CHAPTER 11 SIGNIFICANCE OF THE DERIVATIVE

One aim in this chapter is to justify the time we have spent learning to find the derivative of a function. As we shall see, knowing just a little about f' tells us a lot about f. Extracting information about f from information about f' requires some difficult work, however, and we shall begin with the one theorem which is really easy.

This theorem is concerned with the maximum value of a function on an interval. Although we have used this term informally in Chapter 7, it is worthwhile to be precise, and also more general.

DEFINITION

Let f be a function and A a set of numbers contained in the domain of f. A point x in A is a **maximum point** for f on A if

$$f(x) \geq f(y) \quad \text{for every } y \text{ in } A.$$

The number $f(x)$ itself is called the **maximum value** of f on A (and we also say that f "has its maximum value on A at x").

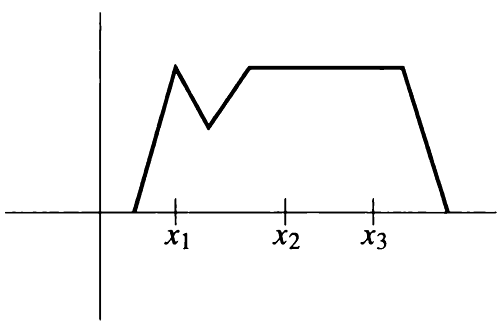


FIGURE 1

Notice that the maximum value of f on A could be $f(x)$ for several different x (Figure 1); in other words, a function f can have several different maximum points on A, although it can have at most one maximum value. Usually we shall be interested in the case where A is a closed interval $[a, b]$; if f is continuous, then Theorem 7-3 guarantees that f does indeed have a maximum value on $[a, b]$.

The definition of a minimum of f on A will be left to you. (One possible definition is the following: f has a minimum on A at x, if $-f$ has a maximum on A at x.)

We are now ready for a theorem which does not even depend upon the existence of least upper bounds.

THEOREM 1 Let f be any function defined on (a, b). If x is a maximum (or a minimum) point for f on (a, b), and f is differentiable at x, then $f'(x) = 0$.
(Notice that we do not assume differentiability, or even continuity, of f at other points.)

PROOF Consider the case where f has a maximum at x. Figure 2 illustrates the simple idea behind the whole argument—secants drawn through points to the left of $(x, f(x))$ have slopes ≥ 0, and secants drawn through points to the right of $(x, f(x))$ have slopes ≤ 0. Analytically, this argument proceeds as follows.

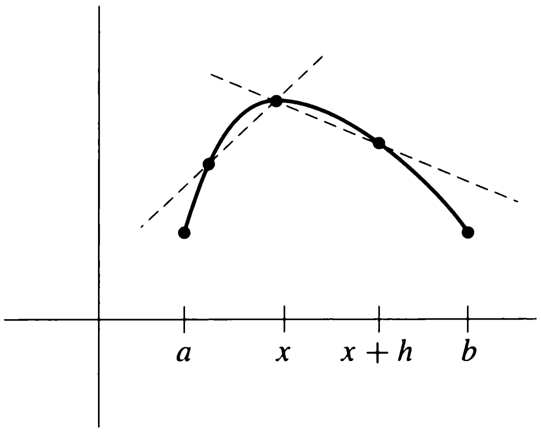


FIGURE 2

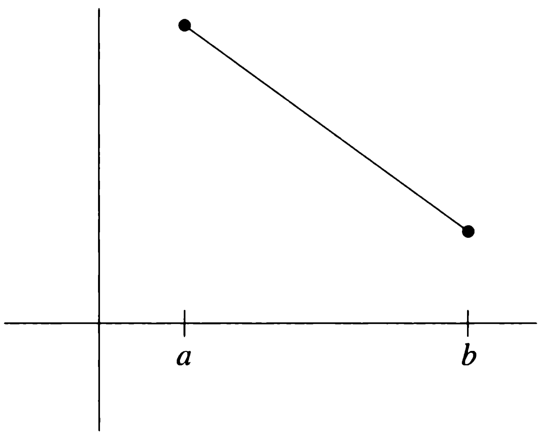


FIGURE 3

If h is any number such that $x + h$ is in (a, b), then

$$f(x) \geq f(x + h),$$

since f has a maximum on (a, b) at x. This means that

$$f(x + h) - f(x) \leq 0.$$

Thus, if $h > 0$ we have

$$\frac{f(x + h) - f(x)}{h} \leq 0.$$

and consequently

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

On the other hand, if $h < 0$, we have

$$\frac{f(x + h) - f(x)}{h} \geq 0,$$

so

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

By hypothesis, f is differentiable at x, so these two limits must be equal, in fact equal to $f'(x)$. This means that

$$f'(x) \leq 0 \quad \text{and} \quad f'(x) \geq 0,$$

from which it follows that $f'(x) = 0$.

The case where f has a minimum at x is left to you (give a one-line proof). ▮

Notice (Figure 3) that we cannot replace (a, b) by $[a, b]$ in the statement of the theorem (unless we add to the hypothesis the condition that x is in (a, b).)

Since $f'(x)$ depends only on the values of f near x, it is almost obvious how to get a stronger version of Theorem 1. We begin with a definition which is illustrated in Figure 4.

DEFINITION

> Let f be a function, and A a set of numbers contained in the domain of f. A point x in A is a **local maximum [minimum] point** for f on A if there is some $\delta > 0$ such that x is a maximum [minimum] point for f on $A \cap (x - \delta, x + \delta)$.

THEOREM 2 If x is a local maximum or minimum for f on (a, b) and f is differentiable at x, then $f'(x) = 0$.

PROOF You should see why this is an easy application of Theorem 1. ▮

The converse of Theorem 2 is definitely not true—it is possible for $f'(x)$ to be 0 even if x is not a local maximum or minimum point for f. The simplest example is provided by the function $f(x) = x^3$; in this case $f'(0) = 0$, but f has no local maximum or minimum anywhere.

Probably the most widespread misconceptions about calculus are concerned with the behavior of a function f near x when $f'(x) = 0$. The point made in the previous paragraph is so quickly forgotten by those who want the world to be simpler than it is, that we will repeat it: the converse of Theorem 2 is *not* true—the condition $f'(x) = 0$ does *not* imply that x is a local maximum or minimum point of f. Precisely for this reason, special terminology has been adopted to describe numbers x which satisfy the condition $f'(x) = 0$.

DEFINITION

A **critical point** of a function f is a number x such that

$$f'(x) = 0.$$

The number $f(x)$ itself is called a **critical value** of f.

The critical values of f, together with a few other numbers, turn out to be the ones which must be considered in order to find the maximum and minimum of a given function f. To the uninitiated, finding the maximum and minimum value of a function represents one of the most intriguing aspects of calculus, and there is no denying that problems of this sort are fun (until you have done your first hundred or so).

Let us consider first the problem of finding the maximum or minimum of f on a closed interval $[a, b]$. (Then, if f is continuous, we can at least be sure that a maximum and minimum value exist.) In order to locate the maximum and minimum of f three kinds of points must be considered:

(1) The critical points of f in $[a, b]$.
(2) The end points a and b.
(3) Points x in $[a, b]$ such that f is not differentiable at x.

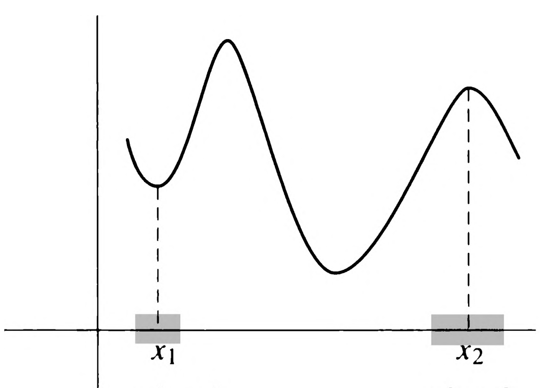


FIGURE 4

If x is a maximum point or a minimum point for f on $[a, b]$, then x must be in one of the three classes listed above: for if x is not in the second or third group, then x *is* in (a, b) and f *is* differentiable at x; consequently $f'(x) = 0$, by Theorem 1, and this means that x is in the first group.

If there are many points in these three categories, finding the maximum and minimum of f may still be a hopeless proposition, but when there are only a few critical points, and only a few points where f is not differentiable, the procedure is fairly straightforward: one simply finds $f(x)$ for each x satisfying $f'(x) = 0$, and $f(x)$ for each x such that f is not differentiable at x and, finally, $f(a)$ and $f(b)$. The biggest of these will be the maximum value of f, and the smallest will be the minimum. A simple example follows.

Suppose we wish to find the maximum and minimum value of the function

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

on the interval $[-1, 2]$. To begin with, we have

$$f'(x) = 3x^2 - 1,$$

so $f'(x) = 0$ when $3x^2 - 1 = 0$, that is, when

$$x = \sqrt{1/3} \quad \text{or} \quad -\sqrt{1/3}.$$

The numbers $\sqrt{1/3}$ and $-\sqrt{1/3}$ both lie in $[-1, 2]$, so the first group of candidates for the location of the maximum and the minimum is

$$(1) \quad \sqrt{1/3}, \ -\sqrt{1/3}.$$

The second group contains the end points of the interval,

$$(2) \quad -1, \ 2.$$

The third group is empty, since f is differentiable everywhere. The final step is to compute

$$\begin{aligned} f(\sqrt{1/3}) &= (\sqrt{1/3})^3 - \sqrt{1/3} = \tfrac{1}{3}\sqrt{1/3} - \sqrt{1/3} = -\tfrac{2}{3}\sqrt{1/3}, \\ f(-\sqrt{1/3}) &= (-\sqrt{1/3})^3 - (-\sqrt{1/3}) = -\tfrac{1}{3}\sqrt{1/3} + \sqrt{1/3} = \tfrac{2}{3}\sqrt{1/3}, \\ f(-1) &= 0, \\ f(2) &= 6. \end{aligned}$$

Clearly the minimum value is $-\frac{2}{3}\sqrt{1/3}$, occurring at $\sqrt{1/3}$, and the maximum value is 6, occurring at 2.

This sort of procedure, if feasible, will always locate the maximum and minimum value of a continuous function on a closed interval. If the function we are dealing with is not continuous, however, or if we are seeking the maximum or minimum on an open interval or the whole line, then we cannot even be sure beforehand that the maximum and minimum values exist, so all the information obtained by this procedure may say nothing. Nevertheless, a little ingenuity will often reveal the nature of things. In Chapter 7 we solved just such a problem when we showed that if n is even, then the function

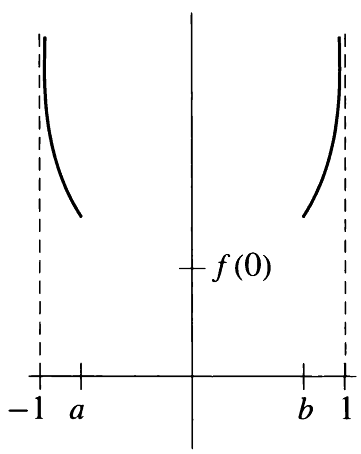


FIGURE 5

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$$

has a minimum value on the whole line. This proves that the minimum value must occur at some number x satisfying

$$0 = f'(x) = nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.$$

If we can solve this equation, and compare the values of $f(x)$ for such x, we can actually find the minimum of f. One more example may be helpful. Suppose we wish to find the maximum and minimum, if they exist, of the function

$$f(x) = \frac{1}{1 - x^2}$$

on the open interval $(-1, 1)$. We have

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

so $f'(x) = 0$ only for $x = 0$. We can see immediately that for x close to 1 or -1 the values of $f(x)$ become arbitrarily large, so f certainly does not have a maximum. This observation also makes it easy to show that f has a minimum at 0. We just note (Figure 5) that there will be numbers a and b, with

$$-1 < a < 0 \quad \text{and} \quad 0 < b < 1,$$

such that $f(x) > f(0)$ for

$$-1 < x \leq a \quad \text{and} \quad b \leq x < 1.$$

This means that the minimum of f on $[a, b]$ is the minimum of f on all of $(-1, 1)$. Now on $[a, b]$ the minimum occurs either at 0 (the only place where $f' = 0$), or at a or b, and a and b have already been ruled out, so the minimum value is $f(0) = 1$.

In solving these problems we purposely did not draw the graphs of $f(x) = x^3 - x$ and $f(x) = 1/(1 - x^2)$, but it is not cheating to draw the graph (Figure 6) as long as you do not rely solely on your picture to prove anything. As a matter of fact, we are now going to discuss a method of sketching the graph of a function that really gives enough information to be used in discussing maxima and minima—in fact we will be able to locate even *local* maxima and minima. This method involves consideration of the sign of $f'(x)$, and relies on some deep theorems.

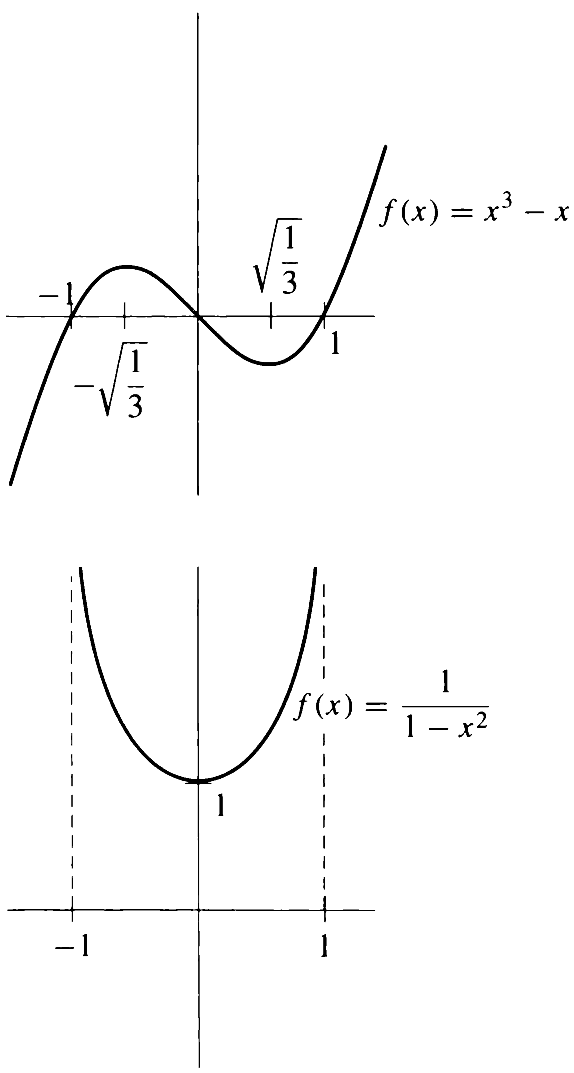


FIGURE 6

The theorems about derivatives which have been proved so far, always yield information about f' in terms of information about f. This is true even of Theorem 1, although this theorem can sometimes be used to determine certain information about f, namely, the location of maxima and minima. When the derivative was first introduced, we emphasized that $f'(x)$ is not $[f(x+h) - f(x)]/h$ for any particular h, but only a limit of these numbers as h approaches 0; this fact becomes painfully relevant when one tries to extract information about f from information about f'. The simplest and most frustrating illustration of the difficulties encountered is afforded by the following question: If $f'(x) = 0$ for all x, must f be a constant function? It is impossible to imagine how f could be anything else, and this conviction is strengthened by considering the physical interpretation—if the velocity of a particle is always 0, surely the particle must be standing still! Nevertheless it is difficult even to begin a proof that only the constant functions satisfy $f'(x) = 0$ for all x. The hypothesis $f'(x) = 0$ only means that

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

and it is not at all obvious how one can use the information about the limit to derive information about the function.

