Difference Equations
to
Differential Equations

Section 3.7

Rolle's Theorem and the Mean Value Theorem

The two theorems which are at the heart of this section draw connections between the instantaneous rate of change and the average rate of change of a function. The Mean Value Theorem, of which Rolle's Theorem is a special case, says that if f is differentiable on an interval, then there is some point in that interval at which the instantaneous rate of change of the function is equal to the average rate of change of the function over the entire interval. For example, if f gives the position of an object moving in a straight line, the Mean Value Theorem says that if the average velocity over some interval of time is 60 miles per hour, then at some time during that interval the object was moving at exactly 60 miles per hour. This is not a surprising fact, but it does turn out to be the key to understanding many useful applications.

Before we turn to a consideration of Rolle's theorem, we need to establish another fundamental result. Suppose an object is thrown vertically into the air so that its position at time t is given by $f(t)$ and its velocity by $v(t) = f'(t)$. Moreover, suppose it reaches its maximum height at time t_0. On its way up, the object is moving in the positive direction, and so $v(t) > 0$ for $t < t_0$; on the way down, the object is moving in the negative direction, and so $v(t) < 0$ for $t > t_0$. It follows, by the Intermediate Value Theorem and the fact that v is a continuous function, that we must have $v(t_0) = 0$. That is, at time t_0, when $f(t)$ reaches its maximum value, we have $f'(t_0) = 0$. This is an extremely useful fact which holds in general for differentiable functions, not only at maximum values but at minimum values as well. Before providing a general demonstration, we first need a few definitions.

Definition A function f is said to have a *local maximum* at a point c if there exists an open interval I containing c such that $f(c) \geq f(x)$ for all x in I. A function f is said to have a *local minimum* at a point c if there exists an open interval I containing c such that $f(c) \leq f(x)$ for all x in I. If f has either a local maximum or a local minimum at c, then we say f has a *local extremum* at c.

In short, f has a local maximum at a point c if the value of f at c is at least as large as the value of f at any nearby point, and f has a local minimum at a point c if the value of f at c is at least as small as the value of f at any nearby point. The next example provides an illustration.

Example Looking at the graph of the function $f(x) = x^3 - 3x$ in Figure 3.7.1, it appears that f has a local maximum of 2 at $x = -1$ and a local minimum of -2 at $x = 1$. We will confirm this observation in Section 3.8.

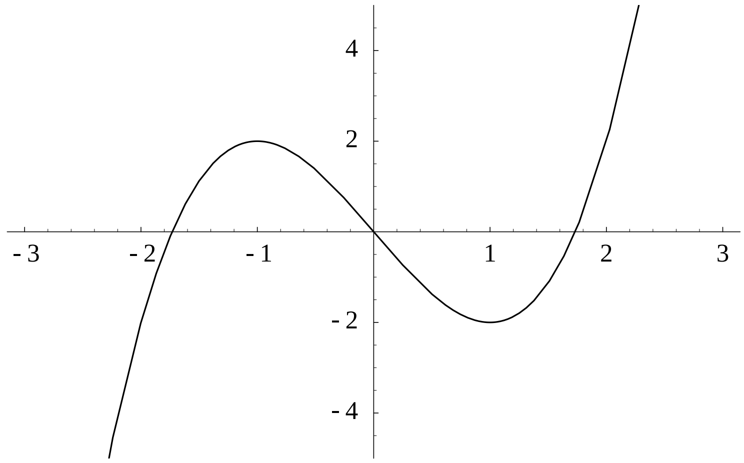

Figure 3.7.1 Graph of $f(x) = x^3 - 3x$

Now suppose f has local maximum at a point c and suppose f is differentiable at c. For small enough $h > 0$, $f(c+h) \leq f(c)$, so

$$f(c+h) - f(c) \leq 0;$$

thus

$$\frac{f(c+h) - f(c)}{h} \leq 0. \tag{3.7.1}$$

Clearly, if each term in a sequence is less than or equal to 0, and the sequence has a limit, then the limit of the sequence must be less than or equal to 0. Hence

