

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation, or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

**PROFESSOR:** Hi. Today, we would like to begin our study of calculus of several real variables. And as I mentioned in our last lecture, I would like to begin it from the traditional point of view.

And I guess I'd like to make a little bit of a short speech even before we begin. There may be a little bit of danger in the way I've been emphasizing the modern approach to mathematics: that one may actually believe that the traditional approach to mathematics left very, very much to be desired. The answer is, it didn't really. That the role of logic in good, traditional mathematics was about the same as it is in good, modern mathematics. That in fact, if we define new mathematics as meaningful mathematics, there really was no difference between the two approaches. Except for the fact that after you have a couple of hundred years of hindsight and can take a subject apart from a different point of view, then obviously it's no great surprise that you can find cleaner ways of really presenting the topic. It's the old cliché of hindsight being better than foresight by a darn sight, or some such thing.

But at any rate, without further ado, let's actually look at the introduction to calculus of several real variables from a traditional point of view. And I've abbreviated that title to simply say, "Calculus of Several Variables." Whereas I say that, think of two things. First of all, it's an introduction. And second of all, we're going to be doing this from the traditional point of view.

Now, what we're assuming is remember, last time we talked about functions of several real variables. We talked about the limit concept, gave you exercises that drilled you on what limits meant, and the like. And as a result, it now makes sense to talk about the calculus, the instantaneous rate of change. Except that you have an

awful lot of variables in here.

In other words, the general case that we'd like to tackle is the case where we have  $w$  being some real number, being a function of the  $n$  independent variables  $x_1$  up to  $x_n$ . And of course, you see, the calculus applies to the  $f$ . The  $w$  used here is to indicate the fact that the output is a real number-- right?

It's like talking about  $f(x)$  of  $x$  versus  $y = f(x)$ . That in a manner of speaking, it's the  $f(x)$  that's important. We won't get into that right now.

What was important, though, about  $y = f(x)$ ? It indicated a graph-- graph. And because it was a graph, we could visualize things pictorially that might have been harder to understand analytically. For this reason, even though  $n$  may be any positive integer whatsoever-- preferably, of course, greater than 1, otherwise we wouldn't have a function of several variables. In fact, that's the definition of a function of several variables-- it's an expression of this type where  $n$  is greater than 1. But in general, and you'll notice this in the text as well, we begin with the special case that  $n$  equals 2. Because, you see, if  $n$  is 2 we need-- what? Two degrees of freedom to plot the independent variables. That would be the  $xy$  plane, for example. A third degree of freedom-- say, the  $z$ -axis, often called  $w$ -axis because of the symbolism.

In other words, in the same way that I could graph a real valued function of a single real variable as a curve in the plane, I can graph a real valued function of two real variables as a surface in three dimensional space. And as a result, when one is adjusting to new language, it's nice to take the case  $n$  equals 2 simply because you can view things pictorially.

The danger, of course, with  $n$  equals 2, is you get so used to the picture that after a while you forget that  $n$  could have been more than 2. So I will try to pay homage to both points of view. I will simply mention here that we will often let  $n$  equal 2 to take advantage of geometry. When we do this, the usual notation is to let  $w$  equal  $f(x,y)$  where  $x$  and  $y$  are independent.

And I should mention why it's so important to talk about independent variables. It'll become clear in a moment. Let me just tell you what independent means, and then I'll show you why it becomes clear in a moment.

To say that  $x$  and  $y$  are independent simply means that you can choose one of the two variables without determining the other. For example, if I were to say let  $y = 2x$ , then certainly  $y$  and  $x$  are not independent, because the choice of  $x$  determines the choice of  $y$ . But to say that  $x$  and  $y$  are independent-- and that's what we mean by two degrees of freedom-- is I can take any value of  $x$  that I want and not be impeded as to the choice of  $y$  that I want. And that's why we say the domain is the  $xy$  plane. I have two degrees of freedom. I can roam anywhere through the plane this way.

