Section 4
Student Guide to Graphical Analysis

Designing a controlled experiment

When scientists set up experiments they often attempt to determine how a given variable affects another variable. This requires the experiment to be designed in such a way that when the experimenter changes one variable, the effects of this change on a second variable can be measured. If any other variable that could affect the second variable were changed, the experimenter would have no way of knowing which variable was responsible for the results. For this reason, scientists always attempt to conduct **controlled experiments**. This is done by choosing only one variable to manipulate in an experiment, observing its effect on a second variable, and *holding all other variables in the experiment constant*.

Suppose you wanted to test how changing the mass of a pendulum affects the time it takes a pendulum to swing back and forth (also known as its period). You must keep all other variables constant. You must make sure the length of the pendulum string does not change. You must make sure that the distance that the pendulum is pulled back (also known as the amplitude) does not change. The length of the pendulum and the amplitude are variables that must be held constant in order to run a controlled experiment. The only thing that you would deliberately change would be the mass of the pendulum. This would then be considered the **independent variable**, because you will decide how much mass to put on the pendulum for each experimental trial. There are three possible outcomes to this experiment. 1. If the mass is increased, the period will increase. 2. If the mass is increased, the period will decrease. 3. If the mass is increased, the period will remain unchanged. Since you are testing the effect of changing the mass on the period, and since the period may depend on the value of the mass, the period is called the **dependent variable**.

In review, there are only two variables that are allowed to change in a well-designed experiment. The variable manipulated by the experimenter (mass in this example) is called the **independent variable.** The **dependent variable** (period in this case) is the one that responds to or depends on the variable that was manipulated. Any other variable that might affect the value of the dependent value must be held constant. We might call these variables **controlled variables.** When an experiment is conducted with one (and only one) independent variable and one (and only one) dependent variable while holding all other variables constant, it is a **controlled experiment**.

Recording Data

How can a scientist determine if two variables are related to one another? First she must collect the data from an experiment. Raw data is recorded in a data table immediately as it is collected in the lab. It is important to build a well-organized data table such as the example shown to the right. Use a ruler to lay it out neatly. You should always record your raw data in ink so that you will not accidentally erase a piece of data that you later decide you need. If you think that a given piece of data is in error, draw a single line through it and recollect the data point. Later, if you decide that the original point was really the correct one, you will still be able to read it.

You will notice that the experimenter has decided to run three trials for each mass value. Is this a good idea? Why? You should also notice that the experimenter has chosen to measure the time for 10 swings instead of the time for 1 swing. Why might she want to do this? Why not?

Pendulum Experiment--Period vs. Mass

Mass (g)	Time for 10 swings (s)
20.0	15.72 15.93 15.28
40.0	15.48 15.83 15.77
60.0	15.39 15.55 15.99
80.0	15.23 15.49 15.89
100.0	15.51 15.43 15.29

Values held constant in the experiment:
Length of string = 76.2 cm
Mass of pendulum = 65.3 g

The independent variable, mass, is given in the first column. Scientists have agreed to consistently place the independent variable in the leftmost column. Whenever something is done as an *agreed upon standard* it is called a **convention**. It is **conventional**, therefore, to place the independent variable in the leftmost column of the data table.

Each column is labeled with the name of the variable being measured and the units of measurement in parentheses below the variable name. Notice that each data entry in a given column is written to the same number of decimal places. The measuring device and technique used in the experiment determines the number of decimal places. In the mass column she recorded mass to the nearest 0.1 g because her balance was calibrated to the nearest 0.1 g. In the "time for 10 swings" column the time was reported to the nearest 0.01 s because the stopwatch gave times to that precision. A case could be made for only reporting the time to the nearest 0.1 s due to reaction time. It is important to exercise good judgment when recording data so as to honestly report how certain you are of your measurements.

It is a good idea to construct the data table before collecting the data. Too often, students will write down data in a disorderly fashion and <u>then</u> try to build their data table. This defeats the purpose of a data table that is to organize and make certain that data is clear and consistent.

Once the raw data has been collected for the experiment, you will proceed to prepare the data for graphing. Before you can graph the data in many experiments, you will need to manipulate the data so that it is ready to graph. For instance, this experiment was designed to test the effect of changing the mass of the pendulum on the period. The period is defined as the time for one swing. The experimenter, in an effort to reduce the error associated with timing the pendulum, decided to measure the time for 10 swings instead of for only one swing. Obviously, a calculation must be done before the period can be determined.

Also, the experimenter took three trials for each mass value. It is unnecessary to plot each of the times for each mass since there should be one representative time for each mass. Multiple trials are usually taken in an experiment where it is difficult to make a given measurement easily. Making multiple measurement (three to five) can help the experimenter determine whether or not the data is representative of the actual value or might be in error. For instance if you were to measure the time for a 40.0 g pendulum five times and the times you recorded were 15.86 s, 15.53 s, 15.47 s, 16.55 s, and 15.72 s you might have an idea that something was wrong with one of your data points. The wise experimenter would then carefully collect another data point for the 40.0 g pendulum. If its value were closer to the other four points than the 16.55 s measured, it would be used in place of the point that seemed out of place. If the value remained in line with the 16.55-s measured, you might need to perform still more trials for the 40.0 g pendulum to determine the appropriate time for its swings.

Usually, when multiple trials are collected for a data point, the trials are averaged to determine a representative value of that data point. This should be done only if the trials seem consistent enough to warrant an average. If you have one or more trials that are significantly different than your others, you need to look for an error in your technique or equipment setup that might be causing the problem. If a problem is found, the data should be recollected for any trials for which the error might have affected the results. Let's say that in the previous example, the retrial yields a time of 15.68 s. Since this is in the range of four out of the five trials, you might be justified in replacing the value of 16.55 with 15.68 s and then using an average of the five consistent trials to represent the data point.

When producing a formal table from which to produce your graph, some of the columns may be exactly the same as your raw data table but others will be the result of calculations made with your raw data. Any entry in your formal table that is the result of a calculation must include an explanation of the column and a sample calculation. For instance, in your formal table it is unnecessary to include each of the trials from your raw data. You can, if appropriate, include the average of your trials for a given data entry, which will reduce the total number of entries in your table. Since you are going to use the average of your three trials for the time for ten swings, you must explain how this column is calculated and show a sample calculation.

