1 Functions

A function is like a sausage machine - you put something in, and you get something out. What's put in can vary. Often it is simply a real number, in which case the function is said to be "single variable" and in other cases it might be several numbers - in which case the function is "multivariable". And what comes out also can vary. Sometimes it is a number - in which case it is called "single valued" and other times a point in two, three or some higher dimensional space. As for the function - it consists of simply some rules for calculation based on the input. In this course we shall consider most often single valued functions of a real variable - that is a real number goes in and a real number comes out.

Mathematics is built around the concept of a function. Functions are used in our efforts to understand the world around us. They form the underpinnings of our complex society - the construction of aircraft and all their complexity, modeling meteorological data, search engines for the internet, video and audio transmission, the exploration of space, medical imaging, the stock market, and so on.

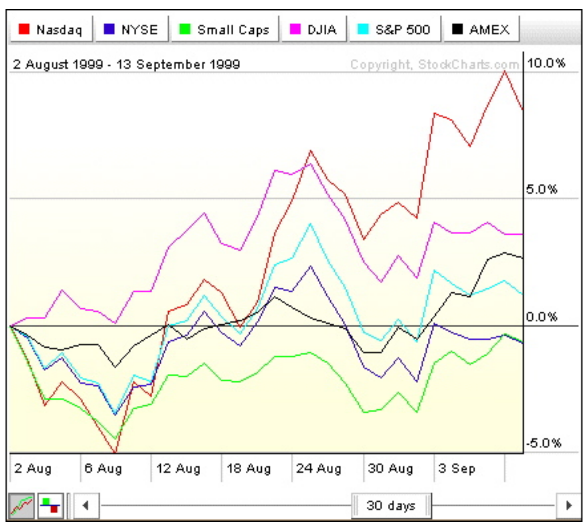

Figure 1: modeling stock market data

In what follows I plan to introduce a variety of different types of functions, polynomial functions, rational functions, exponential and logarithmic functions, and trigonometric functions. I also wish to develop techniques for analysis of functions. That is, once we have described a function that models a particular problem, is there further information that can be derived. For instance, suppose we have been following the market price of a particular stock. Is there a way of predicting price, within some bounds, in the following week? Or more simply, what shape should a manufacturer settle on that would both minimize the metal used in construction and at the same time maximize the volume? But before we get too far into details, I would like to set the stage with some generalities and to give a taste of the bigger picture.

1.1 Generalities

The notion of a function can be phrased in a very general way. This has the advantage in that various types of function are only examples of a general idea. The general definition is amazingly simple - deceptively so perhaps. But that is the beauty of mathematics - simple ideas clearly stated end by having immense power. So here's the definition.

Definition 1 *Given two sets of objects A and B a function f from A to B is simply some mechanism by which every element $a \in A$ is assigned some unique element called $f(a)$ in B. We write $f : A \to B$ to abbreviate all of this.*

Notation

- The set A is referred to as the *domain* of f
- The *range* of f is the set of all objects in B of the form $f(a)$. That is the range of f is the set $\{f(a) : a \in A\}$ and is written simply as $f(A)$

The sets A and B can be very general, but for our purposes, they will normally consist of some subsets of the real numbers, points in the plane, or points in three dimensional space. Points in the plane correspond to all pairs of real numbers (x, y) and the set of all such points is written as $\mathbb{R}^2 = \{(x, y) : x \in \mathbb{R} \text{ and } y \in \mathbb{R}\}$. Points in three dimensional space $\mathbb{R}^3$ similarly correspond to triples (x, y, z) of real numbers and the set of all such points is denoted $\mathbb{R}^3 = \{(x, y, z) : x \in \mathbb{R}, y \in \mathbb{R}, \text{ and } z \in \mathbb{R}\}$. See figure 2.

