The Dot Product

We have seen that points can be thought of as lists of real numbers, and that we can compute/define the distance between two points via the Pythagorean Theorem. But geometry is more than just distances; it also involves angles. The most elegant way to express angles in Cartesian coordinates is via the "dot product".

In this section we will always write \mathbf{x} to refer to the vector $[\mathbf{0}, \mathbf{x}]$. The benefit of this is to greatly simplify the notation. The risk is that we might get confused. Whenever the notation \mathbf{x} appears you will have to figure out from context whether it is supposed to be a point or a vector.

• Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$. Recall that the vectors \mathbf{x} , \mathbf{y} , and $\mathbf{y} - \mathbf{x}$ can be thought of as the edges of a triangle. Now suppose that the vectors \mathbf{x} and \mathbf{y} are perpendicular. What does the Pythagorean Theorem say in this case?

The Pythagorean Theorem says that

$$\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{y} - \mathbf{x}\|^2.$$

On the other hand, if we try to compute $\|\mathbf{y}-\mathbf{x}\|^2$ we get

$$\begin{aligned} \|\mathbf{y} - \mathbf{x}\|^2 &= \|(y_1 - x_1, y_2 - x_2, \dots, y_n - x_x)\|^2 \\ &= (y_1 - x_1)^2 + (y_2 - x_2)^2 + \dots + (y_n - x_n)^2 \\ &= (y_1^2 - 2x_1y_1 + x_1^2) + (y_2^2 - 2x_2y_2 + x_2^2) + \dots + (y_n^2 - 2x_ny_n + x_n^2) \\ &= (x_1^2 + x_2^2 + \dots + x_n^2) + (y_1^2 + y_2^2 + \dots + y_n^2) - 2(x_1y_1 + x_2y_2 + \dots + x_ny_n) \\ &= \|\mathbf{x}\|^2 + \|y\|^2 - 2(x_1y_1 + x_2y_2 + \dots + x_ny_n). \end{aligned}$$

That's strange. We were supposed to get $||x||^2 + ||y||^2$. It must be that the extra term is zero. If the vectors **x** and **y** are perpendicular, then we find that

$$2(x_1y_1 + x_2y_2 + \dots + x_ny_n) = 0$$

$$x_1y_1 + x_2y_2 + \dots + x_ny_n = 0.$$

Congratulations. We have just discovered the dot product.

Definition: Given two vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$ we define their dot product as follows:

$$\mathbf{x} \bullet \mathbf{y} := x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Thus the dot product of two vectors is just a number. At first this definition might seem a bit strange, but we have already proved a theorem about it. If the vectors \mathbf{x} and \mathbf{y} are perpendicular, then their dot product is zero:

$$\mathbf{x} \perp \mathbf{y} \Longrightarrow \mathbf{x} \bullet \mathbf{y} = 0.$$

• Find a formula to express the length of a vector $\mathbf{u} \in \mathbb{R}^n$ in terms of the dot product.

Suppose that $\mathbf{u} = (u_1, u_2, \dots, u_n)$ and recall that the length of the vector is given by

$$\|\mathbf{u}\|^2 = u_1^2 + u_2^2 + \dots + u_n^2.$$

Can we express the right hand side of that equality using the dot product? Yes we can:

$$||u||^2 = \mathbf{u} \bullet \mathbf{u}.$$

• Given vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and a "scalar" $\alpha \in \mathbb{R}$, show that

$$\mathbf{u} \bullet (\mathbf{v} + \alpha \mathbf{w}) = (\mathbf{u} \bullet \mathbf{v}) + \alpha (\mathbf{u} \bullet \mathbf{w}).$$

Hence the dot product behaves very much like a "multiplication" of vectors. This formula will allow us to do future computations in a more mindless way.

