Chapter 4
The Theory of Differentiable Functions

Abstract In this chapter, we put the derivative to work in analyzing functions. We will see how to find optimal values of functions and how to construct polynomial approximations to the function itself.

4.1 The Mean Value Theorem

At the beginning of Chap. 3, we posed a question about traveling with a broken speedometer: "Is there any way of determining the speed of the car from readings on the odometer?" In answering the question, we were led to the concept of the derivative of f at a, $f'(a)$, the instantaneous rate of change in f at a. Next we examine implications of the derivative. The mean value theorem for derivatives provides an important link between the derivative of f on an interval and the behavior of f over the interval. The mean value theorem says that if the distance that we traveled between 2:00 p.m. and 4:00 p.m. is 90 miles, then there must be at least one point in time when we traveled at exactly 45 mph. The conclusion that the average velocity over an interval must be equal to the instantaneous velocity at some point appears to be just common sense, but some of the most commonsensible theorems require somewhat sophisticated proofs. So it is with the mean value theorem, which we now state precisely.

Theorem 4.1. Mean value theorem. *Suppose that a function f is continuous on the closed interval $[a,b]$ and differentiable on the open interval (a,b). Then there exists a number c in the interval (a,b) where*

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

P.D. Lax and M.S. Terrell, *Calculus With Applications*, Undergraduate Texts in Mathematics, 171
DOI 10.1007/978-1-4614-7946-8_4,

The theorem has an interesting geometric interpretation: Given points $(a, f(a))$ and $(b, f(b))$ on the graph, there is a point between them, $(c, f(c))$, at which the tangent is parallel to the secant through $(a, f(a))$ and $(b, f(b))$ (Fig. 4.1).

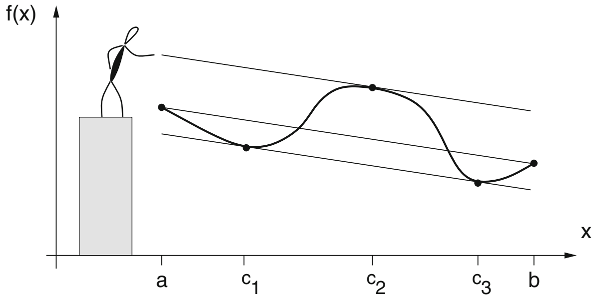

Fig. 4.1 Move the secant vertically as far as possible without breaking contact with the graph

To see how such a point c may be found, take a duplicate copy of the secant line and raise or lower it vertically to the point at which it loses contact with the graph. We notice two things about this point. First: that it occurs at a point $(c, f(c))$ on the graph of f that is farthest from the secant line. Second: that $f'(c) = \dfrac{f(b)-f(a)}{b-a}$. This suggests the key observation in the proof of the mean value theorem: that we find the desired point c by finding the point on the graph that is farthest from the secant line. The proof of the mean value theorem relies on the following result, which is important enough to present separately as a lemma.

Lemma 4.1. *Suppose a function f is defined on an open interval (a,b) and reaches its maximum or minimum at c. If $f'(c)$ exists, then $f'(c) = 0$.*

Proof. We show that $f'(c) = 0$ by eliminating the possibilities that it is positive or negative. Suppose that $f'(c) > 0$. Since the limit of $\dfrac{f(c+h)-f(c)}{h}$ as h tends to 0 exists and is positive, it follows that for h small enough, $\dfrac{f(c+h)-f(c)}{h}$ is also positive. This implies that for all h small enough and *positive*,

$$f(c+h) > f(c).$$

But for h small enough, $c+h$ belongs to (a,b), so that the above inequality violates the assumption that $f(c)$ is the maximum of f on (a,b). Therefore, it is not possible that $f'(c) > 0$.

Similarly, we show that $f'(c) < 0$ is not possible. Suppose that $f'(c)$ is negative. Then for all small h, $\dfrac{f(c+h)-f(c)}{h}$ is also negative. Taking small *negative* h, then the numerator $f(c+h) - f(c)$ is positive, $c+h$ is in (a,b), and

$$f(c+h) > f(c).$$

This contradicts the assumption that $f(c)$ is a maximum.

So only $f'(c) = 0$ is consistent with f achieving its maximum at c. A similar argument shows that if the minimum of f on (a,b) occurs at c, then $f'(c) = 0$. □

Now we are ready to prove the mean value theorem.

Proof. Let $\ell(x) = f(a) + \dfrac{f(b) - f(a)}{b - a}(x - a)$ be the linear function that is the secant through $(a, f(a))$ and $(b, f(b))$. Define d to be the difference between f and ℓ:

$$d(x) = f(x) - \ell(x) = f(x) - \left(f(a) + \frac{f(b) - f(a)}{b - a}(x - a)\right).$$

Since $f(x)$ and $\ell(x)$ have the same values at the endpoints, $d(x)$ is zero at the endpoints a and b. Since both f and ℓ are continuous on $[a,b]$ and differentiable on (a,b), so is d. By the extreme value theorem, Theorem 2.6, d has a maximum and a minimum on $[a,b]$. We consider two possibilities. First, it may happen that both the maximum and the minimum of d occur at the endpoints of the interval. In this case, every $d(x)$ is between $d(a)$ and $d(b)$, but both of these numbers are 0, so this would imply that $d(x) = 0$ and that $f(x) = \ell(x)$. In this case, $f'(x) = \dfrac{f(b) - f(a)}{b - a}$ for all x between a and b. Thus every number c in (a,b) satisfies the requirements of the theorem.

The other possibility is that either the maximum or the minimum of d occurs at some point c in (a,b). Since d is differentiable on (a,b), we have by Lemma 4.1 that $d'(c) = 0$:

$$0 = d'(c) = f'(c) - \ell'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}.$$

it follows that

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

□

Let us see how the mean value theorem provides a way to use information about the derivative on an interval to get information about the function.

Corollary 4.1. *A function whose derivative is zero at every point of an interval is constant on that interval.*

Proof. Let $a \neq b$ be any two points in the interval. The function is differentiable at a and b and every point in between. So by the mean value theorem, there exists at least one c between a and b at which $f(b) - f(a) = f'(c)(b - a)$. Since $f'(c) = 0$ for every c, it follows that $f(a) = f(b)$. Since f has the same value at any two points in the interval, it is constant. □

If two functions have the same derivatives throughout an interval, then $f' - g' = (f - g)' = 0$ and $f - g$ is a constant function. In Chap. 3, we used this result to find all the solutions to the differential equations $y' = y$, $y'' + y = 0$, and $y'' - y = 0$. Here are some additional ways to use this corollary.

Example 4.1. Suppose f is a function for which $f'(x) = 3x^2$. What can f be? One possibility is x^3. Therefore, $f(x) - x^3$ has derivative zero everywhere. According to Corollary 4.1, $f(x) - x^3$ is a constant c. Therefore, $f(x) = x^3 + c$.

Example 4.2. Suppose again that f is a function for which $f'(x) = 3x^2$, and that we now know in addition that $f(1) = 2$. By the previous example, we know that $f(x) = x^3 + c$ for some number c. Since

$$f(1) = 2 = 1^3 + c,$$

it follows that $c = 1$, and $f(x) = x^3 + 1$ is the only function satisfying both requirements.