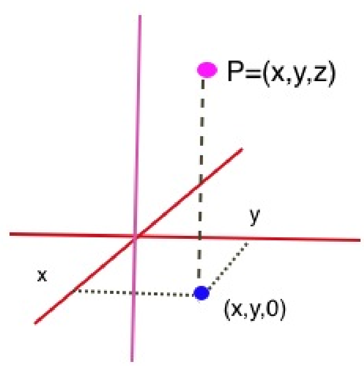

Figure 2: A point in $\mathbb{R}^3$

In the case that A is a subset of the real numbers and x is an arbitrary element of A, x is also said to be an *independent variable*, and if we set $y = f(x)$ then y is said to be the associated *dependent variable.*

If A on the other hand is a subset of the plane, consisting of points $X$of the form $X = (x, y)$ then a function $f : A \to \mathbb{R}$ is a function of two variables (independent variables) x and y. Notice that the letter y in the previous paragraph denoted the output of a function and here it is part of the input. In this case the output or dependent variable is referred to as z, and we write $z = f(x, y)$.

A function is most often defined by stating the rule one uses to obtain the output. For instance the function that squares a number and then adds 2 to the result, is defined by writing $f(x) = x^2 + 2$. The domain of a function if not otherwise specified is the largest collection of numbers for which the definition makes sense. In this case the domain is considered to be the totality of all real numbers - namely the set $\mathbb{R}$. The range is the set of all numbers of the form $x^2 + 2$, where x is an arbitrary real number, and is nothing more than the infinite interval $[2, \infty)$.

Here the equal symbol = in the expression $f(x) = x^2 + 2$ is used to denote a definition not equality as we usually understand it. More specifically the = sign in this context is meant as an abbreviation for the words “is defined by”.

Definition 2 *The graph of a function is the set of all pairs* $(a, f(a))$, *for* $a \in A$.

Observe that in this general definition we have not specified the nature of the elements of A. In the case that A is a subset of the real numbers $\mathbb{R}$, then A contains real numbers x. If A is a subset of the plane, then contains pairs (x, y) of real numbers indicating points of the plane. In other cases A may contain triples, (x, y, z) of real numbers indicating points of three-dimensional space $\mathbb{R}^3$. In the case that A and B are subsets of $\mathbb{R}$, the graph corresponds to some figure in the pane. For the example above of the function $f(x) = x^2+2$ the graph is shown in figure 3. Further examples occur in the following section 1.1.2.

If A is a subset of the plane, then the graph is the set of all points $P = (x, y, f(x, y))$ where (x, y) is an arbitrary element of A, and the third coordinate $f(x, y) = z$ corresponds to the distance of the point P above or below the plane consisting of all points $(x, y, 0)$.

Suppose now that there are two functions $f : A \to B$ and $g : A \to B$. What would it mean for f and g to be equal? A function is a rule for taking input and creating output. So for two functions to be equal they must produce the same output given the same input. In other words $f = g$ means that for all $a \in A$ we must have $f(a) = g(a)$. In order for this to happen the rules for calculation defined by f and by g must be the same. For instance if $f(x) = (x-1)^2$ and $g(x) = x^2-2x+1$, where x is any real number, then through algebraic manipulations we can show that $f(x) = g(x)$ for all $x \in \mathbb{R}$.

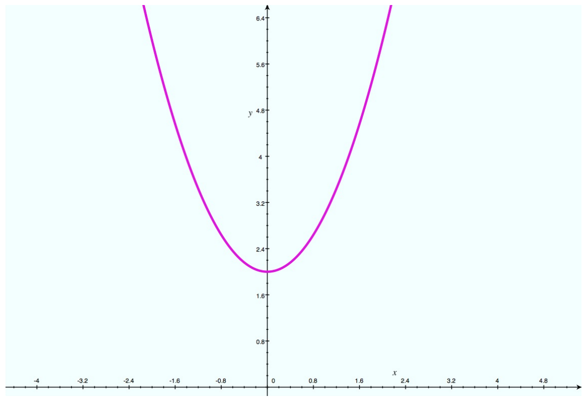

Figure 3: Graph of the function $f(x) = x^2 + 2$

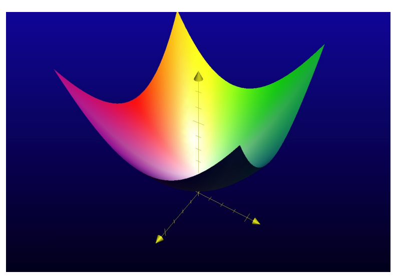

Figure 4: Graph of the function $f(x, y) = x^2 + y^2 + 2$

1.1.1 Equations

The notion of an *equation* is something you have been living with since at least grade nine. What is an equation? We can give lots of examples - but what is it? An equation is simply an expression in which two function definitions are connected with an equal sign - nothing more. For instance,

