

## Mathematics 102 — Fall 1999

### Introduction to functions

Calculus is about **functions**, which are a mathematical abstraction of a common notion. For example, we say that how much you weigh is a function of how much you can eat, or that how far your car can travel without a fuel refill is a function of how large the gas tank is. If we drop a heavy object from a certain height, then its height after that moment is a function of time. *A function expresses a relationship between two quantities.* There are many ways to express this more precisely, and indeed many ways to think of functions. In this course, our operational definition will be this:

- *A function is a rule for getting one number from another.*

In other words, you start with one number, the **input** to the function. You apply the function, and you get back a second number, the **output**. Functions are ubiquitous in modern life, although you might not know it, since computers are incessantly applying functions of various kinds to do what they do. In the language in which computing is discussed, the input is called the **argument**, and the output is called the **return value**.

There are many ways in which the rule defining a function might be formulated. Here are some examples:

- (1) *For an ideal object in free fall from rest at an initial height  $h_0$  near the Earth's surface, the height is a function of time and, according to Galileo's law, at time  $t$  the height will be  $h(t) = h_0 - gt^2/2$  where  $g$  is the gravitational acceleration at the Earth's surface, about 9.8 metres per second per second.*

Here the function is given by a simple formula, but that will not always be the case. On the contrary, sometimes the rule will be rather impractical, perhaps merely theoretical and even just barely imaginable.

- (2) *Galileo's Law for falling objects is in reality only approximate. The air flow around an object is complicated because of statistical and uncontrollable variations. It will cause friction. The object will be buffeted around a bit, and we cannot expect its position to follow Galileo's law precisely. Nonetheless, you can at least imagine keeping a more or less continuous record of the position, say by photographs. Consulting that record would then be the rule.*

It is rather common in science that a complicated experimental situation is well approximated by a simple theoretical model. A well designed experiment will try to separate carefully those phenomena which are accidental from those which are central to his experiment, and to understand the limits of theory.

- (3) *Suppose you are growing bacteria in a Petri dish. You put the dish out at midnight, and as time goes on bacteria drop in and begin to grow. The number of bacteria in the dish is a function of time, although it would be impossible to predict it, and almost impossible to measure it.*

In this case a more practical idea would be to measure the weight of bacteria instead of their population size, and make reasonable assumptions about the relationship between the two.

A functional relationship between two quantities is not symmetrical. The rule that defines the function specifies one of them in terms of the other. The one with which the rule starts—its input—is called the **independent variable**; and the other, the one the rule produces—its output—is called the **dependent variable**.

Sometimes the rule will be applied only to certain numbers. For example, if our rule gives  $1/x$  for a number  $x$ , then this rule makes sense only when  $x \neq 0$ . Likewise, the usual function which gives  $\sqrt{x}$  makes sense only when  $x \geq 0$ . A somewhat different example is the rule which assigns to an integer  $n$  the sum of the first  $n$  integers; in this case the rule is restricted to non-negative integers. And in experiments, if the rule involves consultation of a record, you will expect that only certain values are recorded. For example, if we take photographs of an object as it falls, you will know the position only for those moments when a photograph has been taken.

In any event, it is useful to have a name for the set of numbers for which the rule defining a function is to be applied. It is called its **domain**.

It is also useful to keep in mind that a function in the mathematical sense is not quite the same as a physical one. The position of a falling object is a physical function of time. A position is not a mathematical entity—you can

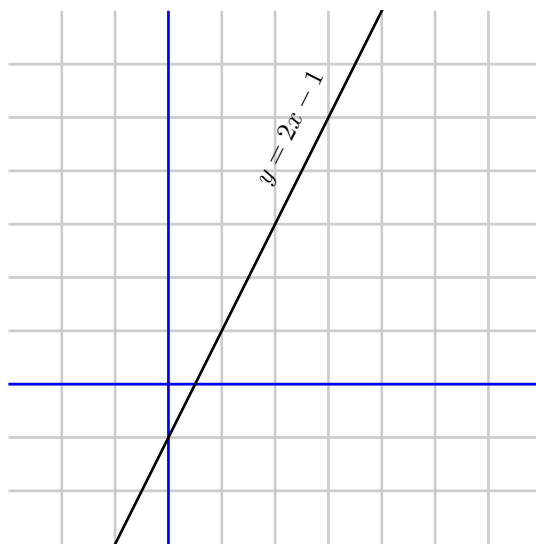
point to it or walk through it or hide it—since mathematics is an abstraction or idealization of reality. But you can also assign coordinates to a position and then, in terms of coordinates, position becomes a mathematical entity. Almost always in science, the first step in introducing mathematics is to choose a system of units to translate physical things to mathematical ones. This is related to something we have already seen in Galileo's experiment, since he worked with ratios, and a measurement in units is just a ratio of some quantity to a standard quantity of the same kind. For example, if we are simply counting bacteria then we are concerned with numbers . . . Well, not quite—how do we count a bacterium in the process of division? Is there such a thing as half a bacterium? But if we are measuring the weight of the bacteria, we are back to a choice of units. We repeat what we have said, because dealing with units is one of the most common nuisances in the use of mathematics in science: *a number in the mathematical sense is always a ratio of similar magnitudes.*

