HERBERT GROSS: Hi. Several times in our course up until now, we have referred to linearity. We referred to it when we talked about systems of linear equations. We referred to it when we talked about systems of linear differential equations. We referred to it when we talked about Laplace transforms. We referred to it when we were mapping the xy plane into the uv plane and talking about differentials.

And now, what we'd like to do in today's lesson is to show how the general concept of a vector space gives us a very nice vehicle to tie together all of these different applications of linearity. In fact, today's lesson is called "Linear Transformations." And by way of a very quick definition, one which I will not bother motivating because we've seen the definition on many, many different occasions. All we're doing now is formalizing it.

Let's suppose I have two vector spaces, V and W, and a mapping f that carries V into W. Remember, these vector spaces are in particular sets. And we've talked about functions that map sets into sets. Now we're mapping a set into a set where the set has a structure, the structure of a vector space.

And what we're saying is that a function that maps a vector space V into a vector space W is called a linear transformation if, given any linear combination of elements in the domain, say C1 alpha 1 plus C2 alpha 2, where C1 and C2 are real numbers, and alpha 1 and alpha 2 are vectors in V, if f of that is C1 f of alpha 1 plus C2 f of alpha 2-- the same linearity definition that we've had in the past. And again, if this is hard for you to visualize, think in terms of pictures. Think of V as, for example, being the xy plane. Think of W as denoting the uv plane.

And we're mapping the xy plane, say, into the uv plane. Not all such mappings will be linear. We studied that when we talked about linear systems of equations.

Now, what we would like to do now is to show how many of the properties of linearity that we've used in the past are immediate consequences of this particular definition. Well, let's take a look at some of the consequences. The first thing I claim is that if f is a linear transformation, it must map the 0 element of V into the 0 element of W. In other words, f acting on 0 maps it into 0.

And the easiest way to prove that is to notice that since 0 is equal to 0 plus 0, by that property, f of 0 is certainly the same as f of 0 plus 0. And now by linearity, f of 0 plus 0 is f of 0 plus f of 0. And now, comparing this with this, remember, these are vectors, the cancellation rule applies for vectors. We then obtain that f of 0 is 0.

And by the way, this is a rather interesting subtlety that comes up. Remember when we talked about linear equations in the plane? In other words, f of x equals mx plus b. That is not the same as a linear transformation. In fact, let me just see if I can scribble something in over here. Look at this.

Suppose I have that f of x is equal to mx plus b. See, that's the equation of a straight line. What is f of 0? If f of x is mx plus b, if I replace x by 0, I get what? 0 times m, which is 0, plus b. And therefore, for f to be a linear transformation according to our definition, this must equal 0. And therefore, what this means is that b itself must equal 0.

In other words, even though we think of a straight line as being linear, to represent a linear transformation, the straight line must pass through the origin. But I don't want to harp on that now. I want to let that go for the exercises because I'm afraid, again, if we spend too much time on details, you will fail to get the overall picture. But at any rate, one property of a linear transformation is that it maps the 0 vector into the 0 vector.

Now, of course, other vectors may be mapped into the 0 vector as well. I mean, we'll talk about that in more detail in the form of a review, in fact, later on. But for the time being, let me show you the next property of linear transformations that are very important. And that is that if V1 and V2 are mapped into 0 by the linear transformation f, then any linear combination of V1 and V2 is also mapped into 0.

And by the way, you may think when I say this-- this may ring a bell and remind you of what we were doing when we talked about homogeneous linear differential equations. And the interesting point is that this is exactly the same as what we were doing there, only now, we're not restricted to be dealing only with linear differential equations. This is now any linear transformation from one vector space into another.

And the proof goes exactly the way it did in the linear differential equation case. Namely, given f of C1V1 plus C2V2, we certainly know that by linearity, that C1 f of V1 plus C2 f of V2. But given that f of V1 is 0, and f of V2 is 0, certainly, this then becomes 0, which is what we wanted to show.

In other words, if you have a linear transformation, the set of elements that are mapped into the 0 element are more than a set. They are themselves a vector space, a subspace of V, a subspace of the domain of f. Because why? Any linear combinations of elements mapped into 0 by f is itself-- any linear combination is itself-- mapped into f.

