(d) Let $I := (1, \infty)$ and consider $\lim_{x \to \infty} (1 + 1/x)^x$, which has the indeterminate form 1^∞. We note that

$$(*) \qquad (1 + 1/x)^x = e^{x \log (1 + 1/x)}.$$

Moreover, we have

$$\lim_{x \to \infty} x \log (1 + 1/x) = \lim_{x \to \infty} \frac{\log (1 + 1/x)}{1/x}$$

$$= \lim_{x \to \infty} \frac{(1 + 1/x)^{-1}(-x^{-2})}{-x^{-2}} = \lim_{x \to \infty} \frac{1}{1 + 1/x} = 1.$$

Since $y \mapsto e^y$ is continuous at $y = 1$, we infer that $\lim_{x \to \infty} (1 + 1/x)^x = e$.

(e) Let $I := (0, \infty)$ and consider $\lim_{x \to 0+} (1 + 1/x)^x$, which has the indeterminate form ∞^0. In view of formula $(*)$, we consider

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

Therefore we have $\lim_{x \to 0+} (1 + 1/x)^x = e^0 = 1$.

Exercises for Section 6.3

1. Suppose that f and g are continuous on $[a, b]$, differentiable on (a, b), that $c \in [a,b]$ and that $g(x) \neq 0$ for $x \in [a, b]$, $x \neq c$. Let $A := \lim_{x \to c} f$ and $B := \lim_{x \to c} g$. If $B = 0$, and if $\lim_{x \to c} f(x)/g(x)$ exists in $\mathbf{R}$, show that we must have $A = 0$. [*Hint:* $f(x) = \{f(x)/g(x)\}g(x)$.]
2. In addition to the suppositions of the preceding exercise, let $g(x) > 0$ for $x \in [a, b]$, $x \neq c$. If $A > 0$ and $B = 0$, prove that we must have $\lim_{x \to c} f(x)/g(x) = \infty$. If $A < 0$ and $B = 0$, prove that we must have $\lim_{x \to c} f(x)/g(x) = -\infty$.
3. Let $f(x) := x^2 \sin (1/x)$ for $0 < x \leqslant 1$ and $f(0) := 0$, and let $g(x) := x^2$ for $x \in [0, 1]$. Then both f and g are differentiable on $[0, 1]$ and $g(x) > 0$ for $x \neq 0$. Show that $\lim_{x \to 0} f(x) = 0 = \lim_{x \to 0} g(x)$ and that $\lim_{x \to 0} f(x)/g(x)$ does not exist.
4. Let $f(x) := x^2$ for x rational, let $f(x) := 0$ for x irrational, and let $g(x) := \sin x$ for $x \in \mathbf{R}$. Use Theorem 6.3.1 to show that $\lim_{x \to 0} f(x)/g(x) = 0$. Explain why Theorem 6.3.3 cannot be used.
5. Let $f(x) := x^2 \sin (1/x)$ for $x \neq 0$, let $f(0) := 0$, and let $g(x) := \sin x$ for $x \in \mathbf{R}$. Show that $\lim f(x)/g(x) = 0$ but that $\lim f'(x)/g'(x)$ does not exist.

6. Evaluate the following limits, where the domain of the quotient is as indicated.
   (a) $\displaystyle\lim_{x \to 0+} \frac{\log (x + 1)}{\sin x}$ $(0, \pi/2)$, (b) $\displaystyle\lim_{x \to 0+} \frac{\tan x}{x}$ $(0, \pi/2)$,
   (c) $\displaystyle\lim_{x \to 0+} \frac{\log \cos x}{x}$ $(0, \pi/2)$, (d) $\displaystyle\lim_{x \to 0+} \frac{\tan x - x}{x^3}$ $(0, \pi/2)$.
7. Evaluate the following limits:
   (a) $\displaystyle\lim_{x \to 0} \frac{\operatorname{Arctan} x}{x}$ $(-\infty, \infty)$, (b) $\displaystyle\lim_{x \to 0} \frac{1}{x(\log x)^2}$ $(0, 1)$,
   (c) $\displaystyle\lim_{x \to 0+} x^3 \log x$ $(0, \infty)$, (d) $\displaystyle\lim_{x \to \infty} \frac{x^3}{e^x}$ $(0, \infty)$.
8. Evaluate the following limits:
   (a) $\displaystyle\lim_{x \to \infty} \frac{\log x}{x^2}$ $(0, \infty)$, (b) $\displaystyle\lim_{x \to \infty} \frac{\log x}{\sqrt{x}}$ $(0, \infty)$,
   (c) $\displaystyle\lim_{x \to 0} x \log \sin x$ $(0, \pi)$, (d) $\displaystyle\lim_{x \to \infty} \frac{x + \log x}{x \log x}$ $(0, \infty)$.
9. Evaluate the following limits:
   (a) $\displaystyle\lim_{x \to 0+} x^{2x}$ $(0, \infty)$, (b) $\displaystyle\lim_{x \to 0} (1 + 3/x)^x$ $(0, \infty)$,
   (c) $\displaystyle\lim_{x \to \infty} (1 + 3/x)^x$ $(0, \infty)$, (d) $\displaystyle\lim_{x \to 0+} \left(\frac{1}{x} - \frac{1}{\operatorname{Arctan} x}\right)$ $(0, \infty)$.
10. Evaluate the following limits:
    (a) $\displaystyle\lim_{x \to \infty} x^{1/x}$ $(0, \infty)$, (b) $\displaystyle\lim_{x \to 0+} (\sin x)^x$ $(0, \pi)$,
    (c) $\displaystyle\lim_{x \to 0+} x^{\sin x}$ $(0, \infty)$, (d) $\displaystyle\lim_{x \to \pi/2-} (\sec x - \tan x)$ $(0, \pi/2)$.
11. Let f be differentiable on $(0, \infty)$ and suppose that $\lim_{x \to \infty} (f(x) + f'(x)) = L$. Show that $\lim_{x \to \infty} f(x) = L$ and $\lim_{x \to \infty} f'(x) = 0$. [*Hint:* $f(x) = e^x f(x)/e^x$.]
12. Try to use L'Hospital's Rule to find

