

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 to view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu. OK, so we're going to continue looking at what happens when we have non-independent variables. So, I'm afraid we don't take deliveries during class time, sorry. Please take a seat, thanks. [LAUGHTER] [APPLAUSE] OK, so Jason, you please claim your package at the end of lecture. OK, so last time we saw how to use Lagrange multipliers to find the minimum or maximum of a function of several variables when the variables are not independent. And, today we're going to try to figure out more about relations between the variables, and how to handle functions that depend on several variables when they're related. So, just to give you an example, in physics, very often, you have functions that depend on pressure, volume, and temperature where pressure, volume, and temperature are actually not independent. But they are related, say, by $PV=nRT$. So, of course, then you can substitute and expressed a function in terms of two of them only, but very often it's convenient to keep all three. But then we have to figure out, what are the rates of change with respect to t , with respect to each other, the rate of change of f with respect to these variables, and so on. So, we have to figure out what we mean by partial derivatives again. So, OK, more generally, let's say just for the sake of notation, I'm going to think of a function of three variables, x , y , z , where the variables are related by some equation, but I will put in the form g of x , y , z equals some constant. OK, so that's the same kind of setup as we had last time, except now we are not just looking for minima and maxima. We are trying to understand partial derivatives. So, the first observation is that if x , y , and z are related, then that means, in principle, we could solve for one of them, and express it as a function of the two others. So, in particular, can we understand even without solving? Maybe we can not solve. Can we understand how the variables are related to each other? So, for example, z , you can think of z as a function of x and y . So, we can ask ourselves, what are the rates of change of z with respect to x , keeping y constant, or with respect to y keeping x constant? And, of course, if we can solve, that we know the formula for this. And then we can compute these guys. But, what if we can't solve? So, how do we find these things without solving? Well, so let's do an example. Let's say that my relation is $x^2 y z^3 = 8$. And, let's say that I'm looking near the point (x, y, z) equals $(2, 3, 1)$. So, let me check 2^2 plus three times one plus 1^3 is indeed eight. OK, but now, if I change x and y a little bit, how does z change? Well, of course I could solve for z in here. It's a cubic equation. There is actually a formula. But that formula is quite complicated. We actually don't want to do that. There's an easier way. So, how can we do it? Well, let's look at the differential -- -- of this constraint quantity. OK, so if we called this g , let's look at dg . So, what's the differential of this? So, the differential of x^2 is $2x dx$ plus, I think there's a $z dy$. There's a $y dz$, and there's also a $3z^2 dz$. OK, you can get this either by implicit differentiation and the product rule, or you could get this just by putting here, here, and here the partial derivatives of this with respect to x , y , and z . OK, any questions about how I got this? No? OK. So, now, what do I do with this? Well, this represents, somehow, variations of g . But, well, I've set this thing equal to eight. And, eight is a constant. So, it doesn't change. So, in fact, well, we can set this to zero because, well, they call this g . Then, g equals eight is constant. That means we set dg equal to zero. OK, so, now let's just plug in some values at this point. That tells us, well, so if x equals two, that's $4dx$ plus z is one. So, dy plus $3z^2 dz$ should be $6dz$ equals zero. And now, this equation, here, tells us about a relation between the changes in x , y , and z near that point. It tells us how you change x and y , well, how z will change. Or, it tells you actually anything you might want to know about the relations between these variables so, for example, you can move dz to that side, and then express dz in terms of dx and dy . Or, you can move dy to that side and express dy in terms of dx and dz , and so on. It tells you at the level of the derivatives how each of the variables depends on the two others. OK, so, just to clarify this: if we want to view z as a function of x and y , then what we will do is we will just move the dz 's to the other side, and it will tell us dz equals minus one over six times $4dx$ plus dy . And, so that should tell you that partial z over partial x is minus four over six. Well, that's minus two thirds, and partial z over partial y is going to be minus one sixth. OK, another way to think about this: when we compute partial z over partial x , that means that actually we keep y constant. OK, let me actually add some subtitles here. So, here that means we keep y constant. And so, if we keep y constant, another way to think about it is we set dy to zero. We set dy equals zero. So if we do that, we get dx equals negative four sixths dx . That tells us the rate of change of z with respect to x . Here, we set x constant. So, that means we set dx equal to zero. And, if we set dx equal to zero, then we have dz equals negative one sixth of dy . That tells us the rate of change of z with respect to y . OK, any questions about that? No? What, yes? Yes, OK, let me explain that again. So we found an expression for dz in terms of dx and dy . That means that this thing, the differential, is the total