Example 4.3. Suppose f is a function for which $f'(x) = -x^{-2}$. What can f be? The domain of f' does not include 0, so two intervals to which we can apply the corollary are x positive, and x negative.

Arguing as in Example 4.1, we conclude that $f(x) = \frac{1}{x} + a$ for x positive, and $f(x) = \frac{1}{x} + b$ for x negative, where a and b are arbitrary numbers.

The mean value theorem also enables us to determine intervals on which a function f is increasing, or on which it is decreasing, by considering the sign of f'.

Corollary 4.2. Criteria for Monotonicity. *If $f' > 0$ on an interval then f is increasing on that interval. If $f' < 0$ on an interval then f is decreasing on that interval. If $f' \geq 0$ on an interval then f is nondecreasing on that interval. If $f' \leq 0$ on an interval thin f is nonincreasing on that interval.*

Proof. Take any two points a and b in the interval such that $a < b$. By the mean value theorem, there exists a point c between a and b such that $f'(c) = \dfrac{f(b) - f(a)}{b - a}$. Since $b > a$, it follows that the sign of $f(b) - f(a)$ is the same as the sign of $f'(c)$. If $f' > 0$, then $f(b) - f(a) > 0$, and f is increasing. If $f' < 0$, then $f(b) - f(a) < 0$, and f is decreasing. The proof for the nonstrict inequalities is analogous. □

4.1a Using the First Derivative for Optimization

Finding Extreme Values on a Closed Interval. The next two examples show how to apply Lemma 4.1 to find extreme values on closed intervals.

Example 4.4. We find the largest and smallest values of

$$f(x) = 2x^3 + 3x^2 - 12x \qquad \text{on } [-4,3].$$

By the extreme value theorem, f must attain a maximum and a minimum on the interval. The extreme values can occur either at the endpoints, $x=-4$ and $x=3$, or in $(-4,3)$. At the endpoints, we have $f(-4)=-32$ and $f(3)=45$. If f attains a maximum or minimum at c in $(-4,3)$, then by Lemma 4.1, the derivative $f'(c)$ must be equal to 0. Next, we identify all points at which the derivative is 0: we have

$$0 = f'(x) = 6x^2 + 6x - 12 = 6(x+2)(x-1)$$

when $x=-2$ or $x=1$. Both of these lie in $(-4,3)$. The possibilities for extreme points are then

$$f(1) = -7, \qquad f(-2) = 20, \qquad f(-4) = -32, \qquad f(3) = 45.$$

The left endpoint yields the smallest value of f on $-4,3$, and the right endpoint yields the largest value. See Fig. 4.2.

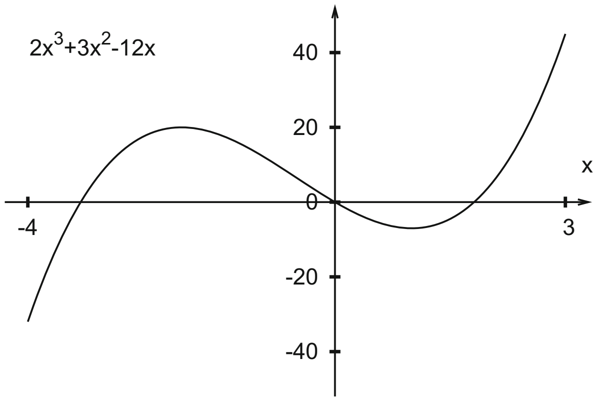

Fig. 4.2 The maximum and minimum of f on $[-4,3]$ occur at the endpoints in Example 4.4

Example 4.5. Suppose we change the domain of $f(x) = 2x^3 + 3x^2 - 12x$ to $[-3,3]$. Then the possibilities are

$$f(1) = -7, \qquad f(-2) = 20, \qquad f(-3) = 9, \qquad f(3) = 45.$$

The maximum is at the right endpoint, and the minimum is now at $x=1$.

Example 4.6. Find the largest and smallest values of

$$f(x) = x^{2/3} \qquad \text{on } [-1,1].$$

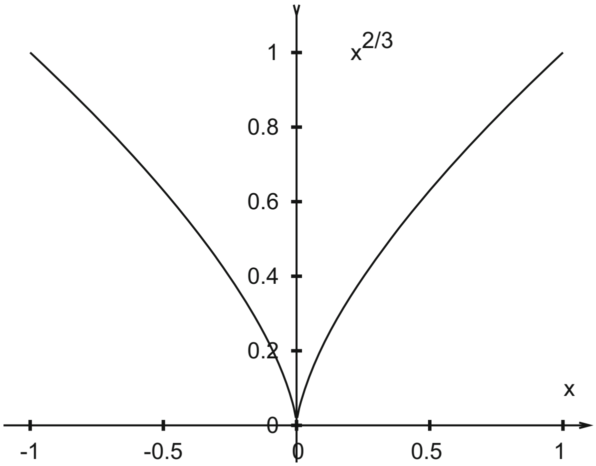

Fig. 4.3 The maximum occurs at both endpoints, and the minimum occurs at a point at which the derivative does not exist, in Example 4.6

The graph of f is shown in Fig. 4.3. We note that f is continuous on $[-1,1]$, and so it has both a maximum and a minimum value on $[-1,1]$. The extreme values can occur at the endpoints or at an interior point c in $(-1,1)$ where $f'(c)=0$ or where $f'(c)$ does not exist. The derivative $f'(x)=\frac{2}{3}x^{-1/3}$ does not exist at $x=0$. There are no points where $f'(x)=0$. So the only candidates for the maximum and minimum are

$$f(-1)=1, \quad f(1)=1, \quad \text{and} \quad f(0)=0.$$

So f attains its maximum value at both endpoints and its minimum at $x=0$, where the derivative fails to exist.

Local and Global Extrema. The graph of the function f in Fig. 4.2 shows that there may be points of interest on a graph that are not the maximum or minimum of f, but are relative or local extrema.

Definition 4.1. Local and global extrema. A function f has a *local maximum* $f(c)$ at c if there is a positive number h such that $f(x)\le f(c)$ whenever $c-h\le x\le c+h$. A function f has a *local minimum* $f(c)$ at c if there is a positive number h such that $f(x)\ge f(c)$ whenever $c-h\le x\le c+h$. A function f has an *absolute, or global, maximum* $f(c)$ at c if $f(x)\le f(c)$ for all x in the domain of f. A function f has an *absolute, or global minimum* $f(c)$ at c if $f(x)\ge f(c)$for all x in the domain of f.

Most functions we encounter in calculus have nonzero derivatives at most points in their domains. Points at which the derivative is zero or does not exist are called *critical points* of the function. Endpoints of the domain (if there are any) and critical points subdivide the domain into smaller intervals on which the derivative is either positive or negative. Next we show how the monotonicity criteria can be used to identify points in the domain of f at which f has extreme values. This result is often called the *first derivative test*.

Theorem 4.2. First derivative test. *Suppose that f is continuous on an interval containing c, and that $f'(x)$ is positive for x less than c, and negative for x greater than c. Then f reaches its maximum on the interval at c.*

A similar characterization holds for the minimum.

