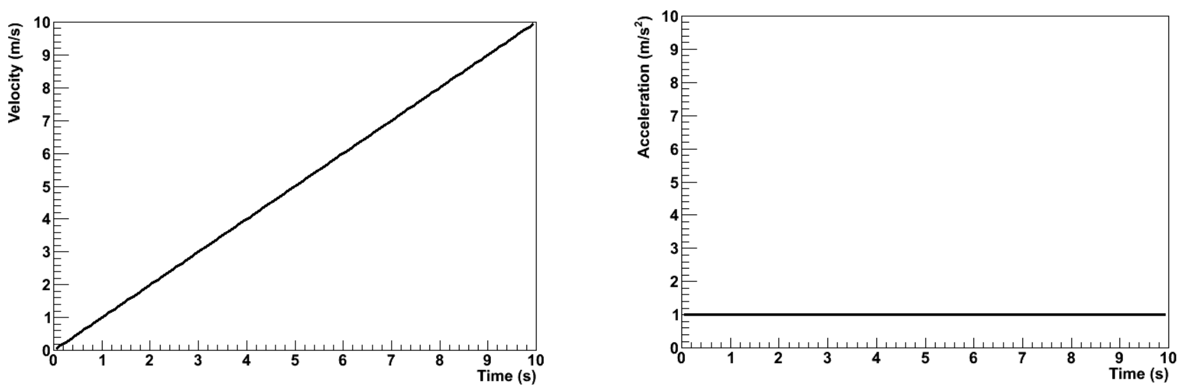
Appendix D: A Review of Graphs

Graphs are visual tools used to represent relationships (or the lack thereof) among numerical quantities in mathematics. In particular, we are interested in the graphs of functions.

What is a graph?

In this course, we will be dealing almost exclusively with graphs of functions. When we graph a quantity A with respect to a quantity B, we mean to put B on the horizontal axis and A on the vertical axis of a two-dimensional region and then to draw a set of points or curve showing the relationship between them. We do not mean to graph any other quantity from which A or B can be determined. For example, a plot of acceleration versus time has acceleration itself, $a(t)$, on the vertical axis, not the corresponding velocity $v(t)$; the time t, of course, goes on the horizontal axis. See Figure 1.



(a) (b)

Figure 1: Graphs of acceleration a and velocity v for an object in 1-dimensional motion with constant acceleration.

Traditionally, we call the vertical axis the " y -" axis; the horizontal axis, the " x -" axis. Please note that there is nothing special about these variables. They are not fixed, and they have no special meaning. If we are graphing, say, a velocity function $v(t)$ with respect to time t, then we do not bother trying to identify $v(t)$ with y or t with x; in that case, we just forget about y and x. This can be particularly important when representing position with the variable x, as we often do in physics. In that case, graphing $x(t)$ with respect to t would give us an x on both the vertical and horizontal axes, which would be extremely confusing. We can even imagine a scenario wherein we should graph a function x of a variable y such that y would be on the horizontal axis and $x(y)$ would be on the vertical axis. In particular, in MotionLab, the variable z, not x, is always used for the horizontal axis; it represents time. Both x and y are plotted on vertical axes as functions of the time z.

There are graphs which are not graphs of functions, e.g. pie graphs. These are not of relevance to this course, but much of what is contained in this document still applies.

Data, Uncertainties, and Fits

When we plot empirical data, it typically comes as a set of ordered pairs (x, y). Instead of plotting a curve, we just draw dots or some other kind of marker at each ordered pair.

Empirical data also typically comes with some uncertainty in the independent and dependent variables of each ordered pair. We need to show these uncertainties on our graph; this helps us to interpret the region of the plane in which the true value represented by a data point might lie. To do this, we attach error bars to our data points. Error bars are line segments passing through a point and representing some confidence interval about it.

