PHYS 500: Research Methodology

Propagation of Errors – The least squares method

Propagation of Errors-a

- In most of the cases in an experiment the direct measurement of some physical quantities is used for the indirect measurement of another physical quantity. For example if we measure the current I through a wire and the voltage V across it then we can estimate its resistance R using the well-known relation $R=V/I$.
- Let's assume that we have measured the voltage V with an error δV (no matter if it is a reading error or the absolute error of the average value) and the current I with an error δI. In this case we do not have a direct measurement of R and an error associated with it. Thus the errors of the voltage and current will have an impact on the error of the resistance.

Propagation of Errors-b

- The mathematical theory of errors gives us the following rule:
- Let's assume we wish to measure the quantity

$$f\left(x, y, z \ldots\right)$$

Where the quantities *x, y, z ,…* have associated errors δ*x, δy, δz ,…* Then the error of the quantity *f* is given by

$$\delta f = \sqrt{\left(\frac{\partial f}{\partial x}\delta x\right)^2 + \left(\frac{\partial f}{\partial y}\delta y\right)^2 + \left(\frac{\partial f}{\partial z}\delta z\right)^2 + \ldots}$$

Propagation of Errors-c

- Thus for the error of the resistance

$$\delta R = \sqrt{\left(\frac{\partial R}{\partial V}\delta V\right)^2 + \left(\frac{\partial R}{\partial I}\delta I\right)^2} =$$

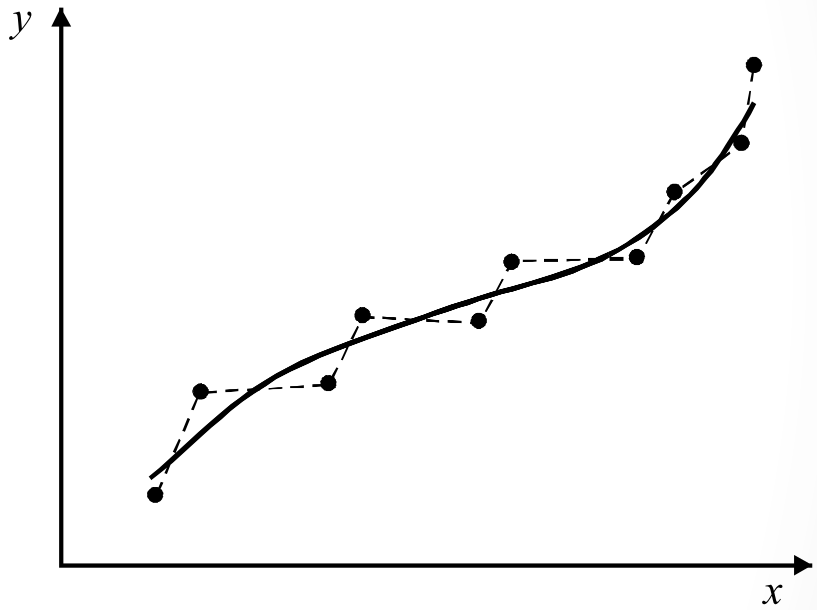
$$\sqrt{\left(\frac{\delta V}{I}\right)^2 + \left(\frac{V}{I^2}\delta I\right)^2}$$

- **Note 1:** When we do the partial differentiations and we need to substitute recordings in the place of *x, y, z* ,...we must put the corresponding results no matter if they are average values or recorded values.

How we draw a curve-a

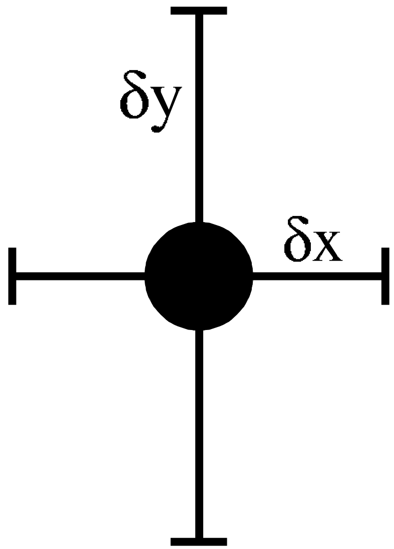
Usually the experimental recordings do not show up smoothly when we put them on a diagram. It is wrong to join them as in the dashed style shown in the figure. This is wrong because it would mean that the physical quantity changes abruptly at some points.

The most reasonable choice would be the smooth solid curve shown in the figure.

