

Reference Guide for PHY191

Bring this Reference Guide to class each week

Contents

- 1. Kaleidagraph Essentials*
- 2. Excel Introduction*
- 3. Significant Figures*
- 4. Uncertainty Calculations*
- 5. Log Plots: Guessing Functional Relations*
- 6. Writing a Good Lab Report*

1. Kaleidagraph Essentials

Histograms: use menus Gallery|Stat|Histogram|*pick data*|New. To **Change Bin Size:** Plot|Plot Options|SpecifyingTheBinSize|OK, then Plot|Axis Options|*enter new bin size*|OK

Statistics Highlight data column, then Function|Statistics and either write down results and |OK, or |Clipboard|OK and then paste where you want them (Plot, Layout, or Word).

y vs. x scatterplot (*x is horizontal!*): Gallery|Linear|Scatter|*select x and y variables*|NewPlot

Curve Fitting CurveFit|General|EditGeneral|Add|NewFit|*give new fitname in box*|Edit|
Entering equation: say want to fit a parabola $y = \text{parabola } a + b x + c x^2$. Kgraph's convention for the first fit parameter is m1 (and so on). So you could enter in the Edit box:

$$m1 + m2*x + m3*x^2; m1=1; m2=1; m3=1$$

where m1 stands for your parameter a, m2 for b, m3 for c. The **extra equations** give *nonzero* guesses for the parameters which the fit routine **must have** to start from. Now, having entered the expression you are fitting, OK|OK|CurveFit|General|*fitname*|*check the box*|OK

If you'd wanted an equation of the form $C + a \cos(kx + b)$, you'd use $m1+m2*\cos(m3*x + m4)$. To **display the fit equation**, Plot|DisplayEquation. This shows you the fit parameters, and their uncertainties (labeled "errors"). The value of m1 is the best fit value of a, etc.

Moving it: *Click ParamsBlock*|Cut(Ctrl-X)|Ctrl-L|Layout|SelectPlot|*plot*|Paste(Ctrl-V)|Drag *ParamsBlock*; Ctrl-L is equivalent to Windows|ShowLayout|*KgraphLayout*

Plot Resizing: Right-click within plot | SetPlotSize | *click both Axis and Frame size* | *Enlarge Frame size Y*, but to less than paper size | drag Fit Parameters block to below the plot.

Residuals: To make a new column in the data window containing fit residuals, *immediately* after you've performed the fit (or redone the fit): CurveFit|General|*fitname*|View|CopyResiduals. Find the right data window by clicking on the small grid in the upper right of the plot window; you can **edit** the data window **column heading** by double-clicking it.

If the function fit doesn't look much like the data, you probably need to give **better initial values** than 1. Do this by CurveFit|General|*fitname*|Define| then edit the initial values in the box:

$$m1 + m2*\cos(m3*x + m4); m1=45; m2=10; m3=30; m4=180$$

Why might these be better values? The C constant (m1) should be near the average y value; the m2 should be the amount above or below the average; m3 should be such that the range Δx of one cycle of x has $\Delta x*m3=360$ degrees (**if Degrees is checked**) or 2π if Radians is checked; and 180 (Degrees) would be appropriate for a maximum the middle of the range.

Remove a previous fit: CurveFit|General|*fitname*|Deselect

Multiple Plots per page: Windows|ShowLayout|*click on layout*| then a Layout menu appears. Layout|SelectPlot|*click on plot*| for the plots you want on the page (2 or 3 will be big enough to read). Then Layout|ArrangeLayout| and select say 2 or 3 rows and one column to arrange your plots. You can drag, resize, or delete the plots; add text; or paste a clipboard.

It's really best to just jump in, but the Kgraph tutorial (examples 2&4) offer more detail.

2. Introduction to EXCEL (and Tutorial)

This introduction will show you the basics you need for this lab. Open *Excel* by double-clicking the icon with the mouse. Then open a new spreadsheet by File | New (and possibly Blank Workbook). Files that end in “.xls,” are Excel spreadsheets. **You can copy and paste data from the *Excel* spreadsheet to the *Kgraph* spreadsheet.** *Kgraph* can open a spreadsheet too, and if it’s just data and single-cell column headers in row 1, it will label the columns.