Sample Calculations:

*Average Time for 10 Swings:
This column, which is the average of the three trials for each data point, was calculated by adding the results from the three trials and dividing the sum by three.

$$\text{Avg. Time} = \frac{15.72\text{ s} + 15.93\text{ s} + 15.28\text{ s}}{3}$$

Avg. Time = 15.64333333 s

Avg. Time = 15.64 s

But since the stopwatch used was only capable of measuring to the nearest 0.01 s, the answer should be rounded to the nearest 0.01 s. Therefore the Avg. Time = 15.64 s

**Period:
The period of a pendulum is defined as the time for the pendulum to make one complete swing. This column was determined by dividing the Average Time for 10 swings by 10 swings, thus giving the average time for 1 swing which is the period.

$$\text{Period} = \frac{\text{Avg. Time for 10 Swings}}{\text{10 Swings}}$$

$$\text{Period} = \frac{15.64\text{ s}}{\text{10 Swings}}$$

Period = 1.564 s, but the stopwatch is only good to 0.01 s so

Period = 1.56 s

The data table to the right has three columns. The first column, mass, is the same as the raw data. In other words it is actually data. The second and third columns are the result of calculations using your data. As such, they are not actually data even though they are based on your data. These columns therefore require you to explain how they were created and to provide sample calculations.

Note that the second column, Average Time for 10 swings, has a * next to it. Likewise the third column, Period, has a ** next to it. These will be used to identify them as columns that are the result of calculations. Always include sample calculations and explanations for any column in your table which is the result of a calculation, no matter how simple.

Pendulum Experiment--Period vs. Mass

Mass (g)	Avg.Time for 10 Swings* (s)	Period** (s)
20.0	15.64	1.56
40.0	15.80	1.58
60.0	15.64	1.56
80.0	15.54	1.55
100.0	15.41	1.54

Values held constant in the experiment:
Length of string = 76.2 cm
Mass of pendulum = 65.3 g

Summary--Characteristics of Good Data Recording

1. Raw data is recorded in ink. Data that you think is "bad" is not destroyed. It is noted but kept in case it is needed for future use.
2. The table for raw data is constructed prior to beginning data collection.
3. The table is laid out neatly using a straightedge.
4. The independent variable is recorded in the leftmost column (by convention).
5. The data table is given a descriptive title that makes it clear, which experiment it represents.
6. Each column of the data table is labeled with the name of the variable it contains.
7. Below (or to the side of) each variable name is the name of the unit of measurement (or its symbol) in parentheses.
8. Data is recorded to an appropriate number of decimal places as determined by the precision of the measuring device or the measuring technique.
9. All columns in the table that are the result of a calculation are clearly explained and sample calculations are shown making it clear how each column in the table was determined.
10. The values held constant in the experiment are described and their values are recorded.

Graphing Data

Once the data is collected, it is necessary to determine the relationship between the two variables in the experiment. You will construct a graph (or sometimes a series of graphs) from your data table in order to determine thee relationship between the independent and dependent variables. An example follows which describes a graph produced from the formal data table created earlier.

For each relationship that is being investigated in your experiment, you should prepare the appropriate graph. In general your graphs in physics are of a type known as scatter graphs. The graphs will be used to give you a conceptual understanding of the relationship between the variables, and will usually also be used to help you formulate a mathematical statement which describes that relationship. Graphs should include each of the elements described below:

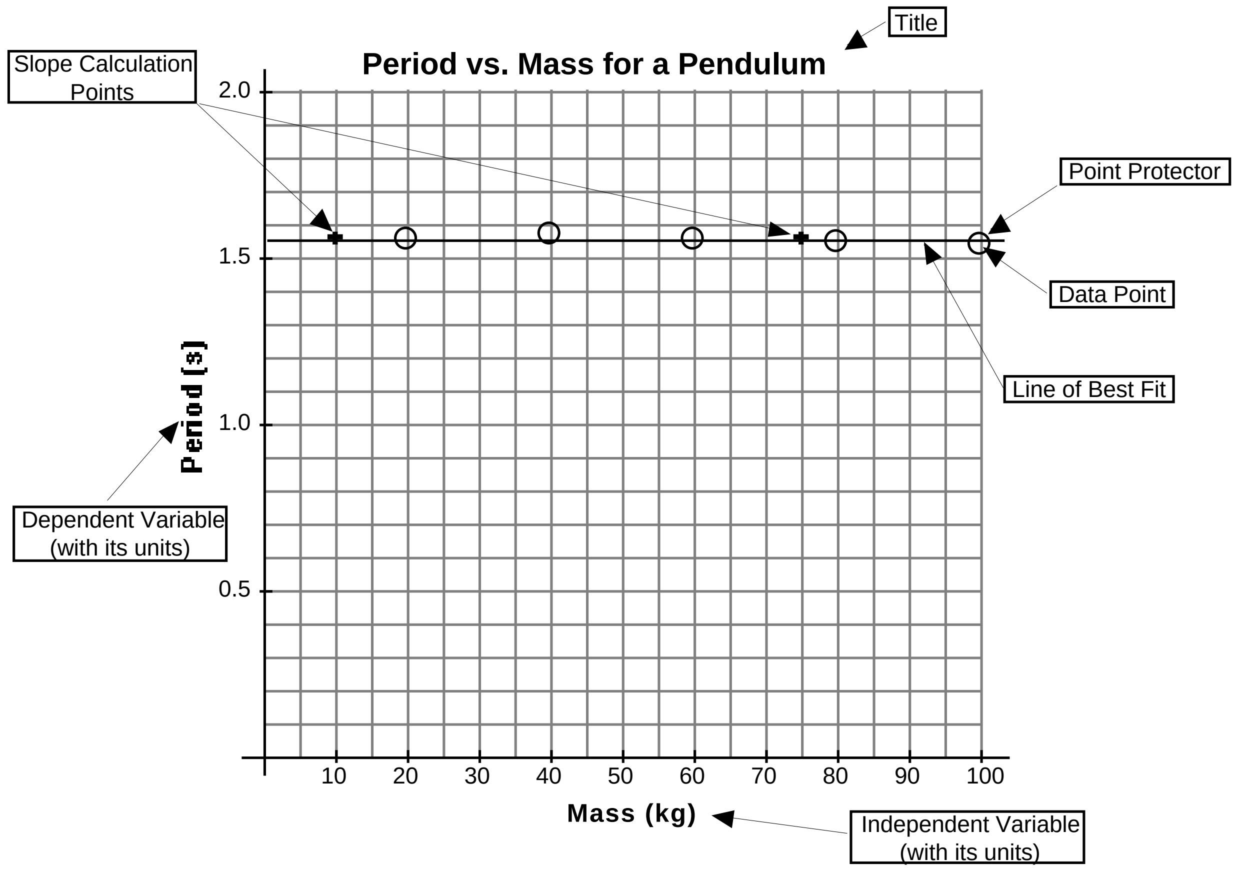
Elements of Good Graphs

- A **title** which describes the experiment. This title should be descriptive of the experiment and should indicate the relationship between the variables. It is conventional to title graphs with DEPENDENT VARIABLE vs. INDEPENDENT VARIABLE. For example, if the experiment was designed to show how changing the mass of a pendulum affects its period, the mass of the pendulum is the independent variable and the period is the dependent variable. A good title might therefore be PERIOD vs. MASS FOR A PENDULUM.
- The graph should **fill the space** allotted for the graph. If you have reserved a whole sheet of graph paper for the graph then it should be as large as the paper and proper scaling techniques permit.
- The graph must be properly **scaled.** The scale for each axis of the graph should always begin at zero. The scale chosen on the axis must be uniform and linear. This means that each square on a given axis <u>must</u> represent the same amount. Obviously each axis for a graph will be scaled independently from the other since they are representing different variables. A given axis must, however, be scaled consistently.
- Scale each axis so as to take up a maximum amount of the space available while still maintaining divisions that will make plotting the graph as easy as possible. Increments on an axis of 1, 2, or 5 are easy to use when plotting points. For larger numbers 10, 20 or 50 or possibly 100, 200, or 500 might work and so on. For smaller numbers 0.1, 0.2, or 0.5 might work or maybe 0.01, 0.02, or 0.05. A good way to choose the scale for an axis is to identify the largest data point that will be plotted on that axis. Then count the number of squares on your graph paper that are available for plotting the variable on that axis. Divide the maximum data value by the number of squares. This will give you smallest value that each square could possible have as its increment. Since the result of this division will most likely not be a convenient number, you should then round up to the nearest convenient value. Once you have chosen a scale, you do not have to label each square with its value. Label enough values on the axis to make it clear what scale you are using on the axis.

> **Scaling Example:** In a given experiment in which the length is being measured, the largest length that was measured was 455.2 m. The graph paper being used has 25 squares that could be used for the length axis. The minimum scaling increment for this graph paper would then be 455.2 m divided by 25 squares or 18.2 m per square. Since 18.2 is not a convenient number to use for the increment (it would be very difficult to plot such a graph) then you should round the value <u>up</u> to the next <u>higher</u> convenient number. This would be 20 m per square. Since each square is worth 20 and the largest value you need to plot on the axis is 455.2, you could label the axis every fifth square with 5 times 20 or 100.

- Each axis should be **labeled** with the **quantity** being measured and the **units** of measurement. The independent variable is plotted on the horizontal (or x) axis and the dependent variable is plotted on the vertical (or y) axis. The units on an axis must be consistent with the variables being plotted on that axis.
- Each data point should be plotted in the proper position. You should plot a point as a small dot at the position of the of the data point and you should circle the data point so that it will not be obscured by your line of best fit. These circles are called point protectors.
- A line of best fit. This line should show the overall tendency (or trend) of your data. If the trend is linear, you should draw a straight line that shows that trend using a straight edge. If the trend is a curve, you should sketch a curve that is your best guess as to the tendency of the data. This line (whether straight or curved) does not have to go through all of the data points and it may, in some cases, not go through any of them.
- Do not, under any circumstances, connect successive data points with a series of straight lines, dot to dot. This makes it difficult to see the overall trend of the data that you are trying to represent.
- If you are plotting the graph by hand, you will choose two points for all linear graphs from which to calculate the slope of the line of best fit. These points should not be data points unless a data point happens to fall perfectly on the line of best fit. Pick two points which are directly on your line of best fit and which are easy to read from the graph. Mark the points you have chosen with a +.
- Do not do other work in the space of your graph such as the slope calculation or other parts of the mathematical analysis.
- If your graph does not yield a straight line, you will be expected to manipulate one (or more) of the axes of your graph, replot the manipulated data, and continue doing this until a straight line results. In general it will probably not take more than three graphs to yield a straight line.

The graph below is an example of a properly prepared graph of the data from the Period vs. Amplitude for a pendulum experiment described earlier. Note that it contains each of the characteristics of good graphs:



Summary--Characteristics of Good Graphs

- It is plotted on a grid (graph paper)
- The axes are highlighted (darker than the rest of the grid lines) and are drawn with a ruler.
- Both axes are labelled with the variable name and its units. Note that we do not label them x or y!
- The independent variable is plotted on the horizontal (x) axis.
- The dependent variable is plotted on the vertical (y) axis.
- The data points have point protectors
- A line of best fit is drawn which shows the trend of the data. The line of best fit may have some points above it, some below it, and some on it. If the trend of the data is linear, the line of best fit is drawn with a ruler. If the trend of the data is curved, a smooth curve should be drawn.
- The graph is clearly titled using the convention dependent variable vs. independent variable.
- The axes are properly scaled so that the graph fits the space, the grids are consistently scaled, and all of the data fits on the graph.
- The slope calculation points are clearly marked with a (+) on the line of best fit.

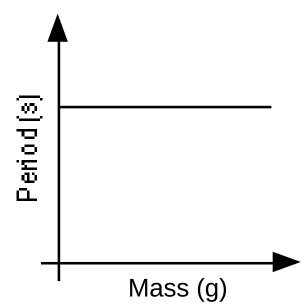
Graphical Analysis and Mathematical Models

Interpreting The Graph

The purpose of doing an experiment in science is to try to find out how nature behaves given certain constraints. In physics this often results in an attempt to try to find the relationship between two variables in a controlled experiment. Sometimes the trend of the data can be loosely determined by looking only at the raw data. The trend becomes more clear when one looks at the graph. In first year physics, most of the graphs we make will represent one of four basic relationships between the variables. These are 1) no relation 2) linear relationships 3) hyperbolic relationships and 4) parabolic relationships. There are many other types of mathematical relationships between variables. However, the vast majority of the experiments you do in an introductory physics course can be analyzed using some combination of the four above. By examining the graph, you should be able to determine whether there is no relation, a linear relation, or possibly a hyperbolic or square relation between the variables. To more specifically describe the relationship between the variables in an experiment, you will be expected to develop an equation. An equation that describes the behavior of a physical system (or any other system for that matter) can be called a **mathematical model**. The information which follows will describe each of the basic types of relationships we tend to see in physics. It also describes the process for arriving at a mathematical model to more fully (and simply) describe the relationship between the variables and the behaviors of physical systems.