OK, let's do it.

$$\mathbf{u} \bullet (\mathbf{v} + \alpha \mathbf{w}) = (u_1, \dots, u_n) \bullet ((v_1, \dots, v_n) + (\alpha w_1, \dots, \alpha w_n))$$

= $(u_1, \dots, u_n) \bullet (v_1 + \alpha w_1, \dots, v_n + \alpha w_n)$
= $u_1(v_1 + \alpha w_1) + \dots + u_n(v_n + \alpha w_n)$
= $(u_1v_1 + \alpha u_1w_1) + \dots + (u_nv_n + \alpha u_nw_n)$
= $(u_1v_1 + \dots + u_nv_n) + \alpha(u_1w_1 + \dots + u_nw_n)$
= $(\mathbf{u} \bullet \mathbf{v}) + \alpha(\mathbf{u} \bullet \mathbf{w}).$

• Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ be vectors and consider again the triangle with sides \mathbf{x}, \mathbf{y} , and $\mathbf{y} - \mathbf{x}$. Use your dot product formulas and mindless computations to show that

$$\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2(\mathbf{x} \bullet \mathbf{y}).$$

Yes, we already did this. But now we can do it more quickly:

$$\|\mathbf{y} - \mathbf{x}\|^{2} = (\mathbf{y} - \mathbf{x}) \bullet (\mathbf{y} - \mathbf{x})$$

= $(\mathbf{y} - \mathbf{x}) \bullet \mathbf{y} - (\mathbf{y} - \mathbf{x}) \bullet \mathbf{x}$
= $\mathbf{y} \bullet \mathbf{y} - \mathbf{x} \bullet \mathbf{y} - (\mathbf{y} \bullet \mathbf{x} - \mathbf{x} \bullet \mathbf{x})$
= $\mathbf{y} \bullet \mathbf{y} - \mathbf{x} \bullet \mathbf{y} - \mathbf{y} \bullet \mathbf{x} + \mathbf{x} \bullet \mathbf{x}$
= $\mathbf{x} \bullet \mathbf{x} + \mathbf{y} \bullet \mathbf{y} - 2(\mathbf{x} \bullet \mathbf{y})$
= $\|\mathbf{x}\|^{2} + \|\mathbf{y}\|^{2} - 2(\mathbf{x} \bullet \mathbf{y}).$

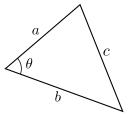
In this proof we used the fact that $\mathbf{x} \bullet \mathbf{y} = \mathbf{y} \bullet \mathbf{x}$. Why is this true?

• On the other hand, if θ is the angle between the vectors **x** and **y**, show that we have

$$\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta.$$

[Hint: Look up the Law of Cosines.]

Consider a triangle with sides of length a, b, c and let θ be the angle between the sides of length a, b.

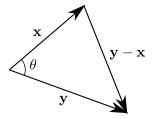


The Law of Cosines then tells us that

$$c^2 = a^2 + b^2 - 2ab\cos\theta.$$

I'm going to ask you for a proof of this later, so I won't spoil the fun by proving it here. Note that the Law of Cosines is a generalization of the Pythagorean Theorem because we have $\cos \theta = 0$ if and only if θ is (plus or minus) a right angle.

Now consider the triangle formed by the vectors \mathbf{x} , \mathbf{y} , $\mathbf{y} - \mathbf{x}$, and let θ be the angle between the vectors \mathbf{x} and \mathbf{y} .



In this case the Law of Cosines says exactly that

$$\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta.$$

• Now you have two different expressions for $\|\mathbf{y} - \mathbf{x}\|^2$. What happens if you compare them?

We get

$$\|\mathbf{x}\|^{2} + \|\mathbf{y}\|^{2} - 2(\mathbf{x} \bullet \mathbf{y}) = \|\mathbf{x}\|^{2} + \|\mathbf{y}\|^{2} - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta$$
$$-2(\mathbf{x} \bullet \mathbf{y}) = -2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta$$
$$\mathbf{x} \bullet \mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|\cos\theta.$$

That looks nice, but what does it mean? It means that the dot product, which we originally defined with a somewhat random-looking algebraic formula, has a very natural geometric interpretation. I'll put a box around this for posterity.