In the examples above, there are two quite different types of function represented: (1) theoretical or mathematical; (2) experimental. Those of the first type are rules related to some mathematical recipe. These are the ones computers might calculate. Those of the second type are data recorded in an experiment of some kind, whether imaginary or real. These are the ones computers will look up in a table. There is another dichotomy of functions, also—into those where the domain is a continuous range, and those where it amounts to a series of isolated values. We shall explain what this second dichotomy means by looking at the simplest examples of each type, here and in the next chapter.

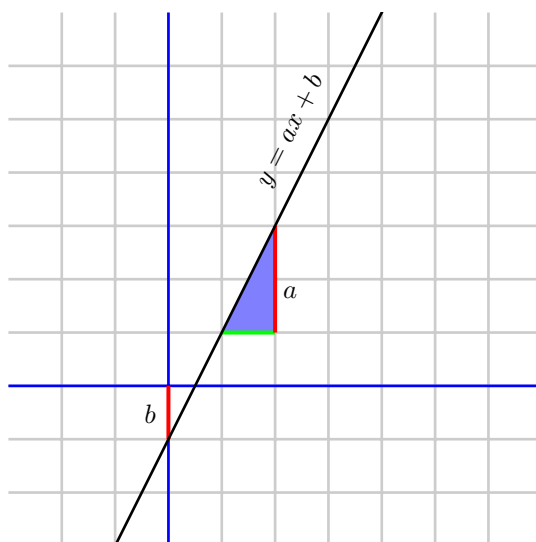
### Linear functions of a continuous variable

A function of a continuous variable is one whose domain includes a continuous range of values. These include functions you are familiar with, the ones given by an algebraic formula—for example,  $1 + 2x$  or  $x^2 - 1$ , where  $x$  can be any number at all. Among the simplest examples are the **linear functions**. Examples are the function whose rule assigns  $f(x) = 1 + 2x$  to  $x$ , or that with  $f(x) = 2 + x$ . The most general form of a linear function is  $ax + b$ , where  $a$  and  $b$  are constants.

We can picture a linear function by a simple graph:



The numbers  $a$  and  $b$  have geometrical significance: (1) If we set  $x = 0$  then  $ax + b = b$ , so that  $b$  is the value of  $f(x)$  for  $x = 0$ . Geometrically, this is the height of the graph where it crosses the  $y$  axis. Therefore  $b$  is called the  **$y$ -intercept** of the graph, (2) If we change  $x$  to  $x + 1$ , then  $ax + b$  changes to  $a(x + 1) + b = (ax + b) + a$ . In other words,  $a$  represents the change in  $f(x)$  corresponding to a change of 1 in the value of  $x$ . Both these features can be pictured:

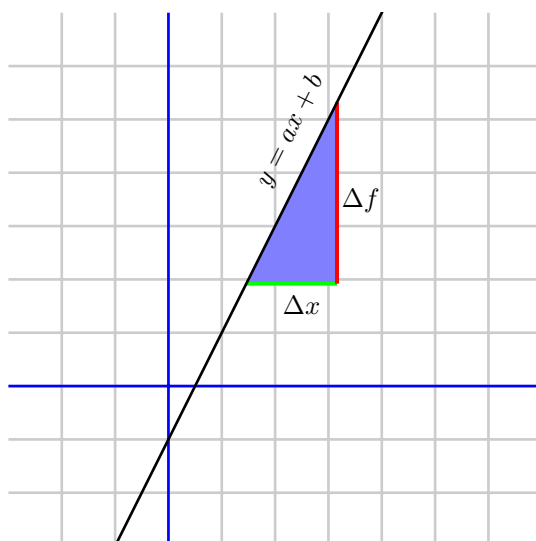


But the change in the value of  $x$  is not restricted to 1. It can be anything. If  $x$  changes by  $h$  then  $ax + b$  changes to  $a(x + h) + b = (ax + b) + ah$ , so the function changes by  $ah$ . That is to say, a change of  $h$  in  $x$  means a change of  $ah$  in  $y$ . In other words,  $a$  is the **relative change** of  $f(x)$  with respect to  $x$ , or the ratio of the change in  $f(x)$  to that in  $x$ .

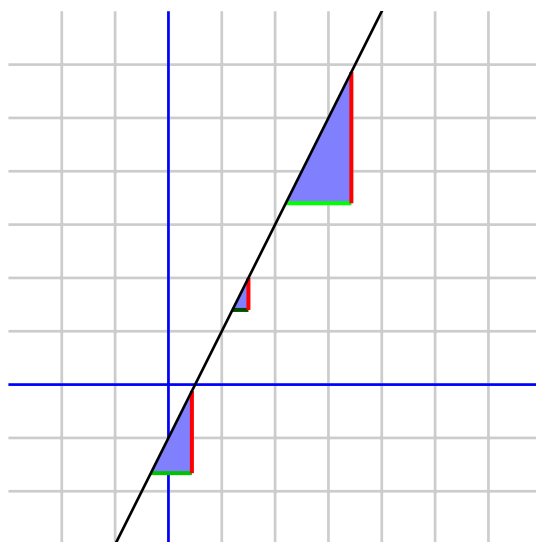
In this course, we shall use a mathematical symbol for ‘change in’, the Greek letter for a capital  $D$ , which is  $\Delta$  (because ‘difference’ starts with ‘d’). So we represent the change in  $f$  by  $\Delta f$  and that in  $x$  by  $\Delta x$ . We can now say that if  $f(x) = ax + b$  is a linear function of  $x$  then

$$a = \frac{\Delta f}{\Delta x}$$

for any change  $\Delta x \neq 0$  in the value of  $x$ . Or, in other words,  $a$  is the **slope** of the graph of  $y = ax + b$ .



It is important to keep in mind that *for linear functions, the change in the value of the function is always proportional to the change in the value of  $x$* . In geometric terms, this means that the triangles in the following figure are all similar.



Linear functions are thus characterized by these two properties: (1) *Increments (changes) in the value of the function depend only on the difference between two values of  $x$ .* (2) *This increment, which depends only on that difference, is proportional to it.*

The slope of a linear function  $f(x) = ax + b$  can be calculated from the values of  $f$  at two distinct values of  $x$  by the formula

$$a = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

And of course  $b$  is the value of  $f$  at  $x = 0$ .

**Exercise 1.** Graph  $y = -1 + 3x$  for  $x = -10$  to  $10$ .

**Exercise 2.** Graph  $y = -1$  for  $x = -10$  to  $10$ .

**Exercise 3.** Graph  $y = -1 - 3x$  for  $x = -10$  to  $10$ .

**Exercise 4.** In geometrical terms, how can you tell whether slope is negative? zero?

### Sampling

In this course we shall deal nearly all the time with two kinds of functions. One of them includes functions of a **discrete** variable, where the possible values of the independent variable are selected from a range of discrete (isolated) values. The other includes functions of a **continuous** variable, such as  $1 + x$  or  $x^2$ , where  $x$  ranges over all numbers. The distinction between these two kinds of functions, those of a discrete variable and those of a continuous variable, is fundamental, but there are relationships between the two which blur that distinction. For example, here is the graph of the temperature of the coffee sitting in a cup, as a function of time. The temperature when the coffee is poured is  $90^\circ$ , and the room temperature is  $24^\circ$ .



