Treatment of Data

Contents

Introduction

This document is an introduction and discussion of most of the error analysis that you will need for this laboratory course. Throughout this course, you will need this packet as a reference.

Experimental Error

Essentially all experimental quantities have an uncertainty associated with them. The only exceptions are a few defined quantities such as the speed of light in a vacuum, which is defined to be 299 792 458 m=s, and the second, which is defined to be the time taken by 9 192 631 770 vibrations of the light of a specified wavelength emitted by a cesium-133 atom. In other words, we have used the fact that the speed of light and this particular cesium transition can be precisely measured, to define the meter. The uncertainty in physical measurements becomes of crucial importance when comparing experimental results with theory. The basis of the scientific method is to test our hypotheses against experimental data. It is therefore very important that you learn how to estimate how large an uncertainty or error to attach to your results due to the uncertainties in your measurements.

Note that estimating uncertainties in measured quantities is entirely different from calculating by what fraction or percentage the measured value differs from the accepted value. Even the sloppy experimenter can sometimes manage to obtain an experimental result very close to the actual value just by blind luck. Furthermore, it may be possible that either no previous measurement of the quantity exists, or the whole point of an experiment may be to test a theory which predicts a value for that quantity. In either case, it is important for the experimenter to be able to answer the following question: When someone with better equipment measures the same quantity, is the new result likely to fall within the quoted uncertainty of the present measurement?

In quoting an uncertainty, there is a standard format which is normally followed. For example, when quoting a measurement for the acceleration due to gravity g you might write:

$$g \pm \delta g = (9.83 \pm 0.05)\ m/s^2$$

Here, g represents the best measurement you have for the acceleration due to gravity, and g is your uncertainty in that value. This expression means that the acceleration due to gravity (as determined by the experiment) most likely falls in the range between $9.78\ m/s^2$ and $9.88\ m/s^2$.

The assignment of probable errors to physical data is not easy. Some sources of error can be estimated fairly accurately; other sources may be difficult or impossible to estimate. The history of physics has many notorious examples of experimenters who have grossly underestimated the errors in their measurements. This is partly the result of human nature hoping for a precise result and partly the result of lack of knowledge. Sometimes there are sources of errors that the experiments didn't know about; sometimes they knew about a source of error but didn't know how to properly estimate the effect on the experiment.

Even if you don't go into scientific research as a career, what you learn about probable errors in this course will be useful to you. We are constantly deluged with data and statistics. Sometimes it is interpreted correctly by the experts; sometimes the experts present an incomplete picture that tends to support their own opinions. In this course you will learn enough to help you understand how to evaluate data.

In the following sections, we give a brief summary of definitions and practical information on errors and treatment of data. It is a bit simplified, but it will serve as a useful introduction to the subject. If you are interested in more details, or in the derivations of the equations and conclusions presented, a list of references is given at the end of this document.

Random error: This type of error presumably arises from random effects, which are unpredictable variations in the conditions of the experiment. Random effects give rise to variations in repeated observations of the quantity being measured. Although it is not possible to compensate for the random error of a measurement results, it can usually be reduced by increasing the number of observations.

Systematic error: This type of error often arises from systematic effects, which are usually associated with a particular instrument or experimental technique. Systematic effects occur in every measurement, and are the same for each observation; thus systematic errors cannot be eliminated by increasing the number of observations. However, if a systematic error arises from a recognized systematic effect and can be quantified, a correction can be applied to compensate for the effect.

In most experiments, a combination of random and systematic errors are present at the same time. The distinction between random errors and systematic errors is often expressed using two words that are usually confused: accuracy and precision.

Accuracy: This is a qualitative description of how close the result of a measurement is to the true value of the quantity being measured. The accuracy of a measurement may not necessarily be known at the time the experiment is performed. Accuracy is more strongly associated with systematic effects, although it can be affected by random effects.

