CALCULUS III

Partial Derivatives

Paul Dawkins

Table of Contents

Preface ii

Chapter 2 : Partial Derivatives 3

Section 2-1 : Limits 5
Section 2-2 : Partial Derivatives 11
Section 2-3 : Interpretations of Partial Derivatives 20
Section 2-4 : Higher Order Partial Derivatives 24
Section 2-5 : Differentials 28
Section 2-6 : Chain Rule 29
Section 2-7 : Directional Derivatives 39

Preface

Here are my online notes for my Calculus III course that I teach here at Lamar University. Despite the fact that these are my “class notes”, they should be accessible to anyone wanting to learn Calculus III or needing a refresher in some of the topics from the class.

These notes do assume that the reader has a good working knowledge of Calculus I topics including limits, derivatives and integration. It also assumes that the reader has a good knowledge of several Calculus II topics including some integration techniques, parametric equations, vectors, and knowledge of three dimensional space.

Here are a couple of warnings to my students who may be here to get a copy of what happened on a day that you missed.

1. Because I wanted to make this a fairly complete set of notes for anyone wanting to learn calculus I have included some material that I do not usually have time to cover in class and because this changes from semester to semester it is not noted here. You will need to find one of your fellow class mates to see if there is something in these notes that wasn’t covered in class.

2. Because I want these notes to provide some more examples for you to read through, I don’t always work the same problems in class as those given in the notes. Likewise, even if I do work some of the problems in here I may work fewer problems in class than are presented here.

3. Sometimes questions in class will lead down paths that are not covered here. I try to anticipate as many of the questions as possible when writing these up, but the reality is that I can’t anticipate all the questions. Sometimes a very good question gets asked in class that leads to insights that I’ve not included here. You should always talk to someone who was in class on the day you missed and compare these notes to their notes and see what the differences are.

4. This is somewhat related to the previous three items, but is important enough to merit its own item. THESE NOTES ARE NOT A SUBSTITUTE FOR ATTENDING CLASS!! Using these notes as a substitute for class is liable to get you in trouble. As already noted not everything in these notes is covered in class and often material or insights not in these notes is covered in class.

Chapter 2 : Partial Derivatives

In Calculus I and in most of Calculus II we concentrated on functions of one variable. In Calculus III we will extend our knowledge of calculus into functions of two or more variables. Despite the fact that this chapter is about derivatives we will start out the chapter with a section on limits of functions of more than one variable. In the remainder of this chapter we will be looking at differentiating functions of more than one variable. As we will see, while there are differences with derivatives of functions of one variable, if you can do derivatives of functions of one variable you shouldn't have any problems differentiating functions of more than one variable. You'll just need to keep one subtlety in mind as we do the work.

Here is a list of topics in this chapter.

Limits – In the section we'll take a quick look at evaluating limits of functions of several variables. We will also see a fairly quick method that can be used, on occasion, for showing that some limits do not exist.

Partial Derivatives – In this section we will the idea of partial derivatives. We will give the formal definition of the partial derivative as well as the standard notations and how to compute them in practice (*i.e.* without the use of the definition). As you will see if you can do derivatives of functions of one variable you won't have much of an issue with partial derivatives. There is only one (very important) subtlety that you need to always keep in mind while computing partial derivatives.

Interpretations of Partial Derivatives – In the section we will take a look at a couple of important interpretations of partial derivatives. First, the always important, rate of change of the function. Although we now have multiple 'directions' in which the function can change (unlike in Calculus I). We will also see that partial derivatives give the slope of tangent lines to the traces of the function.

Higher Order Partial Derivatives – In the section we will take a look at higher order partial derivatives. Unlike Calculus I however, we will have multiple second order derivatives, multiple third order derivatives, *etc*. because we are now working with functions of multiple variables. We will also discuss Clairaut's Theorem to help with some of the work in finding higher order derivatives.

Differentials – In this section we extend the idea of differentials we first saw in Calculus I to functions of several variables.

Chain Rule – In the section we extend the idea of the chain rule to functions of several variables. In particular, we will see that there are multiple variants to the chain rule here all depending on how many variables our function is dependent on and how each of those variables can, in turn, be written in terms of different variables. We will also give a nice method for writing down the chain rule for pretty much any situation you might run into when dealing with functions of multiple variables. In addition, we will derive a very quick way of doing implicit differentiation so we no longer need to go through the process we first did back in Calculus I.

Directional Derivatives – In the section we introduce the concept of directional derivatives. With directional derivatives we can now ask how a function is changing if we allow all the independent variables to change rather than holding all but one constant as we had to do with partial derivatives. In addition, we will define the gradient vector to help with some of the notation and work here. The gradient vector will be very useful in some later sections as well. We will also give a nice fact that will allow us to determine the direction in which a given function is changing the fastest.

Section 2-1 : Limits

In this section we will take a look at limits involving functions of more than one variable. In fact, we will concentrate mostly on limits of functions of two variables, but the ideas can be extended out to functions with more than two variables.

Before getting into this let's briefly recall how limits of functions of one variable work. We say that,

$$\lim_{x \to a} f(x) = L$$

provided,

$$\lim_{x \to a^+} f(x) = \lim_{x \to a^-} f(x) = L$$

Also, recall that,

$$\lim_{x \to a^+} f(x)$$

is a right hand limit and requires us to only look at values of *x* that are greater than *a*. Likewise,

$$\lim_{x \to a^-} f(x)$$

is a left hand limit and requires us to only look at values of *x* that are less than *a*.

In other words, we will have $\lim_{x \to a} f(x) = L$ provided $f(x)$ approaches *L* as we move in towards $x = a$ (without letting $x = a$) from both sides.

Now, notice that in this case there are only two paths that we can take as we move in towards $x = a$. We can either move in from the left or we can move in from the right. Then in order for the limit of a function of one variable to exist the function must be approaching the same value as we take each of these paths in towards $x = a$.

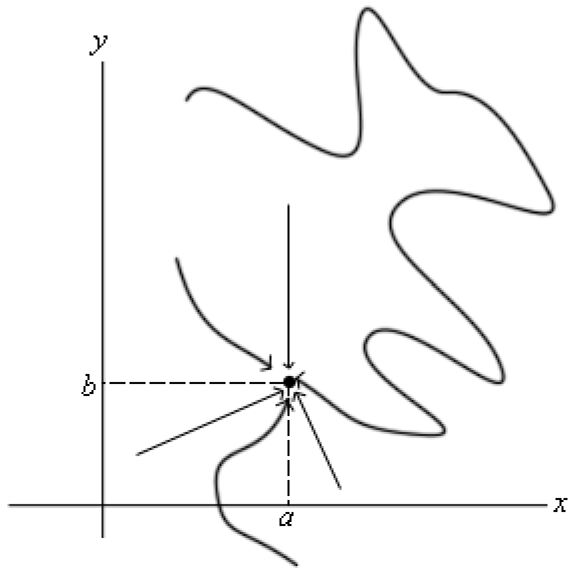
With functions of two variables we will have to do something similar, except this time there is (potentially) going to be a lot more work involved. Let's first address the notation and get a feel for just what we're going to be asking for in these kinds of limits.

We will be asking to take the limit of the function $f(x, y)$ as *x* approaches *a* and as *y* approaches *b*. This can be written in several ways. Here are a couple of the more standard notations.

$$\lim_{\substack{x \to a \\ y \to b}} f(x, y) \qquad\qquad \lim_{(x,y) \to (a,b)} f(x, y)$$

We will use the second notation more often than not in this course. The second notation is also a little more helpful in illustrating what we are really doing here when we are taking a limit. In taking a limit of a function of two variables we are really asking what the value of $f(x, y)$ is doing as we move the point (x, y) in closer and closer to the point (a, b) without actually letting it be (a, b).

Just like with limits of functions of one variable, in order for this limit to exist, the function must be approaching the same value regardless of the path that we take as we move in towards (a,b). The problem that we are immediately faced with is that there are literally an infinite number of paths that we can take as we move in towards (a,b). Here are a few examples of paths that we could take.



We put in a couple of straight line paths as well as a couple of “stranger” paths that aren’t straight line paths. Also, we only included 6 paths here and as you can see simply by varying the slope of the straight line paths there are an infinite number of these and then we would need to consider paths that aren’t straight line paths.

In other words, to show that a limit exists we would technically need to check an infinite number of paths and verify that the function is approaching the same value regardless of the path we are using to approach the point.

Luckily for us however we can use one of the main ideas from Calculus I limits to help us take limits here.

Definition

A function $f(x,y)$ is **continuous** at the point (a,b) if,

$$\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$$

From a graphical standpoint this definition means the same thing as it did when we first saw continuity in Calculus I. A function will be continuous at a point if the graph doesn’t have any holes or breaks at that point.

How can this help us take limits? Well, just as in Calculus I, if you know that a function is continuous at (a,b) then you also know that

$$\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$$

must be true. So, if we know that a function is continuous at a point then all we need to do to take the limit of the function at that point is to plug the point into the function.

All the standard functions that we know to be continuous are still continuous even if we are plugging in more than one variable now. We just need to watch out for division by zero, square roots of negative numbers, logarithms of zero or negative numbers, *etc.*

Note that the idea about paths is one that we shouldn't forget since it is a nice way to determine if a limit doesn't exist. If we can find two paths upon which the function approaches different values as we get near the point then we will know that the limit doesn't exist.

Let's take a look at a couple of examples.

Example 1 Determine if the following limits exist or not. If they do exist give the value of the limit.

(a) $\displaystyle\lim_{(x,y,z)\to(2,1,-1)} 3x^2z + yx\cos(\pi x - \pi z)$

(b) $\displaystyle\lim_{(x,y)\to(5,1)} \frac{xy}{x+y}$

Solution

(a) $\displaystyle\lim_{(x,y,z)\to(2,1,-1)} 3x^2z + yx\cos(\pi x - \pi z)$

Okay, in this case the function is continuous at the point in question and so all we need to do is plug in the values and we're done.

$$\lim_{(x,y,z)\to(2,1,-1)} 3x^2z + yx\cos(\pi x - \pi z) = 3(2)^2(-1) + (1)(2)\cos(2\pi + \pi) = -14$$

(b) $\displaystyle\lim_{(x,y)\to(5,1)} \frac{xy}{x+y}$

In this case the function will not be continuous along the line $y = -x$ since we will get division by zero when this is true. However, for this problem that is not something that we will need to worry about since the point that we are taking the limit at isn't on this line.

Therefore, all that we need to do is plug in the point since the function is continuous at this point.

$$\lim_{(x,y)\to(5,1)} \frac{xy}{x+y} = \frac{5}{6}$$

In the previous example there wasn't really anything to the limits. The functions were continuous at the point in question and so all we had to do was plug in the point. That, of course, will not always be the case so let's work a few examples that are more typical of those you'll see here.

Example 2 Determine if the following limit exist or not. If they do exist give the value of the limit.

$$\lim_{(x,y)\to(1,1)} \frac{2x^2 - xy - y^2}{x^2 - y^2}$$

Solution

In this case the function is not continuous at the point in question (clearly division by zero). However, that does not mean that the limit can't be done. We saw many examples of this in Calculus I where the function was not continuous at the point we were looking at and yet the limit did exist.

In the case of this limit notice that we can factor both the numerator and denominator of the function as follows,

$$\lim_{(x,y)\to(1,1)} \frac{2x^2 - xy - y^2}{x^2 - y^2} = \lim_{(x,y)\to(1,1)} \frac{(2x+y)(x-y)}{(x-y)(x+y)} = \lim_{(x,y)\to(1,1)} \frac{2x+y}{x+y}$$

So, just as we saw in many examples in Calculus I, upon factoring and canceling common factors we arrive at a function that in fact we can take the limit of. So, to finish out this example all we need to do is actually take the limit.

Taking the limit gives,

$$\lim_{(x,y)\to(1,1)} \frac{2x^2 - xy - y^2}{x^2 - y^2} = \lim_{(x,y)\to(1,1)} \frac{2x+y}{x+y} = \frac{3}{2}$$

Before we move on to the next set of examples we should note that the situation in the previous example is what generally happened in many limit examples/problems in Calculus I. In Calculus III however, this tends to be the exception in the examples/problems as the next set of examples will show. In other words, do not expect most of these types of limits to just factor and then exist as they did in Calculus I.

