PROFESSOR: Hi. Our unit to today concerns something called the dot product. Now it's very easy to just mechanically give a definition of the dot product, but in keeping with the spirit both of our game and of our correlation between the real and the abstract world, let's keep in mind that when it came time to define vector addition, we chose, as our definition, a concept already used in the physical world-- namely, that of a resultant vector.

And now what we would like to do is to introduce a further structure in our game of vectors. And by the way, keep in mind that I do not need any more physical interpretations to add more complexities to my game, but that perhaps if we do this, it makes the subject more meaningful, both to the applied person and to the theoretical person. The idea that I wanted to keep in mind for motivating today's lesson was the old high school idea of work equals force times distance in the elementary physics course.

Well rather than talk on like this, let's just take a look and see what's going on. Let's say the lecture is called "The Dot Product," and our physical motivation is the recipe from elementary physics, work equals force times distance.

Now this is a deceptive little formula. First of all, what it really meant was, when we learned this recipe, was that the force was taking place in the same direction as the object was moving. Now what happened next was the following situation. We have an object, say, being moved by a force. The object is being moved along a tabletop, say. The force is represented now by a vector, an arrow. And we're assuming enough friction here so that the block moves along the table and doesn't rise. And the question is, how much work is done on the object as the object moves, say, from this point to this point?

Now without worrying about what motivates this thing physically, the important thing was, is that people observe that physically, the only thing that went into the work was the component of the force in the direction of the displacement. In other words, this was the force that one found had to be multiplied by the displacement. In other words, to find the work being done, one took not the magnitude of F, but rather the component of F in the direction of s. And in this diagram, that's just what? It's the magnitude of F times cosine theta.

And that quantity, which was called the effective force, was then multiplied by the displacement. And to write that in more suggestive form, notice that the work was not the magnitude of the force times the magnitude of the displacement. That was only true in the special case where the force and the displacement were in the same direction. But that rather the work is what? It's the magnitude of the force times the magnitude of the displacement times the cosine of the angle between the force and the displacement.

And notice by the way, the very special cases, that if F and s happen to be parallel and have the same sense, then the angle between F and s is 0. The cosine of 0 is 1, in which case that you have the work is the magnitude of force times the displacement. The other extreme case magnitude-wise is if s and F happen to be perpendicular, in which case the angle, of course, is 90 degrees. The cosine of a 90-degree angle is 0. In which case, what? If the force was at right angles to the displacement, the work was 0.

At any rate then, whether we understand this physical motivation or not is irrelevant. The important point is if we want to keep this structure or this particular motivation in a structural form, we now generalize this as follows. We simply say, let A and B be any two vectors, arrows. And we will define A dot B to be the magnitude of A times the magnitude of B times the cosine of the angle between A and B. And I write this this way to indicate an ordering. In other words, don't think of A and B as being A and B. Think of A as denoting the first and B as denoting the second vector. And what we're saying is the first dotted with the second is the magnitude of the first, times the magnitude of the second, times the cosine of the angle as you rotate the first vector into the second.

And the beauty of having a cosine over here is the fact that what? If you reverse the angle of rotation-- in other words, from B into A-- notice that you change the sign of the angle, but the cosine of theta is the same as cosine minus theta, so no harm is done this particular way. On the other hand, had we been dealing with sine of the angle, as we will in our next lecture, this will make a difference.

But be this as it may, we define A dot B to be the magnitude of A times the magnitude of B times the cosine of the angle between A and B. All right? The two extreme cases being what? If A and B are perpendicular, the dot product is 0. If A and B are parallel, the dot product is equal to the magnitude of the product of the two magnitudes.

Now the only difficult thing here, or what we sometimes call undesirable-- and maybe this is what separates the new three-dimensional geometry from the old three-dimensional geometry-- is that the cosine of an angle is particularly difficult to keep track of, especially when the lines are in three-dimensional space. How do you measure an angle this way? You see, in the plane, it's simple. You draw the thing to scale. But in three space, this can be rather difficult.

