CHAPTER 10 DIFFERENTIATION

The process of finding the derivative of a function is called *differentiation*. From the previous chapter you may have the impression that this process is usually laborious, requires recourse to the definition of the derivative, and depends upon successfully recognizing some limit. It is true that such a procedure is often the only possible approach—if you forget the definition of the derivative you are likely to be lost. Nevertheless, in this chapter we will learn to differentiate a large number of functions, without the necessity of even recalling the definition. A few theorems will provide a mechanical process for differentiating a large class of functions, which are formed from a few simple functions by the process of addition, multiplication, division, and composition. This description should suggest what theorems will be proved. We will first find the derivative of a few simple functions, and then prove theorems about the sum, products, quotients, and compositions of differentiable functions. The first theorem is merely a formal recognition of a computation carried out in the previous chapter.

THEOREM 1 If f is a constant function, $f(x) = c$, then

$$f'(a) = 0 \quad \text{for all numbers } a.$$

PROOF

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} = \lim_{h\to 0} \frac{c-c}{h} = 0. \ \blacksquare$$

The second theorem is also a special case of a computation in the last chapter.

THEOREM 2 If f is the identity function, $f(x) = x$, then

$$f'(a) = 1 \quad \text{for all numbers } a.$$

PROOF

$$\begin{aligned} f'(a) &= \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \\ &= \lim_{h\to 0} \frac{a+h-a}{h} \\ &= \lim_{h\to 0} \frac{h}{h} = 1. \ \blacksquare \end{aligned}$$

The derivative of the sum of two functions is just what one would hope—the sum of the derivatives.

THEOREM 3 If f and g are differentiable at a, then $f+g$ is also differentiable at a, and

$$(f+g)'(a) = f'(a) + g'(a).$$

PROOF

$$\begin{aligned}(f+g)'(a) &= \lim_{h\to 0} \frac{(f+g)(a+h)-(f+g)(a)}{h}\\ &= \lim_{h\to 0} \frac{f(a+h)+g(a+h)-[f(a)+g(a)]}{h}\\ &= \lim_{h\to 0} \left[\frac{f(a+h)-f(a)}{h} + \frac{g(a+h)-g(a)}{h}\right]\\ &= \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} + \lim_{h\to 0} \frac{g(a+h)-g(a)}{h}\\ &= f'(a)+g'(a). \blacksquare\end{aligned}$$

The formula for the derivative of a product is not as simple as one might wish, but it is nevertheless pleasantly symmetric, and the proof requires only a simple algebraic trick, which we have found useful before—a number is not changed if the same quantity is added to and subtracted from it.

THEOREM 4 If f and g are differentiable at a, then $f\cdot g$ is also differentiable at a, and

$$(f\cdot g)'(a) = f'(a)\cdot g(a) + f(a)\cdot g'(a).$$

PROOF

$$\begin{aligned}(f\cdot g)'(a) &= \lim_{h\to 0} \frac{(f\cdot g)(a+h)-(f\cdot g)(a)}{h}\\ &= \lim_{h\to 0} \frac{f(a+h)g(a+h)-f(a)g(a)}{h}\\ &= \lim_{h\to 0} \left[\frac{f(a+h)[g(a+h)-g(a)]}{h} + \frac{[f(a+h)-f(a)]g(a)}{h}\right]\\ &= \lim_{h\to 0} f(a+h)\cdot \lim_{h\to 0} \frac{g(a+h)-g(a)}{h} + \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}\cdot \lim_{h\to 0} g(a)\\ &= f(a)\cdot g'(a) + f'(a)\cdot g(a).\end{aligned}$$

(Notice that we have used Theorem 9-1 to conclude that $\lim_{h\to 0} f(a+h) = f(a)$.) $\blacksquare$

In one special case Theorem 4 simplifies considerably:

THEOREM 5 If $g(x) = cf(x)$ and f is differentiable at a, then g is differentiable at a, and

$$g'(a) = c\cdot f'(a).$$

PROOF If $h(x) = c$, so that $g = h\cdot f$, then by Theorem 4,

$$\begin{aligned}g'(a) &= (h\cdot f)'(a)\\ &= h(a)\cdot f'(a) + h'(a)\cdot f(a)\\ &= c\cdot f'(a) + 0\cdot f(a)\\ &= c\cdot f'(a). \blacksquare\end{aligned}$$

Notice, in particular, that $(-f)'(a) = -f'(a)$, and consequently $(f - g)'(a) = (f + [-g])'(a) = f'(a) - g'(a)$.

To demonstrate what we have already achieved, we will compute the derivative of some more special functions.

THEOREM 6 If $f(x) = x^n$ for some natural number n, then

$$f'(a) = na^{n-1} \qquad \text{for all } a.$$

PROOF The proof will be by induction on n. For $n = 1$ this is simply Theorem 2. Now assume that the theorem is true for n, so that if $f(x) = x^n$, then

$$f'(a) = na^{n-1} \qquad \text{for all } a.$$

Let $g(x) = x^{n+1}$. If $I(x) = x$, the equation $x^{n+1} = x^n \cdot x$ can be written

$$g(x) = f(x) \cdot I(x) \qquad \text{for all } x;$$

thus $g = f \cdot I$. It follows from Theorem 4 that

$$\begin{aligned} g'(a) = (f \cdot I)'(a) &= f'(a) \cdot I(a) + f(a) \cdot I'(a) \\ &= na^{n-1} \cdot a + a^n \cdot 1 \\ &= na^n + a^n \\ &= (n+1)a^n, \qquad \text{for all } a. \end{aligned}$$

This is precisely the case $n + 1$ which we wished to prove. ▮

Putting together the theorems proved so far we can now find f' for f of the form

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0.$$

We obtain

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2 x + a_1.$$

We can also find f'':

$$f''(x) = n(n-1)a_n x^{n-2} + (n-1)(n-2)a_{n-1}x^{n-3} + \cdots + 2a_2.$$

This process can be continued easily. Each differentiation reduces the highest power of x by 1, and eliminates one more a_i. It is a good idea to work out the derivatives f''', $f^{(4)}$, and perhaps $f^{(5)}$, until the pattern becomes quite clear. The last interesting derivative is

$$f^{(n)}(x) = n!a_n;$$

for $k > n$ we have

$$f^{(k)}(x) = 0.$$

Clearly, the next step in our program is to find the derivative of a quotient f/g. It is quite a bit simpler, and, because of Theorem 4, obviously sufficient to find the derivative of $1/g$.