$$f'(c) = \lim_{h \to 0^+} \frac{f(c+h) - f(c)}{h} \leq 0. \tag{3.7.2}$$

Also, for $h < 0$ with $|h|$ small enough, we have $f(c+h) \leq f(c)$, and so

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

However, now, since $h < 0$, we have

$$\frac{f(c+h) - f(c)}{h} \geq 0, \tag{3.7.3}$$

from which it follows that

$$f'(c) = \lim_{h \to 0^-} \frac{f(c+h) - f(c)}{h} \geq 0. \tag{3.7.4}$$

Note that (3.7.1) is saying that secant lines to the right of a local maximum have negative slope, while (3.7.3) is saying that secant lines to the left of a local maximum have positive slope. Now the only way that both (3.7.2) and (3.7.4) can hold at the same time is if $f'(c) = 0$; that is, the only number which is both less than or equal to 0 and greater than

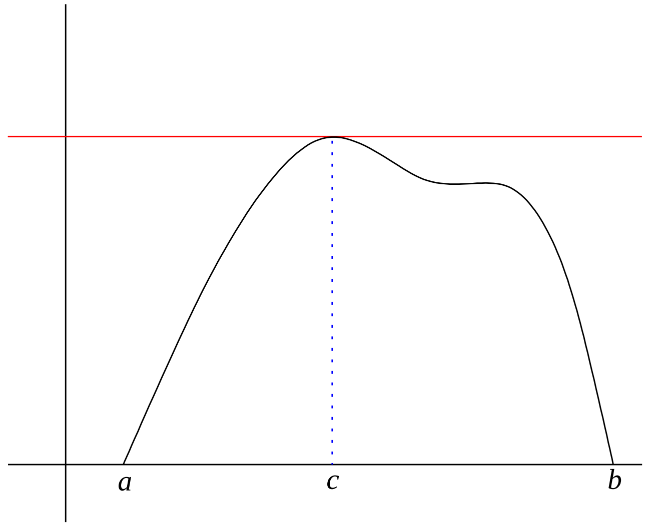

Figure 3.7.2 Illustration of Rolle's Theorem

or equal to 0 is 0 itself. Note that our argument here is just a refinement of our comments about velocity in the previous paragraph.

A similar argument gives the same result if f has a local minimum at c. The following proposition puts these together into one statement.

Proposition If f has a local extremum at c and f is differentiable at c, then $f'(c) = 0$.

Example For $f(x) = x^3 - 3x$, as in the previous example, $f'(x) = 3x^2 - 3$, and so $f'(-1) = 0$ and $f'(1) = 0$, consistent with our observation that f has a local maximum at $x = -1$ and local minimum at $x = 1$. Note, however, that this does not prove that f has local extrema at $x = -1$ and $x = 1$. Indeed, the proposition works in the other direction: if f has a local extremum at c, then $f'(c) = 0$.

This result will be very useful in our work in the next section when we consider the problem of finding the maximum and minimum values of a given function. For our present purpose, consider a function f which is continuous on a closed interval $[a, b]$ and differentiable on the open interval (a, b), with $f(a) = f(b) = 0$. An example of such a function is shown in Figure 3.7.2. By the Extreme Value Theorem of Section 2.5, we know that f must have both a minimum value m and a maximum value M on $[a, b]$. If $m = M = 0$, then $f(x) = 0$ for all x in (a, b), and so $f'(x) = 0$ for all x in (a, b). If either $m \neq 0$ or $M \neq 0$, then f has a local extremum at some point c in (a, b), namely, either a point c for which $f(c) = m$ or a point c for which $f(c) = M$. Hence, by the previous proposition, $f'(c) = 0$. We have thus established the following theorem, credited originally to Michel Rolle (1652-1719).