In other words, let's define n to be the set of all elements eta in V, such that f of eta is 0. And that space is called the null space of f. It's a subspace of the domain of fV. And all we're saying is by our previous remark, if two elements belong to n, their sum belongs to n, and any scalar multiple of an element in n also belongs to n.

So n is indeed a subspace of V. It's a very important subspace. It's called the null space. And one of things about the null space-- and by the way, many people get confused by this name. They say the null space, they think of "null" as being empty, nothing. And they say that null space has nothing in it.

No. Think of it in terms of the mapping. The null space is part of a domain. And what it is-- it's that part of the domain that's mapped into the 0 element of the image. You see? It is not nothing. It's the set of all elements that are mapped into 0. And what we're saying is that that subspace is enough to determine the entire mapping.

Let me show what I mean by that abstractly. I claim that n determines the image of f. Namely, I claim that if I know that f of alpha equals f of beta, and f is a linear transformation, alpha and beta cannot be very random elements of V. In fact, what I claim is that the difference between alpha and beta must be an element of the null space n. Or another way of saying that, I claim that alpha has the form beta plus eta, where eta is an element of the null space n.

And again, this may ring a familiar bell for you when you think in terms of linear differential equations and the like, where we essentially saw that any two solutions of a differential equation, that their difference was a solution of the homogeneous linear differential equation. But again, I don't want to dwell on that now. I want to talk on the most abstract version of this.

How do I know that this is true? Well, look at it. What does it mean to say that f of alpha equals f of beta? It means that f of alpha minus f of beta is 0. But by linearity, f of alpha minus f of beta is the same as f of alpha minus beta.

Therefore, the fact that f of alpha minus f of beta is 0 says that f of alpha minus beta is 0. By definition of the null space n, the fact that alpha minus beta is mapped into 0 means that alpha minus beta belongs to n. And that's the same as saying what? That alpha minus beta is some element in n. Or alpha is beta plus some element in n-- same thing as we asserted.

By the way, another interesting way of seeing one to oneness of a linear transformation is that since the only way that two elements can have the same image is for one of them to equal the other plus something that belongs to n, the only way we can be sure that alpha equals beta is if this element must be 0. You see, alpha equals beta plus 0. We'll say that alpha equals beta. So in particular, if a linear transformation is one to one, the null space must consist of 0 alone. And conversely, if the null space consists of 0 alone, the linear transformation is 1 to 1.

Notice, by the way, the null space can never be empty because by our first property, we show that the 0 vector is always mapped into 0. So at least the 0 vector belongs to the null space. If other vectors belong to the null space, the linear transformation will not be 1 to 1. If 0 is the only vector that belongs to the null space of f, then the linear transformation will be 1 to 1.

Now, I think the time has come to illustrate some of these remarks in terms of concrete examples. For my first example, let's recall that the mapping that maps functions into their derivative is a well-defined linear transformation on the set of differentiable functions. Remember, the derivative of a sum is the sum of the derivatives. The derivative of a scalar times a function is a scalar times the derivative of the function.

In other words, then, the mapping D, which maps a function into its derivative, is a vector space. It's a linear transformation from one vector space to another. In particular, what does the null space look like? The null space of D must be what? All functions f, such that D maps f into 0.

The only way that D of f can be 0 is if f prime is 0. But f prime being identically 0 is the same as saying f equals a constant. Therefore, the null space in this case is the set of all constants.

Notice, by the way, the null space does not consist of 0 alone. It consists of infinitely many numbers, namely all the constants. By the way, that shouldn't be too surprising because after all, the derivative is not a one to one mapping of differentiable functions into the set of functions. Namely, two different functions can have the same derivative.

In fact, what is the property? For D of f to equal D of g, it's necessary and sufficient that f and g differ by a constant. And notice that that's the same as saying what we said before, that for two elements to have the same image under a linear transformation, they must differ by, at most, an element in the null space. That's exactly what happens in this particular example.

As a second example, let's go back to our study of linear differential equations. We wanted to find the solution of the equation L of y equals f of x. Notice that L is a linear transformation mapping functions into functions. The interesting point was, you may recall, that we emphasized the homogeneous equation, where the right-hand side was replaced by 0.