Precision: This is a qualitative description of how well repeated measurements of the same quantity agree with each other. Precision is strongly associated with the smallest units that an instrument can measure as well as random effects, although it can be affected by systematic effects.

Accuracy and precision are qualitative descriptions, not quantitative descriptions. However, you can use numerical values to decide how to describe the accuracy or precision of a measurement. The generic relation between random errors, systematic errors, accuracy, and precision can be depicted in Figure 1.

For example, two students set out separately to measure the acceleration due to gravity near the Earth's surface. The standard acceleration due to gravity has a value defined as $9.806\ m/s^2$. The true value of the acceleration due to gravity may vary by as much as 0.7% of this standard and is location dependent.

Student A: $g \pm \delta g = (9.81 \pm 3.00)\ m/s^2$

Student B: $g \pm \delta g = (8.44 \pm 0.01)\ m/s^2$

Student A has a more accurate result than student B. However, student B has a more precise result than student A. If each student has done many trials to obtain their result, then the results tell us something very important. The results of student A's trials differ widely but average to very near the true value. This may indicate that the measurement error is mainly statistical (random). The results of student B's trials are very consistent, but different from the true value. This may indicate that the measurement error is mainly systematic.

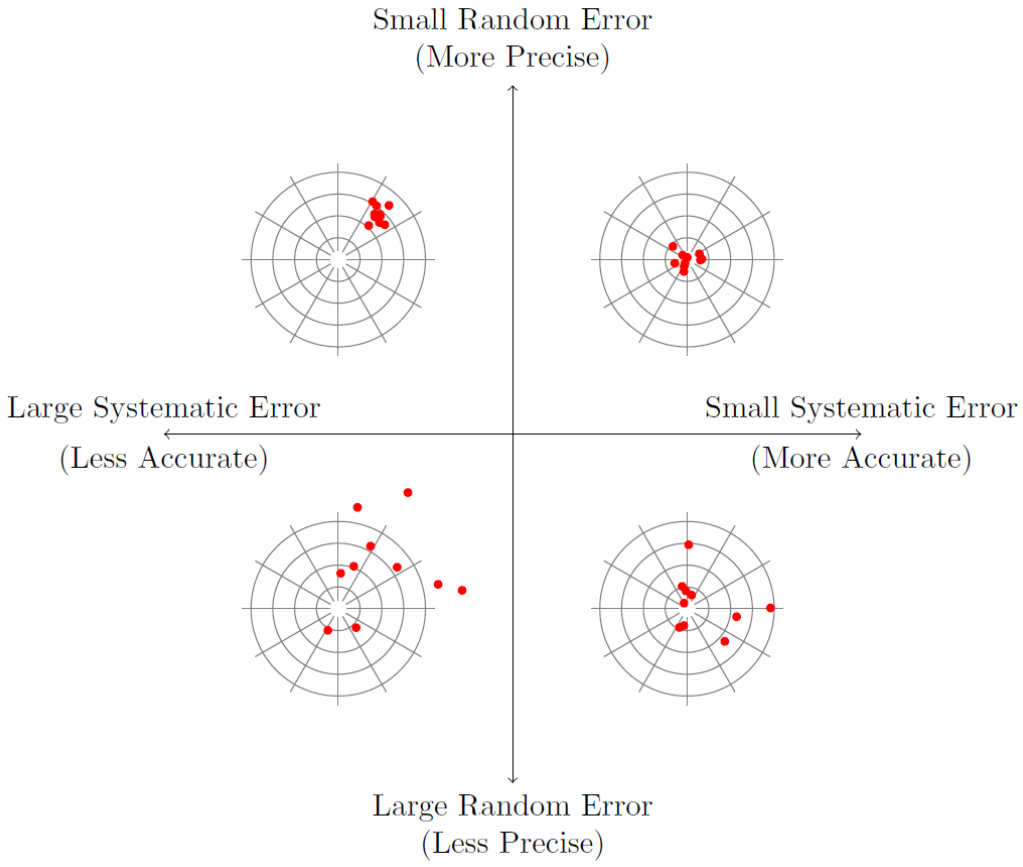

Figure 1: Examples of how systematic and random errors affect measurements and their relationship to precision and accuracy. In each quadrant of the diagram, the center of the circular target represents the true position of an object being measured. Each dot represents the position determined by measurements of the object's position.

