§ 2 FUNCTIONS and VECTOR NOTATION

One way to make the material in Calculus III a bit more coherent is to concentrate on the notion of a function in order to unify the various topics. In its most general setting, a function is an unambiguous “rule” which assigns an element of some “range set” to each member of its “domain set.” We will use the standard notation

$$f : X \to Y$$

to indicate that \(f\) is a function which assigns a member of the set \(Y\) to each member of the set \(X\). This is usually read as “\(f\) is a function from \(X\) to \(Y\).”

Throughout Calculus I and II, all of the functions were of the form \(f : \mathbb{R} \to \mathbb{R}\). In Calculus III, we are studying functions where the domain and/or range is \(\mathbb{R}^n\), for \(n = 2, 3, \ldots\). In chapter 11 of the text, the “space curves” are functions \(f : \mathbb{R} \to \mathbb{R}^n\), where \(n\) is either 2 or 3. Instead of \(f(x)\), we used the notation \(\vec{r}(t)\) to indicate this fact. These functions are sometimes referred to as “vector valued functions of a scalar variable.” In chapters 12 and 13 we are studying functions \(f : \mathbb{R}^n \to \mathbb{R}\), which are called “multivariate functions,” or “functions of several variables,” or “scalar valued functions of a vector variable.” In chapter 14, we will briefly discuss the more general cases of functions \(f : \mathbb{R}^n \to \mathbb{R}^m\).

Vector notation is quite standard for denoting functions \(f : \mathbb{R} \to \mathbb{R}^n\). For example, to define a function \(\vec{r} : \mathbb{R} \to \mathbb{R}^3\), we use something like \(\vec{r}(t) = \langle t, t^2, \sin(t) \rangle\). For functions \(f : \mathbb{R}^n \to \mathbb{R}\), however, there are two different notations that can be used. One system is to simply list the variables: \(f(x, y)\) denotes a function from \(\mathbb{R}^2\) to \(\mathbb{R}\), \(f(x, y, z)\) denotes a domain of \(\mathbb{R}^3\), etc. This is the most convenient way to deal with a specific function. For example, \(f(x, y) = x^2 + y^3\) denotes the “rule” that, for example, assigns to the ordered pair \((1, 2)\) the number \(1^2 + 2^3 = 9\). It is important to realize that the letters \(x\) and \(y\) are just “dummy variables,” and that any letters could actually be used. In Calculus I, the equations \(f(x) = x^2\) or \(f(t) = t^2\) were equivalent definitions for the same “squaring” function. Similarly, \(f(s, t) = s^2 + t^3\) defines the same function that we mentioned above. However, the order in which the variables are listed is important. The function \(g(y, x) = x^2 + y^3\) is different from \(f\), since \(g\) would send \((1, 2)\) to the number \(2^2 + 1^3 = 5\). On the other hand, if we let \(g(y, x) = y^2 + x^3\), then \(g\) is the same function as \(f\) (you should convince yourself of this fact before continuing).

There is a second way to denote a function \(f : \mathbb{R}^n \to \mathbb{R}\). We can use the more compact notation \(f(\vec{x})\) which has a single vector variable \(\vec{x}\) that denotes the entire list of scalar variables (Note that there is no arrow over the \(f\), which indicates that the range is \(\mathbb{R}\).) The textbook briefly mentions this notation, but unfortunately, it rarely uses it. In fact, this notation is not usually convenient when discussing a specific function. For example, it doesn’t really make sense to write \(f(\vec{x}) = x^2 + y^3\) because there is no way to tell from
this equation which order the scalar variables \( x \) and \( y \) should occur in the vector \( \vec{x} \). We could get around these problems by writing either: \( f(\vec{x}) = x^2 + y^3 \) where \( \vec{x} = \langle x, y \rangle \) or \( f(\vec{x}) = (\vec{x} \cdot (1, 0))^2 + (\vec{x} \cdot (0, 1))^3 \). But \( f(x, y) = x^2 + y^3 \) is simpler. The place where vector notation has an advantage is when we want to discuss general properties of functions. For example, consider the idea that “the limit of a sum will equal the sum of the limits.” In Calculus I, this was written as

\[
\lim_{x \to a} (f(x) + g(x)) = \lim_{x \to a} f(x) + \lim_{x \to a} g(x).
\]

