

Proof:

Consider a quantity f to be calculated by multiplying two measured quantities x and y whose uncertainties are σ_x and σ_y, respectively. 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:

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

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

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

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

$$\frac{\sigma_f}{f} = \sqrt{\left(nx^{n-1}y^m\right)^2 \frac{\sigma_x^2}{\left(x^n y^m\right)^2} + \left(mx^n y^{m-1}\right)^2 \frac{\sigma_y^2}{\left(x^n y^m\right)^2}}$$

This expression looks complicated, but it simplifies to the following rather simple result:

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

The result is that the relative (fractional) uncertainty in *f* is the same quadrature (RSS) sum of individual uncertainties in *x* and *y*, but with weights that are proportional to the power laws in the original expression. Examples include:

$f = xy$	$\frac{\sigma_f}{f} = \sqrt{(1)^2\left(\frac{\sigma_x}{x}\right)^2 + (1)^2\left(\frac{\sigma_y}{y}\right)^2}$
$f = x/y$	$\frac{\sigma_f}{f} = \sqrt{(1)^2\left(\frac{\sigma_x}{x}\right)^2 + (-1)^2\left(\frac{\sigma_y}{y}\right)^2}$
$f = xy^2$	$\frac{\sigma_f}{f} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + 4\left(\frac{\sigma_y}{y}\right)^2}$

Note again that the results for multiplication and division are 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 evaluate 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.

The uncertainty estimate from the ULB 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 understand 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.

Appendix II. Practical 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. The

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 seem "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 considered.*

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* despite subtracting the two values. 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

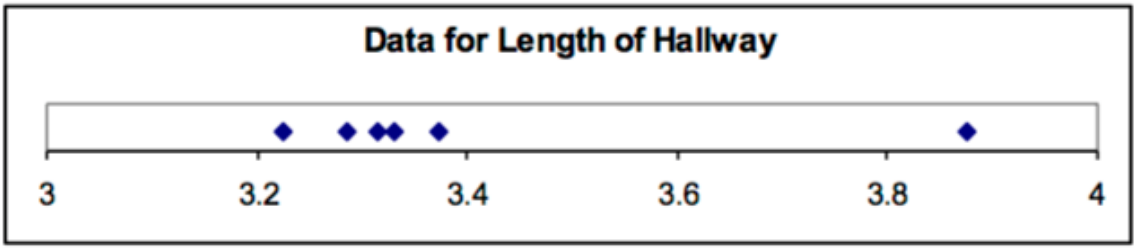
A student uses a protractor to measure an angle of $\theta = 85° \pm 1°$. What should she report for $\sin\theta$?

$\sin(\theta) = 0.996 \pm 0.002$. The uncertainty can be determined from either the upper/lower bound method or propagation of uncertainty using partial derivatives (quadrature). 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 ± 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 mismeasurement (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 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 ± 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/s² (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 result is an acceleration of $a = 2.39 \pm 0.09$ m/s², or 2.4 ± 0.1 m/s².

From the standpoint of upper-lower bound (not the best estimate), we have

$$a_{upper} = 2\frac{d+\delta d}{(t-\delta t)^2} = 2\frac{1.5+0.005\ \text{m}}{(1.12-0.02\ \text{s})^2} = 2.49\ \text{m/s}^2$$

$$a_{lower} = 2\frac{d-\delta d}{(t+\delta t)^2} = 2\frac{1.5-0.005\ \text{m}}{(1.12+0.02\ \text{s})^2} = 2.30\ \text{m/s}^2$$

With the result of $a = 2.39\ {}^{+0.10}_{-0.09}$ m/s². In this case, we see that the ULB and quadrature calculations are virtually identical, with the ULB result being slightly larger, as expected.