So what we would like to do is to eliminate this particular term. We would like to find an expression for A dot B that doesn't involve the cosine of an angle, at least directly. And to do this, we simply draw a little diagram. And notice that even if A and B are three-dimensional vectors, since A and B are lines, if they are not parallel, if they emanate at a common point, they form a plane of their own. Let's call that the plane of the blackboard. The third side of the triangle is either A minus B, or B minus A, depending on where you put the arrowhead here. But we've already discussed that idea.

And the interesting point here is notice that A minus B very subtly includes the angle theta. In other words, imagine the magnitudes of A and B to be fixed. And now fan out A and B. As you fan out A and B, what you're doing is what? Just changing this angle. As you change this angle, notice that A minus B changes. Namely, as these fan out, the vector that joins the two arrowheads here becomes a different vector, both in magnitude and direction. In other words, whether it looks it or not, one of the beauties of our vector notation is that the cosine of the angle theta is indirectly included in A minus B.

But now you see, once we have this triangle here, notice that the law of cosines tells us how to relate the third side of a triangle in terms of two sides and the included angle. OK? And in fact, when we use the Law of Cosines, notice that one of the terms is going to be what? The product of the magnitudes of two sides times the cosine of the included angle. And that, roughly speaking, is just what we mean by A dot B.

So without any further ado, what we do now is we just write the Law of Cosines down here, which says what? The magnitude of A minus B squared is equal to the magnitude of A squared plus the magnitude of B squared minus twice the magnitude of A times the magnitude of B times the cosine of the angle between A and B. And then, you see, we simply recognize that this term here is, by definition, A dot B. We can now take this equation and solve for A dot B-- which, by the way, this is very, very important to notice. I should have pointed this out sooner. But A dot B is a number. It's a product of two magnitudes times the cosine of an angle, which is a number.

This is a numerical equation. We can therefore solve for A dot B. And we wind up with what? That A dot B is the magnitude of A minus B squared minus the magnitude of A squared minus the magnitude of B squared, all divided by 2. And the beauty now is that we have expressed A dot B solely in terms of magnitudes. And notice especially in Cartesian coordinates-- and I'll do that next-- but in terms of Cartesian coordinates, notice that magnitudes are particularly simple to find. We just subtract corresponding components and square, et cetera.

But the important point is that even without Cartesian coordinates, this particular result expressed as A dot B in terms of the magnitudes of A, B, and A minus B, and is a result which is independent of any coordinate system. However-- and this is done very simply in the text, reinforced in our exercises. If you elect to write A, B and A minus B in Cartesian coordinates and use this particularly straightforward recipe, what we wind up with is a rather elegant result-- elegant in terms of simplicity, at least.

And that is-- remember in Cartesian coordinates, we would write A as a1 i plus a2 j plus a3 k. B would be b1 i plus b2 j plus b3 k. And then the beauty is what? That A dot B turns out to be very simply and conveniently a1 b1 plus a2 b2 plus a3 b3. In other words, that to dot A and B, if the vectors are written in Cartesian coordinates-- and this is crucial. If this is not done in Cartesian coordinates, you can get into a heck of a mess. And I have deliberately made an exercise on this unit, get you into that mess, if you fall into that particular trap. But if we have Cartesian coordinates, it turns out that to dot two vectors, you simply do what? Multiply the two i components together. Multiply the two j components together. Multiply the two k components together, and add. All right?

By the way, to show you why this works from a structural point of view without belaboring this point right now, notice that if you were to multiply in the usual sense of the word "multiplication," form the dot product here, you would expect to get nine terms. In other words, each of the terms in A multiplies each of the three terms in B. So that altogether you would expect nine terms. The thing that's rather interesting here is that notice that i dot i, j dot j, and k dot k all happen to be 1. Because after all, the magnitudes of these vectors are each 1. The angle between our i and i is 0, j and j is 0. The angle between k and k is 0. So that i dot i, j dot j, and k dot k are all 1.

