Differentiation of Functions

1 Introduction

Beginning calculus students identify the derivative of a function either in terms of slope or instantaneous rate of change. When thinking of the former they say something like, "$f'(a)$ is the slope of the line that is tangent to the graph of the function f at the point $(a, f(a))$." Can this idea be extended to functions $f : \mathbb{R}^n \to \mathbb{R}$? For example, Figure 1 shows the graph of $f(x, y) = \cos x \sin y$ for $(x, y) \in [-\pi, \pi] \times [-\pi, \pi]$. The point $(0.25, 1.25, f(0.25, 1.25)) \approx (0.25, 1.25, 0.92)$ is represented by the tiny dark circle. What would it even mean to talk about a slope at that point?

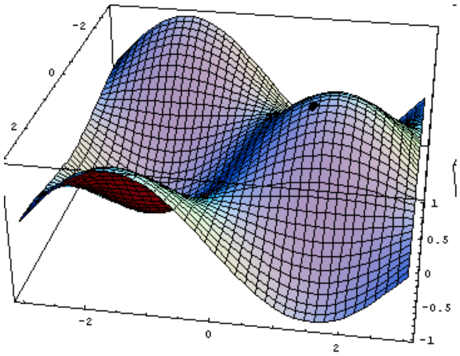

Figure 1

Given a function $f : \mathbb{R}^n \to \mathbb{R}$, we might be tempted to take an algebraic approach in defining the derivative. For functions $f : \mathbb{R} \to \mathbb{R}$ we know that $f'(x_0) = \lim\limits_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0}$, or (equivalently) $f'(x_0) = \lim\limits_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h}$. What would be wrong with defining the derivative at $\overrightarrow{x_0}$ of a function $f : \mathbb{R}^n \to \mathbb{R}$ as $f'\left(\overrightarrow{x_0}\right) = \lim\limits_{\overrightarrow{x} \to \overrightarrow{x_0}} \frac{f(\overrightarrow{x})-f(\overrightarrow{x_0})}{\overrightarrow{x}-\overrightarrow{x_0}}$, or $f'\left(\overrightarrow{x_0}\right) = \lim\limits_{\overrightarrow{h} \to \overrightarrow{0}} \frac{f\left(\overrightarrow{x_0}+\overrightarrow{h}\right)-f(\overrightarrow{x_0})}{\overrightarrow{h}}$?

To answer this question, pay attention to what kind of objects are in the numerator and denominator. The numerator is a number and the denominator is a vector. There is no mechanism for dividing numbers by vectors, so the last expression is meaningless.

Thus, our standard algebraic approach does not generalize to higher dimensions. If we are to get an algebraic rule for the derivative of functions $f : \mathbb{R}^n \to \mathbb{R}$, we will need to revisit elementary calculus and see if we can formulate the derivative in a way that *will* generalize. We begin by looking at functions that are *not* differentiable. Consider the function $f(x) = (x-2)^{\frac{2}{3}} + 1$, where $-5 \le x \le 5$. Figure 2 shows the graph of this function.

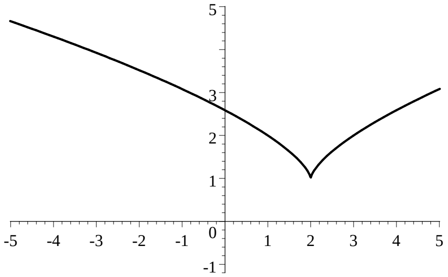

Figure 2

Why isn't $f(x) = (x-2)^{\frac{2}{3}} + 1$ differentiable when $x = 2$? A common answer is that there is no tangent line to the graph of f at the point $(2, f(2)) = (2, 1)$. But this response seems strange, because there are certainly many tangent lines we can draw at that point.

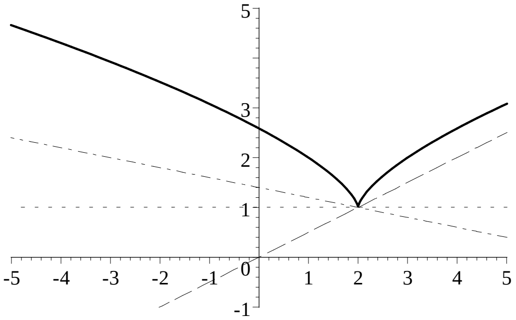

Figure 3

What is the difference between this situation and the case of a function that *is* differentiable? Take this same function, but at the point $x = 1$. We readily compute $f'(x) = \frac{2}{3}(x-2)^{-\frac{1}{3}}$, $f'(1) = -\frac{2}{3}$, and look at the above graph with the tangent line drawn at $(1, 2)$.

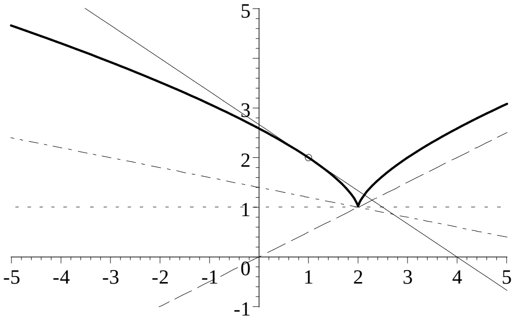

Figure 4

The difference between the tangent line at $x = 1$ and any tangent line we draw at $x = 2$ can be explained by how well the tangent lines "fit" the function at the points in question. For points near $x = 1$, the (solid) tangent line is a very good approximation to the function. For points near $x = 2$, none of the (dotted) tangent lines is an especially good approximation to the function. This notion (goodness of fit) can be translated into algebraic terms, and in such a way that the translation *does* generalize to functions of more than one variable.

The key idea is to look at the limit definition of the derivative in the language of approximations. When we write $f'(x_0) = \lim\limits_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0}$, we mean that if x is sufficiently close to x_0, then $f'(x_0)$ approximately equals $\frac{f(x)-f(x_0)}{x-x_0}$ (Notation: $f'(x_0) \approx \frac{f(x)-f(x_0)}{x-x_0}$). Put another way, if x is sufficiently close to x_0, then $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$. Now, $y = f(x_0) + f'(x_0)(x - x_0)$ is a linear function, and it is the equation of the line tangent to the graph of the function f at the point $(x_0, f(x_0))$. Using the example of $f(x) = (x-2)^{\frac{2}{3}} + 1$ and $x_0 = 1$, we compute the equation of the tangent line to be $y = 2 - \frac{2}{3}(x-1)$. In general, let's designate the linear function we get by $L(x)$, so that $L(x) = f(x_0) + f'(x_0)(x - x_0)$.

When x is near x_0, the function f is only approximated by the function L (unless f is itself a linear function, in which case f would be identical with L). What we need is a way to decide when the approximation is good enough to warrant our dubbing the function with the lofty title of "differentiable at x_0." Note that $f(x) - L(x)$ is the difference in y-values between the function and the tangent line at the point x. We designate this difference by $e(x)$, and call $e(x)$ the error term. We can view this difference pictorially in Figures 5 and 6, using the points $x_0 = 1$ and $x_0 = 2$, respectively. (The dotted line is the tangent line, and the thinner line is the graph of $e(x)$.)

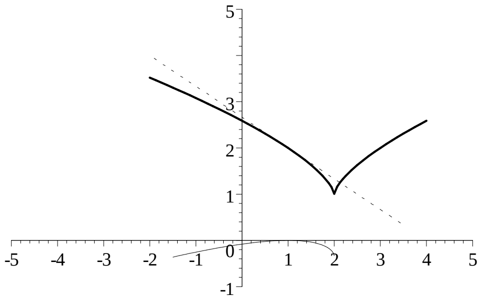

Figure 5

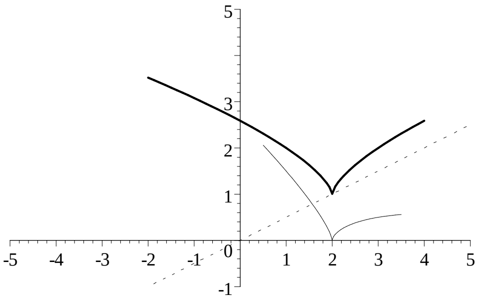

Figure 6

In Figure 5, as $x \to x_0$ the error term $e(x)$ goes to zero much more quickly than does the distance between x and x_0. In other words, $\lim\limits_{x \to x_0} \frac{e(x)}{|x-x_0|} = 0$. This is not the case in Figure 6. To see how this relates to derivatives, note that, for $x \neq x_0$,

$$\begin{aligned}
\frac{e(x)}{x-x_0} &= \frac{f(x)-L(x)}{x-x_0} \\
&= \frac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{x-x_0} \\
&= \frac{f(x)-f(x_0)}{x-x_0} - \frac{f'(x_0)(x-x_0)}{x-x_0} \\
&= \frac{f(x)-f(x_0)}{x-x_0} - f'(x_0),
\end{aligned}$$

and that if f is differentiable at x_0 the last expression *will* approach zero as x approaches x_0.