$$x^2 - 2y + 5 = y^2 + x - 3,$$

where on the left there is the function of two variables x and y defined by $x^2 - 2y + 5$ and on the right the function of two variables defined by $y^2 + x - 1$.

Definition 3 *Given sets A and B and functions $f : A \to B$ and $g : A \to B$, an equation is an expression $f(a) = g(a)$ in which two functions are connected with an equals sign.*

Whenever we have such an equation we are always presented with the problem of finding which values of a actually give equality when substituted in the equation. So with

$$x^2 - 2y + 5 = y^2 + x - 3$$

we can ask which values of the variables x and y give equality. The set of all such points describes the *graph of the equation.* For this example, if $x = 1$ and $y = 2$ are substituted into the equation, there is equality, so hence, the point $(1, 2)$ is on the graph. The complete graph plotted with graphing software is is shown below.

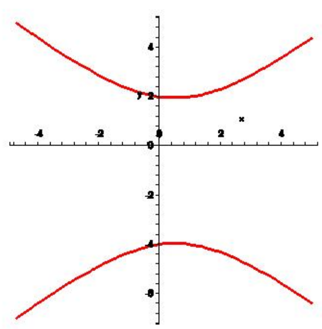

Figure 5: graph of equation $x^2 - 2y + 5 = y^2 + x - 3$

Now every function f of a real variable also describes an equation of two variables $f(x) = y$ and the graph of the equation is precisely the graph of the function. The same is true for functions of several variables. For instance if f is a function of two variables then there is the equation $f(x, y) = z$, and again the graph of the equation is again the graph of the function.

The reverse is not usually true. That is given an equation and its graph, we cannot usually find a function whose graph is the graph of the equation. An example is the equation $x^2 + y^2 = 1$ which is the graph of a circle of radius one centered at $(0, 0)$. If we try to solve for y in terms of x, thereby getting an expression for output y for input x, we see that

$$y = \pm\sqrt{1 - x^2}.$$

This does not describe a function for the reason that given one number as input, the output becomes two numbers, $+\sqrt{1 - x^2}$ and $-\sqrt{1 - x^2}$. Not allowed. Functions must provide unique output.

1.1.2 Examples

Following are a few examples.

- **A constant function:**

  Given some constant say $c = 5$ and letting both A and B be the set real numbers $\mathbb{R}$, we can define a function $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = c$. The graph is just the set of all points in the plane of the form (x, c) where x is any real number. The graph is horizontal line through the point $(0, 5)$ on the y-axis.

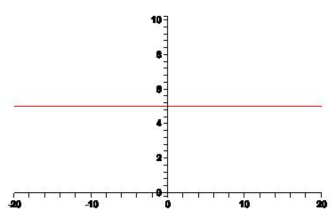

Figure 6: graph of constant function $f(x) = 5$

- **A linear function:**

  For A and B again taken as the set of real numbers $\mathbb{R}$. A line in the plane may be described as the graph of a function $f : \mathbb{R} \to \mathbb{R}$ defined as $f(x) = mx + b$, for constants m and b. The constant b is chosen so that the line intersects the y-axis at the point $(0, b)$ and m measures the slope of the line. More precisely, given any two points on the line m is the vertical distance between the two divided by the horizontal distance. In other words if (x_1, y_1) and (x_2, y_2) are two points on the line, then $m = \frac{y_2 - y_1}{x_2 - x_1}$ is the point.