For a function of two variables, we could write this as

\[
\lim_{(x,y) \to (a,b)} (f(x, y) + g(x, y)) = \lim_{(x,y) \to (a,b)} f(x, y) + \lim_{(x,y) \to (a,b)} g(x, y).
\]

But with vector notation we can instead write

\[
\lim_{\vec{x} \to \vec{a}} (f(\vec{x}) + g(\vec{x})) = \lim_{\vec{x} \to \vec{a}} f(\vec{x}) + \lim_{\vec{x} \to \vec{a}} g(\vec{x})
\]

In addition to being simpler, the vector equation above also makes sense for functions of three or more variables, so a single equation can represent a more general fact. (In the textbook, many concepts are stated twice, once for the two variable case and again for the three variable case.) Also, vector notation helps to emphasize the similarities between concepts from Calculus I and the new ones we are studying.

2. For the sake of accuracy, we should point out a slight error in the above discussion. When we use the notation \( f : X \to Y \), \( X \) must be the domain of the function \( f \), while the set \( Y \) can be any set which contains the range of \( f \). The actual range is the set of elements of \( Y \) that actually get used: \( \text{ran}(f) \overset{\text{def}}{=} \{ y \in Y : y = f(x) \text{ for some } x \in X \} \). So, many of our calculus I functions weren’t really \( f : \mathbb{R} \to \mathbb{R} \). The simple function \( f(x) = 1/x \) is an example, because the the domain is the set of non-zero real numbers. To handle such functions, we should really write \( f : D \to \mathbb{R} \) where \( D \subset \mathbb{R} \). The symbol \( D \) represents the domain of \( f \), and the notation \( D \subset \mathbb{R} \) is read as “\( D \) is a subset of \( \mathbb{R} \).” For the sake of simplicity, however, we will ignore this technicality for now. We will only consider functions \( f : \mathbb{R}^n \to \mathbb{R} \) for now. After we have finished with differentiability we will then discuss how to handle the more general case of functions \( f : D \to \mathbb{R} \) where \( D \subset \mathbb{R}^n \).

2.1 Exercise. Which of the following function definitions are equivalent?

\[
\begin{align*}
 f(\vec{x}) & = (\vec{x} \cdot \vec{x})(\vec{x} \cdot (1, 0)), & f(a, b) & = a^3 + ab^2, & f(y, x) & = x^3 + xy^2 \\
 f(\vec{x}) & = \|\vec{x}\| + 2, & f(x, y) & = x^2 + y^2 + 2, & f(\vec{v}) & = \sqrt{\vec{v} \cdot \vec{v}} + 2
\end{align*}
\]
§ 3 LIMITS and CONTINUITY

For the “Chapter 11” functions \( \vec{r} : \mathbb{R} \rightarrow \mathbb{R}^n \), the notions of limits, continuity, and even differentiability are quite simple. Any such \( \vec{r} \) just consists of an ordered list of \( n \) functions from \( \mathbb{R} \) to \( \mathbb{R} \), and so we can just apply the ideas from Calculus I to each of the “coordinate functions.” For multivariate functions, however, things get more complicated.

For \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) we can use vector notation to write

\[
\lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x}) = L
\]

which should mean the same thing as it did in Calculus I—if \( \vec{x} \) is “very close” to \( \vec{a} \), but not actually equal to \( \vec{a} \), then \( f(\vec{x}) \) is “very close” to \( L \). To define this precisely requires an \( \epsilon \) and a \( \delta \), but with vector notation the definition is almost identical to the one from Calculus I: For every \( \epsilon > 0 \) there exists a \( \delta > 0 \) such that

\[
0 < \|\vec{x} - \vec{a}\| < \delta \implies |f(\vec{x}) - L| < \epsilon.
\]

We can similarly define continuity by the statement \( \lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x}) = L \), or equivalently that for every \( \epsilon > 0 \) there exists a \( \delta > 0 \) such that \( \|\vec{x} - \vec{a}\| < \delta \implies |f(\vec{x}) - f(\vec{a})| < \epsilon \).