Notice that the null space of the mapping L-- see, the null space-- namely what? All functions y that map into 0 with respect to L. That means what? L of y is 0. Well, the null space is just what? The solution set of L of y equals 0. That's just the homogeneous equation.

And to show you what we mean by this, notice what the general theory was for linear differential equations. Namely, the general solution was the homogeneous solution plus a particular solution, noticing that for the homogeneous solution, it was characterized by the fact that it belonged to the null space. Namely, L of y sub h was 0.

The particular solution was just anything in the space, namely L of yp equals f of x, same as we have over here. And so that structure for linear differential equations was not an exception, but rather one concrete illustration of what a really a transformation means in an abstract vector space.

As a third and final and more visual example, let's go back to mapping the xy plane into the uv plane. Let's suppose I take the mapping u equals x plus y, v equals 2x plus 2y. And notice, by the way, that's just another way of saying v equals twice u. The induced function is the one that carries x comma y in the xy plane into u comma v in the uv plane, so my induced function f bar is the one that maps x comma y into x plus y comma 2x plus 2y. w

And you can look at x comma y as being the vector xi plus yj in the xy plane. Or you can look at it as being the point x comma y in the xy plane. I don't care which interpretation you want to use there. But the point is, what we now would like to know in terms of the null space is what points x comma y are mapped into 0, 0 in the uv plane under the mapping f bar? Or in terms of vectors, what vectors xi plus yj are mapped into the 0 vector under the mapping f bar?

Well, notice that for this to be 0, we must have what? That both components are 0. That means that x plus y must be 0 and that 2x plus 2y is 0. But of course, that's the same as this condition.

In other words, f bar of x comma y is 0, meaning what? It's mapped into the origin if and only if x plus y is 0. In other words, the null space of f bar is the set of all points in the plane x comma y for which x plus y equals 0. Or again, if you want to state this vectorially, it's all vectors xi plus yj in the xy plane, for which x plus y is 0.

And what this means pictorially is this. Notice that the set n is nothing more than what we call the line y equals minus x. In the set of all points x comma y, such as x plus y equals 0, is the same as the set of all points x comma y, such that y equals minus x. y equals minus x, or x plus y equals 0, is a straight line through the origin.

Notice again that key point. f of 0 must be 0 for a linear transformation. If that line didn't pass through the origin, the mapping wouldn't be linear.

Now, notice, by the way, if I come back here, the mapping defined by u and v maps the xy plane into the single line v equals 2u. In other words, the xy plane is mapped into the line v equals 2u. And the entire line x plus y equals 0 is mapped into the origin over here.

In other words, the null space here is an entire line. What line is it? It's the line x plus y equals 0. Why is it called the null space? It's called the null space not because it has no points, but rather because its image consists of the single point 0 in the uv plane.

And by the way, notice again the parallelism here and that which would happen when we studied linear differential equations. Let's suppose for the sake of argument that I want to find all points x comma y that map into the point 1 comma 2 in the uv plane. Essentially, what I want is a particular solution here.

I would like to find one point that maps into 1 comma 2. And my claim is that knowing the null space, once I know one such point, I know them all. Well, one point that maps into 1 comma 2 trivially is 0 comma 1. Namely, 0 plus 1 is 1, and twice 0 plus twice 1 is 2. So 0, 1 is mapped into 1 comma 2 by the mapping f bar.

Now the question is, what other points map into 1 comma 2? And the answer is it's every point on the straight line which passes through 0 comma 1 and is parallel to the line x plus y equals 0. In other words, the line x plus y equals 1 is the line which maps into the point 1 comma 2 under the mapping f bar.

And the easiest way to see this, I would say, is vectorially. Pick any point on this line. Look at it as a vector. That vector is the sum of these two vectors. This vector, being a member of the null space, is mapped into-- I'm sorry. I don't want to do it this way. I'd rather do it-- excuse me.

I'd rather do it this way. I want to utilize the fact that I know that this vector is mapped into this vector. So I'll break down this vector into components this way. I'll take one component like this and one like this.