Thus, a function $f : \mathbb{R} \to \mathbb{R}$ is differentiable at the point x_0 just in case there exists a linear function $L : \mathbb{R} \to \mathbb{R}$ such that $f(x) = L(x) + e(x)$, where $\lim\limits_{x \to x_0} \frac{e(x)}{|x-x_0|} = 0$. In other words, there must be a linear function that serves as a very good approximation to the function f for points x near x_0.

Exercise 1 *Find the functions* $L(x)$ *and* $e(x)$ *that show* $f(x) = x^2$ *is differentiable at the point* $x_0 = 3$. *Be sure to justify your assertions by showing* $\lim\limits_{x \to x_0} \frac{e(x)}{|x-x_0|} = 0$.

2 Derivatives in Higher Dimensions

We are now ready to define differentiability in higher dimensions. We focus on $f : \mathbb{R}^2 \to \mathbb{R}$ and leave as an exercise the corresponding definition for n dimensions. Recall that in $\mathbb{R}^2$ a linear function describes a plane and has the general form $L(x, y) = ax + by + c$.

Definition 1 $f : \mathbb{R}^2 \to \mathbb{R}$ *is differentiable at* $\overrightarrow{x_0} \in \mathbb{R}^2$ *provided there exists a linear function* $L(\overrightarrow{x}) = L(x, y) = ax + by + c$ *such that* $f(\overrightarrow{x}) = L(\overrightarrow{x}) + e(\overrightarrow{x})$, *where* $\lim\limits_{\overrightarrow{x} \to \overrightarrow{x_0}} \frac{e(\overrightarrow{x})}{\|\overrightarrow{x} - \overrightarrow{x_0}\|} = 0$.

Geometrically, a function $f : \mathbb{R}^2 \to \mathbb{R}$ is differentiable at the point $\overrightarrow{x_0}$ when there exists a tangent plane that serves as a good approximation to the function f for points $\overrightarrow{x}$ near $\overrightarrow{x_0}$. Figure 7 shows the tangent plane to the graph of $f(x, y) = \cos x \sin y$ at the point $(0.25, 1.25, 0.92)$. If $\overrightarrow{x} = (x, y)$ is near the point $\overrightarrow{x_0} = (0.25, 1.25)$ then $L(x, y)$ is very close to $f(x, y)$, and so the point $(x, y, L(x, y))$ on the tangent plane is very close to the corresponding point $(x, y, f(x, y))$ on the graph of f.

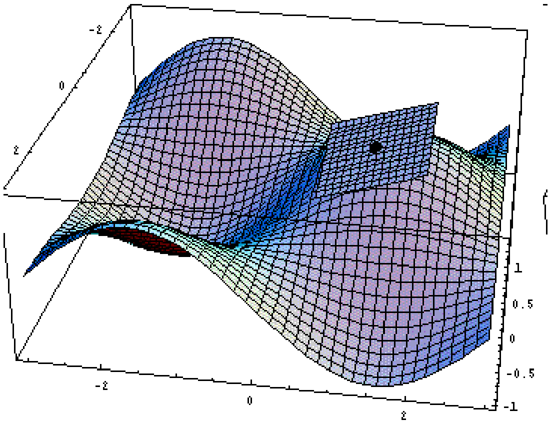

Figure 7

Exercise 2 *Define what it means to say that* $f : \mathbb{R}^n \to \mathbb{R}$ *is differentiable at the point* $\overrightarrow{x_0} \in \mathbb{R}^n$.

In the next section we will see (given a function $f : \mathbb{R}^n \to \mathbb{R}$) how to construct the corresponding function L. We close here with a verification that a given linear function L (meant to serve as an approximation to a function f at a particular point) does indeed satisfy Definition 1.

Example 1 *Show that the function* $f\left(\overrightarrow{x}\right) = f(x, y) = x^2 + 2y^3 + 6$ *is differentiable at* $\overrightarrow{x_0} = (2, 1)$ *by showing that* $L\left(\overrightarrow{x}\right) = L(x, y) = 4x + 6y - 2$ *satisfies the requirements of Definition 1.*

Solution. We compute

$$\begin{aligned}
0 &\le \lim_{\overrightarrow{x} \to \overrightarrow{x_0}} \frac{\left|e\left(\overrightarrow{x}\right)\right|}{\left\|\overrightarrow{x} - \overrightarrow{x_0}\right\|} \\
&= \lim_{\overrightarrow{x} \to \overrightarrow{x_0}} \frac{\left|f\left(\overrightarrow{x}\right) - L\left(\overrightarrow{x}\right)\right|}{\left\|\overrightarrow{x} - \overrightarrow{x_0}\right\|} \\
&= \lim_{\langle x,y\rangle \to \langle 2,1\rangle} \frac{\left|\left(x^2 + 2y^3 + 6\right) - (4x + 6y - 2)\right|}{\left\|\langle x, y\rangle - \langle 2, 1\rangle\right\|} \\
&= \lim_{\langle x,y\rangle \to \langle 2,1\rangle} \frac{\left|x^2 - 4x + 2y^3 - 6y + 8\right|}{\sqrt{(x-2)^2 + (y-1)^2}} \\
&= \lim_{\langle x,y\rangle \to \langle 2,1\rangle} \frac{(x-2)^2 + 2(y-1)^2(y+2)}{\sqrt{(x-2)^2 + (y-1)^2}} \quad \text{(Why can we remove the absolute value signs?)} \\
&\le \lim_{\langle x,y\rangle \to \langle 2,1\rangle} \frac{(x-2)^2 + 8(y-1)^2}{\sqrt{(x-2)^2 + (y-1)^2}} \quad \text{(Provided } y \text{ is close to 1–Why?)} \\
&\le \lim_{\langle x,y\rangle \to \langle 2,1\rangle} \frac{8(x-2)^2 + 8(y-1)^2}{\sqrt{(x-2)^2 + (y-1)^2}} \\
&= \lim_{\langle x,y\rangle \to \langle 2,1\rangle} 8\sqrt{(x-2)^2 + (y-1)^2} = 0.
\end{aligned}$$

From these inequalities it follows that $0 \leq \lim_{\vec{x} \to \vec{x_0}} \frac{|e(\vec{x})|}{\|\vec{x} - \vec{x_0}\|} \leq 0$. The only way this inequality can be satisfied is to have $\lim_{\vec{x} \to \vec{x_0}} \frac{|e(\vec{x})|}{\|\vec{x} - \vec{x_0}\|} = 0$, and this condition is sufficient to guarantee that $\lim_{\vec{x} \to \vec{x_0}} \frac{e(\vec{x})}{\|\vec{x} - \vec{x_0}\|} = 0$.

Exercise 3 *Answer the two questions posed in the above solution.*

Exercise 4 *Show that the function* $f\left(\vec{x}\right) = f(x, y) = 3x^2 + y + 2$ *is differentiable at* $\vec{x_0} = (2, 1)$ *by showing that* $L\left(\vec{x}\right) = L(x, y) = 12x + y - 10$ *satisfies the requirements of Definition 1.*

3 Finding the Derivative

3.1 Motivation

In the last section we defined *when* a function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\vec{x_0}$, namely, when there is a linear function $L : \mathbb{R}^n \to \mathbb{R}$ that is a “good” approximation to f for points $\vec{x}$ near $\vec{x_0}$. We never said, however, *what* the derivative was. For functions $f : \mathbb{R} \to \mathbb{R}$, we have

$$f(x) \approx L(x) = f(x_0) + f'(x_0)(x - x_0), \tag{1}$$

and the derivative itself is the quantity $f'(x_0)$. In this section we seek an analogous expression for L in higher dimensions.