In Calculus I the \( \epsilon - \delta \) definition could be mostly ignored. The limit laws in section 1.3 implied that most “garden variety” functions are continuous at each point in their domain, and thus limits can be evaluated by just plugging in: so \( \lim_{x \rightarrow 2} x^2 + 3x = 2^2 + 6 = 10 \). Furthermore, if \( \lim_{x \rightarrow a} f(x) \neq 0 \) and \( \lim_{x \rightarrow a} g(x) = 0 \) then \( \lim_{x \rightarrow a} f(x)/g(x) \) is undefined. Although it is not clearly stated in our book, the limit laws hold in the multivariate case as well. So does the fact that “non-zero over zero” limits are undefined. Thus, \( \lim_{(x,y) \rightarrow (1,2)} 1/(x^2 + y^3) \) is undefined.

For single variable functions, we had two methods for handling “indeterminate forms” such as \( 0/0 \). The first was to algebraically simplify the expression for \( f \), and the second was to apply L’Hospital’s rule. Except in rare cases, neither of these methods will work for functions of two or more variables. If \( \lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x}) \) really is undefined, then one can often show this by finding two different “paths of approach” to \( \vec{a} \) which yield different limits. A typical example is to consider a function such as \( f(x,y) = \frac{xy}{x^2 + y^2} \) and \( \lim_{\vec{x} \rightarrow \vec{0}} f(\vec{x}) \). Along the \( x \) or \( y \) axes, the limit is zero, but along the line \( y = x \) the limit is \( 1/2 \), so the limit itself is undefined. This resembles the idea from Calculus I of using different right and left hand limits to show a limit does not exist. The difference is that there are now infinitely many possible paths to consider, instead of just two directions.

When \( f \) is indeterminate at \( \vec{x} = \vec{a} \), but the limit actually does exist, it is usually necessary to use some tricky inequalities to show this fact. We will do one example to illustrate this (this example will also be important in the next section). Consider \( \lim_{\vec{x} \rightarrow \vec{0}} f(\vec{x}) \)

Additional notes for Math 281 7 October 25, 2000
for \( f(x, y) = \frac{xy}{\sqrt{x^2 + y^2}} \). If you try several paths, you get a limit of zero in each case, but that is not enough to show the limit is really zero. There is no effective way to consider all possible paths (straight lines are not sufficient). We start the argument with some inequalities, although it will not be immediately clear how they are chosen. Note that \(|x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2}\), and similarly that \(|y| = \sqrt{y^2} \leq \sqrt{x^2 + y^2}\). Now, since \(|x|\) and \(|y|\) are non-negative, \(|xy| = |x||y| \leq (\sqrt{x^2 + y^2})^2\). If follows that \(|f(x, y)| \leq \sqrt{x^2 + y^2}\) or, more concisely, that \(|f(\vec{x})| \leq \|\vec{x}\|\) for all \(\vec{x}\). We can now show that \(\lim_{\vec{x} \to \vec{0}} f(\vec{x}) = 0\) by just setting \(\delta = \epsilon\) in the definition. Alternatively, since \(-\|\vec{x}\| \leq f(\vec{x}) \leq \|\vec{x}\|\), the graph of \(f\) is “squeezed” between the graphs of \(z = \|\vec{x}\|\) and \(z = -\|\vec{x}\|\), and it must therefore approach 0 as \(\vec{x} \to \vec{0}\).

The following theorem justifies using “different paths of approach” to show a limit is undefined:

3.1 Theorem. If \(\vec{r} : \mathbb{R} \to \mathbb{R}^n\) is continuous and \(f : \mathbb{R}^n \to \mathbb{R}\) is continuous, then so is the function from \(\mathbb{R}\) to \(\mathbb{R}\) defined by \(f(\vec{r}(t))\).

3.2 Exercise. Use Theorem 3.1 to fully justify the fact that \(\lim_{\vec{x} \to \vec{0}} \frac{xy}{x^2 + y^2}\) is undefined.