Whereas on the other hand, when you take mixed terms, notice that because i and j, i and k, and j and k are all at right angles, i dot j, j dot k, i dot k are all going to be 0. And that therefore, those other six terms will drop out. In other words, structurally what's happening here is the fact that the three vectors that we're using here all happen to have unit length. And they happen to be mutually perpendicular. If they were not perpendicular, these mixed terms would appear in here. In other words, in general, when you dot two vectors in three space, depending on the coordinate system, you can expect up to nine terms in your answer. But the beauty is that as long as we have Cartesian coordinates, there happens to be a particularly simple, beautiful recipe to compute A dot B.

Now keep in mind, the A dot B that we're talking about here is the same one that we defined before. It's the magnitude of A times the magnitude of B times the cosine of the angle between A and B. All we're saying is that if we use Cartesian coordinates, we can compute it almost as fast as we can read. And let me show you that in terms of some examples.

My first example is the following. Let's imagine that we have three points in Cartesian three space. A is the point (1,2,3), B is the point (2,4,1), and C is the point (3,0,4). We draw the straight lines AB and AC, and we would like to find the angle BAC-- in other words, the angle theta. The first thing that we do-- and this is one of the beauties of how vectors are used in geometry-- is that we vectorize the lines A and B. We put arrow heads on them. That immediately makes them vectors.

We already know from last time how to read the vectors AB and AC. Namely, AB is the vector i plus 2j. See-- just subtract component by component. 2 minus 1 is 1. 4 minus 2 is 2. 1 minus 3 is minus 3, et cetera. So that the vector AB is the vector i plus 2j minus 2k. And the vector AC, working in a similar way, is 2i minus 2j plus k.

Now the beauty is that we can compute these magnitudes very, very quickly by recipe. And we've just learned the recipe for finding A dot B in a hurry. I mean, well in this case, I don't mean A dot B. I mean the vector AB dotted with the vector AC. And going through the computational details here, we square the components of AB, extract the positive square root, and we find very easily that the magnitude of AB is 3. And the hardest part of these problems for me, as a teacher, is to find ones where I find the sum of 3 squares coming out to be a whole number. So I always use the vector 1, 2, 2, because that's a nice vector that way. Similarly, the vector AC also happens to have magnitude 3. OK?

And to find AB dot AC, what do we do? Let's just come back here and make sure we know what we're doing now. We simply dot component by component. It's (1 x 2) + (2x -2) + (-2 x 1). In other words, AB dot AC is (2- 4 -2), which is minus 4. Now using our recipe, we see what? That cosine theta is AB dot AC divided by the product of the magnitudes of AB and AC, from which we very quickly conclude that the cosine of theta is minus 4/9.

And if you're still mixed up as to what that minus sign means, just by way of a quick review of the inverse trigonometric functions, you locate the point (-4, 9) in the xy-plane, and your angle theta is this particular angle here, which means that in terms of principle values, if you look up the angle in the tables whose cosine is 4/9, that will give you this angle here. Subtract that from 180. And that's the angle that you're looking for.

But the beauty is what? That you can now find an angle between two lines in space without having to geometrically worry about what the angle looks like. The algebra in Cartesian coordinates takes care of this by itself. The same thing happens when you're looking for projections in three-dimensional space. Suppose you have a force and a displacement in three-dimensional space, and you want to project a force onto a line, a direction. And our next example shows how the dot product can be used to find projections. Namely here's a vector A, here's a vector B. And I would like to project the vector A onto the vector B. And I would like to find what the length of that projection is.

Well, from elementary trigonometry, I know that the length of this projection is just the magnitude of A times the cosine of theta. And by the way, notice if theta were be greater than 90-degrees, cosine theta would be negative. And the minus sign would not affect the length. It would simply tell us that the projection was in the opposite sense of B. That's all that would mean.

But here's the interesting point. If you look at the magnitude of A times the cosine of theta, it almost looks like a dot product. After all, theta is the angle between A and B. And if the magnitude of B were in here, this would just be A dot B. But the magnitude of B isn't in here. Of course if the magnitude of B happened to be 1, that would be fine. But the magnitude of B might not be 1. And the most honest way to make it 1 is to divide B by its magnitude. And what will that give you? If you divide any vector by its magnitude, that automatically gives you a unit vector having the same direction and sense as the vector that you started with.