The proof follows from the criterion for monotonicity and the definitions of maximum and minimum. We ask you to write it out in Problem 4.6. Figure 4.4 shows some examples. Whether an extremum on an interval is local or absolute depends on whether the interval is the entire domain.

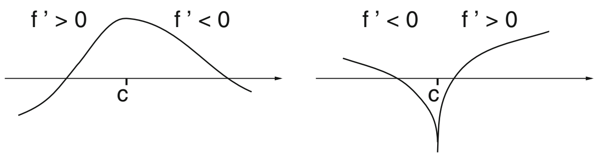

Fig. 4.4 An illustration of the first derivative test. *Left*: f has a maximum at c. *Right*: f has a minimum

Example 4.7. Consider the quadratic function

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

We can rewrite this by "completing the square":

$$f(x) = x^2 + bx + c = \left(x + \frac{1}{2}b\right)^2 - \frac{1}{4}b^2 + c. \tag{4.1}$$

We see at a glance that the minimum of $f(x)$ is achieved at $x = -\frac{1}{2}b$. We next show how to derive this result by calculus. First we look for critical points of f: $f'(x) = 2x + b$ is zero when $x = -\frac{1}{2}b$. Next, we check the sign of the first derivative:

$$f'(x) \begin{cases} < 0 & \text{for } x < -\frac{1}{2}b, \\ = 0 & \text{for } x = -\frac{1}{2}b, \\ > 0 & \text{for } x > -\frac{1}{2}b. \end{cases}$$

By the first derivative test, f achieves an absolute minimum at $x = -\dfrac{1}{2}b$.

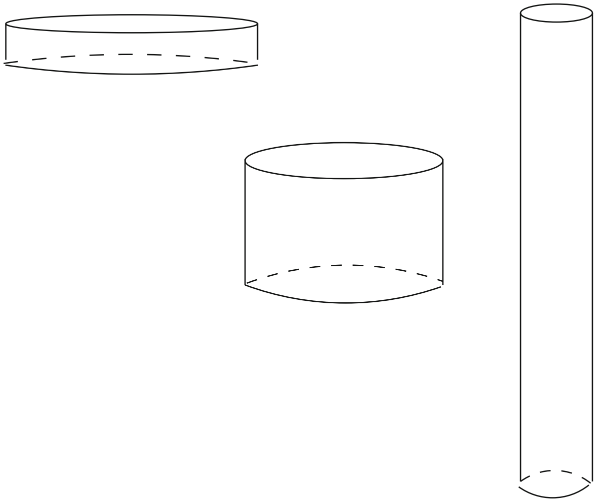

Fig. 4.5 Example 4.8 illustrated. Three cylinders have the same surface area $A = 2\pi$, but unequal volumes. *Top left* : $r = \frac{7}{8}$, $V = 0.6442\ldots$ *Center* : $r = \frac{2}{3}$, $V = 1.1635\ldots$ *Right* : $r = \frac{1}{4}$, $V = 0.7363\ldots$

Example 4.8. We determine the shape of the closed cylinder that has the largest volume among all cylinders of given surface area A.

We plan to accomplish this by expressing the volume as a function of one variable, the radius r. Let h be the height. Then the surface area is

$$A = 2\pi r^2 + 2\pi rh = 2\pi r(r+h), \tag{4.2}$$

and the volume is $V = \pi r^2 h$. We have expressed V as a function of two variables, r and h, but we can eliminate one by means of the constraint (4.2). Solving for h, we see that $h = \dfrac{A}{2\pi r} - r$, so that

$$V = f(r) = \pi r^2 \left(\frac{A}{2\pi r} - r\right) = \frac{Ar}{2} - \pi r^3, \qquad r > 0.$$

The derivative $f'(r) = \frac{1}{2}A - 3\pi r^2$ is zero when $r = r_0 = \sqrt{\dfrac{A}{6\pi}}$. Since $f' > 0$ for smaller values of r, and $f' < 0$ for larger r, f has an absolute maximum at r_0. To determine the shape of this cylinder, evaluate h in terms of r_0: the height of the cylinder of largest volume is

$$h = \frac{A}{2\pi r} - r = \frac{6\pi r_0^2}{2\pi r_0} - r_0 = 2r_0.$$

That is, for cylinders of a given surface area, the volume is greatest when the diameter of the cylinder equals the height. See Fig. 4.5.

The surface area of the cylinder is proportional to the amount of material needed to manufacture a cylindrical container. The above shape is *optimal* in the sense that it encloses the largest volume for a given amount of material. Examine the cans in the supermarket and determine which brands use the optimal shape.

4.1b Using Calculus to Prove Inequalities

Next, we show how calculus makes it easier to derive some inequalities that we obtained before we had calculus.

Exponential Growth. We showed in Theorem 2.10 that as x tends to infinity, the exponential function grows faster than any power of x, in the sense that for a fixed n, $\frac{e^x}{x^n}$ tends to infinity as x tends to infinity. Here is a simple calculus proof of this fact.

Differentiate $f(x) = \frac{e^x}{x^n}$. Using the product rule, we get

$$f'(x) = \frac{e^x}{x^n} - n\frac{e^x}{x^{n+1}} = f(x) - n\frac{f(x)}{x} = f(x)\frac{x-n}{x}. \tag{4.3}$$

This shows that the derivative of $f(x)$ is negative for $0 < x < n$ and is positive for x greater than n. So $f(x)$ is decreasing as x goes from 0 to n and increasing from then on. It follows that $f(x)$ reaches its minimum at $x = n$. This means that

$$f(x) = \frac{e^x}{x^n} \geq \frac{e^n}{n^n}, \qquad (x > 0).$$

Multiply this inequality by x. It follows that

$$\frac{e^x}{x^{n-1}} \geq x\frac{e^n}{n^n}.$$

The number $\frac{e^n}{n^n}$ is fixed by our original choice of n, so the function on the right tends to infinity as x tends to infinity. Therefore, so does the function on the left. Since n is arbitrary, this proves our contention.

The A-G Inequality. Recall from Sect. 1.1c that the arithmetic–geometric inequality says that for any two positive numbers a and b,

$$\frac{a+b}{2} - (ab)^{1/2} > 0, \tag{4.4}$$

unless $a = b$. Let us see how to use calculus to obtain Eq. (4.4). Let a be the smaller of the two numbers: $a < b$. Define the function $f(x)$ to be

$$f(x) = \frac{a+x}{2} - (ax)^{1/2}. \tag{4.5}$$

Then $f(b) = \dfrac{a+b}{2} - (ab)^{1/2}$, $f(a)$ is zero, and the A-G inequality can be formulated thus:

$$f(b) \text{ is greater than } f(a).$$

Since $a < b$, this will follow if we can show that $f(x)$ is an increasing function between a and b. The calculus criterion for a function to be increasing is for its derivative to be positive. The derivative is

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

which is positive for x greater than a. This completes the proof of the A-G inequality for two numbers.

How about three numbers? The A-G inequality for three positive numbers a, b, and c states that

$$\frac{a+b+c}{3} - (abc)^{1/3} > 0, \tag{4.6}$$

unless all three numbers a, b, c are equal.

