Chapter 7

DIFFERENTIATION

7.1 Introduction

Calculus courses succeed very well in conveying an idea of what a derivative is and the students develop many technical skills in computations of derivatives or applications of them. We shall return to the subject of derivatives but with a different objective.

Now we wish to see a little deeper and to understand the basis on which that theory develops. Much of this chapter will appear to be a review of the subject of derivatives with rather more attention paid to the details now, and less to the applications. Some of the more advanced material will be, however, completely new. We start at the beginning, at the rudiments of the theory of derivatives.

7.2 The Derivative

Let f be a function defined on an interval I and let x_0 and x be points of I. Consider the *difference quotient* determined by the points x_0 and x:

$$\frac{f(x)-f(x_0)}{x-x_0}. \tag{1}$$

This difference quotient represents the average rate of change of the function on the interval $[x_0, x]$ (or $[x, x_0]$ if $x < x_0$).

In Figure 7.1 this difference quotient represents the slope of the chord (or secant line) determined by the points $(x, f(x))$ and $(x_0, f(x_0))$.

This same picture allows a physical interpretation. If $f(x)$ represents the distance a point moving on a straight line has moved from

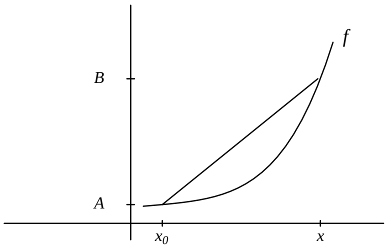


Figure 7.1: The chord determined by $(x, f(x))$ and $(x_0, f(x_0))$.

some fixed point in time x, then $f(x) - f(x_0)$ represents the (net) distance it has moved in the time interval $[x_0, x]$, and the difference quotient (1) represents the average velocity in that time interval.

Suppose now that we fix x_0, and allow x to approach x_0. We learn in elementary calculus that if

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists, then the limit represents the slope of the tangent line to the graph of the function f at the point $(x_0, f(x_0))$. In the setting of motion, the limit represents instantaneous velocity at time x_0.

The derivative owes it origins to these two interpretations in geometry and in the physics of motion but now completely transcends them; the derivative finds applications in nearly every part of mathematics and the sciences.

We shall study the structure of derivatives, but with less concern for computations and applications than we would have seen in our calculus courses. Now we wish to understand the notion and see why it has the properties used in the many computations and applications of the calculus.

7.2.1 Definition of the Derivative

We begin with a familiar definition.

Definition 7.1 Let f be defined on an interval I and let $x_0 \in I$. The *derivative* of f at x_0, denoted by $f'(x_0)$, is defined as

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}, \tag{2}$$

provided either that this limit exists or is infinite. If $f'(x_0)$ is finite we say that f is *differentiable* at x_0. If f is differentiable at every point of a set $E \subset I$, we say that f is *differentiable* on E. When E is all of I, we simply say that f is a *differentiable* function.

Note. We have allowed infinite derivatives and they do play a role in many studies, but differentiable always refers to a finite derivative. Normally the phrase "a derivative exists" also means that that derivative is finite.

Example 7.2 Let $f(x) = x^2$ on $\mathbb{R}$ and let $x_0 \in \mathbb{R}$. If $x \in \mathbb{R}$, $x \neq x_0$, then

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x^2 - x_0^2}{x - x_0} = \frac{(x - x_0)(x + x_0)}{(x - x_0)}.$$

Since $x \neq x_0$, the last expression equals $x + x_0$, so

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} (x + x_0) = 2x_0.$$

This establishes the familiar formula, $f'(x_0) = 2x_0$ for the function $f(x) = x^2$. ◀

Let us take a moment to clarify the definition when the interval I contains one or both of its endpoints. Suppose $I = [a, b]$. For $x_0 = a$ (or $x_0 = b$), the limit in (2) is just a one-sided, or unilateral, limit. The function f is defined only on $[a, b]$ so we cannot consider points outside of that interval.

This brings us to another point. It can happen that a function that is *not* differentiable at a point x_0 does satisfy the requirement of (2) from one side of x_0. This means that the limit in (2) exists as $x \to x_0$ from that side. We present a formal definition.

Definition 7.3 Let f be defined on an interval I and let $x_0 \in I$. The *right-hand derivative* of f at x_0, denoted by $f'_+(x_0)$ is the limit

$$f'_+(x_0) = \lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0},$$

provided that unilateral limit exists or is infinite. Similarly the *left-hand derivative* of f at x_0, $f'_-(x_0)$, is the limit

$$f'_-(x_0) = \lim_{x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0}.$$

Observe, if x_0 is an interior point of I, then $f'(x_0)$ exists if and only if $f'_+(x_0) = f'_-(x_0)$. (Exercise 7:2.8)

Example 7.4 Let $f(x) = |x|$ on $\mathbb{R}$. Let us consider the differentiability of f at $x_0 = 0$. The difference quotient (1) becomes

$$\frac{f(x) - f(0)}{x - 0} = \frac{|x|}{x} = \begin{cases} 1, & \text{if } x > 0 \\ -1, & \text{if } x < 0. \end{cases}$$

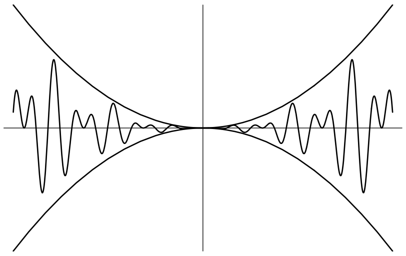


Figure 7.2: A function trapped between x^2 and $-x^2$.

Thus $f'_+(0) = \lim_{x\to x_0+} \frac{|x|}{x} = 1$ while $f'_-(0) = \lim_{x\to x_0-} \frac{|x|}{x} = -1$. The function has different right-hand and left-hand derivatives at $x_0 = 0$ so is not differentiable at $x_0 = 0$. ◀

Example 7.5 (A "trapping principle".) Let f be any function defined in a neighborhood I of zero. Suppose f satisfies the inequality $|f(x)| \leq x^2$ for all $x \in I$. Thus, the graph of f is "trapped" between the parabolas $y = x^2$ and $y = -x^2$. In particular $f(0) = 0$. The difference quotient computed for $x_0 = 0$ becomes

$$\frac{f(x) - f(0)}{x - 0} = \frac{f(x)}{x},$$

from which we calculate

$$\left|\frac{f(x)}{x}\right| \leq \left|\frac{x^2}{x}\right| = |x|, \quad \text{so} \quad \lim_{x\to 0}\left|\frac{f(x)}{x}\right| \leq \lim_{x\to 0} |x| = 0.$$

Thus $\lim_{x\to 0} \frac{f(x)}{x} = 0$. As a result, $f'(0) = 0$. Figure 7.2 illustrates the principle. ◀

Higher Order Derivatives When a function f is differentiable on I, it is possible that its derivative f' is also differentiable. When this is the case, the function $f'' = (f')'$ is called the second derivative of the function f. Inductively, we can define derivatives of all orders: $f^{(n+1)} = (f^{(n)})'$ (provided $f^{(n)}$ is differentiable). When n is small, it is customary to use the convenient notation f'' for $f^{(2)}$, f''' for $f^{(3)}$ etc.

Notation It is useful to have other notations for the derivative of a function f. Common notations are $\frac{df}{dx}$ and $\frac{dy}{dx}$ (when the function is expressed in the form $y = f(x)$). Another notation that is useful is Df. These alternate notations along with slight variations are useful for various calculations. The reader is no doubt familiar with such uses—the convenience of writing

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

when using the chain rule, or viewing "D" as an operator in solving linear differential equations. Notation *can* be important at times. Consider, for example, how difficult it would be to perform a simple arithmetic calculation such as the multiplication (104)(90) using Roman numerals (CIV)(XC)!

Exercises

7:2.1 The reader might be familiar with a slightly different formulation of the definition of derivative. If x_0 is interior to I, then for h sufficiently small, the point $x_0 + h$ is also in I. Show that expression (2) then reduces to

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

Repeat Examples 7.2 and 7.4 using this formulation of the derivative.

7:2.2 Let $c \in \mathbb{R}$. Calculate the derivatives of the functions $g(x) = c$ and $k(x) = x$ directly from the definition of derivative.

7:2.3 Check the differentiability of each of the functions below at $x_0 = 0$.

(a) $f(x) = x|x|$

(b) $f(x) = x \sin x^{-1}$ ($f(0) = 0$)

(c) $f(x) = x^2 \sin x^{-1}$ ($f(0) = 0$)

(d) $f(x) = \begin{cases} x^2, & \text{if } x \text{ rational} \\ 0, & \text{if } x \text{ irrational} \end{cases}$

7:2.4 Let $f(x) = \begin{cases} x^2, & \text{if } x \geq 0 \\ ax, & \text{if } x < 0 \end{cases}$

(a) For which values of a is f differentiable at $x = 0$?

(b) For which values of a is f continuous at $x = 0$?

(c) When f is differentiable at $x = 0$, does $f''(0)$ exist?

7:2.5 For what positive values of p is the function $f(x) = |x|^p$ differentiable at 0?

7:2.6 A function f has a *symmetric derivative* at a point if

$$f_s'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x-h)}{2h}$$

exists. Show that $f_s'(x) = f'(x)$ at any point at which the latter exists but that $f_s'(x)$ may exist even when f is not differentiable at x.

7:2.7 Find all points where $f(x) = \sqrt{1-\cos x}$ is not differentiable and at those points find the one-sided derivatives.

7:2.8 Prove that if x_0 is an interior point of an interval I, then $f'(x_0)$ exists or is infinite if and only if $f_+'(x_0) = f_-'(x_0)$.

7:2.9 Let a function $f : \mathbb{R} \to \mathbb{R}$ be defined by setting $f(1/n) = c_n$ for $n = 1, 2, 3, \ldots$ where $\{c_n\}$ is a given sequence and elsewhere $f(x) = 0$. Find a condition on that sequence so that $f'(0)$ exists.

7:2.10 Let a function $f : \mathbb{R} \to \mathbb{R}$ be defined by setting $f(1/n^2) = c_n$ for $n = 1, 2, 3, \ldots$ where $\{c_n\}$ is a given sequence and elsewhere $f(x) = 0$. Find a condition on that sequence so that $f'(0)$ exists.

7:2.11 Give an example of a function with an infinite derivative at some point. Give an example of a function f with $f_+'(x_0) = \infty$ and $f_-'(x_0) = -\infty$ at some point x_0.

7:2.12 If $f'(x_0) > 0$ for some point x_0 in the interior of the domain of f show that there is a $\delta > 0$ so that

$$f(x) < f(x_0) < f(y)$$

whenever $x_0 - \delta < x < x_0 < y < x_0 + \delta$. Does this assert that f is increasing in the interval $(x_0 - \delta, x_0 + \delta)$?

7:2.13 If f is increasing and differentiable on an interval does it imply that $f'(x) \geq 0$ on that interval? Does it imply that $f'(x) > 0$ on that interval?

7:2.14 Suppose that two functions f and g have the following properties at a point x_0: $f(x_0) = g(x_0)$ and $f(x) \leq g(x)$ for all x in an open interval containing the point x_0. If both $f'(x_0)$ and $g'(x_0)$ exist show that they must be equal. How does this compare to the trapping principle used in Example 7.5 where it seems much more is assumed about the function f.

7:2.15 Suppose that f is a function defined on the real line with the property that $f(x+y) = f(x)f(y)$ for all x, y. Suppose that f is differentiable at 0 and that $f'(0) = 1$. Show that f must be differentiable everywhere and that $f'(x) = f(x)$.

7.2.2 Differentiability and Continuity

A continuous function need not be differentiable (Example 7.4) but the converse is true. Every differentiable function is continuous.

Theorem 7.6 *Let f be defined in a neighborhood I of x_0. If f is differentiable at x_0, then f is continuous at x_0.*

Proof. It suffices to show that $\lim_{x\to x_0}(f(x)-f(x_0))=0$. For $x\neq x_0$,

$$f(x)-f(x_0)=\left(\frac{f(x)-f(x_0)}{x-x_0}\right)(x-x_0).$$

Now

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}=f'(x_0)$$

and $\lim_{x\to x_0}(x-x_0)=0$, so by the product rule for limits

$$\lim_{x\to x_0}(f(x)-f(x_0))=(f'(x_0))(0)=0.$$

■

We can use this theorem in two ways. If we know that a function has a discontinuity at a point then we know immediately that there is no derivative there. On the other hand if we have been able to determine by some means that a function is differentiable at a point then we know automatically that the function must also be continuous at that point.

Exercises

7:2.15 Construct a function on the interval $[0,1]$ that is continuous and is not differentiable at each point of some infinite set.

7:2.16 Suppose that a function has both a right hand and a left hand derivative at a point. What, if anything, can you conclude about the continuity of that function at that point?

7:2.17 Suppose that a function has an infinite derivative at a point. What, if anything, can you conclude about the continuity of that function at that point?

7:2.18 Show that if a function f has a symmetric derivative $f'_s(x_0)$ (see Exercise 7:2.6) then f must be symmetrically continuous at x_0 in the sense that $\lim_{h\to 0}[f(x_0+h)-f(x_0-h)]=0$. Must f in fact be continuous?

7:2.19 If $f'(x_0)=\infty$ does it follow that f must be continuous at x_0 on one side at least?

7:2.20 Find an example of an everywhere differentiable function f so that f' is not everywhere continuous.

7:2.21 Show that a function f that satisfies an inequality of the form

$$|f(x) - f(y)| \leq M\sqrt{|x-y|}$$

for some constant M and all x, y must be everywhere continuous but need not be everywhere differentiable.

7:2.22 The Dirichlet function (see Section 5.2.6) is discontinuous at each rational number. By Theorem 7.6 it follows that this function has no derivative at any rational number. Does it have a derivative at any irrational number?

7.2.3 The Derivative as a Magnification

We offer now one more interpretation of the derivative, this time as a magnification factor. In elementary calculus one often makes use of the geometric content of the graph of a function f. In particular we can view the derivative in terms of slopes of tangent lines to the graph. But the graph of f is a subset of two-dimensional space, while the range of f is a subset of one-dimensional space and, as such, has some additional geometric content.

Suppose f is differentiable on an interval I, and let J be a closed sub-interval of I. The range of f on J will also be a closed interval, because f is differentiable and hence continuous on J, and continuous functions map closed intervals onto closed intervals (Exercise 5:8.8). The expression

$$\frac{|f(J)|}{|J|}$$

represents the amount that the interval J has been expanded (or contracted) under the mapping f.

For example, if $f(x) = x^2$ and $J = [2,3]$, then

$$\frac{|f(J)|}{|J|} = \frac{|[4,9]|}{|[2,3]|} = \frac{5}{1} = 5.$$

Thus the interval $[2,3]$ has been expanded by f to an interval of 5 times its size. If we look only at small intervals then the derivative offers a clue to the size of the magnification factor.