Example 3 Determine if the following limits exist or not. If they do exist give the value of the limit.

(a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2 y^2}{x^4 + 3y^4}$

(b) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^3 y}{x^6 + y^2}$

Solution

(a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2 y^2}{x^4 + 3y^4}$

In this case the function is not continuous at the point in question and so we can't just plug in the point. Also, note that, unlike the previous example, we can't factor this function and do some canceling so that the limit can be taken.

Therefore, since the function is not continuous at the point and because there is no factoring we can do, there is at least a chance that the limit doesn't exist. If we could find two different paths to approach the point that gave different values for the limit then we would know that the limit didn't exist. Two of the more common paths to check are the *x* and *y*-axis so let's try those.

Before actually doing this we need to address just what exactly do we mean when we say that we are going to approach a point along a path. When we approach a point along a path we will do this by either fixing *x* or *y* or by relating *x* and *y* through some function. In this way we can reduce the limit to just a limit involving a single variable which we know how to do from Calculus I.

So, let's see what happens along the *x*-axis. If we are going to approach $(0,0)$ along the *x*-axis we can take advantage of the fact that that along the *x*-axis we know that $y = 0$. This means that, along the *x*-axis, we will plug in $y = 0$ into the function and then take the limit as *x* approaches zero.

$$\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^4+3y^4} = \lim_{(x,0)\to(0,0)} \frac{x^2(0)^2}{x^4+3(0)^4} = \lim_{(x,0)\to(0,0)} 0 = 0$$

So, along the *x*-axis the function will approach zero as we move in towards the origin.

Now, let's try the *y*-axis. Along this axis we have $x = 0$ and so the limit becomes,

$$\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^4+3y^4} = \lim_{(0,y)\to(0,0)} \frac{(0)^2y^2}{(0)^4+3y^4} = \lim_{(0,y)\to(0,0)} 0 = 0$$

So, the same limit along two paths. Don't misread this. This does NOT say that the limit exists and has a value of zero. This only means that the limit happens to have the same value along two paths.

Let's take a look at a third fairly common path to take a look at. In this case we'll move in towards the origin along the path $y = x$. This is what we meant previously about relating *x* and *y* through a function.

To do this we will replace all the *y*'s with *x*'s and then let *x* approach zero. Let's take a look at this limit.

$$\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^4+3y^4} = \lim_{(x,x)\to(0,0)} \frac{x^2x^2}{x^4+3x^4} = \lim_{(x,x)\to(0,0)} \frac{x^4}{4x^4} = \lim_{(x,x)\to(0,0)} \frac{1}{4} = \frac{1}{4}$$

So, a different value from the previous two paths and this means that the limit can't possibly exist.

Note that we can use this idea of moving in towards the origin along a line with the more general path $y = mx$ if we need to.

(b) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^3y}{x^6+y^2}$

Okay, with this last one we again have continuity problems at the origin and again there is no factoring we can do that will allow the limit to be taken. So, again let's see if we can find a couple of paths that give different values of the limit.

First, we will use the path $y = x$. Along this path we have,

$$\lim_{(x,y)\to(0,0)} \frac{x^3 y}{x^6 + y^2} = \lim_{(x,x)\to(0,0)} \frac{x^3 x}{x^6 + x^2} = \lim_{(x,x)\to(0,0)} \frac{x^4}{x^6 + x^2} = \lim_{(x,x)\to(0,0)} \frac{x^2}{x^4 + 1} = 0$$

Now, let's try the path $y = x^3$. Along this path the limit becomes,

$$\lim_{(x,y)\to(0,0)} \frac{x^3 y}{x^6 + y^2} = \lim_{\left(x,x^3\right)\to(0,0)} \frac{x^3 x^3}{x^6 + \left(x^3\right)^2} = \lim_{\left(x,x^3\right)\to(0,0)} \frac{x^6}{2x^6} = \lim_{\left(x,x^3\right)\to(0,0)} \frac{1}{2} = \frac{1}{2}$$

We now have two paths that give different values for the limit and so the limit doesn't exist.

As this limit has shown us we can, and often need, to use paths other than lines like we did in the first part of this example.

So, as we've seen in the previous example limits are a little different here from those we saw in Calculus I. Limits in multiple variables can be quite difficult to evaluate and we've shown several examples where it took a little work just to show that the limit does not exist.

Section 2-2 : Partial Derivatives

Now that we have the brief discussion on limits out of the way we can proceed into taking derivatives of functions of more than one variable. Before we actually start taking derivatives of functions of more than one variable let's recall an important interpretation of derivatives of functions of one variable.

Recall that given a function of one variable, $f\left(x\right)$, the derivative, $f'\left(x\right)$, represents the rate of change of the function as *x* changes. This is an important interpretation of derivatives and we are not going to want to lose it with functions of more than one variable. The problem with functions of more than one variable is that there is more than one variable. In other words, what do we do if we only want one of the variables to change, or if we want more than one of them to change? In fact, if we're going to allow more than one of the variables to change there are then going to be an infinite amount of ways for them to change. For instance, one variable could be changing faster than the other variable(s) in the function. Notice as well that it will be completely possible for the function to be changing differently depending on how we allow one or more of the variables to change.

We will need to develop ways, and notations, for dealing with all of these cases. In this section we are going to concentrate exclusively on only changing one of the variables at a time, while the remaining variable(s) are held fixed. We will deal with allowing multiple variables to change in a later section.

Because we are going to only allow one of the variables to change taking the derivative will now become a fairly simple process. Let's start off this discussion with a fairly simple function.

Let's start with the function $f\left(x,y\right)=2x^2y^3$ and let's determine the rate at which the function is changing at a point, $\left(a,b\right)$, if we hold *y* fixed and allow *x* to vary and if we hold *x* fixed and allow *y* to vary.

We'll start by looking at the case of holding *y* fixed and allowing *x* to vary. Since we are interested in the rate of change of the function at $\left(a,b\right)$ and are holding *y* fixed this means that we are going to always have $y=b$ (if we didn't have this then eventually *y* would have to change in order to get to the point…). Doing this will give us a function involving only *x*'s and we can define a new function as follows,

$$g\left(x\right)=f\left(x,b\right)=2x^2b^3$$

Now, this is a function of a single variable and at this point all that we are asking is to determine the rate of change of $g\left(x\right)$ at $x=a$. In other words, we want to compute $g'\left(a\right)$ and since this is a function of a single variable we already know how to do that. Here is the rate of change of the function at $\left(a,b\right)$ if we hold *y* fixed and allow *x* to vary.

$$g'\left(a\right)=4ab^3$$

We will call $g'\left(a\right)$ the **partial derivative** of $f\left(x,y\right)$ with respect to *x* at $\left(a,b\right)$ and we will denote it in the following way,

$$f_x(a,b) = 4ab^3$$

Now, let's do it the other way. We will now hold *x* fixed and allow *y* to vary. We can do this in a similar way. Since we are holding *x* fixed it must be fixed at $x = a$ and so we can define a new function of *y* and then differentiate this as we've always done with functions of one variable.

Here is the work for this,

$$h(y) = f(a,y) = 2a^2y^3 \qquad \Rightarrow \qquad h'(b) = 6a^2b^2$$

In this case we call $h'(b)$ the **partial derivative** of $f(x,y)$ with respect to *y* at (a,b) and we denote it as follows,

$$f_y(a,b) = 6a^2b^2$$

Note that these two partial derivatives are sometimes called the **first order partial derivatives**. Just as with functions of one variable we can have derivatives of all orders. We will be looking at higher order derivatives in a later section.

Note that the notation for partial derivatives is different than that for derivatives of functions of a single variable. With functions of a single variable we could denote the derivative with a single prime. However, with partial derivatives we will always need to remember the variable that we are differentiating with respect to and so we will subscript the variable that we differentiated with respect to. We will shortly be seeing some alternate notation for partial derivatives as well.

Note as well that we usually don't use the (a,b) notation for partial derivatives as that implies we are working with a specific point which we usually are not doing. The more standard notation is to just continue to use (x,y). So, the partial derivatives from above will more commonly be written as,

$$f_x(x,y) = 4xy^3 \qquad \text{and} \qquad f_y(x,y) = 6x^2y^2$$

Now, as this quick example has shown taking derivatives of functions of more than one variable is done in pretty much the same manner as taking derivatives of a single variable. To compute $f_x(x,y)$ all we need to do is treat all the *y*'s as constants (or numbers) and then differentiate the *x*'s as we've always done. Likewise, to compute $f_y(x,y)$ we will treat all the *x*'s as constants and then differentiate the *y*'s as we are used to doing.

Before we work any examples let's get the formal definition of the partial derivative out of the way as well as some alternate notation.

Since we can think of the two partial derivatives above as derivatives of single variable functions it shouldn't be too surprising that the definition of each is very similar to the definition of the derivative for single variable functions. Here are the formal definitions of the two partial derivatives we looked at above.

$$f_x(x,y) = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h} \qquad f_y(x,y) = \lim_{h\to 0} \frac{f(x,y+h) - f(x,y)}{h}$$

If you recall the Calculus I definition of the limit these should look familiar as they are very close to the Calculus I definition with a (possibly) obvious change.

Now let's take a quick look at some of the possible alternate notations for partial derivatives. Given the function $z = f(x,y)$ the following are all equivalent notations,

$$f_x(x,y) = f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\left(f(x,y)\right) = z_x = \frac{\partial z}{\partial x} = D_x f$$
$$f_y(x,y) = f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\left(f(x,y)\right) = z_y = \frac{\partial z}{\partial y} = D_y f$$

For the fractional notation for the partial derivative notice the difference between the partial derivative and the ordinary derivative from single variable calculus.

$$f(x) \qquad \Rightarrow \qquad f'(x) = \frac{df}{dx}$$
$$f(x,y) \qquad \Rightarrow \qquad f_x(x,y) = \frac{\partial f}{\partial x} \;\; \& \;\; f_y(x,y) = \frac{\partial f}{\partial y}$$

Okay, now let's work some examples. When working these examples always keep in mind that we need to pay very close attention to which variable we are differentiating with respect to. This is important because we are going to treat all other variables as constants and then proceed with the derivative as if it was a function of a single variable. If you can remember this you'll find that doing partial derivatives are not much more difficult that doing derivatives of functions of a single variable as we did in Calculus I.

Example 1 Find all of the first order partial derivatives for the following functions.

(a) $f(x,y) = x^4 + 6\sqrt{y} - 10$

(b) $w = x^2 y - 10y^2 z^3 + 43x - 7\tan(4y)$

(c) $h(s,t) = t^7 \ln(s^2) + \frac{9}{t^3} - \sqrt[7]{s^4}$

(d) $f(x,y) = \cos\left(\frac{4}{x}\right)\mathbf{e}^{x^2 y - 5y^3}$

Solution

(a) $f(x,y) = x^4 + 6\sqrt{y} - 10$

Let's first take the derivative with respect to *x* and remember that as we do so all the *y*'s will be treated as constants. The partial derivative with respect to *x* is,

$$f_x(x,y) = 4x^3$$

Notice that the second and the third term differentiate to zero in this case. It should be clear why the third term differentiated to zero. It's a constant and we know that constants always differentiate to zero. This is also the reason that the second term differentiated to zero. Remember that since we are differentiating with respect to *x* here we are going to treat all *y*'s as constants. That means that terms that only involve *y*'s will be treated as constants and hence will differentiate to zero.

Now, let's take the derivative with respect to *y*. In this case we treat all *x*'s as constants and so the first term involves only *x*'s and so will differentiate to zero, just as the third term will. Here is the partial derivative with respect to *y*.

$$f_y(x,y) = \frac{3}{\sqrt{y}}$$

(b) $w = x^2 y - 10y^2 z^3 + 43x - 7\tan(4y)$

With this function we've got three first order derivatives to compute. Let's do the partial derivative with respect to *x* first. Since we are differentiating with respect to *x* we will treat all *y*'s and all *z*'s as constants. This means that the second and fourth terms will differentiate to zero since they only involve *y*'s and *z*'s.