We begin by looking at a function $f : \mathbb{R}^2 \to \mathbb{R}$ that is differentiable at $\vec{x_0}$. This means there is a function $L\left(\vec{x}\right) = L(x, y) = ax + by + c$ that satisfies the requirements of Definition 1. Now, L is going to be the equation of a tangent plane that touches the graph of the function f at the point of tangency, and hence $f(x_0, y_0) = L(x_0, y_0) = ax_0 + by_0 + c$. Therefore,

$$\begin{aligned} L(x, y) &= (ax_0 + by_0 + c) + (ax + by + c) - (ax_0 + by_0 + c), \text{ or} \\ L(x, y) &= f(x_0, y_0) + \langle a, b \rangle \bullet \langle x - x_0, y - y_0 \rangle, \text{ or} \\ L\left(\vec{x}\right) &= f\left(\vec{x_0}\right) + \langle a, b \rangle \bullet \left(\vec{x} - \vec{x_0}\right). \end{aligned} \tag{2}$$

Equations (1) and (2) jointly clarify *what* the derivative of a function $f : \mathbb{R}^2 \to \mathbb{R}$ should be. It should be a *vector*. In fact, it should equal the vector $\langle a, b \rangle$, and our aesthetic convictions almost compel us to write $\vec{f}\,'\left(\vec{x_0}\right) = \langle a, b \rangle$, because then we have

$$f\left(\vec{x}\right) \approx L\left(\vec{x}\right) = f\left(\vec{x_0}\right) + \vec{f}\,'\left(\vec{x_0}\right) \bullet \left(\vec{x} - \vec{x_0}\right), \tag{3}$$

which is virtually the same form as Equation (1). How sweet it is!

3.2 Calculation Strategy

The notational serendipity of Equation (3) is all well and good, but somewhat useless unless we can come up with a convenient way to determine what the vector $\langle a, b \rangle$ is for a given function $f(x, y)$ at a particular point $\vec{x_0} = (x_0, y_0)$. Referring to Example (1), where $L(x, y) = 4x + 6y - 2$ (and therefore $\vec{f}\,'(2, 1) = \langle 4, 6 \rangle$), we seek a way to calculate the vector $\langle 4, 6 \rangle$.

According to Definition 1, if $f:\mathbb{R}^2\to\mathbb{R}$ is differentiable at $\overrightarrow{x_0}$, then

$$\lim_{\overrightarrow{x}\to\overrightarrow{x_0}}\frac{e\left(\overrightarrow{x}\right)}{\left\|\overrightarrow{x}-\overrightarrow{x_0}\right\|}=0. \tag{4}$$

Keeping in mind that $e\left(\overrightarrow{x}\right)=f\left(\overrightarrow{x}\right)-L\left(\overrightarrow{x}\right)$, and using Equation (2), we see that Equation (4) reduces to

$$\begin{aligned}
0 &= \lim_{\overrightarrow{x}\to\overrightarrow{x_0}}\frac{f\left(\overrightarrow{x}\right)-f\left(\overrightarrow{x_0}\right)-\langle a,b\rangle\bullet\left(\overrightarrow{x}-\overrightarrow{x_0}\right)}{\left\|\overrightarrow{x}-\overrightarrow{x_0}\right\|}\\
&= \lim_{(x,y)\to(x_0,y_0)}\frac{f(x,y)-f(x_0,y_0)-a(x-x_0)-b(y-y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}}\\
&= \lim_{(x,y)\to(x_0,y_0)}\left[\frac{f(x,y)-f(x_0,y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}}-\frac{a(x-x_0)+b(y-y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}}\right],\text{ so}
\end{aligned}$$

$$\lim_{(x,y)\to(x_0,y_0)}\frac{f(x,y)-f(x_0,y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}}=\lim_{(x,y)\to(x_0,y_0)}\frac{a(x-x_0)+b(y-y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}}. \tag{5}$$

Because the limit exists, it equals the same value no matter how (x,y) approaches (x_0,y_0), and we can judiciously choose two different paths of approach that will reveal how to calculate the vector $\langle a,b\rangle$. First, we let (x,y) approach (x_0,y_0) along a line parallel to the x-axis, keeping the y-value fixed at y_0, so that Equation (5) reduces to

$$\begin{aligned}
\lim_{(x,y_0)\to(x_0,y_0)}\frac{f(x,y_0)-f(x_0,y_0)}{\sqrt{(x-x_0)^2}} &= \lim_{(x,y_0)\to(x_0,y_0)}\frac{a(x-x_0)}{\sqrt{(x-x_0)^2}},\text{ or}\\
\lim_{(x,y_0)\to(x_0,y_0)}\frac{f(x,y_0)-f(x_0,y_0)}{|x-x_0|} &= \lim_{(x,y_0)\to(x_0,y_0)}\frac{a(x-x_0)}{|x-x_0|}. \qquad (6)
\end{aligned}$$

The expression $|x-x_0|$ equals $(x-x_0)$ when $x>x_0$ and $-(x-x_0)$ when $x<x_0$. Using this fact, it is possible to show (we leave as an exercise) that

$$\lim_{(x,y_0)\to(x_0,y_0)}\frac{f(x,y_0)-f(x_0,y_0)}{x-x_0}=a. \tag{7}$$

Exercise 5 *Show that Equation (7) is a consequence of Equation (6).*

Similarly, if we let (x,y) approach (x_0,y_0) along a line parallel to the y-axis, keeping the x-value fixed at x_0, then Equation (5) reduces to

$$\lim_{(x_0,y)\to(x_0,y_0)}\frac{f(x_0,y)-f(x_0,y_0)}{y-y_0}=b. \tag{8}$$

If we stare at Equation (7) long enough, we may recognize that the left side is really a friend from beginning calculus. Because y_0 is fixed, the only variable that is changing is x, so the left side is really the plain old derivative in disguise, and could be written as

$$\lim_{x\to x_0}\frac{f(x,y_0)-f(x_0,y_0)}{x-x_0},\text{ or} \tag{9}$$

$$\lim_{h\to 0}\frac{f(x_0+h,y_0)-f(x_0,y_0)}{h}. \tag{10}$$

Example 2 *Apply Expression (9) to the function of Example (1).*

Solution. We calculate $\lim_{x\to x_0} \frac{f(x,y_0)-f(x_0,y_0)}{x-x_0} = \lim_{x\to x_0} \frac{(x^2+2y_0^3+6)-(x_0^2+2y_0^3+6)}{x-x_0} = \lim_{x\to x_0} \frac{x^2-x_0^2}{x-x_0} = \lim_{x\to x_0} (x+x_0) = 2x_0$. When $(x_0, y_0) = (2, 1)$, we get $2x_0 = 4$.

Remark. In Example (1) the linear function approximating $f(x, y) = x^2 + 2y^3 + 6$ was $L(x, y) = ax + by + c = 4x + 6y - 2$. Thus, $\langle a, b\rangle = \langle 4, 6\rangle$, and Expression (9) correctly produced the coefficient $a = 4$, as advertised.

Expressions (9) and (10) are so important that they form the basis for the following definitions.

Definition 2 *We call* $\lim_{h\to 0} \frac{f(x_0+h,y_0)-f(x_0,y_0)}{h}$ *the partial derivative of f with respect to x evaluated at the point* (x_0, y_0), *and we write*

$$f_x(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h}.$$

Definition 3 *Similarly, we call* $\lim_{h\to 0} \frac{f(x_0,y_0+h)-f(x_0,y_0)}{h}$ *the partial derivative of f with respect to y evaluated at the point* (x_0, y_0), *and we write*

$$f_y(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0, y_0+h) - f(x_0, y_0)}{h}.$$

Remark. There is nothing magic about the point (x_0, y_0), and we could just as well talk about the partials evaluated at (x, y), (u, v), or any point in $\mathbb{R}^2$. In other words,

$$\begin{aligned} f_x(x, y) &= \lim_{h\to 0} \frac{f(x+h, y) - f(x, y)}{h} \quad \text{and} \\ f_y(x, y) &= \lim_{h\to 0} \frac{f(x, y+h) - f(x, y)}{h}, \end{aligned}$$

provided, of course, that the limits exist.