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 ± 0.33 m, t = 1.14 ± 0.04 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 $a = 2h/t^2 = 10.387$ m/s². In this case, neither relative uncertainty in time nor distance dominates, so both should be included in a quadrature calculation:

$$\sigma_a = a\sqrt{\left(\frac{\sigma_h}{h}\right)^2 + 4\left(\frac{\sigma_t}{t}\right)^2} = (10.387\ \text{m/s}^2)\sqrt{\left(\frac{0.33}{6.75}\right)^2 + 4\left(\frac{0.04}{1.14}\right)^2} = 0.888\ \text{m/s}^2$$

The final answer would be written as 10.4 ± 0.9 m/s². Note that this answer is higher than the value for g in a vacuum. Although the uncertainty range includes the value 9.8 m/s², 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.

From the standpoint of upper-lower bound (not the best estimate), we have

$$a_{upper} = 2\frac{h+\delta h}{(t-\delta t)^2} = 2\frac{6.75+0.33\ \text{m}}{(1.14-0.04\ \text{s})^2} = 11.70\ \text{m/s}^2$$

$$a_{lower} = 2\frac{h-\delta h}{(t+\delta t)^2} = 2\frac{6.75-0.33\ \text{m}}{(1.14+0.04\ \text{s})^2} = 9.22\ \text{m/s}^2$$

With the result of $a = 10.39\ {}^{+1.31}_{-1.17}$ m/s², or 10.4 ± 1.3 m/s². In this case, we see that the ULB result displays a noticeably larger uncertainty than the quadrature calculation.

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 (κ = 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:

$$C = \frac{Q}{V} = \kappa\varepsilon_0 \frac{A}{d} = 3(8.85\times10^{-12}\frac{\mathrm{C}^2}{\mathrm{Nm}^2})\frac{\pi(0.11\ \mathrm{m})^2}{0.005\ \mathrm{m}} = 202\ \mathrm{pF}$$

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

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 ± 24 pF.

From the standpoint of upper-lower bound (not the best estimate), we have

$$C_{\ e} = \varepsilon_0 \frac{\pi}{4}\frac{(\kappa+\delta\kappa)(D+\delta D)^2}{(d-\delta d)} = (8.85\times10^{-12})\frac{\pi}{4}\frac{(3+0.3)(0.22+0.004)^2}{(0.005-0.0003\ \mathrm{s})} = 224.7\ \mathrm{pF}$$

$$C_{l\ e} = \varepsilon_0 \frac{\pi}{4}\frac{(\kappa-\delta\kappa)(D-\delta D)^2}{(d+\delta d)} = (8.85\times10^{-12})\frac{\pi}{4}\frac{(3-0.3)(0.22-0.004)^2}{(0.005+0.0003\ \mathrm{s})} = 181.6\ \mathrm{pF}$$

With the result of C = 202 +23 –20 pF, or 202 ± 23 pF. In this case, we see that the ULB result displays a somewhat smaller uncertainty than the quadrature calculation.

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 π, 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 π?

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

$$\sigma_m = m\sqrt{\frac{(1/R^2)-1}{n-2}} = (3.1527)\sqrt{\frac{(1/0.9967)-1}{5-2}} = 0.1047$$

The final experimental value for π is then 3.15 ± 0.10, or 3.2 ± 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 π can therefore only be trusted to the extent that –0.0502 (the intercept) is experimentally equivalent to zero. Therefore, the uncertainty in the intercept should 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 Δt that has elapsed is recorded. Since *d* and Δt 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 $v = d/\Delta t$, where d is the flag length and Δt 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 cos30° and then is shortened again by a separate factor (f) by which the timer is defective:

$$incorrect\ speed = \frac{d}{f * cos30° \Delta t}$$

The relative error is given by equation 2: $$relative\ error = \frac{measured - expected}{expected}$$

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

$$Relative\ speed\ error = \frac{\frac{d}{f * cos30° \Delta t} - \frac{d}{\Delta t}}{\frac{d}{\Delta t}} = \frac{1}{f * cos30°} - 1 = +28.3\%\ (too\ fast)$$