§ 4 DIFFERENTIABILITY

Finding the “tangent plane” to a surface \(z = f(x, y)\) at a given point is easy to do, assuming that a well defined tangent plane exists. The equation of the plane will be \(z = m_1 x + m_2 y + b\), where \(m_1\) and \(m_2\) are the values of the partial derivatives (\(\frac{\partial f}{\partial x}\) and \(\frac{\partial f}{\partial y}\) respectively) at the given point, and the constant \(b\) is chosen so that the plane passes through the given point. This is really just a two-variable analog of finding the tangent line to the graph of a single variable function.

One can usually decide if the partial derivatives exists a given point. If either \(\frac{\partial f}{\partial x}\) or \(\frac{\partial f}{\partial y}\) is undefined, then there will not be a well-defined tangent plane, so the surface will not be “smooth” at the given point. (Actually, it also might mean that the tangent plane is vertical, so that it can’t be written as \(z = m_1 x + m_2 y + b\). This same problem occurs with vertical tangent lines.) It would be nice if the existence of \(\frac{\partial f}{\partial x}\) both \(p_{fy}\) at a point would imply that the function was differentiable. This should mean that the graph of \(f\) is smooth at the point—it should have no creases or cusps. But, such is not the case. To see this, one can use the example we mentioned above. Let:

\[
(4.1) \quad f(x, y) = \begin{cases} 
\frac{xy}{\sqrt{x^2 + y^2}} & \text{if } (x, y) \neq \vec{0} \\
0 & \text{if } (x, y) = \vec{0}
\end{cases}
\]

Additional notes for Math 281 8 October 25, 2000
(we need the cases so that \( f(\vec{0}) \) is defined). By the inequalities we worked out in section 3, this function is continuous at the origin. Since \( f \) equals zero on the entire \( x \) axis, it is easy to see that \( \frac{\partial f}{\partial x}(0,0) = 0 \). Similarly \( \frac{\partial f}{\partial y}(0,0) = 0 \). But if \( x > 0 \), then \( f(x,x) = x^2/(\sqrt{2}x) = x/\sqrt{2} \), so in the direction of this ray the tangent plane should have a slope of \( 1/\sqrt{2} \). Even worse, if we also consider negative values for \( x \), then \( f(x,x) = |x|/\sqrt{2} \), so this surface has an absolute value graph sitting in it over the line \( y = x \). So, even though \( f \) is continuous and its partial derivatives exist, it has a “cusp” at the origin. It is not smooth and therefore should not be differentiable.

4.2 Exercise. Use a computer to graph \( \frac{xy}{\sqrt{x^2+y^2}} \), and find a perspective from which the cusp at the origin is clearly visible.

4.3 Exercise. Consider the function \( f \) defined by: \( f(x,y) = \frac{xy}{x^2+y^2} \), when \( (x,y) \neq \vec{0} \) and \( f(\vec{0}) = 0 \). Show that its partial derivative exist at the origin even though it is not even continuous there.

Additional notes for Math 281 9 October 25, 2000
out with remains. Equation 4.1 shows that $\nabla f(\vec{x})$ may exist even though the surface has a cusp at $\vec{x}$. We can now compute derivatives, but we still can’t tell, in general, if a function is differentiable or not!

The key to remedying this problem is differentials. Let’s recall the idea behind differentials for functions of one variable. To make the idea clearer, I’ll avoid using the $dx$ and $dy$ notations. The concept is simply that, if $h$ is “small,” then $f'(x) \approx \frac{f(x+h)-f(x)}{h}$. If we solve this “approximate equation” for $f(x+h)$, we get $f(x+h) \approx f(x) + f'(x)h$. This should also look familiar from Calculus II, since it is just the beginning of the Taylor series for $f$—the Taylor series really just improves on the differential approximation by adding the terms $\frac{f''(x)}{2!}h^2$, $\frac{f'''(x)}{3!}h^3$, etc.