You see, this component here is indeed the same as this vector because these two lines are parallel. So therefore, f of this is f of this plus f of this. By linearity, that's f of this plus f of this. This being a vector in the null space maps into 0, so the image is just whatever f of this happens to be. You see the idea again?

Call this vector, if you will, alpha. Call this vector here beta. f of alpha plus beta is f of alpha, which is 0, because alpha is in the null space, plus f of beta, which is just 1 comma 2. Again, I think I've made too much of a mountain out of a molehill, not because this isn't important, but because you will recall back in our block four treatment of linear equations, linear algebraic equations, we discussed things like this. And I don't want to get too deeply involved in reviewing the details.

What I am interested in seeing is that you see the structure of what a null space really is. And by the way, let's continue on with what the properties of a linear transformation happen to be. Let's leave this go for the time being. That's enough in terms of examples. Let's go on with the properties.

Our fourth property is that if V is a vector space, and we pick a particular basis, u1 up to un, then V is neatly determined. I say "neatly determined" by its effect on the basis vectors u1 up to un. What I mean by neatly determined is the following. Let's do a specific example.

Suppose for the sake of argument that n is 2. Let's look at V, which has as its basis u1 and u2. Pick any vector little v and v. Relative to u1 and u2, v has a unique representation x1u1 plus x2u2.

f of v by the linear property of f-- f being a linear transformation-- is simply what? x1 f of u1 plus x2 f of u2. And notice that once we know what f does to u1 and u2, we now know what it does to every vector in the space. Namely, you see, to find what f does to V, we simply look at the same vector as V, only with u1 and u2 replaced by f of u1 and f of v2.

And by the way, because we can do this, it implies that linear transformations may be represented also very conveniently in terms of matrix notation. And I think that both of these last two remarks, 4 and 5, can be best illustrated in terms of an example. Let me first give you a fairly abstract example so that you see the terminology. And then I'll make this more concrete. And the concrete example will be our finale for today's lesson.

But first of all, by means of an example-- example number 4. Again, let V be a two-dimensional space with u1 and u2 as a basis. Let's assume now for the sake of argument that the linear transformation f maps V into V itself-- in other words, that V and W are equal in this case, just for the sake of argument.

Now, what did I just say? I said before that the transformation is uniquely determined once we know what it does to each element of a basis. What does f do to u2 and u2?

Well, since f is mapping V into V, and since this implies that every vector in V is going to be with respect to a coordinate system u1, u2, what I do know is that whatever f of u1 and f of u2 are, they are linear combinations of the basis vectors of V u1 and u2. So let's say for the sake of argument that f of u1 is this linear combination. f of u2 is this linear combination, where we're using the usual double subscript notation that seems to indicate a matrix treatment coming up. Well, let me show you now what happens next.

Let's suppose we again pick our arbitrary vector and fix it-- some arbitrary vector V in our space V. Let's suppose that relative to u1 and u2, that vector v is x1u1 plus x2u2. Well, we just saw that f of v must therefore be x1 f of u1 plus x2 f of u2. But now, we know what f of u1 and f of u2 look like explicitly as linear combinations of u1 and u2. And therefore, simply by substituting these expressions for f of u1 and f of u2 in here, leaving again the details for you to verify, I find that relative to a 2-tuple notation, where we're using u1 and u2 as our basis vectors, that f of v is x1a11 plus x2a21 comma x1a12 plus x2a22.

Now, this, at first glance, may not seem to suggest a matrix to you. But at second glance, hopefully it will, namely, if I look at the matrix a11, a21, a12, a22. And notice, by the way, now, a very important subtlety that's come up here. In the past, the matrix has always been written a11, a12, a21, a22 to match the coefficients as they occur in these expressions, here.

But notice, I know what the answer has to be. It has to be this. All I'm saying is if you look at this particular matrix, this is what? A11, x1 plus a21, x2. And if you want the matrix to represent the right 2-tuple, it has to be written this way.

By the way, this is a two-row matrix. To make the thing come out exactly right, I simply take the transpose of this. In other words, notice that this, without the transpose, comes out to be two rows and one column, where this is the first row, this is the second row. To make this look exactly right, I put the transpose in here.