But the reason that we want this is as follows. And this is where the traditional mathematics was every bit as logical as the modern mathematics. Let's keep in mind that logic is the art of taking known situations that we know how to handle, you see, and reducing unfamiliar situations to these more familiar situations. The point is that even in traditional mathematics, one knew how to handle the calculus of a single real variable. So what one said was this, given that  $w$  is a function of both  $x$  and  $y$ , let's hold  $y$  fixed, for example, and let  $x$  vary. You see, if I fix  $y$  at some particular value and let  $x$  vary, once I've made that fixed choice of  $y$ , I have  $w$  as a function of  $x$  alone. And as long as  $w$  is a function of  $x$  alone, I already know how to take the derivative of  $w$  with respect to  $x$ .

And by the way, there's no favoritism here between  $x$  and  $y$ . In a similar way, if I hold  $x$  constant and let  $y$  vary, then  $w$  is a function of  $y$  alone. And I can then find the derivative of  $w$  with respect to  $y$ . And these things were called partial derivatives, you see, because it involved holding-- what? All but one of the variables constant. We'll go into that in more detail as we go along.

But the whole idea is-- what? By choosing the variables to be independent, we can hold all of them but one at a time constant. And as long as we're doing that, the resulting function is a function of a single real variable. And this reduces us to part

one of our course.

So at any rate, what we're saying is, fix  $y$  at some constant value-- say,  $y$  equals  $y_0$ . And we then let  $x$  vary as in ordinary one dimensional calculus between  $x_0$  minus the magnitude of  $\Delta x$  and  $x_0$  plus the magnitude of  $\Delta x$ . In other words, we mark off an interval of magnitude,  $\Delta x$ , on either side of  $x_0$  and let  $x$  be any place in here.

Then what we do, is remembering that  $y_0$  is a constant, and that this function, no matter how complicated it looks like, is a function of just  $\Delta x$  alone. We mimic the definition of an ordinary derivative. In other words, we take the limit as  $\Delta x$  approaches 0 of  $f(x_0 + \Delta x, y_0) - f(x_0, y_0)$ , and divide that by  $\Delta x$ .

In other words, with  $y_0$  being treated as a constant, notice that essentially, this is the change in  $f$  as  $x$  varies between  $x_0$  and  $x_0 + \Delta x$ , divided by the change in  $x$ . In other words, it's the limit of an average rate of change of  $f$  of  $x$  with respect to  $x$  for a fixed value of  $y_0$ . And to indicate that, we write this analogously, I would say, to how we write functions of a single real variable.

In other words, we write  $f'_x$ . Instead of  $f'$ , we indicate-- let me put it this way, when you had only one real variable, was there any danger of misinterpreting the prime? After all, with only one independent variable, when you said derivative, it was obvious what variable you were differentiating with respect to. Here, there are two variables,  $x$  and  $y$ . So instead of the prime as a subscript, we write down the variable with respect to which we're taking the derivative. And then we put in the point at which the derivative is being evaluated.

You see, in a similar way, to have defined  $f'_y(x_0, y_0)$ , that would have meant-- what? We hold  $x$  constant at some value  $x_0$ , let  $y$  vary between  $y_0$  and  $y_0 + \Delta y$ , and take the limit of this expression,  $f(x_0, y_0 + \Delta y) - f(x_0, y_0)$ , over  $\Delta y$ . Take the limit of that expression as  $\Delta y$  approaches 0, you see.

And we're going to show this pictorially in a while. After all, that's why we chose the case  $n = 2$ , was so that we could eventually illustrate this thing pictorially. But

for the time being, I prefer not to introduce the picture because I want you to see that even in the case where my input consists of an  $n$ -tuple where  $n$  is greater than 2, or 3, or 4-- whatever you want-- that this thing still makes sense.

In other words, in general, suppose  $w$  is a function of the  $n$  independent variables  $x_1$  up to  $x_n$ . And I want to compute  $f(x_1)$  at the point  $x_1$  equals  $a_1$ -- et cetera--  $x_n$  equals  $a_n$ .

By definition, what I do is I hold  $a_2$  up to  $a_n$  constant. I let  $x_1$  vary from  $a_1$  to  $a_1$  plus  $\Delta x$ . I compute-- what?  $f(a_1 + \Delta x, a_2, \dots, a_n)$ , minus  $f(a_1, a_2, \dots, a_n)$ . Noticing that that numerator is simply the change in  $f$  as  $x_1$  varies from  $a_1$  to  $a_1 + \Delta x$  while all the other variables are being held constant. Which of course I can do if the variables are independent.