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

Put another way, Rolle's theorem says that if f is a differentiable function, then between any two solutions of the equation $f(x) = 0$ there is a point c where $f'(c) = 0$. Used

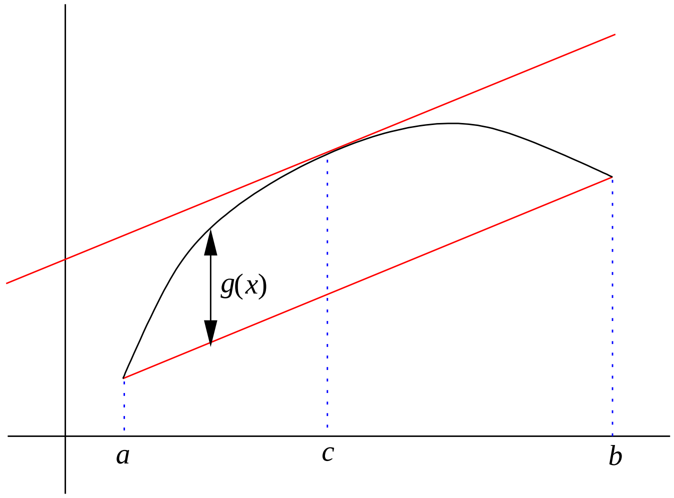

Figure 3.7.3 Illustration of the Mean Value Theorem

in conjunction with the Intermediate Value Theorem, this result can help identify intervals where an equation has a unique solution.

Example Solving the equation

$$x^5 + x^4 = 1 \tag{3.7.5}$$

is equivalent to solving the equation $f(x) = 0$ where $f(x) = x^5 + x^4 - 1$. Since $f(0) = -1$ and $f(1) = 1$, the Intermediate Value Theorem tells us that $f(x) = 0$ has at least one solution in $(0, 1)$. Moreover,

$$f'(x) = 5x^4 + 4x^3,$$

so $f'(x) > 0$ for all x in $(0, 1)$; in particular, there does not exist a point c in $(0, 1)$ such that $f'(c) = 0$. Hence, by Rolle's Theorem, there cannot be two solutions to $f(x) = 0$ in $(0, 1)$. That is, using the Intermediate Value Theorem and Rolle's Theorem together, we are able to conclude that there is exactly one solution to (3.7.5) in the interval $(0, 1)$. We may now use either the bisection algorithm or Newton's method to locate this solution.

Geometrically, Rolle's theorem says if f is a function which satisfies the conditions of the theorem on an interval $[a, b]$, then there is a point c in (a, b) such that the line tangent to the graph of f at $(c, f(c))$ is horizontal. In this case, that means that the line tangent to the graph of f at $(c, f(c))$ is parallel to the line passing through the points $(a, f(a))$ and $(b, f(b))$, as is seen in Figure 3.7.2. Certainly, if we took this picture and rotated or shifted the points $(a, f(a))$ and $(b, f(b))$, rigidly moving the graph with these points, then this conclusion would still follow. That is, if f is continuous on the closed interval $[a, b]$ and differentiable on (a, b), then there must exist a point c in (a, b) such that the line tangent to the graph of f at $(c, f(c))$ is parallel to the line passing through the points $(a, f(a))$ and $(b, f(b))$ (see Figure 3.7.3). In other words, there must be a point c in (a, b) such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{3.7.6}$$

This is the content of the Mean Value Theorem.

Although the above argument for this result seems plausible, we will present a more precise argument. Define a new function g by

$$g(x) = f(x) - S(x), \tag{3.7.7}$$

where

$$S(x) = \left(\frac{f(b) - f(a)}{b - a}\right)(x - a) + f(a). \tag{3.7.8}$$

Geometrically, the graph of S is a line passing through the points $(a, f(a))$ and $(b, f(b))$, and $g(x)$ is the distance from the graph of f to the graph of S above the point x (see Figure 3.7.3). Now g is continuous on $[a, b]$ and differentiable on (a, b); moreover,

$$g(a) = f(a) - S(a) = f(a) - f(a) = 0 \tag{3.7.9}$$

and

$$g(b) = f(b) - S(b) = f(b) - f(b) = 0. \tag{3.7.10}$$

Thus g satisfies the conditions of Rolle's theorem. Hence there exists a point c in (a, b) such that $g'(c) = 0$. But