No Relation

One possible outcome of an experiment is that changing the independent variable will have no effect on the dependent variable. When this happens we say that there is no relationship between the variables. The graph on the previous page is an example of an experiment that demonstrates the fact that there is no relationship between the mass of a pendulum and its period when we hold the length and amplitude constant. Since as the independent variable increases and the dependent variable stays the same, the resulting graph is a horizontal line. The slope of a horizontal line is always zero. Even though such a graph demonstrates no relationship between the variables, and equation of the line can still be determined. To the right is a sketch of a graph that shows no relationship between the period of a pendulum and its mass.



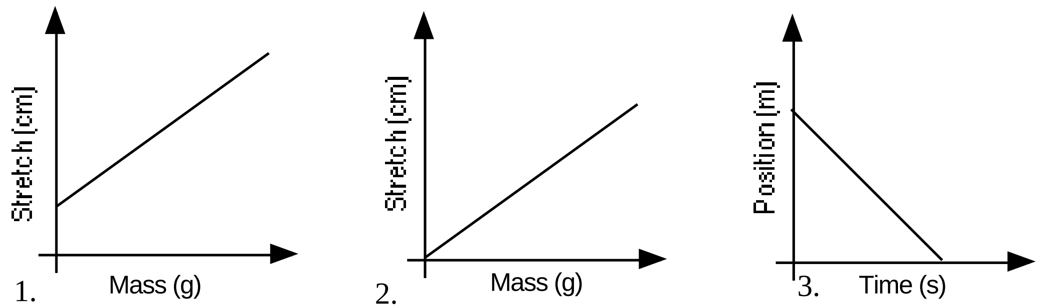
Below is a detailed example of the mathematical analysis of the graph for the Period vs. mass experiment described earlier in this document.

1.	Period → P Mass → m	(definition of variable symbols)
2.	$y = mx + b$	(slope-intercept form of line)
3.	$P = km + b$	(defined variables substituted)
4.	$k = \dfrac{\Delta P}{\Delta m}$	(definition of slope)
5.	$k = \dfrac{P_2 - P_1}{m_2 - m_1}$	
6.	$k = \dfrac{1.56s - 1.56s}{75.0kg - 10.0kg}$	(substitution of marked slope points from graph)
7.	$k = 0$	(result of slope calculation)
8.	$b = 1.56s$	(read from graph)
9.	$P = 0\ m + 1.56s$	(substitution into general equation)
10.	$P = 1.56s$	(final equation resulting from simplification)

Most of the steps in this example should be included in the mathematical analysis of a no relation experiment. Step 2 is really not necessary except as a guide to properly setting up the equation in step 3. Steps 5 and 6 could be omitted if it is clear that the trend of the graph is a horizontal line since the slope of <u>all</u> horizontal lines is zero. It would be acceptable, therefore, to go directly from step 4 to step 7. All other steps are essential and should be included. If you are using the computer to generate the graph and the computer calculates the slope for you, you could also conceivably omit steps 5 and 6 if you were to indicate that the computer calculated the slope for you.

Linear Relations

Another possible outcome of an experiment is that in which the graph forms a straight line with a non-zero slope. We call this type of relationship a linear relation. In a linear relation, equal changes in the independent variable result in corresponding constant changes of dependent variable.



Look at the examples above. You will notice that in each case, as the independent variable changes, the dependent variable also changes by equal amounts. Each one therefore qualifies as a linear relation. You can also see, however, that the relations have significant differences.

There is a significant difference between graphs 1 and 2. Obviously, graph 2 goes through the origin of the graph, while graphs 1 and 3 do not. Mathematically, the point at which a graph crosses the vertical axis is called a **y-intercept**. In the physical world, the y-intercept has some physical meaning. Specifically, it is the value of the dependent variable when the independent variable is zero. While all three graphs are linear relationships, only one of them illustrates a proportional relationship. A **direct proportion** occurs when, as one variable increases by a certain factor, the other variable increases by the same factor. Graphically, therefore, a direct proportion must not only be linear but <u>must</u> also go through the origin of the axes. When one variable is zero, the other variable must also be zero. When one variable doubles, the other variable doubles. When one variable triples, the other variable triples, and so on. Graph 2 is therefore the only example of a direct proportion while graphs 1 and 3 are simply linear relationships.

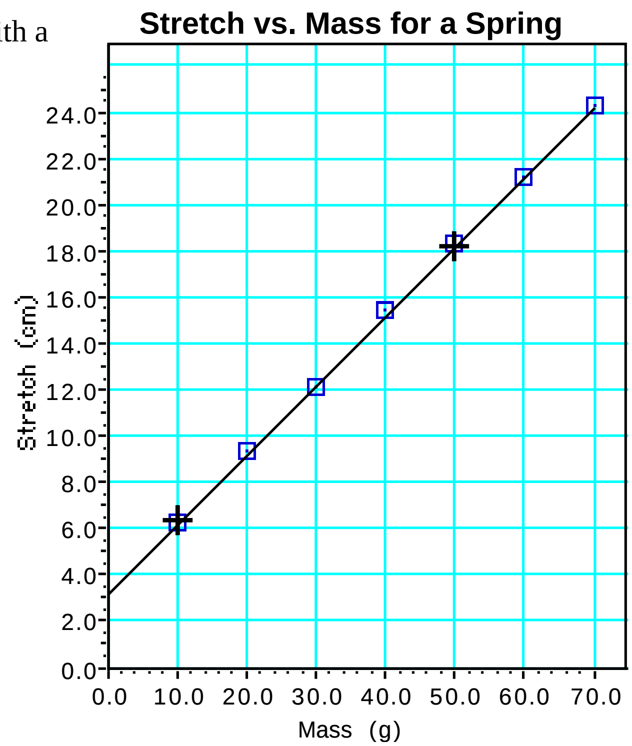
In graph 2, we can say that Stretch is directly proportional to Mass. In mathematical symbols this would be stated:

stretch   mass

Let us now look at a specific example that illustrates how to determine the specific mathematical relationship that is suggested by the linear relationship with a y-intercept.

Linear Relations with a y-intercept

When the data you collect yields a linear graph with a non-zero y-intercept, you will determine the mathematical equation that describes the relationship between the variables with the slope-intercept form of the equation of a line. Consider the following experiment in which the experimenter tests the effect of adding various masses to a spring on the amount that the spring stretches.