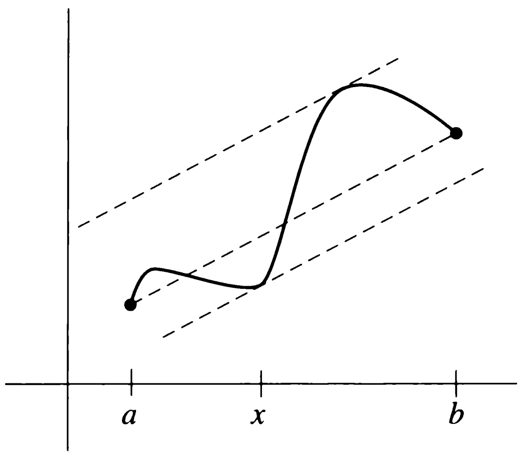


FIGURE 7

The fact that f is a constant function if $f'(x) = 0$ for all x, and many other facts of the same sort, can all be derived from a fundamental theorem, called the Mean Value Theorem, which states much stronger results. Figure 7 makes it plausible that if f is differentiable on $[a, b]$, then there is some x in (a, b) such that

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

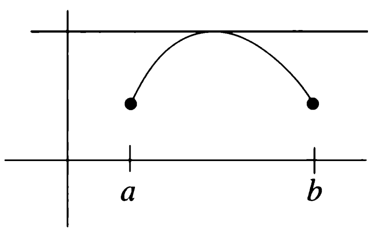


FIGURE 8

Geometrically this means that some tangent line is parallel to the line between $(a, f(a))$ and $(b, f(b))$. The Mean Value Theorem asserts that this is true—there is some x in (a, b) such that $f'(x)$, the instantaneous rate of change of f at x, is exactly equal to the average or "mean" change of f on $[a, b]$, this average change being $[f(b) - f(a)]/[b - a]$. (For example, if you travel 60 miles in one hour, then at some time you must have been traveling exactly 60 miles per hour.) This theorem is one of the most important theoretical tools of calculus—probably the deepest result about derivatives. From this statement you might conclude that the proof is difficult, but there you would be wrong—the hard theorems in this book have occurred long ago, in Chapter 7. It is true that if you try to prove the Mean Value Theorem yourself you will probably fail, but this is neither evidence that the theorem is hard, nor something to be ashamed of. The first proof of the theorem was an achievement, but today we can supply a proof which is quite simple. It helps to begin with a very special case.

THEOREM 3 (ROLLE'S THEOREM) If f is continuous on $[a, b]$ and differentiable on (a, b), and $f(a) = f(b)$, then there is a number x in (a, b) such that $f'(x) = 0$.

PROOF If follows from the continuity of f on $[a, b]$ that f has a maximum and a minimum value on $[a, b]$.

Suppose first that the maximum value occurs at a point x in (a, b). Then $f'(x) = 0$ by Theorem 1, and we are done (Figure 8).

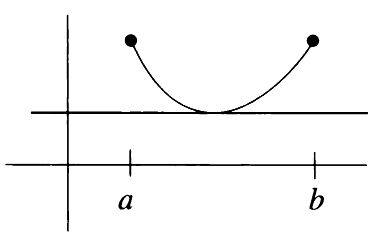


FIGURE 9

Suppose next that the minimum value of f occurs at some point x in (a, b). Then, again, $f'(x) = 0$ by Theorem 1 (Figure 9).

Finally, suppose the maximum and minimum values both occur at the end points. Since $f(a) = f(b)$, the maximum and minimum values of f are equal, so f is a constant function (Figure 10), and for a constant function we can choose any x in (a, b). ▮

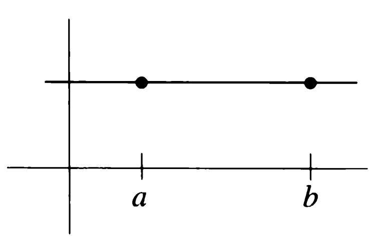


FIGURE 10

Notice that we really needed the hypothesis that f is differentiable everywhere on (a, b) in order to apply Theorem 1. Without this assumption the theorem is false (Figure 11).

You may wonder why a special name should be attached to a theorem as easily proved as Rolle's Theorem. The reason is, that although Rolle's Theorem is a special case of the Mean Value Theorem, it also yields a simple proof of the Mean Value Theorem. In order to prove the Mean Value Theorem we will apply Rolle's Theorem to the function which gives the length of the vertical segment shown in

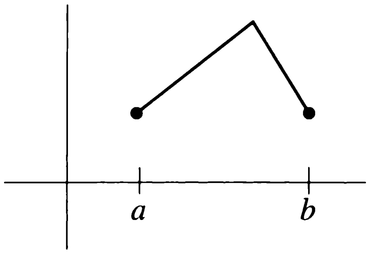


FIGURE 11

Figure 12; this is the difference between $f(x)$, and the height at x of the line L between $(a, f(a))$ and $(b, f(b))$. Since L is the graph of

$$g(x) = \left[\frac{f(b) - f(a)}{b - a}\right](x - a) + f(a),$$

we want to look at

$$f(x) - \left[\frac{f(b) - f(a)}{b - a}\right](x - a) - f(a).$$

As it turns out, the constant $f(a)$ is irrelevant.

THEOREM 4 (THE MEAN VALUE THEOREM) If f is continuous on $[a, b]$ and differentiable on (a, b), then there is a number x in (a, b) such that

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

PROOF Let

$$h(x) = f(x) - \left[\frac{f(b) - f(a)}{b - a}\right](x - a).$$

Clearly, h is continuous on $[a, b]$ and differentiable on (a, b), and

$$\begin{aligned} h(a) &= f(a), \\ h(b) &= f(b) - \left[\frac{f(b) - f(a)}{b - a}\right](x - a) \\ &= f(a). \end{aligned}$$

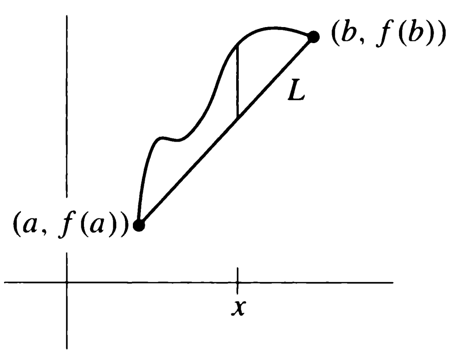


FIGURE 12

Consequently, we may apply Rolle's Theorem to h and conclude that there is some x in (a, b) such that

$$0 = h'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

so that

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

Notice that the Mean Value Theorem still fits into the pattern exhibited by previous theorems—information about f yields information about f'. This information is so strong, however, that we can now go in the other direction.

COROLLARY 1 If f is defined on an interval and $f'(x) = 0$ for all x in the interval, then f is constant on the interval.

PROOF Let a and b be any two points in the interval with $a \neq b$. Then there is some x in

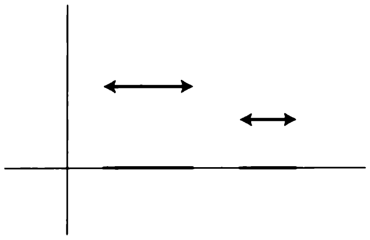


FIGURE 13

(a, b) such that

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

But $f'(x) = 0$ for all x in the interval, so

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

and consequently $f(a) = f(b)$. Thus the value of f at any two points in the interval is the same, i.e., f is constant on the interval. ▮

Naturally, Corollary 1 does not hold for functions defined on two or more intervals (Figure 13).

COROLLARY 2 If f and g are defined on the same interval, and $f'(x) = g'(x)$ for all x in the interval, then there is some number c such that $f = g + c$.

PROOF For all x in the interval we have $(f - g)'(x) = f'(x) - g'(x) = 0$ so, by Corollary 1, there is a number c such that $f - g = c$. ▮

The statement of the next corollary requires some terminology, which is illustrated in Figure 14.

DEFINITION

A function is **increasing** on an interval if $f(a) < f(b)$ whenever a and b are two numbers in the interval with $a < b$. The function f is **decreasing** on an interval if $f(a) > f(b)$ for all a and b in the interval with $a < b$. (We often say simply that f is increasing or decreasing, in which case the interval is understood to be the domain of f.)

COROLLARY 3 If $f'(x) > 0$ for all x in an interval, then f is increasing on the interval; if $f'(x) < 0$ for all x in the interval, then f is decreasing on the interval.

PROOF Consider the case where $f'(x) > 0$. Let a and b be two points in the interval with $a < b$. Then there is some x in (a, b) with

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

But $f'(x) > 0$ for all x in (a, b), so

$$\frac{f(b) - f(a)}{b - a} > 0.$$

Since $b - a > 0$ it follows that $f(b) > f(a)$.

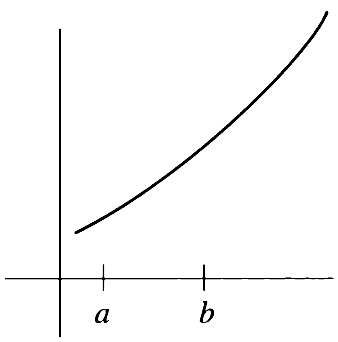
The proof when $f'(x) < 0$ for all x is left to you. ▮

Notice that although the converses of Corollary 1 and Corollary 2 are true (and obvious), the converse of Corollary 3 is not true. If f is increasing, it is easy to see that $f'(x) \geq 0$ for all x, but the equality sign might hold for some x (consider $f(x) = x^3$).

Corollary 3 provides enough information to get a good idea of the graph of a function with a minimal amount of point plotting. Consider, once more, the function $f(x) = x^3 - x$. We have

$$f'(x) = 3x^2 - 1.$$

We have already noted that $f'(x) = 0$ for $x = \sqrt{1/3}$ and $x = -\sqrt{1/3}$, and it is also possible to determine the sign of $f'(x)$ for all other x. Note that $3x^2 - 1 > 0$ precisely when

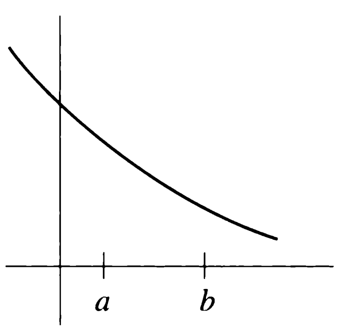


(a) an increasing function

$$\begin{aligned} 3x^2 &> 1 \\ x^2 &> \tfrac{1}{3}, \\ x > \sqrt{1/3} \quad \text{or} \quad x &< -\sqrt{1/3}; \end{aligned}$$

thus $3x^2 - 1 < 0$ precisely when

$$-\sqrt{1/3} < x < \sqrt{1/3}.$$



(b) a decreasing function

FIGURE 14

Thus f is increasing for $x < -\sqrt{1/3}$, decreasing between $-\sqrt{1/3}$ and $\sqrt{1/3}$, and once again increasing for $x > \sqrt{1/3}$. Combining this information with the following facts

(1) $f(-\sqrt{1/3}) = \frac{2}{3}\sqrt{1/3}$,
$\quad f(\sqrt{1/3}) = -\frac{2}{3}\sqrt{1/3}$,
(2) $f(x) = 0$ for $x = -1, 0, 1$,
(3) $f(x)$ gets large as x gets large, and large negative as x gets large negative,

it is possible to sketch a pretty respectable approximation to the graph (Figure 15).

By the way, notice that the intervals on which f increases and decreases could have been found without even bothering to examine the sign of f'. For example, since f' is continuous, and vanishes only at $-\sqrt{1/3}$ and $\sqrt{1/3}$, we know that f' always has the same sign on the interval $(-\sqrt{1/3}, \sqrt{1/3})$. Since $f(-\sqrt{1/3}) > f(\sqrt{1/3})$, it follows that f decreases on this interval. Similarly, f' always has the same sign on $(\sqrt{1/3}, \infty)$ and $f(x)$ is large for large x, so f must be increasing on $(\sqrt{1/3}, \infty)$. Another point worth noting: If f' is continuous, then the sign of f' on the interval between two adjacent critical points can be determined simply by finding the sign of $f'(x)$ for any *one* x in this interval.

Our sketch of the graph of $f(x) = x^3 - x$ contains sufficient information to allow us to say with confidence that $-\sqrt{1/3}$ is a local maximum point, and that $\sqrt{1/3}$ is a local minimum point. In fact, we can give a general scheme for decid-

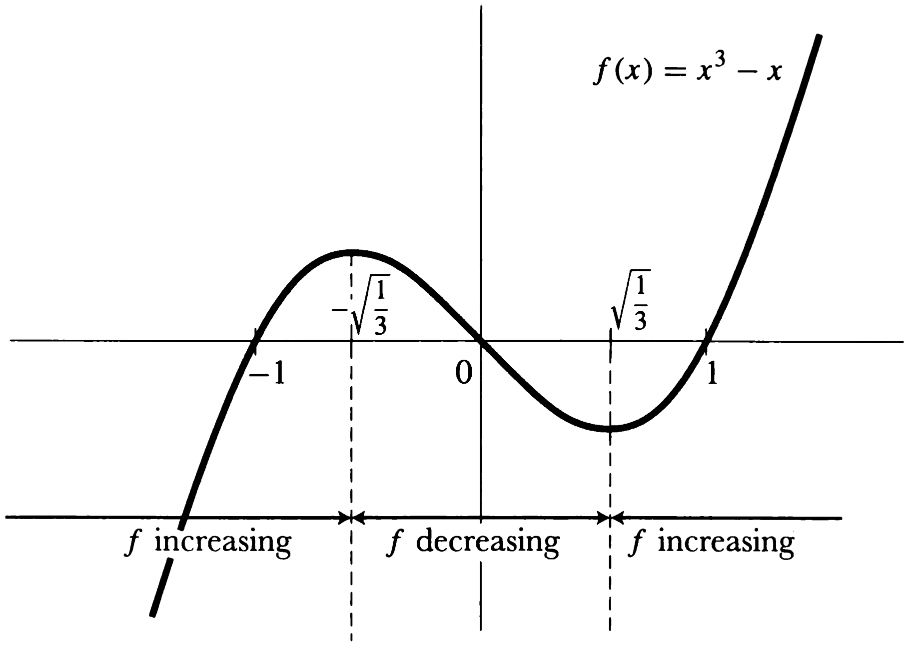


FIGURE 15

ing whether a critical point is a local maximum point, a local minimum point, or neither (Figure 16):

(1) if $f' > 0$ in some interval to the left of x and $f' < 0$ in some interval to the right of x, then x is a local maximum point.
(2) if $f' < 0$ in some interval to the left of x and $f' > 0$ in some interval to the right of x, then x is a local minimum point.
(3) if f' has the same sign in some interval to the left of x as it has in some interval to the right, then x is neither a local maximum nor a local minimum point.

(There is no point in memorizing these rules—you can always draw the pictures yourself.)