  Example: For f defined by $f(x) = 3x + 2$, the graph is below. Notice that the graph passes through the point 2 on the y axis and has slope 3, by which we mean that for

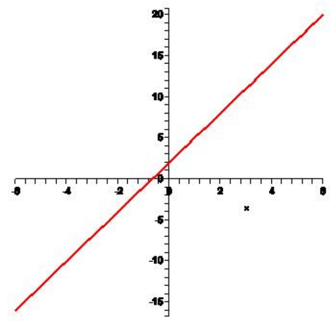

Figure 7: graph of the linear function $f(x) = 3x + 2$

every horizontal unit the graph goes up vertically 3 units.

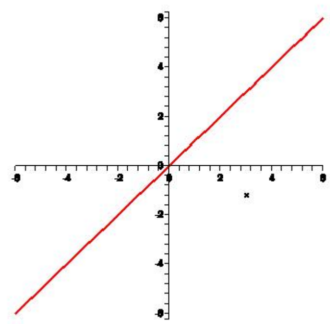

Figure 8: graph of the identity function $f(x) = x$

- **Identity function** $1_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ **:**

  A particular case of a linear function is the *identity* function $1_{\mathbb{R}}(x) = x$. The graph may viewed simply as the graph of the equation $y = x$ which is the set of points (x, x). In other words the graph is the diagonal line passing through the origin $(0, 0)$.

- **Identity function** $1_A : A \to A$ **:**

  The concept of an identity function exists for any set, for if A is any set of objects we can define the function $1_A : A \to A$ by $1_A(x) = x$

- **The absolute value function:**

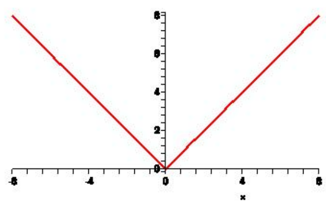

Figure 9: graph of absolute value function

Recall that absolute value function is the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = x, \text{ for } x >= 0$$

$$f(x) = -x, \text{ for } x < 0$$

This means that the graph consists of two pieces, one for $x >= 0$ and one for $x < 0$.

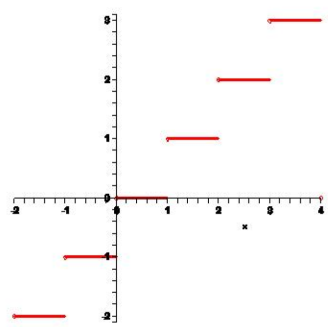

Figure 10: standard step function

- **A step funciton:**

  A step function is another example of a function which has different definitions on different parts of the real line. The standard version is the function $f : \mathbb{R} \to \mathbb{R}$ which is defined on each interval $[n, n+1)$ by $y = f(x) = n$ for $n \leq x < n+1$. That is on each interval $[n, n+1)$ the function takes on the constant value n. This function is also called the *greatest integer function*. This is because the function can be simply defined by saying for any value of x, $f(x)$ = the greatest integer less than or equal to x. The function is written $f(x) = \lfloor x \rfloor$,

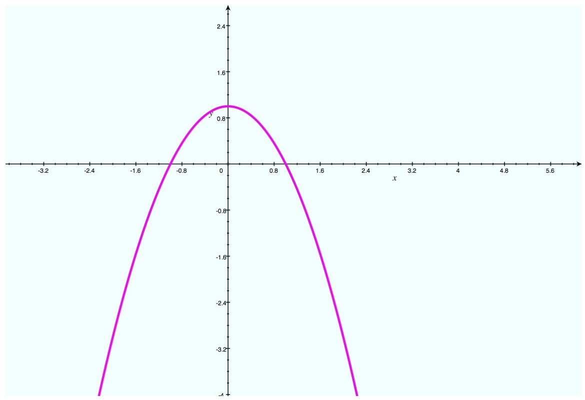

Figure 11: Graph of quadratic function $f(x) = -x^2 + 1$

- **A quadratic function:**

  A general quadratic function $f : \mathbb{R} \to \mathbb{R}$ for constants a, b and c is defined by the function $f(x) = ax^2 + bx + c$. In the case that $a = 1, b = 0$, and $c = 2$ we have the example mentioned before, $f(x) = x^2 + 2$ - see figure 3. If $a = -1, b = 0$, and $c = 1$, we have $f(x) = -x^2 + 1$ and the graph is shown in figure 11. Observe that setting $a = -1$ causes the shape of the curve to be inverted.