I then divide that by  $\Delta x$ . That's the average rate of change. Then I take the limit as  $\Delta x$  approaches 0. And this I can do for any number of variables, provided that they're independent. Keep that in mind, if  $x_n$  could be expressed in terms of the remaining  $x$ 's, how can I hold  $a_n$  constant? Or how can I vary  $a_n$  while I hold the other ones constant? In other words, if the  $a$ 's-- if one of variables depends on the others, they are related in terms of their motion. Keep in mind here, I just chose  $\Delta x$ . Analogous definition is held for  $f_{x_2}$ ,  $f_{x_3}$ , et cetera. OK?

I think again that once you see a few examples, the mystery vanishes. What this thing says in plain English is that when you have a function of several variables and you want to take the derivative with respect to one of those variables, pretend that every one of the other variables was being held constant.

Let me give an example. Let  $w$  be  $f(x_1, x_2, x_3)$ , where the specific  $f$  that I have in mind is obtained as follows. It's  $x_1$  times  $x_2$  times  $x_3$  plus  $e$  to the  $x_1$  power. Now, to take the derivative of this with respect to  $x_1$ , all that this says is treat  $x_2$  and  $x_3$  as if they were constants. Well, look at it. If I treat  $x_2$  and  $x_3$  as if they were constants-- if I differentiate this with respect to  $x_1$ -- all I have left is-- what? This particular constant, which is  $x_2$  times  $x_3$ . And the derivative of  $e$  to the  $x_1$ , with respect to  $x_1$ ,

is  $e$  to the  $x_1$ . Therefore,  $f_{x_1}$  in this case is  $x_2 x_3$ , plus  $e$  to the  $x_1$ .

In particular, if I evaluate  $f_{x_1}$  at the three-tuple  $(1, 2, 3)$ -- namely, I replace  $x_1$  by 1,  $x_2$  by 2,  $x_3$  by 3-- I obtain-- what? 2 times 3 plus  $e$  to the first power. In other words, this is just a number, 6 plus  $e$ .

On the other hand, if I had decided to differentiate this with respect to  $x_2$ , I would've treated  $x_1$  and  $x_3$  as constants. With  $x_1$  and  $x_3$  as constants, the derivative of this with respect to  $x_2$  is just  $x_1 x_3$ . Treating  $x_1$  as a constant, the derivative  $e$  to the  $x_1$  with respect to the variable  $x_3$ , is 0. Because the derivative of any constant is 0.

Notice that this is not really a constant. It's a constant once I've fixed  $x_1$  of the particular value, which is what this particular definition says. In particular,  $f_{x_2}$  of  $(1, 2, 3)$  is 1 times 3, or 3.

I should also mention that sometimes instead of writing  $f_{x_1}$ , we write something that looks like the regular derivative, only we make a funny kind of script  $d$  instead of an ordinary  $d$ . This is read the partial derivative of  $w$  with respect to  $x_1$ . And the relationship between these two notations is rather similar to the relationship that exists between the notations  $dy/dx$  and  $f'$  of  $x$ .

At any rate, if we have the  $n$  independent variables  $x_1$  up to  $x_n$ , and  $f$  is a function of those variables,  $f_{x_1}$ -- et cetera--  $f_{x_n}$  are called the partial derivatives of  $f$  with respect to  $x_1$  up to  $x_n$ , respectively. In other words, this means-- what? The partial derivative of  $f$  with respect to  $x_1$ . That means-- what? Take the ordinary derivative as if  $x_1$  were the only variable. This is the partial of  $f$  with respect to  $x_n$ . That means take the ordinary derivative of  $f$  as if  $x_n$  were the only variable. And you can always do this, provided that your variables are independent.

Because our definitions so closely parallel the structure for what happened in the case of one independent variable-- including the limit theorems, the distance formulas, and what have you-- it turns out that the usual derivative properties still hold. For example, if  $w$  happens to be the function of the two independent variables  $e$  to the  $3x$  plus  $y$  times sine  $2x$  minus  $y$ , and I want to take the derivative of  $w$  with

respect to  $x$ -- meaning differentiate this treating  $y$  as a constant-- notice the treating  $y$  as a constant gives me two functions of-- what?  $x$ . In other words, my function is a product of two functions, each of which depends on  $x$ . So I use the ordinary product rule to differentiate this.