Uncertainties for Measured Values

Multiple Measurements—The Normal Distribution

Suppose we consider a measurement whose result can take on a continuous range of values, i.e. the value is not constrained to being an integer such as in a counting experiment. To be concrete, imagine a very simple experiment. You want to measure the time it takes a ball to fall 1.020 m. To get an accurate value you use a computerized timer that reports values to three decimal places and repeat the measurement 200 times. Figure 2 shows the time data collected from our hypothetical experiment in the form of a histogram. The vertical height of each bar gives the number of measurements that fell within the range of the bin. For example, there were 13 measurements with a falling time between 0.435 s and 0.440 s.

The distribution in Figure 2 is typical of what real data from a well designed and executed experiment might look like. The most notable features are:

1. Measurements with values near the center of the distribution occur more often than measurements with values near the edges of the distribution.
2. The distribution is reasonably symmetric about its center. There is no obvious skewing toward the high or low side.

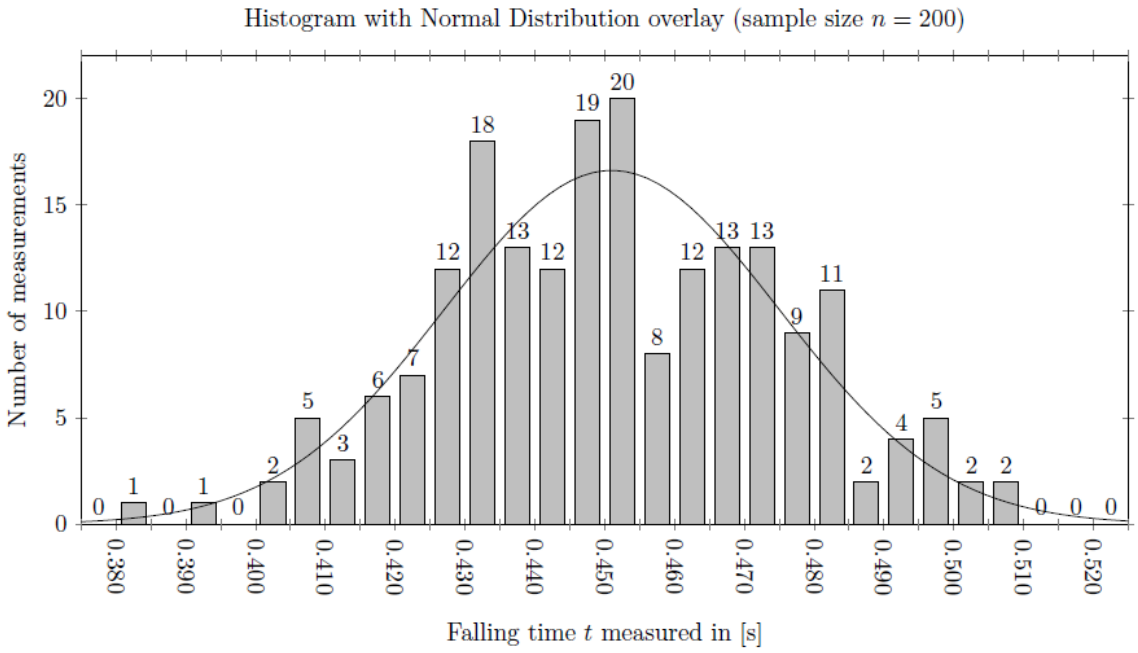

Figure 2: This graph uses 200 points of data that has been grouped into bins of width $0.005\ s$. The mean value of the data is $0.451\ s$, and the standard deviation is $0.024\ s$. The normal distribution curve using these statistics has been overlayed on the histogram.

