Coordinates and Equations for Lines

1. Coordinates on a Line

Given a line, a *one-dimensional coordinate system* on that line is made as follows:

1. First, make the following choices:
   a. choose a point on the line and call it O;
   b. choose a second point P on the line and call OP the "unit interval".
2. Using the unit interval as a unit of measure, assign to each point on the line the real number measure of its "oriented distance" from O. The oriented distance between O and a point X different from O is either positive or negative; it's positive if P and X are on the same side of O and negative if P and X are on different sides.

Fact 1. *A one-dimensional coordinate system constructed as above on a given line creates a bijection (see the following section) between the points on that line and the set of real numbers, $\mathbb{R}$. The point O corresponds to the number 0. If points A and B correspond to numbers a and b, then the distance between A and B, as measured by the unit interval OP of the coordinate system, is $|b - a|$.*

We will treat this fact as a postulate. The idea of a bijection, as well as some other things, may need clarification. Therefore, we review the idea of an (abstract) function.

2. Functions

In this section I will explain the meaning and use of the following mathematical words: *function, domain, codomain, injective, surjective, bijective, inverse function.* I will give intuitive examples of these ideas and I will explain their relevance to one-dimensional coordinate systems. I assume that readers understand what a set is and what it means for a thing to be an element of a set. As usual, "$s \in S$" is shorthand for, "s is an element of S."

All the ideas in this section arise within high school mathematics. The "New Math" of the 1960s attempted to begin teaching these ideas in even lower grades, and it was partially successful in doing so. The precise choice of vocabulary varies from curriculum to curriculum, and the extent to which the ideas are made explicit and fully explained varies. In some high school curricula, all the vocabulary discussed here is carefully defined, fully explained and used extensively. Ideas and vocabulary relating to functions become essential in calculus, and they remain important throughout all of the mathematics beyond calculus, no matter how advanced.

Functions. Suppose S and T are sets. For example, S might be (the set of points on) one line and T (the points on) another. A *function from S to T* is a rule that instructs us to select a certain element of T whenever we are given an element of S. The selection must depend only on the element of s given. It is understood, moreover, that the rule does not change; each time we are given a particular element of S, the rule chooses the same element of T. Finally, the rule must tell us what to pick no matter what element of S we are given.

When we work with functions, we name them. The most common name is f. To make it clear—or serve as a reminder—that f names a function from S to T, mathematicians commonly write

$$f : S \to T.$$

In words, these symbols say, "f is a function from S to T." If $s \in S$ is given, then the element of T that f tells us to pick is written $f(s)$.

If we have reason to discuss several functions, we use additional names such as g, h, *etc.* For particular special purposes, other names are common, *e.g.*, F, T, or ϕ (Greek letter "phi"). On different occasions the same name is often used to name different functions. There is no problem, provided the context is clear. On the other hand, there are some functions that are so commonly used that they have standard names—for example: $\ln : \mathbb{R} \to \mathbb{R}$ (the natural logarithm), $\sin : \mathbb{R} \to \mathbb{R}$ (the sine function from trigonometry).

The set that serves as input—S in the illustration above—is called the *domain of f*. The set of possible outputs—T in the illustration—is called the codomain. The codomain is not the same thing as the range; the *range* of $f : S \to T$ is $\{ f(s) \mid s \in S \}$, which is a subset of T, but may not include all elements of T. The word "codomain" is a more recent addition to the standard mathematical vocabulary than the other words. We need the idea of a codomain to distinguish surjective functions (see below) from functions that are not surjective and, as we shall see, there are several important ideas that depend on our ability to make this distinction.

Example. Let H be the set of all human beings born of woman and let W be the set of all women. (Those who accept the Bible story of creation will believe there were only two human beings not in H, namely Adam and Eve. Other people will have other opinions about membership in H.) Let m be the rule that, given a person in H, points to the woman who bore that person, *i.e.*, that person's mother. Then, to each element of H, m makes a unique assignment of an element of W. Therefore, m is a function from H to W:

$$m : H \to W; \; m(h) \text{ is the mother of person } h.$$

On the other hand, the rule that selects a woman's children is not a function from W to H. There are two problems: 1) some women have no children and 2) some women have several. So selecting children does not create an assignment of a *unique* element of H to *every* element of W.