Now if you think about it, you realize this is probably a fraudulent claim. How can we make a truly continuous record of temperature? This would involve an infinite amount of work, which we certainly didn't have at our

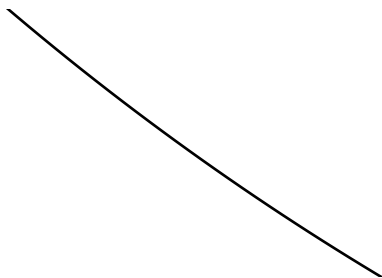
disposal. So we have to be cheating. We haven't really made a continuous record of the temperature—we have measured the temperature at a number of closely spaced points in time, and then connected the recorded points together to make the graph. If you look with a microscope at our graph you will see this:



Realistic data from a continuous process such as an object in free fall or a cooling coffee cup are always produced by **sampling**, which takes measurements at a discrete set of times. The sampling data are very often treated in some way to simulate the original continuous process by **interpolating** the discrete data, to obtain a kind of copy of it which can be manipulated and analyzed with mathematical tools. Sampling and interpolation have been used since the very beginning of science, as Galileo's experiments will confirm, but in the current era they are particularly widespread because computers make it much more practical to do it extensively.

This happens, for example, in recording music. Samples of sound amplitude are taken at a very fast frequency, roughly 44,000 cycles per second, which is a bit more than twice the highest frequency a human being can hear. They are put into a kind of digital format as **bits** and recorded as blips on a CD. The CD player then reconstructs the sound through a process called **sample and hold** because it holds fixed the sampled value until the next time interval. This is the method of interpolation we saw above. It is used to reproduce sound because it is very fast, and the sound quality is satisfactory if not perfect.

Another technique to produce continuous data from discrete data is to **linearly interpolate** the sample, which we shall discuss in a moment. The output looks much smoother, but it requires more calculation. For example, for the cooling coffee cup linear interpolation would give after blow-up



It is just about impossible to tell this from a smooth curve at this resolution.

### Linear interpolation

We have so far described a linear function in terms of slope and  $y$ -intercept. But there are several other ways to obtain one. For a example, a line is determined by two points, and likewise the graph of a linear function is determined by two points on it. Therefore if we are given two points  $(x_0, y_0)$  and  $(x_1, y_1)$  with  $x_0 \neq x_1$  there will be a unique linear function  $f(x) = ax + b$  with  $f(x_0) = y_0$ ,  $f(x_1) = y_1$ . What are  $a$  and  $b$ ? Our assumption means neither more nor less than that we have these conditions:

$$\begin{aligned} ax_0 + b &= y_0 \\ ax_1 + b &= y_1 . \end{aligned}$$

If we subtract one equation from the other we get

$$a = \frac{y_1 - y_0}{x_1 - x_0}$$

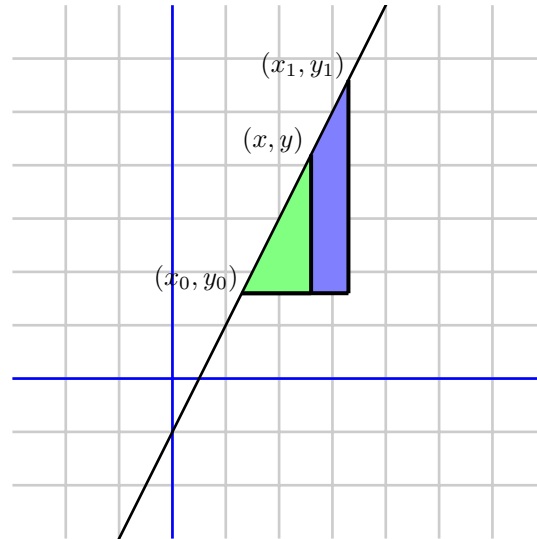
and then solve one of the two to get

$$b = y_0 - ax_0 .$$

We can also write

$$\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0}$$

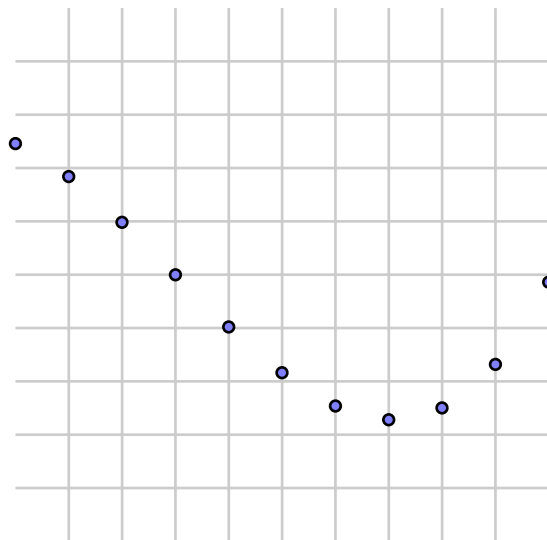
which you can also see from this picture:



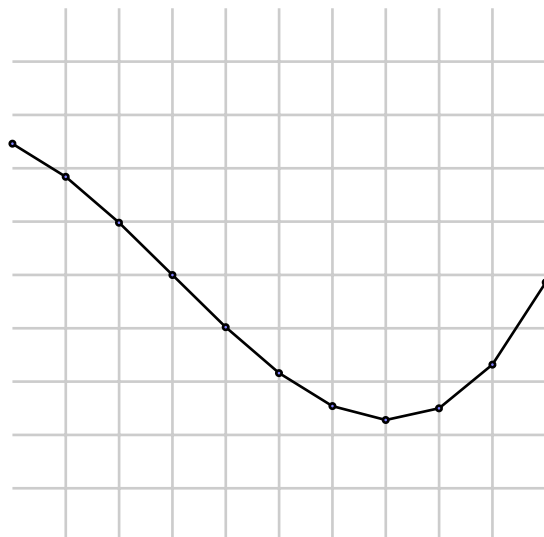
From this in turn you can solve for  $y$  to get

$$y = \left( \frac{y_1 - y_0}{x_1 - x_0} \right) (x - x_0) + y_0 .$$

These formulae are often used in **linear interpolation**. Suppose we are given a set of data points in a sequence, which we put into a plot:



**Interpolation** is what you do to these data points to obtain a function of a continuous variable, or equivalently the graph of a continuous function going through them. There are lots of ways to do this, each valuable in some circumstances. Linear interpolation is one of the most common. It builds the new graph as a polygon connecting the given sequence of points:

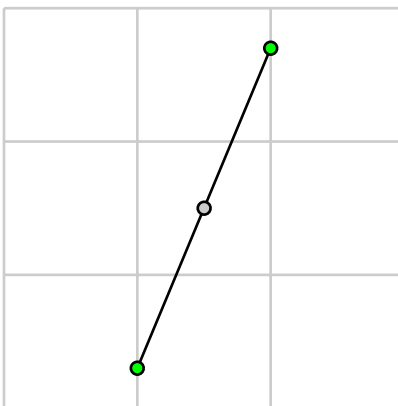


Geometrically, this is a simple process. If we want to calculate values of the new function in between our data points, then we can use the formula we derived earlier in this section. But let's look at an example, starting from scratch.

Suppose we are given just two data points  $(1, 0.3)$  and  $(2, 4.7)$ . This means we are thinking of some function  $f(x)$  with

$$\begin{aligned} f(1) &= 0.3 \\ f(2) &= 2.7 . \end{aligned}$$

Suppose we want to estimate  $f(x)$  for some value of  $x$  between 1 and 2, say 1.5. Using linear interpolation to do this means we want to figure out the  $y$ -coordinate of the middle point in this picture:



The picture explains how we do it. The value 1.5 is half-way from  $x = 1$  to  $x = 2$ , so the value of  $y$  we are looking for is half-way between 0.3 and 2.7. Let  $\Delta y$  be the difference between 0.3 and the  $y$ -value we are seeking, and let  $\Delta x = 0.5$  be the difference in  $x$ -values. Then

$$\frac{\Delta y}{\Delta x} = \frac{2.7 - 0.3}{2 - 1} = 2.4, \quad \Delta y = 2.4 \Delta x = 2.4 \cdot 0.5 = 1.2 .$$

In other words, the difference in  $y$ -value is a simple proportion of the difference in  $x$ -value. The  $y$ -value itself is then

$$0.3 + \Delta y = 1.5 .$$

**Exercise 5.** With the same two data points, interpolate a value for  $x = 1.25$ . Draw the picture that accompanies the calculation.

**Exercise 6.** Here are the data for free fall from 5 metres on the planet Krypton.

---

$t$	$h$
0.0	5.0000
0.1	4.9878
0.2	4.9510
0.3	4.8895
0.4	4.8030
0.5	4.6910

*Use linear interpolation to estimate the height  $h$  at  $t = 0.41$  seconds.*

There are other methods of interpolation, too. We shall see a few more later on. Each method of interpolation has its pluses and minuses, usually a trade-off of accuracy for speed.