THEOREM 7 If g is differentiable at a, and $g(a) \neq 0$, then $1/g$ is differentiable at a, and

$$\left(\frac{1}{g}\right)'(a) = \frac{-g'(a)}{[g(a)]^2}.$$

PROOF Before we even write

$$\frac{\left(\dfrac{1}{g}\right)(a+h) - \left(\dfrac{1}{g}\right)(a)}{h}$$

we must be sure that this expression makes sense—it is necessary to check that $(1/g)(a+h)$ is defined for sufficiently small h. This requires only two observations. Since g is, by hypothesis, differentiable at a, it follows from Theorem 9-1 that g is continuous at a. Since $g(a) \neq 0$, it follows from Theorem 6-3 that there is some $\delta > 0$ such that $g(a+h) \neq 0$ for $|h| < \delta$. Therefore $(1/g)(a+h)$ *does* make sense for small enough h, and we can write

$$\begin{aligned}
\lim_{h\to 0} \frac{\left(\dfrac{1}{g}\right)(a+h) - \left(\dfrac{1}{g}\right)(a)}{h} &= \lim_{h\to 0} \frac{\dfrac{1}{g(a+h)} - \dfrac{1}{g(a)}}{h} \\
&= \lim_{h\to 0} \frac{g(a) - g(a+h)}{h[g(a)\cdot g(a+h)]} \\
&= \lim_{h\to 0} \frac{-[g(a+h) - g(a)]}{h} \cdot \frac{1}{g(a)g(a+h)} \\
&= \lim_{h\to 0} \frac{-[g(a+h) - g(a)]}{h} \cdot \lim_{h\to 0} \frac{1}{g(a)\cdot g(a+h)} \\
&= -g'(a) \cdot \frac{1}{[g(a)]^2}.
\end{aligned}$$

(Notice that we have used continuity of g at a once again.) ▮

The general formula for the derivative of a quotient is now easy to derive. Though not particularly appealing, it is important, and must simply be memorized (I always use the incantation: "bottom times derivative of top, minus top times derivative of bottom, over bottom squared.")

THEOREM 8 If f and g are differentiable at a and $g(a) \neq 0$, then f/g is differentiable at a, and

$$\left(\frac{f}{g}\right)'(a) = \frac{g(a)\cdot f'(a) - f(a)\cdot g'(a)}{[g(a)]^2}.$$

if

$$g(x) = \sin^2 x = \sin x \cdot \sin x,$$

then

$$\begin{aligned} g'(x) &= \sin x \cos x + \cos x \sin x \\ &= 2 \sin x \cos x, \\ g''(x) &= 2[(\sin x)(-\sin x) + \cos x \cos x] \\ &= 2[\cos^2 x - \sin^2 x]; \end{aligned}$$

if

$$h(x) = \cos^2 x = \cos x \cdot \cos x,$$

then

$$\begin{aligned} h'(x) &= (\cos x)(-\sin x) + (-\sin x)\cos x \\ &= -2 \sin x \cos x, \\ h''(x) &= -2[\cos^2 x - \sin^2 x]. \end{aligned}$$

Notice that

$$g'(x) + h'(x) = 0,$$

hardly surprising, since $(g + h)(x) = \sin^2 x + \cos^2 x = 1$. As we would expect, we also have $g''(x) + h''(x) = 0$.

The examples above involved only products of two functions. A function involving triple products can be handled by Theorem 4 also; in fact it can be handled in two ways. Remember that $f \cdot g \cdot h$ is an abbreviation for

$$(f \cdot g) \cdot h \quad \text{or} \quad f \cdot (g \cdot h).$$

Choosing the first of these, for example, we have

$$\begin{aligned} (f \cdot g \cdot h)'(x) &= (f \cdot g)'(x) \cdot h(x) + (f \cdot g)(x)h'(x) \\ &= [f'(x)g(x) + f(x)g'(x)]h(x) + f(x)g(x)h'(x) \\ &= f'(x)g(x)h(x) + f(x)g'(x)h(x) + f(x)g(x)h'(x). \end{aligned}$$

The choice of $f \cdot (g \cdot h)$ would, of course, have given the same result, with a different intermediate step. The final answer is completely symmetric and easily remembered:

> $(f \cdot g \cdot h)'$ is the sum of the three terms obtained by differentiating each of f, g, and h and multiplying by the other two.