After we have plotted data, we often need to try to describe that data with a functional relationship. We call this process “fitting a function to the data” or, more simply, “fitting the data.” There are long, involved statistical algorithms for finding the functions that best fit data, but we won’t go into them here. The basic idea is that we choose a functional form, vary the parameters to make it look like the experimental data, and then see how it turns out. If we can find a set of

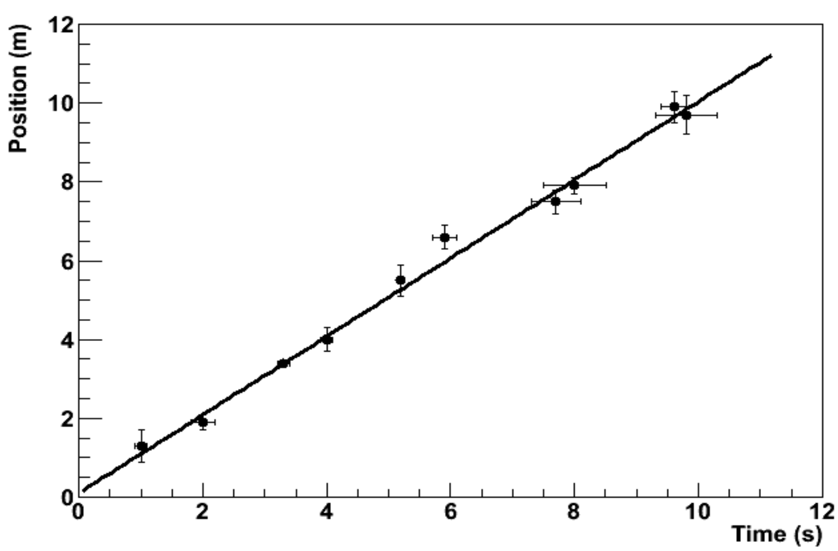


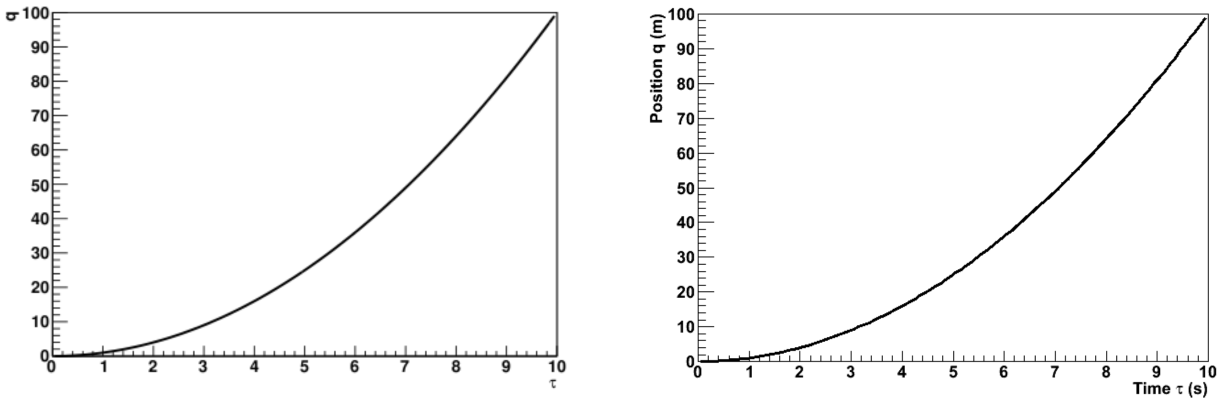
Figure 2: An empirical data set with associated uncertainties and a best-fit line.

parameters that make the function lie very close to most of the data, then we probably chose the right functional form. If not, then we go back and try again. In this class, we will be almost exclusively fitting lines because this is easiest kind of fit to perform by eye. Quite simply, we draw the line through the data points that best models the set of data points in question. The line is not a “line graph;” we do not just connect the dots (That would almost never be a line, anyway, but just a series of line segments.). The line does not actually need to pass through any of the data points. It usually has about half of the points above it and half of the points below it, but this is not a strict requirement. It should pass through the confidence intervals around most of the data points, but it does not need to pass through all of them, particularly if the number of data points is large. Many computer programs capable of producing graphs have built-in algorithms to find the best possible fits of lines and other functions to data sets; it is a good idea to learn how to use a high-quality one.

Making Graphs Say Something

So we now know what a graph is and how to plot it; great. Our graph still doesn’t say much; take the graph in Figure 4(a). What does it mean? Something called q apparently varies quadratically with something called τ, but that is only a mathematical statement, not a physical one. We still need to attach physical meaning to the mathematical relationship that the graph communicates. This is where labels come into play.