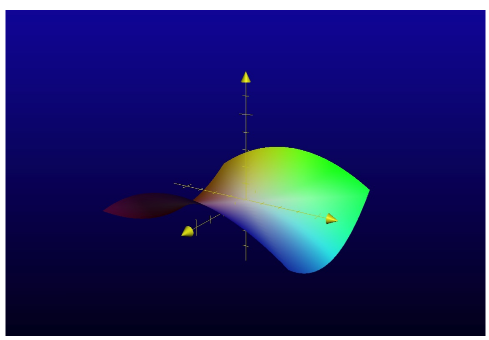

Figure 12: Graph of the multivariable function $f(x, y) = x^2 - y^2$

- **A multivariable function:**

  The graph of a real -valued multivariable function $f : A \to \mathbb{R}$, where A is some subset of the plane $\mathbb{R}^2$ consists of all points in three dimensional space $\mathbb{R}^3$ of the form $\Big(x, y, f(x, y)\Big)$. For the case of the function defined by $f(x, y) = x^2 - y^2$ the graph is shown in figure 12.

1.2 Terminology

In this section terminology is introduced that is useful in understanding the idea of inverse functions.

1.2.1 Injective functions

Given sets A and B a function $f : A \to B$ is said to be *one-to-one* or *injective* if different input gives different output. More precisely there is the following definiton.

Definition 4 *Given sets A and B, to say a function $f : A \to B$ is injective means that for arbitrary points $a_1 \in A$ and $a_2 \in A$, if $a_1 \neq a_2$, then $f(a_1) \neq f(a_2)$.*

This may be expressed with tools of logic as follows:

Definition 5 *To say that a function $f : A \to B$ is injective means that:*

$$\forall a_1 \in A \ \forall a_2 \in A, \ a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2).$$

The quadratic function defined by $f(x) - x^2 + 1$ in the example above is clearly not injective since $f(x) = f(-x)$ for all x, and in particular $f(-1) = f(1)$.

On the other hand in the example of the identity function defined by $f(x) = x$, since the graph of the line is always increasing as we go from left to right, it follows that for $x_1 \neq x_2$, if $x_1 < x_2$, then $f(x_1) < f(x_2)$ and hence $f(x_1) \neq f(x_2)$; similarly if $x_2 < x_1$, then also $f(x_1) \neq f(x_2)$. Thus the identity function is injective.

Horizontal line test: Observe that if A is some subset of $\mathbb{R}$, a function $f : A \to \mathbb{R}$ cannot be one-to one if a horizontal line intersects the graph in more than one spot. If this should happen, it indicates that there are two real numbers which give the same output. And if horizontal lines intersect only once or not at all then we know that distinct input must give distinct output. This is sometimes called the *horizontal line test.* See figure 13.

1.2.2 Surjective and bijective functions

An arbitrary function $f : A \to B$ is said to be *onto* or *surjective* if for each $b \in B$ is the image of some element of A, by which we mean that there is some $a \in A$ such that $f(a) = b$. More precisely there is the following definition.

Definition 6 *Given sets A and B to say a function $f : A \to B$ is surjective means that for each $b \in B$ there is at least one $a \in A$ such that $f(a) = b$, in which case we say that the function f is surjective onto the set B.*

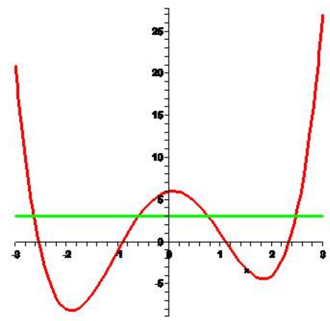

Figure 13: The graph of a function that fails the horizontal line test

Using the language of logic this may be rephrased.

