Appendix 1: Propagation of Errors

All experimental measurements are subject to random error. Propagation of errors determines the effect of the measurement errors on the final results. Errors are expressed in a variety of ways. For n measurements, x_i, the **standard deviation** is the root-mean-squared deviation of the measurements from the mean, $\overline{x}$:

$$s(x) = \left(\frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1} \right)^{1/2} \qquad \overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \text{A1.1}$$

The (n – 1) in the denominator is the number of **degrees of freedom**. The degrees of freedom are given by subtracting the number of extracted parameters from the number of measurements. In the determination of the standard deviation of a set of measurements, the single extracted parameter is the mean, $\overline{x}$. By convention, an error expressed in the form 12.34 ± 0.45 is assumed to be a standard deviation. The relative standard deviation is the standard deviation divided by the value. Multiplying the relative standard deviation by 100% gives the error as a percentage. The variance of x, $s(x)^2$, is the square of the standard deviation. The relative variance is the variance divided by the value squared. Using the example $x = 12.34 \pm 0.45$:

standard deviation:	$s(x) =$	0.45
relative standard deviation:	$\frac{s(x)}{x} =$	$\frac{0.45}{12.34} = 0.0365$ or 3.65%
variance:	$s(x)^2 =$	$(0.45)^2 = 0.202$
relative variance:	$\frac{s(x)^2}{x^2} =$	$\frac{(0.45)^2}{(12.34)^2} = 1.33 \times 10^{-3}$

Average deviations and approximations based on the readability of the instrument are also routinely used in propagation of errors. We will use the symbols δx for the generalized error and $\delta^2 x$ for the generalized variance, whether roughly approximated or carefully calculated from an average or standard deviation. How do measurement errors propagate to the final results?

Consider a single measurement, x, and a final result f(x). Over a narrow range, the function f(x) can be approximated as a linear function with slope (df/dx). An error in x gives a resulting error in f(x), approximated by linear extrapolation, Figure A1.1:

$$\delta f = \frac{df}{dx} \delta x \qquad \text{(fixed-nonrandom error)} \qquad \text{A1.2}$$

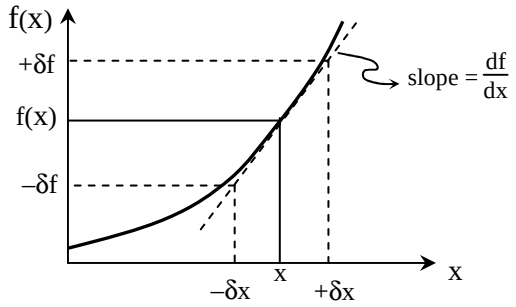

Figure A1.1: The error of the x value is extrapolated linearly to find the error of f(x).

The errors in x are assumed to be distributed randomly. In finding a representation of the average errors, the positive and negative errors cancel upon summation. To avoid cancellation of errors, Eq. A1.2 is recast in terms of the variance, for the same reason that the standard deviation is also based on a sum of squared errors. Each measurement adds a similar term to the error. The general formula for propagation of errors for a two-parameter function f(x,y) is:

$$\delta^2 f = \left(\frac{\partial f}{\partial x}\right)_y^2 \delta^2 x + \left(\frac{\partial f}{\partial y}\right)_x^2 \delta^2 y \qquad \text{A1.3}$$

Additional terms are added for additional measurements. This result is called the **master error relationship**. For example, consider the difference of two numbers, f = x – y. Eq. A1.3 gives:

$$f = x - y: \quad \left(\frac{\partial f}{\partial x}\right)_y = 1 \qquad \left(\frac{\partial f}{\partial y}\right)_x = -1 \qquad \delta^2 f = (1)^2\,\delta^2 x + (-1)^2\,\delta^2 y = \delta^2 x + \delta^2 y \qquad \text{A1.4}$$

The result is identical if addition is chosen as the example. In words, variances add on addition and subtraction. Consider multiplication, f = xy:

$$f = xy: \quad \left(\frac{\partial f}{\partial x}\right)_y = y \qquad \left(\frac{\partial f}{\partial y}\right)_x = x \qquad \delta^2 f = (y)^2\,\delta^2 x + (x)^2\,\delta^2 y \qquad \text{A1.5}$$