This first term contains both *x*'s and *y*'s and so when we differentiate with respect to *x* the *y* will be thought of as a multiplicative constant and so the first term will be differentiated just as the third term will be differentiated.

Here is the partial derivative with respect to *x*.

$$\frac{\partial w}{\partial x} = 2xy + 43$$

Let's now differentiate with respect to *y*. In this case all *x*'s and *z*'s will be treated as constants. This means the third term will differentiate to zero since it contains only *x*'s while the *x*'s in the first term and the *z*'s in the second term will be treated as multiplicative constants. Here is the derivative with respect to *y*.

$$\frac{\partial w}{\partial y} = x^2 - 20yz^3 - 28\sec^2(4y)$$

Finally, let's get the derivative with respect to *z*. Since only one of the terms involve *z*'s this will be the only non-zero term in the derivative. Also, the *y*'s in that term will be treated as multiplicative constants. Here is the derivative with respect to *z*.

$$\frac{\partial w}{\partial z} = -30y^2 z^2$$

(c) $h(s,t) = t^7 \ln(s^2) + \frac{9}{t^3} - \sqrt[7]{s^4}$

With this one we'll not put in the detail of the first two. Before taking the derivative let's rewrite the function a little to help us with the differentiation process.

$$h(s,t) = t^7 \ln(s^2) + 9t^{-3} - s^{\frac{4}{7}}$$

Now, the fact that we're using *s* and *t* here instead of the "standard" *x* and *y* shouldn't be a problem. It will work the same way. Here are the two derivatives for this function.

$$h_s(s,t) = \frac{\partial h}{\partial s} = t^7\left(\frac{2s}{s^2}\right) - \frac{4}{7}s^{-\frac{3}{7}} = \frac{2t^7}{s} - \frac{4}{7}s^{-\frac{3}{7}}$$

$$h_t(s,t) = \frac{\partial h}{\partial t} = 7t^6\ln\left(s^2\right) - 27t^{-4}$$

Remember how to differentiate natural logarithms.

$$\frac{d}{dx}\left(\ln g(x)\right) = \frac{g'(x)}{g(x)}$$

(d) $f(x,y) = \cos\left(\frac{4}{x}\right)\mathbf{e}^{x^2y-5y^3}$

Now, we can't forget the product rule with derivatives. The product rule will work the same way here as it does with functions of one variable. We will just need to be careful to remember which variable we are differentiating with respect to.

Let's start out by differentiating with respect to *x*. In this case both the cosine and the exponential contain *x*'s and so we've really got a product of two functions involving *x*'s and so we'll need to product rule this up. Here is the derivative with respect to *x*.

$$\begin{aligned} f_x(x,y) &= -\sin\left(\frac{4}{x}\right)\left(-\frac{4}{x^2}\right)\mathbf{e}^{x^2y-5y^3} + \cos\left(\frac{4}{x}\right)\mathbf{e}^{x^2y-5y^3}\left(2xy\right) \\ &= \frac{4}{x^2}\sin\left(\frac{4}{x}\right)\mathbf{e}^{x^2y-5y^3} + 2xy\cos\left(\frac{4}{x}\right)\mathbf{e}^{x^2y-5y^3} \end{aligned}$$

Do not forget the chain rule for functions of one variable. We will be looking at the chain rule for some more complicated expressions for multivariable functions in a later section. However, at this point we're treating all the *y*'s as constants and so the chain rule will continue to work as it did back in Calculus I.

Also, don't forget how to differentiate exponential functions,

$$\frac{d}{dx}\left(\mathbf{e}^{f(x)}\right) = f'(x)\mathbf{e}^{f(x)}$$

Now, let's differentiate with respect to *y*. In this case we don't have a product rule to worry about since the only place that the *y* shows up is in the exponential. Therefore, since *x*'s are considered to be constants for this derivative, the cosine in the front will also be thought of as a multiplicative constant. Here is the derivative with respect to *y*.

$$f_y(x,y) = \left(x^2 - 15y^2\right)\cos\left(\frac{4}{x}\right)\mathbf{e}^{x^2y-5y^3}$$

Example 2 Find all of the first order partial derivatives for the following functions.

(a) $z = \dfrac{9u}{u^2 + 5v}$

(b) $g(x, y, z) = \dfrac{x\sin(y)}{z^2}$

(c) $z = \sqrt{x^2 + \ln(5x - 3y^2)}$

Solution

(a) $z = \dfrac{9u}{u^2 + 5v}$

We also can't forget about the quotient rule. Since there isn't too much to this one, we will simply give the derivatives.

$$z_u = \frac{9(u^2 + 5v) - 9u(2u)}{(u^2 + 5v)^2} = \frac{-9u^2 + 45v}{(u^2 + 5v)^2}$$

$$z_v = \frac{(0)(u^2 + 5v) - 9u(5)}{(u^2 + 5v)^2} = \frac{-45u}{(u^2 + 5v)^2}$$

In the case of the derivative with respect to *v* recall that *u*'s are constant and so when we differentiate the numerator we will get zero!

(b) $g(x, y, z) = \dfrac{x\sin(y)}{z^2}$

Now, we do need to be careful however to not use the quotient rule when it doesn't need to be used. In this case we do have a quotient, however, since the *x*'s and *y*'s only appear in the numerator and the *z*'s only appear in the denominator this really isn't a quotient rule problem.

Let's do the derivatives with respect to *x* and *y* first. In both these cases the *z*'s are constants and so the denominator in this is a constant and so we don't really need to worry too much about it. Here are the derivatives for these two cases.

$$g_x(x, y, z) = \frac{\sin(y)}{z^2} \qquad\qquad g_y(x, y, z) = \frac{x\cos(y)}{z^2}$$

Now, in the case of differentiation with respect to *z* we can avoid the quotient rule with a quick rewrite of the function. Here is the rewrite as well as the derivative with respect to *z*.

$$g(x, y, z) = x\sin(y)z^{-2}$$

$$g_z(x, y, z) = -2x\sin(y)z^{-3} = -\frac{2x\sin(y)}{z^3}$$

We went ahead and put the derivative back into the "original" form just so we could say that we did. In practice you probably don't really need to do that.

(c) $z = \sqrt{x^2 + \ln\left(5x - 3y^2\right)}$

In this last part we are just going to do a somewhat messy chain rule problem. However, if you had a good background in Calculus I chain rule this shouldn't be all that difficult of a problem. Here are the two derivatives,

$$\begin{aligned}
z_x &= \frac{1}{2}\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}} \frac{\partial}{\partial x}\left(x^2 + \ln\left(5x - 3y^2\right)\right) \\
&= \frac{1}{2}\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}} \left(2x + \frac{5}{5x - 3y^2}\right) \\
&= \left(x + \frac{5}{2\left(5x - 3y^2\right)}\right)\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}}
\end{aligned}$$

$$\begin{aligned}
z_y &= \frac{1}{2}\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}} \frac{\partial}{\partial y}\left(x^2 + \ln\left(5x - 3y^2\right)\right) \\
&= \frac{1}{2}\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}} \left(\frac{-6y}{5x - 3y^2}\right) \\
&= -\frac{3y}{5x - 3y^2}\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}}
\end{aligned}$$

So, there are some examples of partial derivatives. Hopefully you will agree that as long as we can remember to treat the other variables as constants these work in exactly the same manner that derivatives of functions of one variable do. So, if you can do Calculus I derivatives you shouldn't have too much difficulty in doing basic partial derivatives.

There is one final topic that we need to take a quick look at in this section, implicit differentiation. Before getting into implicit differentiation for multiple variable functions let's first remember how implicit differentiation works for functions of one variable.

Example 3 Find $\frac{dy}{dx}$ for $3y^4 + x^7 = 5x$.

Solution

Remember that the key to this is to always think of *y* as a function of *x*, or $y = y\left(x\right)$ and so whenever we differentiate a term involving *y*'s with respect to *x* we will really need to use the chain rule which will mean that we will add on a $\frac{dy}{dx}$ to that term.

The first step is to differentiate both sides with respect to *x*.

$$12y^3 \frac{dy}{dx} + 7x^6 = 5$$

The final step is to solve for $\frac{dy}{dx}$.

$$\frac{dy}{dx} = \frac{5-7x^6}{12y^3}$$

Now, we did this problem because implicit differentiation works in exactly the same manner with functions of multiple variables. If we have a function in terms of three variables *x*, *y*, and *z* we will assume that *z* is in fact a function of *x* and *y*. In other words, $z = z\left(x, y\right)$. Then whenever we differentiate *z*'s with respect to *x* we will use the chain rule and add on a $\frac{\partial z}{\partial x}$. Likewise, whenever we differentiate *z*'s with respect to *y* we will add on a $\frac{\partial z}{\partial y}$.

Let's take a quick look at a couple of implicit differentiation problems.

Example 4 Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for each of the following functions.

(a) $x^3z^2 - 5xy^5z = x^2 + y^3$

(b) $x^2\sin\left(2y - 5z\right) = 1 + y\cos\left(6zx\right)$

Solution

(a) $x^3z^2 - 5xy^5z = x^2 + y^3$

Let's start with finding $\frac{\partial z}{\partial x}$. We first will differentiate both sides with respect to *x* and remember to add on a $\frac{\partial z}{\partial x}$ whenever we differentiate a *z* from the chain rule.

$$3x^2z^2 + 2x^3z\frac{\partial z}{\partial x} - 5y^5z - 5xy^5\frac{\partial z}{\partial x} = 2x$$

Remember that since we are assuming $z = z\left(x, y\right)$ then any product of *x*'s and *z*'s will be a product and so will need the product rule!

Now, solve for $\frac{\partial z}{\partial x}$.

$$\left(2x^3z - 5xy^5\right)\frac{\partial z}{\partial x} = 2x - 3x^2z^2 + 5y^5z$$

$$\frac{\partial z}{\partial x} = \frac{2x - 3x^2z^2 + 5y^5z}{2x^3z - 5xy^5}$$

Now we'll do the same thing for $\frac{\partial z}{\partial y}$ except this time we'll need to remember to add on a $\frac{\partial z}{\partial y}$ whenever we differentiate a *z* from the chain rule.

$$2x^3 z\frac{\partial z}{\partial y} - 25xy^4 z - 5xy^5\frac{\partial z}{\partial y} = 3y^2$$
$$\left(2x^3 z - 5xy^5\right)\frac{\partial z}{\partial y} = 3y^2 + 25xy^4 z$$
$$\frac{\partial z}{\partial y} = \frac{3y^2 + 25xy^4 z}{2x^3 z - 5xy^5}$$

(b) $x^2\sin\left(2y-5z\right) = 1 + y\cos\left(6zx\right)$

We'll do the same thing for this function as we did in the previous part. First let's find $\frac{\partial z}{\partial x}$.

$$2x\sin\left(2y-5z\right) + x^2\cos\left(2y-5z\right)\left(-5\frac{\partial z}{\partial x}\right) = -y\sin\left(6zx\right)\left(6z + 6x\frac{\partial z}{\partial x}\right)$$

Don't forget to do the chain rule on each of the trig functions and when we are differentiating the inside function on the cosine we will need to also use the product rule. Now let's solve for $\frac{\partial z}{\partial x}$.

$$2x\sin\left(2y-5z\right) - 5\frac{\partial z}{\partial x}x^2\cos\left(2y-5z\right) = -6zy\sin\left(6zx\right) - 6yx\sin\left(6zx\right)\frac{\partial z}{\partial x}$$
$$2x\sin\left(2y-5z\right) + 6zy\sin\left(6zx\right) = \left(5x^2\cos\left(2y-5z\right) - 6yx\sin\left(6zx\right)\right)\frac{\partial z}{\partial x}$$
$$\frac{\partial z}{\partial x} = \frac{2x\sin\left(2y-5z\right) + 6zy\sin\left(6zx\right)}{5x^2\cos\left(2y-5z\right) - 6yx\sin\left(6zx\right)}$$

Now let's take care of $\frac{\partial z}{\partial y}$. This one will be slightly easier than the first one.

$$x^2\cos\left(2y-5z\right)\left(2 - 5\frac{\partial z}{\partial y}\right) = \cos\left(6zx\right) - y\sin\left(6zx\right)\left(6x\frac{\partial z}{\partial y}\right)$$
$$2x^2\cos\left(2y-5z\right) - 5x^2\cos\left(2y-5z\right)\frac{\partial z}{\partial y} = \cos\left(6zx\right) - 6xy\sin\left(6zx\right)\frac{\partial z}{\partial y}$$
$$\left(6xy\sin\left(6zx\right) - 5x^2\cos\left(2y-5z\right)\right)\frac{\partial z}{\partial y} = \cos\left(6zx\right) - 2x^2\cos\left(2y-5z\right)$$
$$\frac{\partial z}{\partial y} = \frac{\cos\left(6zx\right) - 2x^2\cos\left(2y-5z\right)}{6xy\sin\left(6zx\right) - 5x^2\cos\left(2y-5z\right)}$$

There's quite a bit of work to these. We will see an easier way to do implicit differentiation in a later section.