differential of z viewed as a function of x and y . OK, and so the coefficients of dx and dy are the partials. Or, another way to think about it, if you want to know partial z over partial x , it means you set y to be constant. Setting y to be constant means that you will put zero in the place of dy . So, you will be left with dz equals minus four sixths dx . And, that will give you the rate of change of z with respect to x when you keep y constant, OK? So, there are various ways to think about this, but hopefully it makes sense. OK, so how do we think about this in general? Well, if we know that g of x , y , z equals a constant, then dg , which is $g_x dx + g_y dy + g_z dz$ should be set equal to zero. OK, and now we can solve for whichever variable we want to express in terms of the others. So, for example, if we care about z as a function of x and y -- -- we'll get that dz is negative g_x over $g_z dx$ minus g_y over $g_z dy$. And, so if we want partial z over partial x , well, so one way is just to say that's going to be the coefficient of dx in here, or just to write down the other way. We are setting y equals constant. So, that means we set dy equal to zero. And then, we get dz equals negative g_x over $g_z dx$. So, that means partial z over partial x is minus g_x over g_z . And, see, that's a very counterintuitive formula because you have this minus sign that you somehow probably couldn't have seen come if you hadn't actually derived things this way. I mean, it's pretty surprising to see that minus sign come out of nowhere the first time you see it. OK, so now we know how to find the rate of change of constrained variables with respect to each other. You can apply the

same to find, if you want partial x , partial y , or any of them, you can do it. Any questions so far? No? OK, so, before we proceed further, I should probably expose some problem with the notations that we have so far. So, let me try to get you a bit confused, OK? So, let's take a very simple example. Let's say I have a function, f of x , y equals $x y$. OK, so far it doesn't sound very confusing. And then, I can write partial f over partial x . And, I think you all know how to compute it. It's going to be just one. OK, so far we are pretty happy. Now let's do a change of variables. Let's set $x=u$ and $y=u+v$. It's not very complicated change of variables. But let's do it. Then, f in terms of u and v , well, so f , remember f was $x y$ becomes u plus u plus v . That's twice u plus v . What's partial f over partial u ? It's two. So, x and u are the same thing. Partial f over partial x , and partial f over partial u , well, unless you believe that one equals two, they are really not the same thing, OK? So, that's an interesting, slightly strange phenomenon. x equals u , but partial f partial x is not the same as partial f partial u . So, how do we get rid of this contradiction? Well, we have to think a bit more about what these notations mean, OK? So, when we write partial f over partial x , it means that we are varying x , keeping y constant. When we write partial f over partial u , it means we are varying u , keeping v constant. So, varying u or varying x is the same thing. But, keeping v constant, or keeping y constant are not the same thing. If I keep y constant, then when I change x , so when I change u , then v will also have to change so that their sum stays the same. Or, if you prefer the other way around, when I do this one I keep v constant. If I keep v constant and I change u , then y will change. It won't be constant. So, that means, well, life looked quite nice and easy with these notations. But, what's dangerous about them is they are not making explicit what it is exactly that we are keeping constant. OK, so just to write things, so here we change u and x that are the same thing. But we keep y constant, while here we change u , which is still the same thing as x . But, what we keep constant is v , or in terms of x and y , that's y minus x constant. And, that's why they are not the same. So, whenever there's any risk of confusion, OK, so not in the cases that we had before because what we've done until now, we didn't really have a problem. But, in a situation like this, to clarify things, we'll actually say explicitly what it is that we want to keep constant. OK, so what's going to be our new notation? Well, so it's not particularly pleasant because it uses, now, a subscript not to indicate what you are differentiating, but rather what you were holding constant. So, that's quite a conflict of notation with what we had before. I think I can safely blame it on physicists or chemists. OK, so this one means we keep y constant, and partial f over partial u with v held constant, similarly. OK, so now what happens is we no longer have any contradiction. We have partial f over partial x with y constant is different from partial f over partial x with v constant, which is the same as partial f over partial u with v constant. OK, so this guy is one. And these guys are two. So, now we can safely use the fact that x equals u if we are keeping track of what is actually held constant, OK? So now, that's going to be particularly important when we have variables that are related because, let's say now that I have a function that depends on x , y , and z . But, x , y , and z are related. Then, it means that I look at, say, x and y as my independent variables, and z as a function of x and y . Then, it means that when I do partials, say, with respect to x , I will hold y constant. But, I will let z vary as a function of x and y . Or, I could do it the other way around. I could vary x , keep z constant, and let y be a function of x and z . And so, I will need to use this kind of notation to indicate which one I mean. OK, any