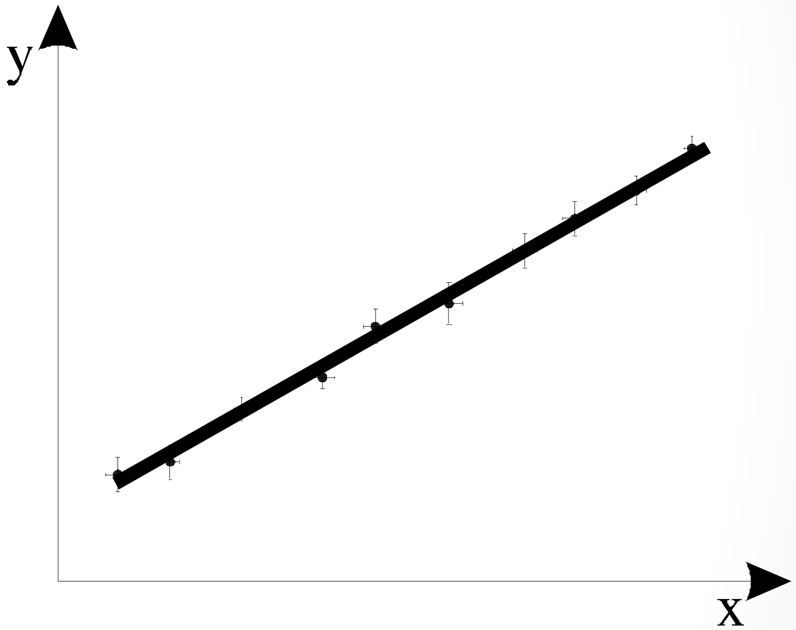


How we draw a curve-b

The previous discussion does not take into account the errors. The errors **must** always be drawn for any point which represent a recording as they are shown in the figure where δx and δy are the errors (reading errors, average value errors or whatever). The relative size of the lines corresponds to the size of the errors in the scales used for x and y.



How we draw a curve-c



Wheb we draw a curve we cannot extend beyond the leftmost or rightmost points even if we know the mathematical form of the curve because we do not know if this is true for the device we operate.

The least-squares method-a

- There are cases where we can draw the best possible curve using mathematical points. This happens when the expected curve is of a known form (e.g. straight line, hyperbola, sinusoidal, exponential).
- The method which we use is called the **least squares-method** because what is needed is that the sum of the squares of the distances of the points from the curve must be minimum.
- In this method is very crucial to define which variable is the independent (normally *x)* and which is the dependent one (normally *y*).
- All the above change if we consider that y is the independent and x is the dependent variable. In this case we get a different least squares curve.

The least-squares method for a straight line-a

- In the case of a straight line this method has as follows:

1. Let's assume we have N recordings with results x_i and y_i .
2. Let's consider that x is the independent variable and y the dependent one
3. The recordings have errors δx, δy respectively.
4. The error δy is the same for all the recordings y_i.
5. For all recordings $\delta x / x_i << \delta y / y_i$.

The least-squares method for a straight line-b

- The expected line is of the form: *y*=A+Bx with A and B and their errors are given by,

$$A = \frac{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i y_i}{D}, \quad B = \frac{N \sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{D}$$

$$D = N \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i \right)^2 \qquad \delta A = \sigma_y \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{D}} \qquad \delta B = \sigma_y \sqrt{\frac{N}{D}}$$

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^{N} \left(y_i - A - Bx_i \right)^2}{N-2}}$$

The least-squares method for a straight line: An example

- We measure the pressure and temperature of a gas and we get the following recordings:

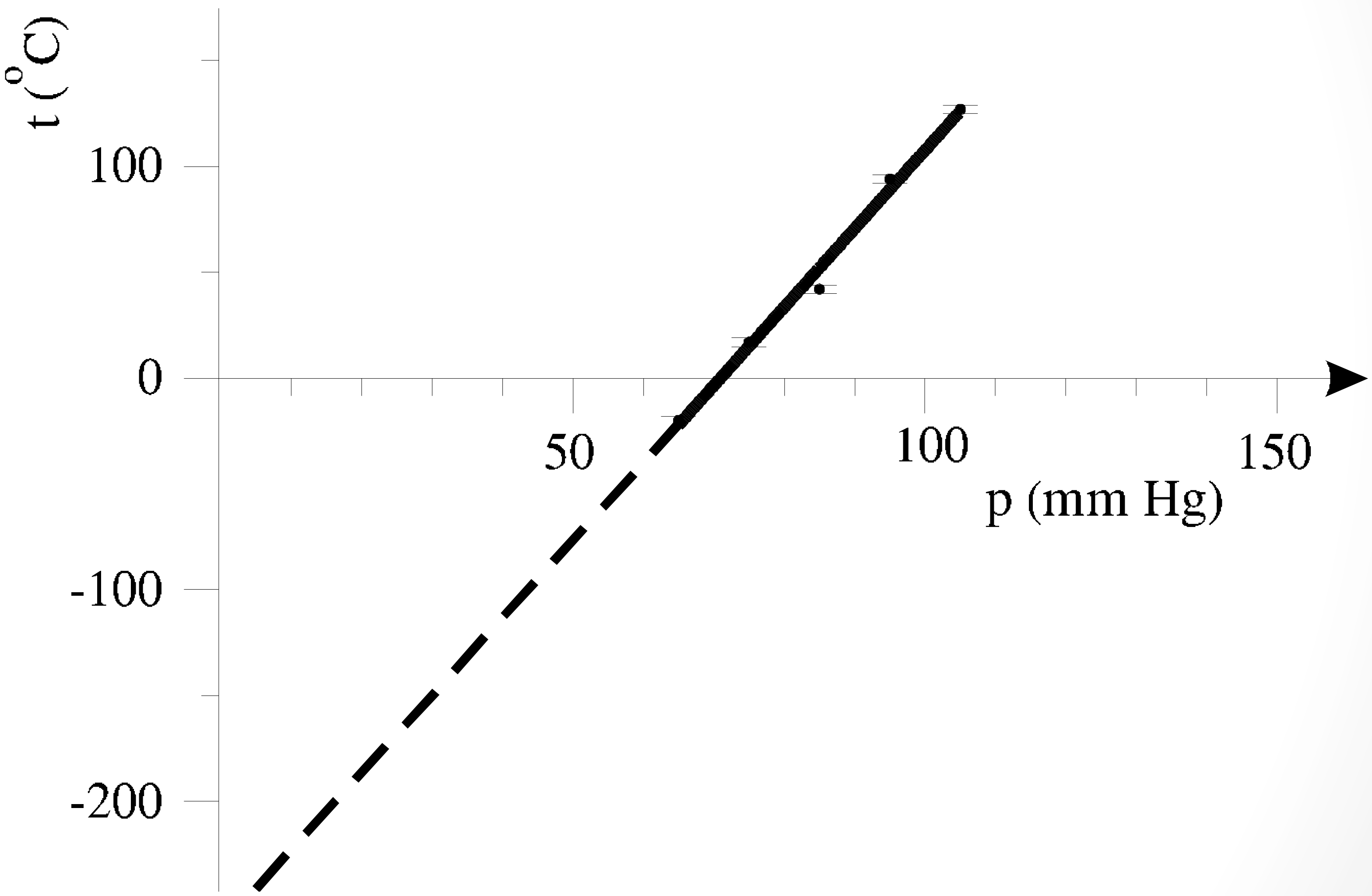
p (mm Hg)	65	75	85	95	105
t (^{o}C)	-20	17	42	94	127

- By applying the least squares method we get:

$$A=-263.35,\ \ B=3.71,\ \ \sigma_t=6.7\approx 7,\ \ \delta A=18,\ \ \delta B=0.22.$$

- So the graph of temperature against pressure is given in the next transparency:

The least-squares method for a straight line: An example



The least-squares method for a parabola

- When we have a sample of points and we wish to approximate it with a parabola of the form:

$$y = a + bx + cx^2$$

- With the constants *a*, *b,* and *c* satisfying the equations:

$$\sum_{i=}^{N} y_i = Na + b\sum_{i=}^{N} x_i + c\sum_{i=}^{N} x_i^2$$

$$\sum_{i=}^{N} x_i y_i = a\sum_{i=}^{N} x_i + b\sum_{i=}^{N} x_i^2 + c\sum_{i=}^{N} x_i^3$$

$$\sum_{i=}^{N} x_i^2 y_i = a\sum_{i=}^{N} x_i^2 + b\sum_{i=}^{N} x_i^3 + c\sum_{i=}^{N} x_i^4$$

The least-squares method for two independent variables

- Let's consider the case where there exists a **linear** relation between the dependent variable *z* and the two dependent variables *x* and *y* of the form

$$z = a + bx + cy$$

with the constants *a*, *b,* and *c*

$$\sum_{i=}^{N} z_i = Na + b\sum_{i=}^{N} x_i + c\sum_{i=}^{N} y_i$$

$$\sum_{i=}^{N} x_i z_i = a\sum_{i=}^{N} x_i + b\sum_{i=}^{N} x_i^2 + c\sum_{i=}^{N} x_i\, y_i$$

$$\sum_{i=}^{N} y_i\, z_i = a\sum_{i=}^{N} y_i + b\sum_{i=}^{N} x_i\, y_i + c\sum_{i=}^{N} x_i^2$$