Uncertainties in Physics: A Brief Reference

This is a brief and incomplete discussion of **error analysis**: the identification and quantitative evaluation of uncertainties in a measurement or experiment. The ultimate goal of error analysis is to determine the range around the measured result where the **true value** is most likely to lie. A good measurement will have uncertainties as small as feasible to constrain the possible values the true value can have, but the uncertainties must also be large enough to account for the imperfections in the experiment.

This discussion of error analysis is incomplete out of necessity; there are many books devoted entirely to the subject, and we cannot hope to learn it all here! So, I will introduce a few methods of error analysis commonly used by scientists that will be helpful in our labs this year.

1 Identifying and Reporting Uncertainties

All experiments are associated with some uncertainty due to the limitations of the procedure and measuring tools. An accurate assessment of the uncertainty is always required when reporting the results of an experiment. Otherwise, there is no way to tell if your result confirms known physics, predicts new physics (and gets you a Nobel Prize!), or is limited by large uncertainties. Uncertainties fall into two categories: random and systematic.

1.1 Identifying Uncertainties

Random uncertainties include the imprecision in reading a meter or scale, uncertainties due to uncontrolled fluctuations in experimental conditions, and uncertainties due to the use of a limited sample size. Generally, for many trials of a measurement, random uncertainties will manifest as a different result in each trial, with the results fluctuating randomly about some mean value. If these fluctuations are small, the experiment is said to be **precise**. Random uncertainties are sometimes referred to as **statistical uncertainties**.

Systematic uncertainties are the result of imperfect calibration of the apparatus or assumptions and approximations made in the experimental design. These effects will shift the result away from the true value in a way that is consistent over repeated trials. In other words, systematic uncertainties change the mean value of the measurement by a fixed number but otherwise give a consistent result from one trial to the next, which makes these effects particularly tricky to detect. If systematic uncertainties are small, the experiment is said to be **accurate**.

Physicists will use the words "error" and "uncertainty" interchangeably. You saw this in the very first sentence of this document in the definition of error analysis. This can be

quite confusing since these words have very different meanings in everyday language. Please be aware that these words truly are synonyms in the physics lab. In particular, a list of "experimental errors" does not mean that the experimenter is listing their mistakes. Also, "human error" or mistakes should never be mentioned in your lab report as potential source of uncertainty. It is expected that you work carefully and diligently to avoid and eliminate mistakes, so that all uncertainties can be attributed to the limitations in the experimental design and apparatus.

1.2 Reporting Uncertainties

Random uncertainties are always present. A well-designed experiment will use a measuring device that is sensitive enough to detect these random fluctuations so that duplicate measurements do not yield identical results . When using devices that require reading a scale or analog meter, the correct procedure is to report your readings by estimating to one tenth of the smallest scale division whenever possible. For example, when using a ruler with millimeter line markings to measure the length of an object, you should report your measurement to the nearest tenth of a millimeter.

Once you've evaluated an uncertainty, there are several ways to report it. The simplest is to report the measurement with the correct number of significant figures. For example, if you report the length of an object to be 7.41 cm, it is implied that the uncertainty is in the last digit and is at least 0.01 cm (or 0.1 mm) but less than 0.09 cm. The obvious downside to this method is that there is some ambiguity. Perhaps, in this example, the experimenter used a ruler with millimeter markings and interpolated to the nearest 0.1 mm so that 0.1 mm is the uncertainty, but we won't know that unless this information is included in the lab report.

It is usually better to explicitly state the uncertainty. This is often done with plus-minus ($\pm$) notation, for example: 7.41 ± 0.03 cm. This indicates that the best measurement is 7.41 cm and the uncertainty is 0.03 cm. The $\pm$ says that the true length is within 0.03 cm of 7.41 cm; in other words, between 7.38 cm and 7.44 cm.

There are two important things to note with this notation. First, the measured value 7.41 cm and the uncertainty 0.03 cm contain the same number of decimal places. This must always be done! Second, the uncertainty of 0.03 cm only contains one significant figure. This is almost always the case - since the uncertainty indicates the degree to which the result is unknown, any figures beyond the first are meaningless. The only exception is if the first significant figure of the uncertainty is 1. In this case, it is acceptable to include a second figure: for example, 7.414 ± 0.013 cm (notice that an extra digit is given with the measured value so that the number of decimal places match).