$$\lim_{x \to \pi/2-} \frac{\tan x}{\sec x}.$$

Then evaluate directly by changing to sines and cosines.

SECTION 6.4 Taylor's Theorem

A very useful technique in the analysis of real functions is the approximation of functions by polynomials. In this section we shall prove a fundamental theorem in this area which goes back to Brook Taylor (1685–1731), although the remainder term was not provided until much later by Joseph-Louis Lagrange (1736–1813).

illustrate the versatility of Taylor's Theorem by briefly discussing some of its applications to numerical estimation, inequalities, extreme values of a function, and convex functions.

Taylor's Theorem can be regarded as an extension of the Mean Value Theorem to "higher order" derivatives. Whereas the Mean Value Theorem relates the values of a function and its first derivative, Taylor's Theorem provides a relation between the values of a function and its higher order derivatives.

Derivatives of order greater than one are obtained by a natural extension of the differentiation process. If the derivative $f'(x)$ of a function f exists at every point x in an interval I containing a point c, then we can consider the existence of the derivative of the function f' at the point c. In case f' has a derivative at the point c, we refer to the resulting number as the **second derivative** of f at c, and we denote this number by $f''(c)$ or by $f^{(2)}(c)$. In similar fashion we define the third derivative $f'''(c) = f^{(3)}(c), \ldots$, and the nth **derivative** $f^{(n)}(c)$, whenever these derivatives exist. It is noted that the existence of the nth derivative at c presumes the existence of the $(n-1)$st derivative in an interval containing c, but we do allow the possibility that c might be an end point of such an interval.

If a function f has an nth derivative at a point x_0, it is not difficult to construct an nth degree polynomial P_n such that $P_n(x_0) = f(x_0)$ and $P_n^{(k)}(x_0) = f^{(k)}(x_0)$ for $k = 1, 2, \ldots, n$. In fact, the polynomial

(1) $$P_n(x) := f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n,$$

has the property that it and its derivatives up to order n agree with the function f and its derivatives up to order n, at the specified point x_0. This polynomial P_n is called the nth **Taylor polynomial for** f **at** x_0. It is natural to expect this polynomial to provide a reasonable approximation to f for points near x_0, but to gauge the quality of the approximation, it is necessary to have information concerning the remainder $R_n := f - P_n$. The following fundamental result provides such information.

6.4.1 Taylor's Theorem *Let $n \in \mathbf{N}$, let $I := [a, b]$, and let $f: I \to \mathbf{R}$ be such that f and its derivatives $f', f'', \ldots, f^{(n)}$ are continuous on I and that $f^{(n+1)}$ exists on (a, b). If $x_0 \in I$, then for any x in I there exists a point c between x and x_0 such that*

(2) $$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}.$$

Proof. Let x_0 and x be given and let J denote the closed interval with end points x_0 and x. We define the function F on J by

$$F(t) := f(x) - f(t) - (x - t)f'(t) - \cdots - \frac{(x - t)^n}{n!}f^{(n)}(t)$$

for $t \in J$. Then an easy calculation shows that we have

$$F'(t) = -\frac{(x - t)^n}{n!}f^{(n+1)}(t).$$

If we define G on J by

$$G(t) := F(t) - \left(\frac{x - t}{x - x_0}\right)^{n+1} F(x_0)$$

for $t \in J$, then $G(x_0) = G(x) = 0$. An application of Rolle's Theorem 6.2.3 yields a point c between x and x_0 such that

$$0 = G'(c) = F'(c) + (n + 1)\frac{(x - c)^n}{(x - x_0)^{n+1}}F(x_0).$$

Hence, we obtain

$$F(x_0) = -\frac{1}{n+1}\frac{(x - x_0)^{n+1}}{(x - c)^n}F'(c)$$

$$= \frac{1}{n+1}\frac{(x - x_0)^{n+1}}{(x - c)^n}\frac{(x - c)^n}{n!}f^{(n+1)}(c) = \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1},$$

which implies the stated result. Q.E.D.