The polynomial functions can all be analyzed in this way, and it is even possible to describe the general form of the graph of such functions. To begin, we need a

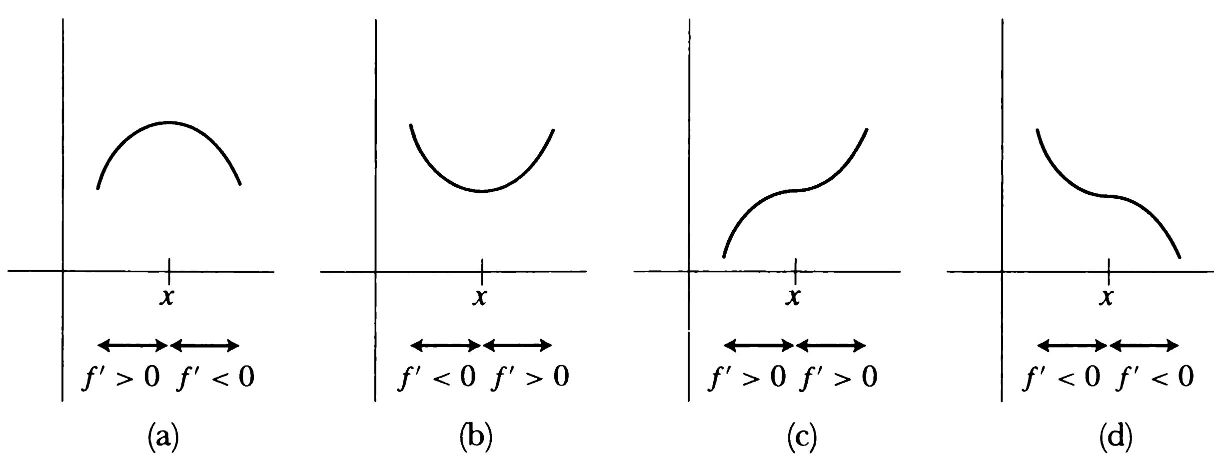


FIGURE 16

result already mentioned in Problem 3-7: If

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0,$$

then f has at most n "roots," i.e., there are at most n numbers x such that $f(x) = 0$. Although this is really an algebraic theorem, calculus can be used to give an easy proof. Notice that if x_1 and x_2 are roots of f (Figure 17), so that $f(x_1) = f(x_2) = 0$, then by Rolle's Theorem there is a number x between x_1 and x_2 such that $f'(x) = 0$. This means that if f has k different roots $x_1 < x_2 < \cdots < x_k$, then f' has at least $k - 1$ different roots: one between x_1 and x_2, one between x_2 and x_3, etc. It is now easy to prove by induction that a polynomial function

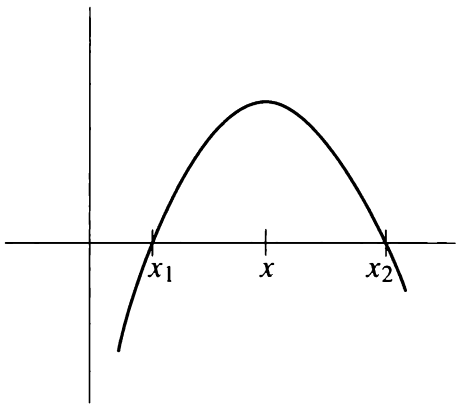


FIGURE 17

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

has at most n roots: The statement is surely true for $n = 1$, and if we assume that it is true for n, then the polynomial

$$g(x) = b_{n+1}x^{n+1} + b_n x^n + \cdots + b_0$$

could not have more than $n + 1$ roots, since if it did, g' would have more than n roots.

With this information it is not hard to describe the graph of

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

The derivative, being a polynomial function of degree $n - 1$, has at most $n - 1$ roots. Therefore f has at most $n - 1$ critical points. Of course, a critical point is not necessarily a local maximum or minimum point, but at any rate, if a and b are adjacent critical points of f, then f' will remain either positive or negative on (a, b), since f' is continuous; consequently, f will be either increasing or decreasing on (a, b). Thus f has at most n regions of decrease or increase.

As a specific example, consider the function

$$f(x) = x^4 - 2x^2.$$

Since

$$f'(x) = 4x^3 - 4x = 4x(x - 1)(x + 1),$$

the critical points of f are -1, 0, and 1, and

$$\begin{aligned} f(-1) &= -1, \\ f(0) &= 0, \\ f(1) &= -1. \end{aligned}$$

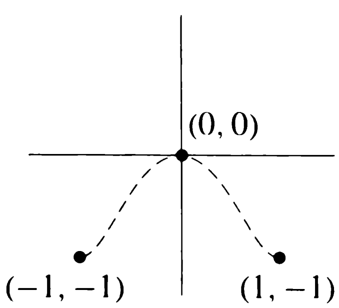


FIGURE 18

The behavior of f on the intervals between the critical points can be determined by one of the methods mentioned before. In particular, we could determine the sign of f' on these intervals simply be examining the formula for $f'(x)$. On the other hand, from the three critical values alone we can see (Figure 18) that f increases on $(-1, 0)$ and decreases on $(0, 1)$. To determine the sign of f' on

$(-\infty, -1)$ and $(1, \infty)$ we can compute

$$f'(-2) = 4 \cdot (-2)^3 - 4 \cdot (-2) = -24,$$
$$f'(2) = 4 \cdot 2^3 - 4 \cdot 2 = 24,$$

and conclude that f is decreasing on $(-\infty, -1)$ and increasing on $(1, \infty)$. These conclusions also follow from the fact that $f(x)$ is large for large x and for large negative x.

We can already produce a good sketch of the graph; two other pieces of information provide the finishing touches (Figure 19). First, it is easy to determine that $f(x) = 0$ for $x = 0, \pm\sqrt{2}$; second, it is clear that f is even, $f(x) = f(-x)$, so the graph is symmetric with respect to the vertical axis. The function $f(x) = x^3 - x$, already sketched in Figure 15, is odd, $f(x) = -f(-x)$, and is consequently symmetric with respect to the origin. Half the work of graph sketching may be saved by noticing these things in the beginning.

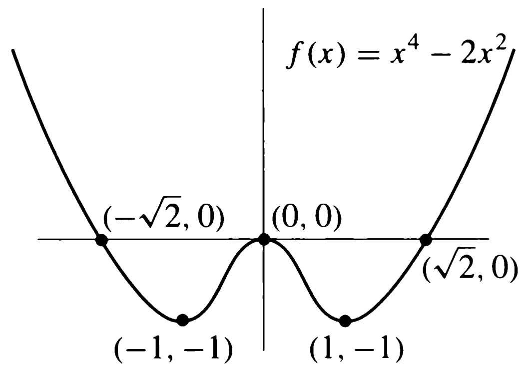


FIGURE 19

Several problems in this and succeeding chapters ask you to sketch the graphs of functions. In each case you should determine

(1) the critical points of f,
(2) the value of f at the critical points,
(3) the sign of f' in the regions between critical points (if this is not already clear),
(4) the numbers x such that $f(x) = 0$ (if possible),
(5) the behavior of $f(x)$ as x becomes large or large negative (if possible).

Finally, bear in mind that a quick check, to see whether the function is odd or even, may save a lot of work.

This sort of analysis, if performed with care, will usually reveal the basic shape of the graph, but sometimes there are special features which require a little more thought. It is impossible to anticipate all of these, but one piece of information is often very important. If f is not defined at certain points (for example, if f is a rational function whose denominator vanishes at some points), then the behavior of f near these points should be determined.

For example, consider the function

$$f(x) = \frac{x^2 - 2x + 2}{x - 1},$$

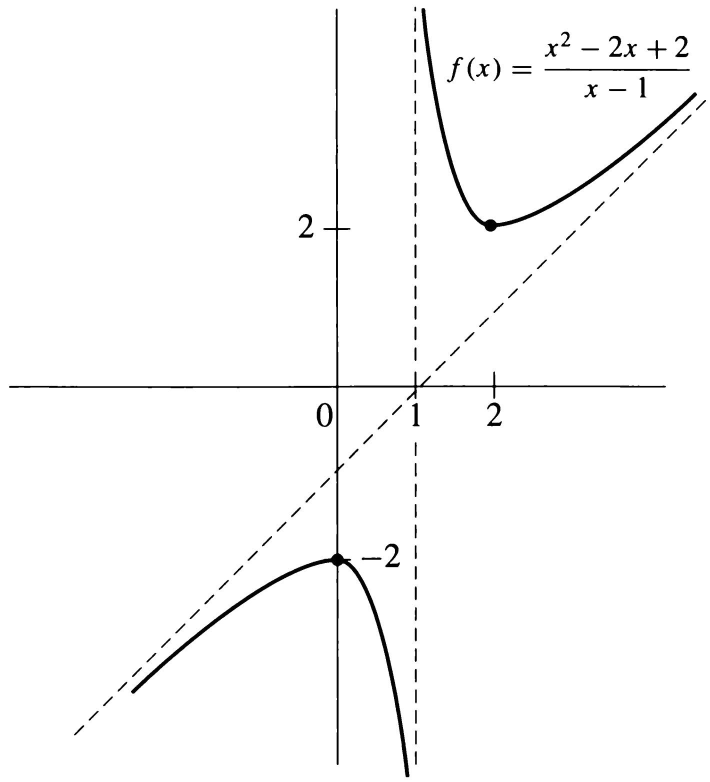


FIGURE 20

of an odd function shoved over 1 unit, and the expression

$$\frac{x^2 - 2x + 2}{x - 1} = \frac{(x-1)^2 + 1}{x - 1}$$

shows that this is indeed the case. However, this is one of those special features which should be investigated only after you have used the other information to get a good idea of the appearance of the graph.

Although the location of local maxima and minima of a function is always revealed by a detailed sketch of its graph, it is usually unnecessary to do so much work. There is a popular test for local maxima and minima which depends on the behavior of the function only at its critical points.

THEOREM 5 Suppose $f'(a) = 0$. If $f''(a) > 0$, then f has a local minimum at a; if $f''(a) < 0$, then f has a local maximum at a.

PROOF By definition,

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

Since $f'(a) = 0$, this can be written

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

which is not defined at 1. We have

$$\begin{aligned} f'(x) &= \frac{(x-1)(2x-2)-(x^2-2x+2)}{(x-1)^2} \\ &= \frac{x(x-2)}{(x-1)^2}. \end{aligned}$$

Thus

(1) the critical points of f are 0, 2.

Moreover,

$$\begin{aligned} (2)\ f(0) &= -2, \\ f(2) &= 2. \end{aligned}$$

Because f is not defined on the whole interval $(0, 2)$, the sign of f' must be determined separately on the intervals $(0, 1)$ and $(1, 2)$, as well as on the intervals $(-\infty, 0)$ and $(2, \infty)$. We can do this by picking particular points in each of these intervals, or simply by staring hard at the formula for f'. Either way we find that

$$\begin{aligned} (3)\ f'(x) &> 0 \quad \text{if} \quad x < 0, \\ f'(x) &< 0 \quad \text{if} \quad 0 < x < 1, \\ f'(x) &< 0 \quad \text{if} \quad 1 < x < 2, \\ f'(x) &> 0 \quad \text{if} \quad 2 < x. \end{aligned}$$

Finally, we must determine the behavior of $f(x)$ as x becomes large or large negative, as well as when x approaches 1 (this information will also give us another way to determine the regions on which f increases and decreases). To examine the behavior as x becomes large we write

$$\frac{x^2-2x+2}{x-1} = x - 1 + \frac{1}{x-1};$$

clearly $f(x)$ is close to $x - 1$ (and slightly larger) when x is large, and $f(x)$ is close to $x - 1$ (but slightly smaller) when x is large negative. The behavior of f near 1 is also easy to determine; since

$$\lim_{x \to 1} (x^2 - 2x + 2) = 1 \neq 0,$$

the fraction

$$\frac{x^2-2x+2}{x-1}$$

becomes large as x approaches 1 from above and large negative as x approaches 1 from below.

All this information may seem a bit overwhelming, but there is only one way that it can be pieced together (Figure 20); be sure that you can account for each feature of the graph.

When this sketch has been completed, we might note that it looks like the graph

Suppose now that $f''(a) > 0$. Then $f'(a+h)/h$ must be positive for sufficiently small h. Therefore:

$$\begin{aligned} &f'(a+h) \text{ must be positive for sufficiently small } h > 0 \\ \text{and } &f'(a+h) \text{ must be negative for sufficiently small } h < 0. \end{aligned}$$

This means (Corollary 3) that f is increasing in some interval to the right of a and f is decreasing in some interval to the left of a. Consequently, f has a local minimum at a.

The proof for the case $f''(a) < 0$ is similar. ▌

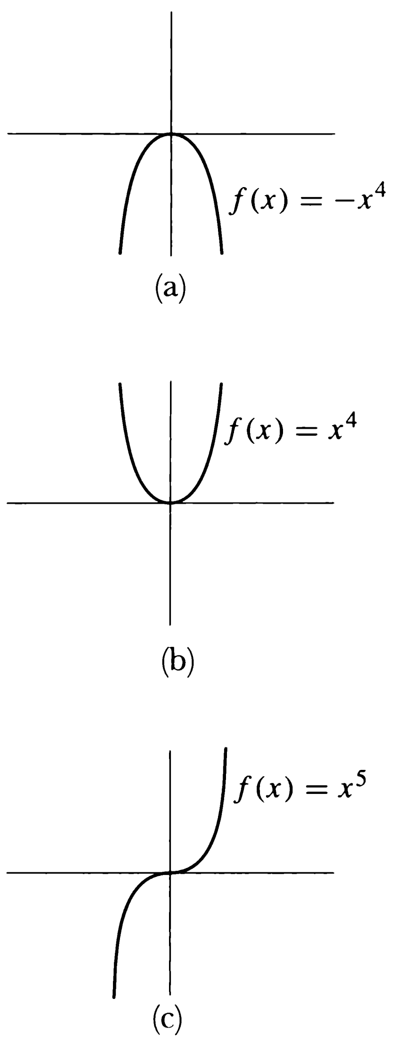


FIGURE 21

Theorem 5 may be applied to the function $f(x) = x^3 - x$, which has already been considered. We have

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

At the critical points, $-\sqrt{1/3}$ and $\sqrt{1/3}$, we have

$$\begin{aligned} f''(-\sqrt{1/3}) &= -6\sqrt{1/3} < 0, \\ f''(\sqrt{1/3}) &= 6\sqrt{1/3} > 0. \end{aligned}$$

Consequently, $-\sqrt{1/3}$ is a local maximum point and $\sqrt{1/3}$ is a local minimum point.

Although Theorem 5 will be found quite useful for polynomial functions, for many functions the second derivative is so complicated that it is easier to consider the sign of the first derivative. Moreover, if a is a critical point of f it may happen that $f''(a) = 0$. In this case, Theorem 5 provides no information: it is possible that a is a local maximum point, a local minimum point, or neither, as shown (Figure 21) by the functions

$$f(x) = -x^4, \quad f(x) = x^4, \quad f(x) = x^5;$$

in each case $f'(0) = f''(0) = 0$, but 0 is a local maximum point for the first, a local minimum point for the second, and neither a local maximum nor minimum point for the third. This point will be pursued further in Part IV.

It is interesting to note that Theorem 5 automatically proves a partial converse of itself.

THEOREM 6 Suppose $f''(a)$ exists. If f has a local minimum at a, then $f''(a) \geq 0$; if f has a local maximum at a, then $f''(a) \leq 0$.

PROOF Suppose f has local minimum at a. If $f''(a) < 0$, then f would also have a local maximum at a, by Theorem 5. Thus f would be constant in some interval containing a, so that $f''(a) = 0$, a contradiction. Thus we must have $f''(a) \geq 0$.

The case of a local maximum is handled similarly. ▌

(This partial converse to Theorem 5 is the best we can hope for: the $\geq$ and $\leq$ signs cannot be replaced by $>$ and $<$, as shown by the functions $f(x) = x^4$ and $f(x) = -x^4$.)

The remainder of this chapter deals, not with graph sketching, or maxima and minima, but with three consequences of the Mean Value Theorem. The first is a simple, but very beautiful, theorem which plays an important role in Chapter 15, and which also sheds light on many examples which have occurred in previous chapters.

THEOREM 7 Suppose that f is continuous at a, and that $f'(x)$ exists for all x in some interval containing a, except perhaps for $x = a$. Suppose, moreover, that $\lim_{x \to a} f'(x)$ exists. Then $f'(a)$ also exists, and

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

PROOF By definition,

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

For sufficiently small $h > 0$ the function f will be continuous on $[a, a+h]$ and differentiable on $(a, a+h)$ (a similar assertion holds for sufficiently small $h < 0$). By the Mean Value Theorem there is a number α_h in $(a, a+h)$ such that

$$\frac{f(a+h) - f(a)}{h} = f'(\alpha_h).$$

Now α_h approaches a as h approaches 0, because α_h is in $(a, a+h)$; since $\lim_{x \to a} f'(x)$ exists, it follows that

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

(It is a good idea to supply a rigorous ε-δ argument for this final step, which we have treated somewhat informally.) ▌

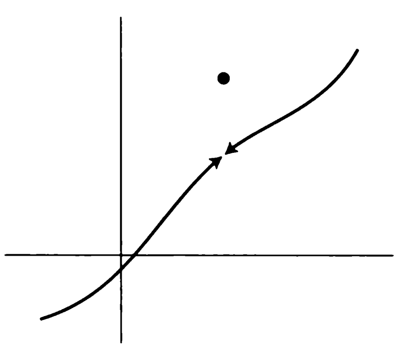


FIGURE 22

Even if f is an everywhere differentiable function, it is still possible for f' to be discontinuous. This happens, for example, if

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

According to Theorem 7, however, the graph of f' can never exhibit a discontinuity of the type shown in Figure 22. Problem 61 outlines the proof of another beautiful theorem which gives further information about the function f', and Problem 62 uses this result to strengthen Theorem 7.

