Making scientific graphs

Irina Novikova

September 4, 2007

Plotting the experimental data graphically is often the most convenient and transparent way to present results of measurements. No matter if you are making plots manually on a graph paper, or using a computer there are several basic rules of doing that.

1 Axes, scales and units

- First choose what variables you will plot along x and y axes. By convention, the *independent* variable is plotted along x axes, and the *dependent* variable is plotted along y. The independent variable is the one experimentalists control (vary) and which causes a change in the dependent variable.

> **Things to consider** When choosing the best way to analyze the data graphically it is important to consider the accuracy of the data fitting. For example, if you are doing graphical analysis manually the only fit you can do reliably is a straight line! That means that you will need to plot your data such that you expect the data to line up. That can be done by changing the variables on the graph. For example, if you need to check the spatial dependence of the Coulomb law $F = kq^2/r^2$. If you plot F as a function of r it would be almost impossible to manually fit the experimental points with the hyperbola and extract fit parameters. However, if you plot F as a function of $1/r^2$ the expected fit is linear, and its parameters can be determined from the graph with good precision.
>
> Most of the computer programs nowadays have relatively sophisticated fitting toolkits able to use many standard fitting function. However, you may need to explore the limitations of the program before making final decision of what variables you will be plotting and fitting.

- Make sure that each axis is properly labeled, including proper units in parentheses. Commonly an axis label will read "Detector displacement d (cm)" or "Stopping potential V (mV)".

- Look at the range of the experimental data before choosing the axis ranges. It is the best if the plotted curve fills the whole graph - then it will be easy to see all the details. For example, the axis range does not have to stretch all the way to zero.

- It is traditional to have the values of thick marks to be divisible by 1, 2, or 5 times some power of 10. For example, division values of 0.02 mW or 5 kΩ are acceptable, but 0.03 mW is not.

2 Data points

- All data points on a graph should be visible, but small enough to show their values unambiguously. If more than one data set are plotted in the same graph, the data points from each sets are plotted differently by using, for example, filled or hollow squares, circles, triangles, *etc.* Practical tip: if two or more curves lie very close to each other try to make their data points as distinguishable as possible. For example, small filled circles, squares and diamonds look very similar; making one of them hollow circles or triangles allows a future reader to identify the right curve quicker.

- Graphically add experimental uncertainties to the data points as bars extending up and down (for uncertainty in y value) or left and right (for uncertainty in x value). The length of the bars must be in scale! If the uncertainty is too small to plot, it is not shown. Don't forget to mention that in the figure caption!

- Don't forget to add a legend if more than one curve are present in the graph.

3   Fitting the data

The next step is fitting the experimental data with known or expected function. Here we'll discuss the procedure for a linear fit, since, as it is mentioned above, this is the only fitting which can be done reliably without computer.

There are two common (and related) purposes for fitting experimental data. First one is to verify a theoretically predicted functional dependence (*i.e.* answering questions like "Is the energy of a photon is really directly proportional to its frequency?"). The second one is to determine average value and experimental uncertainty of the parameters, for example a slope k and an intercept b for a linear fit $y = kx + b$.

- To find the best linear fit for the experimental data points it is necessary for draw a straight line which passes through all the uncertainty intervals of all points in the graph, or at least be as close as possible to all the points. Make sure there is approximately equal number of experimental data points on both sides of the fitting line.

- To find the slope of the fitting line pick to points at the ends of the line (not just two closest data points!), and calculate the average values of the slope and the intercept k_{ave} and b_{ave}.

- There are couple different ways to calculate the uncertainties of the slope and the intercept. Computer usually does the job for you using rather involved least square fit, so it displays final fitting parameters with its absolute uncertainty. When fitting is done by hand one have to add two lines with maximum and minimum slopes $k_{\max}$ and $k_{\min}$ that still fit experimental points reasonably well. Then the uncertainty of the slope Δk is defined as

$$\Delta k = |k_{\max} - k_{\min}|/2,$$

An example of such fitting is shown in Fig. 1. Note, that uncertainty estimated that way is usually larger than what computer calculates, since the method is rather crude.

If the graph is a component of the scientific text (such as lab report, journal article, *etc.*), it must be accompanied by a caption, which provide brief description of the graph content, and, together with the legend, provide specific information about both the displayed data set and the fitting function.
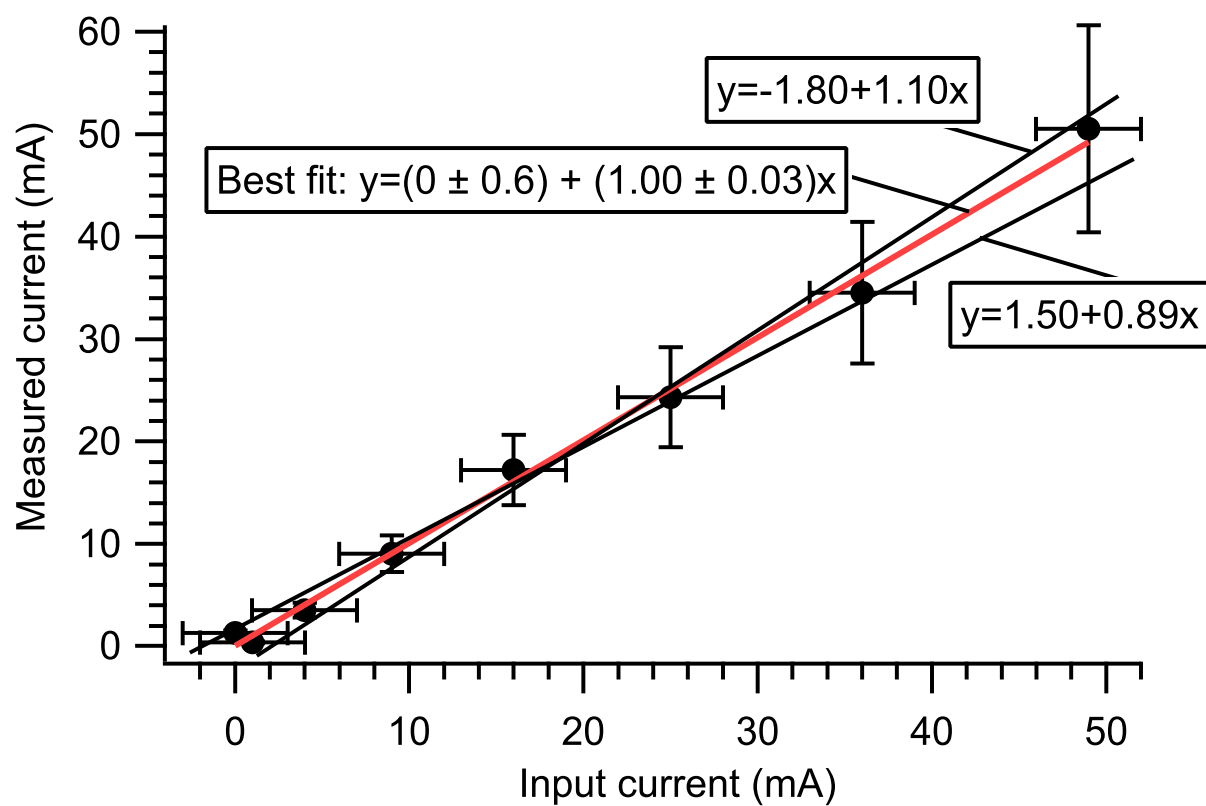
Figure 1: The current measured by the first detector as a function of input current *(example plot). Experimental data with error bars, best fit (done by computer), and maximum and minimum slopes (done manually) are shown.*