It turns out that the differential approximation idea can be used to characterize when a function of one variable is differentiable. In order to do this we have to improve the fuzzy idea of “approximately equal.” This is done by introducing a function for the error in the approximation. Consider a function $f : \mathbb{R} \to \mathbb{R}$, and fix an $x$ and suppose that $m$ is any number such that

$$f(x + h) = f(x) + mh + e(h)$$

holds for all numbers $h$. The function $e(h)$ is measuring the “error” in the “differential-type” approximation, using the arbitrary slope $m$. (Actually, this equation really does nothing more than define $e(h) \overset{\text{def}}{=} f(x+h) - f(x) - mh$, but the above format makes the motivation clearer.) We can easily solve for $m$ (assuming $h \neq 0$), and get that

$$m = \frac{f(x + h) - f(x)}{h} - \frac{e(h)}{h}.$$

Now consider what happens to this equation in the limit as $h \to 0$. There are two important things to notice. On the one hand, if $f'(x)$ exists, and we were to let $m = f'(x)$, then we would have $m = \lim_{h \to 0} e(h) = 0$. On the other hand, if $\lim_{h \to 0} \frac{e(h)}{h} = 0$, then $m = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} - 0$, so $f$ must be differentiable and $m$ must equal $f'(x)$. We have actually just proven an important theorem characterizing differentiability.

**4.4 Theorem.** Let $f : \mathbb{R} \to \mathbb{R}$ be a function and let $x$ be a fixed number. Then $f$ is differentiable at $x$ if, and only if, there is a number $m$ and a function $e : \mathbb{R} \to \mathbb{R}$ such that $\lim_{h \to 0} \frac{e(h)}{h} = 0$ and

$$f(x + h) = f(x) + mh + e(h)$$

holds for all numbers $h$. In the case that such an $m$ and $e$ exist, $m$ must equal $f'(x)$.
It is actually easy to interpret theorem 4.4 in a conceptual way. The condition that
\[ \lim_{h \to 0} \frac{e(h)}{h} = 0 \]
implies that the error \( e(h) \) approaches 0 as \( h \to 0 \). But it really says more
than that. It says that \( e(h) \) approaches 0 quickly—more quickly than \( h \) itself does. Since
we are thinking of \( h \) as the variable, the term \( f(x) \) is a constant and the term \( mh \) is
nothing more than a linear function of \( h \). So, Theorem 4.4 just says that \( f \) is differentiable
when we can find a linear function which does a “very good” job of approximating \( f(x+h) \),
where “very good” really means that the error goes to zero quickly enough for
\[ \lim_{h \to 0} \frac{e(h)}{h} = 0. \]

We can now have the ideas in place to define differentiability for a function of two
variables. Consider a function \( \vec{f}(\vec{x}) \). How can we “linearly” approximate the value of
\( \vec{f}(\vec{x} + \vec{h}) \) for “small” vectors \( \vec{h} \). This linear function should have the graph a graph which
is a plane, using the vector notation we started with it will have an equation of the form
\[ z = \vec{m} \cdot \vec{x} + b. \]
Since \( f(\vec{x} + \vec{0}) = f(\vec{x}) \), we can see that \( b \) must be \( f(\vec{x}) \). This gives an
equation just like the one-dimensional case:
\[ f(\vec{x} + \vec{h}) \approx f(\vec{x}) + \vec{m} \cdot \vec{h} \]
where \( \vec{m} \) is a fixed vector. How can we say that this approximation is “good”? The error
in the approximation is a function \( e(\vec{h}) \), i.e., it is also a function of two variables. We
require that \( e(\vec{h}) \) approach zero quickly by using the condition that
\[ \lim_{\vec{h} \to \vec{0}} \frac{e(\vec{h})}{\|\vec{h}\|} = 0 \]
Which says that \( e(\vec{h}) \) gets small faster than \( \vec{h} \) itself does. Note that we must divide
by the length of \( \vec{h} \), because the quotient \( e(\vec{h})/\vec{h} \) would not make any sense. (A further
justification for introducing \( \|\vec{h}\| \) is the fact that, for functions of a single variable, the
conditions \( \lim_{h \to 0} \frac{e(h)}{h} = 0 \) and \( \lim_{h \to 0} \frac{e(h)}{|h|} = 0 \) are equivalent.)

We can now define differentiability.