Graphs should always have labels on both the horizontal and vertical axes. The labels should be terse but sufficiently descriptive to be unambiguous. Let’s say that q is position and τ is time in Figure 4. If the problem is one-dimensional, then the label “Position” is probably sufficient for the vertical axis (q). If the problem is two-dimensional, then we probably need another qualifier. Let’s say that the object in question is moving in a plane and that q is the vertical component of its position; then “Vertical Position” will probably do the trick. There’s still a problem with our axis labels. Look more closely; where is the object at $\tau = 6\text{s}$? Who knows? We don’t know if the ticks represent seconds, minutes, centuries, femtoseconds, or even some nonlinear measure of time, like humans born. Even if we did, the vertical axis has no units, either. We need for the units of each axis to be clearly indicated if our graph is really to say something. We can tell from Figure 4(b) that the object is at $q = 36\text{m}$ at $\tau = 6\text{s}$. A grain of salt: our prediction graphs will not always need units. For example, if we are asked to draw a graph predicting the relationship of, say, the acceleration due to gravity of an object with respect to its mass, the label “Mass” will do just fine for our horizontal axis. This is because we are not expected to give the precise functional dependence in this situation, only the overall behavior. We don’t know exactly what the acceleration will be at a mass of 10g , and we don’t care. We just need to show whether the variation is increasing, decreasing, constant, linear, quadratic, etc. In this specific case, it might be to our advantage to include units on the vertical axis, though; we can probably predict a specific value of the acceleration, and that value will be meaningless without them.



(a)

(b) Position q with respect to time τ for a mass of 3kg. The acceleration is constant.

Figure 4: Poorly- versus well-labeled and -captioned graphs. The labels and caption make the second graph much easier to interpret.

Every graph we make should also have some sort of title or caption. This helps the reader quickly to interpret the meaning of the graph without having to wonder what it’s trying to say. It particularly helps in documents with lots of graphs. Typically, captions are more useful than just titles. If we have some commentary about a graph, then it is appropriate to put this in a caption, but not a title. Moreover, the first sentence in every caption should serve the same role as a title: to tell the reader what information the graph is trying to show. In fact, if we have an idea for the title of a graph, we can usually just put a period after it and let that be the first “sentence” in a caption. For this reason, it is typically redundant to include both a title and a caption. After the opening statement, the caption should add any information important to the interpretation of a graph that the graph itself does not communicate; this might be an approximation involved, an indication of the value of some quantity not depicted in the graph, the functional form of a fit line, a statement about the errors, etc. Lastly, it is also good explicitly to state any important conclusion that the graph is supposed to support but does not obviously demonstrate. For example, let’s look at Figure 4 again. If we are trying to demonstrate that the acceleration is constant, then we would not need to point this out for a graph of the object’s acceleration with respect to time. Since we did not do that, but apparently had some reason to plot position with respect to time instead, we wrote, “The acceleration is constant.”

Lastly, we should choose the ranges of our axes so that our meaning is clear.

Our axes do not always need to include the origin; this may just make the graph more difficult to interpret. Our data should typically occupy most of the graph to make it easier to interpret; see Figure 5. However, if we are trying to demonstrate a functional form, some extra space beyond any statistical error helps to prove our point; in Figure 5(c), the variation of the dependent with respect to the independent variable is obscured by the random variation of the data. We must be careful not to abuse the power that comes from freedom in