Division of this last equation by $f^2 = x^2y^2$, on both sides, converts the errors to relative variances:

$$f = xy: \quad \frac{\delta^2 f}{f^2} = \frac{\delta^2 x}{x^2} + \frac{\delta^2 y}{y^2} \qquad \text{A1.6}$$

The result is identical if division is chosen as the example. In words, relative variances add on multiplication and division. As a final example, consider logarithmic relationships, f = ln x:

$$f = \ln x: \quad \left(\frac{\partial f}{\partial x}\right)_y = \frac{1}{x} \qquad \delta^2 f = \left(\frac{1}{x}\right)^2 \delta^2 x = \frac{\delta^2 x}{x^2} \qquad \text{A1.7}$$

In words, the variance of ln(x) is the relative variance in the argument. Since there is only one error term, taking the square root of the last equation gives: the standard deviation of ln(x) is the relative standard deviation of the argument, $\delta\ln(x) = \delta x/x$. The results of the master error relationship are summarized by six rules, Table A1.1. For multi-step calculations, the rules are applied sequentially.

Example A1.1: *Error propagation in logarithmic relationships*
The equilibrium constants for a reaction at two different temperatures are $K_1 = 0.03220 \pm 0.00072$ at 298.2 K and $K_2 = 0.4732 \pm 0.0064$ at 353.2 K. Calculate the uncertainty in $\ln(K_2/K_1)$.

Answer: The relative variance in K_2/K_1 is the sum of the relative variances (Rule 2):

$$\begin{aligned} &\text{Relative variance in } K_2 = (0.00072/0.03220)^2 = 5.0\times10^{-4} \\ +\ &\underline{\text{Relative variance in } K_1 = (0.0064/0.4732)^2 \quad = 1.8\times10^{-4}} \\ &\text{Relative variance in } K_2/K_1 \qquad\qquad\qquad\quad = 6.8\times10^{-4} \end{aligned}$$

Then using Rule 3 shows that the variance in $\ln(K_2/K_1)$ is the relative variance in K_2/K_1:

$$\text{Variance in } \ln(K_2/K_1) = \text{Relative variance in } K_2/K_1 = 6.8\times10^{-4}$$

Standard deviation in result $= \sqrt{6.3 \times 10^{-4}} = 0.026$

$\ln(K_2/K_1) = \ln 14.689 = 2.687 \pm 0.026$

Table A1.1: Propagation of Error Rules, expressed with standard deviations.

Rule 1: Variances add on addition or subtraction:	$f = x + y$	$s(f)^2 = s(x)^2 + s(y)^2$
Rule 2: Relative variances add on multiplication or division:	$f = xy$	$\frac{s(f)^2}{f^2} = \frac{s(x)^2}{x^2} + \frac{s(y)^2}{y^2}$
Rule 3: The variance of ln(x) is the relative variance in x:	$f = \ln x$	$s(f)^2 = \frac{s(x)^2}{x^2}$
The variance of log(x) is the (relative variance in x)/$(2.303)^2$:	$f = \log x$	$s(f)^2 = \frac{s(x)^2}{(2.303)^2 x^2}$
Rule 4: The relative variance of e^x is the variance in x:	$f = e^x$	$\frac{s(f)^2}{f^2} = s(x)^2$
The relative variance of 10^x is the (variance in x)$(2.303)^2$:	$f = 10^x$	$\frac{s(f)^2}{f^2} = (2.303)^2 s(x)^2$
Rule 5: The variance of an average of n numbers, assuming equal variances for the x values, $s(x_i)^2 = s(x)^2$:	$f = \frac{\Sigma x_i}{n}$	$s(f)^2 = \frac{s(x)^2}{n}$

Rule 6: In calculations with only one error term, you can work with standard deviations instead of variance.

For homework problems, complete propagation of errors treatments are usually not required. However, an expression of the uncertainty in results is still necessary. For homework problems, significant figures rules are sufficient. Significant figure rules are based on the master error relationship, but implemented as an approximate short-cut to full error propagation. Significant figure rules are discussed in General Chemistry texts in detail:

SF Rule 1: In multiplication and division the number of significant figures in the result is the same as the smallest number of significant figures in the data.