Data Table - Spring Experiment

Mass (g)	Stretch (cm)
10	6.2
20	9.3
30	12.1
40	15.4
50	18.3
60	21.2
70	24.3

A proper mathematical analysis of this graph would include the following:

1.	Stretch → S Mass → m	(definition of variable symbols)
2.	$y = mx + b$	(slope-intercept form of line)
3.	$S = km + b$	(defined variables substituted)
4.	$k = \dfrac{\Delta S}{\Delta m}$	(definition of slope)
5.	$k = \dfrac{S_2 - S_1}{m_2 - m_1}$	
6.	$k = \dfrac{18.0\text{ cm} - 6.0\text{ cm}}{50.0 g - 10.0\ g}$	(substitution of marked slope points from graph)
7.	$k = 0.30\frac{\text{cm}}{\text{g}}$	(result of slope calculation)
8.	$b = 3.2$ cm	(read from graph)
9.	$S = 0.30\frac{\text{cm}}{\text{g}}\ m + 3.2$ cm	(substitution into general equation)

The result of this experiment, then, is a mathematical equation which models the behavior of the spring:

Stretch = 0.30 cm/g · mass + 3.2 cm

With this **mathematical model** we know many characteristics of the spring and can predict its behavior without actually further testing the spring. In models of this type, there is physical significance associated with each value in the equation. For instance, the slope of this graph, 0.30 cm/g, tells us that the spring will stretch 0.30 centimeters for each gram of mass that is added to it. We might call this slope the "wimpiness" of the spring, since if the slope is high it means that the spring stretches a lot when a relatively small mass is placed on it and a low value for the slope means that it takes a lot of mass to get a little stretch.

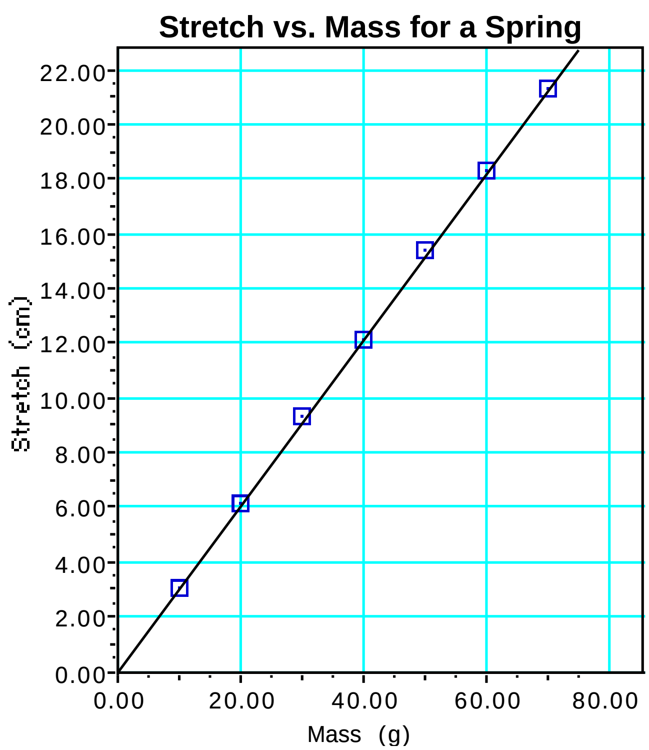
The y-intercept of 3.2 cm tells us that the spring was already stretched 3.2 cm when the experimenter started adding mass to the spring. With this mathematical model, we can determine the stretch of the spring for any value of mass by simply substituting the mass value into the equation. How far would the spring be stretched if 57.2 g of mass were added to the spring? Mathematical models are powerful tools in the study of science and we will use those that you develop experimentally as the basis of many of our studies in physics.

Direct Proportions

When the data you collect yields a line that passes through the origin of the axes (0,0) we call the relationship a direct proportion. This is because for such relationships, changes in one variable result in proportional changes in the other variable. Consider another spring experiment such as the one described earlier. Suppose in this case that the spring initially is unstretched and begins to stretch as soon as any mass is added to it. The data for such an experiment might look like the following:

Data Table - Spring Experiment

Mass (g)	Stretch (cm)
10	3.0
20	6.1
30	9.3
40	12.1
50	15.4
60	18.3
70	21.3



Statistics:	Slope	Y Intercept
Data Set 1	0.31±0.00	0.01±0.09

The mathematical analysis follows:

1.	Stretch → S Mass → m	(definition of variable symbols)
2.	$S \quad m$	(relationship stated as a proportion)
3.	$S = km$	(proportion rewritten as an equation)
4.	$k = \frac{\Delta S}{\Delta m}$	(definition of slope)
5.	$k = 0.30\ \frac{\text{cm}}{\text{g}}$	(slope calculated by computer)
6.	$S = 0.30\ \frac{\text{cm}}{\text{g}}\ m$	(substitution into general equation)

The mathematical model for this experiment states: Stretch = 0.30 cm/g · mass
What is the physical significance of this statement?

Notice that the mathematical analysis of a direct proportion begins with a statement that it is in fact a proportion. Any direct proportion of the form, y x, can be written as an equation of the form y = kx, where k represents the constant of proportionality between the variables. This is also the slope of the graph. The same equation could have been arrived at using the slope-intercept form of the equation of a line. Since the y-intercept term is essentially zero, b drops out of the equation y = kx + b, and the equation becomes y = kx.

When you are evaluating real data, you will need to decide whether or not the graph should go through the origin. Given the limitations of the experimental process, real data will rarely yield a line that goes perfectly through the origin. In the example above, the computer calculated a y-intercept of 0.01 cm ± 0.09 cm. Since the uncertainty (±0.09 cm) in determining the y-intercept exceeds the value of the y-intercept (0.01 cm) it is obviously reasonable to call the y-intercept zero. Other cases may not be so clear cut. The first rule of order when trying to determine whether or not a direct linear relationship is indeed a direct proportion is to ask yourself what would happen to the dependent variable if the independent variable were zero. In many cases you can reason from the physical situation being investigated whether or not the graph should logically go through the origin.