The *Excel* spreadsheet is made up of rectangles called “cells.” To enter text into a cell, click on the cell with the mouse, type the desired text, and press Return or Enter. To execute a **formula** in a cell, you must **always first type “=”** (an **equals** sign), followed by the desired formula. For example, to compute 5+6, type “=5+6” into an empty cell and press “Enter.” The answer, “11,” appears in the cell where you entered the formula. The formula you entered in the cell appears at the top of the screen under the menu bar. This works for any cell: to see a formula, just click on the cell and the formula for that cell appears at the top of the spreadsheet.

Now we will fill a **column A** labeled X with numbers ranging from 0 to 1.4 in steps of 0.2. Do not take the time to do this all by hand! You can have *Excel* do it for you. In cell **A1** enter X and type Enter. In the next cell of the column, **A2**, enter 0, our first value. Then in **A3**, we can give *Excel* a formula we want it to follow in order to fill in the rest of the column. In each cell, we want it to add 0.2 to the value in the cell just above. That will fill in the column in steps of 0.2. Therefore in cell **A3** enter the formula “=A2+0.2” and press return. Highlight 7 cells of the column, starting with the cell that has the formula in it, A3. Then go to Edit | Fill | Down and then let go of the mouse button. Your column should be filled in increments of 0.2.

In the next **column B** (Label it Sin(x) in **B1**) you are going to compute the sine of X using a formula. In **B2** type “=Sin(A2)” and hit Return. By typing **A2** in the parentheses we are telling *Excel* to take the sine of the value (in radians) which it finds in **A2**. Then you can use the same process as above to “fill-down” the entire column. If you do this just as described above, *Excel* knows to start with **A2**, take the sine of each successive value in the A-column, and place the new value in each successive cell of the B-column. Now repeat this same procedure, only this time for cosine X in the **C-column**.

More hints:

- If in doubt, use parentheses to make sure things get calculated in the right order. For example, $3 + 5/2$ results in 5.5, but $(3 + 5)/2$ results in 4. It would have been better to use $3 + (5/2)$ in Excel for the first case.
- To **check your spreadsheet formulas**, type Ctrl-` (that is, hold down the Ctrl key and type the ` key (which has a ~ above it). If you now print your spreadsheet, you can see all the formulas. Save it with a new name, adjust the column width, and use page setup to print landscape, say to fit to 2 pages wide x 1 tall. You can use Print Preview before printing to check.
- Typing Ctrl-` again toggles back to showing the calculated numbers.
- Instead of typing a cell number (say **O27**) you can click on the cell while entering the formula.
- You can change formatting with the % and .00 → .0 and .0 → .00 buttons.
- The formula =A17 copies the contents of cell A17 to the cell containing the formula.

Copying formulas: The menu command Edit | Fill | Down is equivalent to:
 click the cell with the formula cell will be highlighted, with dark borders
 hover mouse on lower right corner mouse pointer turns from a hollow to a solid cross
 drag the mouse down over more cells the formula is “copied”

When the formula is copied, the cell references change automatically! So if your formula in A3 was $A2+.2$, when it is filled or dragged down to cell A4, it becomes $A3 + .2$; if you are doing the same set of calculations on a whole column of numbers, you can set up the formulas in the first row, then highlight and drag down the formulas in several columns at once.

A similar process allows you to copy formulas sideways (across columns), or copy and paste into another column. Put 100 in D2. Next, enter $=D2$ in D3 to copy D2's contents. Now highlight A3:A8, hit copy (Ctrl-c) and then move the mouse to D4 and paste (Ctrl-v). This fills the **D column** with numbers spaced by .2, but starting from 100.

Sometimes this automatic renumbering isn't what you want to do, so you need to be able to stop Excel from adjusting cell numbers. One way to always refer to single cell is to give it a name. Type 10 into cell E1. Highlight E1, then click your mouse in the name box (it will say E1) left of where the formula appears. Type *fred* there and when you use the name *fred* in your formulas and it will always refer to that specific cell. Enter in E3 the formula $=fred^A3$ and copy it down through E4:E9. You should see numbers ranging from 1.58 through 25.12 in the **E column**.