Namely, how do I use the product rule to take the partial of  $w$  with respect to  $x$ ? I treat  $y$  as a constant and I differentiate as if  $x$  is the only variable. I say-- what? It's the first factor  $e^{(3x + y)}$  times the derivative of the second with respect to  $x$ , treating  $y$  as a constant. If I treat  $y$  as a constant, the derivative of  $\sin(2x - y)$  is  $\cos(2x - y)$  times the derivative of what's inside with respect to  $x$ . And that's just 2 plus-- what? The second term times the derivative of the first. And the derivative of  $e$  to the  $3x$  plus  $y$  with respect to  $x$ , treating  $y$  as a constant, is  $e^{(3x + y)}$  times the derivative of  $3x + y$ , with respect to  $x$ , which is just 3. And so I have the partial of  $w$  with respect to  $x$ . In a similar way, I could have found the partial of  $w$  with respect to  $y$ , et cetera.

I should now introduce one little problem that will bother you, maybe even after I help you with this. In fact, I guess what really crushes me is the student who comes up in a live class-- after class-- and says, I understood it until I heard you lecture on it. I hope this doesn't cause that problem. But the danger is-- I tell you look it-- everything that happened for calculus of a single variable happens for calculus of several variables. Then all of a sudden, you'll find an odd textbook. In every textbook, just about, they'll say things like, if you take the partial of the variable  $u$  with respect to  $x$ , you do not get the reciprocal of the partial of  $x$  with respect to  $u$ .

In other words, remember, for one variable,  $dy/dx$ , right?, was the reciprocal of  $dx/dy$ . They say it's not true in several variables. I have written in an accentuated question mark here because I want to explain something very important to you. And by the way, if you understand this, you're 90% of the way home free as far as understanding how calculus of several variables is used in most important applications.

Let me take a very simple case. Let  $u$  equal  $x$  plus  $y$ , and let  $v$  equal  $x$  minus  $y$ , Now

obviously, if  $u$  equals  $x$  plus  $y$ , the partial of  $u$  with respect to  $x$ , if we hold  $y$  constant-- we treat  $x$  as the only variable-- is 1. Right? The partial of  $u$  with respect to  $x$  is 1.

On the other hand, notice that if I add these two equations, the  $y$  term drops out. I get-- what?  $2x$  equals  $u$  plus  $v$ , divide through by 2. That says  $x$  is equal to  $1/2 u$  plus  $1/2 v$ .

Let me now take the partial of  $x$  with respect to  $u$ , holding  $v$  constant, treating  $u$  as the only variable. It's easy for me to see that the partial of  $x$  with respect to  $u$  is  $1/2$ . And certainly, if I now compare these two, it should be clear that this is not the reciprocal of this, or vice versa.

The point that I wanted to mention, though, was the following. That when here-- we said, take the partial of  $u$  with respect to  $x$ -- we were assuming that the independent variables were  $x$  and  $y$ . And when we said, here, take the partial of  $x$  with respect to  $u$ , what did we hold constant? We held  $v$  constant. In other words, I'm not going to keep this habit up very long. But just for the time being, I would like you to get used to the idea of putting in a subscript the variables that are being held constant.

You see, in the textbook when they say this is the case, what they really mean is, if you differentiate  $u$  with respect to  $x$ , holding  $y$  constant, that in general will not be the reciprocal of the derivative of  $x$  with respect to  $u$ , holding  $v$  constant. So I can cross out the question mark now. But the point is, notice that the variables that you're holding constant here are different.

You see, suppose instead I took the partial of  $x$  with respect to  $u$ , holding  $y$  constant. In other words, suppose I solve for  $x$  in terms of  $u$  and  $y$ . That happens to be very easy to do here. Given that  $u$  equals  $x$  plus  $y$ , it follows immediately that  $x$  is equal to  $u$  minus  $y$ . Now, if I take the partial of  $x$  with respect to  $u$ , treating  $y$  as a constant-- so let me do that-- what do I get for an answer? I get 1 for an answer. In other words, the reciprocal of the partial of  $x$  with respect to  $u$ , treating  $y$  as a constant-- as the other independent variable-- is equal to the partial of  $u$  with respect to  $x$  when it's also  $y$  that's being considered as the other variable.