Section 2-3 : Interpretations of Partial Derivatives

This is a fairly short section and is here so we can acknowledge that the two main interpretations of derivatives of functions of a single variable still hold for partial derivatives, with small modifications of course to account of the fact that we now have more than one variable.

The first interpretation we've already seen and is the more important of the two. As with functions of single variables partial derivatives represent the rates of change of the functions as the variables change. As we saw in the previous section, $f_x\left(x,y\right)$ represents the rate of change of the function $f\left(x,y\right)$ as we change *x* and hold *y* fixed while $f_y\left(x,y\right)$ represents the rate of change of $f\left(x,y\right)$ as we change *y* and hold *x* fixed.

Example 1 Determine if $f\left(x,y\right)=\dfrac{x^2}{y^3}$ is increasing or decreasing at $\left(2,5\right)$,

(a) if we allow *x* to vary and hold *y* fixed.
(b) if we allow *y* to vary and hold *x* fixed.

Solution
(a) If we allow *x* to vary and hold *y* fixed.

In this case we will first need $f_x\left(x,y\right)$ and its value at the point.

$$f_x\left(x,y\right)=\frac{2x}{y^3} \qquad \Rightarrow \qquad f_x\left(2,5\right)=\frac{4}{125}>0$$

So, the partial derivative with respect to *x* is positive and so if we hold *y* fixed the function is increasing at $\left(2,5\right)$ as we vary *x*.

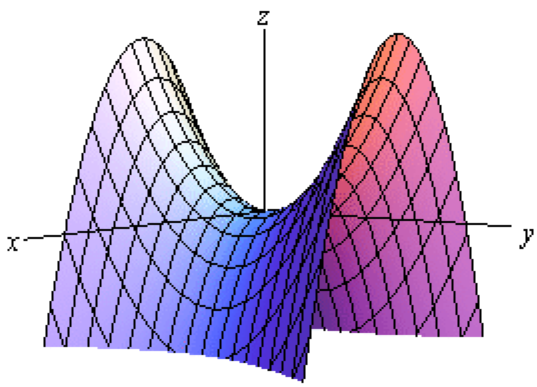
(b) If we allow *y* to vary and hold *x* fixed.

For this part we will need $f_y\left(x,y\right)$ and its value at the point.

$$f_y\left(x,y\right)=-\frac{3x^2}{y^4} \qquad \Rightarrow \qquad f_y\left(2,5\right)=-\frac{12}{625}<0$$

Here the partial derivative with respect to *y* is negative and so the function is decreasing at $\left(2,5\right)$ as we vary *y* and hold *x* fixed.

Note that it is completely possible for a function to be increasing for a fixed *y* and decreasing for a fixed *x* at a point as this example has shown. To see a nice example of this take a look at the following graph.



This is a graph of a hyperbolic paraboloid and at the origin we can see that if we move in along the *y*-axis the graph is increasing and if we move along the *x*-axis the graph is decreasing. So it is completely possible to have a graph both increasing and decreasing at a point depending upon the direction that we move. We should never expect that the function will behave in exactly the same way at a point as each variable changes.

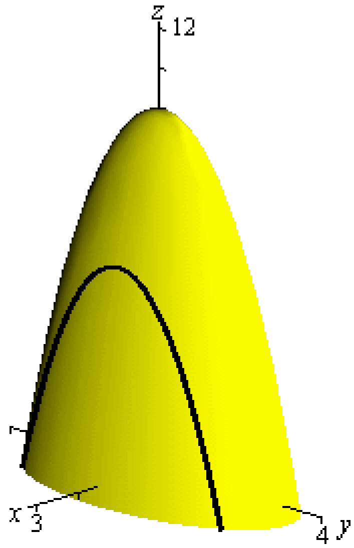
The next interpretation was one of the standard interpretations in a Calculus I class. We know from a Calculus I class that $f'(a)$ represents the slope of the tangent line to $y = f(x)$ at $x = a$. Well, $f_x(a,b)$ and $f_y(a,b)$ also represent the slopes of tangent lines. The difference here is the functions that they represent tangent lines to.

Partial derivatives are the slopes of traces. The partial derivative $f_x(a,b)$ is the slope of the trace of $f(x,y)$ for the plane $y = b$ at the point (a,b). Likewise the partial derivative $f_y(a,b)$ is the slope of the trace of $f(x,y)$ for the plane $x = a$ at the point (a,b).

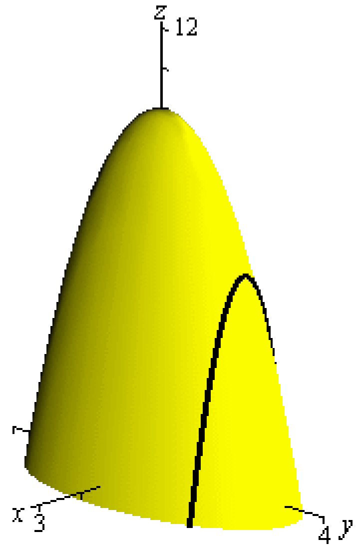
Example 2 Find the slopes of the traces to $z = 10 - 4x^2 - y^2$ at the point $(1,2)$.

Solution

We sketched the traces for the planes $x = 1$ and $y = 2$ in a previous section and these are the two traces for this point. For reference purposes here are the graphs of the traces.



Trace for $x = 1$



Trace for $y = 2$

Next, we'll need the two partial derivatives so we can get the slopes.

$$f_x(x,y) = -8x \qquad\qquad f_y(x,y) = -2y$$

To get the slopes all we need to do is evaluate the partial derivatives at the point in question.

$$f_x(1,2) = -8 \qquad\qquad f_y(1,2) = -4$$

So, the tangent line at $(1,2)$ for the trace to $z = 10 - 4x^2 - y^2$ for the plane $y = 2$ has a slope of -8. Also the tangent line at $(1,2)$ for the trace to $z = 10 - 4x^2 - y^2$ for the plane $x = 1$ has a slope of -4.

Finally, let's briefly talk about getting the equations of the tangent line. Recall that the equation of a line in 3-D space is given by a vector equation. Also, to get the equation we need a point on the line and a vector that is parallel to the line.

The point is easy. Since we know the *x*-*y* coordinates of the point all we need to do is plug this into the equation to get the point. So, the point will be,

$$\left(a, b, f(a,b)\right)$$

The parallel (or tangent) vector is also just as easy. We can write the equation of the surface as a vector function as follows,

$$\vec{r}(x,y) = \langle x, y, z\rangle = \langle x, y, f(x,y)\rangle$$

We know that if we have a vector function of one variable we can get a tangent vector by differentiating the vector function. The same will hold true here. If we differentiate with respect to *x* we will get a tangent vector to traces for the plane $y = b$ (*i.e.* for fixed *y*) and if we differentiate with respect to *y* we will get a tangent vector to traces for the plane $x = a$ (or fixed *x*).

So, here is the tangent vector for traces with fixed *y*.

$$\vec{r}_x(x,y) = \left\langle 1, 0, f_x(x,y) \right\rangle$$

We differentiated each component with respect to *x*. Therefore, the first component becomes a 1 and the second becomes a zero because we are treating *y* as a constant when we differentiate with respect to *x*. The third component is just the partial derivative of the function with respect to *x*.

For traces with fixed *x* the tangent vector is,

$$\vec{r}_y(x,y) = \left\langle 0, 1, f_y(x,y) \right\rangle$$

The equation for the tangent line to traces with fixed *y* is then,

$$\vec{r}(t) = \left\langle a, b, f(a,b) \right\rangle + t\left\langle 1, 0, f_x(a,b) \right\rangle$$

and the tangent line to traces with fixed *x* is,

$$\vec{r}(t) = \left\langle a, b, f(a,b) \right\rangle + t\left\langle 0, 1, f_y(a,b) \right\rangle$$

Example 3 Write down the vector equations of the tangent lines to the traces to $z = 10 - 4x^2 - y^2$ at the point $(1,2)$.

Solution

There really isn't all that much to do with these other than plugging the values and function into the formulas above. We've already computed the derivatives and their values at $(1,2)$ in the previous example and the point on each trace is,

$$\left(1, 2, f(1,2)\right) = (1,2,2)$$

Here is the equation of the tangent line to the trace for the plane $y = 2$.

$$\vec{r}(t) = \left\langle 1,2,2 \right\rangle + t\left\langle 1,0,-8 \right\rangle = \left\langle 1+t, 2, 2-8t \right\rangle$$

Here is the equation of the tangent line to the trace for the plane $x = 1$.

$$\vec{r}(t) = \left\langle 1,2,2 \right\rangle + t\left\langle 0,1,-4 \right\rangle = \left\langle 1, 2+t, 2-4t \right\rangle$$

Section 2-4 : Higher Order Partial Derivatives

Just as we had higher order derivatives with functions of one variable we will also have higher order derivatives of functions of more than one variable. However, this time we will have more options since we do have more than one variable.

Consider the case of a function of two variables, $f(x,y)$ since both of the first order partial derivatives are also functions of *x* and *y* we could in turn differentiate each with respect to *x* or *y*. This means that for the case of a function of two variables there will be a total of four possible second order derivatives. Here they are and the notations that we'll use to denote them.

$$\begin{aligned}
(f_x)_x &= f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} \\
(f_x)_y &= f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} \\
(f_y)_x &= f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} \\
(f_y)_y &= f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}
\end{aligned}$$

The second and third second order partial derivatives are often called mixed partial derivatives since we are taking derivatives with respect to more than one variable. Note as well that the order that we take the derivatives in is given by the notation for each these. If we are using the subscripting notation, *e.g.* f_{xy}, then we will differentiate from left to right. In other words, in this case, we will differentiate first with respect to *x* and then with respect to *y*. With the fractional notation, *e.g.* $\frac{\partial^2 f}{\partial y \partial x}$, it is the opposite. In these cases we differentiate moving along the denominator from right to left. So, again, in this case we differentiate with respect to *x* first and then *y*.

Let's take a quick look at an example.

Example 1 Find all the second order derivatives for $f(x,y) = \cos(2x) - x^2\mathbf{e}^{5y} + 3y^2$.

Solution

We'll first need the first order derivatives so here they are.

$$\begin{aligned}
f_x(x,y) &= -2\sin(2x) - 2x\mathbf{e}^{5y} \\
f_y(x,y) &= -5x^2\mathbf{e}^{5y} + 6y
\end{aligned}$$

Now, let's get the second order derivatives.

$$f_{xx} = -4\cos(2x) - 2\mathbf{e}^{5y}$$
$$f_{xy} = -10x\mathbf{e}^{5y}$$
$$f_{yx} = -10x\mathbf{e}^{5y}$$
$$f_{yy} = -25x^2\mathbf{e}^{5y} + 6$$

Notice that we dropped the (x, y) from the derivatives. This is fairly standard and we will be doing it most of the time from this point on. We will also be dropping it for the first order derivatives in most cases.