SF Rule 2: In addition and subtraction the number of decimal places in the result is the same as the smallest number of decimal places in the data.

SF Rule 3: The number of significant figures in the mantissa of log x is the same as the number of significant figures in x. Use the same rule for ln x. (In log $4.23 \times 10^{-3} = -2.374$, the mantissa is the .374 part.)

SF Rule 4: The number of significant figures in 10^x is the number of significant figures in the mantissa of x. Use the same rule for e^x.

For example, the uncertainty in e^x with x = 3.45 using significant figure rules and in e^x with $x = 3.452 \pm 0.014$ using error propagation are:

Significant figure rules $e^{3.45} = 32.$ (2 SF, 2 SF) Propagation $e^{3.452 \pm 0.014} = 31.56 \pm 0.44$ (or 1.4% error)

Significant figure rules are a poor substitute for careful error propagation.

Appendix 2: Least Square Curve Fitting

Important experimental parameters are often determined from the slope and intercept of linear plots. Curve fitting also provides a statistically valid method for the determination of the precision of results determined from multiple measurements. Consider data plotted in linear form from a given set of n observation pairs, x_i and y_i, Figure A2.1. The x_i values are assumed to be exact, while the y_i values are subject to random error. The object is to find the best fit slope, m, and intercept, b. The calculated fit values, on the best fit line, are $\hat{y}_i = m\,x_i + b$. The residual is the difference between the observed and the fit values:

$$r_i = y_i - \hat{y}_i = y_i - m\,x_i - b \qquad \text{A2.1}$$

x_i	y_i
1.20	1.48
2.34	4.04
3.05	6.23
3.85	6.81
5.06	10.12

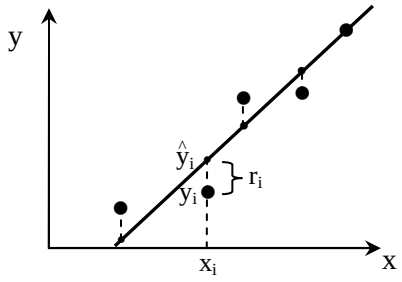

Figure A2.1: The best fit line minimizes the sum of squared residuals.

The best slope and intercept minimize the residuals of the fit values. To avoid the cancellation of positive and negative deviations of the fit values, the sum of squared residuals is used as the criterion for the best fit:

$$S = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} (y_i - mx_i - b)^2 \qquad \text{A2.2}$$

To minimize S the derivatives with respect to the slope and intercept are set equal to zero:

$$\left(\frac{\partial S}{\partial m}\right)_b = 0 = 2\,\Sigma\,(y_i - mx_i - b)(-x_i) \qquad \text{A2.3}$$

$$\left(\frac{\partial S}{\partial b}\right)_m = 0 = 2\,\Sigma\,(y_i - mx_i - b)(-1)) \qquad \text{A2.4}$$

Both equations have a common factor of -2. The summations factor through each term, $\Sigma(a + b) = \Sigma a + \Sigma b$. The final term in Eq. A2.4 gives the sum:

$$-b\sum_{i=1}^{n} 1 = -bn \qquad \text{A2.5}$$

The best fit m and b are the solutions of the resulting simultaneous equations:

$$0 = \Sigma\, y_i x_i - m\,\Sigma\, x_i^2 - b\,\Sigma\, x_i \qquad \text{A2.6}$$

$$0 = \Sigma\, y_i - m\,\Sigma\, x_i - bn \qquad \text{A2.7}$$

The term proportional to the slope can be eliminated by multiplying Eq. A2.6 by $(-\Sigma\, x_i)$ and multiplying Eq. A2.7 by $(\Sigma\, x_i^2)$:

$$0 = (\Sigma\, y_i x_i - m\,\Sigma\, x_i^2 - b\,\Sigma\, x_i)(-\Sigma\, x_i) \qquad \text{A2.8}$$

$$0 = (\Sigma\, y_i - m\,\Sigma\, x_i - bn)(\Sigma\, x_i^2) \qquad \text{A2.9}$$

Adding the resulting equations gives:

$$0 = -\Sigma\, x_i\,\Sigma\, y_i x_i + b\,(\Sigma\, x_i)^2 + \Sigma\, x_i^2\,\Sigma\, y_i - bn\Sigma\, x_i^2 \qquad \text{A2.10}$$