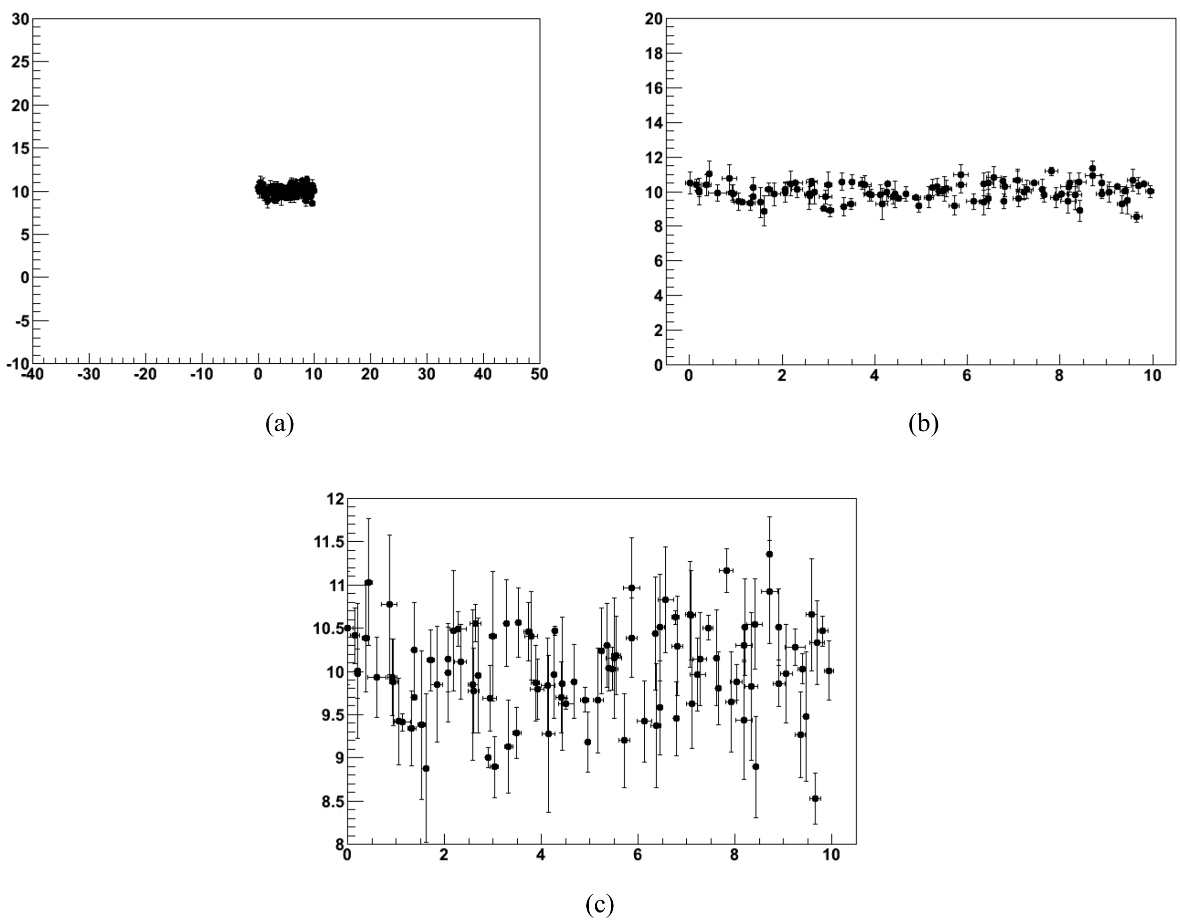


Figure 5: Graphs with too much, just enough, and too little space to be easy to interpret.

plotting our data, however. Graphs can be and frequently are drawn in ways intended to manipulate the perceptions of the audience, and this is a violation of scientific ethics. For example, consider Figure 6. It appears that Candidate B has double the approval of Candidate A, but a quick look at the vertical axis shows that the lead is actually less than one part in seventy. The moral of the story is that our graphs should always be designed to communicate our point, but not to create our point.