questions? No? All right, so let's try to do an example where we have a function that depends on variables that are related. OK, so I don't want to do one with $PV=nRT$ because probably, I mean, if you've seen it, then you've seen too much of it. And, if you haven't seen it, then maybe it's not the best example. So, let's do a geometric example. So, let's look at the area of the triangle. So, let's say I have a triangle, and my variables will be the sides a and b . And the angle here, θ . OK, so what's the area of this triangle? Well, its base times height over two. So, it's one half of the base is a , and the height is $b \sin \theta$. OK, so that's a function of a , b , and θ . Now, let's say, actually, there is a relation between a , b , and θ that I didn't tell you about, namely, actually, I want to assume that it's a right triangle, OK? So, let's now assume it's a right triangle with, let's say, the hypotenuse is b . So, we have the right angle here, actually. So, a is here. b is here. θ is here. So, saying it's a right triangle is the same thing as saying that b equals $\sin \theta$, OK? So that's our constraint. That's the relation between a , b , and θ . And, this is a function of a , b , and θ . And, let's say that we want to understand how the area depends on θ . OK, what's the rate of change of the area of this triangle with respect to θ ? So, I claim there's various answers. I can think of at least three possible answers. So, what can we possibly mean by the rate of change of A with respect to θ ? So, these are all things that we might want to call partial A partial θ . But of course, we'll have to actually use different notations to distinguish them. So, the first way that we actually already know about is if we just forget about the fact that the variables are related, OK? So, if we just think of little a , b , and θ as independent variables, and we just change θ , keeping a and b constant -- So, that's exactly what we meant by partial A , partial θ , right? I'm not putting any constraints. So, just to use some new notation, that would be the rate of change of A with respect to θ , keeping a and b fixed at the same time. Of course, if we are keeping a and b fixed, and we are changing θ , it means we completely ignore this property of being a right triangle. So, in fact, it corresponds to changing the area by changing the angle, keeping these lengths fixed. And, of course, we lose the right angle. When we rotate this side here, but the angle doesn't stay at a right angle. And that one, we know how to compute, right, because it's the one we've been computing all along. So, that means we keep a and b fixed. And then, so let's see, what's the derivatives of A with respect to θ ? It's one half $ab \cos \theta$. OK, now that one we know. Any questions? No? OK, the two other guys will be more interesting. So far, I'm not really doing anything with my constraint. Let's say that actually I do want to keep the right angle. Then, when I change θ , there's two options. One is I keep a constant, and then of course b will have to change because if this width stays the same, then when I change θ , the height increases, and then this side length increases. The other option is to change the angle, keeping b constant. So, actually, this side stays the same length. But then, a has to become a bit shorter. And, of course, the area will change in different ways depending on what I do. So, that's why I said we have three different answers. So, the next one is keep, I forgot which one I said first. Let's say keep a constant. And, that means that b will change. b is going to be some function of a and θ . Well, in fact here, we know what the function is because we can solve the constraint. namely, b is a over $\cos \theta$. But we don't

actually need to know that so that the triangle, so that the right angle, so that we keep a right angle. And, so the name we will have for this is partial a over partial theta with a held constant, OK? And, the fact that I'm not putting b in my subscript there means that actually b will be a dependent variable. It changes in whatever way it has to change so that when theta changes, a stays the same while b changes so that we keep a right triangle. And, the third guy is the one where we actually keep b constant, and now a, we think a as a function of b and theta, and it changes so that we keep the right angle. So actually as a function of b and theta, it's given over there. $A = b \cos \theta$. And so, this guy is called partial a over partial theta with b held constant. OK, so we've just defined them. We don't know yet how to compute these things. That's what we're going to do now. That is the definition, and what these things mean. Is that clear to everyone? Yes, OK. Yes? OK, so the second answer, again, so one way to ask ourselves, how does the area depend on theta, is to say, well, actually look at the area of the right triangle as a function of a and theta only by solving for b. And then, we'll change theta, keep a constant, and ask, how does the area change? So, when we do that, when we change theta and keep a the same, then b has to change so that it stays a right triangle, right, so that this relation still holds. That requires us to change b. So, when we write partial a over partial theta with a constant, it means that, actually, b will be the dependent variable. It depends on a and theta. And so, the area depends on theta, not only because theta is in the formula, but also because b changes, and b is in the formula. Yes? No, no, we don't keep theta constant. We vary theta, right? The goal is to see how things change when I change theta by a little bit. OK, so if I change theta a little bit in this one, if I change theta a little bit and I keep a the