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}, \tag{3.7.11}$$

and so $g'(c) = 0$ implies

$$f'(c) - \frac{f(b) - f(a)}{b - a} = 0, \tag{3.7.12}$$

that is,

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{3.7.13}$$

Mean Value Theorem If f is continuous on $[a, b]$ and differentiable on (a, b), then there exists a point c in (a, b) such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{3.7.14}$$

Increasing and decreasing functions

Similar to the situation with the Intermediate Value Theorem and the Extreme Value Theorem, the Mean Value Theorem is an existence theorem. The point of interest is the existence of c, not in being able to compute a value for c. Although not immediately useful for computations, we will see that the Mean Value Theorem has many important consequences. The first of these, which we will consider now, involves determining when a function is increasing and when it is decreasing.

Definition We say a function f defined on an interval I is *increasing* on I if for every two points u and v in I with $u < v$, $f(u) < f(v)$. We say a function f defined on an interval I is *decreasing* on I if for every two points u and v in I with $u < v$, $f(u) > f(v)$.

Example The function $f(x) = x^2$ is increasing on the interval $[0, \infty)$ since for any two numbers u and v with $0 \leq u < v$, $f(u) = u^2 < v^2 = f(v)$. Moreover, f is decreasing on $(-\infty, 0]$ since for any two numbers u and v with $u < v \leq 0$, $f(u) = u^2 > v^2 = f(v)$.

Now suppose f is differentiable on an interval (a, b) with $f'(x) > 0$ for all x in (a, b). If u and v are two points in (a, b) with $u < v$, then, by the Mean Value Theorem, there exists a point c with $u < c < v$ such that

$$f'(c) = \frac{f(v) - f(u)}{v - u}. \tag{3.7.15}$$

Since c is in (a, b), $f'(c) > 0$, so, using (3.7.15),

$$f(v) - f(u) = f'(c)(v - u) > 0. \tag{3.7.16}$$

Hence $f(v) > f(u)$ and f is increasing on the interval (a, b). Similarly, if $f'(x) < 0$ for all x in (a, b), then we would have $f'(c) < 0$, from which it would follow that $f(v) > f(u)$ and, hence, that f is decreasing on (a, b). In short, to determine the intervals on which a differentiable function is increasing and those on which it is decreasing, we need to look only for the intervals on which the derivative is positive and those on which it is negative, respectively.

Proposition If f is differentiable on (a, b) and $f'(x) > 0$ for all x in (a, b), then f is increasing on (a, b). If f is differentiable on (a, b) and $f'(x) < 0$ for all x in (a, b), then f is decreasing on (a, b).

Geometrically, this proposition is saying that a function is increasing where it has positive slope and decreasing where it has negative slope. This should seem intuitively clear, but it is the Mean Value Theorem which makes the connection between average rates of change and instantaneous rates of change necessary for establishing the result.

Example Suppose $f(x) = 2x^3 + 3x^2 - 12x + 1$. To determine where f is increasing and where it is decreasing, we first find

$$f'(x) = 6x^2 + 6x - 12 = 6(x + 2)(x - 1). \tag{3.7.17}$$

Hence $f'(x) = 0$ only when $x = -2$ or $x = 1$. Since f' is continuous, the Intermediate Value Theorem implies that f' cannot change sign on the intervals $(-\infty, -2)$, $(-2, 1)$, and $(1, \infty)$. Since $f'(-3) = 24 > 0$, it follows that $f'(x) > 0$ for all x in $(-\infty, -2)$. Similarly, since $f'(0) = -12 < 0$, $f'(x) < 0$ for all x in $(-2, 1)$; and, since $f'(2) = 24 > 0$, $f'(x) > 0$ for all x in $(1, \infty)$. It now follows from the previous proposition that f is increasing on the intervals $(-\infty, -2)$ and $(1, \infty)$ and decreasing on the interval $(-2, 1)$.