Rewrite Eq. (4.6) as

$$abc \le \left(\frac{a+b+c}{3}\right)^3$$

and divide by c:

$$ab \le \frac{1}{c}\left(\frac{a+b+c}{3}\right)^3. \tag{4.7}$$

Keep a and b fixed and define the function $f(x)$ as the right side of Eq. (4.7) with c replaced by x, where $x > 0$:

$$f(x) = \frac{1}{x}\left(\frac{a+b+x}{3}\right)^3. \tag{4.8}$$

As x tends to 0, $f(x)$ tends to infinity, and as x tends to infinity, $f(x)$ tends to infinity. Therefore, $f(x)$ attains its absolute minimum for some $x > 0$. We shall use the calculus criterion to find that minimum value. Differentiate f. Using the product rule and chain rule, we get

$$f'(x) = \frac{1}{x}\left(\frac{a+b+x}{3}\right)^2 - \frac{1}{x^2}\left(\frac{a+b+x}{3}\right)^3.$$

We see that $f'(x)$ is zero if

$$x = \frac{a+b+x}{3},$$

so the minimum occurs at $x = \frac{1}{2}(a+b)$. The value of f at this point is

$$f\left(\frac{1}{2}(a+b)\right) = \frac{2}{a+b}\left(\frac{a+b+\frac{1}{2}(a+b)}{3}\right)^3 = \left(\frac{a+b}{2}\right)^2.$$

According to the A-G inequality for $n = 2$, this is greater than or equal to ab. So at its minimum value, $f(x)$ is not less than ab. Therefore, it is not less than ab at any other point c. This completes the proof of the A-G inequality for three numbers.

The A-G inequality can be proved inductively for every n by a similar argument, and we ask you to do so in Problem 4.15.

4.1c A Generalized Mean Value Theorem

At the start of this section, we noted that the mean value theorem guarantees that if there is an interval over which your average velocity was 30 mph, then there was at least one moment in that interval when your velocity was exactly 30 mph. In this section we see that the mean value theorem can be used to prove a somewhat surprising variation: If during an interval of time you have traveled five times as far as your friend, then there has to be at least one moment when you were traveling exactly five times as fast as your friend.

Theorem 4.3. Generalized mean value. *Suppose f and g are differentiable on (a,b) and continuous on $[a,b]$. If $g'(x) \neq 0$ in (a,b), then there exists a point c in (a,b) such that*

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

Proof. Let

$$H(x) = \big(f(x)-f(a)\big)\big(g(b)-g(a)\big) - \big(g(x)-g(a)\big)\big(f(b)-f(a)\big).$$

Then H is differentiable on (a,b) and continuous on $[a,b]$, and $H(a) = H(b) = 0$. By the mean value theorem, there exists a point c in (a,b) where

$$0 = \frac{H(b)-H(a)}{b-a} = H'(c) = f'(c)\big(g(b)-g(a)\big) - g'(c)\big(f(b)-f(a)\big),$$

and so $f'(c)\big(g(b)-g(a)\big) = g'(c)\big(f(b)-f(a)\big)$. Since $g' \neq 0$ in (a,b), neither of $g'(c)$, $\big(g(b)-g(a)\big)$ is 0. To complete the proof, divide both sides by $g'(c)\big(g(b)-g(a)\big)$. □

This variation of the mean value theorem can be used to prove the following technique (l'Hospitals' rule) for evaluating some limits.

Theorem 4.4. *Suppose* $\lim_{x\to a} f(x) = 0$, $\lim_{x\to a} g(x) = 0$, *and that* $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ *exists. Then*

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

Proof. Since $\lim_{x\to a} \dfrac{f'(x)}{g'(x)}$ exists, there is an interval around a (perhaps excluding a) where $f'(x)$ and $g'(x)$ exist and $g'(x) \neq 0$. Define two new functions F and G that agree with f and g for $x \neq a$, and set $F(a) = G(a) = 0$. By Theorem 4.3 applied to F and G, there is a point c between a and x such that

$$\frac{f(x)}{g(x)} = \frac{F(x)}{G(x)} = \frac{F(x)-F(a)}{G(x)-G(a)} = \frac{F'(c)}{G'(c)} = \frac{f'(c)}{g'(c)}.$$

Since c is between a and x and $\lim_{x\to a} \dfrac{f'(x)}{g'(x)}$ exists, it follows that $\lim_{x\to a} \dfrac{f'(x)}{g'(x)} = \lim_{x\to a} \dfrac{f'(c)}{g'(c)} = \lim_{x\to a} \dfrac{f(x)}{g(x)}$. □

Example 4.9. The limit $\lim_{x\to 1} \dfrac{\log x}{x^2-1}$ satisfies $\lim_{x\to 1} \log x = 0$ and $\lim_{x\to 1} (x^2-1) = 0$, and both $\log x$ and x^2-1 are differentiable near 1. Therefore,

$$\lim_{x\to 1} \frac{\log x}{x^2-1} = \lim_{x\to 1} \frac{(\log x)'}{(x^2-1)'},$$

provided that the last limit exists. But it does exist, because

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

Therefore, $\lim_{x\to 1} \dfrac{\log x}{x^2-1} = \dfrac{1}{2}$. See Fig. 4.6.

Another version of this theorem may be found in Problem 4.23.

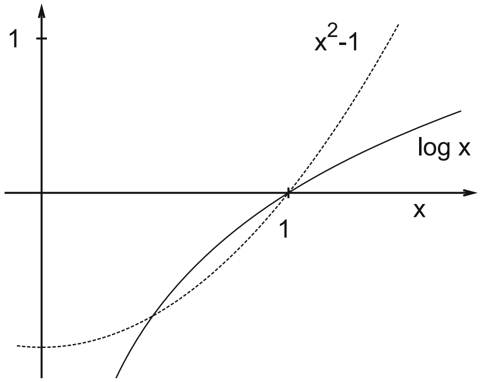

Fig. 4.6 Graphs of the functions in Example 4.9. The ratio of values is equal to the ratio of slopes as x tends to 1

Problems

4.1. Suppose $f(2) = 6$, and $0.4 \le f'(x) \le 0.5$ for x in $[2, 2.2]$. Use the mean value theorem to estimate $f(2.1)$.

4.2. Suppose $g(2) = 6$ and $-0.6 \le g'(x) \le -0.5$ for x in $[1.8, 2]$. Use the mean value theorem to estimate $g(1.8)$.

4.3. If $h'(x) = 2\cos(3x) - 3\sin(2x)$, what could $h(x)$ be? If, in addition, $h(0) = 0$, what could $h(x)$ be?

4.4. If $k'(t) = 2 - 2\mathrm{e}^{-3t}$, what could $k(t)$ be? If, in addition, $k(0) = 0$, what could $k(t)$ be?

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

(a) Find $f'(x)$.
(b) In which interval(s) does f increase?
(c) In which interval(s) does f decrease?
(d) Find the minimum value of f in $[-10, 10]$.
(e) Find the maximum value of f in $[-10, 10]$.

4.6. Justify the following steps to prove the first derivative test, Theorem 4.2. Suppose that $f'(x)$ is positive for $x < c$ and negative for $x > c$. We need to show that f reaches its maximum at c.