Exercise 6 *Use Definition 3 to verify that the coefficient b for the function of Example (1) is indeed equal to 6.*

If $z = f(x, y)$, you will see many of the following notations to indicate partial derivatives.

$$\begin{aligned} f_x(x, y) &= \frac{\partial z}{\partial x} = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} f(x, y) = f_1(x, y) = f^{(1,0)}(x, y) = D_1 f = D_x f, \\ f_y(x, y) &= \frac{\partial z}{\partial y} = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y} f(x, y) = f_2(x, y) = f^{(0,1)}(x, y) = D_2 f = D_y f. \end{aligned}$$

The good news in doing calculations comes from realizing that partial derivatives are nothing more than derivatives. When we take the partial derivative with respect to x, the y-variable stays fixed, so we can treat it as if it were a constant. Similarly, the x-variable acts like a constant when we compute the partial derivative with respect to y.

Example 3 $\frac{\partial}{\partial x}\left(x^5 y^3 + e^{xy}\right) = 5x^4 y^3 + y e^{xy}$.

Example 4 *Evaluate* $f_x(0.25, 1.25)$ *and* $f_y(0.25, 1.25)$ *for the function* $f(x, y) = \cos x \sin y$.

Solution. $f_x(x,y) = -\sin x \sin y$, so $f_x(0.25, 1.25) = -\sin(0.25)\sin(1.25) \approx -0.23$; $f_y(x,y) = \cos x \cos y$, so $f_y(0.25, 1.25) = \cos(0.25)\cos(1.25) \approx 0.31$.

Example 5 *What is the equation of the plane tangent to the graph of* $f(x,y) = \cos x \sin y$ *at the point* $(0.25, 1.25, f(0.25, 1.25)) \approx (0.25, 1.25, 0.92)$?

Solution. The tangent plane is the linear function L given by [see Equation (3)]

$$\begin{aligned} L(x,y) &= f(0.25, 1.25) + \vec{f}\,'(0.25, 1.25) \bullet \langle x - 0.25, y - 1.25 \rangle \\ &\approx 0.92 + \langle -0.23, 0.31 \rangle \bullet \langle x - 0.25, y - 1.25 \rangle \\ &= 0.92 + (-0.23)(x - 0.25) + (0.31)(y - 1.25) \\ &= -0.23x + 0.31y + 0.59. \end{aligned}$$

Exercise 7 *Find the equation of the plane tangent to the graph of* $f(x,y) = x^2 + y^2 + x^2 y$ *at the point* $(2, 3, f(2,3)) = (2, 3, 25)$.

Exercise 8 *Find* $\frac{\partial f}{\partial x}$ *and* $\frac{\partial f}{\partial y}$ *if* $f(x,y) = xe^{-x^2-y^2}$.

3.3 Geometric Interpretation

Imagine the graph a function $f : \mathbb{R}^2 \to \mathbb{R}$, and slice this graph with the plane $y = y_0$. Figure 8 illustrates this process with the function $f(x,y) = 7 - x^2 - y^2$, and the plane $y = 1$.

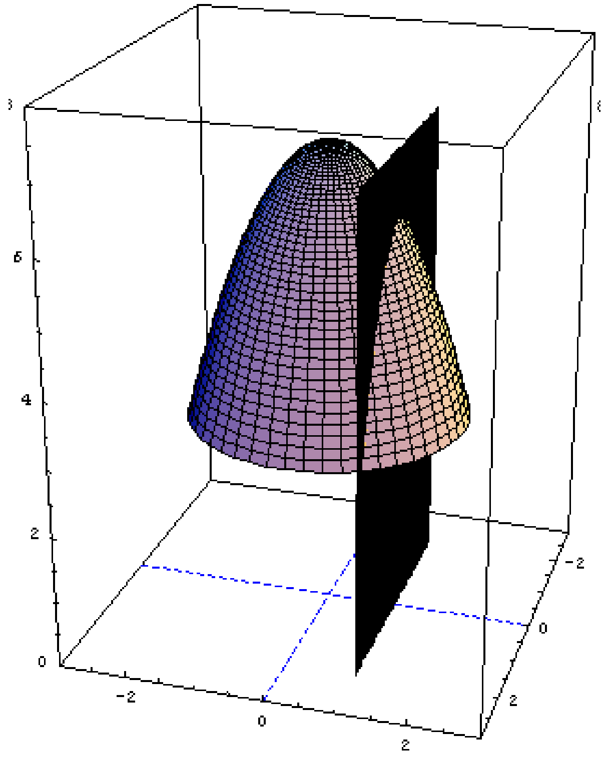

Figure 8

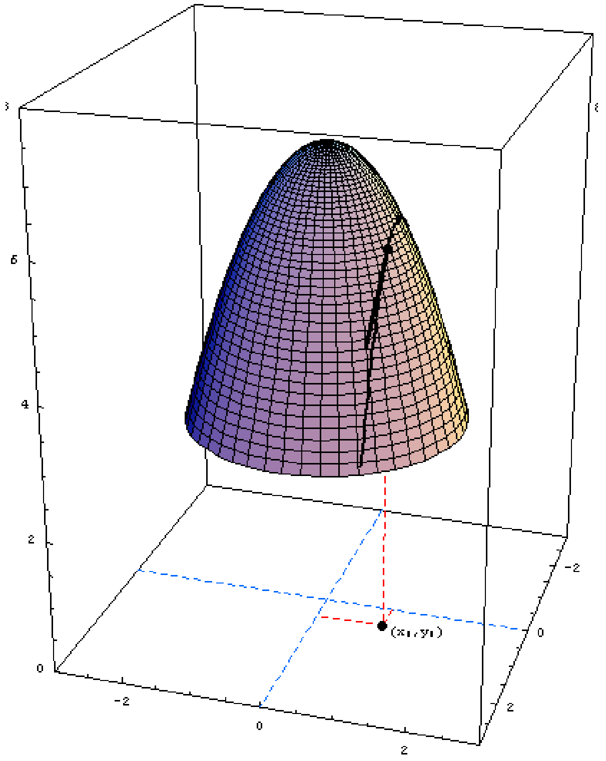

Figure 9

The plane intersects the graph in a surface that is traced out by the vector valued function $\vec{r}(t) = \langle t, y_0, f(t, y_0) \rangle$, for $t \in [a,b]$. Figure 9 illustrates this idea for $f(x,y) = 7 - x^2 - y^2$, with $y_0 = 1$. The dark curved line on the surface of the graph is the path traced by $\vec{r}\,(t) = \langle t, 1, f\,(t, 1)\rangle$ for $-1.7 \leq t \leq 1.7$, so $\vec{r}\,(t) = \left\langle t, 1, 6 - t^2 \right\rangle$. In general, $f_x\,(x_0, y_0)$ is the derivative with respect to x of the function $f\,(x, y_0)$ evaluated at $x = x_0$. It is also the derivative of the third component of the vector valued function $\vec{r}\,(t) = \langle t, y_0, f\,(t, y_0)\rangle$ evaluated at $t = x_0$. In other words, $\vec{r}\,'(x_0) = \left\langle 1, 0, \frac{d}{dt} f\,(t, y_0)\,\big|_{t=x_0} \right\rangle = \langle 1, 0, f_x\,(x_0, y_0)\rangle$. For our specific example, $\vec{r}\,'(t) = \langle 1, 0, -2t\rangle$, so $\vec{r}\,'(x_0) = \langle 1, 0, -2x_0\rangle$. Now, $f_x\,(x, y) = -2x$, so if we set $(x_0, y_0) = \left(\frac{1}{2}, 1\right)$, we get $f_x\,(x_0, y_0) = f_x\left(\frac{1}{2}, 1\right) = -1$. This value is identical to the third component of $\vec{r}\,'(x_0) = \vec{r}\,'\left(\frac{1}{2}\right) = \langle 1, 0, -1\rangle$.

The geometric interpretation of $f_x\left(\frac{1}{2}, 1\right)$ becomes clear by looking at $\vec{r}\,'\left(\frac{1}{2}\right) = \langle 1, 0, -1\rangle$. The latter equation tells us that, at the point $\left(\frac{1}{2}, 1\right)$, every one-unit change in the x direction and zero-unit change in the y direction produces an instantaneous rate of change in the z direction of -1 unit.