For example, if

$$f(x) = x^3 \sin x \cos x,$$

then

$$f'(x) = 3x^2 \sin x \cos x + x^3 \cos x \cos x + x^3(\sin x)(-\sin x).$$

Products of more than 3 functions can be handled similarly. For example, you should have little difficulty deriving the formula

$$\begin{aligned} (f \cdot g \cdot h \cdot k)'(x) = f'(x)g(x)h(x)k(x) &+ f(x)g'(x)h(x)k(x) \\ &+ f(x)g(x)h'(x)k(x) + f(x)g(x)h(x)k'(x). \end{aligned}$$

PROOF Since $f/g = f \cdot (1/g)$ we have

$$\begin{aligned}\left(\frac{f}{g}\right)'(a) &= \left(f \cdot \frac{1}{g}\right)'(a)\\ &= f'(a) \cdot \left(\frac{1}{g}\right)(a) + f(a) \cdot \left(\frac{1}{g}\right)'(a)\\ &= \frac{f'(a)}{g(a)} + \frac{f(a)(-g'(a))}{[g(a)]^2}\\ &= \frac{f'(a) \cdot g(a) - f(a) \cdot g'(a)}{[g(a)]^2}. \blacksquare\end{aligned}$$

We can now differentiate a few more functions. For example,

$$\begin{aligned}&\text{if } f(x) = \frac{x^2-1}{x^2+1}, \text{ then } f'(x) = \frac{(x^2+1)(2x) - (x^2-1)(2x)}{(x^2+1)^2} = \frac{4x}{(x^2+1)^2};\\ &\text{if } f(x) = \frac{x}{x^2+1}, \text{ then } f'(x) = \frac{(x^2+1) - x(2x)}{(x^2+1)^2} = \frac{1-x^2}{(x^2+1)^2};\\ &\text{if } f(x) = \frac{1}{x}, \quad \text{then } f'(x) = -\frac{1}{x^2} = (-1)x^{-2}.\end{aligned}$$

Notice that the last example can be generalized: if

$$f(x) = x^{-n} = \frac{1}{x^n}, \quad \text{for some natural number } n,$$

then

$$f'(x) = \frac{-nx^{n-1}}{x^{2n}} = (-n)x^{-n-1};$$

thus Theorem 6 actually holds both for positive and negative integers. If we interpret $f(x) = x^0$ to mean $f(x) = 1$, and $f'(x) = 0 \cdot x^{-1}$ to mean $f'(x) = 0$, then Theorem 6 is true for $n = 0$ also. (The word "interpret" is necessary because it is not clear how 0^0 should be defined and, in any case, $0 \cdot 0^{-1}$ is meaningless.)

Further progress in differentiation requires the knowledge of the derivatives of certain special functions to be studied later. One of these is the sine function. For the moment we shall divulge, and use, the following information, without proof:

$$\begin{aligned}\sin'(a) &= \cos a \qquad \text{for all } a,\\ \cos'(a) &= -\sin a \qquad \text{for all } a,\end{aligned}$$

This information allows us to differentiate many other functions. For example, if

$$f(x) = x \sin x,$$

then

$$\begin{aligned}f'(x) &= x \cos x + \sin x,\\ f''(x) &= -x \sin x + \cos x + \cos x\\ &= -x \sin x + 2 \cos x;\end{aligned}$$

You might even try to prove (by induction) the general formula:

$$(f_1 \cdot \ldots \cdot f_n)'(x) = \sum_{i=1}^{n} f_1(x) \cdot \ldots \cdot f_{i-1}(x) f_i'(x) f_{i+1}(x) \cdot \ldots \cdot f_n(x).$$

Differentiating the most interesting functions obviously requires a formula for $(f \circ g)'(x)$ in terms of f' and g'. To ensure that $f \circ g$ be differentiable at a, one reasonable hypothesis would seem to be that g be differentiable at a. Since the behavior of $f \circ g$ near a depends on the behavior of f near $g(a)$ (not near a), it also seems reasonable to assume that f is differentiable at $g(a)$. Indeed we shall prove that if g is differentiable at a and f is differentiable at $g(a)$, then $f \circ g$ is differentiable at a, and

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a).$$

This extremely important formula is called the *Chain Rule*, presumable because a composition of functions might be called a "chain" of functions. Notice that $(f \circ g)'$ is practically the product of f' and g', but not quite: f' must be evaluated at $g(a)$ and g' at a. Before attempting to prove this theorem we will try a few applications. Suppose

$$f(x) = \sin x^2.$$

Let us, temporarily, use S to denote the ("squaring") function $S(x) = x^2$. Then

$$f = \sin \circ S.$$

Therefore we have

$$\begin{aligned} f'(x) &= \sin'(S(x)) \cdot S'(x) \\ &= \cos x^2 \cdot 2x. \end{aligned}$$

Quite a different result is obtained if

$$f(x) = \sin^2 x.$$

In this case

$$f = S \circ \sin,$$

so

$$\begin{aligned} f'(x) &= S'(\sin x) \cdot \sin'(x) \\ &= 2 \sin x \cdot \cos x. \end{aligned}$$

Notice that this agrees (as it should) with the result obtained by writing $f = \sin \cdot \sin$ and using the product formula.

Although we have invented a special symbol, S, to name the "squaring" function, it does not take much practice to do problems like this without bothering to write down special symbols for functions, and without even bothering to write down the particular composition which f is—one soon becomes accustomed to taking f apart in one's head. The following differentiations may be used as practice for such mental gymnastics—if you find it necessary to work a few out on paper, by all means do so, but try to develop the knack of writing f' immediately after seeing the definition of f; problems of this sort are so simple that, if you just remember the Chain Rule, there is no thought necessary.

$$\begin{aligned}
&\text{if } f(x) = \sin x^3 &&\text{then } f'(x) = \cos x^3 \cdot 3x^2\\
&f(x) = \sin^3 x &&f'(x) = 3\sin^2 x \cdot \cos x\\
&f(x) = \sin\frac{1}{x} &&f'(x) = \cos\frac{1}{x}\cdot\left(\frac{-1}{x^2}\right)\\
&f(x) = \sin(\sin x) &&f'(x) = \cos(\sin x)\cdot\cos x\\
&f(x) = \sin(x^3+3x^2) &&f'(x) = \cos(x^3+3x^2)\cdot(3x^2+6x)\\
&f(x) = (x^3+3x^2)^{53} &&f'(x) = 53(x^3+3x^2)^{52}\cdot(3x^2+6x).
\end{aligned}$$

