

**ANALYSIS OF EXPERIMENTAL DATA**

Our main concern when working in the lab is whether or not we have obtained the "right" answer. You should be aware, however, that it is not as straightforward as it might seem. Various research groups might have obtained different values for a physical constant. And although these values will differ from each other, they are all to be weighted equally in the absence of any other information. For that reason, **accepted** value is a better term. From now on we will take the comparison between our experimentally determined value and the accepted value as a measure of accuracy. When doing experiments, one must keep watch for two main types of error:

**i. Random Errors;** have a sign associated with them and that sign indicates whether the experimental value is high (+) or low (-) with respect to the accepted value. For example: A musket ball is known to have a mass of 10.00 grams is weighted using a triple beam balance and the average of several trials is found to be 10.01 grams. The accuracy of the measurement is

$$10.01 \text{ g} - 10.00 \text{ g} = +0.01 \text{ g}$$

This quantity is known as an **absolute** error, since the units of the error are the same as the units of the measurement. We might want to compute the **relative** error as well

$$[+0.01 \text{ g} / 10.00 \text{ g}] \times 100 = +0.10 \%$$

Notice this quantity is unitless.

**ii. Systematic Errors;** these deviations can be divided into **determinate** and **indeterminate** errors. Determinate errors are errors which have a definite value that can, in principle, be measured and accounted for. Determinate errors, also called systematic errors, the most common sources are are instrumental, operator, and method errors.

**Determinate errors** are often unidirectional, that is they are all positive or all negative with respect to the accepted value. For example, another student weights the musket ball the next day using the same triple beam balance and the average of several trials is found to be 20.01 grams. The instrument is giving results which are too high. Let's take a look at the instrument...



The failure of the experiment is the fault of a former operator. Someone who used the balance at an earlier time left a Pikachu key-chain hanged on the calibrating screw of the balance. This illustrates a combination of instrumental and operator error; the instrument's calibration is off but the current operator is responsible for checking that there are no foreign objects affecting the measurements.

**Indeterminate Errors**, also called random or statistical errors fluctuate randomly and do not have a definite value. For example, consider the titration data obtained by a student after doing five different measurements using a standard buret...



trial 1: 25.79 mL

trial 2: 25.77 mL

trial 3: 25.79 mL

trial 4: 25.77 mL

trial 5: 25.78 mL

The first three digits are the same in all cases; the last digit has an uncertainty associated with it. This uncertainty is the result of the buret's readability limit. Notice that the buret has ten subdivisions between each mL marking, so the buret can be accurately read to the nearest tenth of a mL and the next digit is uncertain. If we were to cut the last figure off; then the volume would be known only to the nearest tenth of a mL. If we keep that last digit, then the volume is known to the nearest hundredth of a mL; with some uncertainty indicative of the buret's readability limit.

Experimental data is made more meaningful by the inclusion of uncertainties. The value of the uncertainty gives one an idea of the precision inherent in a measurement of an experimental quantity. There are many ways to quantify uncertainty and the method used will depend upon how many measurements are made and on how crucial the reporting of the value of uncertainty is with regard to the interpretation of the experimental data.

### REPORTING EXPERIMENTAL DATA

The students of the physical chemistry lab (this means you) are to report the value of a measured quantity  $x$  as the **mean value**

$$\langle x \rangle = [1/N][\sum_{i=1}^N x_i]$$

where  $x_i$  is the  $i^{\text{th}}$  measurement of  $x$ .

There will be an uncertainty associated with the reported value. The most widely used method for calculating uncertainty is representing it by the standard deviation of a series of measurements. This can be easily accomplished if it is assumed that a gaussian distribution function best describes the spread in the error values for a series of measurements made on a single observable.