4.5 Definition. Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a function and let \( \vec{x} \) be a fixed vector. Then \( f \) is
differentiable at \( \vec{x} \) means that there is a vector \( \vec{m} \) and a function \( e : \mathbb{R}^n \to \mathbb{R} \) such that
\[ f(\vec{x} + \vec{h}) = f(\vec{x}) + \vec{m} \cdot \vec{h} + e(\vec{h}) \]
holds for all vectors \( \vec{h} \), where \( \lim_{\vec{h} \to \vec{0}} \frac{e(\vec{h})}{\|\vec{h}\|} = 0 \). When this occurs, the vector \( \vec{m} \) is denoted
by \( \vec{f}'(\vec{x}) \).

This gives a definition of when a function is differentiable, i.e., it tells us when \( \vec{f}'(\vec{x}) \)
equals. Fortunately for us, we don’t have to use the definition to compute \( \vec{f}'(\vec{x}) \), because
partial derivatives will do that for us. The following theorem says this.

4.6 Theorem. If \( f \) is differentiable at \( \vec{x} \) then \( \vec{f}'(\vec{x}) = \nabla f(\vec{x}) \).

Additional notes for Math 281

11 October 25, 2000
However, if for \( f \) defined in equation 4.1, it turns out that \( \vec{f}'(\vec{0}) \) is undefined even though \( \nabla f(\vec{0}) \) exists. So, it is possible for \( \nabla f(\vec{x}) \) to exist even though \( f \) is not differentiable at \( \vec{x} \). Fortunately, this is an unusual case. The following key theorem says that functions which are “nice enough” will be differentiable.

4.7 Theorem. Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a function and let \( \vec{x} \) be a fixed vector. Suppose that the partial derivative functions for \( f \) are continuous at \( \vec{x} \). Then \( f \) is differentiable at \( \vec{x} \).

So, for a simple function such as \( f(x, y) = x^2 + y^2 \), we have \( \frac{\partial f}{\partial x}(x, y) = 2x \) and \( \frac{\partial f}{\partial y}(x, y) = 2y \) are continuous everywhere and thus \( f \) is differentiable everywhere. There is one more important theorem, which generalizes another fact from Calculus I.

4.8 Theorem. If \( f : \mathbb{R}^2 \to \mathbb{R} \) is differentiable at \( \vec{x} \), then it must be continuous at \( \vec{x} \).

Note that the continuity in Theorem 7 refers to continuity for functions of two (or more) variables, since partial derivatives are multivariate functions. We know that this can be difficult to verify.

Also, you may recall from Calculus I that it is possible for \( f'(a) \) to exist without \( f' \) being continuous at \( x = a \). The simplest example is based on \( x^2 \sin(1/x) \). The same thing can happen here—it is possible for one or more of the partial derivatives of \( f \) to be discontinuous, but for \( f \) to differentiable despite that fact.

4.9 Exercise. Use Theorem 4.8 to argue that the function in exercise 4.3 is not differentiable.

4.10 Exercise. Use Definition 4.5 directly to show that the function in equation 4.1 really is not differentiable at \( \vec{0} \). Is there any other way to argue this fact?

4.11 Exercise. Show that the function in equation 4.1 does not violate Theorem 7. Find an equation for \( \frac{\partial f}{\partial x}(\vec{x}) \) where \( \vec{x} \neq 0 \), and then use two different paths of approach to show that \( \lim_{\vec{x} \to \vec{0}} \frac{\partial f}{\partial x}(\vec{x}) \) is undefined.

To summarize, Theorems 4.6, 4.7, and 4.8 say the following. If \( f \) is differentiable at \( \vec{x} \), the \( f \) must be continuous at \( \vec{x} \) and all the partial derivatives of \( f \) must exist at \( \vec{x} \), and the gradient vector will give the “slopes” for the tangent plane to the surface. On the other hand, if the partial derivatives are continuous at \( \vec{x} \), then \( f \) must be differentiable at \( \vec{x} \). But, there are (rare) difficult cases of continuous functions whose the partial derivatives exist but are not continuous. For such functions, there is no “easy” way to test for differentiability—the definition must be used directly.