Definition 7 *To say a function $f : A \to B$ is surjective means that*

$$\forall\ b \in B, \exists\ a \in A \,\text{such that}\, f(a) = b.$$

Observe that since the range of a function $f : A \to B$ is defined to be all the elements of B of the form $f(a)$ for $a \in A$, it follows that every function is surjective onto its range. In other words a function $f : A \to f(A)$ is always surjective.

Examples :

- The quadratic function defined by $f(x) = x^2 + 1$ has domain the entirety of the real real numbers $\mathbb{R}$ and its range is the interval $[1, \infty$. That is for every number $y \in [1, \infty)$, we can find $x \in R$ such that $x^2 + 1 = y$. How do we do this? - simply solve for x in terms of y, getting two choices for x, $x = +\sqrt{y-1}$ or $x = -\sqrt{y-1}$.
- The rational function $f(x) = \frac{1}{x^2-1}$ on the other hand has as its domain the set all real numbers not equal to one. We write this as $\mathbb{R} - \{1\}$. The range on the other hand is the infinite interval $[-1, \infty)$. How do we know this? - again solve for x in terms of y. Doing this we see that we have two choices for x, $x = +\sqrt{\frac{1}{y} + 1}$ or $x = -\sqrt{\frac{1}{y} + 1}$. In either case we need to know that $\frac{1}{y} + 1$ is non-negative, for it does not make sense to consider square roots of negative numbers in this context. So, knowing that $\frac{1}{y} + 1 \geq 0$, adding -1 to both sides gives: $\frac{1}{y} \geq -1$. Multiplying both sides by y gives: $1 \geq -y$. Finally multiplying both sides by -1 and reversing the inequality (which is always necessary when multiplying both sides of an inequality by a negative number), we get $-1 \leq y$. In other words $y \in [-1, \infty)$.

Combining the notions of what it means for a function to be injective with that of it being surjective, we have the following definition of a bijective function.

Definition 8 *Given sets A and B, a function $f : A \to B$ is bijective if it is both injective and surjective.*

Examples:

- For any set A, the identity function $1_A : A \to A$ is bijective.
- Any linear function $f : \mathbb{R} \to \mathbb{R}$ is bijective.

1.2.3 Composition of functions

The notion of composing functions - that is first doing function and then another - is an easy idea, but one for which we need to establish some conventions, for it is a notion that occurs over and over in occasional complex situations.

Definition 9 *Given two functions $f : A \to B$ and $g : B \to C$, where A, B, C are any sets. The composition of f then g is the function $g \circ f : A \to C$ from A to C defined by $(g \circ f)(x) = g\Big(f(x)\Big)$*

In other words the output of f becomes the input of g

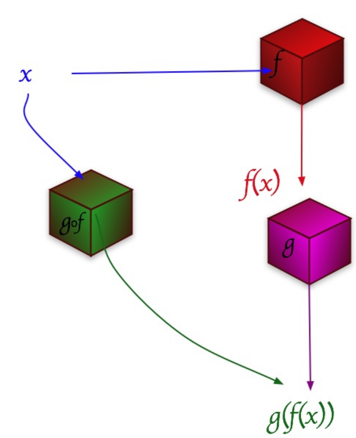

Figure 14: Composition of f then g

What does this mean in practice? For an example lets consider two functions $f : \mathbb{R} \to [0, \infty)$ and $g : [0, \infty) \to [0, \infty)$ defined by $f(x) = x^2 + 1$ and $g(y) = \sqrt{y}$. It follows that

$(g \circ f)(x) = g(f(x)) = \sqrt{x^2 + 1}$.

One of the things that one needs to be careful of when dealing with compositions is the need to make sure that the range of the function f is contained in the domain of the function g. In terms of the definition above, we need to make sure that the range $f(A)$ is contained in the set B. For instance if we had defined f by $f(x) = -x^2$, then for any number other than zero, the composition would not make sense; for example if we were to have $x = 1$, then $f(1) = -1$ and $g(f(1)) = g(-1)$ does not make sense, since g is only defined for non-negative real numbers.