When using $\pm$ notation with scientific notation, it is usually best to use the same power of 10 for both the measurement and uncertainty. For example, if you measure the wavelength of light to be 5.31×10^{-7} m and the uncertainty to be 4×10^{-9} m, you would state the result to be $(5.31 \pm 0.04)\times10^{-7}$ m. Of course, you could avoid scientific notation entirely by using a prefix on the units: 531 ± 4 nm (nanometers) in this example.

2 Assessing Random Uncertainties

Random uncertainties are assessed by repeating a measurement as many times as is practical, without changing any relevant experimental conditions. For example, suppose you measure the time for an object to accelerate (starting from rest) through some distance d. You do this measurement eight times and measure the following times:

TRIAL	Time (s)
1	5.22
2	5.41
3	5.63
4	5.31
5	5.03
6	5.53
7	5.41
8	5.75

If asked to report the best value for the time, a sensible thing to do is to report the *average* (often called the *mean*) value of the time data,[1] where the average time $\bar{t}$ is given by

$$\bar{t} =\equiv \frac{1}{N}\sum_{j=1}^{N} t_j \tag{1}$$

where N is the number of measurements of t (in the example above, $N = 8$). Hence, for the data above,

$$\bar{t} = \frac{(5.22 + 5.41 + 5.63 + 5.31 + 5.03 + 5.53 + 5.41 + 5.75)}{8} = \frac{43.29}{8} = 5.41125 \text{ sec.}$$

Thus, we would report our best estimate for the time as roughly 5.41 seconds. Now, what is the uncertainty on this number? Again, there are several ways to go.

2.1 A conservative approach

The first method is to be very conservative (this has nothing to do with one's political leanings!) and state the maximum and minimum possible values relative to the mean. For example, in the data above, the maximum time is 5.75 seconds, and the minimum time is 5.03 seconds. These values differ from the mean by +0.34 and -0.40 seconds, respectively. So, one could use the higher of these two values (to be the most cautious) and state the time and uncertainty as

$$\bar{t} = (5.4 \pm 0.4) \quad \text{seconds}$$

(be sure you understand why 5.41125 is rounded to one decimal place!). This is perhaps the most straightforward way to get the uncertainty, and it says that we are confident that the true time lies pretty close to 5.4 seconds, but may vary by as much as 0.4 seconds from this time.

[1] There are other ways to calculate the "best" value; the mean is only one of several possibilities.

2.2 Standard deviation

Notice however, that this uncertainty is very cautious; indeed, most measurements lie significantly closer to the mean value. Hence, it is reasonable to quote a smaller uncertainty, so long as you interpret the error bar to give a range where the true value is *likely* to be rather than an absolute statement of possibilities. So here is another way to estimate the uncertainty. This second method is called the standard deviation and is computed by first calculating the deviation of each trial j from the mean; i.e.

$$\Delta t_j = t_j - \bar{t} \tag{2}$$

where $\bar{t}$ is calculated using equation (1). Note that values higher than the mean are positive deviations, and values smaller than than the mean are negative deviations. The table below shows the deviations for the original data:

TRIAL	**Time (s)**	$\mathbf{\Delta t_j}$
1	5.22	-0.19125
2	5.41	-0.00125
3	5.63	+0.21875
4	5.31	-0.10125
5	5.03	-0.38125
6	5.53	+0.11875
7	5.41	-0.00125
8	5.75	+0.33875

Notice that the uncertainty for trials 2 and 7 are not zero; this is because I used the full value for $\bar{t}$ (5.41125 seconds) in calculating Δt_j. It is best to not round error analysis calculations until the very end when reporting the result; see Section 1.2.

At this point, you might be tempted to say "Oh great! So now I just take the deviations, add them together and compute the average value and this will be the uncertainty!" Bad luck. You see, the problem is that if you do this, you will get zero: try it and see! And the problem is even worse than this—the average deviation computed in this way will always be zero because of the definition of the mean and the deviation we have used.

