Graphing

READING: ***Lab Handbook*** **- Section XIII: Representing Data and Experimental Results Using Graphs** (pp. 97-100)

Consider the following guidelines when you are graphing your data:

- **Title**
- **Size:** Graph should be large enough to usefully interpret the data. (Generally, half of a lab notebook page will suffice.)
- **Neatness:** Use a straightedge to draw axes and lines.
- **Axes:**
  - Label with the property represented, the scale, and the units.
  - Put tick marks on the gridlines, at regular and convenient intervals.
  - Use a range appropriate to your data. *Zero* need NOT be on the graph.
  - Put the quantity you control (independent variable) on the x-axis, and the quantity studied on the y-axis (dependent variable).
- **Best fit line or curves:**
  - Do NOT connect-the-dots!
  - Look for a smooth curve or a straight line.
  - ***For points that suggest a linear relationship***, use a straightedge to draw the best average line that fits the points. Some points may be above and some below this line. In some cases, the origin will be a point to consider, but you need NOT force the line to go through the origin.
  - ***For points that suggest a curve***, draw a smooth curve that hits many or most of the points. Crests and troughs in a curve should be approximately symmetrical.
  - If an individual data point does not seem to follow the trend of the rest of the data, (it does not seem to fall on the line or curve) then you may choose to exclude that point in drawing your best-fit trendline.
  - To prevent "crossing-out" on graphs in your notebook, sketch the graph in pencil with the board between the original and the duplicate, then go over it again with a pen and the board properly inserted. Or, make your graphs using Excel or other data analysis software and paste a copy on both the original and duplicate page in your notebook.

XIII

esenting Data and Experime Results Using Graphs

Graphs

Experimental data are often organized and recorded in a table that pairs the experimental conditions, the **independent variable,** with the corresponding observations and measured data, the **dependent variable** (see Figure 36). Frequently a graph helps to establish fundamental relationships between the data. Graphs are also used to predict or estimate data that are difficult to determine experimentally.

A **graph** is a pictorial diagram of data. The simplest and most widely used type for plotting scientific data is the **line graph,** an example of which is shown in