If we continued to make measurements and made the bins very fine, our histogram might begin to look like the smooth bell-shaped curve. This curve is the limiting case in an ideal situation. It is referred to as the normal distribution or Gaussian distribution. Measurement errors that follow this distribution are said to be normally distributed. The normal distribution can be described by its center and its spread. The mathematical theory of the Gaussian curve is important, but much of it is beyond the scope of this handout. What is important is that if the fluctuations in a given quantity are random, then a distribution resembling a normal distribution is usually found and the results of mathematical theory can be used to estimate uncertainties in our data.

For our purposes, suppose we have a data set with n measurements of a value x. We want to state our results for x as $\bar{x} \pm \delta\bar{x}$, where $\bar{x}$ is the **mean** of the sample and $\delta\bar{x}$ is u_x, or the **standard uncertainty in the mean**.

$\bar{x}$: (said "x-bar") This is the mean of the sample and is one way of describing the center of a distribution. The mean $\bar{x}$ is found by taking the (arithmetic) average of all the values in the sample $(x_1, x_2, x_3, \ldots, x_n)$.

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + x_3 + \cdots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

s_x: (Said "S sub X" or just "S X") This is the **standard deviation** of the sample and is one way of describing the spread of a distribution. The standard deviation of the sample s_x is related to the mean $\bar{x}$ and the number of observations n.

$$s_x = \sqrt{\left(\frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)}$$

The standard deviation provides a way to express the standard uncertainty of a measurement:

- $\approx 68\%$ of observations fall within 1 standard deviation of the mean ($\bar{x} \pm s_x$).
- $\approx 95\%$ of observations fall within 2 standard deviations of the mean ($\bar{x} \pm 2s_x$).
- $\approx 99.7\%$ of observations fall within 3 standard deviations of the mean ($\bar{x} \pm 3s_x$).

u_x: This is the **standard uncertainty in the mean**. It describes the spread of mean values taken from multiple iterations of an experiment. This is also known as the standard deviation of the mean or the standard error of the mean—be careful not to confuse it with the standard deviation of the sample. The standard uncertainty in the mean u_x is related to the standard deviation of the sample s_x and the number of observations n.

$$u_x = \frac{s_x}{\sqrt{n}}$$

There is an alternate notation taken from the calculus of variations that will be frequently used in this course:

$$\delta x = u_x$$

The standard uncertainty in the mean can be used to construct a confidence interval similar to those that use the standard deviation of the sample. In PHSX 216N and PHSX 218N, an expression for the range of most likely values is all that will be required:

$$\bar{x} \pm \delta\bar{x}$$

Note: the above discussion assumes that each of the measurements has equal intrinsic accuracy and they are therefore equally weighted in the determination of the mean.

Single Measurement—Digital or Analog Instrument

Often there will be quantities you will only measure once for an experiment, such as the length of a string or the mass of a cylinder. In these cases, the statistics of a distribution are inappropriate because a single measurement does not describe a distribution.

When you make a measurement using a **digital instrument**, the display has a smallest place value—it cannot show any number smaller. Any measured quantity that has a fractional part smaller than what the digital scale can display will be rounded.

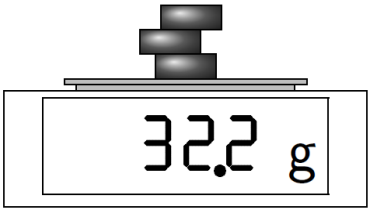

Figure 3: Digital scale reading of the mass of some objects.

Figure 3 shows an example with an electronic scale measure the mass m of some objects as $32.2\ g$. The actual mass of the objects could be any value between $32.15\ g$ and $32.25\ g$ with equal probability; any value within this range is rounded to $32.2\ g$. Thus, this value has an uncertainty of $\delta m = 0.05\ g$. However, in order to keep the number of significant figures in a given result consistent (more on that in a later section), we round this uncertainty up to $\delta m = 0.1\ g$.