Similarly, slicing the graph of f by the plane $x = 1$, produces a curve $\vec{r}\,(t) = \langle 1, t, f\,(1, t)\rangle$, as Figure 10 shows. This time we compute $\vec{r}\,'(y_0) = \left\langle 0, 1, \frac{d}{dt} f\,(x_0, t)\,\big|_{t=y_0} \right\rangle = \left\langle 0, 1, f_y\,(x_0, y_0) \right\rangle$. For $\vec{r}\,(t) = \left\langle 1, t, 6 - t^2 \right\rangle$, we get $\vec{r}\,'(t) = \langle 0, 1, -2t\rangle$, and the third component of this vector is the partial of f with respect to y evaluated at $y = t$. If we set $(x_0, y_0) = \left(1, -\frac{2}{5}\right)$, then $f_y\left(1, -\frac{2}{5}\right) = \frac{4}{5}$, and $\vec{r}\,'\left(-\frac{2}{5}\right) = \left\langle 0, 1, \frac{4}{5} \right\rangle$. The interpretation is that, at the point $\left(1, -\frac{2}{5}\right)$, every zero-unit change in the x direction and one-unit change in the y direction produces an instantaneous rate of change in the z direction of $\frac{4}{5}$ units.

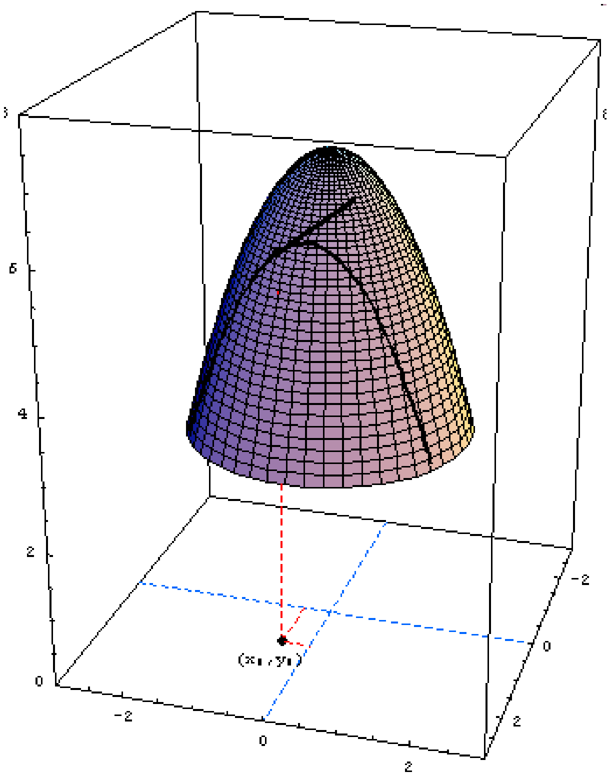

Figure 10

The link http://www.westmont.edu/~howell/courses/ma19/illustrations/pderiv.html gives an animation of Figures 9 and 10.

4 Higher Derivatives and Higher Dimensions

Partial derivatives can be evaluated at various points, so we can consider them to be functions. For example, if $f(x,y) = \sin(xy)$, then $f_x(x,y) = y\cos(xy)$, so $f_x : \mathbb{R}^2 \to \mathbb{R}$. Partial derivatives themselves may have partial derivatives. What notation should we use to indicate the partial derivative with respect to y of the function f_x? A natural choice is f_{xy} because reading left to right gives us the partials in the order that they occur. Taking the partial derivative with respect to x of the function f_y would be expressed as f_{yx}. The game doesn't stop here. We can compute the partial derivative of the function f_{xy} with respect to x, and so forth. If we write $z = f(x,y)$, we can express these higher derivatives in a variety of ways.

$$\begin{aligned}
f_{xx}(x,y) &= \frac{\partial^2 z}{\partial x^2} = \frac{\partial^2 f}{\partial x^2} = D_{xx}f = f^{(2,0)}(x,y) \\
f_{xy}(x,y) &= \frac{\partial^2 z}{\partial y \partial x} = \frac{\partial^2 f}{\partial y \partial x} = D_{xy}f = f^{(1,1)}(x,y) \\
f_{xyy}(x,y) &= \frac{\partial^3 z}{\partial y^2 \partial x} = \frac{\partial^3 f}{\partial y^2 \partial x} = D_{xyy}f = f^{(1,2)}(x,y) \\
f_{xyx}(x,y) &= \frac{\partial^3 z}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial x \partial y \partial x} = D_{xyx}f = f^{(2,1)}(x,y)
\end{aligned}$$

and so forth.

Notice that the higher partial f_{xy} reads left to right, but the symbol $\frac{\partial^2 f}{\partial y \partial x}$ reads right to left. The reason for this apparent anomaly is that $\frac{\partial^2 f}{\partial y \partial x}$ is really shorthand for $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$, in other words, "The partial with respect to y of the partial of f with respect to x."

Notice, also, that there is some ambiguity in the superscript notation. For example, does $f^{(1,1)}(x,y)$ intend to denote f_{xy} or f_{yx}? The two expressions are not always equal, but thanks to the French mathematician Alexis Claude Clairaut (*fl. ca.* 1750), the continuity of f_{xy} and f_{yx} ensures their equality. The following theorem states this fact more precisely.

Theorem 1 (Clairaut's Theorem) *Suppose f_{xy} and f_{yx} are continuous at all points located in a disk centered at the point (a,b). Then $f_{xy}(a,b) = f_{yx}(a,b)$.*

Exercise 9 *Verify Clairaut's theorem for $f(x,y) = x^5y^4 - 3x^2y^3 + 2x^2$ by evaluating the partials.*

What about functions $f : \mathbb{R}^n \to \mathbb{R}$? Fortunately, everything we've said up to this point carries over in a very straightforward way. For example, if $w = f(x,y,z) = x^5 + x^3y^4z^3 + yz^{10}$, then $\frac{\partial w}{\partial y} = 4x^3y^3z^3 + z^{10}$. In general, given $\overrightarrow{x_0} \in \mathbb{R}^n$, and $f : \mathbb{R}^n \to \mathbb{R}$, it can be easily shown that

$$\begin{aligned}
f\left(\overrightarrow{x}\right) &\approx L\left(\overrightarrow{x}\right) = f\left(\overrightarrow{x_0}\right) + \overrightarrow{f}'\left(\overrightarrow{x_0}\right) \bullet \left\langle \overrightarrow{x} - \overrightarrow{x_0} \right\rangle, \text{ where} \\
\overrightarrow{f}'\left(\overrightarrow{x_0}\right) &= \left\langle f_{x_1}\left(\overrightarrow{x_0}\right), f_{x_2}\left(\overrightarrow{x_0}\right), \ldots, f_{x_n}\left(\overrightarrow{x_0}\right) \right\rangle,
\end{aligned}$$

and $f_{x_k}\left(\overrightarrow{x_0}\right)$ designates the partial derivative with respect to the k^{th} coordinate evaluated at $\overrightarrow{x_0}$.

Example 6 *Find $\overrightarrow{f}'(x,y,z)$ if $f(x,y,z) = \frac{x}{y^3+z^4}$.*

Solution. $\vec{f}\,'(x,y,z) = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle = \left\langle \frac{1}{y^3+z^4}, -\frac{3xy^2}{(y^3+z^4)^2}, -\frac{4xz^3}{(y^3+z^4)^2} \right\rangle$.

Remark. Many texts use the symbol $\vec{\nabla} f(\vec{x})$ to designate the vector of partial derivatives, and $\vec{\nabla} f$ is called the gradient of f. Thus, in the last example we would write $\vec{\nabla} f(x,y,z) = \left\langle \frac{1}{y^3+z^4}, -\frac{3xy^2}{(y^3+z^4)^2}, -\frac{4xz^3}{(y^3+z^4)^2} \right\rangle$. It is important, however, not to equate $\vec{\nabla} f(\vec{x})$ with $\vec{f}\,'(\vec{x})$. Certainly, if a function f has a derivative at the point $\vec{x_0}$, then all its partial derivatives exist at $\vec{x_0}$. However,the existence of all the partials at $\vec{x_0}$ is not sufficient to guarantee that the function f is differentiable. Can you describe what such a situation would look like geometrically?

The following theorem tells us when we can be sure a function is differentiable.

Theorem 2 *Suppose* $f : \mathbb{R}^n \to \mathbb{R}$, *and all the partial derivatives exist and are continuous at* $\vec{x_0}$. *Then* f *is differentiable at* $\vec{x_0}$.