If J is a sufficiently small interval having x_0 as an endpoint, then the ratio $|f(J)|/|J|$ is approximately $|f'(x_0)|$, the approximation becoming "exact in the limit". Thus $|f'(x_0)|$ can be viewed as a "magnification factor" of small intervals containing the point x_0. In our illustration with the function $f(x) = x^2$, the magnification factor at $x_0 = 2$ is $f'(2) = 4$. Small intervals about x_0 are magnified by a factor of about 4. At the other endpoint $x_0 = 3$, small intervals about x_0 are magnified by a factor of about 6.

In Exercise 7:2.25 we ask the reader to prove a precise statement covering the preceding discussion.

Exercises

7:2.23 What is the ratio
$$\frac{|f(J)|}{|J|}$$
for the function $f(x) = x^2$ if $J = [2, 2.001]$, $J = [2, 2.0001]$, $J = [2, 2.00001]$?

7:2.24 In this section we have interpreted $f'(x_0)$ as a magnification factor. If $f'(x_0) = 0$ does this mean that small intervals containing the point x_0 are magnified by a factor of 0 when mapped by f?

7:2.25 Let f be differentiable on an interval I and let x_0 be an interior point of I. Make precise the statement below and prove it:
$$\lim_{|J| \to x_0} \frac{|f(J)|}{|J|} = |f'(x_0)|.$$

7.3 Computations of Derivatives

Example 7.2 provides a calculation of the derivative of the function $f(x) = x^2$. The calculation involved direct evaluation of the limit of an appropriate difference quotient. For the function $f(x) = x^2$, this evaluation was straightforward. But limits of difference quotients can be quite complicated. The reader is familiar with certain rules that are useful in calculating derivatives of functions that are 'built up' from functions whose derivatives are known.

In this section we review some of the calculus rules that are commonly used to compute derivatives. We need first to prove the algebraic rules—the sum rule, the product rule and the quotient rule. Then we turn to the chain rule. Finally we look at the power rule. Our viewpoint here is not to practice the computation of derivatives but to build up the theory of derivatives, making sure to see how it depends on work on limits that we proved earlier on.

The various rules we shall obtain in this section should be viewed as aids for computations of derivatives. An understanding of these rules is, of course, necessary for various calculations. But they in no way can substitute for an understanding of the derivative. And they might not be useful in calculating certain derivatives. (For example, derivatives of the functions of Exercise 7:2.3 cannot be calculated at $x_0 = 0$ by using these rules.)

Nonetheless, it is true that one often has a function that can be expressed in terms of several functions via the operations we considered in this section, functions whose derivatives we know. In those cases, the techniques of this section might be very useful.

7.3.1 Algebraic Rules

Functions can be combined algebraically by multiplying by constants, by addition and subtraction, by multiplication and by division. To each of these there is a calculus rule for computing the derivative. We recall that the limit of a sum (a difference, a product, a quotient) is the sum (difference, product, quotient) of the limits. Perhaps we might have thought the same kind of rule would apply to derivatives. The derivative of the sum is indeed the sum of the derivatives, but the derivative of the product is not the product of the derivatives. Nor do quotients work in such a simple way. The reasons for the form of the various rules can be found by writing out the definition of the derivative and following through on the computations.

Theorem 7.7 *Let f and g be defined on an interval I and let $x_0 \in I$. If f and g are differentiable at x_0 then $f+g$ and fg are differentiable at x_0. If $g(x_0) \neq 0$, then f/g is differentiable at x_0. Furthermore, the following formulas are valid:*

1. $(cf)'(x) = cf'(x)$ *for any real number c.*
2. $(f+g)'(x_0) = f'(x_0) + g'(x_0)$.
3. $(fg)'(x_0) = f(x_0)g'(x_0) + g(x_0)f'(x_0)$.
4. $\left(\dfrac{f}{g}\right)'(x_0) = \dfrac{g(x_0)f'(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}$ *(if $g(x_0) \neq 0$).*

Proof. Parts 1 and 2 follow easily from the definition of the derivative and appropriate limit theorems.

To verify part 3, let $h = fg$. Then for each $x \in I$ we have

$$h(x) - h(x_0) = f(x)[g(x) - g(x_0)] + g(x_0)[f(x) - f(x_0)], \text{ so}$$

$$\frac{h(x) - h(x_0)}{x - x_0} = f(x)\frac{g(x) - g(x_0)}{x - x_0} + g(x_0)\frac{f(x) - f(x_0)}{x - x_0}. \tag{3}$$

As $x \to x_0$, $f(x) \to f(x_0)$ since f being differentiable is also continuous. By the definition of the derivative we also know that $\dfrac{g(x)-g(x_0)}{x-x_0} \to g'(x_0)$ and $\dfrac{f(x)-f(x_0)}{x-x_0} \to f'(x_0)$ as $x \to x_0$. We now see from equation (3) that

$$\lim_{x\to x_0} \frac{h(x)-h(x_0)}{x-x_0} = f(x_0)g'(x_0) + g(x_0)f'(x_0),$$

verifying part 3.

Finally, to establish part 4 of the theorem, let $h = f/g$. Straightforward algebraic manipulations show that

$$\frac{h(x)-h(x_0)}{x-x_0} =$$

$$\frac{1}{g(x)g(x_0)}\left[g(x_0)\left(\frac{f(x)-f(x_0)}{x-x_0}\right) - f(x)\left(\frac{g(x)-g(x_0)}{x-x_0}\right)\right]. \quad (4)$$

Now let $x \to x_0$. Since f and g are continuous at x_0, $f(x) \to f(x_0)$ and $g(x) \to g(x_0)$. Thus part 4 of the theorem follows from equation (4), the definition of derivative, and basic limit theorems. ■

Example 7.8 To calculate the derivative of $h(x) = (x^3+1)^2$ we have several ways to proceed.

1. Apply the definition of derivative. The reader may wish to set up the difference quotient and see that a calculation of its limit is a formidable task.

2. Write $h(x) = x^6 + 2x^3 + 1$ and apply the formula $\frac{d}{dx}x^n = nx^{n-1}$ (Exercise 7:3.5) and the rule for sums. Thus we get $h'(x) = 6x^5 + 6x^2$.

3. Use the product rule to obtain
$$h'(x) = (x^3+1)\frac{d}{dx}(x^3+1) + (x^3+1)\frac{d}{dx}(x^3+1).$$
Then, again, use the formula $\frac{d}{dx}x^n = nx^{n-1}$ and the rule for sums to continue:
$$h'(x) = (x^3+1)3x^2 + (x^3+1)3x^2 = 6x^5 + 6x^2.$$

◀

Exercises

7:3.1 Give the details needed in the proof of Theorem 7.7 for the sum rule for derivatives, i.e., $(f+g)'(x_0) = f'(x_0) + g'(x_0)$.

7:3.2 The table gives the values of two functions f and g at certain points. Calculate $(f+g)'(1)$, $(fg)'(1)$ and $(f/g)'(1)$. What can you assert about $(f/g)'(3)$? Is there enough information to calculate $f''(3)$?

x	$f(x)$	$f'(x)$	$g(x)$	$g'(x)$
1	3	3	2	2
2	4	4	4	0
3	6	1	1	0
4	-1	0	1	1
5	2	5	3	3

Figure 7.3: Values of f and g at several points.

7:3.3 Obtain the rule
$$\frac{d}{dx}\frac{1}{f(x)} = -\frac{f'(x)}{f(x)^2}$$
from Theorem 7.7 and also directly from the definition of the derivative.

7:3.4 Obtain the rule for
$$\frac{d}{dx}\left(f(x)\right)^2 = 2f(x)f'(x)$$
from Theorem 7.7 and also directly from the definition of the derivative.

7:3.5 Obtain the formula
$$\frac{d}{dx}x^n = nx^{n-1}$$
for $n = 1, 2, 3, \dots$ by induction.

7:3.6 State and prove a theorem that gives a formula for $f'(x_0)$ when $f = f_1 + f_2 + \cdots + f_n$, and each of the functions $f_1, \dots, f_n$ is differentiable at x_0.

7:3.7 State and prove a theorem that gives a formula for $f'(x_0)$ when $f = f_1 f_2 \dots f_n$, and each of the functions $f_1, \dots, f_n$ is differentiable at x_0.

7:3.8 Show that
$$(fg)''(x_0) = f''(x_0)g(x_0) + 2f'(x_0)g'(x_0) + f(x_0)g''(x_0)$$
under appropriate hypotheses.

7:3.9 Extend Exercise 7:3.8 by obtaining a similar formula for $(fg)'''(x_0)$.

7:3.10 Extend Exercises 7:3.8 and 7:3.9 by obtaining a formula for $(fg)^{(n)}(x_0)$ valid for $n = 1, 2, 3, \dots$.

7.3.2 The Chain Rule

There is another, nonalgebraic, interpretation of Example 7.8 that the reader may recall from calculus courses. We can view the function $h(x) = (x^3+1)^2$ as a *composition* of the function $f(x) = x^3+1$ and $g(u) = u^2$. Thus $h(x) = g \circ f(x)$. The reader is familiar with the *chain rule* that is useful in calculating derivatives of composite functions. In this case the calculation would lead to

$$\begin{aligned} h'(x) &= g'(f(x))f'(x) = g'(x^3+1)3x^2 \\ &= 2(x^3+1)3x^2 = 6x^5 + 6x^2. \end{aligned}$$

In elementary calculus the reader might have preferred to make a substitution and written $u = x^3+1, y = u^2$ so

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = 2(x^3+1)(3x^2) = 6x^5 + 6x^2.$$

The chain rule is the familiar calculus formula

$$\frac{d}{dx}g(f(x)) = g'(f(x))f'(x)$$

for the differentiation of the composition of two functions $g \circ f$ under appropriate assumptions. Calculus students often memorize this in the form

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

by using the new variables $y = g(u)$ and $u = f(x)$.

Let us first try to see why the chain rule should work. Then we'll provide a precise statement and proof of the chain rule. Perhaps the easiest way to "see" the chain rule is by interpreting the derivative as a magnification factor.

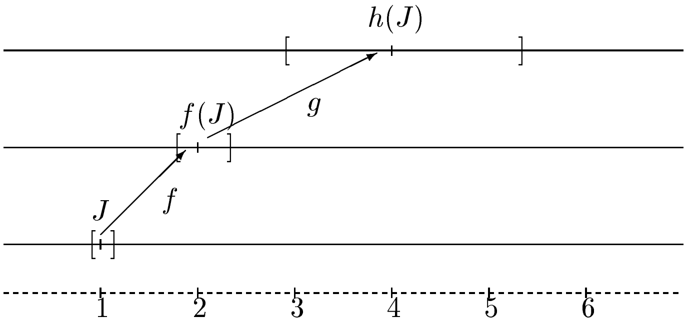
Let f be defined in a neighborhood of x_0 and let g be defined in a neighborhood of $f(x_0)$. If f is differentiable at x_0, then f maps each small interval J containing x_0 onto an interval $f(J)$ containing $f(x_0)$ with $|f(J)|/|J|$ approximately $|f'(x_0)|$. If, also, g is differentiable at $f(x_0)$, then g will map a small interval $f(J)$ containing $f(x_0)$ onto an interval $g(f(J))$ with $|g(f(J))|/|f(J)|$ approximately $|g'(f(x_0))|$. Thus $h = g \circ f$ maps J onto the interval $h(J) = g(f(J))$ and

$$\frac{|h(J)|}{|J|} = \frac{|g(f(J))|}{|f(J)|}\frac{|f(J)|}{|J|}$$

and this is approximately equal to

$$|g'(f(x_0))||f'(x_0)|.$$

In short, the magnification factors $|f'(x_0)|$ and $|g'(f(x_0))|$ multiply to give the magnification factor $|h'(x_0)|$.



$h = g \circ f$. Here $x_0 = 1$, $J = [.9, 1.1]$.

Figure 7.4: f maps J to $f(J)$ and g maps that to $h(J)$.

Example 7.9 Let us relate this discussion to our example $h(x) = (x^3 + 1)^2$. Here $f(x) = x^3 + 1$, $g(x) = x^2$. At $x_0 = 1$ we obtain $f(x_0) = 2$, $f'(x_0) = 3$, $g(f(x_0)) = 4$ and $g'(f(x_0)) = 4$. The function f maps small intervals about $x_0 = 1$ onto ones about three times as long, and in turn, the function g maps those intervals onto ones about four times as long, so the total magnification factor for the function $h = g \circ f$ is about 12 at $x_0 = 1$. Figure 7.4 illustrates. ◀

Proof of the Chain Rule If we wished to formulate a proof of the chain rule based on the preceding discussion we could begin by writing

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \frac{f(x) - f(x_0)}{x - x_0} \tag{5}$$

which compares to our formula

$$\frac{|h(J)|}{|J|} = \frac{|g(f(J))|}{|f(J)|} \frac{|f(J)|}{|J|}.$$

If we let $x \to x_0$ in (5) we would expect to get the desired result

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

And this argument would be valid if f were, for example, strictly increasing. But in order for equation (5) to be valid, we must have $x \neq x_0$ and $f(x) \neq f(x_0)$. When computing the limit of a difference quotient, we can assume $x \neq x_0$, but we can't assume, without additional hypotheses, that if $x \neq x_0$ then $f(x) \neq f(x_0)$. Yet the chain rule applies nonetheless.

The proof is clearer if we separate these two cases. In the simpler case the function does not repeat the value $f(x_0)$ in some neighborhood of x_0. In the harder case the function repeats the value $f(x_0)$ in *every* neighborhood of x_0. Exercise 7:3.11 shows that in that case we must have $f'(x_0) = 0$ and so the chain rule reduces to showing that the composite function $g \circ f$ also has a zero derivative.

Theorem 7.10 (Chain Rule) *Let f be defined on a neighborhood U of x_0 and let g be defined on a neighborhood V for which*

$$f(x_0) \in f(U) \subset V.$$

Suppose f is differentiable at x_0 and g is differentiable at $f(x_0)$. Then the composite function $h = g \circ f$ is differentiable at x_0 and

$$h'(x_0) = (g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Proof. Consider any sequence of distinct points x_n converging to x_0. If we can show that the sequence

$$S_n = \frac{g(f(x_n)) - g(f(x_0))}{x_n - x_0}$$

converges to $g'(f(x_0))f'(x_0)$ for every such sequence then we have obtained our required formula. Note that if $f(x_n) \neq f(x_0)$ then we can write $y_n = f(x_n)$, $y_0 = f(x_0)$ and display S_n as

$$S_n = \frac{g(y_n) - g(y_0)}{y_n - y_0} \times \frac{f(x_n) - f(x_0)}{x_n - x_0}. \tag{6}$$

Seen in this form it becomes obvious that $S_n \to g'(y_0)f'(x_0) = g'(f(x_0))f'(x_0)$ except for the problem that we cannot (as we remarked before beginning our proof) assume that in all cases $f(x_n) \neq f(x_0)$.

Thus we consider two cases. In case (i) we assume that for any sequence of distinct points x_n converging to x_0 there cannot be infinitely many terms with $f(x_n) = f(x_0)$. In that case the chain rule formula is evidently valid.