By the way, remember that we have already seen that to take the transpose of the product of two matrices A, B, it's B transpose A transpose. So the transpose would be what? You transpose this, which is x1, x2, and multiply that by the transpose of this, which is a11, a12, a21, a22. In other words, if you do want this matrix to look exactly like the matrix of coefficients, what we're saying is you are going to have to write the vector x1, x2 on the left-hand side rather than on the right-hand side.

And by the way, that's why in many textbooks that deal with matrix algebra and the like, they always put the variable on the left-hand side. It's because if you put the variable on the left-hand side, you can identify the matrix as it stands with the matrix of coefficients verbatim, here. If you don't do that, if you want to write the x1, x2 on the right, then you have to remember that the matrix that you're multiplying it by has its first column, not its first row, representing f of u1, and its second column representing f of u2. And now, let's illustrate that by means of a still more concrete example.

Let's again let f be a linear transformation that maps V into V. Again, let's assume that V is given in terms of a basis u1 and u2. But let's become more concrete now and replace the as by specific numbers. Let's assume, for example, that we know that f of u1 is 3u1 plus 4u2. f of u2 is 5u1 plus 7u2.

What we're saying now is that by linearity, we know exactly what f does to every vector in V, simply in terms of what it does here. Now, let's see how that works. Let me pick an arbitrary vector in capital V. Let's say, for example, it's 2u1 plus u2. According to my theory of the previous example, what must f of V equal?

Well, what I do is I take the matrix whose first column has 3 and 4 as entries and whose second column has 5 and 7 as entries. And I multiply that by the column matrix 2, 1. And to make this come out to be a row vector, I take the transpose of this.

And that, indeed, does turn out to be what? 3 times 2, 5 times 1. That adds up to 11. 4 times 2 is 8. 7 times 1 is 7. That adds up to 15. That's the vector 11 comma 15, relative to the basis u1 and u2.

Again, if you don't like the idea of having to transpose this, if you want to be able to write the matrix 3, 4, 5, 7 directly, the way you do it is write that matrix directly. But now, write the vector v as a row vector on the left side of this matrix. You see? Again, I'll get the same answer, won't I? 2 times 3 is 6. 1 times 5 is 5. 6 plus 5 is 11. 2 times 4 is 8. 1 times 7 is 7. I get 15.

Now, whichever way I do this, notice that f of v is 11u1 plus 15u2. And notice also, there was nothing sacred about this choice of v. I just picked a concrete example to show you how I can determine what the image of any vector v with respect to the linear transformation f looks like once I know the matrix of coefficients of f relative to the basis u1 and u2.

By the way, in every lecture so far I've been coming back to this rather difficult point that everything seems to depend on what basis we choose. And the interesting point is that whereas a linear transformation is defined independently of any coordinate system, the matrix that represents that linear transformation does indeed depend on the coordinate system. For example, let me just say that. The matrix of f does depend on the basis that you've chosen. And by means of an example, let's let f u1, u2, and v be the same as in the previous exercise. Remember what that is, now.

Now, I'm going to pick a new basis for my vector space v. In fact, let's call capital V the same thing, too. Let all this be the same as in exercise 5, example 5.

Now, I'm going to pick a new basis, alpha 1 and alpha 2. Relative to u1 and u2, alpha 1 is 2u1 plus 3u2. Alpha 2 is u1 plus u2. Since these are not constant multiples of one another, alpha 1 and alpha 2 are linearly independent. Because they're linearly independent, they form a basis because they are what? The dimension of V is 2.

And so any set of two linearly independent vectors in a two-dimensional space will be a basis for that space. So I now have a new basis for V, alpha 1 alpha 2. By the way, in the same way that I can express the alphas in terms of the us, I can invert this by any method that I wish and show that the us can be expressed in terms of the alphas in particularly in this particular manner.

What this means now, of course, is that V is still the same vector space. But I can view it as being with respect to the basis alpha 1, alpha 2. But it's the same vector space that I had back in example 5, when I was viewing this with respect to the basis u1 and u2.