So, we need a way around this trouble. The reason the average deviation is zero is because some of the deviations are positive and some are negative (in just the right amounts so that they sum to zero). So, we can make the deviations positive by first squaring them. Then if we add them together they will not sum to zero, and if we divide by the number of trials minus one[2], we have something called the *variance* of t, or σ_t^2:

$$\sigma_t^2 = \frac{1}{N-1}\sum_{j=1}^{N}(\Delta t_j)^2 \tag{3}$$

Of course, if t has units of seconds, then σ_t^2 has units of seconds-squared, and the obvious way to remedy this situation is to take the square root, thus ending up with the *standard deviation* of t, or σ_t:

[2]We use $N-1$ instead of N for technical reasons that are not important for our purposes.

$$\sigma_t = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}(\Delta t_j)^2} \tag{4}$$

Adding onto the previous data table, we have

TRIAL #	**Time (s)**	$\mathbf{\Delta t_j}$	$\mathbf{(\Delta t_j)^2}$
1	5.22	-0.19125	0.03658
2	5.41	-0.00125	0.000002
3	5.63	+0.21875	0.04785
4	5.31	-0.10125	0.01025
5	5.03	-0.38125	0.14535
6	5.53	+0.11875	0.01410
7	5.41	-0.00125	0.000002
8	5.75	+0.33875	0.11475
		σ_t^2:	**0.05270 sec^2**

The variance is is 0.05270 s^2 and, taking the square root, the standard deviation is 0.22956 seconds. Hence, by the standard deviation method of calculating the uncertainty, we have, as our estimate of t:

$$\bar{t} = (5.4 \pm 0.2) \quad \text{seconds}$$

Notice that this uncertainty is considerably smaller than that derived from the conservative approach. The reasoning is that while occasionally you may measure a time that falls outside of the standard deviation bounds, on average, the majority will fall within this range[3].

2.3 Standard error

The standard deviation quantifies the spread of the data points about the mean. In other words, if you were to measure one additional data point, the standard deviation roughly tells you how far to expect that measurement to be from the mean. This is useful, as it predicts the range of values of future measurements, but it is still a conservative approach. What we usually want is the uncertainty on the mean value itself, rather than the uncertainty on one data point.

The solution is to calculate the *standard error* $\sigma_{\bar{t}}$ of $\bar{t}$, as follows:[4]

$$\sigma_{\bar{t}} = \frac{\sigma_t}{\sqrt{N}} \tag{5}$$

Notice that as the number of measurements N increases, the standard error gets smaller. This is just as you would expect: the uncertainty on the mean gets smaller when you collect more data.

[3]For measured values that follow a normal, or Gaussian, distribution, about 68% will fall within this range.

[4]This formula assumes that your measurements follow a normal distribution. This is usually true, though there are some notable exceptions, such as when counting the number of radioactive decays of an unstable nucleus.

Continuing the example from the previous sections, we divide the standard deviation of t (0.22956 seconds) by $\sqrt{N} = \sqrt{8}$ and calculate the standard error to be 0.081162 seconds. Our new best estimate of t is:

$$\bar{t} = (5.41 \pm 0.08) \text{ seconds}$$

As expected, this uncertainty is smaller than the standard deviation method.

2.4 Which method should I use?

If you do not have many measurements of a quantity, it is probably best to be cautious and use the conservative approach. If you need to predict the range of results of a single future measurement, then the standard deviation is best. But for most cases, you will be interested in the uncertainty on the mean of several trials, which is best determined by the standard error.

QUESTIONS

1. Calculate the standard deviation and standard error for the following data points: x = 0.10 m, 0.08 m, 0.12 m, 0.09 m, 0.10 m, and 0.11 m. **Answers:** $\bar{x} = 0.100$**,** $\sigma_x = 0.014$**,** $\sigma_{\bar{x}} = 0.0058$

2. Suppose that you measure two times with your phone's stopwatch, t=1.51 sec and t=1.51 sec. What—if anything—is wrong with saying that the uncertainty in t is zero? What should you report for the uncertainty in this case?