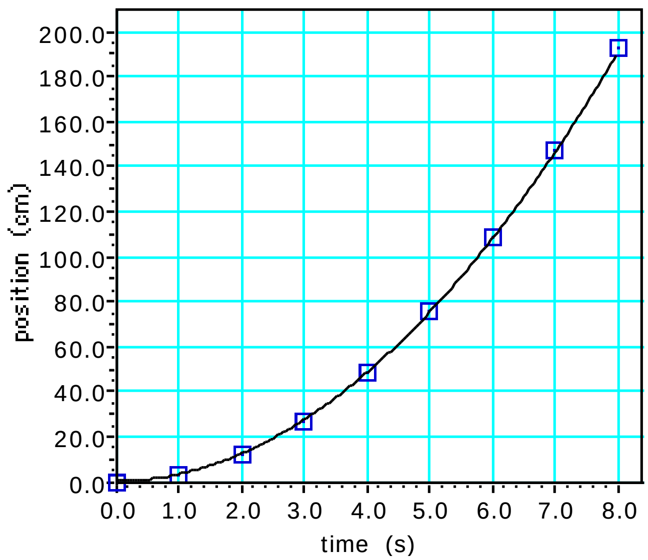
Square Relations—Top Opening Parabolas

When the data you collect is non-linear and looks like it might be parabolic, you will employ a powerful technique called **linearizing the graph** to determine whether or not the graph is a parabola with its vertex at the origin. Consider the data to the right and graph below for an experiment in which a student releases a marble from rest allowing it to go down an inclined track. The student notes its position every 1.0 seconds along the track.

Data Table - Rolling Ball Experiment

time (s)	position (cm)
0	0
1	3.1
2	12.2
3	27
4	47.9
5	75.2
6	108.3
7	146.8

Position vs. Time for a Rolling Ball



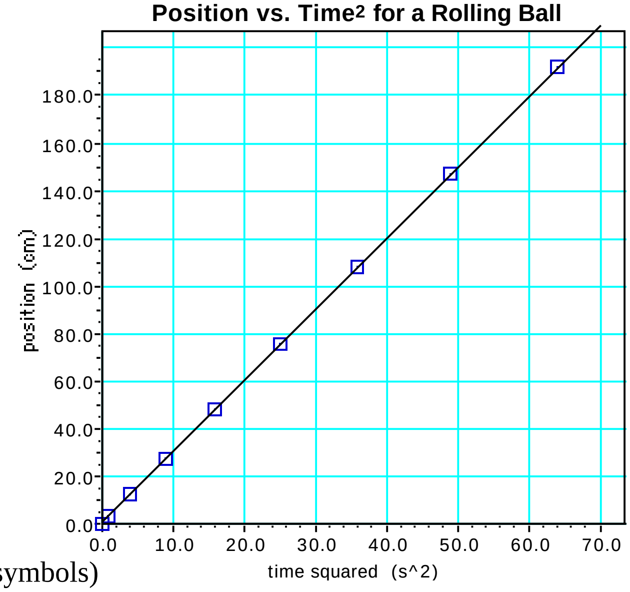
Since the graph is not linear for this experiment, you cannot determine the equation that fits the data using only the techniques shown for the previous graphs. The slope of this graph, for instance, is constantly changing. Notice, however, that the trend of the graph shows that as time increases in constant increments, position increases in greater and greater increments. This is why the graph bends away from the time axis. Since position increases at a greater rate than time, is there any mathematical manipulation of the data that could be performed which would allow us to plot a linear graph? Notice that the graph looks like it might be a parabola. Since the quadratic equations yield parabolas when plotted, and take the form $y = ax^2 + bx + c$, we might look to this form for a hint. First of all, if this is a parabola, it appears to have its vertex at the origin. When the vertex of a parabola is at the origin, the values of b and c are zero and the equation reduces to $y = ax^2$. If we think that this graph is parabolic with a (0, 0) vertex, we might try to manipulate the data based on the form $y = ax^2$. Think about it. If position is increasing at a greater rate than time, isn't it possible that squaring time, which will make it increase at a greater rate, might make it keep up with position? This reasoning is the basis of creating a **test plot**.

Basically, a test plot is a graph made with mathematically manipulated data for the purpose of testing whether or not our guess about a mathematical relation might hold true. Since we think that the graph above may represent a square relationship (parabola), we will make a new table in which we will square every time value. We square time because it is the variable that is not keeping up in this case. We will then plot a graph of position vs. time2 to determine whether or not our hunch was right. If we are correct, our test plot will yield a straight-line position vs. time^2. graph. We can then follow the previously detailed process for linear relations to determine the mathematical model that describes the relationship between position and time for our rolling marble.

Rolling Ball Experiment

time (s)	position (cm)	time^2 (s^2)
0	0	0
1	3.1	1
2	12.2	4
3	27	9
4	47.9	16
5	75.2	25
6	108.3	36
7	146.8	49
8	192.1	64

Position vs. Time^2 for a Rolling Ball



Statistics:	Slope	Y Intercept
Data Set 1	3.0±0.0	0.1±0.1

Note that squaring each of the time values and plotting a new graph of position vs. time^2 yields a linear graph that passes through the origin. This indicates that the position of the ball is proportional to the square of the time that it has rolled from rest. The mathematical analysis of such a graph is the same as for other linear relations.

The resulting equation is indeed that of a parabola.

The mathematical analysis for this experiment follows:

Position → x Time → t	(definition of variable symbols)
$x \quad t^2$	(relationship stated as a proportion)
$x = kt^2$	(proportion written as an equation)
$k = \frac{\Delta x}{\Delta t^2}$	(definition of slope)
$k = 3.0\ cm/s^2$	(slope calculated by computer)
$x = 3.0\ cm/s^2\ \ t^2$	(substitution into general equation)

Our mathematical model that describes the relation between position and time for a marble rolling down an incline is: **position = 3.0 cm/s^2 · time^2**

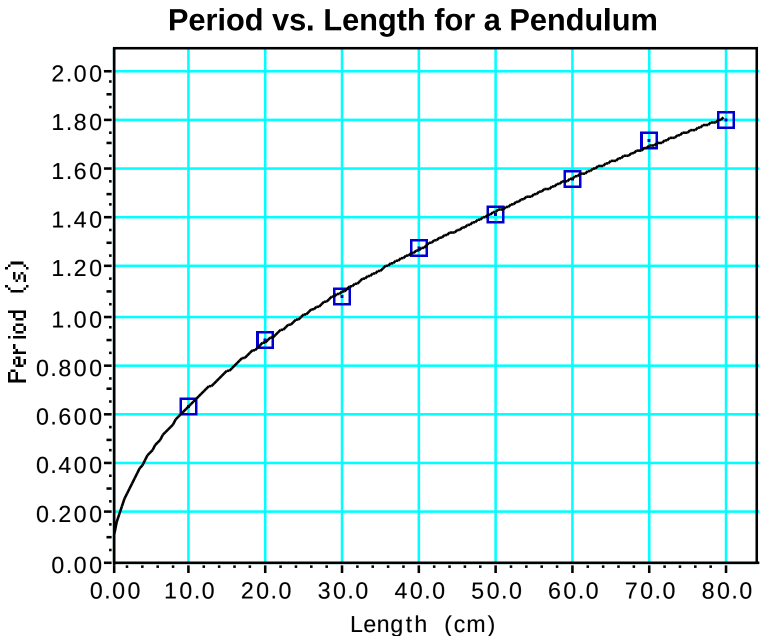
It is important to understand that this equation is not only the equation of the linear graph shown above, but *it is also the equation of the original parabolic graph*. This idea eludes many students. We have arrived at an equation which describes our original (curved) graph by manipulating the data (squaring time) to find the equation of a line.