Discussion

The time measured by a photogate for a flag of length d moving at speed v is t = (dcosθ)/v, where θ is the flag angle measured from the line of travel: for any flag length d, the time measured by the photogate approaches zero as θ approaches 90°. If the photogate itself measures incorrectly by a factor of f, then the time measured is t = f(dcosθ)/v. If the photogate is measuring 10% less than it should, then f = 0.9. We now calculate the relative time error:

$$\textit{Relative time error} = \frac{t_{meas} - t_{expected}}{t_{expected}} = \frac{\frac{fd\cos 30^\circ}{v} - \frac{d}{v}}{\frac{d}{v}} = 0.9\cos 30^\circ - 1 = -.22058$$

*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:*

$$\frac{v_{meas} - v_{expected}}{v_{expected}} = \frac{\frac{d_{meas}}{t_{meas}} - \frac{d_{expected}}{t_{expected}}}{\frac{d_{expected}}{t_{expected}}} = \frac{\frac{d}{0.77942 t_{expected}} - \frac{d}{t_{expected}}}{\frac{d}{t_{expected}}} = \frac{1}{.77942} - 1 = +.283$$

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: $\frac{d\cos 30^\circ - d}{d} = \cos 30^\circ - 1 = -13.4\%$ (–13% to 2 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: $\frac{\frac{d}{\cos 30^\circ \Delta t} - \frac{d}{\Delta t}}{\frac{d}{\Delta t}} = \frac{1}{\cos 30^\circ} - 1 = +15.5\%$

c. Time measurement error: $\frac{\frac{d}{1.1\cos 30^\circ \Delta t} - \frac{d}{\Delta t}}{\frac{d}{\Delta t}} = \frac{1}{1.1\cos 30^\circ} - 1 = +4.97\%$

d. time measurement error: $\frac{\frac{\cos 30^\circ d}{0.9\Delta t} - \frac{d}{\Delta t}}{\frac{d}{\Delta t}} = 1.1\cos 30^\circ - 1 = -4.74\ \%$
Note the wrong sign – answer should result in higher speed if the timer reads low.

e. incorrect: ±13% error on flag length ± 10% error on timer = ±23% or ±3%

f. incorrect: quadrature calculation $\sqrt{(13\%)^2 + (10\%)^2} = \pm 16.4\%$

(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 measures (incorrectly) 10% **more** time than it should – what relative error in speed results?*

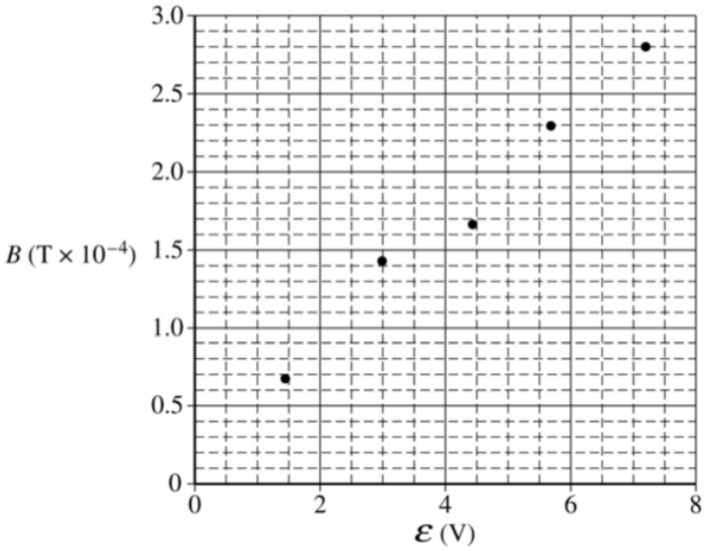
Answer: for a value of f = 1.1, the result is identical to example (c) in the common errors list above:

$$\text{Relative speed error} = \frac{\frac{d}{1.1*cos30°\Delta t} - \frac{d}{\Delta t}}{\frac{d}{\Delta t}} = \frac{1}{1.1*cos30°} - 1 = +4.97\% \text{ (too fast)}$$

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 ε of the power supply is shown to the right (error bars not shown). Theoretically, the value of B should be $\mu_0 nI$, where n is the number of turns per unit length, and I is the current.

