

The Equation of a Line in the Plane

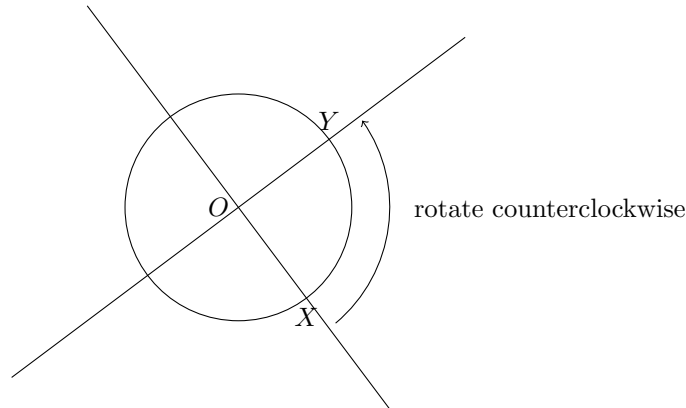
Let’s recall the idea of *the plane* from classical geometry: the plane is like a flat sheet of drawing paper, which extends indefinitely in all directions without bound. It is the playground for lots of serious considerations from high school: points, lines, circles, triangles, rectangles, and various other doodles live in it. I say *in* it rather than *on* it, because all of those objects really have their existence inside the plane. If we were to say “on the plane” then one might think of them as sitting on top of the paper, where a light breeze might move them about. No, those things live inside the plane just as sure as you and I live inside the universe.

And while all of that is evocative and romantic, it doesn’t make doing mathematics any easier. Our aim in this first chapter is to do some concrete mathematics: we want to figure out how to describe a single line in the plane very carefully. To do this, we will use tools that René Descartes taught us: coordinates. Better yet, we will use an update of the idea and introduce *vectors*. Our work has to rely on something, so at some points we will make use of geometry facts you learned in high school. But for all of the points and vectors, angles and dot products, we will go straight to the heart of a single important question:

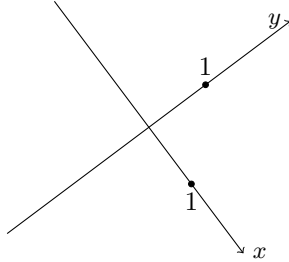
**How can we clearly describe a single
line in the plane with an equation?**

THE EQUATION OF A LINE IN THE PLANE POINTS, VECTORS, AND VECTORS

You have likely seen the idea of *Cartesian coordinates* on the plane before. To be clear, let’s set things down carefully. In the plane, we choose a pair of perpendicular lines which meet at a point O . This special point is called the *origin*. Then, we choose a point X on one of the lines and draw the circle centered at O which passes through X . Note that this circle meets our two lines in two points each, four points total, one of which is the point X . Then, from X , we rotate around the circle by a quarter turn counterclockwise until we hit one of the points on the other line. This new point we will call Y . Are you drawing with me? Here is my picture so far.



We call the line OX the *x-axis* and the line containing Y the *y-axis*. Here comes the amazing part: we declare the circle we used to be of *unit size*, and make the lines OX and OY into number lines! The important part is that the point O should represent 0 on both number lines, and the points X and Y should each represent 1 on their lines. So, instead of marking things with O , X , and Y , we put down marks where X and Y are and label them with 1’s, and add little arrows marked with x and y near the positive “ends” of the lines OX and OY , respectively.



Note that above I have done something a bit silly and let the picture just fall on the paper in an unusual way. I really mean unusual as “not usual.” The usual way arranges the lines on the paper to match our expected horizontal (x) and vertical (y) directions. This isn’t actually required, but it is what everyone always does. So, the more typical picture looks more like this one.

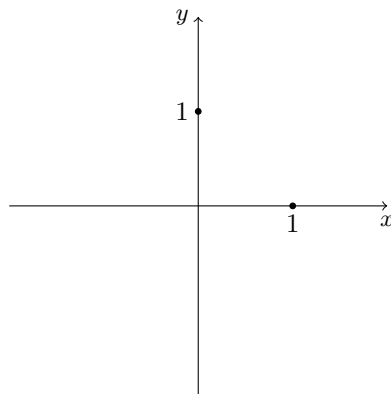


Figure 1: The Standard Cartesian Coordinate System

Now suppose we have some point in the plane, let’s call it P . We can describe the location of P relative to our two lines in a simple way. First, we draw a line through P which is parallel to the y -axis and perpendicular to the x -axis. The foot of this

THE EQUATION OF A LINE IN THE PLANE

perpendicular hits the x -axis at some point A . But this point P' is part of the number line OX , so it has an associated real number, which we will call a . So the point A is instead marked with the label a from this number line.