3 Graphical Analysis and Random Uncertainties

In the physics lab, we often want to test how some quantity of interest (the **dependent variable**) depends on an experimental parameter that is controlled (the **independent variable**). This is usually done by conducting multiple trials of the experiment, with each trial done for a different value of the independent variable. For example, an experimenter may record the electrical current (dependent variable) in a circuit for different values of the input voltage (independent variable) to the circuit.

In this sort of experiment, it is usually best to analyze the data by plotting the dependent variable (on the y axis) vs. the independent variable (on the x axis). By making clever choices for the plotted variables, we can usually make this plot linear so that the measurement is to simply measure the slope and y-intercept of a line of best fit. It usually is not practical to use the random uncertainty analysis methods of section 2 in this sort of measurement,[5] but we can still estimate the random uncertainty of the slope and y-intercept of the linear fit from the way the data points scatter around the fit line, as you will see.

[5] this would require that you take so many measurements that you can produce many plots and measure the slope of each one.

3.1 Data Linearization

For a concrete example, suppose that you want to compute the acceleration of a dense object released from rest as it falls in the gravitational field of the earth. You know from your introductory physics course that for an object released from rest close to the earth's surface, that the distance fallen d as a function of time t is approximately

$$d = \frac{1}{2}gt^2 \tag{6}$$

where g is the acceleration due to gravity.

You set out to determine the experimental value of g by timing how long it takes for a ball bearing to fall through some measured distance. So, being a good experimenter, you measure the time it takes to fall through at least 10 different distances. For each distance, you record five trials, and compute the average time of fall and the uncertainty on the time (via one of the methods outlined in Section 2). You therefore end up with ten trials with uncertainties on the distance of fall and uncertainties on the the time of fall.

Trial	distance ± 0.01m	t ± 0.01 s
1	0.00	0.00
2	0.29	0.18
3	0.23	0.22
4	0.73	0.34
5	1.13	0.41
6	1.82	0.57
7	2.11	0.65
8	3.21	0.77
9	3.95	0.86
10	4.61	0.94

So now, the question is what is the best way to graphically analyze this data, and obtain an experimental value for g? What we want to do is linearize the data so that we get a linear trend when plotted. Then, we will fit the data to a line of best fit and extract a measurement of g from the slope.

To figure out how to plot our data to make it linear, we need to determine a function of distance $F(d)$ and another function of time $G(t)$ so that when we plot $F(d)$ on the y-axis and $G(t)$ on the x-axis, the trend is linear. In other words, we want these functions to satisfy:

$$F(d) = mG(t) + b \tag{7}$$

where m is the slope of the line and b is the y-intercept. Now, look at Eq. (6). If we choose $F(d) = d$, and $G(t) = t^2$, we see that

$$F(d) = \frac{1}{2}gG(t) \tag{8}$$

Therefore, if we were to plot d on the y-axis vs. t^2 on the x-axis, we expect to get a linear trend with a slope equal to $\frac{1}{2}g$ and a y-intercept of zero.

3.2 Uncertainty of a Linear Fit

There are several ways to evaluate the uncertainties on a linear fit. In the first approach, which I'll call the manual method, you will need to construct your plot with vertical and horizontal error bars. In the previous example of a d vs t^2 plot, the vertical error bars indicate the uncertainty on the measurement of d, and the horizontal error bars indicate the uncertainty on t^2 (see the next section for methods to find the uncertainty on t^2 starting from the uncertainty on t). Draw different lines of best fit on the plot that are consistent with most of the error bars. The goal is to find, by eye, the lines of best fit with the smallest and largest possible slopes that are consistent with the uncertainties. In this method, the best fit slope will be the average of the smallest and largest slopes, and the uncertainty will be the difference between the average slope and the extreme values, similar to the conservative method of Section 2.1.

If you use spreadsheet software such as Microsoft Excel or Google Sheets to make your plot, then there is a simpler and more accurate way to find the uncertainty. These programs contain a function called LINEST() which takes as input the data used to construct the plot, and outputs the best fit parameters (e.g., the slope and y-intercept) and their standard errors, among other information. Refer to your spreadsheet software's documentation for LINEST() for more information.