A function like

$$f(x) = \sin^2 x^2 = [\sin x^2]^2,$$

which is the composition of three functions,

$$f = S \circ \sin \circ S,$$

can also be differentiated by the Chain Rule. It is only necessary to remember that a triple composition $f \circ g \circ h$ means $(f \circ g) \circ h$ or $f \circ (g \circ h)$. Thus if

$$f(x) = \sin^2 x^2$$

we can write

$$\begin{aligned}
f &= (S \circ \sin) \circ S,\\
f &= S \circ (\sin \circ S).
\end{aligned}$$

The derivative of either expression can be found by applying the Chain Rule twice; the only doubtful point is whether the two expressions lead to equally simple calculations. As a matter of fact, as any experienced differentiator knows, it is much better to use the second:

$$\boxed{f = S \circ (\sin \circ S).}$$

We can now write down $f'(x)$ in one fell swoop. To begin with, note that the first function to be differentiated is S, so the formula for $f'(x)$ begins

$$f'(x) = 2(\qquad)$$

Inside the parentheses we must put $\sin x^2$, the value at x of the second function, $\sin \circ S$. Thus we begin by writing

$$f'(x) = 2\sin x^2$$

(the parentheses weren't really necessary, after all). We must now multiply this much of the answer by the derivative of $\sin \circ S$ at x; this part is easy—it involves a composition of two functions, which we already know how to handle. We obtain, for the final answer,

$$f'(x) = 2\sin x^2 \cdot \cos x^2 \cdot 2x.$$

The following example is handled similarly. Suppose

$$\boxed{f(x) = \sin(\sin x^2).}$$

Without even bothering to write down f as a composition $g \circ h \circ k$ of three functions, we can see that the left-most one will be sin, so our expression for $f'(x)$ begins

$$f'(x) = \cos(\qquad)$$

Inside the parentheses we must put the value of $h \circ k(x)$; this is simply $\sin x^2$ (what you get from $\sin(\sin x^2)$ by deleting the first sin). So our expression for $f'(x)$ begins

$$f'(x) = \cos(\sin x^2)$$

We can now forget about the first sin in $\sin(\sin x^2)$; we have to multiply what we have so far by the derivative of the function whose value at x is $\sin x^2$—which is again a problem we already know how to solve:

$$f'(x) = \cos(\sin x^2) \cdot \cos x^2 \cdot 2x.$$

Finally, here are the derivatives of some other functions which are the composition of sin and S, as well as some other triple compositions. You can probably just "see" that the answers are correct—if not, try writing out f as a composition:

$$\begin{aligned}
\text{if } f(x) &= \sin((\sin x)^2) & \text{then } f'(x) &= \cos((\sin x)^2) \cdot 2\sin x \cdot \cos x \\
f(x) &= [\sin(\sin x)]^2 & f'(x) &= 2\sin(\sin x) \cdot \cos(\sin x) \cdot \cos x \\
f(x) &= \sin(\sin(\sin x)) & f'(x) &= \cos(\sin(\sin x)) \cdot \cos(\sin x) \cdot \cos x \\
f(x) &= \sin^2(x \sin x) & f'(x) &= 2\sin(x \sin x) \cdot \cos(x \sin x) \\
& & &\qquad \cdot [\sin x + x\cos x] \\
f(x) &= \sin(\sin(x^2 \sin x)) & f'(x) &= \cos(\sin(x^2 \sin x)) \cdot \cos(x^2 \sin x) \\
& & &\qquad \cdot [2x\sin x + x^2\cos x].
\end{aligned}$$

The rule for treating compositions of four (or even more) functions is easy—always (mentally) put in parentheses starting from the right,

$$f \circ (g \circ (h \circ k)),$$

and start reducing the calculation to the derivative of a composition of a smaller number of functions:

$$f'(g(h(k(x))))$$

For example, if

$$f(x) = \sin^2(\sin^2(x)) \qquad \begin{aligned}[t] [f &= S \circ \sin \circ S \circ \sin \\ &= S \circ (\sin \circ (S \circ \sin))] \end{aligned}$$

then

$$f'(x) = 2\sin(\sin^2 x) \cdot \cos(\sin^2 x) \cdot 2\sin x \cdot \cos x;$$

if

$$f(x) = \sin((\sin x^2)^2) \qquad \begin{aligned}[t] [f &= \sin \circ S \circ \sin \circ S \\ &= \sin \circ (S \circ (\sin \circ S))] \end{aligned}$$

then

$$f'(x) = \cos((\sin x^2)^2) \cdot 2 \sin x^2 \cdot \cos x^2 \cdot 2x;$$

if

$$f(x) = \sin^2(\sin(\sin x)) \qquad \text{[fill in yourself, if necessary]}$$

then

$$f'(x) = 2 \sin(\sin(\sin x)) \cdot \cos(\sin(\sin x)) \cdot \cos(\sin x) \cdot \cos x.$$

With these examples as reference, you require only one thing to become a master differentiator—practice. You can be safely turned loose on the exercises at the end of the chapter, and it is now high time that we proved the Chain Rule.

The following argument, while not a proof, indicates some of the tricks one might try, as well as some of the difficulties encountered. We begin, of course, with the definition—

$$\begin{aligned} (f \circ g)'(a) &= \lim_{h \to 0} \frac{(f \circ g)(a+h) - (f \circ g)(a)}{h} \\ &= \lim_{h \to 0} \frac{f(g(a+h)) - f(g(a))}{h}. \end{aligned}$$

Somewhere in here we would like the expression for $g'(a)$. One approach is to put it in by fiat:

$$\lim_{h \to 0} \frac{f(g(a+h)) - f(g(a))}{h} = \lim_{h \to 0} \frac{f(g(a+h)) - f(g(a))}{g(a+h) - g(a)} \cdot \frac{g(a+h) - g(a)}{h}.$$