The next theorem, a generalization of the Mean Value Theorem, is of interest mainly because of its applications.

THEOREM 8 (THE CAUCHY MEAN VALUE THEOREM) If f and g are continuous on $[a, b]$ and differentiable on (a, b), then there is a number x in (a, b) such that

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

(If $g(b) \neq g(a)$, and $g'(x) \neq 0$, this equation can be written

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

Notice that if $g(x) = x$ for all x, then $g'(x) = 1$, and we obtain the Mean Value Theorem. On the other hand, applying the Mean Value Theorem to f and g separately, we find that there are x and y in (a, b) with

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x)}{g'(y)};$$

but there is no guarantee that the x and y found in this way will be equal. These remarks may suggest that the Cauchy Mean Value Theorem will be quite difficult to prove, but actually the simplest of tricks suffices.)

PROOF Let

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

Then h is continuous on $[a, b]$, differentiable on (a, b), and

$$h(a) = f(a)g(b) - g(a)f(b) = h(b).$$

It follows from Rolle's Theorem that $h'(x) = 0$ for some x in (a, b), which means that

$$0 = f'(x)[g(b) - g(a)] - g'(x)[f(b) - f(a)]. \blacksquare$$

The Cauchy Mean Value Theorem is the basic tool needed to prove a theorem which facilitates evaluation of limits of the form

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

when

$$\lim_{x \to a} f(x) = 0 \quad \text{and} \quad \lim_{x \to a} g(x) = 0.$$

In this case, Theorem 5-2 is of no use. Every derivative is a limit of this form, and computing derivatives frequently requires a great deal of work. If some derivatives are known, however, many limits of this form can now be evaluated easily.

THEOREM 9 (L'HÔPITAL'S RULE) Suppose that

$$\lim_{x \to a} f(x) = 0 \quad \text{and} \quad \lim_{x \to a} g(x) = 0,$$

and suppose also that $\lim_{x \to a} f'(x)/g'(x)$ exists. Then $\lim_{x \to a} f(x)/g(x)$ exists, and

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

(Notice that Theorem 7 is a special case.)

PROOF The hypothesis that $\lim_{x \to a} f'(x)/g'(x)$ exists contains two implicit assumptions:

(1) there is an interval $(a - \delta, a + \delta)$ such that $f'(x)$ and $g'(x)$ exist for all x in $(a - \delta, a + \delta)$ except, perhaps, for $x = a$,
(2) in this interval $g'(x) \neq 0$ with, once again, the possible exception of $x = a$.

On the other hand, f and g are not even assumed to be defined at a. If we define $f(a) = g(a) = 0$ (changing the previous values of $f(a)$ and $g(a)$, if necessary), then f and g are continuous at a. If $a < x < a + \delta$, then the Mean Value Theorem and the Cauchy Mean Value Theorem apply to f and g on the interval $[a, x]$ (and a similar statement holds for $a - \delta < x < a$). First applying the Mean Value Theorem to g, we see that $g(x) \neq 0$, for if $g(x) = 0$ there would be some x_1 in (a, x) with $g'(x_1) = 0$, contradicting (2). Now applying the Cauchy Mean Value Theorem to f and g, we see that there is a number α_x in (a, x) such that

$$[f(x) - 0]g'(\alpha_x) = [g(x) - 0]f'(\alpha_x)$$

or

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

Now α_x approaches a as x approaches a, because α_x is in (a, x); since we are assuming that $\lim_{y \to a} f'(y)/g'(y)$ exists, it follows that

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

(Once again, the reader is invited to supply the details of this part of the argument.) ▮

PROBLEMS

1. For each of the following functions, find the maximum and minimum values on the indicated intervals, by finding the points in the interval where the derivative is 0, and comparing the values at these points with the values at the end points.

(i) $f(x) = x^3 - x^2 - 8x + 1$ on $[-2, 2]$.
(ii) $f(x) = x^5 + x + 1$ on $[-1, 1]$.
(iii) $f(x) = 3x^4 - 8x^3 + 6x^2$ on $[-\frac{1}{2}, \frac{1}{2}]$.
(iv) $f(x) = \dfrac{1}{x^5 + x + 1}$ on $[-\frac{1}{2}, 1]$.
(v) $f(x) = \dfrac{x + 1}{x^2 + 1}$ on $[-1, \frac{1}{2}]$.
(vi) $f(x) = \dfrac{x}{x^2 - 1}$ on $[0, 5]$.

2. Now sketch the graph of each of the functions in Problem 1, and find all local maximum and minimum points.

3. Sketch the graphs of the following functions.

(i) $f(x) = x + \dfrac{1}{x}$.

(ii) $f(x) = x + \dfrac{3}{x^2}$.

(iii) $f(x) = \dfrac{x^2}{x^2 - 1}$.

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

4. (a) If $a_1 < \cdots < a_n$, find the minimum value of $f(x) = \sum_{i=1}^{n} (x - a_i)^2$.

*(b) Now find the minimum value of $f(x) = \sum_{i=1}^{n} |x - a_i|$. This is a problem where calculus won't help at all: on the intervals between the a_i's the function f is linear, so that the minimum clearly occurs at one of the a_i, and these are precisely the points where f is not differentiable. However, the answer is easy to find if you consider how $f(x)$ changes as you pass from one such interval to another.

*(c) Let $a > 0$. Show that the maximum value of

$$f(x) = \frac{1}{1 + |x|} + \frac{1}{1 + |x - a|}$$

is $(2 + a)/(1 + a)$. (The derivative can be found on each of the intervals $(-\infty, 0)$, $(0, a)$, and (a, ∞) separately.)

5. For each of the following functions, find all local maximum and minimum points.

(i) $f(x) = \begin{cases} x, & x \neq 3, 5, 7, 9 \\ 5, & x = 3 \\ -3, & x = 5 \\ 9, & x = 7 \\ 7, & x = 9. \end{cases}$

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

(iii) $f(x) = \begin{cases} x, & x \text{ rational} \\ 0, & x \text{ irrational.} \end{cases}$

(iv) $f(x) = \begin{cases} 1, & x = 1/n \text{ for some } n \text{ in } \mathbf{N} \\ 0, & \text{otherwise.} \end{cases}$

(v) $f(x) = \begin{cases} 1, & \text{if the decimal expansion of } x \text{ contains a 5} \\ 0, & \text{otherwise.} \end{cases}$

6. Prove the following (which we often use implicitly): If f is increasing on (a, b) and continuous at a and b, then f is increasing on $[a, b]$. In particular, if f is continuous on $[a, b]$ and $f' > 0$ on (a, b), then f is increasing on $[a, b]$.

7. A straight line is drawn from the point $(0, a)$ to the horizontal axis, and then back to $(1, b)$, as in Figure 23. Prove that the total length is shortest when the angles α and β are equal. (Naturally you must bring a function into the picture: express the length in terms of x, where $(x, 0)$ is the point on the horizontal axis. The dashed line in Figure 23 suggests an alternative geometric proof; in either case the problem can be solved without actually finding the point $(x, 0)$.)

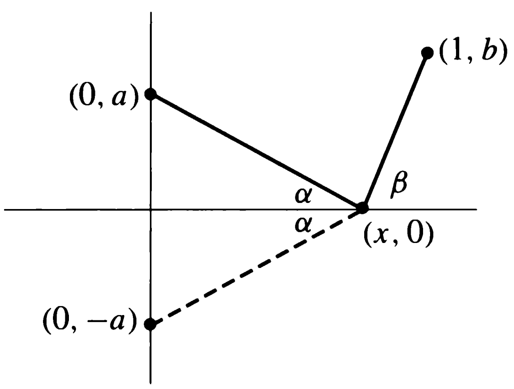


FIGURE 23

8. (a) Let (x_0, y_0) be a point of the plane, and let L be the graph of the function $f(x) = mx + b$. Find the point $\bar{x}$ such that the distance from (x_0, y_0) to $(\bar{x}, f(\bar{x}))$ is smallest. [Notice that minimizing this distance is the same as minimizing its square. This may simplify the computations somewhat.]

(b) Also find $\bar{x}$ by noting that the line from (x_0, y_0) to $(\bar{x}, f(\bar{x}))$ is perpendicular to L.

(c) Find the distance from (x_0, y_0) to L, i.e., the distance from (x_0, y_0) to $(\bar{x}, f(\bar{x}))$. [It will make the computations easier if you first assume that $b = 0$; then apply the result to the graph of $f(x) = mx$ and the point $(x_0, y_0 - b)$.] Compare with Problem 4-22.

(d) Consider a straight line described by the equation $Ax + By + C = 0$ (Problem 4-7). Show that the distance from (x_0, y_0) to this line is $(Ax_0 + By_0 + C)/\sqrt{A^2 + B^2}$.

9. The previous Problem suggests the following question: What is the relationship between the critical points of f and those of f^2?

10. Prove that of all rectangles with given perimeter, the square has the greatest area.

11. Find, among all right circular cylinders of fixed volume V, the one with smallest surface area (counting the areas of the faces at top and bottom, as in Figure 24).

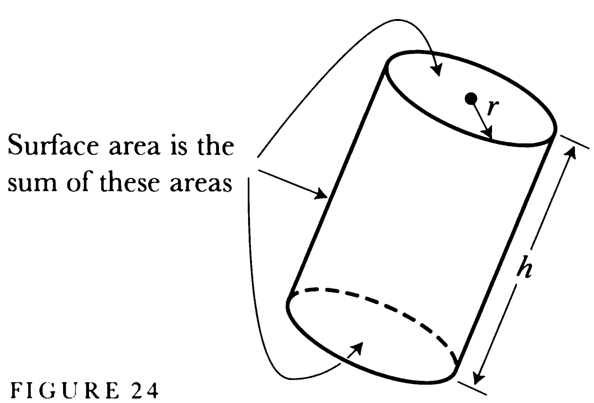


FIGURE 24

12. A right triangle with hypotenuse of length a is rotated about one of its legs to generate a right circular cone. Find the greatest possible volume of such a cone.

13. Show that the sum of a positive number and its reciprocal is at least 2.

14. Find the trapezoid of largest area that can be inscribed in a semicircle of radius a, with one base lying along the diameter.

15. Two hallways, of widths a and b, meet at right angles (Figure 25). What is the greatest possible length of a ladder which can be carried horizontally around the corner?

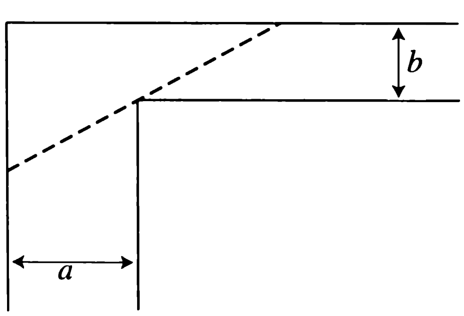


FIGURE 25

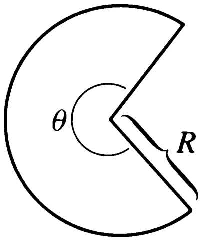


FIGURE 26

16. A garden is to be designed in the shape of a circular sector (Figure 26), with radius R and central angle θ. The garden is to have a fixed area A. For what value of R and θ (in radians) will the length of the fencing around the perimeter be minimized?

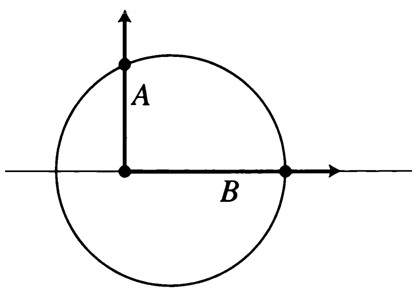


FIGURE 27

17. A right angle is moved along the diameter of a circle of radius a, as shown in Figure 27. What is the greatest possible length $(A + B)$ intercepted on it by the circle?

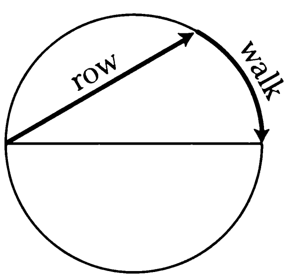


FIGURE 28

18. Ecological Ed must cross a circular lake of radius 1 mile. He can row across at 2 mph or walk around at 4 mph, or he can row part way and walk the rest (Figure 28). What route should he take so as to

(i) see as much scenery as possible?
(ii) cross as quickly as possible?

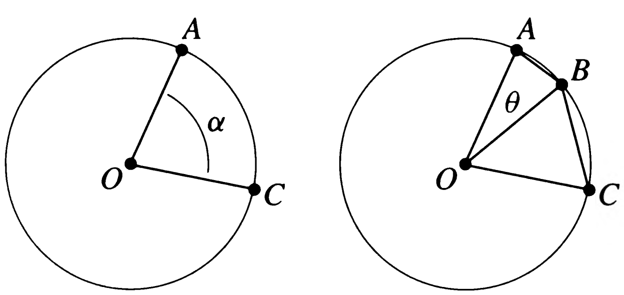


FIGURE 29

19. (a) Consider points A and B on a circle with center O, subtending an angle of $\alpha = \angle AOC$ (Figure 29). How must B be chosen so that the sum of the areas of $\triangle AOB$ and $\triangle BOC$ is a maximum? Hint: Express things in terms of $\theta = \angle AOB$.

(b) Prove that for $n \geq 3$, of all n-gons inscribed in a circle, the regular n-gon has maximum area.

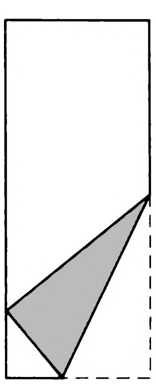


FIGURE 30

***20.** The lower right-hand corner of a page is folded over so that it just touches the left edge of the paper, as in Figure 30. If the width of the paper is α and the page is very long, show that the minimum length of the crease is $3\sqrt{3}\alpha/4$.

21. Figure 31 shows the graph of the *derivative* of f. Find all local maximum and minimum points of f.

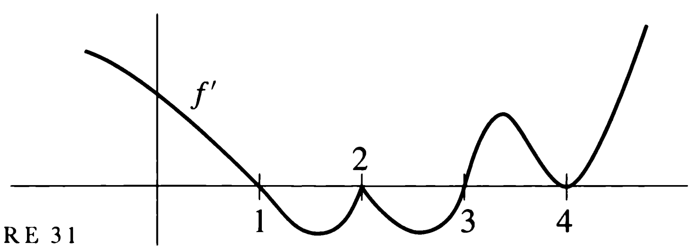


FIGURE 31

22. Suppose that f is a polynomial function, $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$, with critical points -1, 1, 2, 3, 4, and corresponding critical values 6, 1, 2, 4, 3. Sketch the graph of f, distinguishing the cases n even and n odd.

23. (a) Suppose that the critical points of the polynomial function $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ are $-1, 1, 2, 3$, and $f''(-1) = 0$, $f''(1) > 0$, $f''(2) < 0$, $f''(3) = 0$. Sketch the graph of f as accurately as possible on the basis of this information.

(b) Does there exist a polynomial function with the above properties, except that 3 is not a critical point?

24. Describe the graph of a rational function (in very general terms, similar to the text's description of the graph of a polynomial function).

25. (a) Prove that two polynomial functions of degree m and n, respectively, intersect in at most $\max(m, n)$ points.