Note that we could obtain the same information about f' directly from (3.7.17) without evaluating f' and without invoking the Intermediate Value Theorem. Namely, from the facts that $x + 2 < 0$ and $x - 1 < 0$ whenever $x < -2$, we may conclude from the (3.7.17) that $f'(x) > 0$ for all x in $(-\infty, -2)$. Similarly, whenever $-2 < x < 1$, we have $x + 2 > 0$ and $x - 1 < 0$, implying that $f'(x) < 0$ for all x in $(-2, 1)$; and whenever $x > 1$ we have $x + 2 > 0$ and $x - 1 > 0$, implying that $f'(x) > 0$ for all x in $(1, \infty)$.

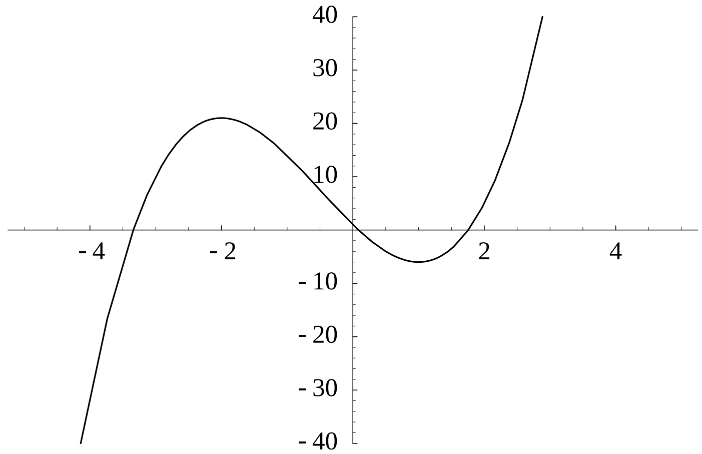

Figure 3.7.4 Graph of $f(x) = 2x^3 + 3x^2 - 12x + 1$

Combining our information on intervals where f is increasing and intervals where f is decreasing with the facts that $f(-3) = 10$, $f(-2) = 21$, $f(0) = 1$, $f(1) = -6$, $f(2) = 5$,

$$\lim_{x \to -\infty} f(x) = \lim_{x \to -\infty} x^3 \left(2 + \frac{3}{x} - \frac{12}{x^2} + \frac{1}{x^3}\right) = -\infty,$$

and

$$\lim_{x \to \infty} f(x) = \lim_{x \to \infty} x^3 \left(2 + \frac{3}{x} - \frac{12}{x^2} + \frac{1}{x^3}\right) = \infty,$$

we can understand why the graph of f looks as it does in Figure 3.7.4.

Example Now consider $f(x) = x^5 - x^3$. Then

$$f'(x) = 5x^4 - 3x^2 = x^2(5x^2 - 3), \tag{3.7.18}$$

so $f'(x) = 0$ when

$$x = -\sqrt{\frac{3}{5}}, x = 0, \text{ or } x = \sqrt{\frac{3}{5}}.$$

Now when

$$x < -\sqrt{\frac{3}{5}},$$

both $x^2 > 0$ and $5x^2 - 3 > 0$, implying, from (3.7.18), that $f'(x) > 0$. For

$$-\sqrt{\frac{3}{5}} < x < 0,$$

$x^2 > 0$, but $5x^2 - 3 < 0$, so $f'(x) < 0$; for

$$0 < x < \sqrt{\frac{3}{5}},$$

$x^2 > 0$ and $5x^2 - 3 < 0$, so $f'(x) < 0$; and for

$$x > \sqrt{\frac{3}{5}},$$

$x^2 > 0$ and $5x^2 - 3 > 0$, so $f'(x) > 0$. Hence f is increasing on

$$\left(-\infty, -\sqrt{\frac{3}{5}}\right)$$

and

$$\left(\sqrt{\frac{3}{5}}, \infty\right),$$

and decreasing on

$$\left(-\sqrt{\frac{3}{5}}, 0\right)$$

and

$$\left(0, \sqrt{\frac{3}{5}}\right).$$

Note that we could have determined the sign of f' on these four intervals by evaluating f' at a point in each interval and then applying the Intermediate value Theorem. For example,