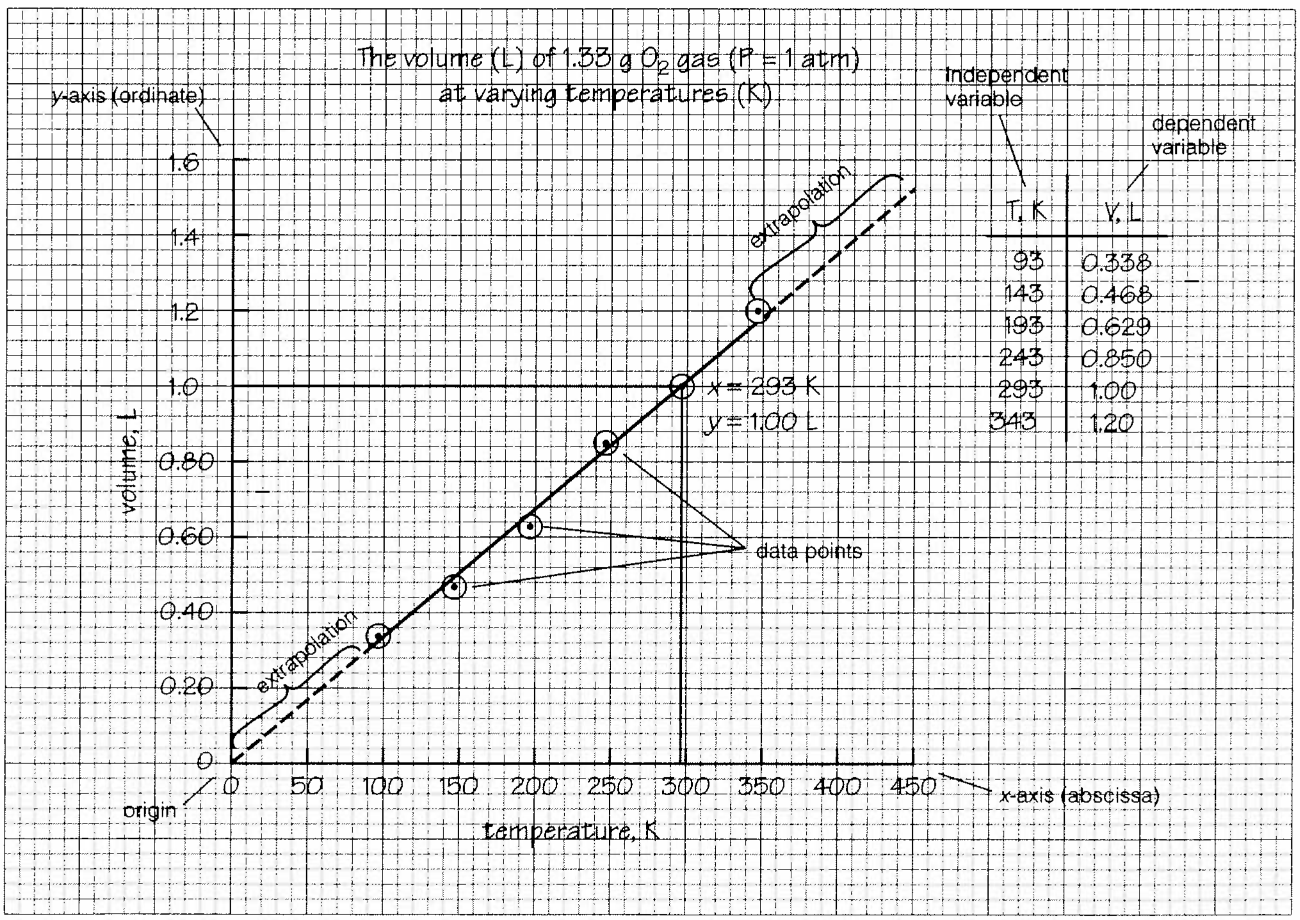

Figure 36 Data table and graph showing the volume (in liters) of 1.33 g O_2 gas (P = 1 atm) at varying temperatures (K)

Figure 36. Line graphs contain a vertical line, called the ***y*-axis,** or **ordinate,** and a horizontal line, called the ***x*-axis,** or **abscissa.** Each axis is divided into a numerical scale of units. Usually, the scale values increase, reading upward on the *y*-axis and to the right on the *x*-axis. Numbers on the *y*-axis are usually values of the dependent variable (volume, in Figure 36). Numbers on the *x*-axis normally represent values of the independent variable (temperature, in Figure 36). If the axes intersect at the point where the values represented on both axes equal 0, this point is called the **origin.** If the extrapolated line drawn through the data points does not pass through the origin, but instead crosses the *y*-axis at some other point, the intersection point is known as the ***y*-intercept.**

Each **data point** plotted on a graph represents the intersection of an *x*-axis value and the corresponding *y*-axis value from the data table for that graph. Data points are used to draw a smooth curve or straight line that best represents all of the data points. Note that the data points are ***not*** connected by a jagged line drawn directly from point to point. Rather, the **best straight line** is drawn, one that has approximately equal numbers of data points on either side of the line, with some data points on the line, as shown in Figure 36.

When preparing a graph, leave plenty of space for plotting data. Draw the axes using a straightedge. Allow space for labels along the left-hand side of the *y*-axis and under the *x*-axis. Construct all graphs so that the lines or curves representing the data points are as long as possible. To do so, select scales for the axes that will cause the graph to extend across a full page of graph paper. This will make the graph easier to interpret.

Based on the range of the data, choose appropriate values to begin and end each axis scale, and decide on the scale-unit size for each axis. Then, mark the axes with the scale units. Note that the units on the two axes do not have to be (and often are not) the same size. Generally, you should not number individual scale units due to space considerations. Rather, select a convenient multiple of the units and mark these multiples at the appropriate intervals. For example, Figure 36 shows *x*-axis unit labels that are multiples of 50.

Label each axis with the property being plotted and the corresponding unit, as shown in Figure 36. Then plot the data points, and draw the best straight line or smooth curve through those points. Title the graph appropriately for easy identification.

Equation and Slope of a Straight Line

The general equation for a straight line is

$$y = mx + b$$

where y is a value on the *y*-axis, x is the corresponding value on the *x*-axis, m is the slope of the line, and b is the point where the line intersects the *y*-axis. For example, a plot of Fahrenheit temperature (*y*-axis) versus Celsius temperature (*x*-axis) yields a straight line for which the slope, m, is 1.8 °F/°C, and the *y*-axis intercept, b, is 32 °F.

Frequently, the slope of a straight-line graph represents an important relationship between the experimental data. For example, the slope of the line for a plot of mass versus volume measurements (g/mL) of several samples of a substance is equal to the density of that substance. The slope of a straight line can be determined by dividing the change in the *y*-axis values of two points on the line by the change in the x-axis values of those two points. For example, the slope of the line in a plot of temperature versus volume (see Figure 36) can be calculated using two points on the line: $T_1 = 93$ K, $V_1 = 0.338$ L and $T_2 = 243$ K and $V_2 = 0.850$ L. The change in temperature (x-axis) is 243 K − 93 K = 150 K. The change in volume (y-axis) is 0.850 L − 0.338 L = 0.512 L. Therefore, the slope of the line can be calculated:

$$\text{slope} = \frac{0.850\text{ L} - 0.338\text{ L}}{243\text{ K} - 93\text{ K}} = \frac{0.512\text{ L}}{150\text{ K}} = 3.41 \times 10^{-3}\text{ L/K}$$

Notice that the units of the slope are the units of the y-axis divided by those of the x-axis. In this case, the units of the slope are L/K, which indicates the volume change of oxygen gas with a change in temperature. If the slope is positive ($m > 0$), as in this case, the line rises from left to right. If the slope is negative ($m < 0$), the line falls from left to right, as in a plot of mass versus time for an evaporating liquid.