Here plotted for the case that  $\mu = 0$  and  $\sigma = 1$  (broad peak) and  $\mu = 0$  and  $\sigma = 0.5$  (narrow peak). The probability function  $P(x)$  is normalized,

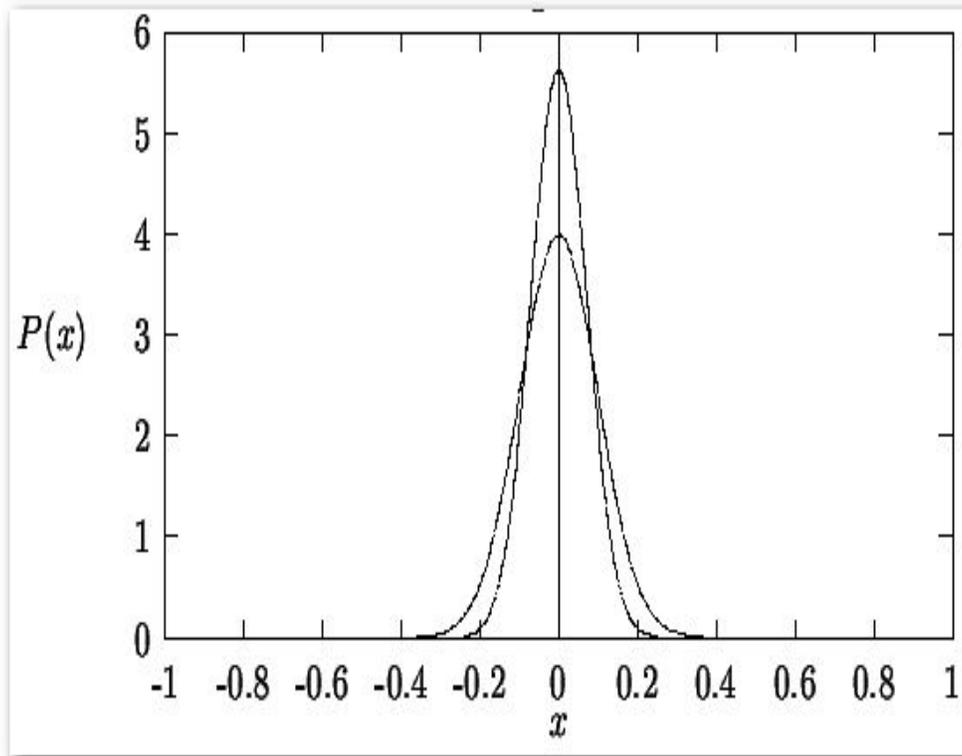
$$\int_{-\infty}^{\infty} dx P(x) = 1$$

implying that  $P(x)$  is a well-defined probability function. The average of  $x$  is given by

$$\langle x \rangle = \int_{-\infty}^{\infty} dx x P(x) = \mu$$

and that the point  $x = \mu$  locates the peak of  $P(x)$  and, by symmetry, is also the mean.

**The Gaussian Distribution,  $P(x) = [1/(\sqrt{2\pi\sigma^2})] \exp[-(x - \mu^2)/(2\sigma^2)]$**



The standard deviation  $\sigma$  can then be obtained from

$$\langle x \rangle^2 - \langle x \rangle^2 = \int_{-\infty}^{\infty} dx P(x) - \mu^2 = \sigma^2$$

Since  $P(x)$  represents the distribution of measured values about the mean  $\mu$ , the result of any measurement has a  $\sim 66\%$  chance of occurring in the range  $\mu \pm \sigma$ ;

$$\int_{\mu - \sigma}^{\mu + \sigma} dx P(x) \sim 0.66$$

and a  $\sim 95\%$  chance of occurring in the range  $\mu \pm 2\sigma$

$$\int_{\mu - 2\sigma}^{\mu + 2\sigma} dx P(x) \sim 0.95$$

The second expression is known as the law of two sigmas.