5 Directional Derivatives

Let's take stock of what we've learned so far.

- A function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\overrightarrow{x_0} \in \mathbb{R}^n$ if and only if there exists a linear function $L : \mathbb{R}^n \to \mathbb{R}$ that is a good approximation to the function f for points $\overrightarrow{x}$ near $\overrightarrow{x_0}$.
- By saying that L is a good approximation to f we mean that $\lim_{\overrightarrow{x} \to \overrightarrow{x_0}} \frac{e(\overrightarrow{x})}{\|\overrightarrow{x} - \overrightarrow{x_0}\|} = 0$, where $e\left(\overrightarrow{x}\right) = f\left(\overrightarrow{x}\right) - L\left(\overrightarrow{x}\right)$. In other words, the error term (that measures the difference between f and L) approaches zero much faster than does the distance between $\overrightarrow{x}$ and $\overrightarrow{x_0}$.
- The linear function L has the form $L\left(\overrightarrow{x}\right) = f\left(\overrightarrow{x_0}\right) + \overrightarrow{m} \bullet \left(\overrightarrow{x} - \overrightarrow{x_0}\right)$, and the vector $\overrightarrow{m}$ has the partial derivatives of f as its components. We define the derivative of f at the point $\overrightarrow{x_0}$, therefore, to be that vector. Thus, $\overrightarrow{m} = \overrightarrow{f}'\left(\overrightarrow{x_0}\right) = \left\langle f_{x_1}\left(\overrightarrow{x_0}\right), f_{x_2}\left(\overrightarrow{x_0}\right), \ldots, f_{x_n}\left(\overrightarrow{x_0}\right)\right\rangle$.
- Thus, if a function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\overrightarrow{x_0}$ we know that

$$\begin{aligned} f\left(\overrightarrow{x}\right) &\approx L\left(\overrightarrow{x}\right) = f\left(\overrightarrow{x_0}\right) + \overrightarrow{f}'\left(\overrightarrow{x_0}\right) \bullet \left\langle \overrightarrow{x} - \overrightarrow{x_0}\right\rangle, \text{ where} \\ \overrightarrow{f}'\left(\overrightarrow{x_0}\right) &= \left\langle f_{x_1}\left(\overrightarrow{x_0}\right), f_{x_2}\left(\overrightarrow{x_0}\right), \ldots, f_{x_n}\left(\overrightarrow{x_0}\right)\right\rangle. \end{aligned}$$

For functions of two variables, where $f'\left(\overrightarrow{x_0}\right) = \langle a, b\rangle$, we saw that $\lim_{\overrightarrow{x} \to \overrightarrow{x_0}} \frac{e(\overrightarrow{x})}{\|\overrightarrow{x} - \overrightarrow{x_0}\|} = 0$ is the same as [see Equation (5)] $\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y)-f(x_0,y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}} = \lim_{(x,y)\to(x_0,y_0)} \frac{a(x-x_0)+b(y-y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}}$. The partial derivatives we looked at in the last section grew out of evaluating these limits as (x, y) approached (x_0, y_0) along straight lines parallel to the x and y axes.

Of course, there are other straight line directions that we can look at. What would happen in $\mathbb{R}^2$ if $\overrightarrow{x} = \langle x, y\rangle$ were to approach $\overrightarrow{x_0} = \langle x_0, y_0\rangle$ along a direction of some *arbitrary* unit vector, say $\overrightarrow{u} = \langle u_1, u_2\rangle$? In this case, we could write $\overrightarrow{x}$ as $\overrightarrow{x_0} + h\overrightarrow{u} = \langle x_0, y_0\rangle + h\langle u_1, u_2\rangle = \langle x_0 + hu_1, y_0 + hu_2\rangle$, and Equation (5) would then become

$$\begin{aligned} \lim_{h\to 0} \frac{f(x_0 + hu_1, y_0 + hu_2) - f(x_0, y_0)}{\sqrt{(hu_1)^2 + (hu_2)^2}} &= \lim_{h\to 0} \frac{a(hu_1) + b(hu_2)}{\sqrt{(hu_1)^2 + (hu_2)^2}}, \text{ or} \\ \lim_{h\to 0} \frac{f\left(\overrightarrow{x_0} + h\overrightarrow{u}\right) - f\left(\overrightarrow{x_0}\right)}{\sqrt{(hu_1)^2 + (hu_2)^2}} &= \lim_{h\to 0} \frac{a(hu_1) + b(hu_2)}{\sqrt{(hu_1)^2 + (hu_2)^2}}, \qquad (11) \end{aligned}$$

where as before $a = f_x(x_0, y_0)$ and $b = f_y(x_0, y_0)$. It is a straightforward exercise to show that this last equation implies

$$\lim_{h\to 0} \frac{f\left(\overrightarrow{x_0} + h\overrightarrow{u}\right) - f\left(\overrightarrow{x_0}\right)}{h} = \langle a, b\rangle \bullet \overrightarrow{u}. \qquad (12)$$

Exercise 10 *Show that Equation (12) is a consequence of Equation (11).*

Happily, our findings apply to functions of more than two variables as well. Below we formalize all of the above with a definition and theorem that apply to functions $f : \mathbb{R}^n \to \mathbb{R}$.

Definition 4 *If $f : \mathbb{R}^n \to \mathbb{R}$, the quantity $f_{\overrightarrow{u}}\left(\overrightarrow{x_0}\right) = \lim_{h\to 0} \frac{f(\overrightarrow{x_0}+h\overrightarrow{u})-f(\overrightarrow{x_0})}{h}$ is called the directional derivative of f with respect to the unit vector $\overrightarrow{u}$ evaluated at $\overrightarrow{x_0}$, provided the limit exists.*

Theorem 3 *If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\overrightarrow{x_0}$, then the directional derivative of f exists in the direction of any unit vector $\overrightarrow{u}$, and $f_{\overrightarrow{u}}\left(\overrightarrow{x_0}\right) = \left\langle f_{x_1}\left(\overrightarrow{x_0}\right), f_{x_2}\left(\overrightarrow{x_0}\right), \ldots, f_{x_n}\left(\overrightarrow{x_0}\right)\right\rangle \bullet \overrightarrow{u}$.*

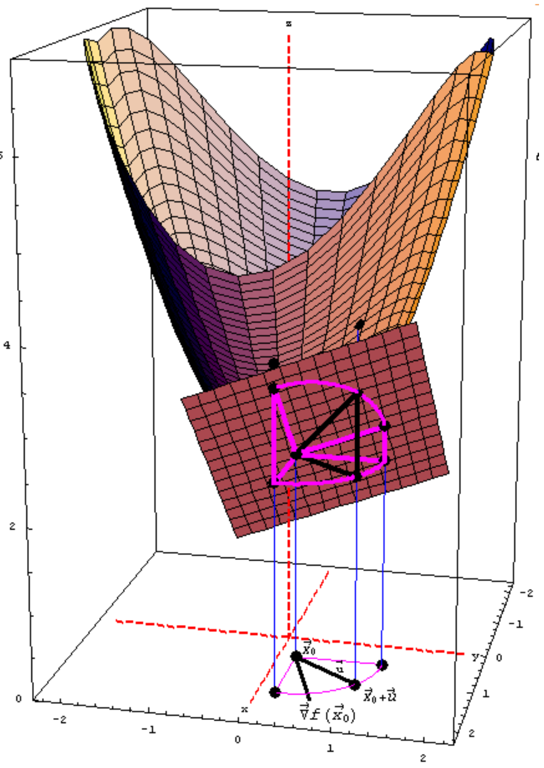

Figure 11

Figure 11 shows: (1) $f(x,y) = x^{\frac{4}{3}} + y^2 + 2$; (2) the point $\overrightarrow{x_0} = (x_0, y_0) = \left(\frac{1}{2}, \frac{1}{5}\right)$; (3) the tangent plane (linear function) approximation to f at the point $(x_0, y_0, f(x_0, y_0)) \approx \left(\frac{1}{2}, \frac{1}{5}, 2.4\right)$; (4) the unit vector $\overrightarrow{u} = \left(\frac{3}{5}, \frac{4}{5}\right)$ drawn from $\overrightarrow{x_0}$; (5) $\overrightarrow{\nabla} f\left(\overrightarrow{x_0}\right) \approx \langle 1.06, 0.4\rangle$. (Recall that $\overrightarrow{\nabla} f$, also called the gradient, is the vector of partial derivatives.)