When the variables outside here match up, the recipes work. Course, that leads to the question, then why do the textbooks tell you that this isn't true? Why do they pick this particular representation rather than this one?

And I have a combination example-- an explanation here. And it's simply this. Let's suppose I was given that  $w$  equals  $e$  to the  $x$  plus  $y$  times cosine  $x$  minus  $y$ . It seems that since the only way the variable appears in the first factor is as  $x$  plus  $y$ , and the only way it appears in the second factor is in the form  $x$  minus  $y$ , a very natural substitution might be to say that  $u$  equals  $x$  plus  $y$  and  $v$  equal  $x$  minus  $y$ .

In other words, I could either visualize  $w$  as being a function of  $x$  and  $y$ -- in which case, it would be  $e$  to the  $x$  plus  $y$  times cosine  $x$  minus  $y$ -- or, I could visualize  $w$  as being a simpler function of the two new variables  $u$  and  $v$ , where  $w$  is then just-- what?  $e$  to the  $u$  times cosine  $v$ .

Notice again, that this formula here is much simpler than this one. This one says-- what? Take  $e$  to of the sum of these two, multiplied by the cosine of the difference of these two. And this one just says-- what? Take  $e$  to the first times the cosine of the second.

But the point is, that you might make this change of variables to simplify your computations. The point is, that either we would treat  $w$  as a function of  $x$  and  $y$ , or we would have treated  $w$  as a function of  $u$  would  $v$ .

And by the way, notice here my use of  $f$  and  $g$ . Notice that  $w$  is a different function of  $x$  and  $y$  than it is of  $u$  and  $v$ . But this notation says-- what? We can consider  $w$  either in terms of  $x$  and  $y$ , or in terms of  $u$  and  $v$ . Hardly ever would we consider  $w$  as a function of  $u$  and  $y$ . In other words, either we would change the variables or we don't. Or, for example, in terms of polar versus Cartesian coordinates, either we use  $x$  and  $y$  or we use  $r$  and  $\theta$ . We don't usually use  $r$  and  $x$ .

In other words-- going back to the textbook example again-- If we look at this particular situation here, all the book is saying is, when you're differentiating with respect to  $n$ , you usually mean that  $y$  is the variable being held constant. When we

differentiate with respect to  $u$ , we usually means that  $v$  is the variable that's being held constant.

Again, more of this is said in our exercises. But for the time being, I want you to see how important it is when you have several variables to keep track of which are the dependent variables, which of the independent variables, and how they're coupled.

I said earlier that one of the nice things about picking  $n$  equals 2 is that you can draw a nice picture of the situation. And what I thought you might like to see is the following. Let's suppose you have that  $w$  is some function of  $x$  and  $y$ . What that means is, we can locate-- for a given value of  $x$  and a given value of  $y$ . In other words, the for  $x$  equals  $x_0$ ,  $y$  equals  $y_0$ , we can think of the point  $(x_0, y_0)$  as being in the  $xy$  plane.  $w$ , which is  $f(x_0, y_0)$ , is just the height to this particular surface. In other words, if I let  $f(x_0, y_0)$  be called  $w_0$ , notice that the function  $f(x, y)$  at the point  $x$  equals  $x_0$ ,  $y$  equals  $y_0$ , graphically corresponds to the point whose coordinates are  $x_0, y_0$ , and  $w_0$ . OK? This is now a surface. And in the same way that we use tangent lines to replace curves when we had one independent variable, in two independent variables, we use tangent planes to replace surfaces.

The question that comes up is, how do you get a tangent plane? And what does this have to do with the partial derivatives? And what I would like show you is the following. First of all, one very natural way of intersecting this surface with a plane-- in other words, when you take the partial of  $w$  with respect to  $y$ -- you're holding  $x$  constant, you're saying, let  $x$  equal  $x_0$ . Notice that in three dimensional space,  $x$  equals  $x_0$  is the equation of this plane. And this plane-- see, what is it the plane  $x$  equals  $x_0$ ? It's the plane that goes through the line  $x$  equals  $x_0$  parallel to the  $wy$  plane. This plane intersects my surface in a particular curve that passes through  $p_0$ . I can talk about the slope of the curve at that particular point. OK?