In case (ii) we assume that there does exist a sequence of distinct points x_n converging to x_0 with infinitely many terms satisfying $f(x_n) = f(x_0)$. In that case (Exercise 7:3.11) we must have $f'(x_0) = 0$ and so, to establish the chain rule, we need to prove that $h'(x_0) = 0$. But in this case for any sequence x_n converging to x_0 either $S_n = 0$ or else S_n can be written in the form of equation (6). It is then clear than $S_n \to 0$ and the proof is complete. ■

Exercises

7:3.11 Show that if $f(x) = f(x_0)$ for some $x \neq x_0$ in every neighborhood of x_0 then either $f'(x_0)$ does not exist or else $f'(x_0) = 0$.

7:3.12 Give an explicit example of functions f and g such that the "proof" of the chain rule based on equation (5) fails.

7:3.13 The heuristic discussion preceding Theorem 7.10 dealt with $|h'(x_0)|$, not with $h'(x_0)$. Explain how the signs of $f'(x_0)$ and $g'(f(x_0))$ affect the discussion. In particular, how can we modify the discussion to get the correct sign for $h'(x_0)$?

7:3.14 Most calculus texts use a proof of Theorem 7.10 based on the following ideas. Define a function G in the neighborhood V of $f(x_0)$ by

$$G(v) = \begin{cases} [g(v) - g(f(x_0))]/[v - f(x_0)], & \text{if } v \neq f(x_0) \\ g'(f(x_0)), & \text{if } v = f(x_0). \end{cases} \tag{7}$$

(a) Show that G is continuous at $f(x_0)$.

(b) Show that $G(v)(v - f(x_0)) = g(v) - g(f(x_0))$ for every $v \in V$ (regardless of whether or not $f(x_0) = v$).

(c) Prove that $\lim_{x \to x_0} \frac{h(x) - h(x_0)}{x - x_0} = g'(f(x_0))f'(x_0)$.

7:3.15 State and prove a theorem that gives a formula for $f'(x_0)$ when $f = f_n \circ f_{n-1} \circ \cdots \circ f_2 \circ f_1$. (Be sure to state all the hypotheses that you need.)

7:3.16 The table gives the values of two functions f and g at certain points. Calculate $(f \circ g)'(1)$ and $(g \circ f)'(1)$. Is there enough information to calculate $(f \circ g)'(3)$ and/or $(g \circ f)'(3)$? How about $\frac{d}{dx}(f^2)(1)$ and $(f \circ f)'(1)$?

x	$f(x)$	$f'(x)$	$g(x)$	$g'(x)$
1	3	3	2	2
2	4	4	4	0
3	6	1	1	0
4	-1	0	1	1
5	2	5	3	3

Figure 7.5: Values of f and g at several points.

7.3.3 Inverse Functions

Suppose that a function $f : I \to J$ has an inverse. This simply means that there is a function g (called the inverse function of f)

that reverses the mapping, i.e. so that if $f(a) = b$ then $g(b) = a$. We can assume that I and J are intervals. Thus f maps the interval I onto the interval J and the inverse function g then maps J back to I. Not all functions have an inverse but we are supposing that this one does.

Suppose too that f is differentiable at a point $x_0 \in I$. Then we would expect from geometric considerations that that the inverse function g should be differentiable at the image point $z_0 = f(x_0) \in J$.

This is entirely elementary. The connection between a function f and its inverse g is given by

$$f(g(x)) = x \text{ for all } x \in J$$

or

$$g(f(x)) = x \text{ for all } x \in I.$$

Using the chain rule on the second of these immediately gives

$$g'(f(x))f'(x) = 1$$

and hence we have the connection

$$g'(f(x)) = \frac{1}{f'(x)}$$

which a geometrical argument could also have found.

Example 7.11 Suppose that the exponential function e^x has been developed and that we have proved that it is differentiable for all values of x and we have the usual formula $\frac{d}{dx}e^x = e^x$. Then, provided we can be sure there is an inverse, a formula for the derivative of that inverse can be found. Let $L(x)$ be the inverse function of $f(x) = e^x$. Then, since we know that $f'(x) = f(x)$

$$L'(f(x)) = \frac{1}{f'(x)} = \frac{1}{f(x)}$$

or, replacing $f(x)$ by another letter, say z, we have

$$L'(z) = \frac{1}{z}.$$

This must be valid for every value z in the domain of L (i.e., every value in the range of f). The reader should recognize the derivative of the function $\ln z$ here. Even so we would still need to justify the existence of the inverse function before we could properly claim to have proved this formula. ◀

We would like a better way to handle inverse functions than presented here. Our observations here allow us to compute the derivative of an inverse but do not assure us that an inverse will exist.

For a theorem that allows us merely to look at the derivative and determine that an inverse exists see Theorem 7.30.

Exercises

7:3.17 Find a formula for the derivative of the function $\sin^{-1} x$ assuming that the usual formula for $\frac{d}{dx}\sin x = \cos x$ has been found.

7:3.18 Find a formula for the derivative of the function $\tan^{-1} x$ assuming that the usual formula for $\frac{d}{dx}\tan x = \sec^2 x$ has been found.

7:3.19 Give a geometric interpretation of the relationship between the slope of the tangent at a point (x_0, y_0) on the graph of $y = f(x)$ and the slope of the tangent at the point (y_0, x_0) on the graph of $y = g(x)$ where g is the inverse of f.

7:3.20 What facts about the function $f(x) = e^x$ would need to be established in order to claim that there is indeed an inverse function? What is the domain and range of that inverse function?

7.3.4 The Power Rule

The power rule is the formula

$$\frac{d}{dx}x^p = px^{p-1}$$

which is the basis for most calculus problems. We have already shown (in Exercise 7:3.5) that

$$\frac{d}{dx}x^n = nx^{n-1}$$

for $n = 1, 2, 3, \ldots$ and for every value of x.

This is easy enough to extend to negative integers. Just interpret for $n = 1, 2, 3, \ldots$ and for every value of $x \neq 0$,

$$\frac{d}{dx}x^{-n} = \frac{d}{dx}\frac{1}{x^n}$$

and, using the quotient rule, we find that again the power rule formula is valid for $p = -1, -2, -3, \ldots$ and any value of x other than 0.

The formula also works for $p = 0$ since we interpret x^0 as the constant 1 (although for $x = 0$ we prefer not to make any claims). Is the formula indeed valid for every value of p, not just for integer values?

Example 7.12 We can verify the power rule formula for $p = 1/2$, i.e., we prove that

$$\frac{d}{dx}\sqrt{x} = \frac{d}{dx}x^{1/2} = \frac{1}{2}x^{1/2-1} = \frac{1}{2\sqrt{x}}.$$

First we must insist that $x > 0$ otherwise $\sqrt{x}$ and the fraction in our formula would not be defined. Now interpret $\sqrt{x}$ as the inverse of the square function $f(x) = x^2$. Specifically let $f(x) = x^2$ for $x > 0$ and $g(x) = \sqrt{x}$ for $x > 0$ and note that $f(g(x)) = g(f(x)) = x$. Thus

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

and so, since $f'(x) = 2x$ and $f'(g(x))g'(x) = 1$ we obtain $2\sqrt{x}g'(x) = 1$ and finally that

$$g'(x) = \frac{1}{2\sqrt{x}}$$

as required if the power rule formula is valid. ◀

Is the power rule

$$\frac{d}{dx}x^p = px^{p-1}$$

valid for all rational values of p? We can handle the case $p = m/n$ for integer m and n by essentially the same methods. We state this as a theorem whose proof is left as an exercise. For irrational p there is also a discussion in the exercises.

Theorem 7.13 *Let $f(x) = x^{m/n}$ for $x > 0$ and integers m, n. Then*

$$f'(x) = \frac{m}{n}x^{m/n-1}.$$

Example 7.14 Every polynomial is differentiable on $\mathbb{R}$ and its derivative can be calculated via term-by-term differentiation, i.e.,

$$\frac{d}{dx}(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) = a_1 + 2a_2x + \cdots + na_nx^{n-1}.$$

This follows from the power rule formula and the rule for sums. Note that the derivative of a polynomial is again a polynomial. ◀

Example 7.15 A rational function is a function $R(x)$ that can be expressed as the quotient of two polynomials,

$$R(x) = \frac{p(x)}{q(x)}.$$

This would be defined at every point at which the denominator $q(x)$ is not equal to zero. Every rational function is differentiable except at those points at which the denominator vanishes. This follows from the previous example showing how to differentiate a polynomial and from the quotient rule. Thus

$$\frac{d}{dx}\left(\frac{p(x)}{q(x)}\right) = \frac{p'(x)q(x) - p(x)q'(x)}{q^2(x)}.$$

Notice that the derivative is another rational function with the same domain since both numerator and denominator are again polynomials. ◀

Exercises

7:3.21 Prove Theorem 7.13.

7:3.22 Show that the power formula is available for all values of p once the formula $\frac{d}{dx}e^x = e^x$ is known.

7:3.23 Let

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n.$$

Compute the sequence of values $p(0)$, $p'(0)$, $p''(0)$, $p'''(0)$, ...

7:3.24 Determine the coefficients of the polynomial

$$p(x) = (1+x)^n = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

by using the formulas that you have obtained in Exercise 7:3.23.

7.4 Continuity of the Derivative?

We have already observed (Theorem 7.6) that if a function f is differentiable on an interval I, then f is also continuous on I. This statement should not be confused with the (incorrect) statement that the derivative, f', is continuous.

Example 7.16 Consider the function f defined on $\mathbb{R}$ by

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

Since $|\sin x^{-1}| \leq 1$ for all $x \neq 0$, $|f(x)| \leq x^2$ for all $x \in \mathbb{R}$. We can now conclude (e.g., from Example 7.5) that $f'(0) = 0$. For $x \neq 0$, we can calculate, as in elementary calculus, that

$$f'(x) = -\cos x^{-1} + 2x \sin x^{-1}.$$

This function f' is continuous at every point $x_0 \neq 0$. At $x_0 = 0$ it is discontinuous. To see this we need only consider an appropriate sequence $x_n \to 0$ and see what happens to $f'(x_n)$. For example, try the sequence

$$x_n = \frac{1}{\pi n}.$$

Since

$$\cos\left(\frac{1}{x_n}\right) = \cos(\pi n)$$

and these numbers are alternately $+1$ and -1 it is clear that $f'(x_n)$ cannot converge. Consequently f' is discontinuous at 0. ◀

Observe that the function f provides an example of a function that is differentiable on all of $\mathbb{R}$, yet f' is discontinuous at a point. It is possible to modify this function to obtain a differentiable function g whose derivative g' is discontinuous at infinitely many points, and even at all the points of the Cantor set (See Exercise 7:4.2).

One might wonder, then, if anything positive could be said about the properties of a derivative f'. It is possible for the derivative of a differentiable function to be discontinuous on a dense set: an example is given later on in Section 9.7[1] We will also show, in Section 7.9, that the function f', while perhaps discontinuous, nonethless shares one significant property of continuous functions: it has the intermediate value property (Darboux property).

Exercises

7:4.1 Give a simple example of a function f differentiable in a deleted neighborhood of x_0 such that $\lim_{x\to x_0} f'(x)$ does not exist.

7:4.2 Let P be a Cantor subset of $[0,1]$ (i.e., P is a nonempty, nowhere dense perfect subset of [0,1]) and let $\{(a_n, b_n)\}$ be the sequence of intervals complementary to P in $(0,1)$. (See Section 6.5.1.) ✂

(a) On each interval $[a_n, b_n]$ construct a differentiable function such that

$$f_n(a_n) = f_n(b_n) = (f_n')_+(a_n) = (f_n')_-(b_n) = 0,$$
$$\limsup_{x\to a_n^+} f'(x) = \limsup_{x\to b_n^-} f_n'(x) = 1,$$
$$\liminf_{x\to a_n^+} f'(x) = \liminf_{x\to b_n^-} f_n'(x) = -1,$$

and $|f_n(x)| \leq (x-a_n)^2(x-b_n)^2$ and $|f_n'(x)|$ is bounded by 1 in each interval $[a_n, b_n]$.

[1] It is not possible for a derivative to be discontinuous at every point. See Bruckner, Bruckner and Thomson, *Real Analysis*, Prentice-Hall (1997), p. 452 (Section 10.4).

(b) Let g be defined on [0,1] by

$$g(x) = \begin{cases} f_n(x), & \text{if } x \in (a_n, b_n), n = 1, 2, \ldots \\ 0, & \text{if } x \in P. \end{cases}$$

Sketch a picture of the graph of g.

(c) Prove g is differentiable on [0,1].

(d) Prove $g'(x) = 0$ for each $x \in P$.

(e) Prove g' is discontinuous at every point of P.

7.5 Local Extrema

We have seen in Section 5.7 that a continuous function defined on a closed interval $[a, b]$ achieves an absolute maximum value and an absolute minimum value on the interval. So there must be points where the maximum and minimum are attained. But how does one go about finding such points?

A familiar process studied in elementary calculus is sometimes useful for locating these extrema when the function is differentiable on (a, b): one looks for critical points, i.e. points where the derivative is zero. We begin with the theorem that forms the basis for this process.

Theorem 7.17 *Let f be defined on an interval I. If f has a local extremum at a point x_0 in the interior of I and f is differentiable at x_0, then $f'(x_0) = 0$.*

Proof. Suppose f has a local maximum at x_0 in the interior of I, the proof for a local minimum being similar. Then there exists $\delta > 0$ such that $[x_0 - \delta, x_0 + \delta] \subset I$ and $f(x) \leq f(x_0)$ for all $x \in [x_0 - \delta, x_0 + \delta]$. Thus

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0 \quad \text{for} \quad x \in (x_0, x_0 + \delta) \quad \text{and} \tag{8}$$

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0 \quad \text{for} \quad x \in (x_0 - \delta, x_0). \tag{9}$$

If $f'(x_0)$ exists, then

$$f'(x_0) = \lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0}. \tag{10}$$

By (8), the first of these limits is at most zero; by (9), the second is at least zero. But by (10), these limits are equal and must therefore be equal to zero. ■

It follows from Theorem 7.17 that a function f that is continuous on $[a, b]$ must achieve its maximum at one (or more) of these types of points:

1. points $x_0 \in (a, b)$ at which $f'(x_0) = 0$.

2. points $x_0 \in (a, b)$ at which f is not differentiable.

3. the points a or b.

We leave it to the reader to provide simple examples of each of these possibilities.

The usual process for locating extrema in elementary calculus thus involves locating points at which f has a zero derivative and comparing the values of f at those points and the points of nondifferentiability (if any) and at the endpoints a and b. In the setting of elementary calculus the situation is usually relatively simple: the function is differentiable, the set on which $f'(x) = 0$ is finite (or contains an interval) and the equation $f'(x) = 0$ is easily solved. Much more complicated situations can occur, of course. The exercises below provide some examples and theorems that indicate just how complicated the set of extrema can be.

Exercises

7:5.1 Give an example of a differentiable function on $\mathbb{R}$ for which $f'(0) = 0$ but 0 is not a local maximum or minimum of f.

7:5.2 Let

$$f(x) = \begin{cases} x^4(2 + \sin x^{-1}), & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

(a) Prove that f is differentiable on $\mathbb{R}$.