Similarly, we draw a line through P parallel to the x -axis and perpendicular to the y -axis. The foot of this perpendicular hits the y -axis at some point B , which is part of the number line OY . We denote the number associated to B by b . Again, the point is labeled with the number b from the number line.

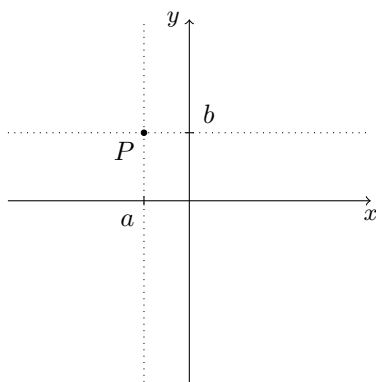


Figure 2: A point P and its coordinates (a, b) .

So, to identify the point P , we can instead give the pair of numbers a and b . These numbers are called the *coordinates* of P . Of course, the order of the coordinates matters, so we make what we call an *ordered pair* (a, b) to keep things straight, where the x -coordinate comes first, and the y -coordinate comes second. Note that in Figure 2 the x -coordinate a is negative, but the y -coordinate b is positive.

This whole process is reversible, too. If we pick a pair of numbers c and d , in order, then we can find a point in the plane which corresponds, and we can do it unambiguously. Find the spot labeled c on the x -axis number line and construct a line perpendicular to the x -axis through this point. Similarly, find the

spot labeled d on the y -axis and construct a line perpendicular to the y -axis through this point. The two lines you just drew will meet in exactly one point Q , and Q will have coordinates (c, d) .

This setup of coordinates on the plane allows us to formalize a wonderful and useful idea from physics, too. Physicists use the concept of a *vector* to describe something (like the wind) which has both magnitude or size (like how fast the air is moving) and direction (which way the air is going). Usually, vectors are drawn as little arrows: the arrow has a direction, and it has a length which represents its magnitude. It is possible to draw vectors which have the same direction but different lengths, and vice versa.

We can use coordinates to describe vectors in the plane, too. Here's how: A physicist's vector V is some arrow in the plane. That arrow has an initial point P and a final point Q . We can write the coordinates of these points as $P = (a, b)$ and $Q = (c, d)$.

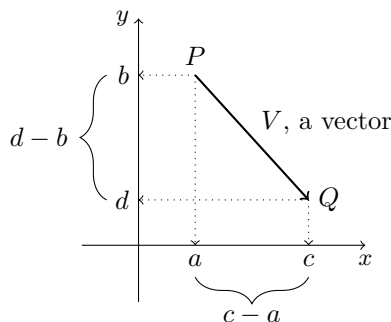


Figure 3: Coordinates for a physicist's vector V .

Then the coordinates of V are taken to be the numbers $c - a$ and $d - b$, which we interpret how much V acts in the directions parallel to the x -axis and the y -axis, respectively. Note that in Figure 3, the y -coordinate is negative, since $b > d$.

These coordinates have a hint of algebraic manipulation in

THE EQUATION OF A LINE IN THE PLANE

them. Those subtractions line up almost like we could write $V = (c - a, d - b) = (c, d) - (a, b) = Q - P$. But V is a vector, and if we write it like $(c - a, d - b)$, it looks like the notation for a point. We should not do that because it could get confusing. Furthermore, that "equation" would mean that we are subtracting points and creating a vector, which is also weird. Still, there is something there. We will return to this idea soon.

For now, let's focus on a bit of ambiguity in the physicist's idea of a vector. Where should that vector be? That is, given the coordinates of a vector in the plane, it is not clear where to draw it! I can slide a vector around the plane, and as long as I keep it parallel to the original, the coordinates won't change. So, unlike with the coordinates of a point, the coordinates of a vector do not uniquely specify the vector.

The mathematician's special fix is this: we simply declare all our vectors to have their initial points, their tails, at the origin O , of our coordinate system. That curtails some of the (admittedly useful) freedom in the physics notion, but it also lets us be more clear.