The displacement on the tangent plane directly above the points $\overrightarrow{x_0}$ and $\overrightarrow{x_0} + \overrightarrow{u}$ is illustrated by the dark slanted line. The slope of this line is $f_{\overrightarrow{u}}\left(\overrightarrow{x_0}\right)$. Because $\overrightarrow{u}$ is a unit vector, the slope is the distance represented by the dark vertical line directly above the point $\overrightarrow{x_0} + \overrightarrow{u}$. In this case we compute $f_{\overrightarrow{u}}\left(\overrightarrow{x_0}\right) \approx \langle 1.06, 0.4\rangle \bullet \left\langle \frac{3}{5}, \frac{4}{5}\right\rangle = 0.956$. The value 0.956 indicates that, at the instant when $\overrightarrow{x_0} = \left(\frac{1}{2}, \frac{1}{5}\right)$, every unit change in the direction of $\overrightarrow{u} = \left(\frac{3}{5}, \frac{4}{5}\right)$ produces 0.956 units of change in the z-values. Alternatively, we could say that, at the instant when $\overrightarrow{x_0} = \left(\frac{1}{2}, \frac{1}{5}\right)$, every $\frac{3}{5}$ unit change in x and $\frac{4}{5}$ unit change in y produces 0.956 units of change in the z-values.

The circular segment drawn at the top of the tangent plane maps out where different displacement vectors above $\overrightarrow{x_0} + \overrightarrow{u}$ would appear for unit vectors $\overrightarrow{u} = \langle \cos\theta, \sin\theta\rangle$, where $0 \le \theta \le \frac{\pi}{2}$.

The maximum value of $f_{\overrightarrow{u}}\left(\overrightarrow{x_0}\right)$ seems to occur when $\overrightarrow{u}$ happens to be in the same direction as $\overrightarrow{\nabla} f\left(\overrightarrow{x_0}\right)$. This is always the case, and explains why $\overrightarrow{\nabla} f$ is called the *gradient* of f.

Theorem 4 *Suppose* $f : \mathbb{R}^n \to \mathbb{R}$ *is differentiable at* $\overrightarrow{x_0}$. *The maximum value for* $f_{\overrightarrow{u}}\left(\overrightarrow{x_0}\right)$ *occurs when* $\overrightarrow{u}$ *is in the same direction as* $\overrightarrow{\nabla} f\left(\overrightarrow{x_0}\right) = \left\langle f_x\left(\overrightarrow{x_0}\right), f_y\left(\overrightarrow{x_0}\right)\right\rangle$, *and the minimum value for* $f_{\overrightarrow{u}}\left(\overrightarrow{x_0}\right)$ *occurs when* $\overrightarrow{u}$ *is in the same direction as* $\overrightarrow{-\nabla} f\left(\overrightarrow{x_0}\right)$.

Proof. $f_{\overrightarrow{u}}\left(\overrightarrow{x_0}\right) = \overrightarrow{\nabla} f\left(\overrightarrow{x_0}\right) \bullet \overrightarrow{u} = \left\|\overrightarrow{\nabla} f\left(\overrightarrow{x_0}\right)\right\| \left\|\overrightarrow{u}\right\| \cos\theta = \left\|\overrightarrow{\nabla} f\left(\overrightarrow{x_0}\right)\right\| \cos\theta$, where θ is the angle between $\overrightarrow{\nabla} f\left(\overrightarrow{x_0}\right)$ and $\overrightarrow{u}$. Now $\left\|\overrightarrow{\nabla} f\left(\overrightarrow{x_0}\right)\right\|$ is constant, so the product $\left\|\overrightarrow{\nabla} f\left(\overrightarrow{x_0}\right)\right\| \cos\theta$ is maximized when $\theta = 0$, and minimized when $\theta = \pi$. ■

Exercise 11 *Evan thinks he can show Michelle what a macho guy he is by jumping barefoot onto a flat bed of burning coals, landing at the point* $(1,\ 2)$, *measured in feet. The temperature (degrees Fahrenheit) at any point* (a, b) *on the bed of coals is given by* $200 - a^3 b - a^4 b^2$, *so he quickly senses the folly of his actions. In what direction should he move so as to have the temperature decrease most rapidly? At what rate will the temperature change if Evan leaves* $(1, 2)$ *heading directly towards the point* $(3, 5)$*?*

6 The Multivariable Chain Rule

6.1 Introduction

For functions of one variable, $f'(c)$ is the slope of the line that is tangent to the graph of the function f at the point $(c,\ f(c))$. The "instantaneous rate of change" perspective of the derivative focuses on how fast the function f is changing at a particular point. For example, if $f'(c) = 4$, we can say that at the instant when $x = c$, the y values are changing four times as fast as the x values. If the function f were to continue to behave as it is at this particular instant, every unit change in x from the point $x = c$ would produce four units of change in the function f.

The chain rule tells us how to differentiate a function h that is defined as a *chain* of two functions, that is, $h(x) = f(g(x)) = (f \circ g)(x)$. The picture below illustrates this concept.

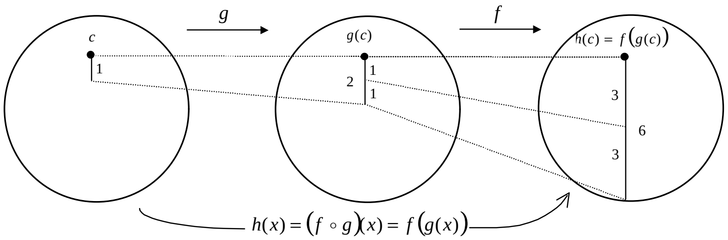

Figure 12

Suppose that $g'(c) = 2$, and $f'(g(c)) = 3$. What could we conclude about the value of $h'(c)$? Using the "instantaneous rate of change" perspective of the derivative, we interpret $g'(c) = 2$ to mean that every unit change in x from the point $x = c$ produces two units of change in the function g. Also, $f'(g(c)) = 3$ means that every unit change in x from the point $x = g(c)$ produces three units of change in the function f. It stands to reason, then, that every unit change in x from the point $x = c$ produces $g'(c)\,f'(g(c)) = (2)(3) = 6$ units change in the function h. Thus, $h'(c) = g'(c)\,f'(g(c))$, or $h'(c) = f'(g(c))\,g'(c)$.

6.2 Higher Dimensions

We now consider $h(t) = f\left(\overrightarrow{g}(t)\right)$, where $\overrightarrow{g} : \mathbb{R} \to \mathbb{R}^2$, and $f : \mathbb{R}^2 \to \mathbb{R}$, as shown.

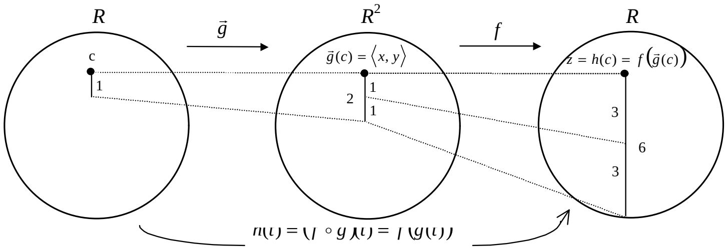

Figure 13

The multivariable chain rule is remarkably similar to the single variable case.

Theorem 5 *Suppose* $\vec{g} : \mathbb{R} \to \mathbb{R}^n$ *is differentiable at* c*, and* $f : \mathbb{R}^n \to \mathbb{R}$ *is differentiable at* $\vec{g}(c)$. *Then the composite function* $h(t) = f\left(\vec{g}(t)\right)$ *is differentiable at* c*, and*

$$h'(c) = \vec{f}'\left(\vec{g}(c)\right) \bullet \vec{g}'(c).$$

Figure 14 illustrates why this formula is true. We have a chain of functions $h(t) = f\left(\vec{g}(t)\right)$, where $\vec{g}(t) = \left\langle -2.5t^3 + 7.2t^2 - 5.9t + 1.7, \frac{4}{3}t - \frac{17}{15} \right\rangle$, $f(x, y) = x^{4/3} + y^2 + 2$, and $c = 1$.