Now let's also notice that, in this case, $f_{xy} = f_{yx}$. This is not by coincidence. If the function is "nice enough" this will always be the case. So, what's "nice enough"? The following theorem tells us.

Clairaut's Theorem

Suppose that *f* is defined on a disk *D* that contains the point (a,b). If the functions f_{xy} and f_{yx} are continuous on this disk then,

$$f_{xy}(a,b) = f_{yx}(a,b)$$

Now, do not get too excited about the disk business and the fact that we gave the theorem for a specific point. In pretty much every example in this class if the two mixed second order partial derivatives are continuous then they will be equal.

Example 2 Verify Clairaut's Theorem for $f(x, y) = x\mathbf{e}^{-x^2y^2}$.

Solution

We'll first need the two first order derivatives.

$$f_x(x, y) = \mathbf{e}^{-x^2y^2} - 2x^2y^2\mathbf{e}^{-x^2y^2}$$
$$f_y(x, y) = -2yx^3\mathbf{e}^{-x^2y^2}$$

Now, compute the two mixed second order partial derivatives.

$$f_{xy}(x, y) = -2yx^2\mathbf{e}^{-x^2y^2} - 4x^2y\mathbf{e}^{-x^2y^2} + 4x^4y^3\mathbf{e}^{-x^2y^2} = -6x^2y\mathbf{e}^{-x^2y^2} + 4x^4y^3\mathbf{e}^{-x^2y^2}$$
$$f_{yx}(x, y) = -6yx^2\mathbf{e}^{-x^2y^2} + 4y^3x^4\mathbf{e}^{-x^2y^2}$$

Sure enough they are the same.

So far we have only looked at second order derivatives. There are, of course, higher order derivatives as well. Here are a couple of the third order partial derivatives of function of two variables.

$$f_{xyx} = \left(f_{xy}\right)_x = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial^3 f}{\partial x \partial y \partial x}$$

$$f_{yxx} = \left(f_{yx}\right)_x = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial^3 f}{\partial x^2 \partial y}$$

Notice as well that for both of these we differentiate once with respect to *y* and twice with respect to *x*. There is also another third order partial derivative in which we can do this, f_{xxy}. There is an extension to Clairaut's Theorem that says if all three of these are continuous then they should all be equal,

$$f_{xxy} = f_{xyx} = f_{yxx}$$

To this point we've only looked at functions of two variables, but everything that we've done to this point will work regardless of the number of variables that we've got in the function and there are natural extensions to Clairaut's theorem to all of these cases as well. For instance,

$$f_{xz}\left(x,y,z\right) = f_{zx}\left(x,y,z\right)$$

provided both of the derivatives are continuous.

In general, we can extend Clairaut's theorem to any function and mixed partial derivatives. The only requirement is that in each derivative we differentiate with respect to each variable the same number of times. In other words, provided we meet the continuity condition, the following will be equal

$$f_{ssrtsrr} = f_{trsrssr}$$

because in each case we differentiate with respect to *t* once, *s* three times and *r* three times.

Let's do a couple of examples with higher (well higher order than two anyway) order derivatives and functions of more than two variables.

Example 3 Find the indicated derivative for each of the following functions.

(a) Find f_{xxyzz} for $f\left(x,y,z\right) = z^3 y^2 \ln\left(x\right)$

(b) Find $\dfrac{\partial^3 f}{\partial y \partial x^2}$ for $f\left(x,y\right) = \mathbf{e}^{xy}$

Solution

(a) Find f_{xxyzz} for $f\left(x,y,z\right) = z^3 y^2 \ln\left(x\right)$

In this case remember that we differentiate from left to right. Here are the derivatives for this part.

$$f_x = \frac{z^3 y^2}{x}$$

$$f_{xx} = -\frac{z^3 y^2}{x^2}$$

$$f_{xxy} = -\frac{2z^3 y}{x^2}$$

$$f_{xxyz} = -\frac{6z^2 y}{x^2}$$
$$f_{xxyzz} = -\frac{12zy}{x^2}$$

(b) Find $\frac{\partial^3 f}{\partial y \partial x^2}$ **for** $f(x,y) = \mathbf{e}^{xy}$

Here we differentiate from right to left. Here are the derivatives for this function.

$$\frac{\partial f}{\partial x} = y\mathbf{e}^{xy}$$
$$\frac{\partial^2 f}{\partial x^2} = y^2\mathbf{e}^{xy}$$
$$\frac{\partial^3 f}{\partial y \partial x^2} = 2y\mathbf{e}^{xy} + xy^2\mathbf{e}^{xy}$$

Section 2-5 : Differentials

This is a very short section and is here simply to acknowledge that just like we had differentials for functions of one variable we also have them for functions of more than one variable. Also, as we've already seen in previous sections, when we move up to more than one variable things work pretty much the same, but there are some small differences.

Given the function $z = f\left(x, y\right)$ the differential *dz* or *df* is given by,

$$dz = f_x\,dx + f_y\,dy \qquad \text{or} \qquad df = f_x\,dx + f_y\,dy$$

There is a natural extension to functions of three or more variables. For instance, given the function $w = g\left(x, y, z\right)$ the differential is given by,

$$dw = g_x\,dx + g_y\,dy + g_z\,dz$$

Let's do a couple of quick examples.

Example 1 Compute the differentials for each of the following functions.

(a) $z = \mathbf{e}^{x^2+y^2}\tan\left(2x\right)$

(b) $u = \dfrac{t^3r^6}{s^2}$

Solution

(a) $z = \mathbf{e}^{x^2+y^2}\tan\left(2x\right)$

There really isn't a whole lot to these outside of some quick differentiation. Here is the differential for the function.

$$dz = \left(2x\mathbf{e}^{x^2+y^2}\tan\left(2x\right) + 2\mathbf{e}^{x^2+y^2}\sec^2\left(2x\right)\right)dx + 2y\mathbf{e}^{x^2+y^2}\tan\left(2x\right)dy$$

(b) $u = \dfrac{t^3r^6}{s^2}$

Here is the differential for this function.

$$du = \frac{3t^2r^6}{s^2}dt + \frac{6t^3r^5}{s^2}dr - \frac{2t^3r^6}{s^3}ds$$

Note that sometimes these differentials are called the **total differentials**.

Section 2-6 : Chain Rule

We've been using the standard chain rule for functions of one variable throughout the last couple of sections. It's now time to extend the chain rule out to more complicated situations. Before we actually do that let's first review the notation for the chain rule for functions of one variable.

The notation that's probably familiar to most people is the following.

$$F\left(x\right)=f\left(g\left(x\right)\right) \qquad\qquad F'\left(x\right)=f'\left(g\left(x\right)\right)g'\left(x\right)$$

There is an alternate notation however that while probably not used much in Calculus I is more convenient at this point because it will match up with the notation that we are going to be using in this section. Here it is.

$$\text{If} \qquad y=f\left(x\right) \qquad \text{and} \qquad x=g\left(t\right) \qquad \text{then} \qquad \frac{dy}{dt}=\frac{dy}{dx}\frac{dx}{dt}$$

Notice that the derivative $\frac{dy}{dt}$ really does make sense here since if we were to plug in for *x* then *y* really would be a function of *t*. One way to remember this form of the chain rule is to note that if we think of the two derivatives on the right side as fractions the *dx*'s will cancel to get the same derivative on both sides.

Okay, now that we've got that out of the way let's move into the more complicated chain rules that we are liable to run across in this course.

As with many topics in multivariable calculus, there are in fact many different formulas depending upon the number of variables that we're dealing with. So, let's start this discussion off with a function of two variables, $z=f\left(x,y\right)$. From this point there are still many different possibilities that we can look at. We will be looking at two distinct cases prior to generalizing the whole idea out.

Case 1 : $z=f\left(x,y\right)$, $x=g\left(t\right)$, $y=h\left(t\right)$ and compute $\frac{dz}{dt}$.

This case is analogous to the standard chain rule from Calculus I that we looked at above. In this case we are going to compute an ordinary derivative since *z* really would be a function of *t* only if we were to substitute in for *x* and *y*.

The chain rule for this case is,

$$\frac{dz}{dt}=\frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}$$

So, basically what we're doing here is differentiating *f* with respect to each variable in it and then multiplying each of these by the derivative of that variable with respect to *t*. The final step is to then add all this up.

Let's take a look at a couple of examples.

Example 1 Compute $\dfrac{dz}{dt}$ for each of the following.

(a) $z = x\mathbf{e}^{xy}$, $x = t^2$, $y = t^{-1}$

(b) $z = x^2y^3 + y\cos x$, $x = \ln\left(t^2\right)$, $y = \sin\left(4t\right)$

Solution

(a) $z = x\mathbf{e}^{xy}$, $x = t^2$, $y = t^{-1}$

There really isn't all that much to do here other than using the formula.

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= \left(\mathbf{e}^{xy} + yx\mathbf{e}^{xy}\right)\left(2t\right) + x^2\mathbf{e}^{xy}\left(-t^{-2}\right) \\ &= 2t\left(\mathbf{e}^{xy} + yx\mathbf{e}^{xy}\right) - t^{-2}x^2\mathbf{e}^{xy}\end{aligned}$$

So, technically we've computed the derivative. However, we should probably go ahead and substitute in for *x* and *y* as well at this point since we've already got *t*'s in the derivative. Doing this gives,

$$\frac{dz}{dt} = 2t\left(\mathbf{e}^{t} + t\mathbf{e}^{t}\right) - t^{-2}t^4\mathbf{e}^{t} = 2t\mathbf{e}^{t} + t^2\mathbf{e}^{t}$$

Note that in this case it might actually have been easier to just substitute in for *x* and *y* in the original function and just compute the derivative as we normally would. For comparison's sake let's do that.

$$z = t^2\mathbf{e}^{t} \qquad \Rightarrow \qquad \frac{dz}{dt} = 2t\mathbf{e}^{t} + t^2\mathbf{e}^{t}$$

The same result for less work. Note however, that often it will actually be more work to do the substitution first.

(b) $z = x^2y^3 + y\cos x$, $x = \ln\left(t^2\right)$, $y = \sin\left(4t\right)$

Okay, in this case it would almost definitely be more work to do the substitution first so we'll use the chain rule first and then substitute.

$$\begin{aligned}\frac{dz}{dt} &= \left(2xy^3 - y\sin x\right)\left(\frac{2}{t}\right) + \left(3x^2y^2 + \cos x\right)\left(4\cos\left(4t\right)\right) \\ &= \frac{4\sin^3\left(4t\right)\ln t^2 - 2\sin\left(4t\right)\sin\left(\ln t^2\right)}{t} + 4\cos\left(4t\right)\left(3\sin^2\left(4t\right)\left[\ln t^2\right]^2 + \cos\left(\ln t^2\right)\right)\end{aligned}$$

Note that sometimes, because of the significant mess of the final answer, we will only simplify the first step a little and leave the answer in terms of *x*, *y*, and *t*. This is dependent upon the situation, class and instructor however so be careful about not substituting in for without first talking to your instructor.

Now, there is a special case that we should take a quick look at before moving on to the next case. Let's suppose that we have the following situation,

$$z = f(x, y) \qquad\qquad y = g(x)$$

In this case the chain rule for $\frac{dz}{dx}$ becomes,

$$\frac{dz}{dx} = \frac{\partial f}{\partial x}\frac{dx}{dx} + \frac{\partial f}{\partial y}\frac{dy}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}$$

In the first term we are using the fact that,

$$\frac{dx}{dx} = \frac{d}{dx}(x) = 1$$

Let's take a quick look at an example.

Example 2 Compute $\frac{dz}{dx}$ for $z = x\ln(xy) + y^3$, $y = \cos(x^2+1)$

Solution

We'll just plug into the formula.

$$\begin{aligned}\frac{dz}{dx} &= \left(\ln(xy) + x\frac{y}{xy}\right) + \left(x\frac{x}{xy} + 3y^2\right)\left(-2x\sin(x^2+1)\right)\\ &= \ln\left(x\cos(x^2+1)\right) + 1 - 2x\sin(x^2+1)\left(\frac{x}{\cos(x^2+1)} + 3\cos^2(x^2+1)\right)\\ &= \ln\left(x\cos(x^2+1)\right) + 1 - 2x^2\tan(x^2+1) - 6x\sin(x^2+1)\cos^2(x^2+1)\end{aligned}$$

Now let's take a look at the second case.