These results for the mass m would thus be stated as:

$$m = (32.2 \pm 0.1)\ g$$

Simply stated, **the uncertainty for a single digital measurement can be stated as a 1 in the measured value's smallest decimal place**.

Another type of instrument you might use for a single measurement would be one with an **analog scale**, such as needle gauges and meter sticks. Generally, a range of measurable values is visible and you have to pick a physical position on the device to read the measurement.

Unlike the digital scale, the analog scale requires your judgment to report a value. For an analog instrument with very small divisions (such as a ruler with millimeter markings), you would likely state your results in multiples of that division (i.e. in millimeters as opposed to, for example, centimeters or half-millimeters). However, Figure 4 shows an example in which you would likely state your results in units *smaller* than the smallest scale division.

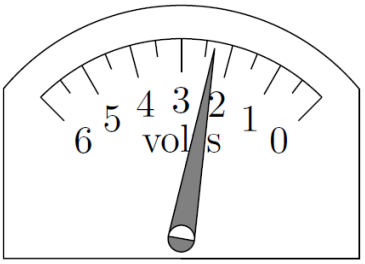

Figure 4: Analog scale reading of an electric potential.

Figure 4 shows an analog voltmeter's needle gauge, which has markings every $0.5\ V$. However, the large separation between tick marks allows for an estimated measurement value more precise than either $2.0\ V$ or $2.5\ V$. You will need to use your best estimate to judge the measurement value. In this example, the reading for the electric potential V might be $V = 2.3\ V$ (although you might estimate it to be $2.4\ V$).

For our purposes, **the uncertainty for a single analog measurement can be stated as a 1 in the stated result's smallest decimal place**. In this example, that would be:

$$V = (2.3 \pm 0.1)\ V$$

In other words, the electric potential is most likely between $2.2\ V$ and $2.4\ V$.

Error Propagation—Uncertainties for Calculated Values

Usually, you cannot make a direct measurement of the quantity you are interested in, but must instead measure other quantities and then calculate the desired quantity from them. Determining the uncertainty in the calculated value can be done if the uncertainties in the measured values are known.

The rules for error propagation can be readily derived using calculus; some results are stated here. The quantity to be calculated is Q and its uncertainty is δQ. The value of Q will depend on measured quantities A and B, which have uncertainties δA and δB. The values c, m, and n are constants, or quantities with negligible uncertainty.

Sum Rule of Error Propagation:

- If $Q = A \pm B$, then

$$\delta Q = \sqrt{(\delta A)^2 + (\delta B)^2}$$

Power-Product Rule of Error Propagation:

- If $Q = cA^m B^n$, then

$$\delta \mathrm{Q} = |Q| \sqrt{\left(m\frac{\delta A}{A}\right)^2 + \left(n\frac{\delta B}{B}\right)^2}$$

Using the Power-Product Rule, it can be shown that:

- If $Q = cA$, then

$$\delta Q = |c|\delta A$$

- If $Q = cA^m$, then

$$\delta Q = \left|cmA^{(m-1)}\right|\delta A$$

"Brute Force" Method of Error Propagation:

If the functional form for Q is not covered by either the Sum Rule or the Power-Product Rule, you can calculate the uncertainty δQ using the "Brute Force" method.

If Q is a function of a single variable A, calculate Q using the mean value of A, then recalculate Q using $A + \delta A$. The positive difference between the two results will give δQ:

- If $Q = f(A)$, then

$$\delta Q = |f(A) - f(A + \delta A)|$$

If Q is a function of more than one variable, vary each separately. Then combine the separate terms in quadrature to find δQ:

- *If* $Q = f(A, B, \ldots)$, then

$$\delta Q_A = |f(A, B, \ldots) - f(A + \delta A, B, \ldots)|$$

$$\delta Q_B = |f(A, B, \ldots) - f(A, B + \delta B, \ldots)|$$

$$\vdots$$

$$\delta Q = \sqrt{(\delta Q_A)^2 + (\delta Q_B)^2 + \cdots}$$

Often, one or two of the terms used to calculate the uncertainty δQ will dominate over the others; the others will give negligible contribution to δQ. Typically, if the dominant term comes from a single measurement from a digital or analog scale, then the experiment should probably be redesigned.