The last equation is solved for the best-fit b. The resulting b is then substituted into Eq. A2.9, which gives the best fit slope, m:

$$b = \frac{\Sigma\, x_i^2\,\Sigma\, y_i - \Sigma\, x_i\,\Sigma\, x_i y_i}{n\Sigma\, x_i^2 - (\Sigma\, x_i)^2} \qquad m = \frac{n\Sigma\, x_i y_i - \Sigma\, x_i\,\Sigma\, y_i}{n\Sigma\, x_i^2 - (\Sigma\, x_i)^2} \qquad \text{A2.11}$$

These equations are universally used by hand-held calculators, Excel, and computer software packages for curve fitting.

How well do we know m and b? The scatter of the data points from the best fit line is characterized by the standard deviation of the y values:

$$s_y = \left(\frac{\Sigma\, r_i^2}{n-2}\right)^{1/2} \qquad \text{(linear)} \qquad \text{A2.12}$$

The n – 2 degrees of freedom in the denominator results because two fit parameters are extracted from the data, the slope and the intercept (compare to Eq. A1.1). The next step is to determine the standard deviations of the slope and intercept. Propagation of errors from the measurement errors in the y_i uses the master error relationship with n error terms, one for each of the y_i values. For example, for the variance of the slope:

$$\delta^2 m = \left(\frac{\partial m}{\partial y_1}\right)^2 \delta^2 y_1 + \left(\frac{\partial m}{\partial y_2}\right)^2 \delta^2 y_2 + \left(\frac{\partial m}{\partial y_2}\right)^3 \delta^2 y_{3+} \ldots \qquad \text{A2.13}$$

We assume that the variance of each data point is the same, $\delta^2 y_i = s_y^2$. Each derivative is the same, giving the standard deviations of the slope and intercept as:

$$s_m = \left(\frac{n}{D}\right)^{1/2} s_y \qquad \text{with} \qquad D = n\Sigma\, x_i^2 - (\Sigma\, x_i)^2 \qquad \text{A2.14}$$

$$s_b = \left(\frac{\Sigma\, x_i^2}{D}\right)^{1/2} s_y \qquad \text{A2.15}$$

where D is the denominator of Eqs. A2.11. D is introduced purely as a convenience factor.

The regression coefficient, R, is used to judge if the data is well-described as a linear relationship:

$$R = \frac{n\Sigma\, x_i y_i - \Sigma\, x_i\,\Sigma\, y_i}{\sqrt{\{n\Sigma\, x_i^2 - (\Sigma\, x_i)^2\}\{n\Sigma\, y_i^2 - (\Sigma\, y_i)^2\}}} \qquad \text{A2.16}$$

R ranges from -1 for strong negative to +1 for strong positive correlation of the x and y values. However, the standard deviations of the slope and intercept are the best statistics for judging goodness-of-fit.

Non-linear Least Squares Curve Fitting: Least square analysis is also applied to non-linear relationships. Consider a non-linear expression with three fit parameters: a, b, and c. For example, $f(x) = a\, e^{-bx} + c$. Non-linear least squares is an iterative process. Guesses of the fit parameters must be made initially: a°, b°, and c°. Improvements of the fit parameters are obtained in a series of iterations. The expression for the improvements in the fit parameters is:

$$\underset{\approx}{A}\,\Delta\underset{\sim}{\alpha} = \underset{\sim}{h} \qquad \text{or}$$