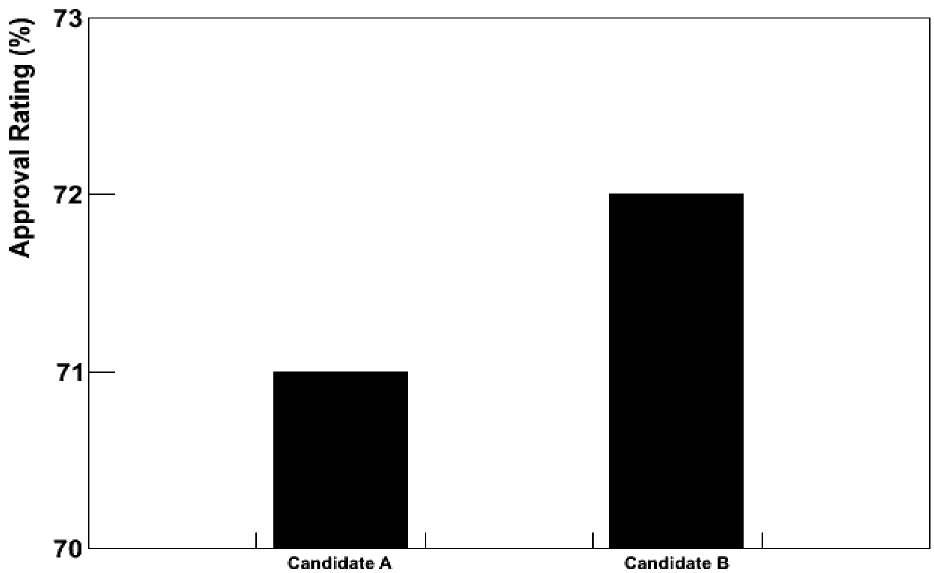


Figure 6: Approval ratings for two candidates in a mayoral race. This graph is designed to mislead the reader into believing that Candidate B has much more support than Candidate A.

Using Linear Relationships to Make Graphs Clear

The easiest kind of graph to interpret is often a line. Our minds are very good at interpreting lines. Unfortunately, data often follow nonlinear relationships, and our minds are not nearly as good at interpreting those. It is sometimes to our advantage to force data to be linear on our graph. There are two ways that we might want to do this in this class; one is with calculus, and the other is by cleverly choosing what quantities to graph.

The “calculus” method is the simpler of the two. Don't let its name fool you: it doesn't actually require any calculus. Let’s say that we want to compare the constant accelerations of two objects, and we have data about their positions and velocities with respect to time. If the accelerations are very similar, then it might be difficult to decide the relationship from the position graphs because we have a hard time detecting fine variations in curvature. It is much easier to compare the accelerations from the velocity graphs because we then just have to look at the slopes of lines; see Figure 7. We call this the “calculus” method because velocity is the first derivative with respect to time of position; we have effectively chosen to plot the derivative of position rather than position itself. We can sometimes use these calculus-based relationships to graph more meaningful quantities than the obvious ones.