(b) Prove that f has an absolute minimum at $x = 0$.

(c) Prove that f' takes on both positive and negative values in every neighborhood of 0.

7:5.3 Let K be the Cantor set and let $\{(a_k, b_k)\}$ be the sequence of intervals complementary to K in [0,1]. For each k, let $c_k = (a_k + b_k)/2$. Define f on [0,1] to be zero on K, $1/k$ at c_k, linear and continuous on each of the intervals. (See Figure 7.6.)

(a) Write equations that represent f on the intervals $[a_k, c_k]$ and $[c_k, b_k]$.

(b) Show that f is continuous on [0,1].

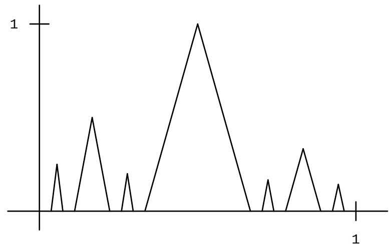


Figure 7.6: Construction of function in Exercise 7:5.3.

(c) Verify that f has minimum zero, achieved at each $x \in K$.

(d) Verify that f has a local maximum at each of the points c_k.

(e) Modify f to a differentiable function with the same set of extrema.

7:5.4 Find all local extrema of the Dirichlet function (see Section 5.2.6) defined on [0,1] by

$$f(x) = \begin{cases} 0, & \text{if } x \text{ is irrational or } x = 0 \\ 1/q, & \text{if } x = p/q,\ \ p, q \in \mathbb{N},\ \ p/q \text{ in lowest terms.} \end{cases}$$

7:5.5 Show that the functions in Exercises 7:5.3 and 7:5.4 have infinitely many maxima, all of them strict. Show that the sets of points at which these functions have a strict maximum is countable.

✂ **7:5.6** Prove that if $f:\mathbb{R}\to\mathbb{R}$, then $\{x : f \text{ achieves a strict maximum at } x\}$ is countable.

✂ **7:5.7** Let $f : \mathbb{R} \to \mathbb{R}$ have the following property: For each $x \in \mathbb{R}$, f achieves a local maximum (not necessarily strict) at x.

(a) Give an example of such an f whose range is infinite.

(b) Prove that for every such f, the range is countable.

7:5.8 There are continuous functions $f:\mathbb{R}\to\mathbb{R}$, even differentiable functions, that are nowhere monotonic. This means that there is *no interval* on which the function is increasing, decreasing or constant. For such functions, the set of maxima as well as the set of minima, is dense in $\mathbb{R}$. Construction of such functions is beyond the scope of this book. (See, e.g., Bruckner, Bruckner and Thomson, *Real Analysis*, Prentice Hall, 1997.)

Show that such a function f maps its set of extrema onto a dense subset of the range of f.

7.6 Mean Value Theorem

There is a close connection between the values of a function and the values of its derivative. In one direction this is trivial since the derivative is defined in terms of the values of the function. The other direction is more subtle. How does information about the derivative provide us with information about the function? One of the keys to providing that information is the mean value theorem.

Suppose f is continuous on an interval $[a, b]$ and is differentiable on (a, b). Consider a point x in (a, b). For $y \in (a, b), \ \ y \neq x$, the difference quotient

$$\frac{f(y) - f(x)}{y - x}$$

represents the slope of the chord determined by the points $(x, f(x))$ and $(y, f(y))$. This slope may or may not be a good approximation to $f'(x)$. If y is sufficiently near x, the approximation will be good; otherwise it may not be. The mean value theorem (MVT) asserts that somewhere in the interval determined by x and y there will be a point at which the derivative is exactly the slope of the given chord. It is the existence of such a point that provides a connection between the values of the function (in this case the value $(f(y) - f(x))/(y - x)$ and the value of the derivative (in this case the value at *some* point between x and y).

7.6.1 Rolle's Theorem

We begin with a preliminary theorem that provides a special case of the mean value theorem.

Theorem 7.18 (Rolle's Theorem) *Let f be continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$ then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. If f is constant on $[a, b]$, then $f'(x) = 0$ for all $x \in (a, b)$, so c can be taken to be any point of (a, b).

Suppose then that f is not constant. Because f is continuous on the compact interval $[a, b]$, f achieves a maximum value M and a minimum value m on $[a, b]$ (Theorem 5.45). Because f is not constant, one of the values M or m is different from $f(a)$ and $f(b)$, say $M > f(a)$. Choose $c \in (a, b)$ such that $f(c) = M$. By Theorem 7.17, $f'(c) = 0$. Since $M > f(a) = f(b)$, $c \neq a$ and $c \neq b$, so $c \in (a, b)$.
■

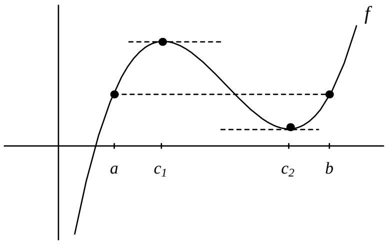


Figure 7.7: Rolle's Theorem (note that $f(a) = f(b)$).

Observe that Rolle's theorem asserts that under our hypotheses, there is a point at which the tangent to the graph of the function is horizontal, and therefore has the same slope as the chord determined by the points $(a, f(a))$ and $(b, f(b))$. (See Figure 7.7.)

There may, of course, be many such points; Rolle's theorem just guarantees the existence of at least one such point. Observe also that we did not require that f be differentiable at the endpoints a and b. The theorem applies to such functions as $f(x) = x \sin x^{-1}$, $f(0) = 0$, on the interval $[0, 1/\pi]$. This function is not differentiable at zero, but it does have an infinite number of points between 0 and $1/\pi$ where the derivative is zero.

Exercises

7:6.1 Apply Rolle's theorem to the function $f(x) = \sqrt{1 - x^2}$ on $[-1, 1]$. Observe that f fails to be differentiable at the endpoints of the interval.

7:6.2 Use Rolle's theorem to explain why the cubic equation

$$x^3 + \alpha x^2 + \beta = 0$$

cannot have more than one solution whenever $\alpha > 0$.

7:6.3 If the nth degree equation

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0$$

has n distinct roots then how many distinct roots does the $(n-1)$st degree equation $p'(x) = 0$ have?

7:6.4 Suppose that $f : \mathbb{R} \to \mathbb{R}$ and both f' and f'' exist everywhere. Show that if f has three zeros then there must be some point ξ so that $f''(\xi) = 0$.

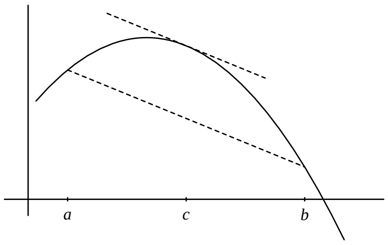


Figure 7.8: Mean Value Theorem ($f'(c)$ is slope of secant).

7:6.5 Let f be continuous on an interval $[a, b]$ and differentiable on (a, b) with a derivative that never is zero. Show that f maps $[a, b]$ one-one onto some other interval.

7:6.6 Let f be continuous on an interval $[a, b]$ and twice differentiable on (a, b) with a second derivative that never is zero. Show that f maps $[a, b]$ two-one onto some other interval, i.e., there are at most two points in $[a, b]$ mapping into any one value in the range of f.

7.6.2 Mean Value Theorem

If we drop the requirement in Rolle's Theorem that $f(a) = f(b)$, we now obtain the result that there is a $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Geometriacally, this states that there exists a point $c \in (a, b)$ for which the tangent to the graph of the function at $(c, f(c))$ is parallel to the chord determined by the points $(a, f(a))$ and $(b, f(b))$. (See Figure 7.8.)

This is the mean value theorem, also known as the law of the mean or the first mean value theorem (because there are other mean value theorems).

Theorem 7.19 (Mean Value Theorem) *Let f be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Proof. We prove this theorem by subtracting from f a function whose graph is the straight line determined by the chord in question, and then applying Rolle's theorem. Let

$$L(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a).$$

We see that $L(a) = f(a)$ and $L(b) = f(b)$. Now let

$$g(x) = f(x) - L(x). \tag{11}$$

Then g is continuous on $[a, b]$, differentiable of (a, b), and satisfies $g(a) = g(b) = 0$.

By Rolle's theorem, there exists $c \in (a, b)$ such that $g'(c) = 0$. Differentiating (11) we see that $f'(c) = L'(c)$. But $L'(c) = \frac{f(b)-f(a)}{b-a}$, so $f'(c) = \frac{f(b)-f(a)}{b-a}$, as was to be proved. ■

Rolle's theorem and the mean value theorem were easy to prove. The proofs relied on the geometric content of the theorems. We suggest that the reader take the time to understand the geometric interpretation of these theorems.

Exercises

7:6.8 A function f is said to satisfy a *Lipschitz condition* on an interval $[a, b]$ if $|f(x) - f(y)| \le M|x - y|$ for all x, y in the interval. Show that if f is assumed to be continuous on $[a, b]$ and differentiable on (a, b) then this condition is equivalent to the derivative f' being bounded on (a, b).

7:6.9 Suppose f satisfies the hypotheses of the mean value theorem on $[a, b]$. Let S be the set of all slopes of chords determined by pairs of points on the graph of f and let D be the set $\{f'(x) : x \in (a, b)\}$.

(a) Prove $S \subset D$.

(b) Give an example to show that D can contain numbers not in S.

7:6.10 Interpreting the slope of a chord as an average rate of change and the derivative as an instantaneous rate of change, what does the mean value theorem say? If a car travels 100 miles in 2 hours, and the position $s(t)$ of the car at time t satisfies the hypotheses of the mean value theorem, can we be sure that there is at least one instant at which the velocity is 50 mph.?

7:6.11 Give an example to show that the conclusion of the mean value theorem can fail if we drop the requirement that f be differentiable at every point in (a, b). Give an example to show that the conclusion can fail if we drop the requirement of continuity at the end points of the interval.

7:6.12 Suppose that f is differentiable on $[0, \infty)$ and that

$$\lim_{x\to\infty} f'(x) = C.$$

Determine

$$\lim_{x\to\infty} [f(x+a) - f(x)].$$

7:6.13 Suppose that f is continuous on $[a, b]$ and differentiable on (a, b). If

$$\lim_{x\to a+} f'(x) = C$$

what can you conclude about the right hand derivative of f at a?

7:6.14 Suppose that f is continuous and that

$$\lim_{x\to x_0} f'(x)$$

exists. What can you conclude about the differentiability of f? What can you conclude about the continuity of f'?

7:6.15 Let $f : [0, \infty) \to \mathbb{R}$ so that f' is decreasing and positive. Show that $\sum_{i=1}^{\infty} f'(i)$ is convergent if and only if f is bounded.

7:6.16 Prove a second order version of the mean value theorem.

> *Let f be continuous on $[a, b]$ and twice differentiable on (a, b). Then there exists $c \in (a, b)$ such that*
>
> $$f(b) = f(a) + (b-a)f'(a) + (b-a)^2 \frac{f''(c)}{2!}.$$

7:6.17 Determine all functions $f : \mathbb{R} \to \mathbb{R}$ that have the property that

$$f'\left(\frac{x+y}{2}\right) = \frac{f(x) - f(y)}{x - y}$$

for every $x \neq y$.

7:6.18 A function is said to be *smooth* at a point x if

$$\lim_{h\to 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} = 0.$$

Show that a smooth function need not be continuous. Show that if f'' is continuous at x then f is smooth at x.

7.6.3 Cauchy's Mean Value Theorem

We can generalize the mean value theorem to curves given parametrically. Suppose f and g are continuous on $[a, b]$ and differentiable on (a, b). Consider the curve given parametrically by

$$x = g(t)\,, \qquad y = f(t) \qquad (t \in [a, b]).$$

As t varies over the interval $[a, b]$, the point (x, y) traces out a curve C joining the points $(g(a), f(a))$ and $(g(b), f(b))$. If $g(a) \neq g(b)$, the slope of the chord determined by these points is $\frac{f(b)-f(a)}{g(b)-g(a)}$. Cauchy's form of the mean value theorem asserts that there is a point (x, y) on C at which the tangent is parallel to the chord in question. We state and prove this theorem.

Theorem 7.20 (Cauchy MVT) *Let f and g be continuous on $[a, b]$ and differentiable on (a, b). Then there exists c such that*

$$[f(b) - f(a)]g'(c) = [g(b) - g(a)]f'(c). \tag{12}$$

Proof. Let

$$\phi(x) = [f(b) - f(a)]g(x) - [g(b) - g(a)]f(x).$$

Then ϕ is continuous on $[a, b]$ and differentiable on (a, b). Furthermore,

$$\phi(a) = f(b)g(a) - f(a)g(b) = \phi(b).$$

By Rolle's theorem, there exists $c \in (a, b)$ for which $\phi'(c) = 0$. It is clear that this point c satisfies (12). ■

Exercises

7:6.17 Use Cauchy's mean value theorem to prove any simple version of L'Hôpital's rule that you can remember from the calculus.

7:6.18 Show that the conclusion of Cauchy's mean value can be put into determinant form as

$$\begin{vmatrix} f(a) & g(a) & 1 \\ f(b) & g(b) & 1 \\ f'(c) & g'(c) & 0 \end{vmatrix} = 0.$$

7:6.19 Formulate and prove a generalized version of Cauchy's mean value whose conclusion is the existence of a point c such that

$$\begin{vmatrix} f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \\ f'(c) & g'(c) & h'(c) \end{vmatrix} = 0.$$

7.7 Monotonicity

In elementary calculus one learns that if $f' \geq 0$ on an interval I, then f is nondecreasing on I. One uses this and related results for a variety of purposes: sketching graphs of functions, locating extrema, etc. In this section we take a closer look at what's involved. We recall some definitions.

Definition 7.21 Let f be real valued on an interval I.

1. If $f(x_1) \leq f(x_2)$ whenever x_1 and x_2 are points in I with $x_1 < x_2$, we say f is *nondecreasing* on I.

2. If the strict inequality $f(x_1) < f(x_2)$ holds, we say f is *increasing*

A similar definition was given for nonincreasing and decreasing functions.

Note. Some authors prefer the terms "increasing" and "strictly increasing" for what we would call nondecreasing and increasing. This has the unfortunate result that constant fucntions are then considered to be both increasing and decreasing. According to our definition we must say that they are both nondecreasing and nonincreasing which sounds more plausible—if something stays constant it is neither going up nor going down). The disadvantage of our usage is the ackwardness the student may at first feel in using the terms (which disappears with practice). It is always safe to say "strictly increasing" for increasing even though it is redundant according to the definition.

The theorems involving monotonicity of functions that one encounters in elementary calculus usually are stated for differentiable functions. But a monotonic function need not be differentiable—or even continuous. For example, if

$$f(x) = \begin{cases} x, & \text{for } x < 0 \\ x+1, & \text{for } x \geq 0, \end{cases}$$

then f is (strictly) increasing on $\mathbb{R}$, but is not continuous at $x = 0$. (For more on discontinuities of monotonic functions see Section 5.9.2.)