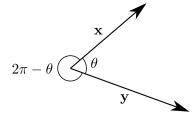
$$\mathbf{x} \bullet \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$$

This formula is super important and it deserves some discussion, so that's what we'll do now.

Discussion:

• Actually there are **two** angles between **x** and **y**. Does this matter?

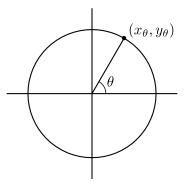
Given two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we can't really refer to **the** angle between them because there are two. We can call the angles θ and $2\pi - \theta$:



However, this doesn't really matter because the dot product can't tell the difference between these two angles. What I mean by this is that the cosines are the same:

$$\cos(\theta) = \cos(2\pi - \theta).$$

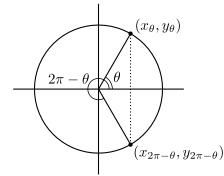
Why is this? Well, we must first remember the **definition** of cosine. You might think that cosine is defined in terms of triangles (adjacent/hypotenuse), but that's not a very good definition because it doesn't work for angles larger than 180° . (Have you ever seen a triangle with an angle larger than 180° ? Me neither.) The modern definition of cosine is based on a **circle**, not a triangle. Consider the unit circle in the Cartesian plane and let (x_{θ}, y_{θ}) be the point on the circle at angle θ measured counterclockwise from the x-axis.



Note that the numbers x_{θ} and y_{θ} are functions of the angle θ . Also note that they are **periodic** functions with period 2π . Wouldn't it be nice to have names for these functions? Yes. We will call them "cosine" and "sine":

$$\cos \theta := x_{\theta},\\ \sin \theta := y_{\theta}.$$

Now we can explain why $\cos(\theta) = \cos(2\pi - \theta)$. It happens because the points (x_{θ}, y_{θ}) and $(x_{2\pi-\theta}, y_{2\pi-\theta})$ are reflections of each other across the *x*-axis, so they have the same *x*-coordinate:



In the future we might **define** the angle between the vectors \mathbf{x} and \mathbf{y} via the equation

$$\theta = \arccos\left(\frac{\mathbf{x} \bullet \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}\right) = \arccos\left(\frac{\mathbf{x} \bullet \mathbf{y}}{\sqrt{\mathbf{x} \bullet \mathbf{x}}\sqrt{\mathbf{y} \bullet \mathbf{y}}}\right).$$

(If you don't like this definition, I'll ask: how would **you** define the angle between two vectors?) Of course this equation doesn't define a unique angle (it defines two angles), but the above discussion shows that this is not a problem.

• We will say that the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are **perpendicular** (and we will write $\mathbf{x} \perp \mathbf{y}$) if the two angles between them are 90° and 270°. Show that $\mathbf{x} \perp \mathbf{y}$ if and only if $\mathbf{x} \bullet \mathbf{y} = 0$.

Assume that neither of \mathbf{x} or \mathbf{y} is the zero vector, otherwise there really isn't an angle between them. Then $\|\mathbf{x}\| \|\mathbf{y}\| \cos \theta = 0$ if and only if $\cos \theta = 0$, and this happens if and only if θ is 90° or 270°. Since $\mathbf{x} \bullet \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$ we conclude that

$$\mathbf{x} \perp \mathbf{y} \Longleftrightarrow \mathbf{x} \bullet \mathbf{y} = 0.$$

You can view this as a grown-up version of the Pythagorean Theorem.

• Interpret the equation $\mathbf{x} \bullet \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$ when $\mathbf{x} = (x)$ and $\mathbf{y} = (y)$ are in \mathbb{R}^1 .