This does not look bad, and it looks even better if we write

$$\begin{aligned} &\lim_{h \to 0} \frac{(f \circ g)(a+h) - (f \circ g)(a)}{h} \\ &\qquad = \lim_{h \to 0} \frac{f(g(a) + [g(a+h) - g(a)]) - f(g(a))}{g(a+h) - g(a)} \cdot \lim_{h \to 0} \frac{g(a+h) - g(a)}{h}. \end{aligned}$$

The second limit is the factor $g'(a)$ which we want. If we let $g(a+h) - g(a) = k$ (to be precise we should write $k(h)$), then the first limit is

$$\lim_{h \to 0} \frac{f(g(a) + k) - f(g(a))}{k}.$$

It looks as if this limit should be $f'(g(a))$, since continuity of g at a implies that k goes to 0 as h does. In fact, one can, and we soon will, make this sort of reasoning precise. There is already a problem, however, which you will have noticed if you are the kind of person who does not divide blindly. Even for $h \neq 0$ we might have $g(a+h) - g(a) = 0$, making the division and multiplication by $g(a+h) - g(a)$ meaningless. True, we only care about small h, but $g(a+h) - g(a)$ could be 0 for arbitrarily small h. The easiest way this can happen is for g to be a constant function, $g(x) = c$. Then $g(a+h) - g(a) = 0$ for all h. In this case, $f \circ g$ is also a constant function, $(f \circ g)(x) = f(c)$, so the Chain Rule does indeed hold:

$$(f \circ g)'(a) = 0 = f'(g(a)) \cdot g'(a).$$

However, there are also nonconstant functions g for which $g(a+h) - g(a) = 0$ for arbitrarily small h. For example, if $a = 0$, the function g might be

$$g(x) = \begin{cases} x^2 \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

In this case, $g'(0) = 0$, as we showed in Chapter 9. If the Chain Rule is correct, we must have $(f \circ g)'(0) = 0$ for any differentiable f, and this is not exactly obvious. A proof of the Chain Rule can be found by considering such recalcitrant functions separately, but it is easier simply to abandon this approach, and use a trick.

THEOREM 9 (THE CHAIN RULE) If g is differentiable at a, and f is differentiable at $g(a)$, then $f \circ g$ is differentiable at a, and

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a).$$

PROOF Define a function ϕ as follows:

$$\phi(h) = \begin{cases} \dfrac{f(g(a+h)) - f(g(a))}{g(a+h) - g(a)}, & \text{if } g(a+h) - g(a) \neq 0 \\ f'(g(a)), & \text{if } g(a+h) - g(a) = 0. \end{cases}$$

It should be intuitively clear that ϕ is continuous at 0: When h is small, $g(a+h) - g(a)$ is also small, so if $g(a+h) - g(a)$ is not zero, then $\phi(h)$ will be close to $f'(g(a))$; and if it is zero, then $\phi(h)$ actually equals $f'(g(a))$, which is even better. Since the continuity of ϕ is the crux of the whole proof we will provide a careful translation of this intuitive argument.

We know that f is differentiable at $g(a)$. This means that

$$\lim_{k \to 0} \frac{f(g(a)+k) - f(g(a))}{k} = f'(g(a)).$$

Thus, if $\varepsilon > 0$ there is some number $\delta' > 0$ such that, for all k,

$$(1) \quad \text{if } 0 < |k| < \delta', \text{ then } \left| \frac{f(g(a)+k) - f(g(a))}{k} - f'(g(a)) \right| < \varepsilon.$$

Now g is differentiable at a, hence continuous at a, so there is a $\delta > 0$ such that, for all h,

$$(2) \quad \text{if } |h| < \delta, \text{ then } |g(a+h) - g(a)| < \delta'.$$

Consider now any h with $|h| < \delta$. If $k = g(a+h) - g(a) \neq 0$, then

$$\phi(h) = \frac{f(g(a+h)) - f(g(a))}{g(a+h) - g(a)} = \frac{f(g(a)+k) - f(g(a))}{k};$$

it follows from (2) that $|k| < \delta'$, and hence from (1) that

$$|\phi(h) - f'(g(a))| < \varepsilon.$$

On the other hand, if $g(a+h) - g(a) = 0$, then $\phi(h) = f'(g(a))$, so it is surely true that

$$|\phi(h) - f'(g(a))| < \varepsilon.$$

We have therefore proved that

$$\lim_{h \to 0} \phi(h) = f'(g(a)),$$

so ϕ is continuous at 0. The rest of the proof is easy. If $h \neq 0$, then we have

$$\frac{f(g(a+h)) - f(g(a))}{h} = \phi(h) \cdot \frac{g(a+h) - g(a)}{h}$$

even if $g(a+h) - g(a) = 0$ (because in that case both sides are 0). Therefore

$$\begin{aligned}(f \circ g)'(a) = \lim_{h \to 0} \frac{f(g(a+h)) - f(g(a))}{h} &= \lim_{h \to 0} \phi(h) \cdot \lim_{h \to 0} \frac{g(a+h) - g(a)}{h} \\ &= f'(g(a)) \cdot g'(a). \ \blacksquare\end{aligned}$$

Now that we can differentiate so many functions so easily we can take another look at the function

$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

In Chapter 9 we showed that $f'(0) = 0$, working straight from the definition (the only possible way). For $x \neq 0$ we can use the methods of this chapter. We have

$$f'(x) = 2x \sin \frac{1}{x} + x^2 \cos \frac{1}{x} \cdot \left(-\frac{1}{x^2}\right);$$

Thus

$$f'(x) = \begin{cases} 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

As this formula reveals, the first derivative f' is indeed badly behaved at 0—it is not even continuous there. If we consider instead

$$f(x) = \begin{cases} x^3 \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0, \end{cases}$$

then

$$f'(x) = \begin{cases} 3x^2 \sin \dfrac{1}{x} - x \cos \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

In this case f' is continuous at 0, but $f''(0)$ does not exist (because the expression $3x^2 \sin 1/x$ defines a function which is differentiable at 0 but the expression $-x \cos 1/x$ does not).