In other words, let u sub B, which is B divided by its magnitude, be the unit vector in the direction-- and by the way, here direction includes sense in the direction of B-- and notice that the unit vector in the direction of B has the same direction as B itself. Therefore, to find the angle between u sub B and A is the same as finding the angle between B and A.

In other words, the kicker now seems to be that I take this length, which is the magnitude of A times cosine theta and rewrite that as follows. It's the magnitude of A-- and now remembering that u sub B has unit length, I just throw that in as a factor-- and theta, being the angle between A and B, is also the angle between A and u sub B. But this, by definition, is A dot u sub B.

You see, in other words, to find the projection of A onto B, all you have to do is dot A with the unit vector in the direction of B. In fact, to summarize that without the u sub B in there, all I'm saying is given two vectors A and B, if you dot A and B and then divide by the magnitude of B, that will be the projection of A in the direction of B. OK? The projection of A in the direction of B. And of course, if you want it the other way around, you have to reverse the roles of A and B. The beauty of this unit in Cartesian coordinates is how easy it is to compute A dot B in Cartesian coordinates.

Oh, another example that you might be interested in that I think is very interesting, and that's the special case where A and B already happen to be unit vectors. If A and B already happen to be unit vectors, then if we use our recipe for the formula for A dot B, we observe that, in this case, by definition of unit vectors, both the magnitudes of A and B are 1. And we find that A dot B is then the cosine of the angle between A and B. Which means that if A and B happen to be unit vectors, as soon as you dot them, you have automatically found the cosine of the angle between the two vectors, which suggests a rather general type of approach. Given any two vectors, divide each by the magnitude, right? That gives you unit vectors. Dot them, and that gives you the cosine of the angle between them. You see?

In particular, and here's an interesting thing-- You know, I don't know if it's that funny, it just struck me as funny. Last night at supper as we were sitting down to eat, my four-year-old looked at me and said, "Dad, did they have baked potatoes when you was a little boy?" And you get the feeling sometimes that people think that the modern world really changed the old in certain basic ways that didn't happen at all.

And one of the interesting points is that long before vector geometry was invented, people were doing three-dimensional geometry using non-vector methods. And one technique that happened to be used were things called directional cosines. Namely, suppose you were given a line in space. OK? As a vector, if you wish to look at it that way-- or if you didn't want to look at it as a vector, imagine the line parallel to the given line that goes through the origin.

As soon as I know what that line looks like, I can compute the angle that it makes with the positive x-axis. I can compute the angle that it makes with the positive y-axis. I can compute the angle that it makes with the positive z-axis. And those three angles uniquely determine the position of the line in space, the direction of the line. OK? And those were called the directional angles. You understand that? That was the three-dimensional analog of slope.

In other words, to find the slope of a line in three-dimensional space, draw the line parallel to that line that goes through the origin, and measure each of the three angles that that line makes with the positive x-, y-, and z-axes. And what the beauty was of the dot product was it just gave us a simpler way of doing that. Namely, if A is any vector, I divide A by its magnitude. That gives me the unit vector in the direction of A.

If I now dot the unit vector in the direction of A with i-- and after all, what is i? i is the unit vector in the direction of the positive x-axis. Since these are both unit vectors, this would be the cosine of the angle between i and A. And that's just what? The cosine of the angle that A makes with the positive x-axis-- traditionally, that angle was called alpha. So u A dot i is simply cosine alpha.

Correspondingly, u A dot j is the cosine of the angle between A and j. That's the cosine of the angle that A makes with the positive y-axis. That angle was called beta. This is cosine beta. And u sub A dot k is cosine gamma where gamma is the angle that A makes with the positive k direction. And these, being the cosines of the angles, these were called the directional cosines, and they yielded the slope of lines. But one must not believe that one needed vectors before he could do three-dimensional geometry. What did happen was that vector techniques greatly simplified many of the aspects of three-dimensional geometry.