4 Propagation of Uncertainties

It is often the case that we are unable to directly measure the desired physical quantity. Instead, we directly measure quantities that have a known relationship with the desired quantity, and we need to use that information to determine the desired quantity and its uncertainty. This problem is an example of **error propagation**, because we need to propagate the error from the directly measured quantities to the desired quantity. In a general error propagation situation, several variables x_1, x_2, ..., x_n with uncertainties Δx_1, Δx_2, ..., Δx_n are measured by the experimenter, who needs to calculate a new quantity y that depends on the measured variables according to a known function f; i.e., $y = f(x_1, x_2, ..., x_n)$.

The method used to compute the uncertainty Δy depends upon the size of the experimental errors on the directly measured quantities. If you were a careful experimenter and used good equipment so that the errors are small, then you can use *differential error propagation*; if your uncertainties in the time were larger, it might make sense to use what I'll call a *worst-case scenario method*.

4.1 Worst-case scenario method - one variable

Let's simplify the general error propagation situation to one that only involves one directly measured quantity x with error Δx. The calculated quantity y depends on x according to a known function $y = f(x)$. If the uncertainty in x is not small, then one could argue that

a more realistic estimate of the uncertainty in y would be to compute the extreme possible values of y that are consistent with the uncertainty on x and report the uncertainty as half their difference. The extreme values of y, which I'll call $y_{\max}$ and $y_{\min}$, are calculated as follows:

$$\begin{aligned} y_{\max} &= f(x + \Delta x) \\ y_{\min} &= f(x - \Delta x) \end{aligned} \tag{9}$$

In other words, $y_{\max}$ is the value of y when x takes on its largest possible value consistent with uncertainties, and $y_{\min}$ corresponds to the smallest possible value of x. The uncertainty Δy on y is

$$\Delta y = \left| \frac{y_{\max} - y_{\min}}{2} \right| \tag{10}$$

Let's see how this method works in an example. Let's say you have measured an angle $\theta = 25 \pm 2^\circ$ and wish to calculate a quantity s that is related to θ according to $s = 5\cos\theta$. Compare this specific problem to the general method above. The measurement θ plays the role of x, the function $f(x)$ is $f(\theta) = 5\cos\theta$, and the calculated quantity s corresponds to y. We start by finding $s_{\max}$ and $s_{\min}$ according to equation (9):

$$\begin{aligned} s_{\max} &= 5\cos(\theta + \Delta\theta) = 5\cos(25^\circ + 2^\circ) = 4.4550 \\ s_{\min} &= 5\cos(\theta - \Delta\theta) = 5\cos(25^\circ - 2^\circ) = 4.6025 \end{aligned} \tag{11}$$

Then we calculate the error Δs:

$$\Delta s = \left| \frac{s_{\max} - s_{\min}}{2} \right| = \left| \frac{4.4550 - 4.6025}{2} \right| = 0.07 \tag{12}$$

Next, calculate the best value of s using the best value of θ:

$$s = 5\cos(25^\circ) = 4.53 \tag{13}$$

We now have everything needed to report the measurement of s with appropriate uncertainty: $s = 4.53 \pm 0.07$.

4.2 Differential error propagation - one variable

This method is optional in PHY 114 and PHY 116 because it assumes familiarity with differential calculus.

As we did in the previous section, let's consider a situation that involves only one directly measured quantity $x \pm \Delta x$ and a calculated quantity y that depends on x according to $y = f(x)$. The standard way to compute the uncertainty on the calculated quantity y, assuming the uncertainty Δx is small, is to use differential error propagation. The differential change in y is simply the total derivative or $dy = \left(\frac{df}{dx}\right) dx$. By assuming that the experimental uncertainty Δx is small, we can replace dx and dy with Δx and Δy to approximate the experimental uncertainty on y as:

$$\Delta y = \left| \left(\frac{df}{dx} \right) \Delta x \right| \tag{14}$$

in which the derivative is evaluated at the measured value of x and the absolute value bars ensure a positive value for the uncertainty.