As you may suspect, increasing the power of x yet again produces another improvement. If

$$f(x) = \begin{cases} x^4 \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0, \end{cases}$$

then

$$f'(x) = \begin{cases} 4x^3 \sin \dfrac{1}{x} - x^2 \cos \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

It is easy to compute, right from the definition, that $(f')'(0) = 0$, and $f''(x)$ is easy to find for $x \neq 0$:

$$f''(x) = \begin{cases} 12x^2 \sin \dfrac{1}{x} - 4x \cos \dfrac{1}{x} - 2x \cos \dfrac{1}{x} - \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

In this case, the *second* derivative f'' is not continuous at 0. By now you may have guessed the pattern, which two of the problems ask you to establish: if

$$f(x) = \begin{cases} x^{2n} \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0, \end{cases}$$

then $f'(0), \dots, f^{(n)}(0)$ exist, but $f^{(n)}$ is not continuous at 0; if

$$f(x) = \begin{cases} x^{2n+1} \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0, \end{cases}$$

then $f'(0), \dots, f^{(n)}(0)$ exist, and $f^{(n)}$ is continuous at 0, but $f^{(n)}$ is not differentiable at 0. These examples may suggest that "reasonable" functions can be characterized by the possession of higher-order derivatives—no matter how hard we try to mask the infinite oscillation of $f(x) = \sin 1/x$, a derivative of sufficiently high order seems able to reveal the underlying irregularity. Unfortunately, we will see later that much worse things can happen.

After all these involved calculations, we will bring this chapter to a close with a minor remark. It is often tempting, and seems more elegant, to write some of the theorems in this chapter as equations about functions, rather than about their values. Thus Theorem 3 might be written

$$(f + g)' = f' + g',$$

Theorem 4 might be written as

$$(f \cdot g)' = f \cdot g' + f' \cdot g,$$

and Theorem 9 often appears in the form

$$(f \circ g)' = (f' \circ g) \cdot g'.$$

Strictly speaking, these equations may be false, because the functions on the left-hand side might have a larger domain than those on the right. Nevertheless, this is hardly worth worrying about. If f and g are differentiable everywhere in their domains, then these equations, and others like them, *are* true, and this is the only case any one cares about.

PROBLEMS

1. As a warm up exercise, find $f'(x)$ for each of the following f. (Don't worry about the domain of f or f'; just get a formula for $f'(x)$ that gives the right answer when it makes sense.)

(i) $f(x) = \sin(x + x^2)$.
(ii) $f(x) = \sin x + \sin x^2$.
(iii) $f(x) = \sin(\cos x)$.
(iv) $f(x) = \sin(\sin x)$.
(v) $f(x) = \sin\left(\dfrac{\cos x}{x}\right)$.
(vi) $f(x) = \dfrac{\sin(\cos x)}{x}$.
(vii) $f(x) = \sin(x + \sin x)$.
(viii) $f(x) = \sin(\cos(\sin x))$.

2. Find $f'(x)$ for each of the following functions f. (It took the author 20 minutes to compute the derivatives for the answer section, and it should not take you much longer. Although rapid calculation is not the goal of mathematics, if you hope to treat theoretical applications of the Chain Rule with aplomb, these concrete applications should be child's play—mathematicians like to pretend that they can't even add, but most of them can when they have to.)

(i) $f(x) = \sin((x+1)^2(x+2))$.
(ii) $f(x) = \sin^3(x^2 + \sin x)$.
(iii) $f(x) = \sin^2((x + \sin x)^2)$.
(iv) $f(x) = \sin\left(\dfrac{x^3}{\cos x^3}\right)$.
(v) $f(x) = \sin(x \sin x) + \sin(\sin x^2)$.
(vi) $f(x) = (\cos x)^{31^2}$.
(vii) $f(x) = \sin^2 x \sin x^2 \sin^2 x^2$.
(viii) $f(x) = \sin^3(\sin^2(\sin x))$.
(ix) $f(x) = (x + \sin^5 x)^6$.
(x) $f(x) = \sin(\sin(\sin(\sin(\sin x))))$.
(xi) $f(x) = \sin((\sin^7 x^7 + 1)^7)$.
(xii) $f(x) = (((x^2 + x)^3 + x)^4 + x)^5$.
(xiii) $f(x) = \sin(x^2 + \sin(x^2 + \sin x^2))$.
(xiv) $f(x) = \sin(6\cos(6\sin(6\cos 6x)))$.

(xv) $f(x) = \dfrac{\sin x^2 \sin^2 x}{1 + \sin x}$.

(xvi) $f(x) = \dfrac{1}{x - \dfrac{2}{x + \sin x}}$.

(xvii) $f(x) = \sin\left(\dfrac{x^3}{\sin\left(\dfrac{x^3}{\sin x}\right)}\right)$.

(xviii) $f(x) = \sin\left(\dfrac{x}{x - \sin\left(\dfrac{x}{x - \sin x}\right)}\right)$.

3. Find the derivatives of the functions tan, cotan, sec, cosec. (You don't have to memorize these formulas, although they will be needed once in a while; if you express your answers in the right way, they will be simple and somewhat symmetrical.)

4. For each of the following functions f, find $f'(f(x))$ (*not* $(f \circ f)'(x)$).

(i) $f(x) = \dfrac{1}{1+x}$.
(ii) $f(x) = \sin x$.
(iii) $f(x) = x^2$.
(iv) $f(x) = 17$.

5. For each of the following functions f, find $f(f'(x))$.

(i) $f(x) = \dfrac{1}{x}$.
(ii) $f(x) = x^2$.
(iii) $f(x) = 17$.
(iv) $f(x) = 17x$.