Let us now address the role of the derivative in the study of monotonicity. We prove a familiar theorem that is the basis for many applications in the calculus. Note that the proof is an easy consequence of the mean value theorem. Indeed this is the most important way in which the mean value theorem is used: it allows information about the derivative to yield statements about the function.

Theorem 7.22 *Let f be differentiable on an interval I.*

(1) *If $f'(x) \geq 0$ for all $x \in I$, then f is nondecreasing on I.*

(2) *If $f'(x) > 0$ for all $x \in I$, then f is increasing on I.*

(3) *If $f'(x) \leq 0$ for all $x \in I$, then f is nonincreasing on I.*

(4) *If $f'(x) < 0$ for all $x \in I$, then f is decreasing on I.*

(5) *If $f'(x) = 0$ for all $x \in I$, then f is constant on I.*

Proof. To prove (1), let $x_1, x_2 \in I$ with $x_1 < x_2$. By the mean value theorem (7.19) there exists $c \in (x_1, x_2)$ such that $f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$. If $f'(c) \geq 0$, then $f(x_2) \geq f(x_1)$. Thus, if $f'(x) \geq 0$ for all $x \in I$, f is nondecreasing on I.

Parts (2), (3) and (4) have similar arguments, and (5) follows immediately from parts (1) and (3). ■

Exercises

7:7.1 Establish the inequality

$$e^x < \frac{1}{1-x}$$

for all $x < 1$.

7:7.2 Suppose that f and g are differentiable functions such that $f' = g$ and $g' = -f$. Show that there exists a number C with the property that

$$[f(x)]^2 + [g(x)]^2 = C$$

for all x.

7:7.3 Suppose f is continuous on (a, c) and $a < b < c$. Suppose also that f is differentiable on (a, b) and on (b, c). Prove that if $f' < 0$ on (a, b) and $f' > 0$ on (b, c), then f has a minimum at b. We need not assume differentiability at b. (Example: $f(x) = |x|$, $b = 0$.)

7:7.4 The hypotheses of Theorem 7.22 require that f be differentiable on all of the interval I. One might think that a positive derivative at a single point also implies that the function is increasing, at least in a neighborhood of that point. This is not true. Consider the function

$$f(x) = \begin{cases} x/2 + x^2 \sin x^{-1}, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

(a) Show that the function $g(x) = x^2 \sin x^{-1}$ ($g(0) = 0$) is everywhere differentiable and that $g'(0) = 0$.

(b) Show that g' is discontinuous at $x = 0$ and that g' takes on values close to ± 1 arbitrarily near 0.

(c) Show that f' takes on both positive and negative values in every neighborhood of zero.

(d) Show that $f'(0) = \frac{1}{2} > 0$, but that f is not increasing in any neighborhood of zero.

(e) Prove that if a function F is differentiable on a neighborhood of x_0 with $F'(x_0) > 0$ and F' is *continuous* at x_0, then F is increasing on some neighborhood of x_0.

(f) Why does the example $f(x)$ given here not contradict part (e)?

7:7.5 Let f be differentiable on $[0,\infty)$ and suppose that $f(0)=0$ and that the derivative f' is an increasing function on $[0,\infty)$. Show that

$$\frac{f(x)}{x} < \frac{f(y)}{y}$$

for all $0 < x < y$.

7:7.6 Suppose that $f,\ g:\mathbb{R}\to\mathbb{R}$ and both have continuous derivatives and the determinant

$$\phi(x) = \begin{vmatrix} f(x) & g(x)) \\ f'(x) & g'(x) \end{vmatrix}$$

is never zero. Show that between any two zeros of f there must be a zero of g.

7.8 Dini Derivates

We observed in Example 7.4 that the function $f(x)=|x|$ does not have a derivative at the point $x=0$, but does have the unilateral derivatives $f'_+(0)=1$ and $f'_-(0)=-1$. It is not difficult to construct continuous functions that don't have even unilateral derivatives at a point.

Example 7.23 Consider the function

$$f(x) = \begin{cases} |x|\left|\cos x^{-1}\right|, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

(See Figure 7.23.) Since $|\cos x^{-1}| \leq 1$ for all $x \neq 0$,

$$\lim_{x\to 0} f(x) = 0 = f(0)$$

so f is continuous at $x=0$. It is clear that f is continuous at all other points in $\mathbb{R}$, so f is a continuous function.

The oscillatory behavior of f is such that the sets

$$\left\{x : \left|\cos x^{-1}\right| = 1\right\} \quad \text{and} \quad \left\{x : \left|\cos x^{-1}\right| = 0\right\}$$

both have zero as a bilateral limit point. Thus each of the sets

$$\{x : f(x) = |x|\} \quad \text{and} \quad \{x : f(x) = 0\}$$

has zero as bilateral limit point. Inspection of the difference quotient reveals that

$$\limsup_{x\to 0+} \frac{f(x)-f(0)}{x-0} = 1, \text{ while } \liminf_{x\to 0+} \frac{f(x)-f(0)}{x-0} = 0,$$

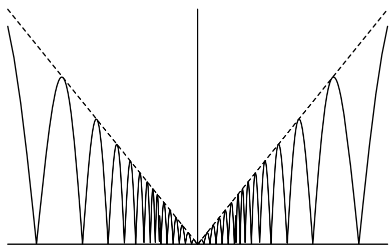


Figure 7.9: Graph of $f(x) = |x \cos x^{-1}|$.

so f'_+ does not exist at $x = 0$. Similarly, $f'_-(0)$ does not exist. The limits that are required to exist for f to have a derivative, or a unilateral derivative, don't exist at $x = 0$. ◀

Example 7.24 A function defined on an interval I may fail to have a derivative, even a unilateral derivative, at every point. Let

$$g(x) = \begin{cases} 0, & \text{if } x \text{ is rational,} \\ 1, & \text{if } x \text{ is irrational.} \end{cases}$$

Since g is everywhere discontinuous on both sides, g has no derivative and no unilateral derivative at any point. ◀

There are, also, *continuous* functions that fail to have a unilateral derivative, finite or infinite, at even a single point. Such functions are difficult to construct, the first construction having been given by Besicovitch in 1925.

Now the derivative, when it exists, plays an important role in analysis, and it is useful to have a substitute when it doesn't exist. Many good substitutes have been developed for certain situations. Perhaps the simplest such substitutes are the Dini derivates. These exist at every point for every function defined on an open interval.

Definition 7.25 Let f be real-valued in a neighborhood of x_0. We define the four *Dini derivates* of f at x_0 by

1. [Upper right Dini derivate]
$$D^+f(x_0) = \limsup_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0}$$

2. [Lower right Dini derivate]
$$D_+f(x_0) = \liminf_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0}$$

3. [Upper left Dini derivate]

$$D^{-}f(x_0) = \limsup_{x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0}$$

4. [Lower left Dini derivate]

$$D_{-}f(x_0) = \liminf_{x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0}.$$

Example 7.26 For the function $f(x) = |x|\left|\cos x^{-1}\right|,\ f(0) = 0$, we calculate that

$$D^{+}f(x_0) = 1, D_{+}f(x_0) = 0, D^{-}f(x_0) = 0, D_{-}f(x_0) = -1.$$

Everywhere else $f'(x)$ exists and all four Dini derivatives have that value. ◀

Example 7.27 The function

$$g(x) = \begin{cases} 0, & \text{if } x \text{ is rational,} \\ 1, & \text{if } x \text{ is irrational.} \end{cases}$$

has at every rational x

$$D^{+}g(x) = 0, D_{+}g(x) = -\infty, D^{-}g(x) = \infty, D_{-}g(x) = 0.$$

For x irrational there are similar values for the Dini derivates (see Exercise 7:8.1a). ◀

It is easy to check that a function f has a derivative at a point x_0 if and only if all four Dini derivates are equal at that point, and a unilateral derivative at x_0 if the two Dini derivates from that side are equal (see Exercise 7:8.2).

We end this section with an illustration of the way in which knowledge about a Dini derivate can substitute for that of the ordinary derivative. We prove a theorem about monotonicity. The reader is familiar with the fact that if f is differentiable on an interval $[a, b]$ and $f'(x) > 0$ for all $x \in [a, b]$, then f is an increasing function on $[a, b]$. (We provided a formal proof in Section 7.7.)

Here is a generalization of that theorem.

Theorem 7.28 *Let f be continuous on $[a, b]$. If $D^{+}f(x) > 0$ for all $x \in [a, b)$, then f is increasing on $[a, b]$.*

Proof. Let us first show that f is nondecreasing on $[a, b]$. We prove this by contradiction. If f fails to be nondecreasing on $[a, b]$, there exist points c and d such that $a \leq c < d \leq b$ and $f(c) > f(d)$. Let y be any point in the interval $(f(d), f(c))$.

Since f is continuous on $[a, b]$, it possesses the intermediate value property. Thus from Theorem 5.48 (or more precisely from the version of that theorem given as Exercise 5:8.7(a)) there exists a point $t \in (c, d)$ such that $f(t) = y$. Thus the set $\{x : f(x) = y)\} \cap [c, d]$ is non empty. Let $x_0 = \sup\{x : c \leq x \leq d \text{ and } f(x) = y\}$. Now, $f(d) < y$ and f is continuous, from which it follows that $x_0 < d$. Thus $f(x) < y$ for $x \in (x_0, d]$. Furthermore, the set $\{x : f(x) = y\}$ is closed (because f is continuous), so $f(x_0) = y$.

But this implies $D^+f(x_0) \leq 0$.This contradicts our hypothesis that $D^+f(x) > 0$ for all $x \in [a, b)$. This contradiction completes the proof that f is nondecreasing.

Now we wish to show that it is in fact increasing. If not then there must be some subinterval in which the function is constant. But at every point interior to that interval we would have $f'(x) = 0$ and so it would be impossible for $D^+f(x) > 0$ at such points. ■

Exercises

7:8.1 Calculate the four Dini derivates for each of the following functions at the given point.

(a)
$$g(x) = \begin{cases} 1, & \text{if } x \text{ is rational} \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$
for $x = \pi$.

(b) $h(x) = x \sin x^{-1}$ $(h(0) = 0)$ at $x = 0$.

(c) $f(x) = x \sin x^{-1}$ $(f(0) = 5)$ at $x = 0$.

(d)
$$u(x) = \begin{cases} x^2, & \text{if } x \text{ is rational} \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$
at $x = 0$ and at $x = 1$.

7:8.2 Prove f has a derivative at x_0 if and only if
$$D^+f(x_0) = D_+f(x_0) = D^-f(x_0) = D_-f(x_0).$$
In that case, $f'(x_0)$ is the common value of the Dini derivates at x_0. (We assume that f is defined in a neighborhood of x_0.)

7:8.3 **(Derived Numbers)** The Dini derivates are sometimes called "extreme unilateral derived numbers". Let $\lambda \in [-\infty, \infty]$. Then λ is a *derived number* for f at x_0 if there exists a sequence $\{x_k\}$ with $\lim_{k\to\infty} x_k = x_0$ such that
$$\lambda = \lim_{k\to\infty} \frac{f(x_k) - f(x_0)}{x_k - x_0}.$$

(a) For the function $f(x) = \left|x \cos x^{-1}\right|$, $f(0) = 0$, show that every number in the interval $[-1, 1]$ is a derived number for f at $x = 0$. Show that the two extreme derived numbers from the right are 0 and 1, and the two from the left are -1 and 0.

(b) Show that a function has a derivative at a point if and only if all derived numbers at that point coincide.

(c) Let $f : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \mathbb{R}$. Prove that if f is continuous on $\mathbb{R}$ then the set of derived numbers of f at x_0 consists of either one or two closed intervals (that might be degenerate or unbounded). Give examples to illustrate the various possibilities.

7:8.4 Let $f, g : \mathbb{R} \to \mathbb{R}$.

(a) Prove that $D^+(f + g)(x) \leq D^+f(x) + D^+g(x)$.

(b) Give an example to illustrate that the inequality in (a) can be strict.

(c) State and prove the analogue of part (a) for the lower right derivate D_+f.

7:8.5 Generalize Theorem 7.17 to the following:

> *If f achieves a local maximum at x_0, then $D^+f(x_0) \leq 0$ and $D_-f(x_0) \geq 0$.*

Illustrate the result with a function that is not differentiable at x_0.

7:8.6 Prove a variant of Theorem 7.28 that assumes that the Dini derivate $D^+f(x) > 0$ for all x in $[a, b)$ except for x in some countable set.

7:8.7 Prove a variant of Theorem 7.28: If f is continuous and $D^+f(x) \geq 0$ for all $x \in [a, b)$ then f is nondecreasing on $[a, b]$. ✂

7:8.8 Prove yet another (more subtle) variant of Theorem 7.28: If f is continuous and $D^+f(x) > 0$ for all $x \in [a, b)$ except for x in some countable set then f is increasing on $[a, b]$. ✂

7:8.9 (a) Give an example of a function $f : \mathbb{R} \to \mathbb{R}$ such that $D^+f(x) = \infty$ for all $x \in \mathbb{R}$. ✂

(b) Prove that no continuous function can have $D^+f(x) = \infty$ for all $x \in \mathbb{R}$.

7:8.10 (a) Give an example of a function f such that $D^+f < D_-f$ on an infinite set. ✂

(b) Show that $\{x : D^+f(x) < D_-f(x)\}$ cannot be uncountable.

7.9 The Darboux Property of the Derivative

Suppose f is differentiable on an interval $[a, b]$. We argued in the proof of Rolle's theorem (7.18) that if $f(a) = f(b)$, then there exists a point $c \in (a, b)$ at which f achieves an extremum. At this point c we have $f'(c) = 0$.

A different hypothesis can lead to the same conclusion. Suppose f is differentiable on $[a, b]$ and $f'(a) < 0 < f'(b)$ (or $f'(b) < 0 < f'(a)$). Once again the extreme value f achieves must occur at a point c in the *interior* of $[a, b]$, (why?), and at this point we must have $f'(c) = 0$. This observation is a special case of the following theorem first proved by Darboux in 1875.

Theorem 7.29 *Let f be differentiable on an interval I. Suppose $a, b \in I$; $a < b$, and $f'(a) \neq f'(b)$. Let γ be any number between $f'(a)$ and $f'(b)$. Then there exists $c \in (a, b)$ such that $f'(c) = \gamma$.*

Proof. Let $g(x) = f(x) - \gamma x$. If $f'(a) < q < f'(b)$, then $g'(a) = f'(a) - \gamma < 0$ and $g'(b) = f'(b) - \gamma > 0$. The discussion preceding the statement of the theorem shows that there exists $c \in (a, b)$ such that $g'(c) = 0$. For this c we have

$$f'(c) = g'(c) + \gamma = \gamma,$$

completing the proof for the case $f'(a) < f'(b)$.

The proof when $f'(a) > f'(b)$ is similar. ∎

The reader might have noted that Theorem 7.29 is exactly the statement that the derivative of a differentiable function has the Darboux property (i.e., the intermediate value property) that we established for continuous functions in Section 5.8. The derivative f' of a differentiable function f need not be continuous, of course. The result does imply however that f' cannot have jump discontinuities and cannot have removable discontinuities.