Rounding, Significant Figures, and Scientific Notation

Do not round numbers until you arrive at a final result. Rounding between steps when calculating can introduce mistakes in your calculations. If you do have to round between steps, be sure to carry several extra digits to avoid rounding errors. Be aware that some operations are more sensitive to rounding than others.

In general, most uncertainties will only have one or two significant figures. Unlike usual rounding procedures where you can round up or down, uncertainties should always be rounded up.

The number of significant figures you give in any result ($x \pm \delta x$) should be consistent. This means the smallest place value in both numbers is the same.

incorrect:	9.5876543 ± 0.67789796	10 ± 0.0141026
incorrect:	9.5876543 ± 0.68	10 ± 0.015
correct:	9.59 ± 0.68	10.000 ± 0.015

Additionally, if you express your values in scientific notation, then they should have the same power of ten. This means the mean value will be in proper scientific notation, however the uncertainty will likely not be in proper scientific notation. The underlying reason for this practice is ease of readability. Do NOT use computer E notation unless you are communicating with a computer or discussing how to enter numbers into a computer.

incorrect:	$1.6 \times 10^{-19} \pm 1.2 \times 10^{-21}$	WRONG:	$1.6\text{E-}19 \pm 1.2\text{E-}19$
incorrect:	$1.602 \times 10^{-19} \pm 1.2 \times 10^{-21}$	WRONG:	$1.602\text{E-}19 \pm 1.2\text{E-}19$
correct:	$1.602 \times 10^{-19} \pm 0.012 \times 10^{-19}$	WRONG:	$1.602\text{E-}19 \pm 0.012\text{E-}19$
correct:	$(1.602 \pm 0.012) \times 10^{-19}$	WRONG:	$(1.602 \pm 0.012)\text{E-}19$

Making Graphs

Graphs are a very useful means of exhibiting data so they can be readily visualized. Graphs can be used for smoothing out data and often provide a convenient means of analyzing data to provide a "best fit" value of some physical quantity. It is assumed you have had some experience with graphing data and this section will deal mainly with the finer points.

Choosing the scale: The scale of your graph is often a matter of usefulness and aesthetics. Often you will need to make more than one attempt to find a scale that is satisfactory. The resolution of your graph should reflect the uncertainty of the data; if the uncertainties are on the order of 1% of the measured values, then you will want a finer resolution to the graph's scale, whereas if they are on the order of 20% then the scale can be coarser. If you were to analyze the data based on the visual representation of the data, then the scale can greatly affect your results.

Error bars: Usually the quantities you are plotting will have an uncertainty associated with them. The uncertainty is indicated by drawing a line through the point to indicate the uncertainty. For example, if the velocity of an object is determined to be $v_x \pm \delta v_x = (1.80 \pm 0.15)\ m/s$, you would plot the point at $1.80\ m/s$ and the vertical bar through that point would extent from $1.65\ m/s$ to $1.95\ m/s$.

The Best Fit Curve: Perhaps in the past you did data analysis by drawing lines by hand, but in this course all data will be analyzed by numerical means, such as a linear regression. In practice, graphical analysis is often used to see if your data are consistent with a theory and perhaps to determine the value of some constant in

the theoretical expression which best fits your data. For example, an object dropped from rest is expected to have a velocity v at time t of

$$v = gt$$

where g is the acceleration due to gravity. This equation describes the velocity as a linear function of time. The value of g could be determined by graphing different experimental values for v and t, then determining the slope of the resulting line.

