Analyzing Experiment Data

This document will provide some extra guidance for getting organized and performing your analysis of an experiment's results. Consider also the AnalysisPages document on the Physics 6 website.

> **The proper analysis of a typical experiment will be involved, difficult, and time-consuming. You must give yourself adequate time to do it correctly.** *Focus on the physics and use the results to answer (or pose) interesting questions!* Make sure that you evaluate both the data and the relevant theory quantitatively and critically.

Generally, the process of organizing your thoughts and preparing for an efficient, effective analysis will involve the following steps (usually in the order presented):

1. Review the *Experiment Notes* to remind yourself of the purpose of the experiment, the relevant theory and how it will apply specifically to this experiment, and what fundamental constants (free parameters) of the theory, if any, should be determined as a result of the analysis. **Is the primary purpose of the experiment to investigate the validity and accuracy of the theory, or to use the theory and your data to accurately determine the value of a particular fundamental constant, or both?**
2. **Review how the data were acquired, what parameters were controlled but not necessarily measured, and what data were measured. What were the units of these various quantities? How do these quantities relate to the theory? To the theory's fundamental constants?**
3. During the conduct of the experiment, were any preliminary estimates of the theory's accuracy and of its free parameters calculated? Was there any clear evidence of potential inadequacy of the theory to accurately describe the experiment's results? Did the apparatus perform adequately, or were there potentially problematic data points?
4. What are the possible sources of error in the determination of the quantities, either measured of controlled, during the experiment? Which data or other experimental quantities have demonstrated lack of repeatability (noise)? What are the sources of systematic errors which may be relevant during your analysis – errors due to instrument calibration, measurements of the apparatus geometry or electrical characteristics, adjustments to the apparatus, etc. For which of these systematic error sources do you have uncertainties? Which errors can be included as additional free parameters in an augmented version of the theory? If necessary, review the *Data Analysis* text for help with how these various errors should be addressed during your analysis.
5. **Now consider how your conclusions should read – how you will organize and state your final assessment of the theory's relevance and accuracy, how you should state the values and uncertainties of free parameters, and how you should address the limitations of the experiment and offer suggestions for further study.**

6. **With this description of how your conclusions should be organized and stated, map out the logical flow and the steps required to proceed from the data to the conclusions.** How will you get the data into *Mathematica* and *CurveFit*? What algebraic transformations or other data modifications may be required before performing fits of the theory to the data? How will the optimized fit parameters map to the free parameters of the theory? How will fit residuals tell you about the theory's accuracy? How will systematic error uncertainties be incorporated into the final analysis of your results?

> Your analysis should be much more detailed and thorough than what would be found in a typical research paper in a scientific journal. In the interests of brevity, most journal articles leave out many of the actual calculations and judgements of the experimenter. The data presented are often a subset of the full data set to particularly reinforce the writers' conclusions. **This "cherry picking" of the data has no place in your analysis!**

Once you've gotten organized by going through the above steps, you can begin the detailed work to get from the data to the conclusions. Load the data into *Mathematica* and *CurveFit* and perform the necessary transformations to get it ready to be fit. Incorporate uncertainties due to noise (repeatability) by analyzing the data point scatter.

Use CurveFit to fit the appropriate function to the data using chi-squared minimization. Consider the fit residuals and evaluate the quality of the fit. Are there potential outliers in the data which need to be addressed? In this case are there in-lab comments which might help you evaluate the quality of these data points?

What is the overall accuracy of the theory over the range of parameter space you've fit? Do the fit residuals and the fit's reduced chi-squared values indicate that the theory may need improvement? Should the fitting function be augmented by adding a constant or a linear function to the fundamental theory? Should one or more subsets of the data be fit for which the theory may prove to be more accurate? Try fitting these augmented fitting functions or reduced data ranges as well. **As you investigate these issues, your assessment of the theory's fit to the data should be coming into clearer focus, and you should now know how you will organize and present these results in a way that will lead to your final conclusions.**

> Does the data show some behavior that you didn't expect to observe and that the theory doesn't address? Do not leave a "wiggle on a plot" unnoticed if it may be more than just noise. What could be going on? Is it an artifact of the apparatus or instrumentation, or might it be something more fundamental? Even if you cannot definitively explain the cause of this behavior, make sure you call attention to it. Speculations of possible causes are certainly encouraged!

Now you are ready to use the optimized fit parameter values and their uncertainties to calculate values for the theory's free parameters and any experimental systematic errors included as parameters of the fitting function. Organize the calculations which transform the fit parameter values and uncertainties into your relevant final constant determinations in a way that is easy for the reader to follow, especially the calculations required for error propagation and the inclusion of other systematic uncertainties.

Organizing and writing up your analysis and conclusions

Once you've completed the data transformations, fits, and final calculations, then you should organize the relevant data tables, fit results and plots, and calculations into a final analysis document which is clear, complete, properly labeled, and organized.

Use plenty of space and clearly highlight important findings and conclusions regarding the adequacy of the relevant theory and estimates of numerical values for the theory's free parameters. Always write down the formulas you use, and always include units with numerical quantities. Label graphs and their axes; clearly differentiate real data from theoretical curves or fits.

Do not forget to include any unexplained behavior in the data or limitations of the theory. Use such observations to suggest areas for further research.