We shall use the notation P_n for the nth Taylor polynomial (1) of f, and R_n for the remainder. Thus we may write the conclusion of Taylor's Theorem as $f(x) = P_n(x) + R_n(x)$ where R_n is given by

(3) $$R_n(x) := \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}$$

for some point c between x and x_0. This formula for R_n is referred to as the **Lagrange form** (or the **derivative form**) of the remainder. Many other expressions for R_n are known; one is in terms of integration and will be given later. (See

Applications of Taylor's Theorem

The remainder term R_n in Taylor's Theorem can be used to estimate the error in approximating a function by its Taylor polynomial P_n. If the number n is prescribed, then the question of the accuracy of the approximation arises. On the other hand, if a certain accuracy is specified, then the question of finding a suitable value of n is germane. The following examples illustrate how one responds to these questions.

6.4.2 Examples (a) Use Taylor's Theorem with $n = 2$ to approximate $\sqrt[3]{1+x}$, $x > -1$.

We take the function $f(x) := (1+x)^{1/3}$, the point $x_0 = 0$, and $n = 2$. Since $f'(x) = \frac{1}{3}(1+x)^{-2/3}$ and $f''(x) = \frac{1}{3}(-\frac{2}{3})(1+x)^{-5/3}$, we have $f'(0) = \frac{1}{3}$ and $f''(0) = -2/9$. Thus we obtain

$$f(x) = P_2(x) + R_2(x) = 1 + \tfrac{1}{3}x - \tfrac{1}{9}x^2 + R_2(x),$$

where $R_2(x) = \dfrac{f'''(c)}{3!}x^3 = \dfrac{5}{81}(1+c)^{-8/3}x^3$ for some point c between 0 and x. If we let $x = 0.3$, we get the approximation $P_2(0.3) = 1.09$ for $\sqrt[3]{1.3}$. Moreover, since $c > 0$ in this case, then $(1+c)^{-8/3} < 1$ and so the error is at most

$$R_2(0.3) \leqslant \frac{5}{81}\left(\frac{3}{10}\right)^3 = \frac{1}{600} < 0.17 \times 10^{-2}.$$

Hence, we have $|\sqrt[3]{1.3} - 1.09| < 0.5 \times 10^{-2}$, so that two decimal place accuracy is assured.

(b) Approximate the number e with error less than 10^{-5}.

We shall consider the function $g(x) := e^x$ and take $x_0 = 0$ and $x = 1$ in Taylor's Theorem. We need to determine n so that $|R_n(1)| < 10^{-5}$. To do so, we shall use the fact that $g'(x) = e^x$ and the initial bound of $e^x \leqslant 3$ for $0 \leqslant x \leqslant 1$.

Since $g'(x) = e^x$, it follows that $g^{(k)}(x) = e^x$ for all $k \in N$, and therefore $g^{(k)}(0) = 1$ for all $k \in N$. Consequently the nth Taylor polynomial is given by

$$P_n(x) := 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$$

and the remainder for $x = 1$ is given by $R_n(1) = e^c/(n+1)!$ for some c satisfying $0 < c < 1$. Since $e^c < 3$, we seek a value of n such that $3/(n+1)! < 10^{-5}$. A calculation reveals that $9! = 362{,}880 > 3 \times 10^5$ so that the value $n = 8$ will provide the desired accuracy; moreover, since $8! = 40{,}320$, no smaller value of n

will be certain to suffice. Thus, we obtain

$$e \approx P_8(1) = 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{8!} = 2.718\,28$$

with error less than 10^{-5}.

Taylor's Theorem can also be used to derive inequalities.

6.4.3 Examples (a) $1 - \frac{1}{2}x^2 \leqslant \cos x$ for all $x \in R$.

Using $f(x) := \cos x$ and $x_0 = 0$ in Taylor's Theorem, we obtain

$$\cos x = 1 - \tfrac{1}{2}x^2 + R_2(x),$$

where for some c between 0 and x we have

$$R_2(x) = \frac{f'''(c)}{3!}x^3 = \frac{\sin c}{6}x^3.$$

If $0 \leqslant x \leqslant \pi$, then $0 \leqslant c < \pi$; since c and x^3 are both positive, we have $R_2(x) \geqslant 0$. Also, if $-\pi \leqslant x \leqslant 0$, then $-\pi \leqslant c \leqslant 0$; since $\sin c$ and x^3 are both negative, we again have $R_2(x) \geqslant 0$. Therefore, we see that $1 - \frac{1}{2}x^2 \leqslant \cos x$ for $|x| \leqslant \pi$. If $|x| \geqslant \pi$, then $1 - \frac{1}{2}x^2 < -3 \leqslant \cos x$ and the inequality is trivially valid. Hence, the inequality holds for all $x \in R$.