This should be easy right? Vectors in \mathbb{R}^1 are pretty much the same as numbers. First note that the dot product is just multiplication:

$$\mathbf{x} \bullet \mathbf{y} = xy.$$

Now what about the other side of the equation? What is the length of the vector $\mathbf{x} = (x)$? Using the formula $\|\mathbf{x}\|^2 = \mathbf{x} \bullet \mathbf{x}$ gives

$$\|\mathbf{x}\| = +\sqrt{\mathbf{x} \bullet \mathbf{x}} = +\sqrt{x^2}.$$

Do you recognize this? You might be tempted to say $+\sqrt{x^2} = x$, but that is not correct. If x is a **negative** number then we can't have $x = +\sqrt{x^2}$, because $+\sqrt{x^2}$ is defined to be a **positive** number. Instead we have a piecewise-defined function

$$+\sqrt{x^2} = \begin{cases} x & \text{if } x \ge 0, \\ -x & \text{if } x < 0. \end{cases}$$

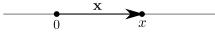
Do you recognize this? This is just the definition of the **absolute value** of a number:

$$\|\mathbf{x}\| = +\sqrt{x^2} = |x|.$$

I guess that explains the notation we use for the length of a vector. OK, so now we have the equation

$$\mathbf{x} \bullet \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$$
$$xy = |x||y| \cos \theta.$$

What sense can you make out of this? Note that x and y are just real numbers. What does it mean to talk about the **angle** between real numbers? Aha! We have to remember that $\mathbf{x} = (x)$ is not just a number but an arrow pointing from the number 0 to the number x:

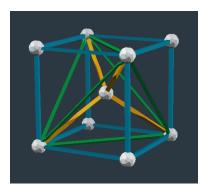


The vector points to the right if x is positive and points to the left if x is negative. Thus we have $\theta = 0$ when x and y have the same sign, and $\theta = 180^{\circ}$ when x and y have opposite signs. Finally, the equation $xy = |x||y| \cos \theta$ can be translated as

$$xy = \begin{cases} |x||y| & \text{if } x \text{ and } y \text{ have the same sign,} \\ -|x||y| & \text{if } x \text{ and } y \text{ have different signs.} \end{cases}$$

This is true isn't it? You learned long ago that "negative times positive is negative" and "negative times negative is positive". (Did you ever wonder **why**?) Now we see that these facts can be explained in terms of vectors and dot products.

• A molecule of methane is formed by four hydrogen atoms surrounding a central carbon atom. Let θ be the angle between any two of the hydrogen atoms. Use dot products to compute θ . [Hint: Look at the Zome model.] For reasons of symmetry, the hydrogen atoms will lie at the four vertices of a regular tetrahedron with the carbon atom at the center. We can build this tetrahedron out of Zometools using green struts. Then we can connect the center to the four vertices using yellow struts. Our goal is to compute the angle between two yellow struts.



But how can we do this? We need to choose a suitable coordinate system. I have a recommendation: Let the center of the tetrahedron be (0,0,0) and consider the four vertices of the tetrahedron as half of the vertices of a cube. We can build this cube with blue struts. So, what are the coordinates of the vertices of the cube? There is an obvious choice:

$$\begin{array}{ccc} (1,1,1) \\ (1,1,-1) & (1,-1,1) & (-1,1,1) \\ (1,-1,-1) & (-1,1,-1) & (-1,-1,1) \\ & (-1,-1,-1) \end{array}$$

If we suppose that the vertex closest to us in the picture is (1, 1, 1) then the four vertices of our green tetrahedron are

$$(1,1,1), (1,-1,-1), (-1,1,-1), (-1,-1,1).$$

The angle between any two of these vectors is the same, so we can choose any two we want. Let's compute the angle between $\mathbf{x} = (1, 1, 1)$ and $\mathbf{y} = (1, -1, -1)$. Our favorite formula says

$$\mathbf{x} \bullet \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$$

(1,1,1) • (1,-1,-1) = $\|(1,1,1)\| \|(1,-1,-1)\| \cos \theta$
1 · 1 + 1(-1) + 1(-1) = $\sqrt{1^2 + 1^2 + 1^2} \sqrt{1^2 + (-1)^2 + (-1)^2} \cos \theta$
1 - 1 - 1 = $\sqrt{3}\sqrt{3} \cos \theta$
-1 = 3 cos θ .

Therefore the angle is $\theta = \arccos(-1/3)$. This evaluates to $109.4712207\cdots$ degrees, but most chemistry books will lie to you and just say 109.5° .