$$f'(-1) = 2 > 0,$$

$$f'\left(-\frac{1}{5}\right) = -\frac{14}{125} < 0,$$

$$f'\left(\frac{1}{5}\right) = -\frac{14}{125} < 0,$$

and

$$f'(1) = 2 > 0.$$

As in the previous example, if we combine this information with the facts

$$f(-1) = 0,$$

$$f\left(-\sqrt{\frac{3}{5}}\right) = \frac{6}{25}\sqrt{\frac{3}{5}},$$

$$f(0) = 0,$$

$$f\left(\sqrt{\frac{3}{5}}\right) = -\frac{6}{25}\sqrt{\frac{3}{5}},$$

$$f(1) = 0,$$

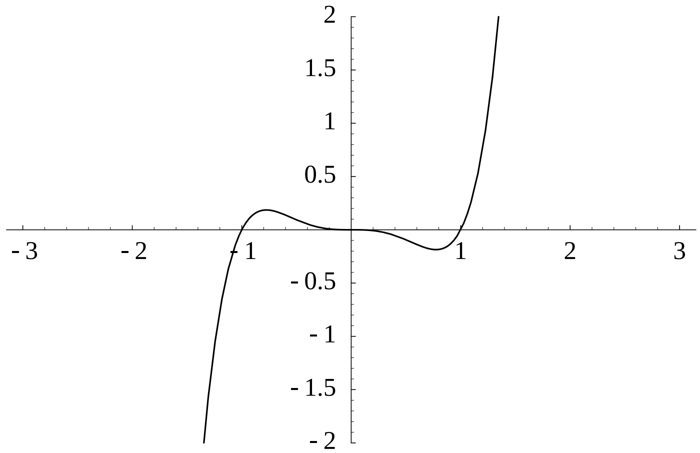

Figure 3.7.5 Graph of $f(x) = x^5 - x^3$

$$\lim_{x \to -\infty} f(x) = \lim_{x \to -\infty} x^5 \left(1 - \frac{1}{x^2}\right) = -\infty,$$

and

$$\lim_{x \to \infty} f(x) = \lim_{x \to \infty} x^5 \left(1 - \frac{1}{x^2}\right) = \infty,$$

we can understand why the graph of f looks as it does in Figure 3.7.5.

Antiderivatives

We will close this section with a look at one more important application of the Mean Value Theorem. Although not needed for our current discussion, our result will be very useful in the next chapter. We begin with a definition.

Definition If F and f are functions defined on an open interval (a, b) such that $F'(x) = f(x)$ for all x in (a, b), then we call F an *antiderivative* of f.

In other words, an antiderivative of a function f is another function whose derivative is f. Although a given function f has at most one derivative, it is possible to have more than one antiderivative, as the next example demonstrates.

Example $F(x) = x^3$ is an antiderivative of $f(x) = 3x^2$ on $(-\infty, \infty)$. However, note that $G(x) = x^3 + 4$ is also an antiderivative of f. In fact, given any constant k,

$$H(x) = x^3 + k \tag{3.7.19}$$

is an antiderivative of f. This should not be too surprising since specifying the derivative of a function fixes only the slope of its graph, and the graphs of the functions in (3.7.19) are all in a sense parallel to each other.

This example shows that a given function may have an infinite number of antiderivatives. However, note that the difference of any two these antiderivatives is a constant. We will now show that this is always the case, and, in particular, that (3.7.19) specifies all possible antiderivatives of $f(x) = 3x^2$.