(b) For any $k \in N$, and for all $x > 0$, we have

$$x - \frac{1}{2}x^2 + \cdots - \frac{1}{2k}x^{2k} < \log(1+x) < x - \frac{1}{2}x^2 + \cdots + \frac{1}{2k+1}x^{2k+1}.$$

Using the fact that the derivative of $\log(1+x)$ is $1/(1+x)$ for $x > 0$, we see that the nth Taylor polynomial for $\log(1+x)$ with $x_0 = 0$ is

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

and the remainder is given by

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

for some c satisfying $0 < c < x$. Thus for any $x > 0$, if $n = 2k$ is even, then we have $R_{2k}(x) > 0$; and if $n = 2k+1$ is odd, then we have R

Relative Extrema

It was established in Theorem 6.2.1 that if a function $f: I \to \mathbf{R}$ is differentiable at a point c interior to the interval I, then a necessary condition for f to have a relative extremum at c is that $f'(c) = 0$. One way to determine whether f has a relative maximum or relative minimum [or neither] at c, is to use the First Derivative Test 6.2.8. Higher order derivatives, if they exist, can also be used in this determination, as we now show.

6.4.4 Theorem *Let I be an interval, let x_0 be an interior point of I, and let $n \geqslant 2$. Suppose that the derivatives $f', f'', \ldots, f^{(n)}$ exist and are continuous in a neighborhood of x_0 and that $f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0$, but $f^{(n)}(x_0) \neq 0$.*

(i) If n is even and $f^{(n)}(x_0) > 0$, then f has a relative minimum at x_0.

(ii) If n is even and $f^{(n)}(x_0) < 0$, then f has a relative maximum at x_0.

(iii) If n is odd, then f has neither a relative minimum nor relative maximum at x_0.

Proof. Applying Taylor's Theorem at x_0, we find that for $x \in I$ we have

$$f(x) = P_{n-1}(x) + R_{n-1}(x) = f(x_0) + \frac{f^{(n)}(c)}{n!}(x - x_0)^n,$$

where c is some point between x_0 and x. Since $f^{(n)}$ is continuous, if $f^{(n)}(x_0) \neq 0$, then there exists an interval U containing x_0 such that $f^{(n)}(x)$ will have the same sign as $f^{(n)}(x_0)$ for $x \in U$. If $x \in U$, then the point c also belongs to U and consequently $f^{(n)}(c)$ and $f^{(n)}(x_0)$ will have the same sign.

(i) If n is even and $f^{(n)}(x_0) > 0$, then for $x \in U$ we have $f^{(n)}(c) > 0$ and $(x - x_0)^n \geqslant 0$ so that $R_{n-1}(x) \geqslant 0$. Hence, $f(x) \geqslant f(x_0)$ for $x \in U$, and therefore f has a relative minimum at x_0.

(ii) If n is even and $f^{(n)}(x_0) < 0$, then it follows that $R_{n-1}(x) \leqslant 0$ for $x \in U$, so that $f(x) \leqslant f(x_0)$ for $x \in U$. Therefore, f has a relative maximum at x_0.

(iii) If n is odd, then $(x - x_0)^n$ is positive if $x > x_0$ and negative if $x < x_0$. Consequently, if $x \in U$, then $R_{n-1}(x)$ will have opposite signs to the left and to the right of x_0. Therefore, f has neither a relative minimum nor a relative maximum at x_0. Q.E.D.

Convex Functions

The notion of convexity plays an important role in a number of areas, particularly in the modern theory of optimization. We shall briefly look at convex functions of one real variable and their relation to differentiation. The basic results,

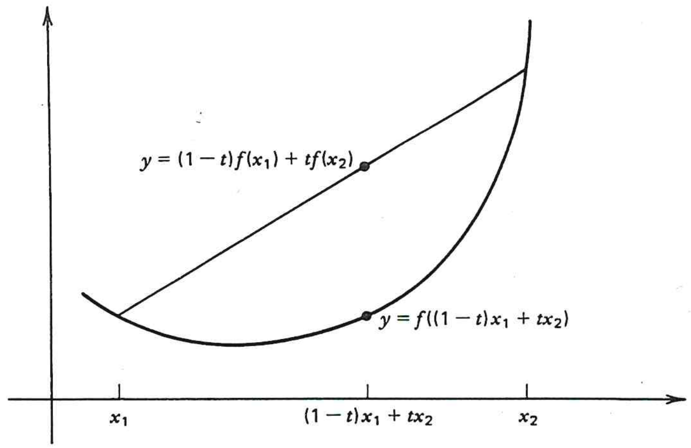

FIGURE 6.4.1 A convex function.

6.4.5 Definition Let $I \subseteq \mathbf{R}$ be an interval. A function $f: I \to \mathbf{R}$ is said to be **convex** on I if for any t satisfying $0 \leqslant t \leqslant 1$ and any points x_1, x_2 in I, we have

$$f((1-t)x_1 + tx_2) \leqslant (1-t)f(x_1) + tf(x_2).$$

Note that if $x_1 < x_2$, then as t ranges from 0 to 1, the point $(1-t)x_1 + tx_2$ traverses the interval from x_1 to x_2. Thus if f is convex on I and if $x_1, x_2 \in I$, then the chord joining any two points $(x_1, f(x_1))$ and $(x_2, f(x_2))$ on the graph of f lies above the graph of f. (See Figure 6.4.1.)