<u>**Case 2 :**</u> $z = f(x, y)$, $x = g(s,t)$, $y = h(s,t)$ and compute $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.

In this case if we were to substitute in for *x* and *y* we would get that *z* is a function of *s* and *t* and so it makes sense that we would be computing partial derivatives here and that there would be two of them.

Here is the chain rule for both of these cases.

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \qquad\qquad \frac{\partial z}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}$$

So, not surprisingly, these are very similar to the first case that we looked at. Here is a quick example of this kind of chain rule.

Example 3 Find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$ for $z = \mathbf{e}^{2r}\sin(3\theta)$, $r = st - t^2$, $\theta = \sqrt{s^2+t^2}$.

Solution

Here is the chain rule for $\frac{\partial z}{\partial s}$.

$$\begin{aligned}\frac{\partial z}{\partial s} &= \left(2\mathbf{e}^{2r}\sin(3\theta)\right)(t) + \left(3\mathbf{e}^{2r}\cos(3\theta)\right)\frac{s}{\sqrt{s^2+t^2}} \\ &= t\left(2\mathbf{e}^{2\left(st-t^2\right)}\sin\left(3\sqrt{s^2+t^2}\right)\right) + \frac{3s\mathbf{e}^{2\left(st-t^2\right)}\cos\left(3\sqrt{s^2+t^2}\right)}{\sqrt{s^2+t^2}}\end{aligned}$$

Now the chain rule for $\frac{\partial z}{\partial t}$.

$$\begin{aligned}\frac{\partial z}{\partial t} &= \left(2\mathbf{e}^{2r}\sin(3\theta)\right)(s-2t) + \left(3\mathbf{e}^{2r}\cos(3\theta)\right)\frac{t}{\sqrt{s^2+t^2}} \\ &= (s-2t)\left(2\mathbf{e}^{2\left(st-t^2\right)}\sin\left(3\sqrt{s^2+t^2}\right)\right) + \frac{3t\mathbf{e}^{2\left(st-t^2\right)}\cos\left(3\sqrt{s^2+t^2}\right)}{\sqrt{s^2+t^2}}\end{aligned}$$

Okay, now that we've seen a couple of cases for the chain rule let's see the general version of the chain rule.

Chain Rule

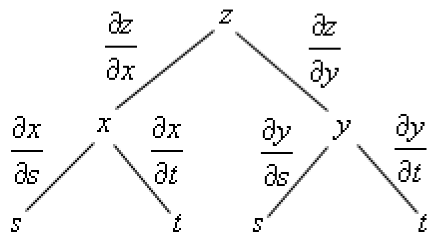
Suppose that *z* is a function of *n* variables, $x_1, x_2, \ldots, x_n$, and that each of these variables are in turn functions of *m* variables, $t_1, t_2, \ldots, t_m$. Then for any variable t_i, $i = 1, 2, \ldots, m$ we have the following,

$$\frac{\partial z}{\partial t_i} = \frac{\partial z}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial z}{\partial x_2}\frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial z}{\partial x_n}\frac{\partial x_n}{\partial t_i}$$

Wow. That's a lot to remember. There is actually an easier way to construct all the chain rules that we've discussed in the section or will look at in later examples. We can build up a **tree diagram** that will give us the chain rule for any situation. To see how these work let's go back and take a look at the chain rule for $\frac{\partial z}{\partial s}$ given that $z = f(x,y)$, $x = g(s,t)$, $y = h(s,t)$. We already know what this is, but it may help to illustrate the tree diagram if we already know the answer. For reference here is the chain rule for this case,

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

Here is the tree diagram for this case.



We start at the top with the function itself and the branch out from that point. The first set of branches is for the variables in the function. From each of these endpoints we put down a further set of branches that gives the variables that both *x* and *y* are a function of. We connect each letter with a line and each line represents a partial derivative as shown. Note that the letter in the numerator of the partial derivative is the upper "node" of the tree and the letter in the denominator of the partial derivative is the lower "node" of the tree.

To use this to get the chain rule we start at the bottom and for each branch that ends with the variable we want to take the derivative with respect to (*s* in this case) we move up the tree until we hit the top multiplying the derivatives that we see along that set of branches. Once we've done this for each branch that ends at *s*, we then add the results up to get the chain rule for that given situation.

Note that we don't always put the derivatives in the tree. Some of the trees get a little large/messy and so we won't put in the derivatives. Just remember what derivative should be on each branch and you'll be okay without actually writing them down.

Let's write down some chain rules.

Example 4 Use a tree diagram to write down the chain rule for the given derivatives.

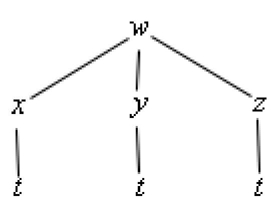
(a) $\dfrac{dw}{dt}$ for $w = f\left(x,y,z\right)$, $x = g_1\left(t\right)$, $y = g_2\left(t\right)$, and $z = g_3\left(t\right)$

(b) $\dfrac{\partial w}{\partial r}$ for $w = f\left(x,y,z\right)$, $x = g_1\left(s,t,r\right)$, $y = g_2\left(s,t,r\right)$, and $z = g_3\left(s,t,r\right)$

Solution

(a) $\dfrac{dw}{dt}$ for $w = f\left(x,y,z\right)$, $x = g_1\left(t\right)$, $y = g_2\left(t\right)$, and $z = g_3\left(t\right)$

So, we'll first need the tree diagram so let's get that.



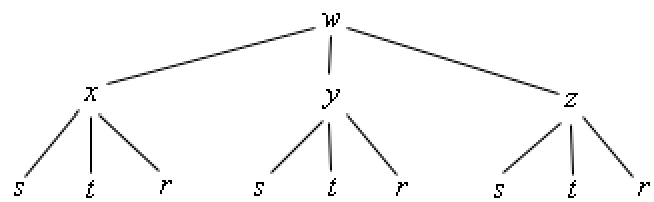
From this it looks like the chain rule for this case should be,

$$\frac{dw}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}$$

which is really just a natural extension to the two variable case that we saw above.

(b) $\frac{\partial w}{\partial r}$ for $w = f(x, y, z)$, $x = g_1(s,t,r)$, $y = g_2(s,t,r)$, and $z = g_3(s,t,r)$

Here is the tree diagram for this situation.



From this it looks like the derivative will be,

$$\frac{\partial w}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial r}$$

So, provided we can write down the tree diagram, and these aren't usually too bad to write down, we can do the chain rule for any set up that we might run across.

We've now seen how to take first derivatives of these more complicated situations, but what about higher order derivatives? How do we do those? It's probably easiest to see how to deal with these with an example.

Example 5 Compute $\frac{\partial^2 f}{\partial \theta^2}$ for $f(x, y)$ if $x = r\cos\theta$ and $y = r\sin\theta$.

Solution

We will need the first derivative before we can even think about finding the second derivative so let's get that. This situation falls into the second case that we looked at above so we don't need a new tree diagram. Here is the first derivative.

$$\begin{aligned}\frac{\partial f}{\partial \theta} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta} \\ &= -r\sin(\theta)\frac{\partial f}{\partial x} + r\cos(\theta)\frac{\partial f}{\partial y}\end{aligned}$$

Okay, now we know that the second derivative is,

$$\frac{\partial^2 f}{\partial \theta^2} = \frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial \theta}\right) = \frac{\partial}{\partial \theta}\left(-r\sin(\theta)\frac{\partial f}{\partial x} + r\cos(\theta)\frac{\partial f}{\partial y}\right)$$

The issue here is to correctly deal with this derivative. Since the two first order derivatives, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$, are both functions of *x* and *y* which are in turn functions of *r* and θ both of these terms are products. So, the using the product rule gives the following,

$$\frac{\partial^2 f}{\partial \theta^2} = -r\cos(\theta)\frac{\partial f}{\partial x} - r\sin(\theta)\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right) - r\sin(\theta)\frac{\partial f}{\partial y} + r\cos(\theta)\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial y}\right)$$

We now need to determine what $\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right)$ and $\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial y}\right)$ will be. These are both chain rule problems again since both of the derivatives are functions of *x* and *y* and we want to take the derivative with respect to θ.

Before we do these let's rewrite the first chain rule that we did above a little.

$$\frac{\partial}{\partial \theta}(f) = -r\sin(\theta)\frac{\partial}{\partial x}(f) + r\cos(\theta)\frac{\partial}{\partial y}(f) \qquad (1)$$

Note that all we've done is change the notation for the derivative a little. With the first chain rule written in this way we can think of (1) as a formula for differentiating any function of *x* and *y* with respect to θ provided we have $x = r\cos\theta$ and $y = r\sin\theta$.

This however is exactly what we need to do the two new derivatives we need above. Both of the first order partial derivatives, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$, are functions of *x* and *y* and $x = r\cos\theta$ and $y = r\sin\theta$ so we can use (1) to compute these derivatives.

To do this we'll simply replace all the *f* 's in (1) with the first order partial derivative that we want to differentiate. At that point all we need to do is a little notational work and we'll get the formula that we're after.

Here is the use of (1) to compute $\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right)$.

$$\begin{aligned}\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right) &= -r\sin(\theta)\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) + r\cos(\theta)\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\\ &= -r\sin(\theta)\frac{\partial^2 f}{\partial x^2} + r\cos(\theta)\frac{\partial^2 f}{\partial y \partial x}\end{aligned}$$

Here is the computation for $\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial y}\right)$.

$$\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right)=-r\sin(\theta)\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)+r\cos(\theta)\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$
$$=-r\sin(\theta)\frac{\partial^2 f}{\partial x\partial y}+r\cos(\theta)\frac{\partial^2 f}{\partial y^2}$$

The final step is to plug these back into the second derivative and do some simplifying.

$$\begin{aligned}\frac{\partial^2 f}{\partial\theta^2}&=-r\cos(\theta)\frac{\partial f}{\partial x}-r\sin(\theta)\left(-r\sin(\theta)\frac{\partial^2 f}{\partial x^2}+r\cos(\theta)\frac{\partial^2 f}{\partial y\partial x}\right)-\\&\qquad r\sin(\theta)\frac{\partial f}{\partial y}+r\cos(\theta)\left(-r\sin(\theta)\frac{\partial^2 f}{\partial x\partial y}+r\cos(\theta)\frac{\partial^2 f}{\partial y^2}\right)\\&=-r\cos(\theta)\frac{\partial f}{\partial x}+r^2\sin^2(\theta)\frac{\partial^2 f}{\partial x^2}-r^2\sin(\theta)\cos(\theta)\frac{\partial^2 f}{\partial y\partial x}-\\&\qquad r\sin(\theta)\frac{\partial f}{\partial y}-r^2\sin(\theta)\cos(\theta)\frac{\partial^2 f}{\partial x\partial y}+r^2\cos^2(\theta)\frac{\partial^2 f}{\partial y^2}\\&=-r\cos(\theta)\frac{\partial f}{\partial x}-r\sin(\theta)\frac{\partial f}{\partial y}+r^2\sin^2(\theta)\frac{\partial^2 f}{\partial x^2}-\\&\qquad 2r^2\sin(\theta)\cos(\theta)\frac{\partial^2 f}{\partial y\partial x}+r^2\cos^2(\theta)\frac{\partial^2 f}{\partial y^2}\end{aligned}$$

It's long and fairly messy but there it is.

The final topic in this section is a revisiting of implicit differentiation. With these forms of the chain rule implicit differentiation actually becomes a fairly simple process. Let's start out with the implicit differentiation that we saw in a Calculus I course.