First consider a function F defined on an open interval (a, b) for which $F'(x) = 0$ for all x in (a, b). That is, F is an antiderivative of the function which is 0 for all values of x in (a, b). Now if u and v are any two points in (a, b), then, by the Mean Value Theorem,

$$\frac{F(v) - F(u)}{v - u} = F'(c) \tag{3.7.20}$$

for some c in (a, b). But $F'(c) = 0$, so

$$\frac{F(v) - F(u)}{v - u} = 0, \tag{3.7.21}$$

which implies $F(u) = F(v)$. If we let $k = F(u)$ for a fixed u in (a, b), we now have $F(v) = F(u) = k$ for all v in (a, b). In other words, if the derivative of a function is 0 on an open interval, then the function must be constant on that interval.

Now suppose F and G are two functions defined on an open interval (a, b) for which $F'(x) = G'(x)$ for all x in (a, b). Let $H(x) = F(x) - G(x)$ for all x in (a, b). Then

$$H'(x) = F'(x) - G'(x) = 0 \tag{3.7.22}$$

for all x in (a, b). Then, by what we have just shown, there exists a constant k for which

$$k = H(x) = F(x) - G(x) \tag{3.7.23}$$

for all x in (a, b). Hence if F an G have the same derivative on an open interval, that is, are antiderivatives of the same function, then they can differ only by a constant.

Proposition If F and G are both antiderivatives of f on an open interval (a, b), then there exists an constant k such that

$$F(x) = G(x) + k \tag{3.7.24}$$

for all x in (a, b).

Example Since

$$\frac{d}{dx}\sin(x) = \cos(x),$$

we know that $G(x) = \sin(x)$ is an antiderivative of $f(x) = \cos(x)$ on $(-\infty, \infty)$. Thus if F is any antiderivative of f, then

$$F(x) = \sin(x) + k \tag{3.7.25}$$

for some constant k. In other words, functions of the form given in (3.7.25) are the only antiderivatives of $f(x) = \cos(x)$. Figure 3.7.6 shows the graphs of (3.7.25) for nine different values of k. Although each of these graphs is the graph of a different function, they are

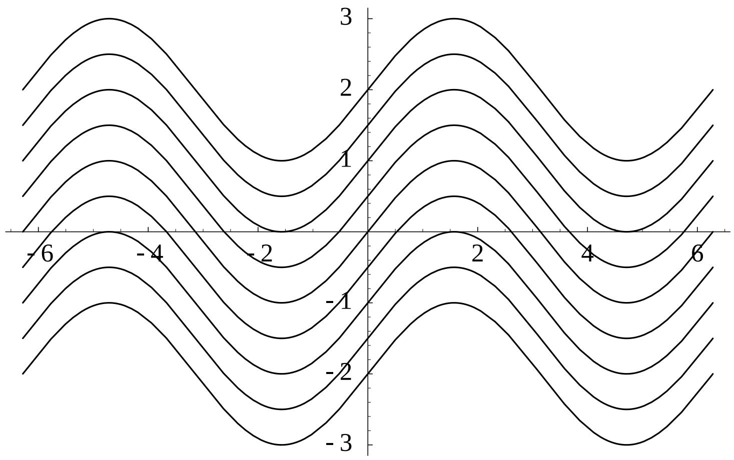

Figure 3.7.6 Graphs of $F(x) = \sin(x) + k$ for different values of k

parallel to one another in the sense that they all have the same slope at any given value of x, namely, $\cos(x)$.