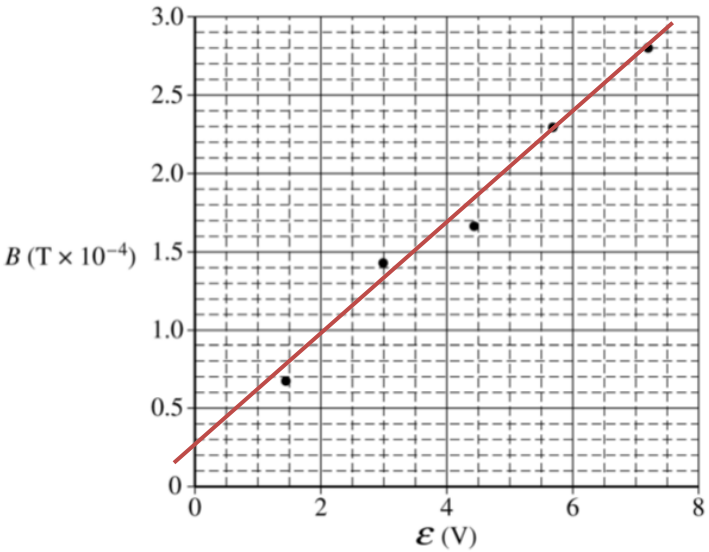


One of the students notes that the horizontal component of the magnetic field of the Earth is 2.5×10^{-5} T. Is there evidence from the graph that the Earth's magnetic field systematically affected their measurement? What can you say about the orientation of the solenoid?

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

where ε is the voltage (emf) applied to the solenoid and R is the resistance of the coil. Therefore, the plot of B vs ε 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.



XVI. Combining Type A and Type B Uncertainties I

Refer to the Combining and Reporting Uncertainties section of this document, in particular the discussion of Type A and Type B uncertainties. Let us reconsider Example I of the appendix. Suppose we have found a trustworthy digital balance with a precision of 0.1 g, and its measurement of a mass is 495.3 g with an accuracy also of 0.1 g – how should we report the result?

We don't know the details of how the balance comes to the result it displays. Imagine that the balance has an extra digit. The value of 495.3 with the extra digit could be anywhere between 495.2-something and 495.3-something – this is the effect of the limited precision of the device, and we don't know how the balance's microcontroller programming applies rounding rules. Adding the accuracy into the mix in an upper-lower bound sense means the "true" value could be anywhere between 495.1-something and 495.4-something. In an upper-lower bound sense then, we would most safely cite this as 493.3 ± 0.2 g.

In contrast, we know that the upper-lower bound is an overestimate, because it assumes that every measurement could be at the outer limits of the possible values. This assumption is not typically correct (e.g., due to a random noise component). A better approach is to combine the precision and uncertainty in a sum-of-squares (quadrature) sense:

$$\sqrt{(0.1\ \text{g})^2 + (0.1\ \text{g})^2} = 0.1414\ \text{g}$$

Unfortunately, we are faced here with a slight conundrum. Strict rounding rules would allow us to cite this as 0.1 g, but we sometimes make exceptions when the leading digit is 1, so to be safe we might cite the uncertainty as either 0.14 g or 0.2 g. When faced with this issue, you should explain your choice (which is why we always ask you to explain your answers!).