A convex function need not be differentiable at every point, as the example $f(x) := |x|$, $x \in \mathbf{R}$, reveals. However, it can be shown that if I is an open interval and if $f: I \to \mathbf{R}$ is convex on I, then the left and right derivatives of f exist at every point of I. As a consequence, it follows that a convex function on an open interval is necessarily continuous. We shall not verify the preceding assertions, nor shall we develop many other interesting properties of convex functions. Rather, we shall restrict ourselves to establishing the connection between a convex function f and its second derivative f'', assuming that f'' exists.

6.4.6 Theorem *Let I be an open interval and suppose that $f: I \to \mathbf{R}$ has a second derivative on I. Then f is a convex function on I if and only if $f''(x) \geqslant 0$ for all $x \in I$.*

Proof. To prove the necessity of the condition, we shall make use of the fact that the second derivative is given by the limit

$$(*) \qquad f''(a) = \lim \frac{f(a+h) - 2f(a) + f(a-h)}{h^2}$$

for each $a \in I$. (See Exercise 6.4.16.) Given $a \in I$, let h be such that $a + h$ and $a - h$ belong to I. Then $a = \frac{1}{2}((a+h) + (a-h))$, and since f is convex on I, we have

$$f(a) = f(\tfrac{1}{2}(a+h) + \tfrac{1}{2}(a-h)) \leqslant \tfrac{1}{2}f(a+h) + \tfrac{1}{2}f(a-h).$$

Therefore, we have $f(a+h) - 2f(a) + f(a-h) \geqslant 0$. Since $h^2 > 0$ for all $h \neq 0$, we see that the limit in $(*)$ must be nonnegative. Hence, we obtain $f''(a) \geqslant 0$ for any $a \in I$.

To prove the sufficiency of the condition we shall use Taylor's Theorem. Let x_1, x_2 be any two points of I, let $0 < t < 1$, and let $x_0 = (1-t)x_1 + tx_2$. Applying Taylor's Theorem to f at x_0 we obtain a point c_1 between x_0 and x_1 such that

$$f(x_1) = f(x_0) + f'(x_0)(x_1 - x_0) + \tfrac{1}{2}f''(c_1)(x_1 - x_0)^2,$$

and a point c_2 between x_0 and x_2 such that

$$f(x_2) = f(x_0) + f'(x_0)(x_2 - x_0) + \tfrac{1}{2}f''(c_2)(x_2 - x_0)^2.$$

If f'' is nonnegative on I, then the term

$$R := \tfrac{1}{2}(1-t)f''(c_1)(x_1 - x_0)^2 + \tfrac{1}{2}tf''(c_2)(x_2 - x_0)^2$$

is also nonnegative. Thus we obtain

$$\begin{aligned}(1-t)f(x_1) + tf(x_2) &= f(x_0) + f'(x_0)((1-t)x_1 + tx_2 - x_0) \\ &\quad + \tfrac{1}{2}(1-t)f''(c_1)(x_1 - x_0)^2 + \tfrac{1}{2}tf''(c_2)(x_2 - x_0)^2 \\ &= f(x_0) + R \\ &\geqslant f(x_0) = f((1-t)x_1 + tx_2).\end{aligned}$$

Hence, we see that f is a convex function on I. Q.E.D.

Newton's Method

It is often desirable to estimate a solution of an equation with a high degree of accuracy. The method of interval bisection, used in the proof of the Location of Roots Theorem 5.3.5, provides one estimation procedure, but it has the disadvantage of converging to a solution rather slowly. A method that often results in much more rapid convergence is based on the geometric idea of successively approximating a curve by tangent lines. The method is named after its discoverer, Isaac Newton.

Let f be a differentiable function that has a zero at r and let x_1 be an initial estimate of r. The line tangent to the graph at $(x_1, f(x_1))$ has the equation

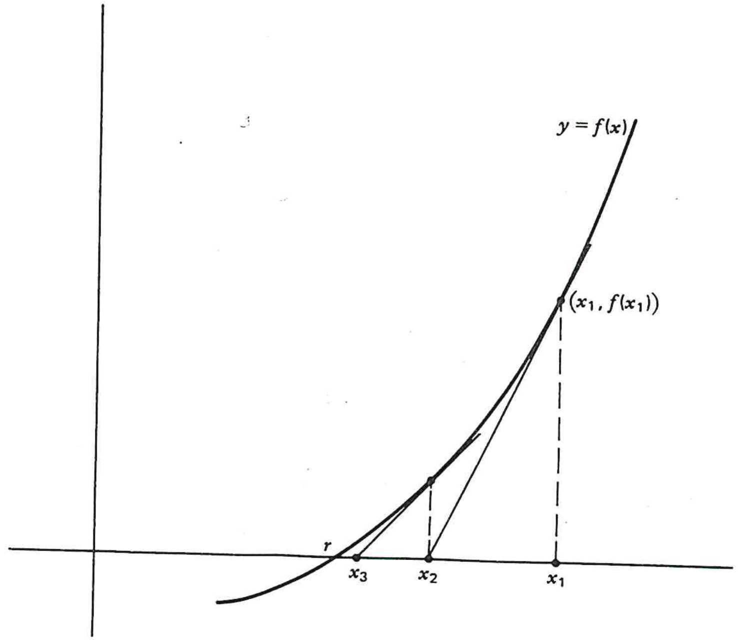

FIGURE 6.4.2 Newton's Method.