In a similar way, I could've sliced the surface by the plane  $y$  equals  $y_0$ . In other words, hold  $y$  constant. Well, you see, to hold  $y$  constant, that means I, again, draw my plane. It intersects the surface in a different curve, but that curve must also go through the point  $p_0$ . That's the same  $p_0$  that's over here, because after all, the

point on the surface that is directly above  $x_0, y_0$ , is  $p_0$ . And that point doesn't change, no matter what plane you slice this thing with. OK? And so again, I can talk about the slope of the tangent to this particular curve of intersection.

By the way, one very brief aside before I continue on with this. One of the things that makes functions of several variables so difficult is-- I would like you to observe, because this will become the backbone of our future investigations-- that these two particular planes that I drew to intersect my surface were very special planes. One of them was parallel-- they both went through the point  $x_0, y_0$ . But one happened to be parallel to the  $wy$  plane and the other one happened to be parallel to the  $wx$  plane. Notice that, in general, I could have passed infinitely many different planes through  $x_0, y_0$ . Each of which would have intersected the surface in a different curve.

And this is what leads later to the more generalized concept of directional derivatives that we will talk about in more detail in future lectures. But I simply mention the word now because in the reading material in this assignment, some mention is made of directional derivatives. In other words, I wanted to point out that to take partial derivatives, we are interested in two special directions, even though there are other directions.

The question that comes up, of course, is where do the partials come in here? And secondly, knowing where partials come up in here, how do we find the equation of a tangent plane? And all I would like you to see from here is that I'm going to use the same old technique as always. Given these two tangent lines, I'm going to vectorize them. I am then going to find the equation of a plane that passes through those two vectors. To do that, I'm going to have to find the normal to that plane, et cetera. And I'm going to wind up with the equation of a tangent plane.

In more slow motion, all I'm saying is if I now take this picture and draw it over here-- notice the representation-- it looks as if the curve that I get is in the  $wy$  plane. It's actually in a plane parallel to the  $wy$  plane characterized by  $x$  equals  $x_0$ .

And by the way, notice even though  $x$  equals  $x_0$  was chosen so that  $x_0$  was fixed, I

hope it's clear to you that if I let  $x_0$  vary and I picked different values of  $x_0$ , the curve that I get here will, in general, differ. You see, what I'm saying is if I take slices parallel to the  $wy$  plane for a particular surface, the slices that I get-- the shape-- will depend on what the particular value of  $x_0$  is. But that's not the crucial issue right now. The point that I'm driving at is, that if I vectorize the tangent line and call that  $v_1$ , what does that tangent line look like?

First of all, let's see what its slope is. Now again, if you didn't see this over here and all I showed was this diagram, you would say, hey, that slope is just the derivative of  $w$  with respect to  $y$ . But we write that the partial of  $w$  with respect to  $y$  in deference to the fact that  $w$  is not a function of  $y$  alone-- that  $w$  depends on both  $x$  and  $y$ , and the reason that we got this curve was that we fixed  $x$  at some particular value. So this is really-- what? Not the derivative of  $w$  with respect to  $y$  evaluated at  $y$  equals  $y_0$ . It's the partial of  $w$  with respect to  $y$  evaluated at-- what?  $(x_0, y_0)$ .

All right, at any rate, now that we have the slope, notice that the slope of a vector can always be viewed as the  $k$  component. In this case, the  $k$  component divided by the  $j$  component. If I make the  $j$  component 1, the slope will just be the  $k$  component. In other words, a vector which is tangent to the curve-- and one that's easy to write down-- is  $j$  plus the partial of  $w$  with respect to  $y$ , evaluated at  $(x_0, y_0)$ , times  $k$ . And notice, by the way, that this is a number. That once I take the partial derivative and evaluate it at the point, I have a number. So this is-- what? A constant vector once  $x_0$  and  $y_0$  are fixed.