Well let's leave this part for a moment and close for today by coming back to our game idea. Remember that ultimately, all we will ever use, once we get started with our game, all we will ever use are the structural properties. Now I've gone through these in the notes. I've gone through them-- well you go through them with me in the text or with yourselves in the text. Let me just point out certain properties of the dot product that are shared by regular arithmetic as well.

For example, A dot B equals B dot A. The dot product is commutative. Why is that? Well think of what A dot B is. It's the magnitude of A times the magnitude of B times the cosine of the angle between A and B. But that's the same as what? The magnitude of B times the magnitude of A. After all, numbers, we know are commutative when you multiply them. And the cosine of the angle between A and B is the same as the cosine of the angle between B and A. So these are equal numerical quantities.

Again, without going through the proof here, it turns out that if you want to dot a vector with the sum of two given vectors, the distributive property holds. Namely, A dot B plus C is A dot B plus A dot C. By the way, if you did want to prove this, all you would have to do if you couldn't see it geometrically, is to argue as follows. You say, you know? The easiest way to add and dot vectors is in Cartesian coordinates.

So let me prove that this result is true in Cartesian coordinates. Carry out the details, and if it works in Cartesian coordinates, since the result doesn't depend on the coordinate system, the result must be true, regardless of the coordinate system. But it's a very simple exercise to actually write A, B, and C in terms of i, j, k components, compute both sides of this expression, and show that they're numerically equal. I say "numerically" because it's crucial to notice that both expressions on either side of the equal sign are numbers. B plus C is a vector. A is a vector. When you dot two vectors, you get a number.

Finally, a scalar multiple of a vector dotted with another vector has the property that you can leave the scalar multiple outside. In other words, you can first dot the two vectors and then multiply by the scalar. In other words, in a way, you don't have to worry about voice inflection when you have a scalar multiple. And we will talk more about these as we go along.

However, what's very crucial is to notice that the dot product does have some difficulties not associated with ordinary multiplication. So I say, beware. For example, somebody might say, I wonder if the dot product is associative. I wonder if A dot B dotted with C is the same as A dotted with B dotted with C. And this is nonsensical. I wanted to do this with great dramatic gesture, but I probably would have broken two fingers against the board here. Let's just cross this out, so that you won't be inclined to remember that. This not only is false, it's stupid. And the reason that it's stupid is that it's nonsensical, that these things don't make sense. Namely, the dot product has been defined to be an operation between two vectors that yields a number.

Notice that as soon as you dot A and B, you no longer have a vector. You have a number. And you cannot dot a number with a vector. In other words, neither (A dot B) dot C nor A dot (B dot C) is defined. Because you see, a number is never dotted with a vector. All right?

And finally, a closing note, as we've already seen, if A is perpendicular to B, then the cosine of the angle between A and B is 0. And that says that if A is perpendicular to B, then A dot B is 0. In fact, this is one of the most common usages on the elementary level, of the dot product, is to prove that two vectors are perpendicular.

But whereas that's a nice property, what causes great hardship here is to notice the following, in particular, that if A dot B is 0, we cannot conclude that either A is the 0 vector or B is the 0 vector. For example, i dot j is 0. But neither i nor j is the 0 vector. You see, with ordinary arithmetic, we had the cancellation rule, things that said if the product of two numbers is 0, at least one of the factors must be 0. With the dot product, this need not be true.

The thing I want you to get from this lesson more than anything else, other than the applications that you can get from the book and from the exercises, is to learn to get a feeling for the structure. Don't be upset that certain vector properties are different than arithmetic properties, and certain ones are the same. Notice in terms of the game, we take our rules as they may apply and just carry them out towards inescapable conclusions. But I think that will become clearer as you read the text and do the exercises.

And until next time when we'll talk about a new vector product, let's just say, so long.

Funding for the publication of this video was provided by the Gabriella and Paul Rosenbaum Foundation. Help OCW continue to provide free and open access to MIT courses by making a donation at ocw.mit.edu/donate.