Both the mean value theorem and Theorem 7.29 give information about the range of the derivative f' of a differentiable function f. The mean value theorem implies that the range of f' includes all slopes of chords determined by the graph of f on the interval of definition of f. Theorem 7.29 tells us that this range is actually an interval. This interval may be unbounded and, if bounded, may or may not contain its end points. (See Exercise 7:9.1.)

Derivative of an Inverse Function Theorem 7.29 allows us to establish a familiar theorem about differentiating inverse functions.

Theorem 7.30 *Suppose f is differentiable on an interval I and for all $x \in I$, $f'(x) \neq 0$. Then*

(i) *f is one-to-one on I,*

(ii) *f^{-1} is differentiable on $J = f(I)$,*

(iii) $(f^{-1})'(f(x)) = \dfrac{1}{f'(x)}$ *for all $x \in I$.*

Proof. By Theorem 7.29 either $f'(x) > 0$ for all $x \in I$ or $f'(x) < 0$ for all $x \in I$. In either case, f is either increasing or decreasing on I, and is thus one-to-one, establishing (i).

To verify (ii) and (iii) observe first that f^{-1} is continuous, since f is continuous and strictly monotonic (see Exercise 5:9.17). Let $y_0 \in J$ and let $x_0 = f^{-1}(y_0)$. We wish to show that $(f^{-1})'(y_0)$ exists and has value $1/(f'(x_0))$. For $x \in I$, write $y = f(x)$, so $x = f^{-1}(y)$.

Consider the difference quotient

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)}.$$

As $y \to y_0$, $x \to x_0$, because the function f^{-1} is continuous. Thus

$$\lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{x \to x_0} \frac{1}{\left(\frac{f(x) - f(x_0)}{x - x_0}\right)} = \frac{1}{f'(x_0)}.$$

■

Exercises

7:9.1 Let f be differentiable on $[a, b]$ and let $\mathcal{R}(f')$ denote the range of f' on $[a, b]$. Give examples to illustrate that $\mathcal{R}(f')$ can be

(a) a closed interval.

(b) an open interval.

(c) a half-open interval.

(d) an unbounded interval.

7:9.2 (a) Give an example of a differentiable function f such that $f'(x_0) \neq \lim_{x \to x_0} f'(x)$.

(b) Show that if f is defined in a neighborhood of x_0 and $\lim_{x \to x_0} f'(x)$ exists and is finite, then f is differentiable at x_0 and f' is continuous at x_0.

7:9.3 Most classes of functions we have encountered are closed under the operations of addition and multiplication (e.g. polynomials, continuous functions, differentiable functions). The class of derivatives is closed under addition, but behaves very badly with respect to multiplication. Consider for example the pair of functions F and G defined on $\mathbb{R}$ by

$$\begin{aligned} F(x) &= x^2 \sin \frac{1}{x^3}, (F(0)=0), \text{ and} \\ G(x) &= x^2 \cos \frac{1}{x^3}, (G(0)=0). \end{aligned}$$

Verify each of the following statements:

(a) F and G are differentiable on $\mathbb{R}$.

(b) The functions FG' and GF' are bounded functions.

(c)
$$F(x)G'(x) - F'(x)G(x) = \begin{cases} 3, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

(d) One at least of the functions FG' or GF' must fail to be a derivative.

Thus, even the product of a differentiable function F with a derivative G' need not be a derivative.

7:9.4 Show, in contrast to Exercise 7:9.3, that if a function f has a continuous derivative on $\mathbb{R}$ and g is differentiable, then fg' is a derivative.

7:9.5 Let $f : [0,1] \to \mathbb{R}$ be a continuous function that is differentiable on $(0,1)$ and with $f(0)=0$ and $f(1)=1$. Show there must exist distinct numbers ξ_1 and ξ_2 in that interval such that

$$f'(\xi_1)f'(\xi_2) = 1.$$

7:9.6 Prove or disprove that if $f:\mathbb{R}\to\mathbb{R}$ is differentiable and monotonic, then f' must be continuous on $\mathbb{R}$.

7.10 Convexity

In elementary calculus one studies functions that are concave-up or concave-down on an interval. A knowledge of the intervals on which a function is concave-up or concave-down is useful for such purposes as sketching the graph of the equation $y = f(x)$ and studying extrema of the function. (Figure 7.10 illustrates.)

In the setting of elementary calculus the functions usually have second derivatives on the intervals involved. In that setting one can define a function as being concave-up on an interval I if $f'' \geq 0$ on

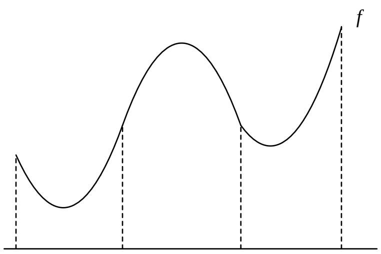


Figure 7.10: Concave up/down/up.

I, and concave-down if $f'' \leq 0$ on I. Definitions involving the first derivative, but not the second, can also be given: f is concave-up on I if f' is increasing on I, concave-down if f' is decreasing on I. Equivalently, f is concave-up if the graph of f lies "above" (more precisely "not below") each of its tangent lines, concave-down if the graph lies below (not above) each of its tangent lines.

The geometric properties one wishes to capture when one says a function is concave-up or down do not depend on differentiability properties. The condition is that the graph should lie below (or above) all its chords. The definitions below make this concept precise. We shall follow the common practice of using the terms "convex" and "concave" in place of the terms "concave-up" and "concave-down".

Definition 7.31 Let f be defined on an interval I. If for all $x_1, x_2 \in I$ and $\alpha \in [0, 1]$ the inequality

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \qquad (13)$$

is satisfied, we say that f is *convex* on I. If the reverse inequality in (13) applies, we say that f is *concave* on I. If the inequalities are strict for all $\alpha \in (0, 1)$ we say f is *strictly convex* or *strictly concave* on I.

For example the function $f(x) = |x|$ is convex, but not strictly convex on $\mathbb{R}$. Strict convexity implies the graph of f has no line segments in it. Note that the function $f(x) = |x|$ is not differentiable at $x = 0$.

The geometric condition defining convexity does imply a great deal of regularity of a function. Our first objective is to address this issue. We begin with some simple geometric considerations.

Suppose f is convex on an open interval I. Let x_1 and x_2 be

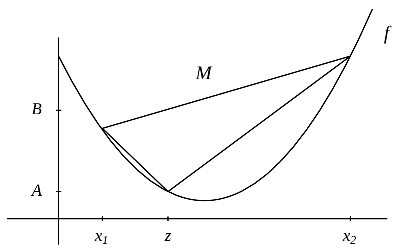


Figure 7.11: Comparison of the three slopes in the inequalities (14).

points in I with $x_1 < x_2$. The chord determined by the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$ defines a linear function M on $[x_1, x_2]$: if $x = \alpha x_1 + (1-\alpha)x_2$ then

$$M(x) = \alpha f(x_1) + (1-\alpha)f(x_2).$$

The definition of convex states that $f(x) \leq M(x)$ for all $x \in [x_1, x_2]$ and that $M(x_1) = f(x_1)$ and $M(x_2) = f(x_2)$. Now let $z \in (x_1, x_2)$ Then

$$\frac{f(z)-f(x_1)}{z-x_1} \leq \frac{M(z)-M(x_1)}{z-x_1} = \frac{M(x_2)-M(z)}{x_2-z} \leq \frac{f(x_2)-f(z)}{x_2-z} \tag{14}$$

(Figure 7.11 illustrates.)

Thus, the chord determined by f and the points x_1 and x_2 has a slope between the slopes of the chord determined by x_1 and z and the chord determined by z and x_2.

The inequalities (14) have a number of useful consequences:

1. For fixed $x \in I$, $(f(x+h)-f(x))/h$ is a nondecreasing function of h on some interval $(0, \delta)$ Thus
$$\lim_{h \to 0+} \frac{f(x+h)-f(x)}{h} = \inf_{h>0} \frac{f(x+h)-f(x)}{h}$$
exists or possibly is $-\infty$. That it is in fact finite can be shown by using (14) again to get a finite lower bound, since
$$(f(x')-f(x))/(x'-x) \leq (f(x+h)-f(x))/h$$
for any $x' \in I$ with $x' < x$. Thus f has a right derivative $f'_+(x)$ at x. Similarly f has a finite left derivative at x.

2. If $x, y \in I$ and $x < y$, then $f'_+(x) \leq f'_+(y)$.

From observation 1 we infer that

$$(f(x+h) - f(x))/h \leq (f(y+h) - f(y))/h$$

whenever $h > 0$ and $x+h, y+h$ are in I. Thus f'_+ is a nondecreasing function. Similarly f'_- is a nondecreasing function.

3. It is also clear from (14) that $f'_-(x) \leq f'_+(x)$ for all $x \in I$.

4. f is continuous on I. To see this, observe that since both one sided derivatives exist at every point the function must be continuous on both sides, hence continuous.

We summarize the preceding discussion as a theorem.

Theorem 7.32 *Let f be convex on an open interval I. Then*

(i) *f has finite left and right derivatives at each point of I. Each of these unilateral derivatives is a nondecreasing function of x on I, and*

$$f'_-(x) \leq f'_+(x) \quad \textit{for all } x \in I. \tag{15}$$

(ii) *f is continuous on I*

Note. If f is convex on a *closed* interval $[a, b]$, some of the results do not apply at the endpoints a and b. (See Exercise 7:10.3.) Note, too, that the corresponding results are valid for concave functions on I, the unilateral derivatives now being decreasing functions of x and the inequality in (15) being reversed.

We can now obtain the characterizations of convex functions familiar from elementary calculus.

Corollary 7.33 *Let f be defined on an open interval I.*

(i) *If f is differentiable on I, then f is convex on I if and only if f' is nondecreasing on I.*

(ii) *If f is twice differentiable on I, then f is convex on I if and only if $f'' \geq 0$ on I.*

We leave the verification of Corollary 7.33 as Exercise 7:10.4.

Exercises

7:10.1 Suppose f is convex on an open interval I. Prove that f is differentiable except on a countable set.

7:10.2 Suppose f is convex on an open interval I. Prove that if f is differentiable on I, then f' is continuous on I.

7:10.3 Theorem 7.32 pertains to functions that are convex on an open interval. Discuss the extent to which the results of the theorem hold when f is convex on a *closed* interval $[a, b]$. In particular, determine whether continuity of f at the endpoints of the interval follows from the definition. Must $f'_+(a)$ and $f'_-(b)$ be finite?

7:10.4 Prove Corollary 7.33.

7:10.5 Let f be convex on an open interval (a, b). Show that f does not have a strict maximum value.

7:10.6 Let f be defined and continuous on an open interval (a, b). Show that f is convex there if and only if there do not exist real numbers α and β such that the function $f(x) + \alpha x + \beta$ has a strict maximum value in (a, b).

7:10.7 Let $A = \{a_1, a_2, a_3, \dots\}$ be any countable set of real numbers. Let

$$f(x) = \sum_{1}^{\infty} \frac{|x - a_k|}{10^k}.$$

Prove that f is convex on $\mathbb{R}$, differentiable on $\mathbb{R} \setminus A$ and non-differentiable on A.

7:10.8 (Inflection Points) In elementary calculus one studies inflection points. The definitions one finds try to capture the idea that at such a point the sense of concavity changes from strict "up to down" or vice-versa. Here are three common definitions that apply to differentiable functions. In each case f is defined on an open interval (a, b) containing the point x_0. The point x_0 is an *inflection point* for f if there exists an open interval $I \subset (a, b)$ such that on I

(Definition A) f' strictly increases on one side of x_0 and strictly decreases on the other side.

(Definition B) f' attains a strict maximum or minimum at x_0.

(Definition C) The tangent line to the graph of f at $(x_0, f(x_0))$ lies below the graph of f on one side of x_0 and above on the other side.

(a) Prove that if f satisfies Definition A at x_0, then it satisfies Definition B at x_0.

(b) Prove that if f satisfies Definition B at x_0, then it satisfies Definition C at x_0.

(c) Give an example of a function satisfying Definition B at x_0, but not satisfying Definition A.

(d) Give an example of an infinitely differentiable function satisfying Definition C at x_0, but not satisfying Definition B.

(e) Which of the three definitions states that the sense of concavity of f is "up" on one side of x_0 and "down" on the other?

The three definitions above are not equivalent even for infinitely differentiable functions. They are, however, equivalent for *analytic* functions, that is functions represented by power series (a topic we cover in Chapter 10). Since the scope of elementary calculus is more or less limited to functions that are analytic on the intervals on which the functions are concave up or down, one might argue that on that level, the definition to take is the one the author finds simplest to develop. We should mention, however, that there are differentiable functions that are not concave up or down on any interval!

7.11 L'Hôpital's Rule

Suppose f and g are defined in a deleted neighborhood of x_0 and $\lim_{x\to x_0} f(x) = A$, $\lim_{x\to x_0} g(x) = B$. According to our usual theory of limits we then have

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)} = \frac{A}{B},$$

unless $B = 0$.

But what happens if $B = 0$ which is often the case? A number of possibilities exist: if $B = 0$ and $A \neq 0$ then the limit does not exist. The most interesting case remains: if both A and B are zero then the limiting behavior depends on the rates at which $f(x)$ and $g(x)$ approach zero.

Example 7.34 Consider

$$\lim_{x\to 0} \frac{6x}{3x} = \lim_{x\to 0} \frac{6}{3} = 2.$$

Look at this simple example geometrically. For $x \neq 0$, the height $6x$ is twice that of the height $3x$. The straight line $y = 6x$ approaches zero at twice the rate that the line $y = 3x$ does. ◀

Example 7.35 Now consider the slightly more complicated limit

$$\lim_{x\to 0} \frac{6x + x^2}{3x + 5x^3}.$$

If we divide the numerator and denominator by $x \neq 0$, we see that the limit is the same as

$$\lim_{x\to 0} \frac{6 + x}{3 + 5x^2}.$$

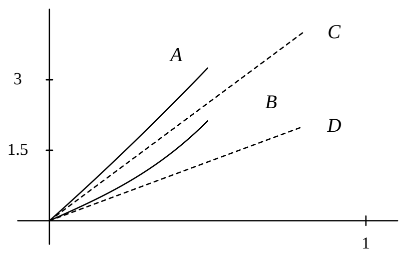


Figure 7.12: Comparison of the rates in Example 7.35.

This last limit can be calculated by our usual elementary methods as equaling $6/3 = 2$. Here, for $x \neq 0$ near zero, the height $6x + x^2$ is approximately $6x$, while the height of $3x + 5x^3$ is approximately $3x$, that is, the desired ratio is approximately 2. (See Figure 7.12 where there is an illustration showing graphs and the values of $6x + x^2$, $6x$, $3x + 5x^3$, and $3x$ at $x = .1$.) Again, the numerator approaches zero at about twice the rate that the denominator does.