$y = f(x_1) + f'(x_1)(x - x_1)$, and crosses the x-axis at the point

$$x_2 := x_1 - \frac{f(x_1)}{f'(x_1)}.$$

(See Figure 6.4.2.) If we replace x_1 by the second estimate x_2, then we obtain a point x_3, and so on. At the nth iteration we get the point x_{n+1} from the point x_n by the formula

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}.$$

Under suitable hypotheses, the sequence (x_n) will converge rapidly to a root of the equation $f(x) = 0$, as we now show. The key tool in establishing the rapid rate of convergence is Taylor's Theorem.

6.4.7 Newton's Method *Let $I := [a, b]$ and let $f: I \to \mathbf{R}$ be twice differentiable on I. Suppose that $f(a)f(b) < 0$ and that there are constants m, M such that $|f'(x)| \geqslant m > 0$ and $|f''(x)| \leqslant M$ for all $x \in I$ and let $K := M/2m$. Then there exists a subinterval I^**

sequence (x_n) *defined by*

$$(*) \qquad x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)} \qquad \text{for} \quad n \in N,$$

belongs to I^* *and* (x_n) *converges to* r. *Moreover*

$$(**) \qquad |x_{n+1} - r| \leqslant K|x_n - r|^2 \qquad \text{for} \quad n \in N.$$

Proof. Since $f(a)f(b) < 0$, the numbers $f(a)$ and $f(b)$ have opposite signs; hence by Theorem 5.3.5 there exists $r \in I$ such that $f(r) = 0$. Since f' is never zero on I, it follows from Rolle's Theorem 6.2.3 that f does not vanish at any other point of I.

We now let $x' \in I$ be arbitrary; by Taylor's Theorem there exists a point c' between x' and r such that

$$0 = f(r) = f(x') + f'(x')(r - x') + \tfrac{1}{2}f''(c')(r - x')^2,$$

from which it follows that

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

If x'' is the number defined from x' by "the Newton procedure":

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

then an elementary calculation shows that

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

whence it follows that

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

Since $c' \in I$, the assumed bounds on f' and f'' hold and, setting $K := M/2m$, we obtain the inequality

$$(\#) \qquad |x'' - r| \leqslant K|x' - r|^2.$$

We now choose $\delta > 0$ so small that $\delta < 1/K$ and that the interval $I^* := [r - \delta, r + \delta]$ is contained in I. If $x_n \in I^*$, then $|x_n - r| \leqslant \delta$ and it follows from (#)

that $|x_{n+1} - r| \leqslant K|x_n - r|^2 \leqslant K\delta^2 < \delta$; hence $x_n \in I^*$ implies that $x_{n+1} \in I^*$. Therefore if $x_1 \in I^*$, we infer that $x_n \in I^*$ for all $n \in N$. Also if $x_1 \in I^*$, then an elementary induction argument using (#) shows that $|x_{n+1} - r| < (K\delta)^n|x_1 - r|$ for $n \in N$. But since $K\delta < 1$ this proves that $\lim x_n = r$. Q.E.D.

6.4.8 Example We shall illustrate Newton's Method by using it to approximate $\sqrt{2}$. If we let $f(x) := x^2 - 2$ for $x \in \mathbf{R}$, then we seek the positive root of the equation $f(x) = 0$. Since $f'(x) = 2x$, the iteration formula is

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

$$= x_n - \frac{x_n^2 - 2}{2x_n} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

If we take $x_1 := 1$ as our initial estimate, we obtain the successive values $x_2 = 3/2 = 1.5$, $x_3 = 17/12 = 1.416\,666\ldots$, $x_4 = 577/408 = 1.414\,215\ldots$, and $x_5 = 665\,857/470\,832 = 1.414\,213\,562\,374\ldots$, which is correct to eleven places.

Remarks (1) If we let $e_n := x_n - r$ be the error in approximating r, then inequality $(**)$ can be written in the form $|Ke_{n+1}| \leqslant |Ke_n|^2$. Consequently, if $|Ke_n| < 10^{-m}$ then $|Ke_{n+1}| < 10^{-2m}$ so that the number of significant digits in Ke_n has been doubled. Because of this doubling, the sequence generated by Newton's Method is said to converge "quadratically".

(2) In practice, when Newton's Method is programmed for a computer, one often makes an initial guess x_1 and lets the computer run. If x_1 is poorly chosen, or if the root is too near the end point of I, the procedure may not converge to a zero of f. A variety of possible difficulties are illustrated in Figure 6.4.3 through 6.4.5. One familiar strategy is to use the bisection method to arrive at a fairly close estimate of the root and then to switch to Newton's Method for the coup de grâce.

