06 January 2026

UNIVERSITY OF TORONTO

Advanced Physics Laboratory

Data Analysis Exercise
2025/2026

This assignment has been created and modified by APL coordinators.

Objective

This assignment is required of all first-time students in the Advanced Physics Laboratory to ensure one has the basic skills needed for Advanced Lab data analysis. It is not intended to be long or onerous. It tests whether one has adequate software for nonlinear fitting of models to data, and if they can think holistically and communicate clearly about uncertainty.

To complete this exercise, one needs to be able to analyse data with heteroskedastic[1] (and possibly non-Gaussian) random uncertainties in both x and y, which may also have calibration and other possible systematic uncertainties.

The primary goal of this assignment is to help students be ready to analyse real data from their experiments.

- **Do the assignment as soon as possible, before you have real experimental data.**
- **Do not hesitate to contact the Lab Coordinator – preferably well in advance of the due date – if you have any questions or need assistance. We will be happy to help you.**

The Assignment

You are given a graphics file with ellipses whose horizontal and vertical axes correspond to x and y values generated from an unknown (by you) function that is either:

Normal/Gaussian

$$y(x) = y_\mu e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

or Lognormal

$$y(x) = \frac{y_\mu}{x\sigma}\, e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}} \tag{2}$$

or Laplacian

$$y(x) = y_\mu e^{-\frac{|x-\mu|}{\sigma}} \tag{3}$$

[1] "*heteroskedastic*" just means the uncertainties of different data points are not all the same size.

[2] If you are curious about the accuracy of uncertainty estimates in published research, feel free to look at Why OUTLIERS are good for science, D. Bailey, Significance 15 (2018) 14-19, or
Not Normal: the uncertainties of scientific measurements, D. Bailey, R. Soc. Open Sci.4 (2017) 160600.

Print out the file onto a sheet of paper and carefully measure the lengths of the horizontal and vertical axes of the ellipses with a ruler. You are expected to estimate the uncertainties on these measurements, calculate the (x, y) values, fit the values to a model (based on the information above), and identify (if possible) which of the equations best describes your data. You should also discuss the sources of uncertainty in the measurements and results and how accurately you evaluated those uncertainties.[2]

Details

Your data can be found at http://www.physics.utoronto.ca/apl/simdata/ellipsesXY.jpg, where **you must replace the "XY" with the last two digits of your student number**.

Each ellipse is labeled with a number (n) from 1 to 12 and a scale factor (s). The graphics file also includes a "Calibration Square" box that allows you to convert your measurements to the units used by the equations (1), (2), or (3). To get the (x, y) values, you must multiply your measured lengths by the (x, y) calibration factors determined from the calibration box and the scale factor (s) given for each ellipse.

Example: You measure the horizontal and vertical axes of an ellipse to be 23.4 mm and 7.2 mm, and the ellipse scale factor is given as $s = 2$, and your file has a "3 unit Calibration Square" that you measure to be 9.8 mm square, then (x, y) = (23.4, 7.2) mm × 2 × (3 units /9.8 mm) = (14.33, 4.41) units.

You will, of course, also estimate the uncertainty in x and y. Don't forget that any uncertainty in measuring the calibration box will contribute to a "systematic" uncertainty common to all measurements of the ellipses. Be sure you are clear (in both your head and in your report) about what your uncertainties mean, e.g. are they the standard deviation of an expected Gaussian variation, or the full range of maximum possible values.

Submit a short and well-organized paper report on your measurements and analysis.

What to Report

1) Open the report with reporting which graphics file you analyzed (e.g. "ellipses27.jpg") and description of the ruler. Append the printed sheet you measured with your raw measurements (before calibration scaling) and their estimated uncertainties as well as the code you used at the end. Make sure your code and ellipses sheet are in the same document.
2) Describe how you measured the ellipses. Include any technical information (e.g. a description of your ruler) that someone would need to duplicate your measurements.
3) Clearly state how you estimated the measurement uncertainties.
4) A short description of your analysis. Append any code, graphs, figures, and fit results. Make sure you consider uncertainties in both x and y directions. You can use code repository available to APL students and pay attention to both uncertainties for residuals and their uncertainties.
5) A judgment about whether your data is consistent with any of equations (1), (2), or (3), none, all, or some? Present any fits used to support this judgement. State how good or bad the various fits are, based on appropriate metrics (usually $\chi 2$) and plots (regular and residual). Discuss both the fit and residual patterns.
6) Report your best estimates, with uncertainties, for the parameters of your best fits of the equation parameters, e.g. μ, σ, y_{μ}
7) Do you think your estimated measurement uncertainties were accurate, or too large or too small?
8) A short discussion of any issues that arose in your analysis, and any thoughts on how the measurement or analysis could be improved if you had more time or could get more data.

(Continued on next page)

Other Information

Rules

- The assignment is worth 4% of the course grade. Late submissions will be penalized 25% per calendar day. (The assignment is less useful to you if you don't finish it before you finish analysing your first experiment.)
- The expected length of the report is 1-2 pages, **excluding ellipse sheet with measurements, figures, graphs, and code**.
- Reminder: **It is plagiarism to submit any work without proper attribution of code or other help.** It is fine for you to talk to other students about this assignment, but any substantial discussions or shared analysis code must be acknowledged in your report. The goal is to ensure that you have the necessary skills, so simply copying another student's methods without understanding them is a very bad idea. We may notice very similar answers, so under-acknowledgement could lead to an unpleasant meeting with the Lab Coordinator and possible serious consequences.

Resources and Tools

- Notes on data analysis and error calculation and other on-line material are available at http://www.physics.utoronto.ca/apl/links.htm.
- It is up to you to choose a fitting program for this assignment. The computers in the Advanced Physics Labs have Python and Excel and are connected to a free printer in MP251. Excel can be very useful for entering data and preliminary plotting, but it is not adequate for the level of analysis expected in the Advanced Physics Lab. You are allowed to use any software that is up to the task, but Python is strongly recommended if you are not already very familiar with other software that has sufficient capabilities.
- Useful data fitting examples, including code for data with uncertainties in both *x* and *y*, can be found at http://www.physics.utoronto.ca/apl/python/index.htm - repository.
- **Do not hesitate to contact the Lab Coordinator – preferably well in advance of the due date – if you have any questions or need assistance. We will be very happy to help you.**

Note

- Using fancy image analysis methods to fit the dimensions of the fuzzy ellipses is a waste of your valuable time and will not get you any extra points. Just use a ruler or cursor to measure their width and height.