We can be more precise by calculating these rates exactly. Let $f(x) = 6x + x^2$, $g(x) = 3x + 5x^3$. Then

$$\begin{aligned} f'(x) &= 6 + 2x, \quad f'(0) = 6 \\ g'(x) &= 3 + 5x^2, \quad g'(0) = 3. \end{aligned}$$

This makes precise our statement that the numerator approaches zero twice as fast as the denominator does. ◀

Let us try to generalize from these two examples. Suppose f and g are differentiable in a neighborhood of $x = a$ and that $f(a) = g(a) = 0$. Consider the following calculations and what conditions on f and g are required to make them valid.

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{\left(\frac{f(x)-f(a)}{x-a}\right)}{\left(\frac{g(x)-g(a)}{x-a}\right)} \xrightarrow{x \to a} \frac{f'(a)}{g'(a)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}. \quad (16)$$

If these calculations are valid, they show that under these assumptions ($f(a) = g(a) = 0$ and both $f'(a)$ and $g'(a)$ exist) we should be able to claim that

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

The reader should check the various conditions that must be met to justify the calculations: $g(x)$ cannot equal zero at any point of the neighborhood in question (other than a); nor can $g(x) = g(a)$, (for $x \neq a$); $f(a)$ and $g(a)$ must equal zero (for the first equality), and f'/g' must be continuous at $x = a$ (for the last equality).

The calculations (16) provide a simple proof of a rudimentary form for a method of computing limits known as L'Hôpital's Rule. We say "rudimentary", because some of the conditions we assumed are not needed for the conclusion

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

Just because the rule is named after L'Hôpital one should not assume that he was the discoverer; this is a common feature in the way things are labelled in our subject.

7.11.1 L'Hôpital's Rule: $\frac{0}{0}$ Form

Our first theorem provides a version of the rule identical with our introductory remarks but under weaker assumptions.

Theorem 7.36 (L'Hôpital's Rule: $\frac{0}{0}$ Form) *Let f and g be differentiable in a deleted neighborhood N of $x = a$. Suppose*

1. $\lim_{x \to a} f(x) = 0$,
2. $\lim_{x \to a} g(x) = 0$,
3. *For every $x \in N$, $g'(x) \neq 0$, and*
4. $\lim_{x \to a} \frac{f'(x)}{g'(x)}$ *exists.*

Then $\displaystyle\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}$.

Proof. Our hypotheses do not require f and g to be defined at $x = a$. But we can in any case define (or redefine) f and g at $x = a$ by $f(a) = g(a) = 0$. Because of assumptions 1 and 2, this results in continuous functions defined on the full neighborhood $N \cup \{a\}$ of the point $x = a$. We can now apply Cauchy's form of the mean value theorem (7.20).

Suppose $x \in N$ and $a < x$. By Theorem 7.20 there exists $c = c_x$ in (a, x) such that

$$[f(x) - f(a)]g'(c_x) = [g(x) - g(a)]f'(c_x). \tag{17}$$

Since $f(a) = g(a) = 0$, (17) becomes

$$f(x)g'(c_x) = g(x)f'(c_x). \tag{18}$$

Equation (18) is valid for $x > a$ in N. We would like to express (18) in the form

$$\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)}. \tag{19}$$

To justify (19) we show that $g(x)$ is never zero in $N \cap \{x : x > a\}$. (That $g'(c_x)$ is never zero in N is our hypothesis 3.) If for some $x \in N$, $x > a$, we have $g(x) = 0$, then by Rolle's theorem there would exist a point $t \in (a, x)$ such that $g'(t) = 0$, contradicting hypothesis 2. Thus equation (19) is valid for all $N \cap \{x : x > a\}$. A similar argument shows that if $x \in N$, $x < a$, then there exists $c_x \in (x, a)$ such that (19) holds.

Now as $x \to a$, c_x also approaches a, since c_x is between a and x. Thus

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f(c_x)}{g(c_x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)},$$

since the last limit exists by hypothesis 4. ■

Observe that we did not require f to be defined at $x = a$, nor did we require that f'/g' be continuous at $x = a$.

Example 7.37 Let us use L'Hôpital's Rule to evaluate

$$\lim_{x\to 0} \ln(1+x)/x.$$

Let $f(x) = \ln(1+x)$, $g(x) = x$. Then

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} g(x) = 0, f'(x) = \frac{1}{1+x}, g'(x) = 1.$$

Thus

$$\lim_{x\to 0} \frac{\ln(1+x)}{x} = \lim_{x\to 0} \frac{1}{1+x} = 1.$$

◀

It is important to observe that L'Hôpital's Rule does *not* imply that, under hypotheses (i) (ii) and (iii), if $\lim_{x\to 0} f(x)/g(x)$ exists, then $\lim_{x\to 0} f'(x)/g'(x)$ also exists. Exercise 7:11.5 provides an example to illustrate this.

We refer to this theorem as the "$\frac{0}{0}$ Form" for obvious reasons. There is also a version of the form $\frac{\infty}{\infty}$, (see Theorem 7.40). In addition, other modifications are possible. The point a can be replaced with $a = \infty$ or $a = -\infty$, (Theorem 7.39), and the results are valid for one-sided limits. (Our proof of Theorem 7.36 actually established that fact since we considered the case $x > a$ and $x < a$ separately.) Various other "indeterminate forms" — ones for which the limit depends on the rates at which component parts approach their separate limits — can be manipulated to make use of L'Hôpital's Rule possible.

Here is an example in which the forms "1^∞" and "$1^{-\infty}$" come into play. Observe that the function whose limit we wish to calculate is of the form $f(x)^{g(x)}$ where $f(x) \to 1$ as $x \to a$ but $g(x) \to \infty$ as $x \to a+$ and $g(x) \to -\infty$ as $x \to a-$.

Example 7.38 Evaluate $\lim_{x\to 0}(1+x)^{2/x}$. This expression is of the form 1^∞ (when $x > 0$). To calculate $\lim_{x\to 0}(1+x)^{2/x}$, write

$$y = (1+x)^{2/x}, z = \ln y = \frac{2}{x}\ln(1+x).$$

Now the numerator and denominator of the function z satisfy the hypotheses of L'Hôpital's Rule. Thus

$$\lim_{x\to 0} z = \lim_{x\to 0}\frac{2\ln(1+x)}{x} = \lim_{x\to 0}\frac{2}{1+x} = 2.$$

Since $\lim_{x\to 0} z = 2$, $\lim_{x\to 0} y = e^2$. ◀

7.11.2 L'Hôpital's Rule as $x \to \infty$

We proved Theorem 7.36 under the assumption that $a \in \mathbb{R}$, but the theorem is valid when $a = -\infty$ or $a = +\infty$. In this case we are, of course, dealing with one-sided limits. As before, the relation

$$\lim_{x\to\infty}\frac{f'(x)}{g'(x)} = L$$

implies something about relative rates of growth of the functions $f(x)$ and $g(x)$ as $x \to \infty$. We can base a proof of the versions of L'Hôpital's Rule that have $a = \infty$ (or $-\infty$) on Theorem 7.36 by a simple transformation.

Theorem 7.39 *Let f, g be differentiable on some interval $(-\infty, b)$. Suppose*

(i) $\lim_{x\to-\infty} f(x) = 0$,

(ii) $\lim_{x\to-\infty} g(x) = 0$,

(iii) *For every $x \in (-\infty, b)$, $g'(x) \neq 0$, and*

(iv) $\lim_{x\to-\infty}\frac{f'(x)}{g'(x)}$ *exists.*

Then $\displaystyle\lim_{x\to-\infty}\frac{f(x)}{g(x)} = \lim_{x\to-\infty}\frac{f'(x)}{g'(x)}$. *A similar result holds when we replace ∞ by $-\infty$ in the hypotheses.*

Proof. Let $x = -1/t$. Then, as $t \to 0+$, $x \to -\infty$ and vice-versa. Define functions F and G by

$$F(t) = f\left(-\frac{1}{t}\right) \quad \text{and} \quad G(t) = g\left(-\frac{1}{t}\right).$$

Both functions F and G are defined on some interval $(0, \delta)$. We verify easily that

$$\lim_{t\to 0+} F(t) = \lim_{t\to 0+} G(t) = 0$$

and that

$$\lim_{t\to 0+} \frac{F'(t)}{G'(t)} = \lim_{x\to -\infty} \frac{f'(x)}{g'(x)}. \tag{20}$$

Using Theorem 7.36 we infer

$$\lim_{t\to 0+} \frac{F'(t)}{G'(t)} = \lim_{t\to 0+} \frac{F(t)}{G(t)} = \lim_{t\to 0+} \frac{f(-\frac{1}{t})}{g(-\frac{1}{t})} = \lim_{x\to -\infty} \frac{f(x)}{g(x)}. \tag{21}$$

The result follows from (20) and (21) ■

7.11.3 L'Hôpital's Rule: $\frac{\infty}{\infty}$ Form

When $f(x) \to \infty$ and $g(x) \to \infty$ as $x \to a$ we obtain the indeterminate form $\frac{\infty}{\infty}$. L'Hôpital's theorem then takes the form below. Note, however, that we don't require $f(x) \to \infty$ in our hypotheses, or even that $f(x)$ approaches any limit.

Theorem 7.40 *Let f and g be differentiable on a deleted neighborhood N of $x = a$.*

(i) $\lim_{x\to a} g(x) = \infty$.

(ii) *For every $x \in N$ $g'(x) \neq 0$.*

(iii) $\lim_{x\to a} f'(x)/g'(x)$ *exists.*

Then $\displaystyle\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$. *The analogous statements are valid if $a = \pm\infty$ or $\lim_{x\to a} g(x) = -\infty$.*

Proof. We prove the main part of Theorem 7.40 under the assumption that $\lim_{x\to a} f'(x)/g'(x)$ is finite. The case that the limit is infinite as well as variants are left as Exercises 7:11.6 and 7:11.7. It suffices to consider the case of right-hand limits, the proof for left-hand limits being similar. Let $L = \lim_{x\to a+} f'(x)/g'(x)$. We will show that if $p < L < q$, then there exists $\delta > 0$ such that for $x \in (a, a+\delta)$, $p < f(x)/g(x) < q$. Since p and q are arbitrary (subject to the restriction $p < L < q$), we can then conclude $\lim_{x\to a+} f'(x)/g'(x) = L$ as required.

Choose $r \in (L, q)$. By (iii) and the definition of L there exists δ_1 such that $f'(x)/g'(x) < r$ whenever $x \in (a, a+\delta_1)$. If $a < x < y < a+\delta_1$, then we infer from Theorem 7.20, Cauchy's form of the mean value theorem, and our assumption (ii) that there exists $c \in (x, y)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(c)}{g'(c)} < r. \tag{22}$$

Fix y in (22). Since $\lim_{x\to a+} g(x) = \infty$, there exists $\delta_2 > 0$ such that $a + \delta_2 < y$ and such that $g(x) > g(y)$ and $g(x) > 0$ if $a < x < a + \delta_2$. We then have $(g(x) - g(y))/g(x) > 0$ for $x \in (a, a+\delta_2)$, so we can multiply both sides of the inequality (22) by $(g(x) - g(y))/g(x)$ obtaining

$$\frac{f(x)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)} \text{ for } x \in (a, a+\delta_2). \tag{23}$$

Now let $x \to a+$. Then $g(x) \overset{x\to a+}{\to} \infty$ by assumption (i). Since r, $g(y)$ and $f(y)$ are constants, the second and third terms on the right side of (23) approach zero. It now follows from the inequality $r < q$ that there exists $\delta_3 \in (0, \delta_2)$ such that

$$\frac{f(x)}{g(x)} < q \quad \text{whenever} \quad a < x < a + \delta_3. \tag{24}$$

In a similar fashion we find a $\delta_4 > 0$ such that

$$\frac{f(x)}{g(x)} > p \quad \text{whenever} \quad a < x < a + \delta_4. \tag{25}$$

If we let $\delta = \min(\delta_3, \delta_4)$, we have shown that

$$p < \frac{f(x)}{g(x)} < q \quad \text{whenever} \quad x \in (a, a+\delta).$$

Since p and q were arbitrary numbers satisfying $p < L < q$, our conclusion

$$\lim_{x\to a+} \frac{f(x)}{g(x)} = L = \lim_{x\to a+} \frac{f'(x)}{g'(x)}$$

follows. ■

Exercises

7:11.1 Consider the function

$$f(x) = (3^x - 2^x)/x$$

defined everywhere except at $x0$.

(a) What value should be assigned to $f(0)$ in order that f be everywhere continuous?

(b) Does $f'(0)$ exist if this value is assigned to $f(0)$?

(c) Would it be correct to calculate $f'(0)$ by computing instead $f'(x)$ by the usual rules of the calculus and finding $\lim_{x\to 0} f'(x)$.

7:11.2 Suppose f and g are defined in a deleted neighborhood of x_0 and $\lim_{x\to x_0} f(x) = A \neq 0$ and $\lim_{x\to x_0} g(x) = 0$. Show that

$$\lim_{x\to x_0} \left|\frac{f(x)}{g(x)}\right| = \infty.$$

7:11.3 Discuss the limiting behavior as $x \to 0$ for each of the functions below.

(a) $\dfrac{1}{x}$ (b) $\dfrac{1}{x^2}$

(c) $\dfrac{1}{\sin x}$ (d) $\dfrac{1}{x\sin x^{-1}}$

7:11.4 Evaluate each of the following limits

(a) $\displaystyle\lim_{x\to 0} \frac{e^x - \cos x}{x}$

(b) $\displaystyle\lim_{t\to 0} \frac{\sin t - t}{t^3}$

(c) $\displaystyle\lim_{u\to 1} \frac{u^5 + 5u - 6}{2u^5 + 8u - 10}$

7:11.5 Let $f(x) = x^2 \sin x^{-1}$, $g(x) = x$.

Show that $\displaystyle\lim_{x\to 0} \frac{f(x)}{g(x)} = 0$ but $\displaystyle\lim_{x\to 0} \frac{f'(x)}{g'(x)}$ does not exist.

7:11.6 The proof we provided for Theorem 7.40 required that $\lim_{x\to a} f'(x)/g'(x)$ be finite. Prove the result holds if this limit is infinite.

7:11.7 Prove the part of Theorem 7.40 dealing with $a = \pm\infty$ or $\lim_{x\to a} g(x) = -\infty$.

7:11.8 Evaluate the following limits.

(a) $\displaystyle\lim_{x\to\infty} \frac{x^3}{e^x}$ (b) $\displaystyle\lim_{x\to\infty} \frac{\ln x}{x}$

(c) $\displaystyle\lim_{x\to 0+} x\ln x$ (d) $\displaystyle\lim_{x\to 0+} x^x$

7:11.9 This exercise gives information about the relative rates of increase of certain types of functions. Prove that for each positive number p,

$$\lim_{x\to\infty} \frac{\ln x}{x^p} = \lim_{x\to\infty} \frac{x^p}{e^x} = 0.$$

7:11.10 Give an example of functions f and g defined on $\mathbb{R}$ such that

$$\lim_{x\to\infty} g(x) = \infty,\ \limsup_{x\to\infty} f(x) = \infty,\ \liminf_{x\to\infty} f(x) = -\infty$$

and Theorem 7.40 applies.