Problems

1. Explain why the equation $\cos(x) = x$ has exactly one solution in the interval $[0, 1]$.
2. Explain why the equation $x^4 - 2x^2 = 2$ has exactly one solution in the interval $[1, 2]$ and exactly one solution in the interval $[-2, -1]$.
3. Suppose f is continuous on $[a, b]$ and differentiable on (a, b). Moreover, suppose there is a constant M such that $|f'(x)| \leq M$ for all x in (a, b). Show that
$$|f(v) - f(u)| \leq M|v - u|$$
for all u and v in $[a, b]$.
4. Use Problem 3 to show that $|\sin(x) - \sin(y)| \leq |x - y|$ for all values of x and y.
5. For each of the following functions, identify the intervals where the function is increasing and where it is decreasing. Use this information to sketch the graph.
   (a) $f(x) = x^2 - 3$
   (b) $g(t) = 3t^2 + t - 6$
   (c) $h(z) = \dfrac{1}{z - 1}$
   (d) $f(x) = \dfrac{1}{x^2 + 1}$
   (e) $f(t) = \dfrac{t}{t^2 + 1}$
   (f) $g(x) = x^4 - x^3$
   (g) $y(t) = \dfrac{t}{t^2 - 4}$
   (h) $f(x) = 4x^5 - 15x^4 - 20x^3 + 110x^2 - 120x$
6. Let $f(x) = |x|$. Then
$$\frac{f(1) - f(-1)}{1 - (-1)} = \frac{0}{2} = 0,$$

but there does not exist a point c in $(-1, 1)$ such that $f'(c) = 0$. Does this contradict the Mean Value Theorem?

7. Suppose f is continuous on $[a, b]$ and differentiable on (a, b).
   (a) Show that if $f'(x) > 0$ for all x in (a, b), then f is increasing on $[a, b]$.
   (b) Show that if $f'(x) < 0$ for all x in (a, b), then f is decreasing on $[a, b]$.

8. Suppose f and g are continuous on $[a, b]$, differentiable on (a, b), $f(a) = g(a)$, and $f'(x) < g'(x)$ for all x in (a, b). Show that $f(b) < g(b)$.

9. Show that $\sqrt{1+x} < 1 + \frac{1}{2}x$ for $x > 0$.

10. Show that $a + \frac{1}{a} < b + \frac{1}{b}$ whenever $1 < a < b$.

11. Find antiderivatives for the following functions.
    (a) $f(x) = 2x$
    (b) $g(t) = t^2$
    (c) $g(x) = \sin(x)$
    (d) $f(z) = \sin(2z)$
    (e) $h(x) = x^2 - 3x$
    (f) $f(x) = 3\cos(4x)$

12. Find all antiderivatives of $f(x) = 3x^2 - 3$ and plot the graphs of six of them.

13. Find all antiderivatives of $g(t) = \sin(2t)$ and plot the graphs of six of them.

14. If $f(x) = -\sin^2(x)$ and $g(x) = \cos^2(x)$, then $f'(x) = g'(x)$. What does this imply about the relationship between the functions f and g?

15. If $f(t) = \tan^2(t)$ and $g(t) = \sec^2(t)$, then $f'(t) = g'(t)$. Thus $f(t) = g(t) + k$ for some constant k. Evaluate f and g at $t = 0$ in order to determine k.

16. Suppose f is differentiable on an open interval containing the closed interval $[a, b]$.
    (a) Show that for any x in (a, b),
    $$f(x) = f(a) + f'(c)(x - a)$$
    for some point c with $a < c < x$.
    (b) Let f'' denote the *second derivative* of f. That is,
    $$f''(x) = \frac{d}{dx} f'(x).$$
    Assuming that f' is continuous on $[a, b]$ and differentiable on (a, b), show that there exists a point d with $a < d < c$ such that
    $$f(x) = f(a) + f'(a)(x - a) + f''(d)(c - a)(x - a).$$
    (c) Compare the results in (a) and (b) to the statement that $f(x) \approx T(x)$ for x close to a, where T is the best affine approximation to f at a.

(d) Let $h = x - a$. Show that

$$f(a+h) - T(a+h) = f''(d)(c-a)h.$$

(e) Assuming f'' is continuous on $[a, b]$, show that

$$\left| \frac{f(a+h) - T(a+h)}{h^2} \right| \leq M$$

for some constant M. This statement means that the remainder function

$$R(h) = f(a+h) - T(a+h)$$

is $O(h^2)$. That is, $R(h)$ goes to 0 as least as fast as h^2. Note that this is a stronger statement than the statement that $R(h)$ is $o(h)$.