When your spreadsheet is complete, make sure you save and print!

Common operations in Excel

you can substitute cells for numbers with these operations

What?	Mathematics	Excel Equation	Value
Addition	$11 + 12$	$= 11 + 12$	23
Subtraction	$29 - 21$	$= 29 - 21$	8
Multiplication	30×15	$= 30 * 15$	450
Division	$44 / 12$	$= 44/12$	3.66666667
Example	$3 + \frac{4}{5 \times 2} - 3 \times 7$	$= 3 + (4/(5*2) - (3*7))$	17.6
Power	6^3	$= 6^{.3}$	1.7117
Exponential	$e^{6.1}$	$= \exp(6.1)$	445.86
Square Root	$\sqrt{3^2 + 4^2}$	$= \text{sqrt} (3^2 + 4^2)$	5
Sine or Cosine	$\text{Sin}(x)$ or $\text{Cos}(x)$	$= \text{sin}(x)$ or $= \text{cos}(x)$	$x = .5$ (radians): $\text{sin} = .479$
Pi	$\text{Cos} (2 \pi x)$	$= \text{cos}(2*\text{pi}()*x)$.809 for $x = .1$

Operations on groups of cells: use as an example cells A3...A9 containing .2, .4, 1.4
The group can be denoted by A3:A9, or you can highlight a group of cells while entering the formula

Sum	$\sum a_i$	$= \text{sum}(A3:A9)$	5.6
Mean		$= \text{average}(A3:A9)$	0.8
Standard Deviation		$= \text{stdev}(A3:A9)$.316

3. Significant Figures

In calculations, it is always important to distinguish significant figures (s.f.) from insignificant figures in the final presented numbers. The rules in Taylor section 2.2 describe things rather well. We add the following comments.

1. The smallest digit in the **final** reported value should be of the same order as the smallest significant digit of the uncertainty. Since the error of the measurement is only an estimate of the uncertainty of the measurement, we keep only the two most significant digits (when available); we keep 2 s.f. so we can calculate t values to 2 s.f. Examples:

Incorrect	Correct
517.436 m \pm 12.34 cm (avoid mixed units)	517.44 \pm 0.12 m
24.3441364 \pm 0.0023 m/s	24.3442 \pm 0.0023 m/s
12385 s \pm 241 s	12390 \pm 240 s or better $(1.239 \pm .024) \times 10^4$ s

2. **Perform intermediate calculations to a few more figures** than you know to be significant (this avoids inadvertent loss of significance) and apply rule 1 for final results.

Rules 3-5 describe which digits are the **minimum** that are known to be significant. These rules also help when you have no explicit uncertainty calculation at the end, and are just calculating from inputs with a fixed number of significant digits.

3. When adding or subtracting two numbers, the result should have the same number of the significant digits after the decimal point as the least precise summand. Example:

4. When multiplying or dividing, the result should have the same total number of the significant digits as the least precise multiplier. Example:

$$1234.3 \times 23.45 \approx 28940$$

5. For other operations (raising to power, square root, exponent, sine) a rough and ready rule is to keep at least as many significant digits in the final result as you had in the input. Example:

$$\text{Sin}(3.567 \text{ radians}) = - .4127$$

Complex functions can increase **or** decrease the number of actual significant digits by 1, depending on the function and the argument. For example

$$\text{Sin}(3.567 \text{ radians}) = - .4127; \quad \text{Sin}(3.568) = -.4136 \quad \text{lost nearly 1 digit: } .4127 \pm .0009$$

$$\text{But} \quad \sqrt{6.567} = 2.5626; \quad \sqrt{6.568} = 2.5628 \quad \text{gained 1 digit: } 2.5626 \pm .0002$$

Raising to powers > 1 may lose significant digits, and powers < 1 may gain them.