Note that this process highlights a method by which we can decide if and when to discard either the precision or the accuracy as a contributor to the answer. The relative contributions of precision and accuracy are quantitative under the radical. If the ***square*** *of either is significantly larger or smaller than the other, it can be ignored; in this example, neither can be. See the examples in the Quadrature section of this document for more discussion.*

XVII. Combining Type A and Type B Uncertainties II

A student uses a stopwatch to measures three instances of the time it takes a ball to fall to the ground. The student is faced with the following uncertainties:

1. The stopwatch itself has a precision of 0.01 s (the stopwatch cannot do better than 0.01 s, but it is unknown if that is a representation of how accurate the stopwatch is, or merely the limit of the display).
2. The standard deviation of the three measurements is 0.03 s. This uncertainty is due to random variations in the execution of the experiment (Type A).
3. After the experiment is over, the student realizes that they did not carefully consider or try to correct for their reaction time in pressing the stopwatch start and stop button. After some research, it is discovered that human reaction times can be as large as 0.2 sec and could vary by 10% from person to person or from event to event. This uncertainty is classified as systematic (Type B).

How should the uncertainty in this measurement be expressed?

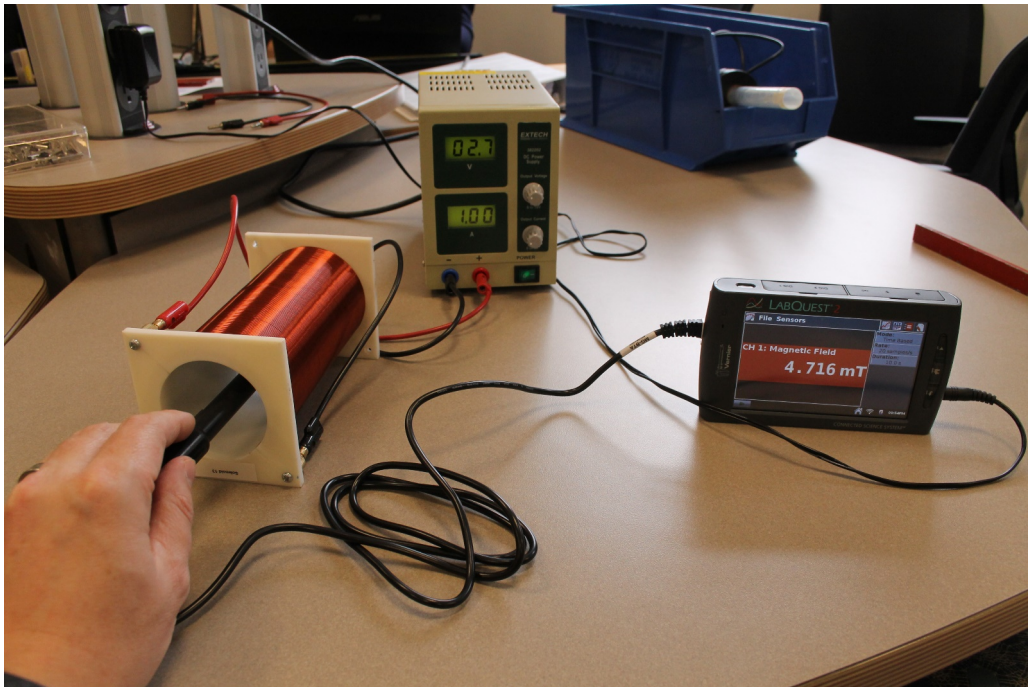
Note first that the reaction time likely affects both the start and the stop of the stopwatch; therefore, the 0.2 sec "error" due to reaction time is not likely to result in a significant uncertainty in the elapsed time between start and stop. In contrast, the reaction time itself can vary by 10% of this value, or 0.02 s.
The correct approach to combining these separate uncertainties is a root sum of squares:

$$\textit{total uncertainty} = \sqrt{(0.01\ s)^2 + (0.03\ s)^2 + (0.02\ s)^2}\ = 0.037\ s\ =\ 0.04\ s\ \textit{(1 sig fig)}$$

Note that if the experiment is repeated with better instruments, more trials, or a more careful control of the systematic effects, any of the uncertainty contributions could be reduced so as to be negligible in a root-mean-square sense.