Graphing Programs

Computerized graphing software automatically makes graphs that would be time consuming to create manually. Graphing programs can also investigate relationships between the independent and dependent variables, beyond just the plot of these variables. A graphing calculator or computer graphing program can plot a **linear regression line,** which is the best straight line passing through, or as close as possible to, all the data points on a graph. The calculator or computer program may also be able to determine the correlation between the data points and the linear regression line. The **correlation** is a value between 0 and 1 that indicates how well the linear regression fits the data points—that is, how close the data points are to forming a straight line. A high correlation value (close to 1) indicates a close fit between the data and the linear regression line; the plot has little **scatter** to the data. The lower the correlation value, the more scattered the data, with fewer points on or near the straight line.

Graphing Exercise

The following graphing exercise uses both manual graphing and Microsoft® Excel 2003 to plot mass versus time data for a sample of acetone, $(CH_3)_2CO$, a volatile liquid, in an open container. The data are given in the following table.

Time, min	Mass Acetone, g	Time, min	Mass Acetone, g
0.0	8.860	6.0	8.765
1.0	8.844	7.0	8.748
2.0	8.828	8.0	8.734
3.0	8.814	9.0	8.719
4.0	8.796	10.0	8.703
5.0	8.782		

Manual Graphing

1. Using these time and mass data, prepare an accurate ***hand-drawn*** graph of mass (y-axis) versus time (x-axis) on the graph paper provided at the end of this book.
2. Draw the best straight line through the data points, and calculate the slope of the line.

Microsoft® Excel 2003 Graphing*

Note: If Excel 2003 is not available, Excel 2000 can be used.

Three typestyles are used in the following directions: (1) **boldface** is used for system menus, categories within menus, dialog boxes, and cell references; (2) *italics* is used for menu choices and keyboard entries for menu choices; and (3) UPPERCASE indicates individual keystrokes, including OK and NEXT.

Data Entry Using the Time and Mass Data for Acetone

1. Open a new file in Excel and enter a descriptive title for your graph in cell **A1.** Press ENTER on your keyboard when you are finished.
2. Enter a label and unit (*time, minutes*) in cell **A3.** Do the same for (*mass, grams*) in cell **B3.** Don't worry if the labels and units extend beyond the cell boundaries; you can correct that later. Excel assumes that the values in the left-hand data column are for the independent variable and are to be plotted on the x-axis.
3. Enter the time data in cells **A4–A14** and the corresponding mass data in cells **B4–B14.** To display all values to three decimal points, click in cell **A4** and drag to cell **B14** to highlight all the time and mass entries. Now select **Format** and **Cells.** A **Number** tab will appear, and with it, a **Category.** Choose *Scientific* from the **Category** menu, then *3 Decimal Places,* and click OK.

Creating the Graph

4. Click and drag to highlight cells **A4** to **B14.**
5. Click **ChartWizard** on the toolbar, or select *Chart* from the **Insert** drop-down menu. From the **Chart Type** menu, choose **Chart type:** *XY (Scatter)*; from **Chart sub-type,** select *Smooth Lines without Markers*. Click NEXT.
6. The **Chart Source Data** menu shows a preview of the plot of mass versus time. Check to make sure the menu entries read **Data range:** = *Sheet1! $A$4:$B$14* and **Series in:** *Columns*. Make any necessary corrections, then click NEXT.
7. Click the **Titles** tab and enter a descriptive **Chart title.** Enter a label and units for the x-axis in the **Value (*X*) Axis** box, and a label and units for the y-axis in the **Value (*Y*) Axis** box. Click the **Gridlines** tab and deselect any gridlines that have been selected by ChartWizard. Click NEXT.

*This exercise is adapted from one found in MISC 877, *Introduction to Computer-Based Graphical Analysis,* by M. L. Gillette, H. A. Neidig, and J. R. Crook, © 2000 by Chemical Education Resources. A more complete treatment of graphing using Microsoft® Excel 97 is contained in this title.

8. Choose **As object in:** *Sheet 1* from the **Chart Location** menu. Click FINISH to display the graph.

9. The plot should appear roughly linear. To confirm its linearity, click on any data point using the right mouse button. This should highlight all data points, and a shortcut menu should appear. From this menu choose **Add Trendline.** From the **Add Trendline** menu, choose the **Type** tab, and select **Trend/ Regression type:** *Linear*. From the same menu select the **Options** tab and click *Display equation on chart* and *Display R-squared value on chart*. Select *Automatic* from the **Trendline Name** submenu. Click OK to finish.

The line now passing through the data points is the linear regression line. The equation for the linear regression line appears inside the chart area. The equation is in the form $y = mx + b$, where m is the slope of the line. The units of the slope are the y-axis unit (grams) divided by the x-axis unit (minutes). The value of R^2 is the correlation value.