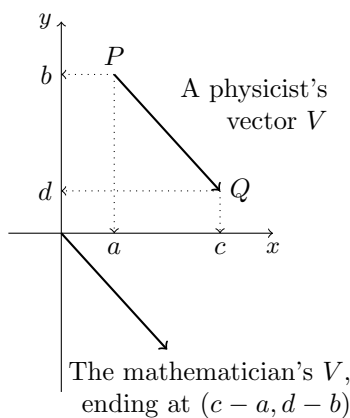
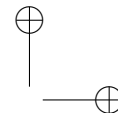
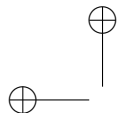


Figure 4: The physicist's vector vs. the mathematician's vector.

But it pays to keep in mind that the physicists conception of



the vector V with coordinates $c - a$ and $d - b$ could be one of many different arrows, while the mathematician’s vector V is the arrow from the point $O = (0, 0)$ to the point $(c - a, d - b)$.

Now we have circled back around to a muddle. If a mathematician’s vector is always based at O , we only need to specify where the head of the vector is... which is just a point. So, how is a vector supposed to be different from a point, again?

This confusion of three different, shifting, partially-overlapping interpretations causes some trouble to the new learner. Professionals tend to pass back and forth between these and use them flexibly to get results. Once you have gotten used to the ideas, you will, too. You should watch out for these instances where the words point and vector get interchanged. If they cause you trouble, remember that we have three different things, which are closely related.

For now, the simplest way to handle things is like this:

- A point is a location in the plane, and represented by coordinates in the form of an ordered pair of numbers (a, b) .
- Ignore the physicist’s version of the word vector as much as possible.
- A vector is an arrow based at the origin, which can be specified by the coordinates of its head. To keep this separate from the idea of a point, we will write it differently, with the numbers stacked vertically like this: $\begin{pmatrix} a \\ b \end{pmatrix}$.

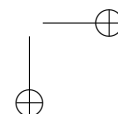
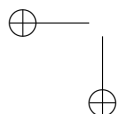
With this in mind, we make our first official definition.

1 Definition. *A 2-vector is a vertical stack of 2 real numbers, like so:*

$$v = \begin{pmatrix} a \\ b \end{pmatrix}.$$

The collection of all such 2-vectors is called the plane, and written with this notation: \mathbb{R}^2 .

The notation \mathbb{R}^2 is often read “arr-two,” and many people use that language interchangeably with “the plane.”



THE EQUATION OF A LINE IN THE PLANE

VECTOR ALGEBRA

Let’s return to that glimpse of subtraction we saw in Figure 3. We saw there that for points $P = (a, b)$ and $Q = (c, d)$, the vector V from P to Q has coordinates $c - a$ and $d - b$. This looks almost like we subtracted the points to get $Q - P = V$. Can we use that? The weird part is that it mixes up points and vectors. So, we will just change viewpoints, and instead think of P and Q as (mathematician’s) vectors:

$$P = \begin{pmatrix} a \\ b \end{pmatrix}, \quad Q = \begin{pmatrix} c \\ d \end{pmatrix}$$

If we put these together on the plane with the physicist’s vector V from P to Q and the mathematician’s V we see a wonderful triangle, and an extra vector. So we see a way to talk about

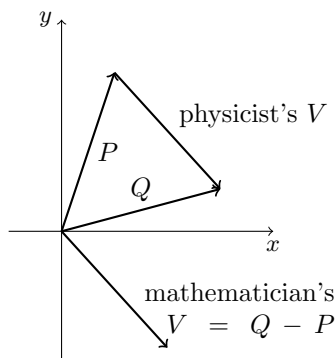
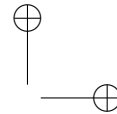
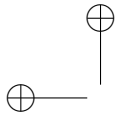


Figure 5: Subtraction of vectors

subtracting vectors: Given two vectors P and Q as above, their *difference* is the vector

$$V = Q - P = \begin{pmatrix} c \\ d \end{pmatrix} - \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} c - a \\ d - b \end{pmatrix}.$$



Geometrically, we draw the arrow from P to Q and then translate it down so that its tail is at the origin $O = (0, 0)$. Of course, the order of P and Q in this operation matters. If we switch them, we get an arrow pointing in the opposite direction.

If we can subtract vectors, surely we can add vectors. How would that work? Algebraically, if $V = Q - P$, then we expect $Q = V + P$ by rearranging. That would mean

$$Q = \begin{pmatrix} c \\ d \end{pmatrix} = V + P = \begin{pmatrix} c - a \\ d - b \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix},$$

which all fits. It looks like addition should be defined coordinate-by-coordinate.