Square Relations--Side Opening Parabolas

Consider an experiment in which a student investigates the effect of changing the length of a pendulum on the time required for the pendulum to make on swing (or the period).

Data Table - Pendulum Experiment

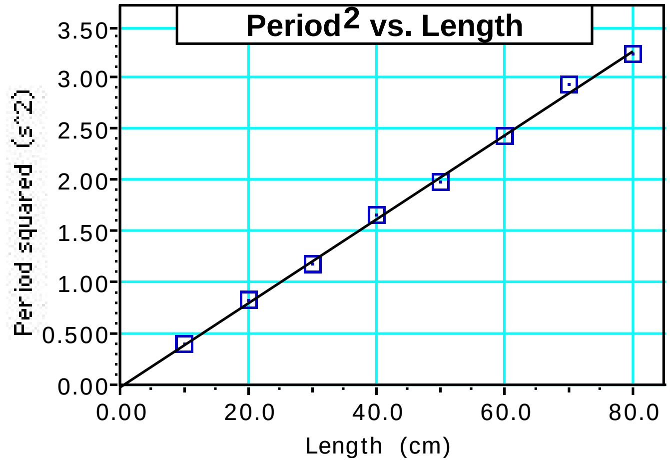
Length (cm)	Period (s)
10	0.63
20	0.90
30	1.08
40	1.28
50	1.41
60	1.56
70	1.71
80	1.80



Notice that this graph resembles graph 3 and that it could be a parabola which opens to the right. Once again, since the graph is not linear, you can't determine the equation directly. The trend of the graph shows that as the length of the pendulum increases in constant increments, the period increases in decreasing increments. This is why the graph bends away from the period axis. Since length increases at a greater rate than period, what mathematical manipulation of the data would you perform which might allow us to plot a linear graph? Since squaring the variable which was increasing at the lower rate worked in the previous example, why not try it again. This time the reasoning might be: If length is increasing at a greater rate than time, isn't it possible that squaring period, which will make it increase at a greater rate, might make it keep up with length? Based on our speculation that this might be a sideways opening parabola, we will make a new table in which we will square every period value. We square period because it is the variable which is not keeping up in this case. Our test plot will then be Period^2 vs. Length. How will we know if this the mathematical relationship we have predicted is correct? Of course! If the new graph is linear and passes through the origin, we can say that period^2 is proportional to length. If the graph is the straight line through the origin that we expected, we can t determine the mathematical model which describes the relationship between period and time for our pendulum.

Pendulum Experiment

Length (cm)	period (s)	period^2 (s^2)
10	0.63	0.40
20	0.90	0.81
30	1.08	1.17
40	1.28	1.64
50	1.41	1.99
60	1.56	2.43
70	1.71	2.92
80	1.80	3.24



Statistics:	Slope	Y Intercept
Data Set 1	0.0412±0.000626	-0.0301±0.0316

Since the resulting test plot is linear and passes through the origin, our prediction is confirmed. The original graph was indeed a sideways opening parabola. To determine the mathematical model that describes the relationship between the Period and Length of a pendulum, we will follow the standard procedure as outlined below:

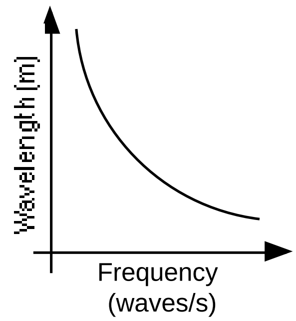
Period → T Length → L	(definition of variable symbols)
$T^2 \quad L$	(relationship stated as a proportion)
$T^2 = kL$	(proportion rewirtten as an equation)
$k = \dfrac{\Delta T^2}{\Delta L}$	(definition of slope)
$k = 0.041 \frac{s^2}{cm}$	(slope calculated by computer)
$T^2 = 0.041 \frac{s^2}{cm} L$	(substitution into general equation)

The resulting mathematical model is: **Period² = 0.041 s²/cm · Length**

Remember that this equation describes not only the linear test plot but is also the equation of the original parabolic graph.

Inverse Relations

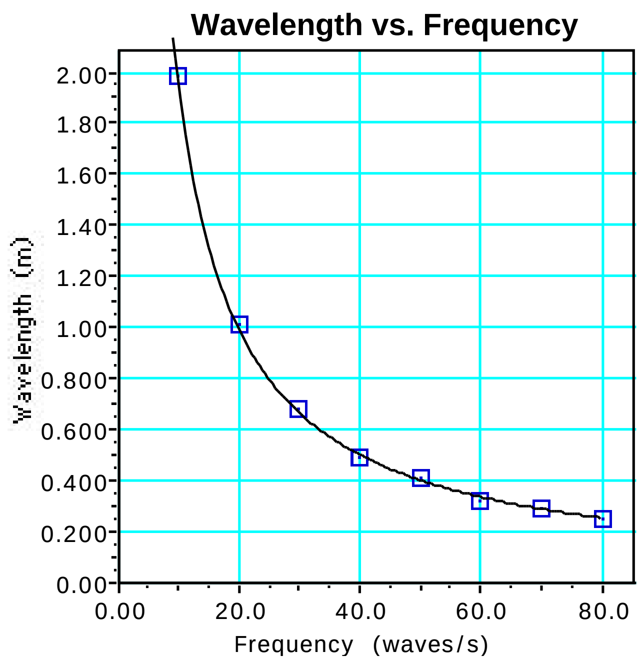
The final type of fundamental relationship that we will study is the inverse relation. An inverse relation is basically the opposite of a direct relation. In an inverse relation, as the independent variable increases the dependent variable decreases. This can take multiple forms, but the most common type that you will encounter in this physics course looks like the sketch to the right.