6. More subtle points: when you average many measurements, the standard deviation of the mean becomes small compared to the least significant digit of individual measurements, indicating that the mean has more significant digits than the individual measurements. This will normally be covered by following rules 1 and 2. Similarly, if you wish **to calculate a t value to 2 significant digits** (in order to compare it with 2.0, as suggested in the Uncertainty Calculations section below), you will need to **estimate your uncertainty to 2 significant digits**.

4. Uncertainty Calculations: Critical Things to Know

Justify your uncertainty: Give a specific reason you chose δx as the uncertainty for the measurement of x . See examples in Taylor §1.5; §3.1-3.2; and §1.6, §4.1-4.6 for standard deviation and standard deviation of a mean for repeatable measurements.

Compatibility (§1.3, 2.4-2.5): The whole point of quantitative measurement with uncertainties is to test hypotheses, and compare results. Say you measure q , and you compare it to p (the expected value). Define the discrepancy as the difference of your result from the result expected by some hypothesis:

$$D = q - p = \text{measured} - \text{expected}$$

The best way to describe the degree of discrepancy of p and q is in terms of the number of standard deviations (the “ t value”) of their difference from expectations:

$$t = D / \delta D \quad \delta D \text{ is the uncertainty of } D \text{ (its standard deviation, for Gaussian}$$

uncertainties).

The “**two standard deviations**” rule says p and q are compatible as long as $|t| \leq 2$.

Typically $\delta D = \sqrt{(\delta q)^2 + (\delta p)^2}$; or just δq if p is well known (so δp is tiny). Best practice is to calculate t , then say something like “the difference is 1.6 times its uncertainty, so the measurements are compatible by the 2 standard deviation rule.” If $|t| > 2$, we would call p and q statistically incompatible, or call their difference statistically significant. To obtain a t value of 2 significant digits, keep 2 s.f. in uncertainties.

If your uncertainties are Gaussian, and correctly estimated, and the assumptions (hypothesis) leading to the expected value are also correct, a $|t| > 2$ deviation would occur by chance only about 5% of the time. So large $|t|$ values suggest real disagreement from what you expected, while small $|t|$ values are compatible—not proven to disagree. If you measure poorly (δD is large), your result is compatible with almost anything: not a very useful measurement.

Occasionally we use a simpler criterion compares $|D|$ with $\delta q + \delta p$ (the worst case for δD , but allowing only 1 standard deviation difference): this is just “do the error bars touch”.

We are often also interested in the **fractional deviation** the measured value from what we expected, which is just $D/p = (q-p)/p$. The $D(\%)$ (**% deviation** or **% difference**) is the same thing expressed in percent. D/p or $D(\%)$ is all we can report if we don’t know δD . But just because the percent difference is small, does not necessarily make it insignificantly different statistically. That’s what the t criterion is for.

Know the Uncertainty Calculation Formulae (§3.3-3.7; 3.11) on inside covers of Taylor, and how/when to use them. Some hints:

For $q = x \pm y$, x , y , q , dx , dy , and dq must have the **same units** (so I can add $q + dq$: error bars) The fraction uncertainties $\delta q/q$, $\delta x/x$, $\delta y/y$ all have NO UNITS (can write as a fraction, or as %, but watch the factor of 100!)

But to get δq , **don’t forget to multiply $q \times (\delta q/q)$**

How to check your calculations to see if they make sense:

$$q = x + y \quad \text{always must have:} \quad \delta q > \max(\delta x, \delta y)$$

$$q = x * y \text{ or } x/y \text{ always must have} \quad \delta q/q > \max(\delta x/x, \delta y/y)$$

Independent measurement: no relationship in the *imperfections* between the measurements; e.g. 2 students measure the same distance each with a different, but good, ruler. A measurement dominated by a systematic error (same shrunken ruler used by both students) would produce results that aren’t independent. See Chapter 4; needed to apply Chapter 3 formulas.

Random: you expect to get slightly different values each time you measure it: due to reading uncertainties, varying judgments, uncontrollable factors, or inherent properties of the measurement.