(a) Explain why f is increasing for all $x < c$, and why f is decreasing for all $x > c$.
(b) Use the continuity of f at c and the fact that f is increasing when $x < c$ to explain why $f(x)$ cannot be greater than $f(c)$ for any $x < c$. Similarly explain why $f(x)$ cannot be greater than $f(c)$ for any $x > c$.
(c) Explain why $f(c)$ is a maximum on S.

(d) Revise the argument to show that if $f'(x)$ is negative for all x less than c and $f'(x)$ is positive for all x greater than c, then f reaches its minimum at $f(c)$.

4.7. Rework Problem 1.10 using calculus.

4.8. Find the maximum and minimum values of the function

$$f(x) = 2x^3 - 3x^2 - 12x + 8$$

on each of the following intervals.

(a) $[-2.5, 4]$
(b) $[-2, 3]$
(c) $[-2.25, 3.75]$

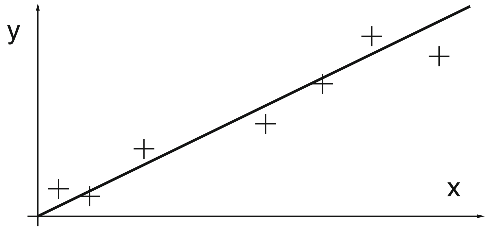

Fig. 4.7 In Problem 4.9 we find a slope to fit given data

4.9. Suppose an experiment is carried out to determine the value of the constant m in the equation

$$y = mx$$

relating two physical quantities. Let $(x_1, y_1),\ (x_2, y_2), \ldots, (x_n, y_n)$ be the measured values. Find the value of m, in terms of the x_i and y_i, that minimizes E, the sum of the squares of the errors between the observed measurements and the linear function $y = mx$:

$$E = (y_1 - mx_1)^2 + (y_2 - mx_2)^2 + \cdots + (y_n - mx_n)^2.$$

See Fig. 4.7.

4.10. Consider an open cardboard box whose bottom is a square of edge length x, and whose height is y. The volume V and surface area S of the box are given by

$$V = x^2 y, \qquad S = x^2 + 4xy.$$

Among all boxes with given volume, find the one with smallest surface area. Show that this box is squat, i.e., $y < x$.

4.11. Consider a particle of unit mass moving on a number line whose position at time t is given by $x(t) = 3t - t^2$. Find the time when the particle's position x is maximal.

4.12. Find the point on the graph of $y = \frac{1}{2}x^2$ closest to the point $(6,0)$.

4.13. What is the largest amount by which a positive number can exceed its cube?

4.14. Find the positive number x such that the sum of x and its reciprocal is as small as possible,

(a) Using calculus, and
(b) By the A-G inequality.

4.15. Use calculus to prove, by induction, the A-G inequality for n positive numbers.

4.16. Let $w_1,\dots,w_n$ be positive numbers whose sum is 1, and $a_1,\dots,a_n$ any positive numbers. Prove an extension of the A-G inequality:

$$a_1^{w_1}a_2^{w_2}\cdots a_n^{w_n} \le w_1a_1 + w_2a_2 + \cdots + w_na_n,$$

with equality only in the case $a_1 = a_2 = \cdots = a_n$. Try an inductive proof with one of the a's as the variable.

4.17. Suppose $g'(x) \le h'(x)$ for $0 < x$ and $g(0) = h(0)$. Prove that $g(x) \le h(x)$ for $0 < x$.

4.18. Here we apply Problem 4.17 to find polynomial bounds for the cosine and sine.

(a) Show that $g'(x) \le h'(x)$ for $g(x) = \sin x$ and $h(x) = x$ and deduce that

$$\sin x \le x \qquad (x > 0). \tag{4.9}$$

(b) Rewrite Eq. (4.9) as $\left(-\cos x\right)' \le \left(\frac{x^2}{2} - 1\right)'$ and deduce that $1 - \frac{x^2}{2} \le \cos x$.
(c) Continue along these lines to derive

$$1 - \frac{x^2}{2} \le \cos x \le 1 - \frac{x^2}{2} + \frac{x^4}{4!},$$

and in particular, estimate $\cos(0.2)$ with a tolerance of 0.001.
(d) Extend the previous argument to derive

$$1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} \le \cos x \le 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!}.$$

4.19. Denote the exponential function e^x by $e(x)$. Use $e' = e$ and Problem 4.17 to show the following.

(a) For $x > 0$, $1 < e(x)$.
(b) $1 < e'(x)$ for $0 < x$, and deduce that $1 + x < e(x)$.
(c) Rewrite this as $1 + x < e'(x)$, and deduce that $1 + x + \frac{1}{2}x^2 < e(x)$ for $0 < x$.
(d) For all n and all $x > 0$, $1 + x + \frac{x^2}{2} + \cdots + \frac{x^n}{n!} < e(x)$.

4.20. Evaluate $\lim_{x\to 0} \frac{\sin x}{x}$, first using the mean value theorem to write $\sin x = \sin x - \sin 0 = \cos(c)x$, and then using Theorem 4.4.

4.21. Evaluate $\lim_{x\to 0} \frac{x}{e^x - 1}$, first by recognizing the quotient as a reciprocal derivative, and then using Theorem 4.4.

4.22. Evaluate $\lim_{x\to 0} \frac{\sin x - x\cos x}{x^3}$ using Theorem 4.4 twice.

4.23. Substitute $f(x) = F(1/x)$ and $g(x) = G(1/x)$ into Theorem 4.4 to prove the following version of Theorem 4.4.

Suppose $\lim_{y\to\infty} F(y) = 0$, $\lim_{y\to\infty} G(y) = 0$, and that $\lim_{y\to\infty} \frac{F'(y)}{G'(y)}$ exists. Then

$$\lim_{y\to\infty} \frac{F(y)}{G(y)} = \lim_{y\to\infty} \frac{F'(y)}{G'(y)}.$$

You will need to take $a = 0$ in the theorem. Explain how to extend f and g as odd functions, so that the theorem can be applied.

4.24. Use the result of Problem 4.23 and the exponential growth theorem where needed to evaluate the following limits.

(a) $\lim_{y\to\infty} e^{-1/y}$

(b) $\lim_{y\to\infty} y^2 e^{-y}$

(c) $\lim_{y\to\infty} \frac{e^{-y}}{1 - e^{-1/y}}$

4.2 Higher Derivatives

Many of the functions f we have presented in examples so far have the property that their derivatives f' also turned out to be differentiable. Such functions are called *twice differentiable*. Similarly, we define a three-times differentiable function f as one whose second derivative is differentiable.

Definition 4.2. A function f is called *n times differentiable* at x if its $(n-1)$st derivative is differentiable at x. The resulting function is called the nth derivative of f and is denoted by

$$f^{(n)} \quad \text{or} \quad \frac{d^n f}{dx^n}.$$

Definition 4.3. A function f is called *continuously differentiable* on an interval if f' exists and is continuous on the interval. A function f is *n times continuously differentiable* if the nth derivative exists and is continuous on the interval.

As we saw in Chap. 3, if $x(t)$ denotes the position of a particle at time t, then the rate of change of position $\frac{dx}{dt}$ is the velocity v of the particle. The derivative of velocity is called *acceleration*. Thus

$$\text{acceleration} = \frac{dv}{dt} = \frac{d^2x}{dt^2};$$

in words, *acceleration is the second derivative of position*.