6. Find f' in terms of g' if

(i) $f(x) = g(x + g(a))$.
(ii) $f(x) = g(x \cdot g(a))$.
(iii) $f(x) = g(x + g(x))$.
(iv) $f(x) = g(x)(x - a)$.
(v) $f(x) = g(a)(x - a)$.
(vi) $f(x + 3) = g(x^2)$.

7. (a) A circular object is increasing in size in some unspecified manner, but it is known that when the radius is 6, the rate of change of the radius is 4. Find the rate of change of the area when the radius is 6. (If $r(t)$ and $A(t)$ represent the radius and the area at time t, then the functions r and A satisfy $A = \pi r^2$; a straightforward use of the Chain Rule is called for.)

***21.** Prove that if $f^{(n)}(g(a))$ and $g^{(n)}(a)$ both exist, then $(f \circ g)^{(n)}(a)$ exists. A little experimentation should convince you that it is unwise to seek a formula for $(f \circ g)^{(n)}(a)$. In order to prove that $(f \circ g)^{(n)}(a)$ exists you will therefore have to devise a reasonable assertion about $(f \circ g)^{(n)}(a)$ which can be proved by induction. Try something like: "$(f \circ g)^{(n)}(a)$ exists and is a sum of terms each of which is a product of terms of the form . . . ."

22. (a) If $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_0$, find a function g such that $g' = f$. Find another.

(b) If

$$f(x) = \frac{b_2}{x^2} + \frac{b_3}{x^3} + \cdots + \frac{b_m}{x^m},$$

find a function g with $g' = f$.

(c) Is there a function

$$f(x) = a_nx^n + \cdots + a_0 + \frac{b_1}{x} + \cdots + \frac{b_m}{x^m}$$

such that $f'(x) = 1/x$?

23. Show that there is a polynomial function f of degree n such that

(a) $f'(x) = 0$ for precisely $n - 1$ numbers x.
(b) $f'(x) = 0$ for no x, if n is odd.
(c) $f'(x) = 0$ for exactly one x, if n is even.
(d) $f'(x) = 0$ for exactly k numbers x, if $n - k$ is odd.

24. (a) The number a is called a **double root** of the polynomial function f if $f(x) = (x - a)^2g(x)$ for some polynomial function g. Prove that a is a double root of f if and only if a is a root of both f and f'.

(b) When does $f(x) = ax^2 + bx + c$ ($a \neq 0$) have a double root? What does the condition say geometrically?

25. If f is differentiable at a, let $d(x) = f(x) - f'(a)(x - a) - f(a)$. Find $d'(a)$. In connection with Problem 24, this gives another solution for Problem 9-20.

***26.** This problem is a companion to Problem 3-6. Let $a_1, \dots, a_n$ and $b_1, \dots, b_n$ be given numbers.

(a) If $x_1, \dots, x_n$ are distinct numbers, prove that there is a polynomial function f of degree $2n - 1$, such that $f(x_j) = f'(x_j) = 0$ for $j \neq i$, and $f(x_i) = a_i$ and $f'(x_i) = b_i$. Hint: Remember Problem 24.

(b) Prove that there is a polynomial function f of degree $2n - 1$ with $f(x_i) = a_i$ and $f'(x_i) = b_i$ for all i.

***27.** Suppose that a and b are two consecutive roots of a polynomial function f, but that a and b are not double roots, so that we can write $f(x) = (x - a)(x - b)g(x)$ where $g(a) \neq 0$ and $g(b) \neq 0$.

(a) Prove that $g(a)$ and $g(b)$ have the same sign. (Remember that a and b are consecutive roots.)

14. Prove similarly that the tangent lines to an ellipse or hyperbola intersect these sets only once.

15. If $f + g$ is differentiable at a, are f and g necessarily differentiable at a? If $f \cdot g$ and f are differentiable at a, what conditions on f imply that g is differentiable at a?

16. (a) Prove that if f is differentiable at a, then $|f|$ is also differentiable at a, provided that $f(a) \neq 0$.
(b) Give a counterexample if $f(a) = 0$.
(c) Prove that if f and g are differentiable at a, then the functions $\max(f, g)$ and $\min(f, g)$ are differentiable at a, provided that $f(a) \neq g(a)$.
(d) Give a counterexample if $f(a) = g(a)$.

17. Give an example of functions f and g such that g takes on all values, and $f \circ g$ and g are differentiable, but f isn't differentiable. (The problem becomes trivial if we don't require that g takes on all values; g could just be a constant function, or a function that only takes on values in some interval (a, b), in which case the behavior of f outside of (a, b) would be irrelevant.)

18. (a) If $g = f^2$ find a formula for g' (involving f').
(b) If $g = (f')^2$, find a formula for g' (involving f'').
(c) Suppose that the function $f > 0$ has the property that

$$(f')^2 = f + \frac{1}{f^3}.$$

Find a formula for f'' in terms of f. (In addition to simple calculations, a bit of care is needed at one point.)

19. If f is three times differentiable and $f'(x) \neq 0$, the *Schwarzian derivative* of f at x is defined to be

$$\mathcal{D}f(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2.$$

(a) Show that

$$\mathcal{D}(f \circ g) = [\mathcal{D}f \circ g] \cdot g'^2 + \mathcal{D}g.$$

(b) Show that if $f(x) = \dfrac{ax + b}{cx + d}$, with $ad - bc \neq 0$, then $\mathcal{D}f = 0$. Consequently, $\mathcal{D}(f \circ g) = \mathcal{D}g$.

20. Suppose that $f^{(n)}(a)$ and $g^{(n)}(a)$ exist. Prove *Leibniz's formula*:

$$(f \cdot g)^{(n)}(a) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(a) \cdot g^{(n-k)}(a).$$

(b) Suppose that we are now informed that the circular object we have been watching is really the cross section of a spherical object. Find the rate of change of the *volume* when the radius is 6. (You will clearly need to know a formula for the volume of a sphere; in case you have forgotten, the volume is $\frac{4}{3}\pi$ times the cube of the radius.)