(b) For each m and n exhibit two polynomial functions of degree m and n which intersect $\max(m, n)$ times.

26. Suppose f is a polynomial function of degree n with $f \geq 0$ (so n must be even). Prove that $f + f' + f'' + \cdots + f^{(n)} \geq 0$.

***27.** (a) Suppose that the polynomial function $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ has exactly k critical points and $f''(x) \neq 0$ for all critical points x. Show that $n - k$ is odd.

(b) For each n, show that if $n - k$ is odd, then there is a polynomial function f of degree n with k critical points, at each of which f'' is non-zero.

(c) Suppose that the polynomial function $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ has k_1 local maximum points and k_2 local minimum points. Show that $k_2 = k_1 + 1$ if n is even, and $k_2 = k_1$ if n is odd.

(d) Let n, k_1, k_2 be three integers with $k_2 = k_1 + 1$ if n is even, and $k_2 = k_1$ if n is odd, and $k_1 + k_2 < n$. Show that there is a polynomial function f of degree n, with k_1 local maximum points and k_2 local minimum points.

Hint: Pick $a_1 < a_2 < \cdots < a_{k_1+k_2}$ and try $f'(x) = \prod_{i=1}^{k_1+k_2} (x - a_i) \cdot (1 + x^2)^l$ for an appropriate number l.

28. (a) Prove that if $f'(x) \geq M$ for all x in $[a, b]$, then $f(b) \geq f(a) + M(b - a)$.

(b) Prove that if $f'(x) \leq M$ for all x in $[a, b]$, then $f(b) \leq f(a) + M(b - a)$.

(c) Formulate a similar theorem when $|f'(x)| \leq M$ for all x in $[a, b]$.

29. Suppose that $f'(x) \geq M > 0$ for all x in $[0, 1]$. Show that there is an interval of length $\frac{1}{4}$ on which $|f| \geq M/4$.

30. (a) Suppose that $f'(x) > g'(x)$ for all x, and that $f(a) = g(a)$. Show that $f(x) > g(x)$ for $x > a$ and $f(x) < g(x)$ for $x < a$.

(b) Show by an example that these conclusions do not follow without the hypothesis $f(a) = g(a)$.

(c) Suppose that $f(a) = g(a)$, that $f'(x) \geq g'(x)$ for all x, and that $f'(x_0) > g'(x_0)$ for some $x_0 > a$. Show that $f(x) > g(x)$ for all $x \geq x_0$.

31. Find all functions f such that

(a) $f'(x) = \sin x$.

(b) $f''(x) = x^3$.

(c) $f'''(x) = x + x^2$.

32. Although it is true that a weight dropped from rest will fall $s(t) = 16t^2$ feet after t seconds, this experimental fact does not mention the behavior of weights which are thrown upwards or downwards. On the other hand, the law $s''(t) = 32$ is always true and has just enough ambiguity to account for the behavior of a weight released from any height, with any initial velocity. For simplicity let us agree to measure heights upwards from ground level; in this case velocities are positive for rising bodies and negative for falling bodies, and all bodies fall according to the law $s''(t) = -32$.

(a) Show that s is of the form $s(t) = -16t^2 + \alpha t + \beta$.
(b) By setting $t = 0$ in the formula for s, and then in the formula for s', show that $s(t) = -16t^2 + v_0 t + s_0$, where s_0 is the height from which the body is released at time 0, and v_0 is the velocity with which it is released.
(c) A weight is thrown upwards with velocity v feet per second, at ground level. How high will it go? ("How high" means "what is the maximum height for all times".) What is its velocity at the moment it achieves its greatest height? What is its acceleration at that moment? When will it hit the ground again? What will its velocity be when it hits the ground again?

33. A cannon ball is shot from the ground with velocity v at an angle α (Figure 32) so that it has a vertical component of velocity $v \sin\alpha$ and a horizontal component $v\cos\alpha$. Its distance $s(t)$ above the ground obeys the law $s(t) = -16t^2 + (v\sin\alpha)t$, while its horizontal velocity remains constantly $v\cos\alpha$.

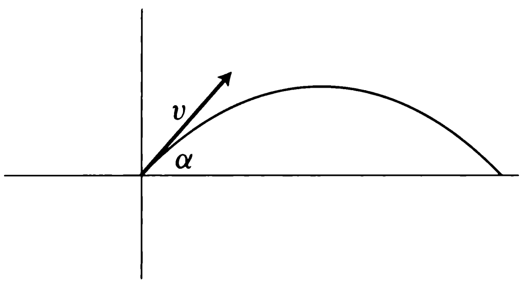


FIGURE 32

(a) Show that the path of the cannon ball is a parabola (find the position at each time t, and show that these points lie on a parabola).
(b) Find the angle α which will maximize the horizontal distance traveled by the cannon ball before striking the ground.

34. (a) Give an example of a function f for which $\lim_{x\to\infty} f(x)$ exists, but $\lim_{x\to\infty} f'(x)$ does not exist.
(b) Prove that if $\lim_{x\to\infty} f(x)$ and $\lim_{x\to\infty} f'(x)$ both exist, then $\lim_{x\to\infty} f'(x) = 0$.
(c) Prove that if $\lim_{x\to\infty} f(x)$ exists and $\lim_{x\to\infty} f''(x)$ exists, then $\lim_{x\to\infty} f''(x) = 0$. (See also Problem 20-22.)

35. Suppose that f and g are two differentiable functions which satisfy $fg' - f'g = 0$. Prove that if $f(a) = 0$ and $g(a) \neq 0$, then $f(x) = 0$ for all x in an interval around a. Hint: On any interval where f/g is defined, show that it is constant.

36. Suppose that $|f(x) - f(y)| \le |x - y|^n$ for $n > 1$. Prove that f is constant by considering f'. Compare with Problem 3-20.

37. A function f is *Lipschitz of order* α at x if there is a constant C such that

$$(*) \qquad |f(x) - f(y)| \le C|x - y|^{\alpha}$$

for all y in an interval around x. The function f is *Lipschitz of order* α *on an interval* if $(*)$ holds for all x and y in the interval.

(a) If f is Lipschitz of order $\alpha > 0$ at x, then f is continuous at x.
(b) If f is Lipschitz of order $\alpha > 0$ on an interval, then f is uniformly continuous on this interval (see Chapter 8, Appendix).
(c) If f is differentiable at x, then f is Lipschitz of order 1 at x. Is the converse true?
(d) If f is differentiable on $[a, b]$, is f Lipschitz of order 1 on $[a, b]$?
(e) If f is Lipschitz of order $\alpha > 1$ on $[a, b]$, then f is constant on $[a, b]$.

38. Prove that if

$$\frac{a_0}{1} + \frac{a_1}{2} + \cdots + \frac{a_n}{n+1} = 0,$$

then

$$a_0 + a_1 x + \cdots + a_n x^n = 0$$

for some x in $(0, 1)$.

39. Prove that the polynomial function $f_m(x) = x^3 - 3x + m$ never has two roots in $[0, 1]$, no matter what m may be. (This is an easy consequence of Rolle's Theorem. It is instructive, after giving an analytic proof, to graph f_0 and f_2, and consider where the graph of f_m lies in relation to them.)

40. Suppose that f is continuous and differentiable on $[0, 1]$, that $f(x)$ is in $[0, 1]$ for each x, and that $f'(x) \neq 1$ for all x in $[0, 1]$. Show that there is exactly one number x in $[0, 1]$ such that $f(x) = x$. (Half of this problem has been done already, in Problem 7-11.)

41. (a) Prove that the function $f(x) = x^2 - \cos x$ satisfies $f(x) = 0$ for precisely two numbers x.
(b) Prove the same for the function $f(x) = x^2 - x \sin x - \cos x$.
*(c) Prove this also for the function $f(x) = 2x^2 - x \sin x - \cos^2 x$. (Some preliminary estimates will be useful to restrict the possible location of the zeros of f.)

***42.** (a) Prove that if f is a twice differentiable function with $f(0) = 0$ and $f(1) = 1$ and $f'(0) = f'(1) = 0$, then $|f''(x)| \geq 4$ for some x in $(0, 1)$. In more picturesque terms: A particle which travels a unit distance in a unit time, and starts and ends with velocity 0, has at some time an acceleration ≥ 4. Hint: Prove that either $f''(x) \geq 4$ for some x in $(0, \frac{1}{2})$, or else $f''(x) \leq -4$ for some x in $(\frac{1}{2}, 1)$.
(b) Show that in fact we must have $|f''(x)| > 4$ for some x in $(0, 1)$.

43. Suppose that f is a function such that $f'(x) = 1/x$ for all $x > 0$ and $f(1) = 0$. Prove that $f(xy) = f(x) + f(y)$ for all $x, y > 0$. Hint: Find $g'(x)$ when $g(x) = f(xy)$.

44. Suppose that f satisfies

$$f''(x) + f'(x)g(x) - f(x) = 0$$

for some function g. Prove that if f is 0 at two points, then f is 0 on the interval between them. Hint: Use Theorem 6.

45. Suppose that f is continuous on $[a, b]$, that it is n-times differentiable on (a, b), and that $f(x) = 0$ for $n+1$ different x in $[a, b]$. Prove that $f^{(n)}(x) = 0$ for some x in (a, b).

46. Let $x_1, \dots, x_{n+1}$ be arbitrary points in $[a, b]$, and let

$$Q(x) = \prod_{i=1}^{n+1} (x - x_i).$$

Suppose that f is $(n + 1)$-times differentiable and that P is a polynomial function of degree $\le n$ such that $P(x_i) = f(x_i)$ for $i = 1, \dots, n + 1$ (see Problem 3-6). Show that for each x in $[a, b]$ there is a number c in (a, b) such that

$$f(x) - P(x) = Q(x) \cdot \frac{f^{(n+1)}(c)}{(n+1)!}.$$

Hint: Consider the function

$$F(t) = Q(x)[f(t) - P(t)] - Q(t)[f(x) - P(x)].$$

Show that F is zero at $n + 2$ different points in $[a, b]$, and use Problem 45.

47. Prove that

$$\tfrac{1}{9} < \sqrt{66} - 8 < \tfrac{1}{8}$$

(without computing $\sqrt{66}$ to 2 decimal places!).

48. Prove the following slight generalization of the Mean Value Theorem: If f is continuous and differentiable on (a, b) and $\lim_{y \to a^+} f(y)$ and $\lim_{y \to b^-} f(y)$ exist, then there is some x in (a, b) such that

$$f'(x) = \frac{\lim_{y \to b^-} f(y) - \lim_{y \to a^+} f(y)}{b - a}.$$

(Your proof should begin: "This is a trivial consequence of the Mean Value Theorem because . . . ".)

49. Prove that the conclusion of the Cauchy Mean Value Theorem can be written in the form

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x)}{g'(x)},$$

under the additional assumptions that $g(b) \neq g(a)$ and that $f'(x)$ and $g'(x)$ are never simultaneously 0 on (a, b).

50. Prove that if f and g are continuous on $[a, b]$ and differentiable on (a, b), and $g'(x) \neq 0$ for x in (a, b), then there is some x in (a, b) with

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

Hint: Multiply out first, to see what this really says.

51. What is wrong with the following use of l'Hôpital's Rule:

$$\lim_{x \to 1} \frac{x^3 + x - 2}{x^2 - 3x + 2} = \lim_{x \to 1} \frac{3x^2 + 1}{2x - 3} = \lim_{x \to 1} \frac{6x}{2} = 3.$$

(The limit is actually -4.)

52. Find the following limits:

(i) $\lim_{x \to 0} \dfrac{x}{\tan x}$.

(ii) $\lim_{x \to 0} \dfrac{\cos^2 x - 1}{x^2}$.

53. Find $f'(0)$ if

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

and $g(0) = g'(0) = 0$ and $g''(0) = 17$.

54. Prove the following forms of l'Hôpital's Rule (none requiring any essentially new reasoning).

(a) If $\lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x) = 0$, and $\lim_{x \to a^+} f'(x)/g'(x) = l$, then $\lim_{x \to a^+} f(x)/g(x) = l$ (and similarly for limits from below).

(b) If $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0$, and $\lim_{x \to a} f'(x)/g'(x) = \infty$, then $\lim_{x \to a} f(x)/g(x) = \infty$ (and similarly for $-\infty$, or if $x \to a$ is replaced by $x \to a^+$ or $x \to a^-$).

(c) If $\lim_{x \to \infty} f(x) = \lim_{x \to \infty} g(x) = 0$, and $\lim_{x \to \infty} f'(x)/g'(x) = l$, then $\lim_{x \to \infty} f(x)/g(x) = l$ (and similarly for $-\infty$). Hint: Consider $\lim_{x \to 0^+} f(1/x)/g(1/x)$.

(d) If $\lim_{x \to \infty} f(x) = \lim_{x \to \infty} g(x) = 0$, and $\lim_{x \to \infty} f'(x)/g'(x) = \infty$, then $\lim_{x \to \infty} f(x)/g(x) = \infty$.

55. There is another form of l'Hôpital's Rule which requires more than algebraic manipulations: If $\lim_{x \to \infty} f(x) = \lim_{x \to \infty} g(x) = \infty$, and $\lim_{x \to \infty} f'(x)/g'(x) = l$, then $\lim_{x \to \infty} f(x)/g(x) = l$. Prove this as follows.

(a) For every $\varepsilon > 0$ there is a number a such that

$$\left| \frac{f'(x)}{g'(x)} - l \right| < \varepsilon \quad \text{for } x > a.$$

Apply the Cauchy Mean Value Theorem to f and g on $[a, x]$ to show that

$$\left| \frac{f(x) - f(a)}{g(x) - g(a)} - l \right| < \varepsilon \quad \text{for } x > a.$$

(Why can we assume $g(x) - g(a) \neq 0$?)

(b) Now write

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(a)}{g(x)-g(a)} \cdot \frac{f(x)}{f(x)-f(a)} \cdot \frac{g(x)-g(a)}{g(x)}$$

(why can we assume that $f(x) - f(a) \neq 0$ for large x?) and conclude that

$$\left| \frac{f(x)}{g(x)} - l \right| < 2\varepsilon \quad \text{for sufficiently large } x.$$

56. To complete the orgy of variations on l'Hôpital's Rule, use Problem 55 to prove a few more cases of the following general statement (there are so many possibilities that you should select just a few, if any, that interest you):

If $\lim\limits_{x\to[\ ]} f(x) = \lim\limits_{x\to[\ ]} g(x) = \{\quad\}$ and $\lim\limits_{x\to[\ ]} f'(x)/g'(x) = (\quad)$, then $\lim\limits_{x\to[\ ]} f(x)/g(x) = (\quad)$. Here [ ] can be a or a^+ or a^- or ∞ or $-\infty$, and { } can be 0 or ∞ or $-\infty$, and ( ) can be l or ∞ or $-\infty$.

57. If f and g are differentiable and $\lim\limits_{x\to a} f(x)/g(x)$ exists, does it follow that $\lim\limits_{x\to a} f'(x)/g'(x)$ exists (a converse to l'Hôpital's Rule)?

58. Prove that if f' is increasing, then every tangent line of f intersects the graph of f only once. (In particular, this is true for the function $f(x) = x^n$ if n is even.)

59. Redo Problem 10-18 (c) when

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

(Why is this problem is this chapter?)

***60.** (a) Suppose that f is differentiable on $[a, b]$. Prove that if the minimum of f on $[a, b]$ is at a, then $f'(a) \geq 0$, and if it is at b, then $f'(b) \leq 0$. (One half of the proof of Theorem 1 will go through.)

(b) Suppose that $f'(a) < 0$ and $f'(b) > 0$. Show that $f'(x) = 0$ for some x in (a, b). Hint: Consider the minimum of f on $[a, b]$; why must it be somewhere in (a, b)?

(c) Prove that if $f'(a) < c < f'(b)$, then $f'(x) = c$ for some x in (a, b). (This result is known as Darboux's Theorem. Note that we are *not* assuming that f' is continuous.) Hint: Cook up an appropriate function to which part (b) may be applied.