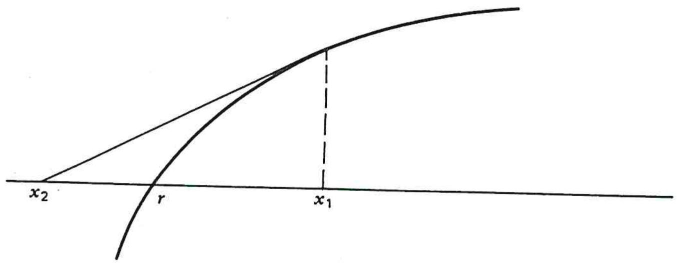

FIGURE 6.4.3

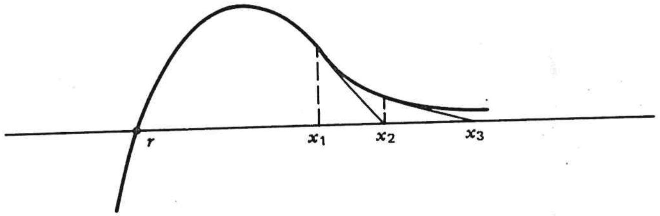

FIGURE 6.4.4 $x_n \to \infty$.

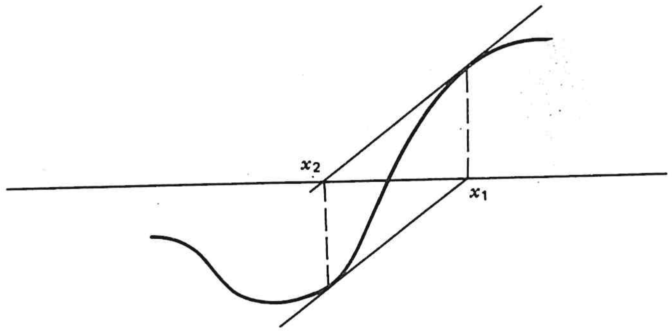

FIGURE 6.4.5 (x_n) oscillates between x_1 and x_2.

Exercises for Section 6.4

1. Let $f(x) := \cos ax$ for $x \in \boldsymbol{R}$ where $a \neq 0$. Find $f^{(n)}(x)$ for $n \in \boldsymbol{N}$, $x \in \boldsymbol{R}$.
2. Let $g(x) := |x^3|$ for $x \in \boldsymbol{R}$. Find $g'(x)$ and $g''(x)$ for $x \in \boldsymbol{R}$, and $g'''(x)$ for $x \neq 0$. Show that $g'''(0)$ does not exist.
3. Use induction to prove Leibniz's rule for the nth derivative of a product:

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

4. Show that if $x > 0$, then $1 + \frac{1}{2}x - \frac{1}{8}x^2 \leqslant \sqrt{1+x} \leqslant 1 + \frac{1}{2}x$.
5. Use the preceding exercise to approximate $\sqrt{1.2}$ and $\sqrt{2}$. What is the best accuracy you can be sure of, using this inequality?
6. Use Taylor's Theorem with $n = 2$ to obtain more accurate approximations for $\sqrt{1.2}$ and $\sqrt{2}$.
7. If $x > 0$ show that $|(1+x)^{1/3} - (1 + \frac{1}{3}x - \frac{1}{9}x^2)| \leqslant (5/81)x^3$. Use this inequality to approximate $\sqrt[3]{1.2}$ and $\sqrt[3]{2}$.
8. If $f(x) := e^x$, show that the remainder term in Taylor's Theorem converges to zero as $n \to \infty$, for each fixed x_0 and x. [*Hint:* See Theorem 3.2.11.]
9. If $g(x) := \sin x$, show that the remainder term in Taylor's Theorem converges to 0 as $n \to \infty$ for each fixed x_0 and x.

10. Let $h(x) := e^{-1/x^2}$ for $x \neq 0$ and $h(0) := 0$. Show that $h^{(n)}(0) = 0$ for all $n \in \boldsymbol{N}$. Conclude that the remainder term in Taylor's Theorem for $x_0 = 0$ does *not* converge to zero as $n \to \infty$ for $x \neq 0$. [*Hint:* By L'Hospital's Rule, $\lim_{x \to 0} h(x)/x^k = 0$ for any $k \in \boldsymbol{N}$. Use Exercise 3 to calculate $h^{(n)}(x)$ for $x \neq 0$.]
11. If $x \in [0,1]$ and $n \in \boldsymbol{N}$, show that

$$\left| \log(1+x) - \left(x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + (-1)^{n-1}\frac{x^n}{n} \right) \right| < \frac{x^{n+1}}{n+1}.$$

Use this to approximate $\log 1.5$ with an error less than 0.01. Less than 0.001.

12. We wish to approximate sin by a polynomial on $[-1,1]$ so that the error is less than 0.001. Show that we have

$$\left| \sin x - \left(x - \frac{x^3}{6} + \frac{x^5}{120} \right) \right| < \frac{1}{5040} \qquad \text{for} \quad |x| \leqslant 1.$$

13. Calculate e correct to 7 decimal places.
14. Determine whether or not $x = 0$ is a point of relative extremum of the following functions:
(a) $f(x) := x^3 + 2$, (b) $g(x) := \sin x - x$,
(c) $h(x) := \sin x + \frac{1}{6}x^3$, (d) $k(x) := \cos x - 1 + \frac{1}{2}x^2$.
15. Let f be continuous on $[a,b]$ and assume the second derivative f'' exists on (a,b). Suppose that the graph of f and the line segment joining the points $(a, f(a))$ and $(b, f(b))$ intersect at a point $(x_0, f(x_0))$ where $a < x_0 < b$. Show that there exists a point $c \in (a,b)$ such that $f''(c) = 0$.
16. Let $I \subseteq \boldsymbol{R}$ be an open interval, let $f\colon I \to \boldsymbol{R}$ be differentiable on I, and suppose $f''(a)$ exists at $a \in I$. Show that

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

Give an example where this limit exists, but the function does not have a second derivative at a.