If the function m seems natural to you, then you may skip what I have to say in this paragraph. If on the other hand you find m troubling, then perhaps it is for the following reason. One usually thinks of mothers as coming prior to their children and as playing a very important role in their creation. So, going from child to mother, as the function m does, may seem unnatural. In addition to this, it is not unusual for people to think of the domain of a function as something that comes prior to the output. Sometimes functions are even spoken of as if they "create" their outputs from their inputs. This is another reason why one might be tempted to think of the set of mothers—rather than the set of children—as the domain for a function that expresses motherhood. Be this as it may, child-to-mother (not mother-to-child) is the direction we need to go in order to be able to view the relationship between mothers and children as a function. The reason, as we have seen, is that a woman may have several children—or none. Thus, given a woman, we do not have a unique way to pass to a child. On the other hand, each human born of woman has exactly one mother, so the passage in the direction child-to-mother is "functional". Perhaps the lesson to take away from all this is that mathematical ideas sometimes force us to go against the grain of some of our natural instincts. This may be one of the reasons why math is hard to learn, and it may help to explain why some people are uncomfortable with it.

Example. One of the most important examples of a function is a coordinate system. As we have just described, a coordinate system on the line is a rule for selecting a number when a point on the line is chosen. The rule is completely determined when the points O and P referred to above are chosen, provided of course that you know how to measure. (The measurement process itself involves a lot of ideas, but right now we shall assume that readers are clear enough about measuring.) Given a line ℓ and points O and P on ℓ, measurement assigns a number to each point. We call that rule $\overline{x}$. Thus, in accordance with our conventions for functions, we write

$$\overline{x} : \ell \to \mathbb{R}.$$

When we read this, we understand it to be a reminder that $\overline{x}$ is a rule for passing from points on ℓ to numbers. We need to remember that the rule depends on O and P. A different initial choice of O and P would have resulted in a different coordinate function.

Injective functions. Suppose S and T are sets and $f : S \to T$. We say f is *injective* if $f(s_1) = f(s_2)$ implies $s_1 = s_2$. In other words, f is injective when it picks different elements of T whenever different elements of S are supplied as inputs.

Example. The function $m : H \to W$ which we described above is *not* injective. My brother, John F. Madden makes this impossible, since although

$$m(\text{James J. Madden}) = m(\text{John F. Madden}),$$

it is not the case that

$$\text{James J. Madden} = \text{John F. Madden}.$$

My brother Thomas E. Madden is a further reason why the mother function m is not injective. We all have the same mother, Margaret R. Madden. (Of course, pointing to any pair of people with the same mother would show that m is not injective. I chose the examples dearest to me.)

Example. Assume O and P on line ℓ have been chosen. The corresponding coordinate function $\overline{x} : \ell \to \mathbb{R}$ is injective, since it takes different points to different numbers. Indeed, if $\overline{x}(A) = \overline{x}(B)$, then A and B are on the same side of O and are the same distance form O, so they are the same point: $A = B$.

Surjective functions. Suppose S and T are sets and $f : S \to T$. We say f is *surjective* if: for every element t in the set T, there is s in S such that $f(s) = t$. In other words, f is surjective if every element of the codomain T "gets hit" when we feed f all the elements of the domain. Yet another way to say this: the range of f is the whole codomain.

Example. The function $m : H \to W$ which we described above is *not* surjective. The problem is, there are women without children. For example, Queen Elizabeth was childless, so there is no h such that $m(h) =$ Queen Elizabeth. Queen Elizabeth is in the codomain of m (the set of all women) but she is not in the range of m (the set of all mothers).

Example. The coordinate function $\overline{x} : \ell \to \mathbb{R}$ is surjective, since at any given oriented distance from O there is a point. Thus, if any number a is given, one can find a point A on ℓ such that $\overline{x}(A) = a$.

Bijective functions. A function that is both surjective and injective is said to be bijective.

Example. Let G be the set of girls at a high-school dance and let B be the set of boys. If each girl picks a boy (only one!) with whom to dance, no two girls pick the same boy and every boy is picked by some girl, then we have a bijection $f : G \to B$. Here $f(x)$ denotes the boy who was picked by Ms. x.

The inverse of a function. Suppose $f : S \to T$ is a function and $g : T \to S$ is a function. We say that g is *an inverse of* f if the following conditions are satisfied:

i) for any $s \in S$, $g(f(s)) = s$, and

ii) for any $t \in T$, $f(g(t)) = t$.