We will start with a function in the form $F(x,y)=0$ (if it's not in this form simply move everything to one side of the equal sign to get it into this form) where $y=y(x)$. In a Calculus I course we were then asked to compute $\frac{dy}{dx}$ and this was often a fairly messy process. Using the chain rule from this section however we can get a nice simple formula for doing this. We'll start by differentiating both sides with respect to *x*. This will mean using the chain rule on the left side and the right side will, of course, differentiate to zero. Here are the results of that.

$$F_x+F_y\frac{dy}{dx}=0 \qquad\Rightarrow\qquad \frac{dy}{dx}=-\frac{F_x}{F_y}$$

As shown, all we need to do next is solve for $\frac{dy}{dx}$ and we've now got a very nice formula to use for implicit differentiation. Note as well that in order to simplify the formula we switched back to using the subscript notation for the derivatives.

Let's check out a quick example.

Example 6 Find $\frac{dy}{dx}$ for $x\cos(3y) + x^3y^5 = 3x - \mathbf{e}^{xy}$.

Solution

The first step is to get a zero on one side of the equal sign and that's easy enough to do.

$$x\cos(3y) + x^3y^5 - 3x + \mathbf{e}^{xy} = 0$$

Now, the function on the left is $F(x, y)$ in our formula so all we need to do is use the formula to find the derivative.

$$\frac{dy}{dx} = -\frac{\cos(3y) + 3x^2y^5 - 3 + y\mathbf{e}^{xy}}{-3x\sin(3y) + 5x^3y^4 + x\mathbf{e}^{xy}}$$

There we go. It would have taken much longer to do this using the old Calculus I way of doing this.

We can also do something similar to handle the types of implicit differentiation problems involving partial derivatives like those we saw when we first introduced partial derivatives. In these cases we will start off with a function in the form $F(x, y, z) = 0$ and assume that $z = f(x, y)$ and we want to find $\frac{\partial z}{\partial x}$ and/or $\frac{\partial z}{\partial y}$.

Let's start by trying to find $\frac{\partial z}{\partial x}$. We will differentiate both sides with respect to *x* and we'll need to remember that we're going to be treating *y* as a constant. Also, the left side will require the chain rule. Here is this derivative.

$$\frac{\partial F}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0$$

Now, we have the following,

$$\frac{\partial x}{\partial x} = 1 \qquad \text{and} \qquad \frac{\partial y}{\partial x} = 0$$

The first is because we are just differentiating *x* with respect to *x* and we know that is 1. The second is because we are treating the *y* as a constant and so it will differentiate to zero.

Plugging these in and solving for $\frac{\partial z}{\partial x}$ gives,

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}$$

A similar argument can be used to show that,

$$\frac{\partial z}{\partial y} = -\frac{F_y}{F_z}$$

As with the one variable case we switched to the subscripting notation for derivatives to simplify the formulas. Let's take a quick look at an example of this.

Example 7 Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for $x^2 \sin(2y - 5z) = 1 + y\cos(6zx)$.

Solution

This was one of the functions that we used the old implicit differentiation on back in the Partial Derivatives section. You might want to go back and see the difference between the two.

First let's get everything on one side.

$$x^2 \sin(2y - 5z) - 1 - y\cos(6zx) = 0$$

Now, the function on the left is $F(x, y, z)$ and so all that we need to do is use the formulas developed above to find the derivatives.

$$\frac{\partial z}{\partial x} = -\frac{2x\sin(2y - 5z) + 6yz\sin(6zx)}{-5x^2\cos(2y - 5z) + 6yx\sin(6zx)}$$

$$\frac{\partial z}{\partial y} = -\frac{2x^2\cos(2y - 5z) - \cos(6zx)}{-5x^2\cos(2y - 5z) + 6yx\sin(6zx)}$$

If you go back and compare these answers to those that we found the first time around you will notice that they might appear to be different. However, if you take into account the minus sign that sits in the front of our answers here you will see that they are in fact the same.

Section 2-7 : Directional Derivatives

To this point we've only looked at the two partial derivatives $f_x(x,y)$ and $f_y(x,y)$. Recall that these derivatives represent the rate of change of *f* as we vary *x* (holding *y* fixed) and as we vary *y* (holding *x* fixed) respectively. We now need to discuss how to find the rate of change of *f* if we allow both *x* and *y* to change simultaneously. The problem here is that there are many ways to allow both *x* and *y* to change. For instance, one could be changing faster than the other and then there is also the issue of whether or not each is increasing or decreasing. So, before we get into finding the rate of change we need to get a couple of preliminary ideas taken care of first. The main idea that we need to look at is just how are we going to define the changing of *x* and/or *y*.

Let's start off by supposing that we wanted the rate of change of *f* at a particular point, say (x_0, y_0). Let's also suppose that both *x* and *y* are increasing and that, in this case, *x* is increasing twice as fast as *y* is increasing. So, as *y* increases one unit of measure *x* will increase two units of measure.

To help us see how we're going to define this change let's suppose that a particle is sitting at (x_0, y_0) and the particle will move in the direction given by the changing *x* and *y*. Therefore, the particle will move off in a direction of increasing *x* and *y* and the *x* coordinate of the point will increase twice as fast as the *y* coordinate. Now that we're thinking of this changing *x* and *y* as a direction of movement we can get a way of defining the change. We know from Calculus II that vectors can be used to define a direction and so the particle, at this point, can be said to be moving in the direction,

$$\vec{v} = \langle 2,1 \rangle$$

Since this vector can be used to define how a particle at a point is changing we can also use it describe how *x* and/or *y* is changing at a point. For our example we will say that we want the rate of change of *f* in the direction of $\vec{v} = \langle 2,1 \rangle$. In this way we will know that *x* is increasing twice as fast as *y* is. There is still a small problem with this however. There are many vectors that point in the same direction. For instance, all of the following vectors point in the same direction as $\vec{v} = \langle 2,1 \rangle$.

$$\vec{v} = \left\langle \frac{1}{5}, \frac{1}{10} \right\rangle \qquad \vec{v} = \langle 6,3 \rangle \qquad \vec{v} = \left\langle \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right\rangle$$

We need a way to consistently find the rate of change of a function in a given direction. We will do this by insisting that the vector that defines the direction of change be a unit vector. Recall that a unit vector is a vector with length, or magnitude, of 1. This means that for the example that we started off thinking about we would want to use

$$\vec{v} = \left\langle \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right\rangle$$

since this is the unit vector that points in the direction of change.

For reference purposes recall that the magnitude or length of the vector $\vec{v} = \langle a,b,c \rangle$ is given by,

$$\|\vec{v}\| = \sqrt{a^2 + b^2 + c^2}$$

For two dimensional vectors we drop the *c* from the formula.

Sometimes we will give the direction of changing *x* and *y* as an angle. For instance, we may say that we want the rate of change of *f* in the direction of $\theta = \frac{\pi}{3}$. The unit vector that points in this direction is given by,

$$\vec{u} = \langle \cos\theta, \sin\theta \rangle$$

Okay, now that we know how to define the direction of changing *x* and *y* its time to start talking about finding the rate of change of *f* in this direction. Let's start off with the official definition.

Definition

The rate of change of $f(x,y)$ in the direction of the unit vector $\vec{u} = \langle a,b \rangle$ is called the **directional derivative** and is denoted by $D_{\vec{u}} f(x,y)$. The definition of the directional derivative is,

$$D_{\vec{u}} f(x,y) = \lim_{h \to 0} \frac{f(x+ah, y+bh) - f(x,y)}{h}$$

So, the definition of the directional derivative is very similar to the definition of partial derivatives. However, in practice this can be a very difficult limit to compute so we need an easier way of taking directional derivatives. It's actually fairly simple to derive an equivalent formula for taking directional derivatives.

To see how we can do this let's define a new function of a single variable,

$$g(z) = f(x_0 + az, y_0 + bz)$$

where x_0, y_0, *a*, and *b* are some fixed numbers. Note that this really is a function of a single variable now since *z* is the only letter that is not representing a fixed number.

Then by the definition of the derivative for functions of a single variable we have,

$$g'(z) = \lim_{h \to 0} \frac{g(z+h) - g(z)}{h}$$

and the derivative at $z = 0$ is given by,

$$g'(0) = \lim_{h \to 0} \frac{g(h) - g(0)}{h}$$

If we now substitute in for $g(z)$ we get,

$$g'(0) = \lim_{h \to 0} \frac{g(h) - g(0)}{h} = \lim_{h \to 0} \frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h} = D_{\vec{u}} f(x_0, y_0)$$

So, it looks like we have the following relationship.

$$g'(0) = D_{\vec{u}} f(x_0, y_0) \qquad (1)$$

Now, let's look at this from another perspective. Let's rewrite $g(z)$ as follows,

$$g(z) = f(x,y) \quad \text{where} \quad x = x_0 + az \quad \text{and} \quad y = y_0 + bz$$

We can now use the chain rule from the previous section to compute,

$$g'(z) = \frac{dg}{dz} = \frac{\partial f}{\partial x}\frac{dx}{dz} + \frac{\partial f}{\partial y}\frac{dy}{dz} = f_x(x,y)\,a + f_y(x,y)\,b$$

So, from the chain rule we get the following relationship.

$$g'(z) = f_x(x,y)\,a + f_y(x,y)\,b \tag{2}$$

If we now take $z = 0$ we will get that $x = x_0$ and $y = y_0$ (from how we defined *x* and *y* above) and plug these into (2) we get,

$$g'(0) = f_x(x_0,y_0)\,a + f_y(x_0,y_0)\,b \tag{3}$$

Now, simply equate (1) and (3) to get that,

$$D_{\vec{u}}\,f(x_0,y_0) = g'(0) = f_x(x_0,y_0)\,a + f_y(x_0,y_0)\,b$$

If we now go back to allowing *x* and *y* to be any number we get the following formula for computing directional derivatives.

$$\boxed{D_{\vec{u}}\,f(x,y) = f_x(x,y)\,a + f_y(x,y)\,b}$$

This is much simpler than the limit definition. Also note that this definition assumed that we were working with functions of two variables. There are similar formulas that can be derived by the same type of argument for functions with more than two variables. For instance, the directional derivative of $f(x,y,z)$ in the direction of the unit vector $\vec{u} = \langle a,b,c\rangle$ is given by,

$$\boxed{D_{\vec{u}}\,f(x,y,z) = f_x(x,y,z)\,a + f_y(x,y,z)\,b + f_z(x,y,z)\,c}$$

Let's work a couple of examples.

Example 1 Find each of the directional derivatives.

(a) $D_{\vec{u}} f(2,0)$ where $f(x,y) = x\mathbf{e}^{xy} + y$ and $\vec{u}$ is the unit vector in the direction of $\theta = \frac{2\pi}{3}$.

(b) $D_{\vec{u}} f(x,y,z)$ where $f(x,y,z) = x^2 z + y^3 z^2 - xyz$ in the direction of $\vec{v} = \langle -1,0,3 \rangle$.