• Find the equation of the **line** in \mathbb{R}^2 that contains the point (0,0) and is perpendicular to the vector $\mathbf{n} = (a, b)$. We call \mathbf{n} the "normal vector" of the line. When are two such lines perpendicular?

Let (x, y) be any point in the Cartesian plane \mathbb{R}^2 and observe that the **point** (x, y) is on the line if and only if the **vector** (x, y) is orthogonal to the vector $\mathbf{n} = (a, b)$. By the above discussion we know that this happens if and only if

$$(a,b) \bullet (x,y) = 0$$
$$ax + by = 0.$$

so this is the equation of the line. This confirms what we already knew.

When are the lines ax + by = 0 and a'x + b'y = 0 perpendicular to each other? This is the same as asking when the **vectors** (a, b) and (a', b') are perpendicular to each other, and we know that this happens if and only if

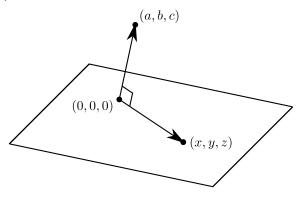
$$(a,b) \bullet (a',b') = 0$$
$$aa' + bb' = 0.$$

Again, this just confirms something we already know. But last time the answer was mysterious and this time it makes perfect sense in terms of the dot product.

• Find the equation of the **plane** in \mathbb{R}^3 that contains the point (0, 0, 0) and is perpendicular to the vector $\mathbf{n} = (a, b, c)$.

Remember when we tried to do this and failed? Well, now we have the appropriate technology. The first thing is to realize that a plane in \mathbb{R}^3 containing the point (0, 0, 0) is uniquely determined by a normal vector. Once we're realized this, the rest is easy.

Note that then **point** (x, y, z) is contained in the plane if and only if the **vector** (x, y, z) is perpendicular to (a, b, c).



Thus the point (x, y, z) is in the plane if and only if

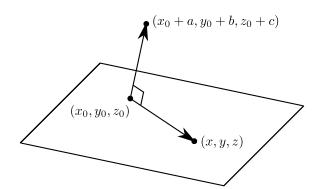
$$a, b, c) \bullet (x, y, z) = 0$$

 $ax + by + cz = 0.$

This is the equation of the plane. You see that the equation of a plane is just screaming the dot product at us because it is the correct tool for the job.

• Find the equation of the plane in \mathbb{R}^3 that contains the point $(x_0, y_0, z_0$ and is perpendicular to the vector $\mathbf{n} = (a, b, c)$.

We can reuse the diagram from the previous problem, but we have to relabel it.



Note that the **point** (x, y, z) is in the plane if and only if the **vector**

 $[(x_0, y_0, z_0), (x, y, z)] = [(0, 0, 0), (x - x_0, y - y_0, z - z_0)]$

is perpendicular to the vector

$$[(x_0, y_0, z_0), (x_0 + a, y_0 + b, z_0 + c)] = [(0, 0, 0), (a, b, c)].$$

And note that this happens if and only if

$$(a, b, c) \bullet (x - x_0, y - y_0, z - z_0) = 0$$

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

$$ax + by + cz - (ax_0 + by_0 + cz_0) = 0.$$

This is the equation of the plane. Thus the general equation of a plane in \mathbb{R}^3 is

ax + by + cz + d = 0,

where $a, b, c, d \in \mathbb{R}$ are any real numbers that determine the properties of the plane. For example, the plane will contain the point (0, 0, 0) if and only if d = 0.

• We know that a **line** in \mathbb{R}^3 cannot be expressed with a single equation. So how *can* we describe it? Consider the line that is the intersection of the two planes

$$x + y + z + 1 = 0,$$

$$x + 2y + 3z + 4 = 0.$$

Try to "solve" these two equations to come up with an explicit description of the line.