1.2.4 Inverse functions

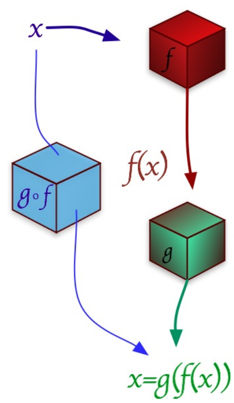

Figure 15: g is the inverse of f

The inverse of a function f is another function g with the property that if you do f followed by g or the reverse, g followed by f, then in each case you get back to where you started from. More formally we have the following definition

Definition 10 *Given a function $f : A \to B$, a second function $g : B \to A$ is the inverse of f, and also f is the inverse of g, provided*

1. $(g \circ f)(x) = x$ *for each* $x \in A$
2. $(f \circ g)(y) = y$ *for each* $y \in B$

or more succinctly: $(g \circ f) = 1_A$ *and* $(f \circ g) = 1_B$.

It is important to note that an identity function $1_A : A \to A$ must be bijective, and consequently if we have two functions f and g such that $g \circ f = 1_A$ is the identity, then

f must be injective or else the composition could not possibly be injective and further g must be surjective or else the composition could not be surjective. These ideas will be expanded on in the exercises. If we apply this insight to the case in which f is the inverse of g, we have $g \circ f = 1_A$ and $f \circ g = 1_B$, so that it follows that both f and g must be bijective.

Here are some examples -

1. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 2x + 1$. It is not hard to convince yourself that $f$must be bijective. Its inverse g can be easily calculated - set $y = 2x + 1$ and solve for x in terms of y. We get $x = \frac{1}{2}y - \frac{1}{2}$. If we define $g : \mathbb{R} \to \mathbb{R}$ by $g(y) = \frac{1}{2}y - \frac{1}{2}$, it follows that g is the inverse of f - just do the calculations and show that:
   - $g(f(x)) = x$ for all x
   - $f(g(y)) = y$ for all y.
2. Let $f : [0, \infty) \to [0, \infty)$ be defined by $f(x) = \sqrt{x}$. Again it is evident that f is bijective. It is certainly injective, since the square root of different numbers give different answers, and it is also surjective, since any positive number is the square root of some other number. But to find the inverse, we do as before. Let $y = +\sqrt{x}$ and solve for x. This gives $x = y^2$. We then define $g : [0, \infty) \to [0, \infty)$ by $g(y) = y^2$, and calculations show that g is the inverse of f - namely $g(f(x)) = g(+\sqrt{x}) = (+\sqrt{x})^2 = x$, for all non-negative x. Likewise f is the inverse of g as the following calculations show, $f(g(y)) = f(y^2) = +\sqrt{y^2} = y$, for all non-negtive y.

Finally, if $f : A \to B$ is a function that has an inverse $g : B \to A$, where A and B are intervals, there is an easy way of figuring out the graph of the g from looking only at the graph of f. Here's how we do it. Suppose (x, y) is a point of the graph of f. This means that

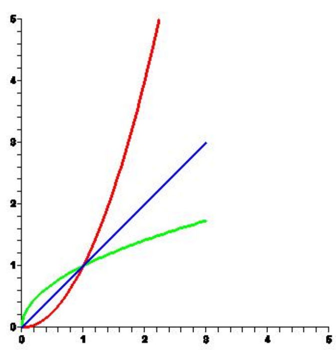

Figure 16: Reflection of $f(x) = x^2$ in red about diagonal in blue gives the graph of inverse $g(x) = \sqrt{x}$ in green

$f(x) = y$, and since g is the inverse of f, we know that $g(y) = x$, which tells us that (y, x) is on the graph of g. The final step then is to recognize that (y, x) is simply the reflection about the diagonal line $y = x$ of the point (x, y); this follows from a simple exercise in analytic geometry, which you will find as one of the exercises. See figure 16. The result can be expressed as follows:

Remark 11 *Given a function $f : A \to B$ with an inverse $g : B \to A$, where A and B are intervals, the graph of g is obtained reflecting the graph of f about the diagonal line $y = x$*