Let's apply this method to the example of the previous section, in which you have measured an angle $\theta = 25 \pm 2^\circ$ and wish to calculate a quantity s that is related to θ according to $s = f(\theta) = 5\cos\theta$. Before plugging in to Equation (14), we need to convert the angle measurements to radians[6]. The conversion gives $\theta = 0.4363 \pm 0.0349$ radians (I keep a few extra digits to make sure that the conversion calculation doesn't introduce rounding errors. Next, using Equation (14), we have:

$$\Delta s = \left| \left(\frac{df}{d\theta} \right) \Delta\theta \right| = |(-5\sin(0.4363 \text{ rad}))\,\Delta\theta| = |(-2.113)\,(0.0349)| = 0.07 \tag{15}$$

As in the previous section, we find $s = 4.53 \pm 0.07$, which demonstrates that the uncertainties are small enough that the differential error propagation method is suitable for this example.

4.3 Propagating multiple sources of error

Now we are ready to generalize to the situation of a desired quantity y that depends on two or more measured quantities that each have their own error. As a concrete example, imagine that you want to measure the surface area A of a rectangular table by measuring its length l and width w. The surface area is determined by the equation

$$A = lw \tag{16}$$

and l and w each have an uncertainty Δl and Δw that contributes to the total uncertainty on the surface area ΔA. To determine ΔA, first calculate the uncertainty due to l only (call this ΔA_l, meaning the "uncertainty on A due to the uncertainty on l") while assuming that w is fixed at its measured value, using differential error propagation or the worst-case scenario method described above. Next, calculate ΔA_w (the uncertainty on A due to the uncertainty on w) using the same method. Then combine the uncertainties by adding them *in quadrature*[7] to find the total uncertainty on A:

$$\Delta A = \sqrt{\Delta A_l^2 + \Delta A_w^2}$$

For each additional source of "input" error, there will be another Δ term that is squared underneath the square root sign.

Let's practice this method with a situation in which $l = 2.4 \pm 0.2$ cm and $w = 4.1 \pm 0.4$ cm. We get the following using the worst-case scenario approach:

[6]The reason is that the differential $d\theta$ is outside the trig function

[7]This relies on the underlying uncertainties following a Gaussian distribution. If that is not the case, then another method might be better. We won't consider non-Gaussian cases in this class.

$$
\begin{aligned}
A_{\max,l} &= w\,(l + \Delta l) = (4.1\text{ cm})\,(2.4 + 0.2\text{ cm}) = 10.66\text{ cm}^2 \\
A_{\min,l} &= w\,(l - \Delta l) = (4.1\text{ cm})\,(2.4 - 0.2\text{ cm}) = 9.02\text{ cm}^2 \\
\Delta A_l &= \left|\frac{A_{\max,l} - A_{\min,l}}{2}\right| = 0.82\text{ cm}^2
\end{aligned}
$$

$$
\begin{aligned}
A_{\max,w} &= l\,(w + \Delta w) = (2.4\text{ cm})\,(4.1 + 0.4\text{ cm}) = 10.80\text{ cm}^2 \\
A_{\min,w} &= l\,(w - \Delta w) = (2.4\text{ cm})\,(4.1 - 0.4\text{ cm}) = 8.88\text{ cm}^2 \\
\Delta A_w &= \left|\frac{A_{\max,w} - A_{\min,w}}{2}\right| = 0.96\text{ cm}^2
\end{aligned}
\tag{17}
$$

$$
\Delta A = \sqrt{\Delta A_l^2 + \Delta A_w^2} = 1.26\text{ cm}^2 \approx 1.3\text{ cm}^2
$$

The first three lines calculate the error on A due solely to the error on l, and the second three lines calculate the error on A due solely to the error on w. The last line calculates the total uncertainty. Now we just need to find the best measured value of A using $A = lw = 4.1\text{ cm} \times 2.4\text{ cm} = 9.84\text{ cm}^2$. The measurement of the area is $A = 9.8 \pm 1.3\text{ cm}^2$.