Example. Suppose the girls choose boys bijectively. Then for each boy, there is exactly one girl who chooses him. This gives a function $g : B \to G$, where for a given boy y, $g(y)$ is the girl who chooses him. A moment of thought about the meanings shows that $g(f(x)) = x$ and $f(g(y)) = y$.

Lemma. *A function $f : S \to T$ has an inverse if and only if it is bijective.*

Proof. If $f : S \to T$ is bijective, then for every $s \in S$, there is exactly one $t \in T$ such that $f(s) = t$. This means that we have a function from T to S that "undoes" the action of f, and if we call this function g, then it satisfies conditions *i*) and *ii*). Conversely, suppose $g : T \to S$ is an inverse of f. We show that f is injective. Suppose $f(s_1) = f(s_2)$. Then $s_1 = g(f(s_1)) = g(f(s_2)) = s_2$, where the first and third equalities follow from the fact that g is an inverse for f and the middle equality is true because g is a function. We show that f is surjective. Suppose $t \in T$. Then $f(g(t)) = t$. But $g(t) \in S$. Q.E.D.

Example. Fix points O and P on line ℓ to make a function $\overline{x} : \ell \to \mathbb{R}$. Since this function is bijective, it has an inverse. We will call the inverse $\overline{P} : \mathbb{R} \to \ell$. Given a number a, $\overline{P}(a)$ is the corresponding point. We have $\overline{x}(\overline{P}(a)) = a$ for any $a \in \mathbb{R}$ and $\overline{P}(\overline{x}(A)) = A$ for any point $A \in \ell$. (The letter "P" in the name of the function $\overline{P}$ is not a reference to the point P that I chose in order to create the coordinate system. It is just a coincidence that I used the letter twice for two different purposes. But observe that $\overline{P}(1) = P$ and $\overline{x}(P) = 1$.)

2.1. Problem. Consider the following kinship relationships: father-son, father-eldest son, grandfather-grandson, maternal grandfather-grandson, brother-brother, youngest brother-oldest brother, self-self. Associated with some of these relationships there is a function. If there is a function, describe its domain, codomain and the rule that determines it. Determine whether it is injective and whether it is surjective.

2.2. Problem. Let $f : S \to T$ be a function. The graph of f is the set of ordered pairs $\{ (s,t) \mid t = f(s) \}$. a) Suppose (s,t) and (s,t') are two elements of the graph of f with the same first entry. Explain why it must be the case that $t = t'$, using the conditions defining the concept of a function which were given in the first paragraph of this section. b) The "vertical line test" says that a subset W of the plane is the graph of a function if and only if no vertical line meets W in more than one point. Explain the relationship of the vertical line test to part a). In particular, suppose W passes the vertical line test. What is the domain of the function associated with W? What is its codomain?

3. Coordinates on a plane.

Given a plane, a coordinate system on it is made as follows:

1. First, make the following choices:
   a. choose two lines in the plane that are not parallel; call the point of intersection O; name one of the lines the "$\overline{x}$-axis" and name the other the "$\overline{y}$-axis";
   b. choose points P and Q, different from O on the $\overline{x}$- and $\overline{y}$-axes, respectively.
2. Make a one-dimensional coordinate system on each of the lines, using OP as the unit interval on the $\overline{x}$-axis and OQ as the unit interval on the $\overline{y}$-axis.
3. Given a point A in the plane, assign coordinates to it as follows: Draw the line through A parallel to the $\overline{y}$-axis. This line crosses the $\overline{x}$-axis. Determine the coordinate of this point in the one-dimensional system on the $\overline{x}$-axis. This is called the $\overline{x}$-coordinate of A and it will be denoted $\overline{x}(A)$. The $\overline{y}$-coordinate of A is determined in the same way, except that the roles of $\overline{x}$ and $\overline{y}$ are exchanged.

Remark. Note that the Parallel Postulate guarantees that there is *only one* way to draw the parallel lines used above.