Now the thing that comes up now is can I convert everything from example 5 into the language of the alphas? For example, we knew in example 5 that v was 2u1 plus u2. But now, knowing what u1 and u2 look like in terms of the alphas, I simply make this substitution, replacing u1 and u2 by what they look like in terms of the alphas. And I conclude that in terms of the alphas, v is minus alpha 1 plus 4 alpha 2.

In other words, notice that v is the 2-tuple minus 1 comma 4 relative to the basis alpha 1, alpha 2. But if the 2-tuple 2 comma 1 relative to u1 and u2 is the basis-- but that we already knew from the previous lecture that this was the case, that the 2-tuple depended on the basis that was chosen. But now, I'd like to show you that the matrix also depends on the basis chosen.

Well, how do we write f of v relative to alpha 1 and alpha 2? According to my general theory, all I have to know is what alpha 1 and alpha 2 get mapped into. And then I'm home free. In other words, I would like to know what f of alpha 1 and f of alpha 2 look like, but now as linear combinations of alpha 1 and alpha 2.

All right, how do I do this? How do I express f of alpha 1 and f of alpha 2 as linear combinations of alpha 1 and alpha 2? Well, from the previous example, the easiest thing to do is to express f of alpha 1 and f of alpha 2 in terms of u1 and u2. Namely, f of alpha 1 is what? The matrix relative to u1 and u2 is 3, 5, 4, 7.

Alpha 1 written in u1 and u2 components is 2u1 plus 3u2. So f of alpha 1 is the 2-tuple 21 comma 29, where this is relative to u1 and u2. In other words, f of alpha 1 is 21u1 plus 29u2.

We know what alpha 1 and alpha 2 look like in terms of alpha 1 and alpha 2. We just were talking about that. Leaving the details to you, replacing u1 and u2 by what they're equal to in terms of alpha 1 alpha 2, we see that f of alpha 1 is 8 alpha 1 plus 5 alpha 2. We have now, you see, expressed f of alpha 1 as a 2-tuple relative to alpha 1 and alpha 2.

In a similar way, f of alpha 2 is given by this, which, again, replacing u1 and u2 by what they're equal to in terms of the alphas, turns out to be this. So I now know the matrix of f in terms of what it does to alpha 1 and alpha 2. In fact, what will that matrix do?

Writing it in terms of the transposed idea here, the first column of my matrix will be 8, 5. The second column will be 3, 2. So the matrix now of f is 8, 3, 5, 2. I multiply that by the vector v, which as we just saw over here, is minus 1 comma 4.

If I carry out this operation, I have what? Minus 8 plus 12, which is 4, minus 5 plus 8, which is 3. So f of v is now the 2-tuple 4 comma 3 relative to the basis alpha 1, alpha 2. In other words, f of v is 4 alpha 1 plus 3 alpha 2. And remembering that alpha 1 was 2u1 plus 3u2, and that alpha 2 was u1 plus u2, I can now as a check convert this into us.

And I then find what? What is this? 8 plus 3 is 11. 12 plus 3 is 15. f of v is, indeed, 11u1 plus 15u2, which does check with our result of the example-- what number was it? Example 5, the previous example, where we did the same problem in terms of the us.

And what I want you to see is this. Obviously, v is a fixed vector. f is a fixed linear transformation. So f of v must be the same vector, no matter what the basis is.

But in terms of one basis, f of v is the 2-tuple 4 comma 3. In terms of another basis, it's the 2-tuple 11 comma 15. In terms of one basis, f is represented by the matrix 8, 3, 5, 2. In terms of another basis, f is represented by the matrix 3, 5, 4, 7.

So what you see, indeed, is nothing more than this-- that there are two difficult problems that we have to deal with. One problem, which is relatively easy, once you get the basic ideas down, is the fact that linear transformations have nothing to do with the basis. You see? It's defined for a vector space.

But the difficult point is that in many cases we are dealing with a specific basis. In the middle of a problem, we change from one basis to another. Consequently, the basis that represents the linear transformation will, in general, change, as we go from one representation to another. And this is where a lot of computational difficulty seems to creep in.

Again, I will hammer this home during the exercises. But for now, what I hope you have gotten out of this lecture is the overall view of what a linear transformation means and how, in terms of vector spaces, the concept of linearity finds a very nice unified home to live in. At any rate then, until next time, goodbye.

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.