The geometric interpretation of the second derivative is no less interesting than the physical interpretation. We note that a linear function $f(x) = mx + b$ has second derivative zero. Therefore, a function with nonzero second derivative is not linear. Since a linear function can be characterized as one whose graph is a straight line, it follows that if $f'' \neq 0$, then the graph of f is not flat but curved. This fact suggests that the size of $f''(x)$ measures in some sense the deviation of the graph of f from a straight line at the point x.

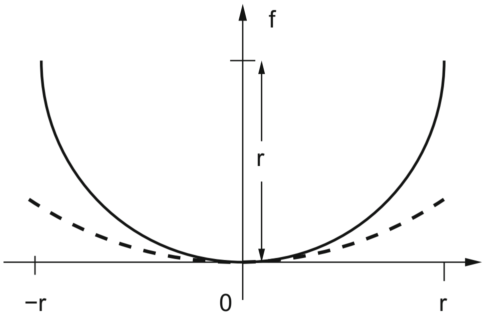

Fig. 4.8 The second derivative illustrated: large r corresponds to small curvature. The dotted arc has a larger value of r. See Example 4.10

Example 4.10. The graph of the function

$$f(x) = r - \sqrt{r^2 - x^2}, \qquad -r < x < r,$$

is a semicircle of radius r; see Fig. 4.8. The larger the value of r, the closer this semicircle lies to the x-axis, for values of x in a fixed interval about the origin, as illustrated in Fig. 4.8. We have

$$f'(x) = \frac{x}{\sqrt{r^2 - x^2}}, \qquad f''(x) = \frac{1}{\sqrt{r^2 - x^2}} + \frac{x^2}{(r^2 - x^2)^{3/2}} = \frac{r^2}{(r^2 - x^2)^{3/2}}.$$

The value of f'' at $x = 0$ is $f''(0) = \dfrac{r^2}{(r^2)^{3/2}} = \dfrac{1}{r}$. The larger r is, the smaller is the value of $f''(0)$, so in this case, the smallness of $f''(0)$ indeed indicates that the graph of f is close to a straight line.

What Does f'' Tell Us About f? The goal of this section is to explain the following result: *over a short interval, every twice continuously differentiable function can be exceedingly well approximated by a quadratic polynomial*. Note the simplification implied here, that a complicated function can sometimes be replaced by a simple one.

We shall use the monotonicity criterion to relate knowledge about f'' to knowledge about f. For example, if $f'' > 0$, then by monotonicity, f' is increasing. Graphically, this means that the slopes of the tangents to the graph of f are increasing as you move from left to right. Similarly, if $f'' < 0$, then f' is decreasing, and the slopes of the tangents to the graph of f are decreasing as you move from left to right. In Fig. 4.9, some quick sketches of tangent lines with increasing (and decreasing) slopes suggest that the graph of the underlying function opens upward if $f'' > 0$, and that it opens downward if $f'' < 0$.

Rather than trust a few sketches, we shall investigate this question: if we have information about f'', what can we say about f itself? Suppose we have an estimate for f'',

$$m \le f''(x) \le M \quad \text{on } [a,b]. \tag{4.10}$$

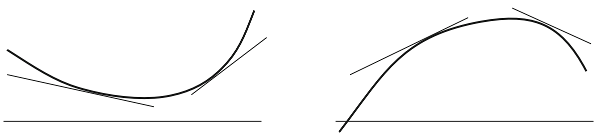

Fig. 4.9 *Left*: $f'' > 0$ and increasing slopes. *Right*: $f'' < 0$ and decreasing slopes

The inequality on the right is equivalent to $M - f''(x) \ge 0$. Note that $M - f''(x)$ is the derivative of $Mx - f'(x)$. So by the monotonicity criterion, $Mx - f'(x)$ is a nondecreasing function, and we conclude that

$$Ma - f'(a) \le Mx - f'(x) \text{ on } [a,b].$$

This inequality can be rewritten as follows:

$$f'(x) - f'(a) \le M(x-a) \text{ on } [a,b].$$

Note that the function on left-hand side is the derivative of $f(x) - xf'(a)$, and the function on the right-hand side is the derivative of $\frac{1}{2}M(x-a)^2$. Taking their difference, again by the monotonicity criterion it follows that $\frac{1}{2}M(x-a)^2 - (f(x) - xf'(a))$ is a nondecreasing function. Since a is less than or equal to x, we have

$$\frac{1}{2}M(a-a)^2-(f(a)-af'(a)) \le \frac{M}{2}(x-a)^2-(f(x)-xf'(a)).$$

Rewrite this last inequality by taking the term $f(x)$ to the left-hand side and all other terms to the right-hand side, giving

$$f(x) \le f(a)+f'(a)(x-a)+\frac{M}{2}(x-a)^2.$$

Remark. This is the first step in our stated goal, since the function on the right-hand side is a quadratic polynomial.

By an analogous argument we can deduce from $m \le f''(x)$ and repeated uses of monotonicity that

$$f(a)+f'(a)(x-a)+\frac{m}{2}(x-a)^2 \le f(x)$$

for all x in $[a,b]$. We can combine the two inequalities into one statement. If $f''(x)$ is bounded below by m and above by M on $[a,b]$, then f itself is bounded below and above by two quadratic polynomials:

$$f(a)+f'(a)(x-a)+\frac{m}{2}(x-a)^2 \le f(x) \le f(a)+f'(a)(x-a)+\frac{M}{2}(x-a)^2. \quad (4.11)$$

The upper and lower bounds differ inasmuch as one contains the constant m and the other M. It follows that there is a number H between m and M such that

$$f(x)=f(a)+f'(a)(x-a)+\frac{H}{2}(x-a)^2. \quad (4.12)$$

Suppose next that f'' is continuous on $[a,b]$ and that m and M are the minimum and maximum values of f'' on the interval. Take $x=b$ in Eq. (4.12). It follows again from the intermediate value theorem that there is a point c between a and b such that

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

This equation provides a rich source of observations about f, and we obtain the following generalization of the mean value theorem:

Theorem 4.5. Linear approximation. *Let f be twice continuously differentiable on an interval containing a and b. Then there is a point c between a and b such that*

$$f(b)=f(a)+f'(a)(b-a)+\frac{f''(c)}{2}(b-a)^2. \quad (4.13)$$

We proved the linear approximation theorem for $a<b$. It is also true for $a>b$. The proof is outlined in Problem 4.32.

Example 4.11. Let us use Theorem 4.5 to estimate $\log(1.1)$. Let $f(x) = \log(1+x)$. Then $f'(x) = \dfrac{1}{1+x}$ and $f''(x) = -\dfrac{1}{(1+x)^2}$. Taking $a = 0$ and $b = 0.1$, we get $f(0) = 0$, $f'(0) = 1$, and

$$\log(1.1) = 0 + 1(0.1-0) - \frac{1}{(1+c)^2}\frac{(0.1)^2}{2},$$

where c is a number between 0 and 0.1. Since $-\dfrac{1}{(1)^2} \le f''(c) \le -\dfrac{1}{(1.1)^2}$, we get

$$0.095 = 0.1 - \frac{0.01}{2} \le \log(1.1) \le 0.1 - \frac{1}{(1.1)^2}\frac{0.01}{2} = 0.0958\ldots < 0.096.$$

See what your calculator says about the natural logarithm of 1.1.