Suppose, however, that your data consisted of various values of falling distances h and falling times t of a ball dropped from rest. Then, we expect

$$h = \frac{1}{2}gt^2$$

This is the equation for a parabola which would be more difficult to fit graphically than a linear equation. In fact, many functions you might wish to model are not linear. Since solving linear equations is usually easier than non-linear equations, there are techniques for re-expressing various classes of equations as linear equations. Linearizing a function means to re-express the parameters in the equation so that the resulting function describes a straight line. Some functions cannot be linearized and must be solved with advanced techniques. However, you will find with most of the labs for PHSX 216N and PHSX 218N that linear fits are possible.

Linearizing Quadratic, Exponential, and Power Law Functions

Linearizing Quadratic Functions

For a quadratic function, linearizing requires no special mathematical transformations. You can group the parts of the equation so they correspond to the different parts of the slope-intercept form of the equation of a line.

General equation of a line: $y = m x + b$

Meaning of each variable: [vertical variable] = [slope] [horizontal variable] + [vertical intercept]

$$h = \left(\frac{g}{2}\right) t^2$$

Now the squared time t^2 is the quantity to be plotted on the horizontal axis. After determining the slope of line fit to the data, the acceleration due to gravity can then be found.

$$[slope] = \frac{g}{2}$$

$$2 \times [slope] = g$$

You will notice the value of the intercept is predicted to be zero before you even perform the experiment. After finding the best fit line, you will discover the intercept is unlikely to be zero. The size of the deviation between the experimental value of the intercept and the theoretical value of the intercept can sometimes be used as a clue to the accuracy of the experiment.

Linearizing Exponential Functions

In some cases the quantity that is to be determined by experiment is in the exponent. Two common examples of this are radioactive decay (exponential) and the period of a simple pendulum (power law). In radioactive decay, the number of counts per second R is expected to vary with time as

$$R = Ce^{-\lambda t}$$

where C and λ are quantities to be determined by measuring R as a function of time.

Fitting the data directly with an exponential equation to determine λ would be tedious. However, taking the natural logarithm of both sides of the equation results in an expression that can be simplified to something less cumbersome.

$$\ln(R) = \ln\left(C\mathrm{e}^{-\lambda t}\right)$$

$$\ln(R) = \ln(C) + \ln\left(e^{-\lambda t}\right)$$

$$\ln(R) = \ln(C) + (-\lambda t)$$

This equation has the same form as the slope-intercept form of the equation of a line. The slope of this line is equal to $-\lambda$ and the vertical intercept is equal to $\ln(C)$.

$$y = m\ x + b$$

$$\ln(R) = (-\lambda)\ t + \ln(C)$$

Linearizing Power Law Functions

A power law is a function of the form

$$y = ax^m$$

where the coefficient a and the exponent m are to be determined. Just as with the exponential function, taking the natural logarithm of both sides proves to be useful.

$$\ln(y) = \ln(ax^m)$$

$$\ln(y) = \ln(a) + \ln(x^m)$$

$$\ln(y) = \ln(a) + m\ln(x)$$

Again, the equation now has a form that corresponds to the slope-intercept form of the equation of a line. The slope of the line is equal to m and the vertical intercept is equal to $ln(a)$/

$$y = m\ x + b$$

$$ln(y) = m\ ln(x) + ln(a)$$

The Linear Fitting of Data

You will be provided with Python plotting templates, which you will update with your actual data to be plotted and which will in turn use the Linear Least Squares Fitting Data method to calculate the slope and the vertical intercept of the "best fit" line, each with their own uncertainties.

The "goodness of fit" will also be calculated for you; if the values you determined for the slope and intercept give a line with a good fit, then you would expect the "goodness of fit" value to be close to the number of plotted data points n. However, it is likely that you get a value different from n. If you get values that are either **much greater** than or **much less** than n, then you should try to determine if there are any corrections that you can make until you have a value close to n.