$$\begin{pmatrix} \Sigma\left(\frac{\partial f}{\partial a}\right)^2 & \Sigma\left(\frac{\partial f}{\partial a}\right)\left(\frac{\partial f}{\partial b}\right) & \Sigma\left(\frac{\partial f}{\partial a}\right)\left(\frac{\partial f}{\partial c}\right) \\ \Sigma\left(\frac{\partial f}{\partial b}\right)\left(\frac{\partial f}{\partial a}\right) & \Sigma\left(\frac{\partial f}{\partial b}\right)^2 & \Sigma\left(\frac{\partial f}{\partial b}\right)\left(\frac{\partial f}{\partial c}\right) \\ \Sigma\left(\frac{\partial f}{\partial c}\right)\left(\frac{\partial f}{\partial a}\right) & \Sigma\left(\frac{\partial f}{\partial c}\right)\left(\frac{\partial f}{\partial b}\right) & \Sigma\left(\frac{\partial f}{\partial c}\right)^2 \end{pmatrix} \begin{pmatrix} \Delta a \\ \Delta b \\ \Delta c \end{pmatrix} = \begin{pmatrix} \Sigma\left(\frac{\partial f}{\partial a}\right)(y_i - \hat{y}_i) \\ \Sigma\left(\frac{\partial f}{\partial b}\right)(y_i - \hat{y}_i) \\ \Sigma\left(\frac{\partial f}{\partial c}\right)(y_i - \hat{y}_i) \end{pmatrix} \qquad \text{A2.17}$$

where $\Delta a = a - a°$, $\Delta b = b - b°$, $\Delta c = c - c°$ are the differences between the improved values and the old guesses. The fit values, $\hat{y}_i$, are the points on the fit curve evaluated with the old guess parameters. This matrix equation is easily solved for Δa, Δb, and Δc using matrix inversion, Eq. 2.8.19, as $\Delta\underset{\sim}{\alpha} = \underset{\approx}{A}^{-1}\,\underset{\sim}{h}$. Then improved values are calculated, for example $a = a° + \Delta a$. For the $f(x) = a\, e^{-bx} + c$ example, the derivatives are:

$$\left(\frac{\partial f}{\partial a}\right)_{b,c} = e^{-bx} \qquad \left(\frac{\partial f}{\partial b}\right)_{a,c} = -ax\, e^{-bx} \qquad \left(\frac{\partial f}{\partial c}\right)_{a,b} = 1 \qquad \text{A2.18}$$

which are evaluated at each point x_i in the sums. The sums are over all data points, i = 1 to n.

The power of the technique is that the uncertainties of the fit parameters are generated in a statistically valid way. The variances in the fit parameters are given by the diagonal elements of the inverse, $\underset{\approx}{A}^{-1}$, and the variance of the y values, s_y (Eq. A2.12 with (n – 3) in the denominator):

$$s_a^2 = [\underset{\approx}{A}^{-1}]_{11}\, s_y^2 \qquad s_b^2 = [\underset{\approx}{A}^{-1}]_{22}\, s_y^2 \qquad s_c^2 = [\underset{\approx}{A}^{-1}]_{33}\, s_y^2 \qquad \text{A2.19}$$

Correlations between fit parameters are also determined using the off-diagonal elements of $\underset{\approx}{A}^{-1}$:

$$R_{ab} = \frac{[\underset{\approx}{A}^{-1}]_{12}}{([\underset{\approx}{A}^{-1}]_{11}\,[\underset{\approx}{A}^{-1}]_{22})^{½}} \qquad R_{ac} = \frac{[\underset{\approx}{A}^{-1}]_{13}}{([\underset{\approx}{A}^{-1}]_{11}\,[\underset{\approx}{A}^{-1}]_{33})^{½}} \qquad R_{bc} = \frac{[\underset{\approx}{A}^{-1}]_{23}}{([\underset{\approx}{A}^{-1}]_{22}\,[\underset{\approx}{A}^{-1}]_{33})^{½}} \qquad \text{A2.20}$$

The correlation coefficients between the fit parameters are important for validating the results.[1] For example for R_{bc}, a correlation coefficient between the fit parameters of 0 means that the error in b has no effect on the error in c. For a correlation coefficient of ±1, the errors in b and c are completely correlated. For high correlation coefficients, a small change in the value of one data point, caused by experimental error, gives a large change in the fit values of both b and c. Correlation coefficients larger than ~0.95 mean that neither parameter estimate is valid.

An applet for "Non-Linear Least Squares Curve Fitting," is available on the text Web site or companion CD, in a three- or four-parameter version; 22 common functional forms are supported. Chapters 3-6, 18, and 20 provide examples of the use of non-linear curve fitting.

Literature Cited:

1. W. R. Draper, H. Smith, *Applied Regression Analysis, 2nd Ed.*, Wiley, New York, NY, 1981, Chapt 10, pp. 487-8.