Fact 2. *A coordinate system constructed as above on a plane* $\mathbf{P}$ *creates a bijection between the points in* $\mathbf{P}$ *and the set* $\mathbb{R}^2$ *of ordered pairs of real numbers.*

Proof. Let $c : \mathbf{P} \to \mathbb{R}^2$ be the coordinate function; in other words, for any point A in the plane,

$$c(A) := (\overline{x}(A), \overline{y}(A)).$$

We want to show that c is a bijection. The lemma in the previous section tells us that in order to do this, it suffices to produce an inverse for c. So, let $\overline{P} : \mathbb{R}^2 \to \mathbf{P}$ be the function defined as follows: if a and b are numbers, then draw the line parallel to the $\overline{y}$-axis through the point on the $\overline{x}$-axis that has coordinate a (in the one-dimensional system determined by O and P. Then draw the line parallel to the $\overline{x}$-axis through the point on the $\overline{y}$-axis that has coordinate b (in the one-dimensional system determined by O and Q. $\overline{P}((a,b))$ is the point where these lines intersect. We will show that this is the inverse of c. To do this, we must show:

i) $c(\overline{P}((a,b))) = (a,b)$ for all $(a,b) \in \mathbb{R}^2$, and
ii) $\overline{P}(c(A)) = A$ for all $A \in \mathbf{P}$.

Now, the lines drawn upon being given (a,b) in order to determine the point $\overline{P}((a,b))$ are the same lines that will be drawn when we look for the coordinates of $\overline{P}((a,b))$. This shows that the $\overline{x}$- and $\overline{y}$-coordinates of $\overline{P}((a,b))$ are a and b, respectively, and thus $i)$ is proved. Conversely, the lines drawn upon being given A in order to determine its coordinates are the same as the lines that will be drawn in order to determine the point $\overline{P}(c(A))$, and this shows that $ii)$ is true. Q.E.D.

4. Slope Between Points

Choose a plane **P** and in it choose a coordinate system. As above, the coordinate functions will be denoted $\overline{x}$ and $\overline{y}$ and the function that takes pairs of numbers to points will be denoted $\overline{P}$. This system will remain the same throughout the rest of this discussion (sections 4, 5 and 6).

Definition. *Suppose A and B are points and thy do not have the same $\overline{x}$-coordinate. Then we define the slope from A to B to be the number:*

$$m(A,B) := \frac{\overline{y}(B)-\overline{y}(A)}{\overline{x}(B)-\overline{x}(A)}.$$

4.1. Show: If $A = \overline{P}(x_1, y_1)$ and $B = \overline{P}(x_2, y_2)$, and $x_1 \neq x_2$, then $m(A,B) = \frac{y_2-y_1}{x_2-x_1}$.
4.2. Show: For any points A and B with different $\overline{x}$-coordinates, $m(A,B) = m(B,A)$.
4.3. Show: Let A, B and C be three points with different $\overline{x}$-coordinates. If $m(A,B) = m(B,C)$, then $m(A,B) = m(A,C)$. Indeed, if two of the numbers $m(A,B)$, $m(B,C)$ and $m(A,C)$ are equal, then all three are equal.

Our next task will be to prove the following.

Proposition. *Suppose A, B, A' and B' are any four points in the plane. Also suppose that A and B have different $\overline{x}$-coordinates and that A' and B' also have different $\overline{x}$-coordinates. Then $m(A,B) = m(A',B')$ if and only if line AB is parallel to line $A'B'$.*

To prove this proposition, we have two tasks:
a) Show: $AB \parallel A'B'$ implies $m(A,B) = m(A',B')$.
b) Show: $m(A,B) = m(A',B')$ implies $AB \parallel A'B'$.

To accomplish these tasks, we begin by treating a special and particularly easy case.
4.4. Show: a) If line AB is parallel to the $\overline{x}$-axis, then $m(A,B) = 0$. b) If $m(A,B) = 0$ then line AB is parallel to the $\overline{x}$-axis.

Problem 4.4 takes care of the special case of the proposition that arises when $m(A,B) = 0 = m(A',B')$ or when $AB \parallel A'B'$. Thus, from now on we only need to think about the cases where $m(A,B)$ is nonzero, or when AB is not parallel to the $\overline{x}$-axis.
4.5. Show: Assuming that AB is not parallel to the $\overline{x}$-axis, $AB \parallel A'B'$ implies $m(A,B) = m(A',B')$.
4.6. Show: Assuming that $m(A,B) > 0$, $m(A,B) = m(A',B')$ implies $AB \parallel A'B'$.
4.7. Show: Assuming that $m(A,B) < 0$, $m(A,B) = m(A',B')$ implies $AB \parallel A'B'$.
With 4.7, the proof of the proposition above is complete.
4.8. Show: If A, B and C are any three points with distinct $\overline{x}$-coordinates, then they all lie on the same line if and only if $m(A,B) = m(B,C)$.