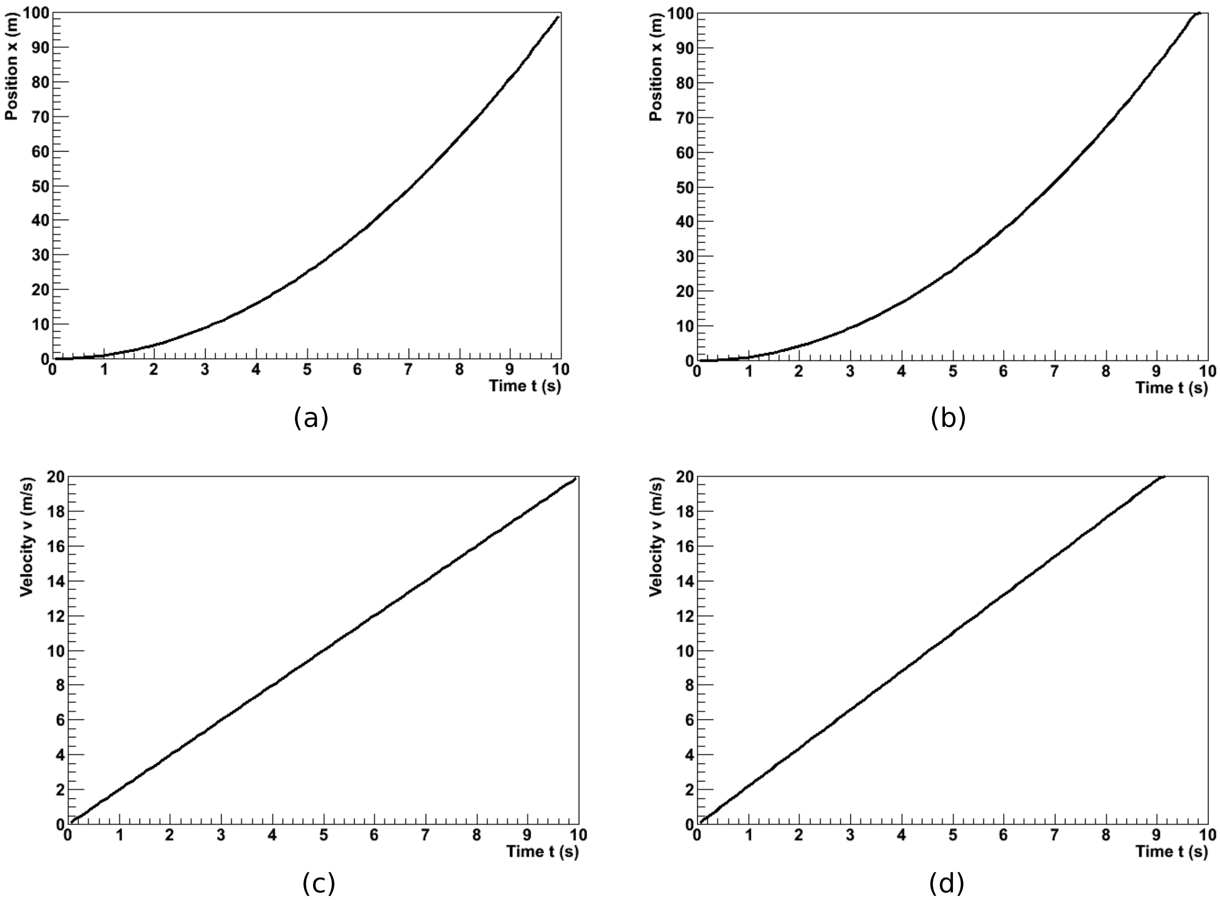


Figure 7: Position and velocity with respect to time for an objects with slightly different accelerations. The difference is easier to see in the velocity graphs.

The other method is creatively named “linearization.” Essentially, it amounts to choosing non-obvious quantities for the independent and/or dependent variables in a graph in such a way that the result graph will be a line. An easy example of this is, once again, an object moving with a constant acceleration, like one of those in Figure 7. Instead of taking the derivative and plotting the velocity, we might have chosen to graph the position with respect to $t^2/2$; because the initial velocity for this object happened to be 0 , this would also have produced a graph with a constant slope.

The Bottom Line

Ultimately, graphs exist to communicate information. This is the objective that we should have in mind when we create them. If our graph can effectively communicate our point to our readers, then it has accomplished its purpose.

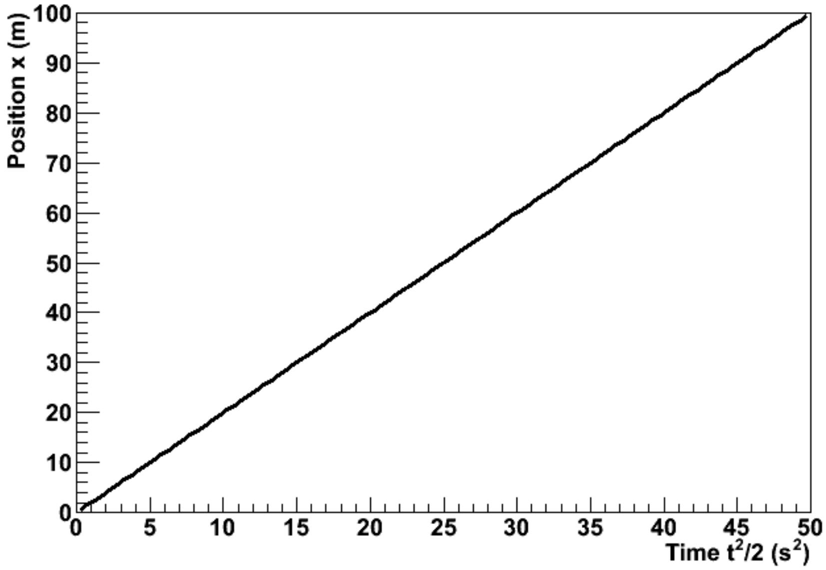


Figure 8: The position of the first object from Figure 7 plotted with respect to $t^2/2$. The relationship has been linearized.