The standard deviation is related to the width of the distribution of these errors. If  $\sigma$  has a small value, the probability function is narrow indicating that the probability of large errors is small. If  $\sigma$  has a large value, the probability function is broad indicating a high degree of probable errors. The above equations imply that  $\sigma$  can only be known exactly for a number of measurements  $N = \infty$ . Since the human life span is not long enough to make an infinite number of measurements, we must rely on estimates of the true  $\sigma$ , the best estimates of this parameter are given by the square root of the variance ( $S^2 =$  the variance,  $S =$  square root of the variance)

$$S^2 = [1/N-1][\sum_{i=1}^N (x_i - \langle x \rangle)^2]$$

The variance can also be obtained from

$$S^2 = [1/N-1][(\sum x_i^2) - (\sum \langle x \rangle_i^2)]$$

Thus for multiple measurements  $N$ , the value that should be reported in the mean value and the uncertainty that should be reported is the 95% confidence level of the mean  $\lambda_{95}$  given by

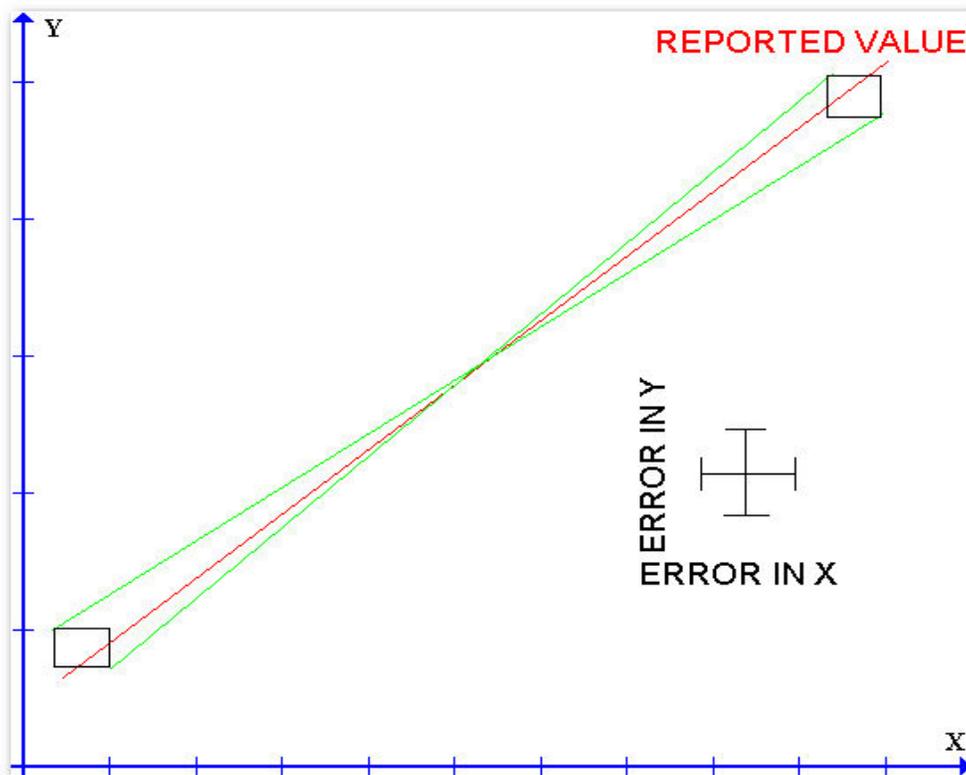
$$\lambda_{95} = [tS/\sqrt{N}]$$

where  $t$  is the student's  $t$ . For different numbers of independent measurements  $[N-1]$  the student's  $t$  has the following values...

N-1	90%	95%	97.5%	99.5%
1	3.07766	6.31371	12.7062	63.656
2	1.88562	2.91999	4.30265	9.92482
3	1.63774	2.35336	3.18243	5.84089
4	1.53321	2.13185	2.77644	4.60393
5	1.47588	2.01505	2.57058	4.03212
10	1.37218	1.81246	2.22814	3.16922
30	1.31042	1.69726	2.04227	2.74999
100	1.29007	1.66023	1.98397	2.62589
$\infty$	1.28156	1.64487	1.95999	2.57584

**GRAPHICAL ANALYSIS - METHOD OF LIMITING SLOPES**

The "method of limiting slopes" is the best technique for determining the uncertainty in the measurement of the slope and intercept of a linear plot. The following figure illustrates the method.