Example 4.12. Let $f(x) = \log(1+x)$. We approximate f by two quadratic polynomials on $[0, 0.5]$. From Example 4.11, we have $f(0) = 0$, $f'(0) = 1$, the minimum of $f''(x)$ on $[0, 0.5]$ is -1, and the maximum is $-\dfrac{1}{(1+0.5)^2} = -\dfrac{4}{9}$. Therefore (See Fig. 4.10),

$$x - \frac{x^2}{2} \le \log(1+x) \le x - \frac{4}{9}\frac{x^2}{2}, \quad (0 \le x \le 0.5).$$

In the linear approximation theorem, suppose b is close to a. Then c is even closer to a, and since f'' is continuous, $f''(c)$ is close to $f''(a)$. We express this by writing

$$f''(c) = f''(a) + s,$$

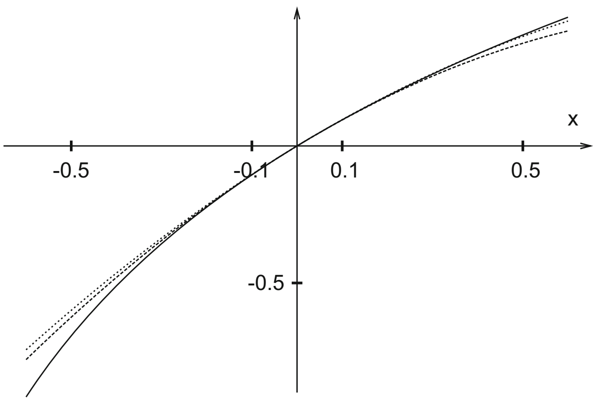

Fig. 4.10 The graphs of $\log(1+x)$ and the two quadratic polynomials of Example 4.12

where s denotes a quantity that is small when b is close to a. Substituting this into Eq. (4.13), we get

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

This formula shows that for b close to a, the first three terms on the right are a very good approximation to $f(b)$. Therefore, it follows from the linear approximation theorem, as we have stated, that *over a short interval, every twice continuously differentiable function can be exceedingly well approximated by a quadratic polynomial*.

4.2a Second Derivative Test

The linear approximation theorem, Theorem 4.5, has applications to optimization. The next two theorems are sometimes referred to as the *second derivative test* for local extrema:

Theorem 4.6. Local minimum theorem. *Let f be a twice continuously differentiable function on an open interval containing a, and suppose that $f'(a) = 0$ and $f''(a) > 0$. Then f has a local minimum at a, i.e.,*

$$f(a) < f(b)$$

for all points $b \neq a$ sufficiently close to a.

Proof. We have $f''(a) > 0$, so by the continuity of f'', $f''(x) > 0$ for all x close enough to a. Choose b so close to a that $f''(c) > 0$ for all c between a and b. According to the linear approximation theorem, since $f'(a) = 0$ and $f''(c) > 0$, we get

$$f(b) = f(a) + \frac{f''(c)}{2}(b-a)^2 > f(a),$$

as asserted. □

We suggest a way for you to prove the analogous maximum theorem in Problem 4.38:

Theorem 4.7. Local maximum theorem. *Let f be a twice continuously differentiable function defined on an open interval containing a, and suppose that $f'(a) = 0$ and $f''(a) < 0$. Then f has a local maximum at a, i.e.,*

$$f(a) > f(b)$$

for all points $b \neq a$ sufficiently close to a.

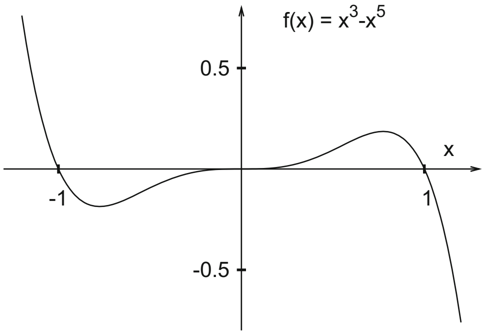

Fig. 4.11 A local maximum and a local minimum, in Example 4.13

Example 4.13. The polynomial $f(x) = x^3 - x^5$ has $f'(x) = x^2(3 - 5x^2)$, which is 0 at three numbers $x_1 = -\sqrt{\frac{3}{5}}$, $x_2 = 0$, and $x_3 = \sqrt{\frac{3}{5}}$. The second derivative is $f''(x) = 6x - 20x^3 = 2x(3 - 10x^2)$. So

$$f''(x_1) = -2\sqrt{\frac{3}{5}}(3-6) > 0, \;\; f''(x_2) = 0, \;\; f''(x_3) = 2\sqrt{\frac{3}{5}}(3-6) < 0.$$

We conclude that f has a local minimum at x_1 and a local maximum at x_3. However, $f''(x_2) = 0$ does not give any information about the possibility of local extrema at x_2. The graph of f is drawn in Fig. 4.11.

4.2b Convex Functions

We give further applications of the linear approximation theorem.

Suppose f'' is nonnegative in an interval containing a and b. Then the last term on the right in Eq. (4.13) is nonnegative, so omitting it yields the inequality

$$f(b) \geq f(a) + f'(a)(b-a).$$

This inequality has a striking geometric interpretation. We notice that the quantity on the right is the value at b of the linear function

$$l(x) = f(a) + f'(a)(x-a).$$

The graph of this linear function is the line tangent to the graph of f at $(a, f(a))$. So

$$f(b) \geq f(a) + f'(a)(b-a)$$

asserts that the graph of f lies above its tangent lines.

Definition 4.4. A function whose graph lies above its tangents is called *convex*.

In this language, the inequality says that *every function whose second derivative is positive is convex.* Convex functions have another interesting property:

Theorem 4.8. Convexity theorem. *Let f be a twice continuously differentiable function on $[a,b]$, and suppose that $f'' > 0$ there. Then for every x satisfying $a < x < b$,*

$$f(x) < f(b)\frac{x-a}{b-a} + f(a)\frac{b-x}{b-a}. \tag{4.14}$$

This theorem has an illuminating geometric interpretation. Denote by $\ell(x)$ the function on the right-hand side of inequality (4.14). Then ℓ is a linear function whose values at $x = a$ and at $x = b$ agree with the values of f at these points. Thus the graph of ℓ is the *secant line* of f on $[a,b]$. Therefore, inequality (4.14) says that *the graph of a convex function f on an interval $[a,b]$ lies below the secant line* (Fig. 4.12).