(c) Now suppose that the rate of change of the area of the circular cross section is 5 when the radius is 3. Find the rate of change of the volume when the radius is 3. You should be able to do this problem in two ways: first, by using the formulas for the area and volume in terms of the radius; and then by expressing the volume in terms of the area (to use this method you will need Problem 9-3).

8. The area between two varying concentric circles is at all times 9π in^2. The rate of change of the area of the larger circle is 10π in^2/sec. How fast is the circumference of the smaller circle changing when it has area 16π in^2?

9. Particle A moves along the positive horizontal axis, and particle B along the graph of $f(x) = -\sqrt{3}x$, $x \le 0$. At a certain time, A is at the point $(5,0)$ and moving with speed 3 units/sec; and B is at a distance of 3 units from the origin and moving with speed 4 units/sec. At what rate is the distance between A and B changing?

10. Let $f(x) = x^2 \sin 1/x$ for $x \neq 0$, and let $f(0) = 0$. Suppose also that h and k are two functions such that

$$\begin{array}{ll} h'(x) = \sin^2(\sin(x+1)) & \qquad k'(x) = f(x+1) \\ h(0) = 3 & \qquad k(0) = 0. \end{array}$$

Find

(i) $(f \circ h)'(0)$.
(ii) $(k \circ f)'(0)$.
(iii) $\alpha'(x^2)$, where $\alpha(x) = h(x^2)$. Exercise great care.

11. Find $f'(0)$ if

$$f(x) = \begin{cases} g(x)\sin\dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0, \end{cases}$$

and

$$g(0) = g'(0) = 0.$$

12. Using the derivative of $f(x) = 1/x$, as found in Problem 9-1, find $(1/g)'(x)$ by the Chain Rule.

13. (a) Using Problem 9-3, find $f'(x)$ for $-1 < x < 1$, if $f(x) = \sqrt{1-x^2}$.

(b) Prove that the tangent line to the graph of f at $(a, \sqrt{1-a^2})$ intersects the graph only at that point (and thus show that the elementary geometry definition of the tangent line coincides with ours).

(b) Prove that there is some number x with $a < x < b$ and $f'(x) = 0$. (Also draw a picture to illustrate this fact.) Hint: Compare the sign of $f'(a)$ and $f'(b)$.

(c) Now prove the same fact, even if a and b are multiple roots. Hint: If $f(x) = (x-a)^m(x-b)^n g(x)$ where $g(a) \neq 0$ and $g(b) \neq 0$, consider the polynomial function $h(x) = f'(x)/(x-a)^{m-1}(x-b)^{n-1}$.

This theorem was proved by the French mathematician Rolle, in connection with the problem of approximating roots of polynomials, but the result was not originally stated in terms of derivatives. In fact, Rolle was one of the mathematicians who never accepted the new notions of calculus. This was not such a pigheaded attitude, in view of the fact that for one hundred years no one could define limits in terms that did not verge on the mystic, but on the whole history has been particularly kind to Rolle; his name has become attached to a much more general result, to appear in the next chapter, which forms the basis for the most important theoretical results of calculus.

28. Suppose that $f(x) = xg(x)$ for some function g which is continuous at 0. Prove that f is differentiable at 0, and find $f'(0)$ in terms of g.

29. Suppose f is differentiable at 0, and that $f(0) = 0$. Prove that $f(x) = xg(x)$ for some function g which is continuous at 0. Hint: What happens if you try to write $g(x) = f(x)/x$?

30. If $f(x) = x^{-n}$ for n in $\mathbf{N}$, prove that

$$\begin{aligned} f^{(k)}(x) &= (-1)^k \frac{(n+k-1)!}{(n-1)!} x^{-n-k} \\ &= (-1)^k k! \binom{n+k-1}{k} x^{-n-k}, \qquad \text{for } x \neq 0. \end{aligned}$$

***31.** Prove that it is impossible to write $x = f(x)g(x)$ where f and g are differentiable and $f(0) = g(0) = 0$. Hint: Differentiate.

32. What is $f^{(k)}(x)$ if

(a) $f(x) = 1/(x-a)^n$?

*(b) $f(x) = 1/(x^2-1)$?

***33.** Let $f(x) = x^{2n} \sin 1/x$ if $x \neq 0$, and let $f(0) = 0$. Prove that $f'(0), \dots, f^{(n)}(0)$ exist, and that $f^{(n)}$ is not continuous at 0. (You will encounter the same basic difficulty as that in Problem 21.)

***34.** Let $f(x) = x^{2n+1} \sin 1/x$ if $x \neq 0$, and let $f(0) = 0$. Prove that $f'(0), \dots, f^{(n)}(0)$ exist, that $f^{(n)}$ is continuous at 0, and that $f^{(n)}$ is not differentiable at 0.

35. In Leibnizian notation the Chain Rule ought to read:

$$\frac{df(g(x))}{dx} = \left.\frac{df(y)}{dy}\right|_{y=g(x)} \cdot \frac{dg(x)}{dx}.$$

Instead, one usually finds the following statement: "Let $y = g(x)$ and $z = f(y)$. Then

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}."$$

Notice that the z in dz/dx denotes the composite function $f \circ g$, while the z in dz/dy denotes the function f; it is also understood that dz/dy will be "an expression involving y," and that in the final answer $g(x)$ must be substituted for y. In each of the following cases, find dz/dx by using this formula; then compare with Problem 1.

(i) $z = \sin y, \quad y = x + x^2.$
(ii) $z = \sin y, \quad y = \cos x.$
(iii) $z = \sin u, \quad u = \sin x.$
(iv) $z = \sin v, \quad v = \cos u, \quad u = \sin x.$