Standard Deviation and Standard Deviation of the Mean

The standard deviation (σ) is a measure of the uncertainty of any single measurement. The standard deviation of the mean, $\sigma_m = \sigma/\sqrt{N}$, is a measure of the uncertainty of an average of N such measurements. Clearly, the average is better known than a single measurement.

Examples

q = x/y Often easiest to do in terms of %

$$x = 10 \quad \delta x = .10 \quad y = 2.7 \quad \delta y = .20 \quad \text{so } q = 3.7$$

$$\delta q/q = \sqrt{(1\% + 8\%)^2} \approx 8\% \quad \text{so } \delta q \approx .08 \times q \approx .3 \quad (\text{notice } 8\% \rightarrow .08, \text{ the factor of } 100)$$

Whip out your calculator now: Let's try $r=10$ and $\delta r = .1$, so what's the fractional error for r ?
 $\delta r/r = 1\%$ Now say $q=r^2$ then what's $\delta q/q = ?$

From Eq 3.23, 3.26:

$$\delta q / q = (|dq/dr| \delta r) / r^2 = 2 \delta r / r = 2\%$$

For comparison, you can calculate changes in q directly (the most general way, rather than the Chapter 3 formulas, which rely on first derivative approximations):

$$(q + \delta q)/q = (r + \delta r)^2 / r^2 = 102.01/100 = 1.0201 = (q + \delta q) / q, \text{ so } \delta q/q = 2.01\% \text{ (same as } \delta q \rightarrow 0)$$

A More Complicated Example Calculation (See Step by Step: Taylor Chapter 3.8)

$$q = x^2 y + z^{1/3} \quad x = 10 \pm .1 \quad y = 20 \pm .2 \quad z = 10000 \pm 1800$$

$$\delta x/x = 1\% \quad \delta y/y = 1\% \quad \delta z/z = 18\%$$

$$\text{let } w = z^{1/3} = 15.8 \quad x^2 y = 2000 \quad \text{and } q = 2015.8$$

Let's start with the product term: $x^2 y$

$$\delta (x^2 y) / (x^2 y) = \sqrt{(\delta x^2/x^2)^2 + (\delta y/y)^2} = \sqrt{(2 \times 1\%)^2 + (1\%)^2} = 2.2\% \approx 2\%$$

notice we have used $\delta x^2 / x^2 = 2 \delta x/x$: **the 2 goes inside the parentheses!**

$$\text{so } \delta x^2 y = x^2 y \times (\delta x^2 y / x^2 y) = 2000 \times (2\%) = 40$$

$$\text{Now } \delta w/w = 1/3 (\delta z/z) = 1/3 \times 18\% = 6\%, \text{ so } \delta w = 6\% \times w \approx .9$$

notice that 6% is NOT rounded up to 10%, nor is .948 rounded up to 1
 in each instance we keep the first significant digit, though a real calculation, it might make sense to keep two digits or even an extra digit. Here w has *more* significant digits than z : $15.8 \pm .9$ compared to $(10.0 \pm 1.8) \times 10^3$!

$$\text{Finally, since } q = x^2 y + w, \delta q = \sqrt{(40)^2 + (.9)^2} \approx 40$$

So $q = 2015.8 \pm 40$, or $2020 \pm 40 = (2.02 \pm .04) \times 10^3$ after significant figures.

5. Log Plots and Guessing Functional Relationships

In Kgraph, from Plot | Axis Options you can choose Linear (default) or Log (logarithmic) for either the x or y axis. Let's see how this might help you: For this discussion, we will assume logarithmic means natural logarithms, log to the base e. See also Taylor §8.6. In Kgraph you can select Log from Plot | Axis Options, or make a new data column with, e.g. $\ln(c5)$ or $\log(c5)$. When using log scales, you should also use Plot | Axis Options to choose minima and maxima so that your displayed data fills much of the plot area, not just a small corner.

1) Semi-log plots

Suppose the true relationship is something like $y = y(x) = e^{-ax}$ (the constant a might be positive or negative). If you plot y on a log scale and x on a linear scale, this is a semi-log plot (logarithmic on only one axis). Then you would be looking at $\ln(y)$ vs. x , and substituting the true relationship $y(x)$, your plot would display $-ax$ vs. x .