XVIII. Revisiting the Solenoid

The details of a solenoid experiment are shown below:



A voltage of 2.7 VDC applied to the solenoid produces a current of 1.00 A. A probe measures the magnetic field in the center of the solenoid as 4.716 ± 0.004 mT . The voltage and current measurements have an accuracy of ± 2.5%.

Using only the measured current and magnetic field and the theoretical equation for the magnetic field in the solenoid, calculate the number of windings per unit length and its uncertainty. (Result 1).

$$n = \frac{B}{\mu_0 I} = \frac{4.716 \pm 0.004 \times 10^{-3}\ T}{(1.257 \times 10^{-6}\ H/m)(1.00 \pm 0.03\ A)} = 3752 \pm 119 \text{ per meter}$$

This result is an upper-lower bound; from a quadrature point of view, the relative uncertainty in n is:

$$\frac{\sigma_n}{n} = \sqrt{\left(\frac{\sigma_B}{B}\right)^2 + \left(\frac{\sigma_I}{I}\right)^2} = \sqrt{(0.0008)^2 + (0.03)^2} = 0.03$$

In contrast to the previous example, we can ignore the contribution to uncertainty from the field measurement because it is much smaller than the current measurement uncertainty. The result is σ_n = 113 per meter, which is smaller (marginally) than the upper-lower bound result, as expected.

Appendix III. Summary of Important Tips for Uncertainty in PHYS 118/119

1. Precision and accuracy are *unrelated*. Precision is how many significant figures are retained in any value. Accuracy is how close a value is to the *true* value. An answer can be very precise and very inaccurate. Likewise, an accurate answer may or may not be precise. *Refer to the discussion surrounding Figure 1*.
2. The uncertainty in a measurement has the same units as the measurement. In contrast, the *relative* uncertainty is a ratio of the uncertainty to the measured value and is expressed as a percentage. *Refer to the discussion surrounding Equation 1*.
3. The term *human error* is vague and should be avoided in discussions or presentations. Although it's possible that a measurement was incorrectly made, such an error is a form of *systematic uncertainty* and should be referred to as such.
4. An *average* of several values (at least 3 for PHYS 118/119) can be taken as the *expected* or *measured* value of a quantity. *Refer to the discussion surrounding Table 1*.
5. The *standard deviation* of several values (at least 3 for PHYS 118/119) can be taken as the *uncertainty* of the *expected* or *measured* value of a quantity if there is no knowledge of possible systematic effects. If systematic effects are known to have been *eliminated*, the *standard error* is likely to be a better calculation of the uncertainty. *Refer to the discussion surrounding Table 2, Figure 2, and Figure 3*.
6. Significant figures can be used to estimate uncertainty, but PHYS 118/119 *requires* uncertainty to be computed as a separate value. If you are performing your own experiments with limited trials, round your uncertainty to 1 sig fig and match the precision of the measured value to the precision of the uncertainty. *Refer to the discussion surrounding Table 2*.
7. If you are working on an exam or homework problem that cites uncertainty to more than 1 sig fig, take it at face value – you did not perform that experiment!
8. Upper-lower bound calculations are acceptable for uncertainty propagation; however, quadrature calculations are more elegant and more statistically sound. Quadrature-like calculations (root sum of squares, or RSS) are required to combine uncertainties from different sources, especially in the following two cases:
    - combining random and systematic uncertainties.
    - combining any uncertainties and instrument precisions.

    Refer to the example in the section on Combining and Reporting Uncertainties.

9. Unless otherwise directed, ***always*** attempt to linearize data so that a linear regression can be applied to find a slope, intercept, and their uncertainties. *Any other approach is not considered best practice and will be judged as inferior.*