7.12 Taylor Polynomials

Suppose f is continuous on an open interval I and $c \in I$. The constant function $g(x) = f(c)$ approximates f closely when x is sufficiently close to the point c, but may or may not provide a good approximation elsewhere. If f is differentiable on I, then we see from the mean value theorem (Theorem 7.19) that for each $x \in I$ $(x \neq c)$ there exists z between x and c such that

$$f(x) = f(c) + f'(z)(x - c).$$

The expression $R_0(x) = f'(z)(x - c) = f(x) - f(c)$ provides the size of the error obtained in approximating the function f by the zero-degree polynomial (i.e. constant) $P_0(x) = f(c)$.

We do not expect a constant function to be a good approximation to a given continuous function in general. But our acquaintance with Taylor series (as presented in elementary calculus courses) suggests that if a function is sufficiently differentiable, it can be approximated well by polynomials of sufficiently high degree.

Suppose we wish to approximate f by a polynomial P_n of degree n. In order for the polynomial P_n to have a chance to approximate f well in a neighborhood of a point c, we should require

$$P_n(c) = f(c), P_n'(c) = f'(c), \ldots, P_n^{(n)}(c) = f^{(n)}(c).$$

In that case we at least guarantee that P_n "starts out" with the correct value, the correct rate of change etc. to give it a chance to approximate f well in some neighborhood of c. The test however is this. Write

$$f(x) = P_n(x) + R_n(x).$$

Is it true that the "error" or "remainder" $R_n(x)$ is small when $x \in I$?

In order to answer this sort of question, it would be useful to have workable forms for this error term $R_n(x)$. We present two forms for the remainder. The first is due to Lagrange, who obtained Theorem 7.41 below in 1797. He used integration methods to prove the theorem. We provide a popular and more modern proof based on the mean value theorem.

Theorem 7.41 (Lagrange) *Let f possess at least $n+1$ derivatives on an open interval I and let $c \in I$. Let*

$$P_n(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n$$

and let $R_n(x) = f(x) - P_n(x)$. Then for each $x \in I$ there exists z between x and c ($z = c$ if $x = c$) such that

$$R_n(x) = \frac{f^{(n+1)}(z)}{(n+1)!}(x-c)^{n+1}.$$

Note. The function P_n is called the *nth Taylor polynomial for f*. The reader will recognize P_n as the nth partial sum of the Taylor series studied in elementary calculus. (See also Chapter 10.) The function R_n is called the *remainder* or *error* function between f and P_n. If P_n is to be a good approximation to f, then R_n must be small in absolute value.

Observe that $P_n(c) = f(c)$ and that

$$P_n^{(k)}(c) = f^{(k)}(c) \quad \text{for} \quad k = 0, 1, 2, \ldots, n.$$

Observe also that the mean value theorem is the special case of Theorem 7.41 obtained by taking $n = 0$.

Proof. Fix $x \in I$. Then there is a number M (depending on x, of course) such that

$$f(x) = P_n(x) + M(x-c)^{n+1}.$$

We wish to show that $M = (f^{(n+1)}(z))/n!$ for some z between x and c.

Consider the function g defined on I by

$$\begin{aligned} g(t) &= f(t) - P_n(t) - M(t-c)^{n+1} \\ &= R_n(t) - M(t-c)^{n+1}. \end{aligned}$$

Now P_n is a polynomial of degree at most n, so $P_n^{(n+1)}(t) = 0$ for all $t \in I$. Thus

$$g^{(n+1)}(t) = f^{(n+1)}(t) - (n+1)!M \quad \text{for all } t \in I. \tag{26}$$

Also, since $f^{(k)}(c) = P_n^{(k)}(c)$ for $k = 1, 2, \ldots, n$, we readily see that

$$g^{(k)}(c) = 0 \quad \text{for } k = 0, 1, 2, \ldots, n. \tag{27}$$

Suppose now that $x > c$, the case $x < c$ having a similar proof, and the case $x = c$ being obvious. We apply Rolle's theorem to the interval $[c, x]$. Since $g(x) = 0$ (by definition of M) and $g(c) = 0$ (by (27)), there exists $z_1 \in (c, x)$ such that$g'(z_1) = 0$.

Now apply Rolle's theorem to g' on the interval $[c, z_1]$, obtaining a point $z_2 \in (c, z_1)$ such that $g''(z_2) = 0$.

Continuing in this way we use (27) and Rolle's theorem repeatedly to obtain a point $z_n \in (c, z_{n-1})$ such that $g^{(n)}(z_n) = 0$. Finally, we apply the Rolle's theorem to the function $g^{(n)}$ on the interval $[c, z_n]$. We obtain a point $z \in (c, z_n)$ such that $g^{(n+1)}(z) = 0$. From (26) we deduce $f^{(n+1)}(z) = (n+1)!M$, completing the proof. ■

Lagrange's result expresses the error term R_n in a particular way. It provides a sense of the error in approximating f by P_n. Note that we do not get an exact statement of the error term since it is given in terms of the value $f^{(n+1)}(z)$ at *some* point z. But if we know a little bit about the function $f^{(n+1)}$ on the interval in question we might be able to say that this error is not very large.

Example 7.42 Suppose we wish to approximate the function $f(x) = \sin x$ on the interval $[-a, a]$ by a Taylor polynomial of degree 3, with $c = 0$. Here

$f'(x) = \cos x$, $f''(x) = -\sin x$, $f'''(x) = -\cos x$ and $f^{(4)}(x) = \sin x$.

Thus

$$\begin{aligned} P_3(x) &= \cos(0)x - \frac{\sin(0)}{2!}x^2 - \frac{\cos(0)}{3!}x^3 = x - \frac{x^3}{3!} \quad \text{and} \\ R_3(x) &= \frac{\sin z}{4!}x^4 \quad \text{for some } z \text{ in } [-a, a]. \end{aligned}$$

The exact error depends on which z makes this all true. But since $|\sin z| \leq 1$ no matter where z is we get immediately that

$$|R_3(x)| \leq a^4/4! = a^4/24,$$

so P_3 approximates f to within $a^4/24$ on the interval $[-a, a]$. For a small, the approximation should be sufficient for the purposes at hand. For large a, a higher degree polynomial can produce the desired accuracy, since

$$|R_n(x)| \leq \frac{|x^{n+1}|}{(n+1)!}.$$

◀

Various other forms for the error term R_n are useful. The so-called Integral form is one of them. We state this form without proof. We assume the reader is familiar with the integral as studied in calculus courses. (This theorem may be omitted until one has studied Chapter 8.)

Theorem 7.43 (Integral Form of Remainder) *Suppose f possesses at least $n+1$ derivatives on an open interval I and $f^{(n+1)}$*

is Riemann integrable on every closed interval contained in I. Let $c \in I$, Then

$$R_n(x) = \frac{1}{n!}\int_c^x f^{(n+1)}(t)(x-t)^n\,dt \quad \textit{for all } x \in I.$$

We shall see this form of the remainder again in Chapter 10 when we study Taylor series.

Exercises

7:12.1 Exhibit the Taylor polynomial about $x = 0$ of degree n for the function $f(x) = e^x$. Find n so that $|R_n(x)| \leq .0001$ for all $x \in [0, 2]$.

7:12.2 Show that if f is a polynomial of degree n, then it is its own Taylor polynomial of degree n with $c = 0$.

7:12.3 Calculate the Taylor polynomial of degree 5 with $c = 1$ for the functions below:

(a) x^5

(b) $\ln x$

7:12.4 Let $f(x) = \frac{1}{x+2}, c = -1, n = 2$.

Show that $\frac{1}{x+2} = 1 - (x+1) + (x+1)^2 + R_3$ where $R_3 = -\dfrac{(x+1)^3}{(2+z)^4}$ for some z between x and -1.

7:12.5 Let $f(x) = \ln(1+x)$, $c = 0$, $(x > -1)$. Show that

$$f(x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots + (-1)^{n-1}\frac{x^n}{n} + R_n$$

where $R_n = \dfrac{(-1)^n}{n+1}\left(\dfrac{x}{1+z}\right)^{n+1}$ (z between 0 and x). Estimate R_n on the interval $[0, .1]$.

7:12.6 Just because a function possesses derivatives of all orders on an interval I does not guarantee that some Taylor polynomial approximates f in a neighborhood of some point of I. Let

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

(a) Show that f has derivatives of all orders and that $f^{(k)}(0) = 0$ for $k = 0, 1, 2, \ldots$.

(b) Write down the polynomial P_n with $c = 0$.

(c) Write down Lagrange's form for the remainder of order n. Observe its magnitude and take the time to understand why P_n is not a good approximation for f on any interval I, no matter how large n is.

7.13 Additional Problems for Chapter 7

7:13.1 (Straddled derivatives) Let $f:\mathbb{R}\to\mathbb{R}$ and let $x_0 \in \mathbb{R}$. Prove f is differentiable at x_0 if and only if

$$\lim_{u\to x_0-,\, v\to x_0+} \frac{f(v)-f(u)}{v-u}$$

exists (finite), and in this case, $f'(x_0)$ equals this limit.

7:13.2 (Unstraddled derivatives) Let $f:\mathbb{R}\to\mathbb{R}$ and let $x_0 \in \mathbb{R}$. We say f is strongly differentiable at x_0 if

$$\lim_{u\to x_0,\, v\to x_0,\, u\neq v} \frac{f(v)-f(u)}{v-u}$$

exists.

(a) Show that a differentiable function need not be strongly differentiable everywhere.

(b) Show that a strongly differentiable function must be differentiable.

(c) If f is strongly differentiable at a point x_0 show that f' must be continuous there.

7:13.3 Let p be a polynomial of the nth degree that is everywhere nonnegative. Show that

$$p(x) + p'(x) + p''(x) + \cdots + p^{(n)}(x) \geq 0$$

for all x.

7:13.4 Suppose that f is continuous on $[0,1]$ and differentiable on $(0,1)$ with $f(0) = 0$ and $f(1) = 1$. For every integer n show that there must exist n distinct points $\xi_1, \xi_2, \ldots, \xi_n$ in that interval so that

$$\sum_{k=1}^{n} \frac{1}{f'(\xi_k)} = n$$

7:13.5 Show that there exists precisely one real number α with the property that for every function f differentiable on $[0,1]$ and satisfying $f(0) = 0$ and $f(1) = 1$ there exists a number ξ in $(0,1)$ (which depends, in general, on f) so that

$$f'(\xi) = \alpha\xi.$$

7:13.6 Let f be a continuous function. Show that the set of points where f is differentiable but not strongly differentiable (as defined in Exercise 7:13.2) is of the first category. ✂

7:13.7 Let f be a continuous function on an open interval (a,b). Show that f is convex if and only if

$$\limsup_{h\to 0} \frac{f(x+h)+f(x-h)-2f(x)}{h^2} \geq 0$$

for all $x \in (a, b)$¿

7:13.8 Let f be continuous on an interval (a, b).

(a) Prove that the four Dini derivates of f and the difference quotient $\frac{f(y)-f(x)}{y-x}$ $(x \neq y \in (a, b))$ have the same bounds.

(b) Prove that if one of the Dini derivates is continuous at a point x_0, then f is differentiable at x_0.

(c) Show by example that the statements in the first two parts can fail for discontinuous functions.

7:13.9 (Wronskians) The Wronskian of two differentiable functions f and g is the determinant

$$W(f, g) = \begin{vmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{vmatrix}.$$

Prove that if $W(f, g)$ does not vanish on an interval I and $f(x_1) = f(x_2) = 0$ for points $x_1 < x_2$ in I, then there exists $x_3 \in (x_1, x_2)$ such that $g(x_3) = 0$. (The functions $f(x) = \sin x, g(x) = \cos x$ furnish an example.)

7:13.10 (Denjoy-Young-Saks Theorem) The theorem with this name is a far reaching theorem relating the four Dini Deruvates D^+f, D_+f, D^-f and D_-f. It was proved independently by Grace Chisolm Young and Arnaud Denjoy for continuous functions in 1916 and 1915, respecively. Young then extended the result to a larger class of functions called measurable functions. Finally S. Saks proved the theorem for all finite functions in 1924. Here is their theorem.

> **Theorem (Denjoy-Young-Saks)** Let f be an arbitrary finite function defined on $[a, b]$. Then except for a set of measure zero every point $x \in [a, b]$ is in one of four sets:
>
> (1) A_1 on which f has a finite derivative.
>
> (2) A_2 on which $D^+f = D_-f$ (finite), $D^-f = \infty$ and $D_+f = -\infty$.
>
> (3) A_3 on which $D^-f = D_+f$ (finite), $D^+f = \infty$ and $D_-f = -\infty$.
>
> (4) A_4 on which $D^-f = D^+f = \infty$ and $D_-f = D_+f = -\infty$.

(a) Sketch a picture illustrating points in the sets A_2, A_3 and A_4. To which set does $x = 0$ belong when $f(x) = \sqrt{|x|} \sin x^{-1}$, $f(0) = 0$.

(b) Use DYS to prove that an increasing function f has a finite derivative except on a set of measure zero.

(c) Use DYS to show that if all derived numbers of f are finite except on a set of measure zero, then f is differentiable except on a set of measure zero.

(d) Use DYS to show that, for every finite function f, the set $\{x : f'(x) = \infty\}$ has measure zero.

7:13.11 Let f be a continuous function on an interval $[a, b]$ with a second derivative at all points in (a, b). Let $a < x < b$. Show that there exists a point $\xi \in (a, b)$ so that

$$\frac{\frac{f(x)-f(a)}{x-a} - \frac{f(b)-f(a)}{b-a}}{x-b} = \tfrac{1}{2}f''(\xi).$$

7:13.12 Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function such that $f(0) = 0$ and $|f'(x)| < |f(x)|$ for all $x \in \mathbb{R}$. Show that f is identically zero.

7:13.13 Let $f : \mathbb{R} \to \mathbb{R}$ have a third derivative that exists at all points. Suppose that

$$\lim_{x\to\infty} f(x)$$

exists and that

$$\lim_{x\to\infty} f'''(x) = 0.$$

Show that

$$\lim_{x\to\infty} f'(x) = \lim_{x\to\infty} f''(x) = 0.$$

7:13.14 Let f be defined on an interval I of length at least 2 and suppose that f'' exists there. If $|f(x)| \leq 1$ and $|f''(x)| \leq 1$ for all $x \in I$ show that $|f'(x)| \leq 2$ on the interval.

7:13.15 Let $f : \mathbb{R} \to \mathbb{R}$ be infinitely differentiable and suppose that

$$f\left(\frac{1}{n}\right) = \frac{n^2}{n^2+1}$$

for all $n = 1, 2, 3, \ldots$. Determine the values of $f'(0)$, $f''(0)$, $f'''(0)$, $f^{(4)}(0)$, $\ldots$.

7:13.16 Let $f : \mathbb{R} \to \mathbb{R}$ have a third derivative that exists at all points. Show that there must exist at least one point ξ for which

$$f(\xi)f'(\xi)f''(\xi)f'''(\xi) \geq 0.$$