In a similar way, my vector  $v_2$ , which is for my curve parallel to  $wx$  plane-- the slope of that curve at the point  $p_0$  is not  $dw/dx$ -- it's the partial of  $w$  with respect to  $x$ , evaluated  $x_0, y_0$ . Again, why? Because even though this looks like  $w$  is a function of  $x$  alone, here, it was only that way because we took the second variable and, in a sense, froze it at the value  $y_0$ .

At any rate, in an analogous manner,  $v_2$  turns out to be what vector? It's  $i$  plus the partial of  $w$  with respect to  $x$  evaluated  $(x_0, y_0)$ , times  $k$ . In other words, the slope is the partial of  $w$  with respect to  $x$ . And it's in the  $ik$  plane. In other words, the  $xw$

plane.

Now that I have my two vectors, to find the plane that passes through them, all I need is-- what? The normal vector and a point in the plane. Well, one of the nice things about having studied Cartesian-- the cross product here-- is that I can now take the cross product of  $v_1$  and  $v_2$ . Sparing you the details, remember what I do here. I just write down  $i, j, k$ .  $v_1$  has these components,  $v_2$  has these components. I now expand this determinant in the usual way. And I get-- what? This particular vector. In other words, the  $i$  component is the partial of  $w$  with respect to  $x$ . The  $j$  component is a partial of  $w$  with respect to  $y$ . Both evaluated at the point  $(x_0, y_0)$ . And the  $k$  component is minus 1.

Now, to find the equation of a tangent plane, all I have to do is-- what? Take the standard form of the plane. In other words, the coefficients will be these. And then I take  $(x_0, y_0, w_0)$ , which is a point in my plane. And then the equation of the tangent plane becomes, quite simply, this expression here. And in fact, if I rewrite that, notice I can transpose the  $(w - w_0)$  term. That is the change in  $w$ . But to remind us that we're in the tangent plane, we don't write this as  $\Delta w$ , we write it as  $\Delta w_{\text{tan}}$ . And notice that the change in  $w$  to the tangent plane is just the partial of  $w$  with respect to  $x$  times  $\Delta x$ , plus the partial of  $w$  with respect to  $y$  times  $\Delta y$ .

And hopefully, this starts to seem familiar to you. This should almost look like the equation of  $\Delta y_{\text{tan}}$ , only in two dimensions. Remember when we had  $\Delta y_{\text{tan}}$  in part one of the course?

See, what this really says is-- what? This is the change in  $w$  with respect to  $x$ , multiplied by the total change in  $x$ . In other words, this term is-- what? It's the change in  $w$  due to the change in  $x$ , alone. This term is the change in  $w$ , with respect to  $y$ , times the total change in  $y$ . So this term tells you the change in  $w$  due to the change in  $y$ , alone. And if you add these two up, since  $x$  and  $y$  are independent, this should give you the total change in  $w$ .

The reason, of course, that you don't get the total change in  $w$ , but rather,  $\Delta w_{\text{tan}}$ , is the fact that if these expressions here are not constant numbers as  $\Delta x$

and  $\Delta y$  of varying, these numbers, here, are varying. These are variables. But you see, once you fix these at these values, these become numbers, and now you know what  $\Delta w_{\tan}$  is.

Now, in some of the homework problems in the text, what we're going to do is just practice the geometry itself. We're going to find equations of tangent planes. The thing that's very important to us in future lectures from a structural point of view, is that the equation of the tangent plane-- in other words, the equation for  $\Delta w_{\tan}$  - is far more simple to handle than the equation for the true  $\Delta w$ . You see, the equation for  $\Delta w_{\tan}$  just has  $\Delta x$  and  $\Delta y$  appearing to the first power. Whereas, if you try to find  $\Delta w$  exactly for an arbitrary function of  $x$  and  $y$ , this can become a very, very messy thing.

In short,  $\Delta w_{\tan}$  will play for two and more independent variables the same role that  $\Delta y_{\tan}$  played for us in part one of this course as a differential.

And what we'll be discussing next, or at least trying to get at the root of, is just what does play the role of a differential when you're dealing with calculus of several variables? And what we're going to find is that, again, the computations become messy enough so that even though the structure stays fairly nice, there are some wrinkles that come up that will take us a considerable amount of time to untangle.

But we'll worry about that when the time comes. And so until that 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](https://ocw.mit.edu/donate).