same, then b has to change also in some way. There's a right triangle. And then, because theta and b change, that causes the area to change. OK, so maybe I should re-explain that again. So, theta changes. A is constant. But, we have the constraint, $a = b \cos \theta$. That means that b changes. And then, the question is, how does A change? Well, it will change in part because theta changes, and in part because b changes. But, we want to know how it depends on theta in this situation. Yes? Ah, that's a very good question. So, what about, I don't keep a and b constant? Well, then there's too many choices. So I have to decide actually how I'm going to change things. See, if I just say I have this relation, that means I have two independent variables left, whichever two of the three I want. But, I still have to specify two of them to say exactly which triangle I mean. So, I cannot ask myself just how will it change if I change theta and do random things with a and b? It depends what I do with a and b. Of course, I could choose to change them simultaneously, but then I have to specify how exactly I'm going to do that. Ah, yes, if you wanted to, indeed, we could also change things in such a way that the third side remains constant. And that would be, yet, a different way to attack the problem. I mean, we don't have good notation for this, here, because we didn't give it a name. But, yeah, I mean, we could. We could call this guy c, and then we'd have a different formula, and so on. So, I mean, I'm not looking at it for simplicity. But, you could have many more. I mean, in general, you will want, once you have a set of nice, natural variables, you will want to look mostly at situations where one of the variables changes. Some of them are held fixed, and then some dependent variable does whatever it must so that the constraint keeps holding. OK, so let's try to compute one of them. Let's say I decide that we will compute this one. OK, let's see how we can compute partial a, partial theta with a held fixed. [APPLAUSE] OK, so let's try to compute partial A, partial theta with a held constant. So, let's see three different ways of doing that. So, let me start with method zero. OK, it's not a real method. That's why I'm not getting a positive number. So, that one is just, we solve for b, and we remove b from the formulas. OK, so here it works well because we know how to solve for b. But I'm not considering this to be a real method because in general we don't know how to do that. I mean, in the beginning I had this relation that was an equation of degree three. You don't really want to solve your equation for the dependent variable usually. Here, we can. So, solve for b and substitute. So, how do we do that? Well, the constraint is $a = b \cos \theta$. That means b is $a / \cos \theta$. Some of you know that as a secant theta. That's the same. And now, if we express the area in terms of a and theta only, A is one half of $ab \sin \theta$, sorry, $ab \sin \theta$ is now one half of $a^2 \sin \theta / \cos \theta$. Or, if you prefer, one half of $a^2 \tan \theta$. Well, now that it's only a function of a and theta, I know what it means to take the partial derivative with respect to theta, keeping a constant. I know how to do it. So, partial A over partial theta, a held constant, well, if a is a constant, then I get this one half a^2 coming out times, what's the derivative of tangent? Secant squared, very good. If you're European and you've never heard of secant, that's one over cosine. And, if you know the derivative as one plus tangent squared, that's the same thing. And, it's also correct. OK, so, that's one way of doing it. But, as I've already said, it doesn't get us very far if we don't know how to solve for b. We really used the fact that we could solve for b and get rid of it. So, there's two systematic methods, and let's say the basic rule is that you should give both of them a chance. You should see which one you prefer, and you should be able to use one or the other on the exam. OK, most likely you'll actually have a choice between one or the other. It will be up to you to decide which one you want to use. But, you cannot use solving in substitution. That's not fair. OK, so the first one is to use differentials. By the way, in the notes they are called also method one and method two. I'm not promising that I have the same one, am I? I mean, I might have one and two switched. It doesn't really matter. So, how do we do things using differentials? Well, first, we know that we want to keep a fixed, and that means that we'll set da equal to zero, OK? The second thing that we want to do is we want to look at the constraint. The constraint is $a = b \cos \theta$. And, we want to differentiate that. Well, differentiate the left-hand side. You get da . And, differentiate the right-hand side as a function of b and theta. You should get, well, how many db 's? Well, that's the rate of change with respect to b. That's $\cos \theta db - b \sin \theta d\theta$. That's a product rule applied to b times $\cos \theta$. So -- Well, now, if we have a constraint that's relating da , db , and $d\theta$, OK, so that's actually what we did, right, that's the same sort of thing as what we did at the beginning when we related dx , dy , and dz . That's really the same thing, except now the variables are a, b, and theta. Now, we know that also we are keeping a fixed. So actually, we set this equal to zero. So, we have zero equals $da = \cos \theta db - b \sin \theta d\theta$. That means that actually we know how to solve for db . OK, so $\cos \theta db = b \sin \theta d\theta$ or $db = b \tan \theta d\theta$. OK, so in fact, what we found, if you want, is the rate of change of b with respect to theta. Why do we care?