61. Suppose that f is differentiable in some interval containing a, but that f' is discontinuous at a. Prove the following:

(a) The one-sided limits $\lim\limits_{x\to a^+} f'(x)$ and $\lim\limits_{x\to a^-} f'(x)$ cannot both exist. (This is just a minor variation on Theorem 7.)

(b) These one-sided limits cannot both exist even if we allow limits with the value $+\infty$ or $-\infty$. Hint: Use Darboux's Theorem (Problem 60).

***62.** It is easy to find a function f such that $|f|$ is differentiable but f is not. For example, we can choose $f(x) = 1$ for x rational and $f(x) = -1$ for x irrational. In this example f is not even continuous, nor is this a mere coincidence: Prove that if $|f|$ is differentiable at a, and f is continuous at a, then f is also differentiable at a. Hint: It suffices to consider only a with $f(a) = 0$. Why? In this case, what must $|f|'(a)$ be?

***63.** (a) Let $y \neq 0$ and let n be even. Prove that $x^n + y^n = (x + y)^n$ only when $x = 0$. Hint: If $x_0{}^n + y^n = (x_0 + y)^n$, apply Rolle's Theorem to $f(x) = x^n + y^n - (x + y)^n$ on $[0, x_0]$.

(b) Prove that if $y \neq 0$ and n is odd, then $x^n + y^n = (x + y)^n$ only if $x = 0$ or $x = -y$.

64. Suppose that $f(0) = 0$ and f' is increasing. Prove that the function $g(x) = f(x)/x$ is increasing on $(0, \infty)$. Hint: Obviously you should look at $g'(x)$. Prove that it is positive by applying the Mean Value Theorem to f on the right interval (it will help to remember that the hypothesis $f(0) = 0$ is essential, as shown by the function $f(x) = 1 + x^2$).

65. Use derivatives to prove that if $n \geq 1$, then

$$(1 + x)^n > 1 + nx \quad \text{for } -1 < x < 0 \text{ and } 0 < x$$

(notice that equality holds for $x = 0$).

66. Let $f(x) = x^4 \sin^2 1/x$ for $x \neq 0$, and let $f(0) = 0$ (Figure 33).

(a) Prove that 0 is a local minimum point for f.

(b) Prove that $f'(0) = f''(0) = 0$.

This function thus provides another example to show that Theorem 6 cannot be improved. It also illustrates a subtlety about maxima and minima that often goes unnoticed: a function may not be increasing in any interval to the right of a local minimum point, nor decreasing in any interval to the left.

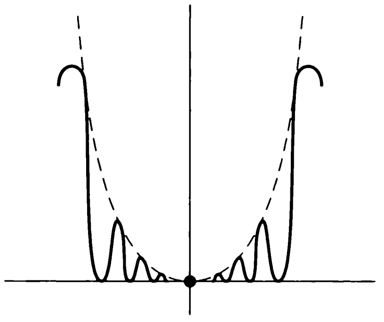


FIGURE 33

***67.** (a) Prove that if $f'(a) > 0$ and f' is continuous at a, then f is increasing in some interval containing a.

The next two parts of this problem show that continuity of f' is essential.

(b) If $g(x) = x^2 \sin 1/x$, show that there are numbers x arbitrarily close to 0 with $g'(x) = 1$ and also with $g'(x) = -1$.

(c) Suppose $0 < \alpha < 1$. Let $f(x) = \alpha x + x^2 \sin 1/x$ for $x \neq 0$, and let $f(0) = 0$ (see Figure 34). Show that f is not increasing in any open interval containing 0, by showing that in any interval there are points x with $f'(x) > 0$ and also points x with $f'(x) < 0$.

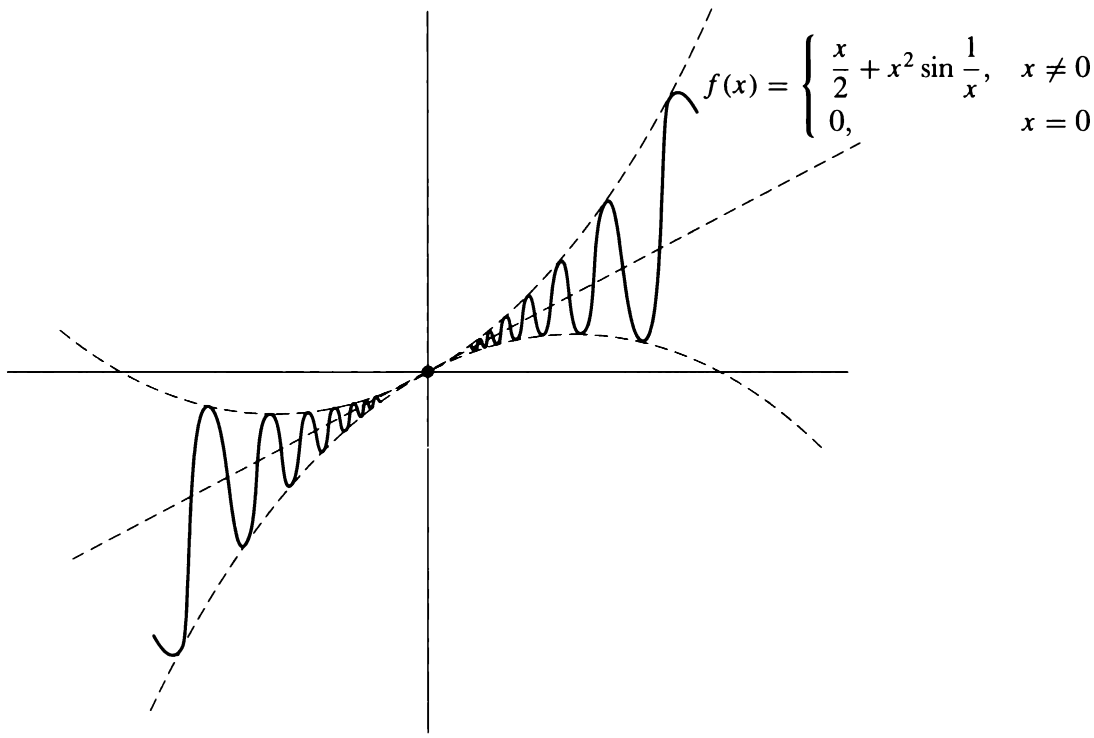


FIGURE 34

The behavior of f for $\alpha \geq 1$, which is much more difficult to analyze, is discussed in the next problem.

****68.** Let $f(x) = \alpha x + x^2 \sin 1/x$ for $x \neq 0$, and let $f(0) = 0$. In order to find the sign of $f'(x)$ when $\alpha \geq 1$ it is necessary to decide if $2x \sin 1/x - \cos 1/x$ is < -1 for any numbers x close to 0. It is a little more convenient to consider the function $g(y) = 2(\sin y)/y - \cos y$ for $y \neq 0$; we want to know if $g(y) < -1$ for large y. This question is quite delicate; the most significant part of $g(y)$ is $-\cos y$, which does reach the value -1, but this happens only when $\sin y = 0$, and it is not at all clear whether g itself can have values < -1. The obvious approach to this problem is to find the local minimum values of g. Unfortunately, it is impossible to solve the equation $g'(y) = 0$ explicitly, so more ingenuity is required.

(a) Show that if $g'(y) = 0$, then

$$\cos y = (\sin y)\left(\frac{2 - y^2}{2y}\right),$$

and conclude that

$$g(y) = (\sin y)\left(\frac{2 + y^2}{2y}\right).$$

(b) Now show that if $g'(y) = 0$, then

$$\sin^2 y = \frac{4y^2}{4+y^4},$$

and conclude that

$$|g(y)| = \frac{2+y^2}{\sqrt{4+y^4}}.$$

(c) Using the fact that $(2+y^2)/\sqrt{4+y^4} > 1$, show that if $\alpha = 1$, then f is not increasing in any interval around 0.

(d) Using the fact that $\lim_{y\to\infty} (2+y^2)/\sqrt{4+y^4} = 1$, show that if $\alpha > 1$, then f *is* increasing in some interval around 0.

****69.** A function f is **increasing at a** if there is some number $\delta > 0$ such that

$$f(x) > f(a) \quad \text{if} \quad a < x < a+\delta$$

and

$$f(x) < f(a) \quad \text{if} \quad a-\delta < x < a.$$

Notice that this does *not* mean that f is increasing in the interval $(a-\delta, a+\delta)$; for example, the function shown in Figure 34 is increasing at 0, but is not an increasing function in any open interval containing 0.

(a) Suppose that f is continuous on $[0, 1]$ and that f is increasing at a for every a in $[0, 1]$. Prove that f is increasing on $[0, 1]$. (First convince yourself that there is something to be proved.) Hint: For $0 < b < 1$, prove that the minimum of f on $[b, 1]$ must be at b.

(b) Prove part (a) without the assumption that f is continuous, by considering for each b in $[0, 1]$ the set $S_b = \{x : f(y) \geq f(b) \text{ for all } y \text{ in } [b, x]\}$. (This part of the problem is not necessary for the other parts.) Hint: Prove that $S_b = \{x : b \leq x \leq 1\}$ by considering $\sup S_b$.

(c) If f is increasing at a and f is differentiable at a, prove that $f'(a) \geq 0$ (this is easy).

(d) If $f'(a) > 0$, prove that f is increasing at a (go right back to the definition of $f'(a)$).

(e) Use parts (a) and (d) to show, without using the Mean Value Theorem, that if f is continuous on $[0, 1]$ and $f'(a) > 0$ for all a in $[0, 1]$, then f is increasing on $[0, 1]$.

(f) Suppose that f is continuous on $[0, 1]$ and $f'(a) = 0$ for all a in $(0, 1)$. Apply part (e) to the function $g(x) = f(x) + \varepsilon x$ to show that $f(1) - f(0) > -\varepsilon$. Similarly, show that $f(1) - f(0) < \varepsilon$ by considering $h(x) = \varepsilon x - f(x)$. Conclude that $f(0) = f(1)$.

This particular proof that a function with zero derivative must be constant has many points in common with a proof of H. A. Schwarz, which may be the first rigorous proof ever given. Its discoverer, at least, seemed to think it was. See his exuberant letter in reference [54] of the Suggested Reading.

****70.** (a) If f is a constant function, then every point is a local maximum point for f. It is quite possible for this to happen even if f is not a constant function: for example, if $f(x) = 0$ for $x < 0$ and $f(x) = 1$ for $x \geq 0$. But prove, using Problem 8-4, that if f is continuous on $[a, b]$ and every point of $[a, b]$ is a local maximum point, then f is a constant function. The same result holds, of course, if every point of $[a, b]$ is a local minimum point.

(b) Suppose now that every point is either a local maximum or a local minimum point for the continuous function f (but we don't preclude the possibility that some points are local maxima while others are local minima). Prove that f is constant, as follows. Suppose that $f(a_0) < f(b_0)$. We can assume that $f(a_0) < f(x) < f(b_0)$ for $a_0 < x < b_0$. (Why?) Using Theorem 1 of the Appendix to Chapter 8, partition $[a_0, b_0]$ into intervals on which $\sup f - \inf f < (f(b_0) - f(a_0))/2$; also choose the lengths of these intervals to be less than $(b_0 - a_0)/2$. Then there is one such interval $[a_1, b_1]$ with $a_0 < a_1 < b_1 < b_0$ and $f(a_1) < f(b_1)$. (Why?) Continue inductively and use the Nested Interval Theorem (Problem 8-14) to find a point x that cannot be a local maximum or minimum.

****71.** (a) A point x is called a **strict maximum point** for f on A if $f(x) > f(y)$ for all y in A with $y \neq x$ (compare with the definition of an ordinary maximum point). A **local strict maximum point** is defined in the obvious way. Find all local strict maximum points of the function

$$f(x) = \begin{cases} 0, & x \text{ irrational} \\ \dfrac{1}{q}, & x = \dfrac{p}{q} \text{ in lowest terms.} \end{cases}$$

It seems quite unlikely that a function can have a local strict maximum at *every* point (although the above example might give one pause for thought). Prove this as follows.

(b) Suppose that every point is a local strict maximum point for f. Let x_1 be any number and choose $a_1 < x_1 < b_1$ with $b_1 - a_1 < 1$ such that $f(x_1) > f(x)$ for all x in $[a_1, b_1]$. Let $x_2 \neq x_1$ be any point in (a_1, b_1) and choose $a_1 \leq a_2 < x_2 < b_2 \leq b_1$ with $b_2 - a_2 < \frac{1}{2}$ such that $f(x_2) > f(x)$ for all x in $[a_2, b_2]$. Continue in this way, and use the Nested Interval Theorem (Problem 8-14) to obtain a contradiction.

APPENDIX. CONVEXITY AND CONCAVITY

Although the graph of a function can be sketched quite accurately on the basis of the information provided by the derivative, some subtle aspects of the graph are revealed only by examining the second derivative. These details were purposely omitted previously because graph sketching is complicated enough without worrying about them, and the additional information obtained is often not worth the effort. Also, correct proofs of the relevant facts are sufficiently difficult to be placed in an appendix. Despite these discouraging remarks, the information presented here is well worth assimilating, because the notions of convexity and concavity are far more important than as mere aids to graph sketching. Moreover, the proofs have a pleasantly geometric flavor not often found in calculus theorems. Indeed, the basic definition is geometric in nature (see Figure 1).

DEFINITION 1

A function f is **convex** on an interval, if for all a and b in the interval, the line segment joining $(a, f(a))$ and $(b, f(b))$ lies above the graph of f.

The geometric condition appearing in this definition can be expressed in an analytic way that is sometimes more useful in proofs. The straight line between $(a, f(a))$ and $(b, f(b))$ is the graph of the function g defined by

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

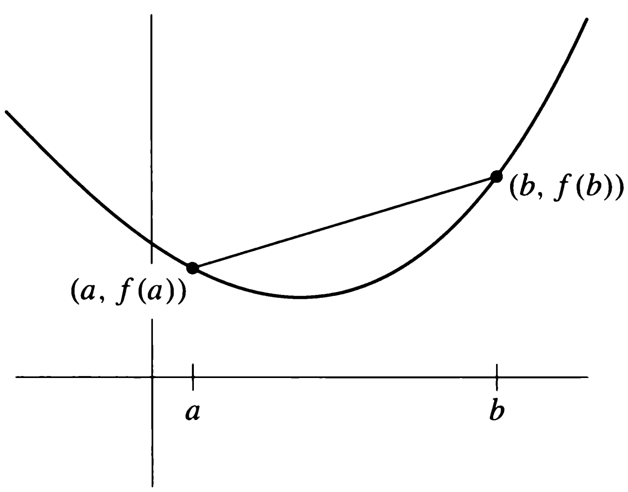


FIGURE 1

This line lies above the graph of f at x if $g(x) > f(x)$, that is, if

$$\frac{f(b) - f(a)}{b - a}(x - a) + f(a) > f(x)$$

or

$$\frac{f(b) - f(a)}{b - a}(x - a) > f(x) - f(a)$$

or

$$\frac{f(b) - f(a)}{b - a} > \frac{f(x) - f(a)}{x - a}.$$

We therefore have an equivalent definition of convexity.