The technique consists of drawing a rectangle around every data point where the dimensions of the rectangle represent the uncertainties of the measured quantities. A best straight line is

then drawn throughout the individual points. Two other lines which represent the maximum or minimum slopes are then drawn so that both lines pass through every rectangle. The differences in the slopes or intercepts of the limiting lines can be taken as twice the uncertainty in the slope or intercept of the best line.

### PROPAGATION OF ERRORS

Suppose you are determining the density of a gold block, the density is a function of Mass and Volume

$$\rho = f(M, L, H, W)$$

Each of the measured quantities has an error  $[+/-\Delta M, +/-\Delta L, +/-\Delta H, +/-\Delta W]$  associated with it. These errors will be carried through in some way to the error in the density  $+/-\Delta\rho$ . For example, consider a small change in  $\rho$  from a small error in  $M$

$$\Delta\rho \sim [\partial\rho/\partial M] \times \Delta M$$

where  $\partial\rho/\partial M$  is the partial differential of  $\rho$  with respect to  $M$ , the total error is obtained from

$$\Delta\rho \sim |[\partial\rho/\partial M] \times \Delta M| + |[\partial\rho/\partial L] \times \Delta L| + |[\partial\rho/\partial H] \times \Delta H| + |[\partial\rho/\partial W] \times \Delta W|$$

The error for a given quantity is:

### For Addition and Subtraction

- [1] Add the squares of the uncertainties
- [2] Take the square root of the sum

Example 1, add the following lengths,

$$w = (4.52 \pm 0.02) \text{ cm}, x = (2.0 \pm 0.2) \text{ cm}, y = (3.0 \pm 0.6) \text{ cm}.$$

$$\text{Sum} = 4.52 \text{ cm} + 2.0 \text{ cm} + 3.0 \text{ cm} = 9.52 \text{ cm}$$

$$\begin{aligned} \text{Uncertainty} &= \sqrt{(0.02)^2 + (0.2)^2 + (0.6)^2} \\ &= \sqrt{0.0004 + 0.04 + 0.36} \\ &= 0.63 = 0.6 \end{aligned}$$

$$\text{Answer} = 9.5 \pm 0.6 \text{ cm}$$

Example 2

A particle rolls from position  $w = 9.3 \pm 0.2 \text{ m}$  to position as  $x = 14.4 \pm 0.3 \text{ m}$ . Compute the displacement and uncertainty.

$$\text{Displacement} = 14.4 \text{ m} - 9.3 \text{ m} = 5.1 \text{ m}$$

$$\begin{aligned} \text{Uncertainty} &= \sqrt{(0.2)^2 + (0.3)^2} \\ &= \sqrt{0.04 + 0.09} \\ &= 0.3605 = 0.4 \end{aligned}$$

$$\text{Answer} = 5.1 \pm 0.4 \text{ cm}$$

**For Multiplication and Division**

- [1] divide the uncertainties by the quantities
- [2] Square the division results
- [3] Add the squares
- [4] Take the square root of the sum
- [5] multiply the uncertainty by the quantity

Example 1, compute the area of a rectangular plane of dimensions,  
 $w = (4.52 \pm 0.02) \text{ cm}$ ,  $x = (2.0 \pm 0.2) \text{ cm}$ .  
 $\text{Area} = 4.52 \text{ cm} \times 2.0 \text{ cm} = 9.04 \text{ cm}^2$

$$\begin{aligned} \text{Uncertainty} &= \text{Area} \times \sqrt{\frac{(0.02/4.52)^2 + (0.2/2.0)^2}{}} \\ &= \text{Area} \times \sqrt{(1.96 \times 10^{-5}) + (0.01)} \\ &= 9.04 \times 0.1 = 0.9 \end{aligned}$$