Consider an experiment where students make waves on a spring by shaking the spring at a certain frequency, and measuring the resulting wavelength of the waves.

Data Table - Wave Experiment

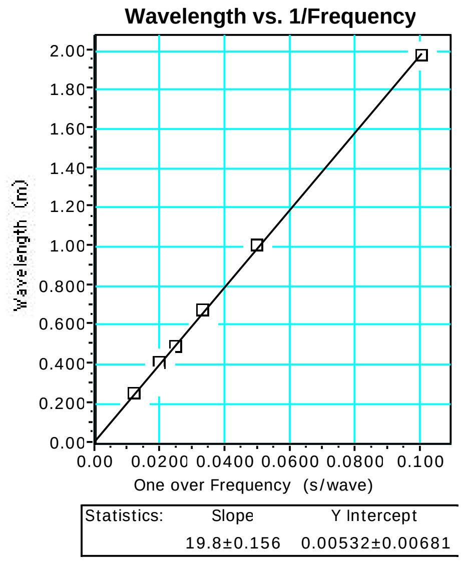
Wavelength (m)	Frequency (waves/s)
10	1.98
20	1.01
30	0.68
40	0.49
50	0.41
60	0.32
70	0.29
80	0.25



The data from this experiment indicate an inverse relation between Wavelength and Frequency. As Frequency increases, Wavelength decreases. How do we determine an equation for this obviously nonlinear graph? If you said, we need to make a test plot, you are headed in the right direction. But what mathematical manipulation of the data might possibly linearize the graph? How can we make both variables change in the same direction? The answer might be to take the reciprocal of one of the variables. This will cause the variable you have manipulated to decrease if it was increasing or to increase if it was decreasing. The convention that we will follow for our test plots of inverse relations is to take the reciprocal of the independent variable. We will therefore divide one by every frequency value to make a column of one over frequency. Remember that whatever manipulation you make of the physical quantity, you must also make of the units for that physical quantity. The reciprocal of waves/s is s/wave. You will plot a new graph of Wavelength vs. 1/Frequency.

Waves Experiment

Frequency (waves/s)	Wavelength (m)	1/Frequency (s/wave)
10	1.98	0.1000
20	1.01	0.0500
30	0.68	0.0333
40	0.49	0.0250
50	0.41	0.0200
60	0.32	0.0167
70	0.29	0.0143
80	0.25	0.0125



Wavelength → W (definition of variable symbols)
Frequency → f

$W \quad \frac{1}{f}$ (statement of relation as proportion)

$W = k \; \frac{1}{f}$ (relation rewritten as equation)

$k = \frac{\Delta W}{\Delta \frac{1}{f}}$ (definition of slope)

$k = 19.8 \; \frac{\frac{m}{wave}}{\frac{s}{wave}}$ (slope calculated by computer)

$k = 19.8 \frac{m}{s}$ (slope units simplified)

$W = 19.8 \frac{m}{s} \; \frac{1}{f}$ (substitution of slope into general equation)

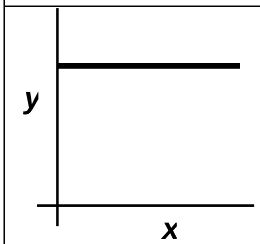
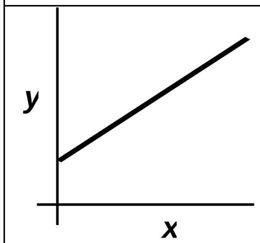
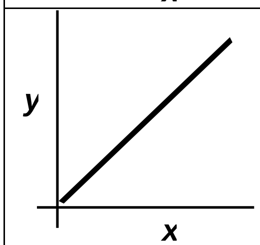
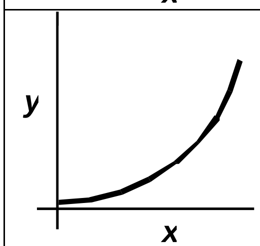
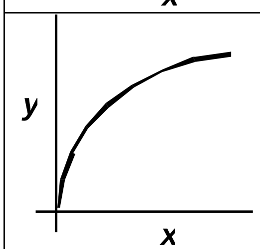
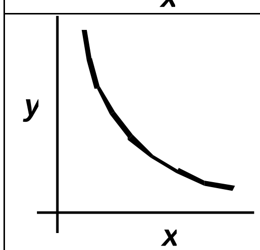
In this experiment, since Wavelength is directly proportional to 1/Frequency, we can also say that Wavelength is inversely proportional to Frequency. Such a relation is known as an inverse proportion.

The mathematical model that describes the relationship between Wavelength and Frequency for this experiment is: Wavelength = 19.8 m/s · 1/Frequency

While the relations described above do not describe every possible physical situation that might be encountered, it does serve as the basis of a great many of them and will cover nearly every situation you are likely to encounter in an introductory physics course. Sometimes your test plot will not yield a linear relation initially, but it might suggest yet another test plot. In most situations in this course, a maximum of three graphs will allow you to linearize the data and obtain a mathematical model.

Summary--Mathematical Models from Graphs

One of the most effective tools for the visual evaluation of data is a graph. The independent variable is conventionally placed on the x-axis. The dependent variable is placed on the y-axis. The ability to interpret what the graph means is an essential skill. You will be expected to learn to describe the relationship between the variables on a graph in two ways. One way will be to give a written statement (in words) of the general relationship between the two variables. The second is to develop an equation that will describe the relationship between these variables mathematically. We will call this equation a mathematical model of the physical relationship

Graph Shape	***Written Relationship***	***Modification Required to Linearize Graph***	***Mathematical Model***
	**No Relation:** As x increases, y remains the same. There is no relationship between the variables.	**No Modification Required**	**y = b**
	**Linear Relation:** Equal changes in x result in equal changes in y. y varies at a constant rate with respect to x.	**No Modification Required**	**y = kx + b**
	**Direct Proportion:** As x increases, y increases by the same factor. y is directly proportional to x.	**No Modification Required**	**y x** **y = kx**
	**Parabolic Relation:** As x increases, y increases at a greater rate. y is proportional to the square of x	**Graph** **y vs. x^2**	**y x^2** **y = kx^2**
	**Parabolic Relation:** As x increases, y increases at a lower rate. y^2 is proportional x	**Graph** **y^2 vs. x**	**y^2 x** **y^2 = kx**
	**Inverse Proportion** As x increases, y decreases y is inversely proportional to x	**Graph** **y vs. 1/x**	**y 1/x** **y = k(1/x)**