That is, **exponential relationships produce straight lines on a semi-log plot** and the slope is equal to the coefficient of x in the exponent. To find the slope, you could perform either the nonlinear fit $y = \text{Exp}(-a x)$, or a linear fit to the semi-log plot, fitting $\ln(y) = -a x$. These fits will give somewhat different answers as equal weighting in y assumes different uncertainties than equal weighting in $\ln(y)$.

2) Log-Log plots

Suppose now the relationship between y and x is of the form $y^p = a x^q$. A linear relationship ($y = a x$, with no additive constant) is a special case with $p = q = 1$. Assuming y and x are positive, you can take the p -th root and recast of this relationship as $y = y(x) = b x^u$, where $b = a^{1/p}$ and $u = q/p$.

If you plot both x and y on a logarithmic scale, it's a log-log plot. Then you would be looking at $\ln(y)$ vs. $\ln(x)$, which, on substituting the true dependence $y(x)$ would yield $\ln(b) + u \ln(x)$ vs. $\ln(x)$

That is, **power relationships** of the form $y^p = a x^q$ **produce straight lines on a log-log plot**. The slope is the power u , to which x is raised to give y . For $y = x^2$, on a linear plot you get a parabola—not straight. But a $\log y$ vs. $\log x$ plot (for $x > 0$!) gives a straight line of slope $u=2$.

You can find this slope by rise over run by measuring distances directly on the graph:

$$u = [\ln(y_1) - \ln(y_0)] / [\ln(x_1) - \ln(x_0)]$$

This works directly from the plot, as long as the y and x scales are the same. It's the same as

$$u = \ln(y_1/y_0) / \ln(x_1/x_0)$$

which makes it obvious that the units cancel, and all that matters is that a given factor (e, or 10, say) takes up the same amount of space on the y and x axes. You could also use either formula with coordinates of points on the line.

Or you could perform either the nonlinear fit $y = b x^u$, or a linear fit to the log-log plot, fitting $\ln(y) = \ln(b) + u \ln(x)$. These fits will give somewhat different answers as equal weighting in x or y assumes different uncertainties than equal weighting in $\ln(x)$ or $\ln(y)$.

To really test these relationships well, you should ideally:

- Compare the fit quality with an alternative, such as a straight line

- Have more than 3 data points

- Have data covering at least a factor of e^2 (≈ 7.4) to 10 in range of x and y .

6. Writing a good lab report

The Notebook: Write what section number (say 4.1.1) of the lab write-up you are working on. Record your original measurements (with units) in your lab book. Include your estimates of uncertainty and their justification for the measured quantities (it might be more uncertain than the finest you can read the instrument!). All your work should be in the lab notebook - including any mistakes or duplicate measurements. Your lab notebook records what exactly you did, including any false steps. Errors should be neatly crossed out and a note should be recorded in the lab book indicating the nature of the mistake. Write in pen—no pencils or erasers in lab. This is the method used by practicing scientists for the recording of their experimental measurements. However, the lab notebook is for your reference and does not need to be included in the lab report.

Analysis of results with Kaleidagraph and Excel:

Calculate derived quantities from the original data and uncertainties

Plot data with correct labeling and uncertainties (error bars)

Class info in upper right corner, a descriptive title, labeled axes with units, etc.

Titles should distinguish graphs from each other.

Check that your results make sense: smooth graphs; consistency in tables

Find best fit lines, curves, and fit results as needed

Testing for statistical consistency: learn to use formulae in Taylor:

Use uncertainties to compare two results (calculate the t value)

Use uncertainties to compare a result with an expected value (calculate the t value)

Lab reports: The goal is to clearly communicate your analysis and results. Use the report forms (in Word format) and insert data in tables and answers to questions at the indicated section numbers. Attach Excel spreadsheets and Kaleidagraph plots as separate pages immediately after the section that requests them. **DO NOT make your grader hunt for things!** When answering questions in asked in the lab manual, make quantitative comparisons when possible, but use clear, complete sentences at all times.