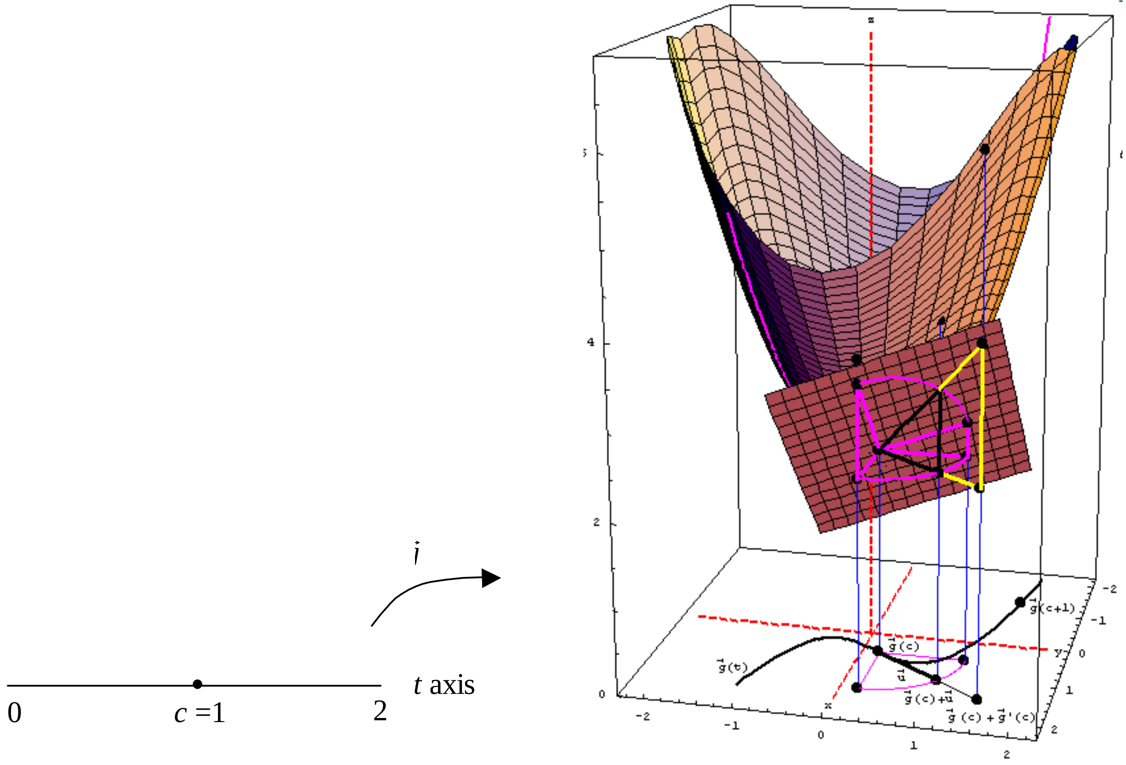

Figure 14

With $c = 1$ we compute $\vec{g}(c) = \langle 0.5, 0.2 \rangle$, $\vec{g}'(c) = \left\langle 1, \frac{4}{3} \right\rangle$, and $\left\| \vec{g}'(c) \right\| = \frac{5}{3}$. Thus, the unit vector in the direction of $\vec{g}'(c)$ is $\vec{u} = \frac{\vec{g}'(c)}{\left\| \vec{g}'(c) \right\|} = \left\langle \frac{3}{5}, \frac{4}{5} \right\rangle$.

Now, every unit change in t from the point $t = c$ produces $\left\| \vec{g}'(c) \right\| = \frac{5}{3}$ units change in the function $\vec{g}$ *in the direction of* $\vec{g}'(c)$. But every unit change from the point $\vec{g}(c)$ in the direction of $\vec{g}'(c)$ produces $f_{\vec{u}}\left(\vec{g}(c)\right)$ units change in the function f, which equals

$$f_{\vec{u}}\left(\vec{g}(c)\right) = \vec{f}'\left(\vec{g}(c)\right) \bullet \vec{u} = \vec{f}'\left(\vec{g}(c)\right) \bullet \frac{\vec{g}'(c)}{\left\| \vec{g}'(c) \right\|}.$$

Hence, the total instantaneous change in the function h (per every unit change from the point $t = c$) is

$$h'(c) = \left[\vec{f}'\left(\vec{g}(c)\right) \bullet \frac{\vec{g}'(c)}{\left\| \vec{g}'(c) \right\|} \right] \left\| \vec{g}'(c) \right\| = \vec{f}'\left(\vec{g}(c)\right) \bullet \vec{g}'(c),$$

which is what the multivariable chain rule stipulates. For these particular functions, we compute $\vec{f}'(x, y) = \left\langle \frac{4}{3}x^{1/3}, 2y \right\rangle$, so $\vec{f}'\left(\vec{g}(c)\right) = \vec{f}'(0.5, 0.2) \approx \langle 1.06, 0.4 \rangle$. We finally get $h'(c) =$

$h'(1) = \overrightarrow{f}'\left(\overrightarrow{g}(1)\right) \bullet \overrightarrow{g}'(1) \approx \langle 1.06, 0.4\rangle \bullet \left\langle 1, \frac{4}{3}\right\rangle \approx 1.59$. Thus, every unit change from the point $c = 1$ on the t-axis produces a change of 1.59 units on the z-axis. In other words, the instantaneous rate of change in the function h from the point $h(1) = f\left(\overrightarrow{g}(1)\right)$ amounts to 1.59 units for every unit change from the point $c = 1$ on the t-axis. The vertical "yellow" line directly above the vector $\overrightarrow{g}(c) + \overrightarrow{g}'(c)$ in Figure 14, then, must equal 1.59.

6.3 Leibniz Notation

The Leibniz notation for $h'(t)$ is similar to that for single variable calculus. We let z represent the output of the function f and also consider z as a function of the variable t. We then have $\overrightarrow{f}'\left(\overrightarrow{g}(t)\right) = \overrightarrow{f}'(x, y) = \langle f_x, f_y\rangle = \left\langle \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right\rangle$ and $\overrightarrow{g}'(t) = \left\langle \frac{dx}{dt}, \frac{dy}{dt}\right\rangle$, so that

$$\frac{dz}{dt} = \left\langle \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right\rangle \bullet \left\langle \frac{dx}{dt}, \frac{dy}{dt}\right\rangle = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

Great care must be taken in using this approach, however, as the main weaknesses in Leibniz notation are that the functions chained together get obscured, and it is not clear at what points the derivatives are evaluated. Again, for our particular functions we have $\langle x, y\rangle = \overrightarrow{g}(t) = \left\langle -2.5t^3 + 7.2t^2 - 5.9t + 1.7, \frac{4}{3}t - \frac{17}{15}\right\rangle$ and $z = f(x, y) = x^{4/3} + y^2 + 2$. Thus, $\frac{dx}{dt} = -7.5t^2 + 14.4t - 5.9$, $\frac{dy}{dt} = \frac{4}{3}$, $\frac{\partial z}{\partial x} = \frac{4}{3}x^{1/3}$, and $\frac{\partial z}{\partial y} = 2y$. This gives $\frac{dz}{dt} = \left(\frac{4}{3}x^{1/3}\right)\left(-7.5t^2 + 14.4t - 5.9\right) + (2y)\left(\frac{4}{3}\right)$, which we evaluate for $t = 1$, $x = 0.5$, and $y = 0.2$. Do you see *why* we evaluate at these points?

A slight advantage in Leibniz notation is that it more easily tells us how to piece the various functions of a chain together. Suppose that w is the output of a function of three variables, say $w = f(x, y, z)$, and that the vector $\langle x, y, z\rangle$ is the output of a function $\overrightarrow{g}(s, t) : \mathbb{R}^2 \to \mathbb{R}^3$, so that $\overrightarrow{g}(s, t) = \langle x, y, z\rangle$. Then w depends on s and t because $w = f\left(\overrightarrow{g}(s, t)\right)$, and

$$\begin{aligned}\frac{\partial w}{\partial s} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s},\\ \frac{\partial w}{\partial t} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t}.\end{aligned}$$

Exercise 12 *Text, page 974, problem 13.*

Exercise 13 *Text, page 974, problem 14.*

Exercise 14 *Text, page 975, problem 37.*

Exercise 15 *Text, page 975, problem 40.*