2 Definition (Addition of Vectors). *Let $P = \begin{pmatrix} a \\ b \end{pmatrix}$ and $Q = \begin{pmatrix} c \\ d \end{pmatrix}$ be two vectors in \mathbb{R}^2 . Their sum is the vector*

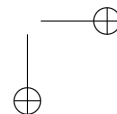
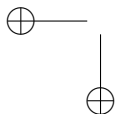
$$P + Q = \begin{pmatrix} a + c \\ b + d \end{pmatrix}.$$

3 Theorem. *Addition of vectors satisfies the same rules as addition of real numbers:*

- *when adding more than two vectors, it doesn't matter which operations you do first: $(P + Q) + R = P + (Q + R)$;*
- *one can add in either order $P + Q = Q + P$;*
- *the origin O is a “zero” since $P + O = O + P = P$;*
- *for each vector P , there is an opposite vector $-P$ so that $P + (-P) = O$.*

Remember that you are supposed to read actively. You can draw all of these pictures and try out all of these things with specific examples that you invent. You should check these by making examples and working out the details. Can you also draw the pictures which go with your examples?

But what about subtracting geometrically? In Figure 5, I have a strong desire to complete the figure by joining the loose end of V to the rest of the figure. If we draw the arrow from the head of V to the head of Q , we get this:



THE EQUATION OF A LINE IN THE PLANE

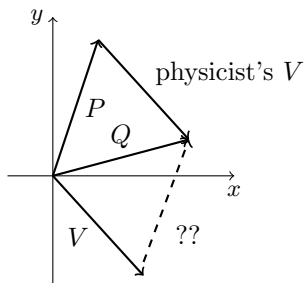


Figure 6: Subtraction of vectors

What should the label on the dashed vector in Figure 6 be? Just as the physicist’s V and the mathematician’s V are parallel, this new vector is parallel to the mathematician’s vector P . So the dashed vector must be a physicist’s version of P . Then we see $Q = V + P$.

Now we know how to add geometrically: to add two vectors A and B , translate B until its tail is on the head of A , then draw a new vector $A + B$ as the vector from the tail of A to the head of this translated B . It just repurposes the structure of Figure 6.

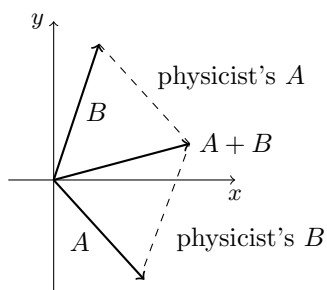


Figure 7: Geometric Addition of Vectors

This is sometimes called the *parallelogram rule* for addition of vectors.

There is another useful operation on vectors called *scalar multiplication*. This corresponds to stretching (or shrinking) a vector, without changing its direction. The terminology comes from physics (again) where a *scalar* quantity is just a number, and not a vector. So “scalar multiplication” means to multiply a vector by a scalar.

4 Definition (Scalar Multiplication for vectors). *Let $P = \begin{pmatrix} a \\ b \end{pmatrix}$ be a vector in \mathbb{R}^2 , and let λ be a real number. Then the scalar multiple λP is defined to be*

$$\lambda P = \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix}$$

If you have never seen the symbol λ before, it is an old Greek letter pronounced “lamb-duh.” It is traditional to use it in linear algebra in lots of places. Welcome to the λ -club.

Again, this operation has some important similarities to the familiar multiplication of numbers, but because it combines a scalar (a number) with a vector (not a number, exactly) to produce another vector (again, not a number) things are a little different.

5 Theorem. *Suppose that P and Q are vectors, and λ and μ are numbers. Scalar multiplication has the following properties:*

- *Scalar multiplication distributes over addition: $\lambda(P + Q) = \lambda P + \lambda Q$;*
- *Scalar multiplication and regular multiplication can be done in either order: $\lambda(\mu P) = (\lambda\mu)P$;*
- *if $\lambda = 0$, then $\lambda P = 0P = O$ is the zero vector O .*

LINES AS PARAMETRIC OBJECTS

- lines in \mathbb{R}^2 as parametric objects: through the origin, and not
- sketching lines in \mathbb{R}^2

LENGTHS AND ANGLES IN THE PLANE

- lengths, angles, and the dot product in \mathbb{R}^2

THE EQUATION OF A LINE IN THE PLANE

- normal vector to a line in \mathbb{R}^2
- duality
- points vs physics vectors vs math vectors
- the idea of \mathbb{R}^2
- sketching points and vectors in the plane

LINES AND EQUATIONS

- equation of a line, three methods: elimination, from two pts via similar triangles, from geometry of dot product
- families of parallel lines
- sketching a line from an equation

CONCLUSION: THE BIG THEOREM