17. Suppose that $I \subseteq \boldsymbol{R}$ is an open interval and that $f''(x) \geqslant 0$ for all $x \in I$. If $c \in I$, show that the part of the graph of f on I is never below the tangent line to the graph at $(c, f(c))$.
18. Let $I \subseteq \boldsymbol{R}$ be an interval and let $c \in I$. Suppose that f and g are defined on I and that the derivatives $f^{(n)}, g^{(n)}$ exist and are continuous on I. If $f^{(k)}(c) = 0$ and $g^{(k)}(c) = 0$ for $k = 0, 1, \ldots, n-1$, but $g^{(n)}(c) \neq 0$, show that

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \frac{f^{(n)}(c)}{g^{(n)}(c)}.$$

19. Show that the function $f(x) := x^3 - 2x - 5$ has a zero r in the interval $I := [2, 2.2]$. If $x_1 := 2$ and if we define the sequence (x_n) using the Newton procedure, show that $|x_{n+1} - r| \leqslant (0.7)|x_n - r|^2$. Show that x_4 is accurate to within six decimal places.

20. Approximate the real zeros of $g(x) := x^4 - x - 3$.
21. Approximate the real zeros of $h(x) := x^3 - x - 1$. Apply Newton's Method starting with the initial choices (a) $x_1 := 2$, (b) $x_1 := 0$, (c) $x_1 = -2$. Explain what happens.
22. The equation $\log x = x - 2$ has two solutions. Approximate them using Newton's Method. What happens if $x_1 := \frac{1}{2}$ is the initial point?
23. The function $f(x) := 8x^3 - 8x^2 + 1$ has two zeros in $[0, 1]$. Approximate them, using Newton's Method, with the starting points (a) $x_1 := \frac{1}{8}$, and (b) $x_1 := \frac{1}{4}$. Explain what happens.
24. Approximate the solution of the equation $x = \cos x$, accurate to within six decimals.
25. Let $I := [a, b]$ and let $f: I \to \mathbf{R}$ be differentiable on I. Suppose that $f(a) < 0 < f(b)$ and that there exist m, M such that $0 < m < f'(x) \leqslant M$ for $x \in I$. Let $x_1 \in I$ be arbitrary and define $x_{n+1} := x_n - f(x_n)/M$ for $n \in N$. Show that the sequence (x_n) is well defined and converges to the unique zero $r \in I$ of f and that

$$|x_{n+1} - r| \leqslant (1 - m/M)^n |x_1 - r| \leqslant (1 - m/M)^n |f(x_1)|/m.$$

[*Hint:* If $\varphi(x) := x - f(x)/M$, show that $0 \leqslant \varphi'(x) \leqslant 1 - m/M < 1$ and that $\varphi(I) \subseteq I$.]
26. Apply the result in the preceding exercise to some of the functions considered in Exercises 19–24.

CHAPTER SEVEN

THE RIEMANN INTEGRAL

We have already mentioned the developments, during the 1630s, by Fermat and Descartes leading to analytic geometry and the theory of the derivative. However, the subject we know as calculus did not begin to take shape until the late 1660s when Isaac Newton (1642–1727) created his theory of "fluxions" and invented the method of "inverse tangents" to find areas under curves. The reversal of the process for finding tangent lines to find areas was also discovered in the 1680s by Gottfried Leibniz (1646–1716), who was unaware of Newton's unpublished work and who arrived at the discovery by a very different route. Leibniz introduced the terminology "calculus differentialis" and "calculus integralis", since finding tangent lines involved differences and finding areas involved summations. Thus, they had discovered that integration, being a process of summation, was inverse to the operation of differentiation.

During a century and a half of development and refinement of techniques, calculus consisted of these paired operations and their applications, primarily to physical problems. In the 1850s, Bernhard Riemann (1826–1866) adopted a new and different viewpoint. He separated the concept of integration from its companion, differentiation, and examined the motivating summation and limit process of finding areas by itself. He broadened the scope by considering all functions on an interval for which this process of "integration" could be defined: the class of "integrable" functions. The Fundamental Theorem of Calculus became a result that held only for a restricted set of integrable functions. The viewpoint of Riemann led others to invent other integration theories, the most significant being Lebesgue's theory of integration.

In this chapter we shall begin by defining the concept of Riemann integrability of functions on an interval by means of upper and lower sums. In Section 7.2, we shall discuss the basic properties of the integral and the class of integrable functions on an interval. The Fundamental Theorem of Calculus is the principal topic in Section 7.3. Other approaches to the Riemann integral are discussed in Section 7.4 and their equivalence is established, and a brief introduction to "improper integrals" is presented. In Section 7.5 we discuss several methods of