DEFINITION 2

A function f is **convex** on an interval if for a, x, and b in the interval with $a < x < b$ we have

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a}.$$

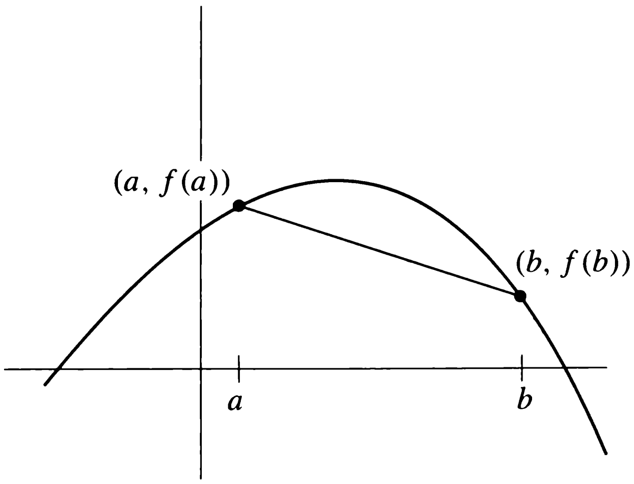


FIGURE 2

If the word "over" in Definition 1 is replaced by "under" or, equivalently, if the inequality in Definition 2 is replaced by

$$\frac{f(x) - f(a)}{x - a} > \frac{f(b) - f(a)}{b - a},$$

we obtain the definition of a **concave** function (Figure 2). It is not hard to see that the concave functions are precisely the ones of the form $-f$, where f is convex. For this reason, the next three theorems about convex functions have immediate corollaries about concave functions, so simple that we will not even bother to state them.

Figure 3 shows some tangent lines of a convex function. Two things seem to be true:

(1) The graph of f lies above the tangent line at $(a, f(a))$ except at the point $(a, f(a))$ itself (this point is called the **point of contact** of the tangent line).
(2) If $a < b$, then the slope of the tangent line at $(a, f(a))$ is less than the slope of the tangent line at $(b, f(b))$; that is, f' is increasing.

As a matter of fact these observations are true, and the proofs are not difficult.

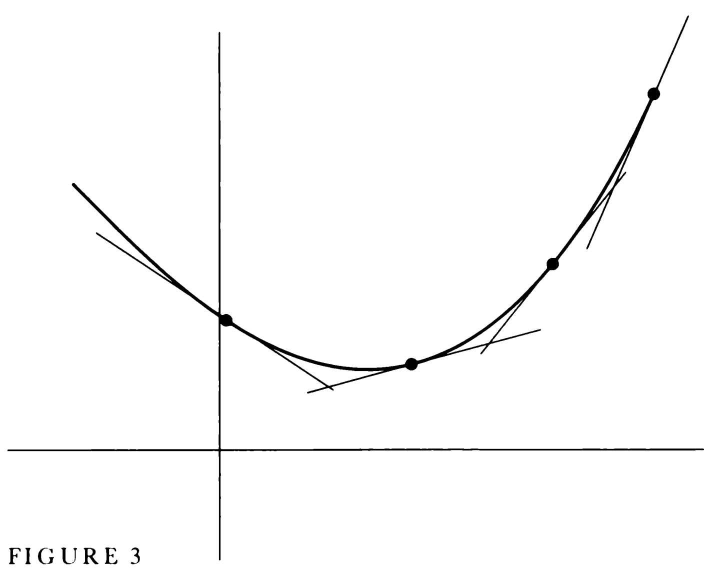


FIGURE 3

THEOREM 1 Let f be convex. If f is differentiable at a, then the graph of f lies above the tangent line through $(a, f(a))$, except at $(a, f(a))$ itself. If $a < b$ and f is differentiable at a and b, then $f'(a) < f'(b)$.

PROOF If $0 < h_1 < h_2$, then as Figure 4 indicates,

$$(1) \quad \frac{f(a + h_1) - f(a)}{h_1} < \frac{f(a + h_2) - f(a)}{h_2}.$$

A nonpictorial proof can be derived immediately from Definition 2 applied to $a < a + h_1 < a + h_2$. Inequality (1) shows that the values of

$$\frac{f(a + h) - f(a)}{h}$$

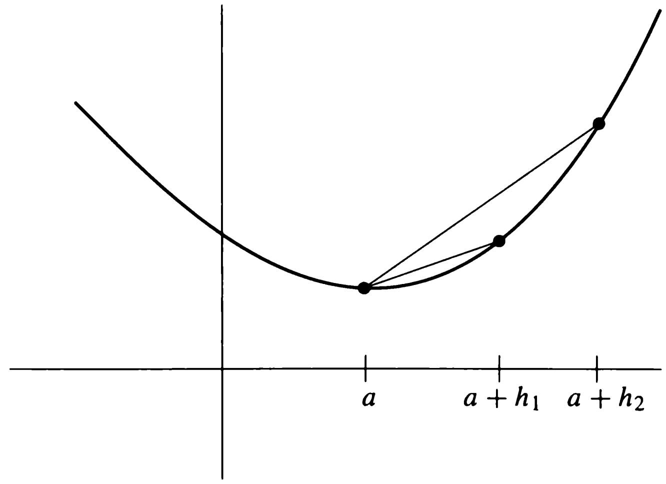


FIGURE 4

decrease as $h \to 0^+$. Consequently,

$$f'(a) < \frac{f(a+h) - f(a)}{h} \quad \text{for } h > 0$$

(in fact $f'(a)$ is the greatest lower bound of all these numbers). But this means that for $h > 0$ the secant line through $(a, f(a))$ and $(a+h, f(a+h))$ has larger slope than the tangent line, which implies that $(a+h, f(a+h))$ lies above the tangent line (an analytic translation of this argument is easily supplied).

For negative h there is a similar situation (Figure 5): if $h_2 < h_1 < 0$, then

$$\frac{f(a+h_1) - f(a)}{h_1} > \frac{f(a+h_2) - f(a)}{h_2}.$$

This shows that the slope of the tangent line is greater than

$$\frac{f(a+h) - f(a)}{h} \quad \text{for } h < 0$$

(in fact $f'(a)$ is the least upper bound of all these numbers), so that $f(a+h)$ lies above the tangent line if $h < 0$. This proves the first part of the theorem.

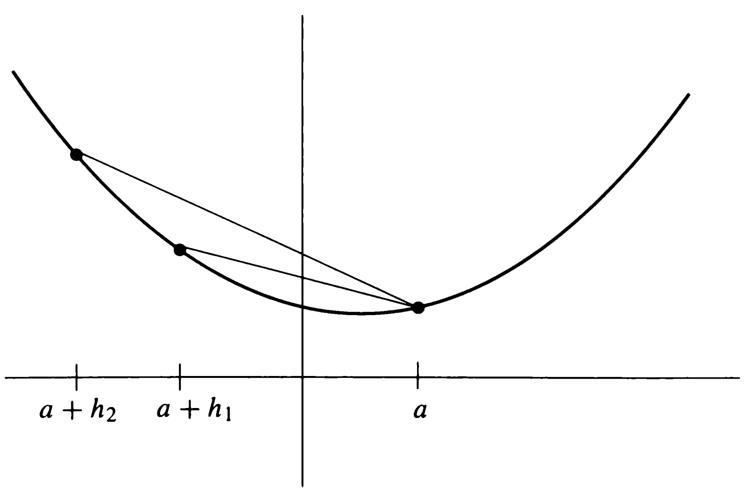


FIGURE 5

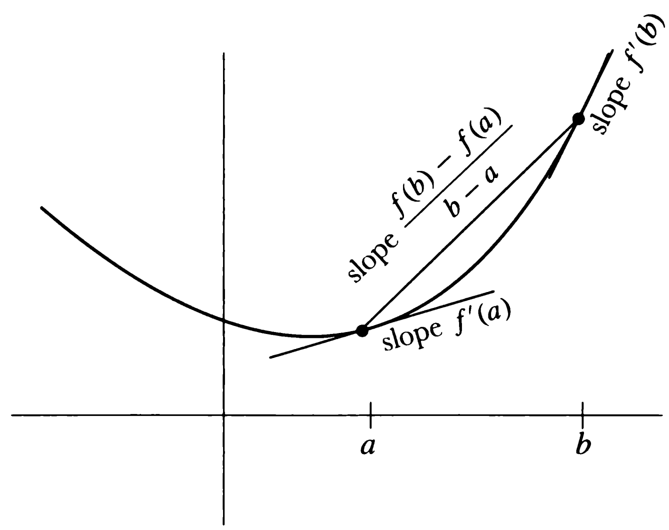


FIGURE 6

Now suppose that $a < b$. Then, as we have already seen (Figure 6),

$$\begin{aligned} f'(a) &< \frac{f(a+(b-a))-f(a)}{b-a} \qquad \text{since } b-a>0 \\ &= \frac{f(b)-f(a)}{b-a} \end{aligned}$$

and

$$\begin{aligned} f'(b) &> \frac{f(b+(a-b))-f(b)}{a-b} \qquad \text{since } a-b<0 \\ &= \frac{f(a)-f(b)}{a-b} = \frac{f(b)-f(a)}{b-a}. \end{aligned}$$

Combining these inequalities, we obtain $f'(a) < f'(b)$. ▌

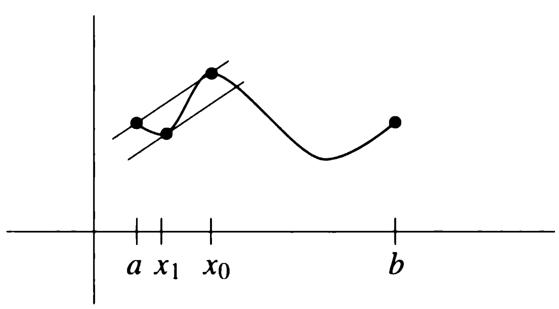


FIGURE 7

Theorem 1 has two converses. Here the proofs will be a little more difficult. We begin with a lemma that plays the same role in the next theorem that Rolle's Theorem plays in the proof of the Mean Value Theorem. It states that if f' is increasing, then the graph of f lies below any secant line *which happens to be horizontal.*

LEMMA Suppose f is differentiable and f' is increasing. If $a < b$ and $f(a) = f(b)$, then $f(x) < f(a) = f(b)$ for $a < x < b$.

PROOF Suppose that $f(x) \geq f(a) = f(b)$ for some x in (a, b). Then the maximum of f on $[a, b]$ occurs at some point x_0 in (a, b) with $f(x_0) \geq f(a)$ and, of course, $f'(x_0) = 0$ (Figure 7). On the other hand, applying the Mean Value Theorem to the interval $[a, x_0]$, we find that there is x_1 with $a < x_1 < x_0$ and

$$f'(x_1) = \frac{f(x_0)-f(a)}{x_0-a} \geq 0,$$

contradicting the fact that f' is increasing. ▌

We now attack the general case by the same sort of algebraic machinations that we used in the proof of the Mean Value Theorem.

THEOREM 2 If f is differentiable and f' is increasing, then f is convex.

PROOF Let $a < b$. Define g by

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

It is easy to see that g' is also increasing; moreover, $g(a) = g(b) = f(a)$. Applying the lemma to g we conclude that

$$g(x) < f(a) \quad \text{if} \quad a < x < b.$$

In other words, if $a < x < b$, then

$$f(x) - \frac{f(b) - f(a)}{b - a}(x - a) < f(a)$$

or

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a}.$$

Hence f is convex. ▌

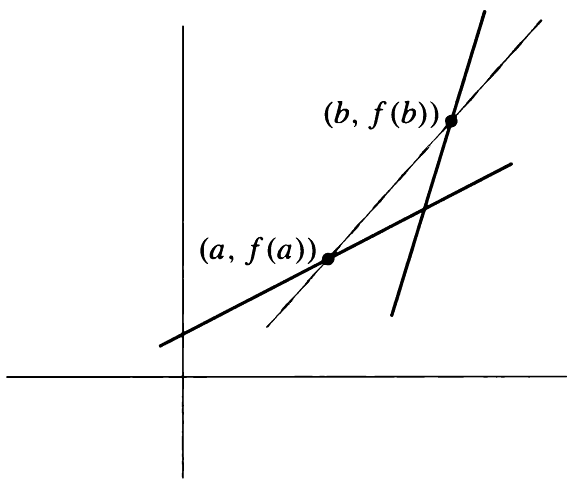


FIGURE 8

THEOREM 3 If f is differentiable and the graph of f lies above each tangent line except at the point of contact, then f is convex.

PROOF Let $a < b$. It is clear from Figure 8 that if $(b, f(b))$ lies above the tangent line at $(a, f(a))$, and $(a, f(a))$ lies above the tangent line at $(b, f(b))$, then the slope of the tangent line at $(b, f(b))$ must be larger than the slope of the tangent line at $(a, f(a))$. The following argument just says this with equations.

Since the tangent line at $(a, f(a))$ is the graph of the function

$$g(x) = f'(a)(x - a) + f(a),$$

and since $(b, f(b))$ lies above the tangent line, we have

$$(1) \quad f(b) > f'(a)(b - a) + f(a).$$

Similarly, since the tangent line at $(b, f(b))$ is the graph of

$$h(x) = f'(b)(x - b) + f(b),$$

and $(a, f(a))$ lies above the tangent line at $(b, f(b))$, we have

$$(2) \quad f(a) > f'(b)(a - b) + f(b).$$

It follows from (1) and (2) that $f'(a) < f'(b)$.

It now follows from Theorem 2 that f is convex. ▌

If a function f has a reasonable second derivative, the information given in these theorems can be used to discover the regions in which f is convex or concave.

Consider, for example, the function

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

For this function,

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

Thus $f'(x) = 0$ only for $x = 0$, and $f(0) = 1$, while

$$\begin{aligned} f'(x) &> 0 \quad \text{if} \quad x < 0, \\ f'(x) &< 0 \quad \text{if} \quad x > 0. \end{aligned}$$

Moreover,

$$\begin{aligned} &f(x) > 0 \quad \text{for all } x, \\ &f(x) \to 0 \quad \text{as } x \to \infty \text{ or } -\infty, \\ &f \text{ is even.} \end{aligned}$$

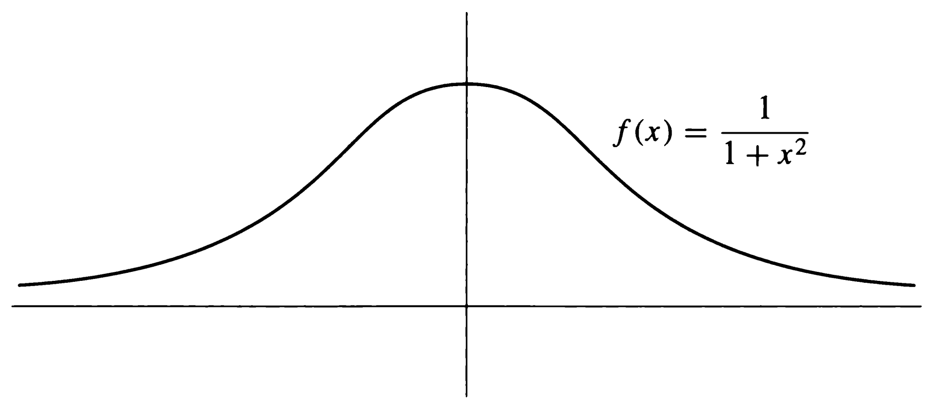


FIGURE 9

The graph of f therefore looks something like Figure 9. We now compute

$$\begin{aligned} f''(x) &= \frac{(1+x^2)^2(-2) + 2x \cdot [2(1+x^2) \cdot 2x]}{(1+x^2)^4} \\ &= \frac{2(3x^2-1)}{(1+x^2)^3}. \end{aligned}$$

It is not hard to determine the sign of $f''(x)$. Note first that $f''(x) = 0$ only when $x = \sqrt{1/3}$ or $-\sqrt{1/3}$. Since f'' is clearly continuous, it must keep the same sign on each of the sets

$$\begin{gathered} (-\infty, -\sqrt{1/3}\,), \\ (-\sqrt{1/3}, \sqrt{1/3}\,), \\ (\sqrt{1/3}, \infty). \end{gathered}$$

Since we easily compute, for example, that

$$\begin{aligned} f''(-1) &= \tfrac{1}{2} > 0, \\ f''(0) &= -2 < 0, \\ f''(1) &= \tfrac{1}{2} > 0, \end{aligned}$$

we conclude that

$$\begin{aligned} &f'' > 0 \text{ on } (-\infty, -\sqrt{1/3}\,) \text{ and } (\sqrt{1/3}, \infty), \\ &f'' < 0 \text{ on } (-\sqrt{1/3}, \sqrt{1/3}\,). \end{aligned}$$

Since $f'' > 0$ means f' is increasing, it follows from Theorem 2 that f is convex on $(-\infty, -\sqrt{1/3}\,)$ and $(\sqrt{1/3}, \infty)$, while on $(-\sqrt{1/3}, \sqrt{1/3}\,)$ f is concave (Figure 10).