Solution

(a) $D_{\vec{u}} f(2,0)$ **where** $f(x,y) = x\mathbf{e}^{xy} + y$ **and** $\vec{u}$ **is the unit vector in the direction of** $\theta = \frac{2\pi}{3}$**.**

We'll first find $D_{\vec{u}} f(x,y)$ and then use this a formula for finding $D_{\vec{u}} f(2,0)$. The unit vector giving the direction is,

$$\vec{u} = \left\langle \cos\left(\frac{2\pi}{3}\right), \sin\left(\frac{2\pi}{3}\right) \right\rangle = \left\langle -\frac{1}{2}, \frac{\sqrt{3}}{2} \right\rangle$$

So, the directional derivative is,

$$D_{\vec{u}} f(x,y) = \left(-\frac{1}{2}\right)\left(\mathbf{e}^{xy} + xy\mathbf{e}^{xy}\right) + \left(\frac{\sqrt{3}}{2}\right)\left(x^2\mathbf{e}^{xy} + 1\right)$$

Now, plugging in the point in question gives,

$$D_{\vec{u}} f(2,0) = \left(-\frac{1}{2}\right)(1) + \left(\frac{\sqrt{3}}{2}\right)(5) = \frac{5\sqrt{3}-1}{2}$$

(b) $D_{\vec{u}} f(x,y,z)$ **where** $f(x,y,z) = x^2 z + y^3 z^2 - xyz$ **in the direction of** $\vec{v} = \langle -1,0,3 \rangle$**.**

In this case let's first check to see if the direction vector is a unit vector or not and if it isn't convert it into one. To do this all we need to do is compute its magnitude.

$$\|\vec{v}\| = \sqrt{1+0+9} = \sqrt{10} \neq 1$$

So, it's not a unit vector. Recall that we can convert any vector into a unit vector that points in the same direction by dividing the vector by its magnitude. So, the unit vector that we need is,

$$\vec{u} = \frac{1}{\sqrt{10}}\langle -1,0,3 \rangle = \left\langle -\frac{1}{\sqrt{10}}, 0, \frac{3}{\sqrt{10}} \right\rangle$$

The directional derivative is then,

$$\begin{aligned} D_{\vec{u}} f(x,y,z) &= \left(-\frac{1}{\sqrt{10}}\right)(2xz - yz) + (0)\left(3y^2z^2 - xz\right) + \left(\frac{3}{\sqrt{10}}\right)\left(x^2 + 2y^3 z - xy\right) \\ &= \frac{1}{\sqrt{10}}\left(3x^2 + 6y^3 z - 3xy - 2xz + yz\right) \end{aligned}$$

There is another form of the formula that we used to get the directional derivative that is a little nicer and somewhat more compact. It is also a much more general formula that will encompass both of the formulas above.

Let's start with the second one and notice that we can write it as follows,

$$\begin{aligned} D_{\vec{u}} f(x,y,z) &= f_x(x,y,z)a + f_y(x,y,z)b + f_z(x,y,z)c \\ &= \left\langle f_x, f_y, f_z \right\rangle \bullet \left\langle a,b,c \right\rangle \end{aligned}$$

In other words, we can write the directional derivative as a dot product and notice that the second vector is nothing more than the unit vector $\vec{u}$ that gives the direction of change. Also, if we had used the version for functions of two variables the third component wouldn't be there, but other than that the formula would be the same.

Now let's give a name and notation to the first vector in the dot product since this vector will show up fairly regularly throughout this course (and in other courses). The **gradient of *f*** or **gradient vector of *f*** is defined to be,

$$\nabla f = \left\langle f_x, f_y, f_z \right\rangle \qquad \text{or} \qquad \nabla f = \left\langle f_x, f_y \right\rangle$$

Or, if we want to use the standard basis vectors the gradient is,

$$\nabla f = f_x \vec{i} + f_y \vec{j} + f_z \vec{k} \qquad \text{or} \qquad \nabla f = f_x \vec{i} + f_y \vec{j}$$

The definition is only shown for functions of two or three variables, however there is a natural extension to functions of any number of variables that we'd like.

With the definition of the gradient we can now say that the directional derivative is given by,

$$D_{\vec{u}} f = \nabla f \bullet \vec{u}$$

where we will no longer show the variable and use this formula for any number of variables. Note as well that we will sometimes use the following notation,

$$D_{\vec{u}} f(\vec{x}) = \nabla f \bullet \vec{u}$$

where $\vec{x} = \left\langle x,y,z \right\rangle$ or $\vec{x} = \left\langle x,y \right\rangle$ as needed. This notation will be used when we want to note the variables in some way, but don't really want to restrict ourselves to a particular number of variables. In other words, $\vec{x}$ will be used to represent as many variables as we need in the formula and we will most often use this notation when we are already using vectors or vector notation in the problem/formula.

Let's work a couple of examples using this formula of the directional derivative.

Example 2 Find each of the directional derivative.

(a) $D_{\vec{u}}f\left(\vec{x}\right)$ for $f\left(x,y\right)=x\cos\left(y\right)$ in the direction of $\vec{v}=\left\langle 2,1\right\rangle$.

(b) $D_{\vec{u}}f\left(\vec{x}\right)$ for $f\left(x,y,z\right)=\sin\left(yz\right)+\ln\left(x^2\right)$ at $\left(1,1,\pi\right)$ in the direction of $\vec{v}=\left\langle 1,1,-1\right\rangle$.

Solution

(a) $D_{\vec{u}}f\left(\vec{x}\right)$ **for** $f\left(x,y\right)=x\cos\left(y\right)$ **in the direction of** $\vec{v}=\left\langle 2,1\right\rangle$**.**

Let's first compute the gradient for this function.

$$\nabla f=\left\langle \cos\left(y\right),-x\sin\left(y\right)\right\rangle$$

Also, as we saw earlier in this section the unit vector for this direction is,

$$\vec{u}=\left\langle \frac{2}{\sqrt{5}},\frac{1}{\sqrt{5}}\right\rangle$$

The directional derivative is then,

$$\begin{aligned} D_{\vec{u}}f\left(\vec{x}\right)&=\left\langle \cos\left(y\right),-x\sin\left(y\right)\right\rangle\bullet\left\langle \frac{2}{\sqrt{5}},\frac{1}{\sqrt{5}}\right\rangle\\ &=\frac{1}{\sqrt{5}}\left(2\cos\left(y\right)-x\sin\left(y\right)\right)\end{aligned}$$

(b) $D_{\vec{u}}f\left(\vec{x}\right)$ **for** $f\left(x,y,z\right)=\sin\left(yz\right)+\ln\left(x^2\right)$ **at** $\left(1,1,\pi\right)$ **in the direction of** $\vec{v}=\left\langle 1,1,-1\right\rangle$**.**

In this case are asking for the directional derivative at a particular point. To do this we will first compute the gradient, evaluate it at the point in question and then do the dot product. So, let's get the gradient.

$$\begin{aligned}\nabla f\left(x,y,z\right)&=\left\langle \frac{2}{x},z\cos\left(yz\right),y\cos\left(yz\right)\right\rangle\\ \nabla f\left(1,1,\pi\right)&=\left\langle \frac{2}{1},\pi\cos\left(\pi\right),\cos\left(\pi\right)\right\rangle=\left\langle 2,-\pi,-1\right\rangle\end{aligned}$$

Next, we need the unit vector for the direction,

$$\left\|\vec{v}\right\|=\sqrt{3}\qquad\qquad \vec{u}=\left\langle \frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},-\frac{1}{\sqrt{3}}\right\rangle$$

Finally, the directional derivative at the point in question is,

$$D_{\vec{u}} f(1,1,\pi) = \langle 2, -\pi, -1 \rangle \bullet \left\langle \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}} \right\rangle$$
$$= \frac{1}{\sqrt{3}}(2 - \pi + 1)$$
$$= \frac{3-\pi}{\sqrt{3}}$$

Before proceeding let's note that the first order partial derivatives that we were looking at in the majority of the section can be thought of as special cases of the directional derivatives. For instance, f_x can be thought of as the directional derivative of *f* in the direction of $\vec{u} = \langle 1, 0 \rangle$ or $\vec{u} = \langle 1, 0, 0 \rangle$, depending on the number of variables that we're working with. The same can be done for f_y and f_z

We will close out this section with a couple of nice facts about the gradient vector. The first tells us how to determine the maximum rate of change of a function at a point and the direction that we need to move in order to achieve that maximum rate of change.

Theorem

The maximum value of $D_{\vec{u}} f(\vec{x})$ (and hence then the maximum rate of change of the function $f(\vec{x})$) is given by $\left\| \nabla f(\vec{x}) \right\|$ and will occur in the direction given by $\nabla f(\vec{x})$.

Proof

This is a really simple proof. First, if we start with the dot product form $D_{\vec{u}} f(\vec{x})$ and use a nice fact about dot products as well as the fact that $\vec{u}$ is a unit vector we get,

$$D_{\vec{u}} f = \nabla f \bullet \vec{u} = \left\| \nabla f \right\| \left\| \vec{u} \right\| \cos\theta = \left\| \nabla f \right\| \cos\theta$$

where θ is the angle between the gradient and $\vec{u}$.

Now the largest possible value of $\cos\theta$ is 1 which occurs at $\theta = 0$. Therefore the maximum value of $D_{\vec{u}} f(\vec{x})$ is $\left\| \nabla f(\vec{x}) \right\|$ Also, the maximum value occurs when the angle between the gradient and $\vec{u}$ is zero, or in other words when $\vec{u}$ is pointing in the same direction as the gradient, $\nabla f(\vec{x})$.

Let's take a quick look at an example.

Example 3 Suppose that the height of a hill above sea level is given by $z = 1000 - 0.01x^2 - 0.02y^2$. If you are at the point $(60,100)$ in what direction is the elevation changing fastest? What is the maximum rate of change of the elevation at this point?

Solution
First, you will hopefully recall from the Quadric Surfaces section that this is an elliptic paraboloid that opens downward. So even though most hills aren't this symmetrical it will at least be vaguely hill shaped and so the question makes at least a little sense.

Now on to the problem. There are a couple of questions to answer here, but using the theorem makes answering them very simple. We'll first need the gradient vector.

$$\nabla f(\vec{x}) = \langle -0.02x, -0.04y \rangle$$

The maximum rate of change of the elevation will then occur in the direction of

$$\nabla f(60,100) = \langle -1.2, -4 \rangle$$

The maximum rate of change of the elevation at this point is,

$$\left\| \nabla f(60,100) \right\| = \sqrt{(-1.2)^2 + (4)^2} = \sqrt{17.44} = 4.176$$

Before leaving this example let's note that we're at the point $(60,100)$ and the direction of greatest rate of change of the elevation at this point is given by the vector $\langle -1.2, -4 \rangle$. Since both of the components are negative it looks like the direction of maximum rate of change points up the hill towards the center rather than away from the hill.

The second fact about the gradient vector that we need to give in this section will be very convenient in some later sections.

Fact

The gradient vector $\nabla f(x_0, y_0)$ is orthogonal (or perpendicular) to the level curve $f(x,y) = k$ at the point (x_0, y_0). Likewise, the gradient vector $\nabla f(x_0, y_0, z_0)$ is orthogonal to the level surface $f(x,y,z) = k$ at the point (x_0, y_0, z_0).

Proof

We're going to do the proof for the $\mathbb{R}^3$ case. The proof for the $\mathbb{R}^2$ case is identical. We'll also need some notation out of the way to make life easier for us let's let *S* be the level surface given by $f(x,y,z) = k$ and let $P = (x_0, y_0, z_0)$. Note as well that *P* will be on *S*.

Now, let *C* be any curve on *S* that contains *P*. Let $\vec{r}\left(t\right)=\left\langle x\left(t\right),y\left(t\right),z\left(t\right)\right\rangle$ be the vector equation for *C* and suppose that t_0 be the value of *t* such that $\vec{r}\left(t_0\right)=\left\langle x_0,y_0,z_0\right\rangle$. In other words, t_0 be the value of *t* that gives *P*.

Because *C* lies on *S* we know that points on *C* must satisfy the equation for *S*. Or,

$$f\left(x(t),y(t),z(t)\right)=k$$

Next, let's use the Chain Rule on this to get,

$$\frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}+\frac{\partial f}{\partial z}\frac{dz}{dt}=0$$

Notice that $\nabla f=\left\langle f_x,f_y,f_z\right\rangle$ and $\vec{r}'\left(t\right)=\left\langle x'\left(t\right),y'\left(t\right),z'\left(t\right)\right\rangle$ so this becomes,

$$\nabla f\bullet\vec{r}'\left(t\right)=0$$

At, $t=t_0$ this is,

$$\nabla f\left(x_0,y_0,z_0\right)\bullet\vec{r}'\left(t_0\right)=0$$

This then tells us that the gradient vector at *P* , $\nabla f\left(x_0,y_0,z_0\right)$, is orthogonal to the tangent vector, $\vec{r}'\left(t_0\right)$, to any curve *C* that passes through *P* and on the surface *S* and so must also be orthogonal to the surface *S*.

As we will be seeing in later sections we are often going to be needing vectors that are orthogonal to a surface or curve and using this fact we will know that all we need to do is compute a gradient vector and we will get the orthogonal vector that we need. We will see the first application of this in the next chapter.