Well, we care because let's look, now, at dA , the function that we want to look at. OK, so the function is A equals one half $ab \sin \theta$. Well, then, dA , so we had to use the product rule carefully, or we use the partials. So, the coefficient of da will be partial with respect to a . That's one half $b \sin \theta da$ plus coefficient of db will be one half $a \sin \theta db$ plus coefficient of $d\theta$ will be one half $ab \cos \theta d\theta$. But now, what do I do with that? Well, first I said a is constant. So, da is zero. Second, well, actually we don't like b at all, right? We want to view a as a function of θ . So, well, maybe we actually want to use this formula for db that we found in here. OK, and then we'll be left only with $d\theta$ s, which is what we want. So, if we plug this one into that one, we get da equals one half $a \sin \theta$ times $b \tan \theta d\theta$ plus one half $ab \cos \theta d\theta$. And, if we collect these things together, we get one half of ab times $\sin \theta \tan \theta + \cos \theta d\theta$. And, if you know your trig, but you'll see that this is $\sin^2 \theta / \cos \theta + \cos \theta$. That's the same as $\sec \theta$. So, now you have expressed da as something times $d\theta$. Well, that coefficient is the rate of change of A with respect to θ with the understanding that we are keeping a fixed, and letting b vary as a dependent variable. Not enough space: sorry. OK, in case it's clearer for you, let's think about it backwards. So, we wanted to find how A changes. To find how A changes, we write da . But now, this tells us how A depends on a , b , and θ . Well, we know actually we want to keep a constant. So, we set this to be zero. θ , well, we are very happy because we want to express things in terms of θ . db we want to get rid of. How do we get rid of db ? Well, we do that by figuring out how b depends on θ when a is fixed. And, we do that by differentiating the constraint equation, and setting da equal to zero. OK, so -- I guess to summarize the method, we wrote dA in terms of da , db , $d\theta$. Then, we say that a is constant means we set da equals zero. And, the third thing is that because, well, we differentiate the constraint. And, we can solve for db in terms of $d\theta$. And then, we plug into dA , and we get the answer. OK, oops. So, here's another method to do the same thing differently is to use the chain rule. So, we can use the chain rule with dependent variables, OK? So, what does the chain rule tell us? The chain rule tells us, so we will want to differentiate -- -- the formula for A with respect to θ holding a constant. So, I claim, well, what does the chain rule tell us? It tells us that, well, when we change things, A changes because of the changes in the variables. So, part of it is that A depends on θ and θ changes. How fast does θ change? Well, you could call that the rate of change of θ with respect to θ with a constant. But of course, how fast does θ depend to itself? The answer is one. So, that's pretty easy. Plus, then we have the partial derivative, formal partial derivative, of A with respect to a times the rate of change of a in our situation. Well, how does a change if a is constant? Well, it doesn't change. And then, there is A_b , the formal partial derivative times, sorry, the rate of change of b . OK, and how do we find this one? Well, here we have to use the constraint. OK, and we can find this one from the constraint as we've seen at the beginning either by differentiating the constraint, or by using the chain rule on the constraint. So, of course the calculations are exactly the same. See, this is the same formula as the one over there, just dividing everything by partial θ and with subscripts a . But, if it's easier to think about it this way, then that's also valid. OK, so tomorrow we are going to review for the test, so I'm going to tell you a bit more about this also as we go over one practice problem on that.