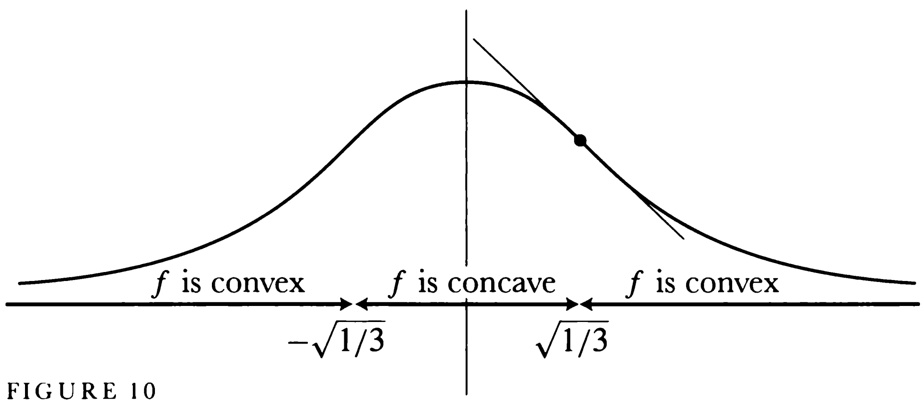


FIGURE 10

Notice that at $(\sqrt{1/3}, \frac{3}{4})$ the tangent line lies below the part of the graph to the right, since f is convex on $(\sqrt{1/3}, \infty)$, and above the part of the graph to the left, since f is concave on $(-\sqrt{1/3}, \sqrt{1/3}\,)$; thus the tangent line crosses the graph. In general, a number a is called an **inflection point** of f if the tangent line to the graph of f at $(a, f(a))$ crosses the graph; thus $\sqrt{1/3}$ and $-\sqrt{1/3}$ are inflection points of $f(x) = 1/(1+x^2)$. Note that the condition $f''(a) = 0$ does *not* ensure that a is an inflection point of f; for example, if $f(x) = x^4$, then $f''(0) = 0$, but f is convex, so the tangent line at $(0, 0)$ certainly doesn't cross the graph of f. In order to conclude that a is an inflection point of a function f, we need to know that f'' has different signs to the left and right of a.

This example illustrates the procedure which may be used to analyze any function f. After the graph has been sketched, using the information provided by f', the zeros of f'' are computed and the sign of f'' is determined on the intervals between consecutive zeros. On intervals where $f'' > 0$ the function is convex; on intervals where $f'' < 0$ the function is concave. Knowledge of the regions of convexity and concavity of f can often prevent absurd misinterpretation of other data about f. Several functions, which can be analyzed in this way, are given in the problems, which also contain some theoretical questions.

To round out our discussion of convexity and concavity, we will prove one further result that you may already have begun to suspect. We have seen that convex and concave functions have the property that every tangent line intersects the graph just once; a few drawings will probably convince you that no other functions have this property, but the only proof I know is rather tricky.

THEOREM 4 If f is differentiable on an interval and intersects each of its tangent lines just once, then f is either convex or concave on that interval.

PROOF There are two parts to the proof.

(1) First we claim that no straight line can intersect the graph of f in *three* different points. Suppose, on the contrary, that some straight line did intersect the graph of f at $(a, f(a))$, $(b, f(b))$ and $(c, f(c))$, with $a < b < c$ (Figure 11). Then we would have

$$(1) \qquad \frac{f(b)-f(a)}{b-a} = \frac{f(c)-f(a)}{c-a}.$$

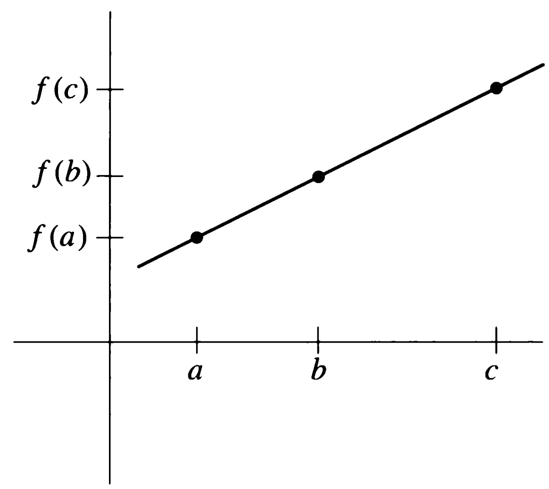


FIGURE 11

Consider the function

$$g(x) = \frac{f(x)-f(a)}{x-a} \qquad \text{for } x \text{ in } [b, c].$$

Equation (1) says that $g(b) = g(c)$. So by Rolle's Theorem, there is some number x in (b, c) where $0 = g'(x)$, and thus

$$0 = (x-a)f'(x) - [f(x) - f(a)]$$

or

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

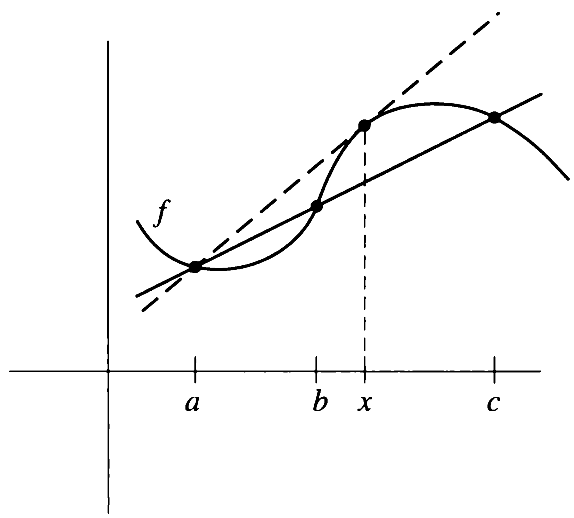


FIGURE 12

But this says (Figure 12) that the tangent line at $(x, f(x))$ passes through $(a, f(a))$, contradicting the hypotheses.

(2) Suppose that $a_0 < b_0 < c_0$ and $a_1 < b_1 < c_1$ are points in the interval. Let

$$\begin{aligned} x_t &= (1-t)a_0 + ta_1 \\ y_t &= (1-t)b_0 + tb_1 \qquad 0 \le t \le 1. \\ z_t &= (1-t)c_0 + tc_1 \end{aligned}$$

Then $x_0 = a_0$ and $x_1 = a_1$ and (Problem 4-2) the points x_t all lie between a_0 and a_1, with analogous statements for y_t and z_t. Moreover,

$$x_t < y_t < z_t \qquad \text{for} \qquad 0 \le t \le 1.$$

Now consider the function

$$g(t) = \frac{f(y_t)-f(x_t)}{y_t - x_t} - \frac{f(z_t)-f(x_t)}{z_t - x_t} \qquad \text{for } 0 \le t \le 1.$$

By step (1), $g(t) \ne 0$ for all t in $[0, 1]$. So either $g(t) > 0$ for all t in $[0, 1]$ or $g(t) < 0$ for all t in $[0, 1]$. Thus, either f is convex or f is concave. ▮

PROBLEMS

1. Sketch, indicating regions of convexity and concavity and points of inflection, the functions in Problem 11-1 (consider (iv) as double starred).

2. Figure 30 in Chapter 11 shows the graph of f'. Sketch the graph of f.

3. Show that f is convex on an interval if and only if for all x and y in the interval we have

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y), \quad \text{for } 0 < t < 1.$$

(This is just a restatement of the definition, but a useful one.)

4. (a) Prove that if f and g are convex and f is increasing, then $f \circ g$ is convex. (It will be easiest to use Problem 3.)
(b) Give an example where $g \circ f$ is not convex.
(c) Suppose that f and g are twice differentiable. Give another proof of the result of part (a) by considering second derivatives.

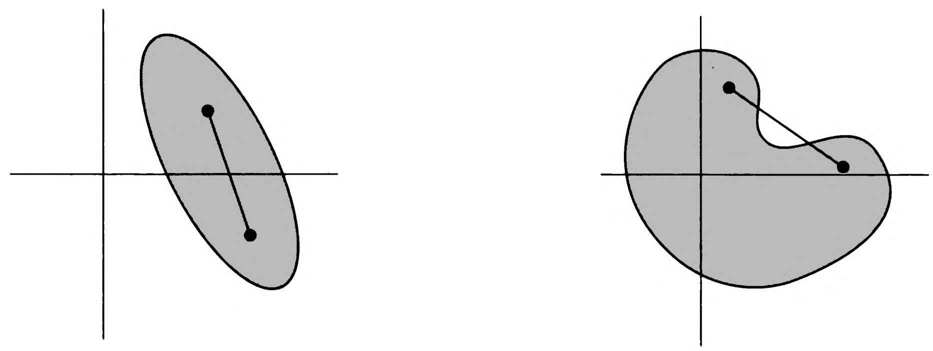
5. (a) Suppose that f is differentiable and convex on an interval. Show that either f is increasing, or else f is decreasing, or else there is a number c such that f is decreasing to the left of c and increasing to the right of c.
(b) Use this fact to give another proof of the result in Problem 4(a) when f and g are (one-time) differentiable. (You will have to be a little careful when comparing $f'(g(x))$ and $f'(g(y))$ for $x < y$.)
(c) Prove the result in part (a) without assuming f differentiable. You will have to keep track of several different cases, but no particularly clever ideas are needed. Begin by showing that if $a < b$ and $f(a) < f(b)$, then f is increasing to the right of b; and if $f(a) > f(b)$, then f is decreasing to the left of a.

***6.** Let f be a twice-differentiable function with the following properties: $f(x) > 0$ for $x \geq 0$, and f is decreasing, and $f'(0) = 0$. Prove that $f''(x) = 0$ for some $x > 0$ (so that in reasonable cases f will have an inflection point at x—an example is given by $f(x) = 1/(1+x^2)$). Every hypothesis in this theorem is essential, as shown by $f(x) = 1 - x^2$, which is not positive for all x; by $f(x) = x^2$, which is not decreasing; and by $f(x) = 1/(x+1)$, which does not satisfy $f'(0) = 0$. Hint: Choose $x_0 > 0$ with $f'(x_0) < 0$. We cannot have $f'(y) \leq f'(x_0)$ for all $y > x_0$. Why not? So $f'(x_1) > f'(x_0)$ for some $x_1 > x_0$. Consider f' on $[0, x_1]$.

***7.** (a) Prove that if f is convex, then $f([x+y]/2) < [f(x) + f(y)]/2$.
(b) Suppose that f satisfies this condition. Show that $f(kx + (1-k)y) < kf(x) + (1-k)f(y)$ whenever k is a rational number, between 0 and 1, of the form $m/2^n$. Hint: Part (a) is the special case $n = 1$. Use induction, employing part (a) at each step.
(c) Suppose that f satisfies the condition in part (a) and f is continuous. Show that f is convex.

12. Find two convex functions f and g such that $f(x) = g(x)$ if and only if x is an integer. Hint: First find an example where g is merely weakly convex, and then modify it, using the result of Problem 9 as a guide.

13. A set A of points in the plane is called *convex* if A contains the line segment joining any two points in it (Figure 14). For a function f, let A_f be the set of points (x, y) with $y \geq f(x)$, so that A_f is the set of points on or above the graph of f. Show that A_f is convex if and only if f is weakly convex, in the terminology of the previous problem. Further information on convex sets will be found in reference [18] of the Suggested Reading.



(a) a convex subset of the plane (b) a non-convex subset of the plane

FIGURE 14

***8.** For $n > 1$, let $p_1, \dots, p_n$ be *positive* numbers with $\sum_{i=1}^{n} p_i = 1$.

(a) For any numbers $x_1, \dots, x_n$ show that $\sum_{i=1}^{n} p_i x_i$ lies between the smallest and the largest x_i.

(b) Show the same for $(1/t)\sum_{i=1}^{n-1} p_i x_i$, where $t = \sum_{i=1}^{n-1} p_i$.

(c) Prove *Jensen's inequality:* If f is convex, then $f\left(\sum_{i=1}^{n} p_i x_i\right) < \sum_{i=1}^{n} p_i f(x_i)$. Hint: Use Problem 3, noting that $p_n = 1 - t$. (Part (b) is needed to show that $(1/t)\sum_{i=1}^{n-1} p_i x_i$ is in the domain of f if $x_1, \dots, x_n$ are.)

***9.** (a) For any function f, the right-hand derivative, $\lim_{h\to 0^+} [f(a+h) - f(a)]/h$, is denoted by $f'_+(a)$, and the left-hand derivative is denoted by $f'_-(a)$. The proof of Theorem 1 actually shows that $f'_+(a)$ and $f'_-(a)$ always exist if f is convex on some open interval containing a. Check this assertion, and also show that f'_+ and f'_- are increasing, and that $f'_-(a) \le f'_+(a)$.

(b) Conversely, suppose that f is convex on $[a, b]$ and g is convex on $[b, c]$, with $f(b) = g(b)$ and $f'_-(b) \le g'_+(b)$ (Figure 13 (a)). If we define h on $[a, c]$ to be f on $[a, b]$ and g on $[b, c]$, show that h is convex on $[a, c]$. Hint: Given P and Q on opposite sides of $O = (b, f(b))$, as in Figure 13 (b), compare the slope of OQ with that of PO.

(c) Show that if f is convex, then $f'_+(a) = f'_-(a)$ if and only if f'_+ is continuous at a. (Thus f is differentiable precisely when f'_+ is continuous.) Hint: $[f(b) - f(a)]/(b - a)$ is close to $f'_-(a)$ for $b < a$ close to a, and $f'_+(b)$ is less than this quotient.

***10.** (a) Prove that a convex function on $\mathbf{R}$, or on any open interval, must be continuous.

(b) Give an example of a convex function on a closed interval that is *not* continuous, and explain exactly what kinds of discontinuities are possible.

11. Call a function f *weakly convex* on an interval if for $a < b < c$ in this interval we have

$$\frac{f(x) - f(a)}{x - a} \le \frac{f(b) - f(a)}{b - a}.$$

(a) Show that a weakly convex function is convex if and only if its graph contains no straight line segments. (Sometimes a weakly convex function is simply called "convex," while convex functions in our sense are called "strictly convex".)

(b) Reformulate the theorems of this section for weakly convex functions.

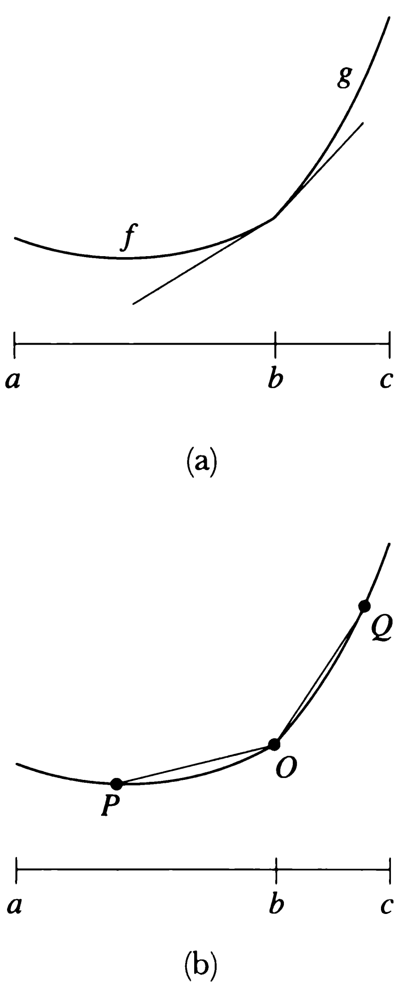


FIGURE 13