Answer =  $9.0 \pm 0.9 \text{ cm}^2$

**Example 2**

A particle travels a distance  $w = 5.1 \pm 0.4 \text{ m}$  in  $0.4 \pm 0.1 \text{ sec}$ . Compute the velocity and uncertainty.  
 $\text{Velocity} = 5.1 / 0.4 = 12.75 \text{ m/s}$

$$\begin{aligned} \text{Uncertainty} &= \text{Velocity} \times \sqrt{\frac{(0.4/5.1)^2 + (0.1/0.4)^2}{}} \\ \text{Uncertainty} &= \text{Velocity} \times \sqrt{(0.0062) + (0.0625)} \\ &= 12.75 \times 0.262 = 3.34 \end{aligned}$$

Answer =  $12.75 \pm 3.34 \text{ m/s}$

**For Multiplication with a constant**

- [1] Compute the uncertainty with all quantities except the constant
- [2] multiply the uncertainty by the absolute value of the constant

**Example 1**

Certain reaction yields  $25.4 \pm 0.2 \text{ g}$  of carbon 12. Compute the number of mol and uncertainty.  
 $n = 25.4 \text{ g} \times 12 \text{ g/mol} = 304.8 \text{ mol}$   
 Note that 12 g/mol is an exact constant

$$\begin{aligned} \text{Uncertainty} &= 12 \times 0.2 \\ &= 2.4 \end{aligned}$$

Answer =  $304.8 \pm 2.4 \text{ mol}$

Example 2

A cylindrical container has a radius of  $5.6 \pm 0.1$  m and a height of  $7.4 \pm 0.1$  m. Compute the volume and uncertainty.

$$\text{Volume} = \pi(5.6)^2 \times 7.4 = 729 \text{ m}^3$$

$$\text{Uncertainty} = \pi \times \text{Vol} \times \sqrt{(0.2/5.6)^2 + (0.1/7.4)^2}$$

$$\text{Uncertainty} = \pi \times \text{Vol} \times \sqrt{(3.2 \times 10^{-4}) + (1.8 \times 10^{-4})}$$

$$\text{Uncertainty} = \pi \times 729 \times 0.022 = 51$$

$$\text{Answer} = 729 \pm 51 \text{ m}^3$$

The uncertainty in R for general functions of one or more observables. If R is a function of X and Y, written as  $R(X,Y)$ , then the uncertainty in R is obtained by taking the partial derivatives of R with respect to each variable, multiplication with the uncertainty in that variable, and addition of these individual terms in quadrature.

$$R = R(X, Y, \dots)$$

$$\delta R = \sqrt{\left(\frac{\partial R}{\partial X} \cdot \delta X\right)^2 + \left(\frac{\partial R}{\partial Y} \cdot \delta Y\right)^2 + \dots}$$

**Error Propagation, Main Points**

For **Addition and Subtraction** of quantities with uncertainties-

- 1) **Add absolute uncertainties** (in quadrature) to get a single uncertainty.
- 2) **Convert** absolute uncertainty **to fractional** (per cent) **uncertainty**

For **Multiplication and Division** of quantities with uncertainties

- 1) **Convert** absolute uncertainties **to fractional** (percent) **uncertainties**
- 2) **Add fractional uncertainties** (in quadrature) to get a single uncertainty.
- 3) **Calculate** (absolute) **uncertainty** of product.

For uncertainty in **a quantity raised to a power**

- 1) **Convert** absolute uncertainty to fractional (percent) uncertainty.
- 2) **Multiply** the **fractional uncertainty** by the **power** (e.g., two if quantity is squared).
- 3) **Calculate** (absolute) **uncertainty** of this quantity.