First we might solve each equation for x and then compare them to get

$$-y - z - 1 = -2y - 3z - 4$$
$$y + 2z + 3 = 0.$$

We say that we have **eliminated** the x variable. That's progress, right? Alternatively, we can eliminate y from the original two equations as follows:

$$-x - z - 1 = -\frac{1}{2}x - \frac{3}{2}z - 2$$
$$2x + 2z + 2 = x + 3z + 4$$
$$x - z - 2 = 0.$$

Or we can eliminate z:

$$-x - y - 1 = -\frac{1}{3}x - \frac{2}{3}y - \frac{4}{3}$$
$$3x + 3y + 3 = x + 2y + 4$$
$$2x + y - 1 = 0.$$

Unfortunately, this is the best we can do. We can't eliminate any more variables because we don't have any more equations to use. The three equations

$$2x + y + 0 = 0,$$

$$x + 0 - z = 2,$$

$$0 + y + 2z = -3$$

represent the "shadows" of our line when it is projected on the x, y-plane, x, z-plane, and y, z-plane, respectively.

What can we do next? Are we done? The trick is to admit defeat in a rather dramatic way by defining a **free parameter**. In this case we can take any of x, y, or z be a parameter so we will just define t := z and just admit that we will never the know the value of t (it is "free"). The benefit of this is that we can now completely solve the problem, in terms of t. We have

$$x = 2 + t$$
$$y = -3 - 2t$$
$$z = t.$$

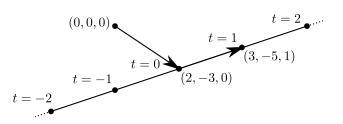
Even better, we can write this as an equation of vectors:

$$(x, y, z) = (2, -3, 0) + t(1, -2, 1).$$

This is often called the "equation of the line", even though

- (a) this equation is not unique, and
- (b) it is actually infinitely many equations; one for each value of t.

Nonetheless, it is a very good way to describe a line because it immediately tells us the important facts: This is the line that starts at the point (2, -3, 0) and from there it moves in the direction of the vector (1, -2, 1). We say that the line is "parametrized" by t.



• Now that we know how to describe lines, we can easily do geometric computations in three dimensions. For practice, find the point of intersection of the plane x + 2y + z = 0 and the line (x, y, z) = (1, 0, 0) + t(1, 2, 3).

The "equation" of the line says (x, y, z) = (1 + t, 2t, 3t). Plugging these values into the equation of the plane gives

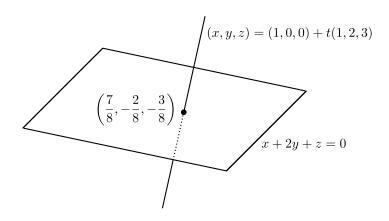
$$x + 2y + z = 0$$

(1+t) + 2(2t) + (3t) = 0
1 + t + 4t + 3t = 0
1 + 8t = 0
t = -1/8.

Finally, this value of t corresponds to the point

$$(x, y, z) = (1, 0, 0) - \frac{1}{8}(1, 2, 3) = \left(\frac{7}{8}, -\frac{2}{3}, -\frac{3}{8}\right).$$

See here:



While we're at it, what's the **angle** between the line and the plane? Well, the **normal** vector for the plane is (1, 2, 1) and the **direction vector** for the line is (1, 2, 3). The angle between these two vectors is

$$\arccos\left(\frac{(1,2,1)\bullet(1,2,3)}{\sqrt{(1,2,1)\bullet(1,2,1)}\sqrt{(1,2,3)\bullet(1,2,3)}}\right) = \arccos\left(\frac{8}{\sqrt{6}\sqrt{14}}\right) \approx 29.2^{\circ}.$$

So you can say that the line is 29.2° from the vertical, or 60.8° from the horizontal. It's hard to see this in the picture, especially since I made absolutely no attempt to show this in the picture.

Now you should be able to compute intersections between lines, planes, and spheres in three dimensional space, should the need ever arise. Computing the intersections between planes is actually quite useful; there is an entire subject devoted to it called **linear algebra**. Linear algebra is the tool to use when you need to compute the intersection of 37 random hyperplanes in 112 dimensional space.