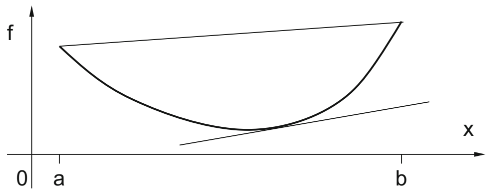

Fig. 4.12 The graph of a convex function lies above its tangent lines and below the secant on $[a,b]$

Proof. We wish to show that $f(x) - \ell(x) \le 0$ on $[a,b]$. According to the extreme value theorem, Theorem 2.6, $f - \ell$ reaches a maximum at some point c in $[a,b]$. The point c could be either at an endpoint or in (a,b). We show now that c is not in (a,b). For if it were, then the first derivative of $f - \ell$ would be zero at c. The second derivative of $f(x) - \ell(x)$ at c is given by

$$f''(c) - \ell''(c) = f''(c) - 0,$$

since ℓ is linear. We have assumed that f'' is positive. According to the local minimum theorem, Theorem 4.6, the function $f - \ell$ has a local minimum at c. This shows that the point c where the maximum occurs cannot be in the interior of $[a,b]$.

The only alternative remaining is that c is one of the endpoints. At an endpoint, f and ℓ have the same value. This shows that the maximum of $f - \ell$ is 0, and that at all points x of $[a,b]$ other than the endpoints,

$$f(x) - \ell(x) < 0.$$

This completes the proof of the convexity theorem. □

Definition 4.5. A function whose graph lies *below* its tangent is called *concave*.

The following analogues of results for convex functions hold: *every function whose second derivative is negative is concave, and the graph of a concave function lies above its secant* (Fig. 4.13).

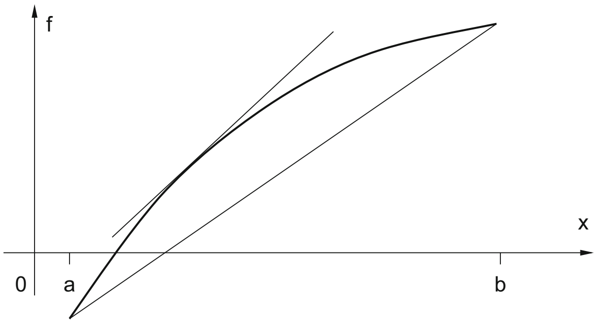

Fig. 4.13 The graph of a concave function f lies above the secant on $[a,b]$, and below each of the tangent lines

Example 4.14. We have seen in Example 4.11 that the second derivative of the function $\log(1+x)$ is negative. It follows that $\log(1+x)$ is a concave function.

Problems

4.25. A particle has position $x = f(t)$, and at time $t = 0$, the position and velocity are 0 and 3, respectively.. The acceleration is between 9.8 and 9.81 for all t. Give bounds on $f(t)$.

4.26. Recall from the chain rule that if f and g are differentiable inverse functions, $f(g(x)) = x$, then

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

Find a relation for the second derivatives.

4.27. Find all local extreme values of $f(x) = 2x^3 - 3x^2 + 12x$. On what intervals is f convex? concave? Sketch a graph of f based on this information.

4.28. Over which intervals are the following functions convex?

(a) $f(x) = 5x^4 - 3x^3 + x^2 - 1$
(b) $f(x) = \dfrac{x+1}{x-1}$
(c) $f(x) = \sqrt{x}$
(d) $f(x) = \dfrac{1}{\sqrt{x}}$
(e) $f(x) = \sqrt{1-x^2}$
(f) $f(x) = \mathrm{e}^{-x^2}$

4.29. Are the linear approximations to $f(x) = x^2 - 3x + 5$ above or below the graph?

4.30. Is the secant line for $f(x) = -x^2 - 3x + 5$ on $[0,7]$ above or below the graph?

4.31. Find an interval $(0,b)$ where $\mathrm{e}^{-1/x}$ is convex. Sketch the graph on $(0,\infty)$.

4.32. We proved the linear approximation theorem, Theorem 4.5, for a twice continuously differentiable function f on an interval containing a and b, where $a < b$,

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(c)}{2}(b-a)^2. \tag{4.15}$$

In this problem we show how the case $a > b$ follows from this. Given f'' continuous on $[a,b]$, define the function g as $g(x) = f(a+b-x)$.

(a) Show that g is defined in the interval $[a,b]$.
(b) Show that g'' is continuous in $[a,b]$.
(c) Show that

$$g(a) = f(b), \quad g'(a) = -f'(b), \quad g''(a) = f''(b),$$
$$g(b) = f(a), \quad g'(b) = -f'(a), \quad g''(b) = f''(a).$$

(d) Write equation (4.15) for the function g. Then use results from part (c) to conclude that Eq. (4.15) holds for $b < a$.

4.33. Is e^f convex when f is convex?

4.34. Give an example of convex functions f and g for which $f \circ g$ is not convex.

4.35. Show, using the linear approximation theorem, that for f'' continuous on an interval containing a and b,

$$\frac{f(a)+f(b)}{2} \quad \text{differs from } f\left(\frac{a+b}{2}\right) \quad \text{by less than } \frac{M}{8}(b-a)^2,$$

where M is an upper bound for $|f''|$ on $[a,b]$.

4.36. Suppose f'' is continuous on an interval that contains a and b. Use the linear approximation theorem to explain why

$$\frac{f(b)-f(a)}{b-a} = \frac{f'(b)+f'(a)}{2} + s(b-a),$$

where s is small when b is close to a.

4.37. Let f have continuous first and second derivatives in $a < x < b$. Prove that

(a) $f'(x) = \lim_{h\to 0} \dfrac{f(x+h)-f(x-h)}{2h}$
(b) $f''(x) = \lim_{h\to 0} \dfrac{f(x+h)-2f(x)+f(x-h)}{h^2}$

4.38. Prove Theorem 4.7 by applying Theorem 4.6 to the function $-f$.

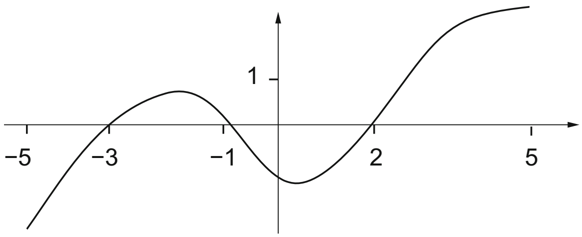

Fig. 4.14 The graph of f in Problems 4.39 and 4.40

4.39. The graph of a function f is given on $[-5,5]$ in Fig. 4.14. Use the graph to find, approximately, the intervals on which $f' > 0$, $f' < 0$, $f'' > 0$, $f'' < 0$.

4.40. Use approximations (see Problem 4.37)

$$f'(x) \approx \frac{f(x+h)-f(x-h)}{2h}, \qquad f''(x) \approx \frac{f(x+h)-2f(x)+f(x-h)}{h^2},$$

with $h = 1$, to estimate $f'(-1)$ and $f''(0.5)$ for the function graphed in Fig. 4.14.

4.41. Let $f(x) = e^{-\frac{x^2}{2}}$ for all x and $g(x) = e^{-1/x}$ for $x > 0$.

(a) Use your calculator or computer to graph f and g.
(b) Use calculus to find the intervals on which f is increasing, decreasing, convex, concave, and locate any extreme values or critical points.
(c) Use calculus to find the intervals on which g is increasing, decreasing, convex, concave.

4.3 Taylor's Theorem

We saw in Sect. 4.2 that bounds on the second derivative, $m \le f''(x) \le M$ in $[a,b]$, enabled us to find two quadratic polynomial functions that bound f:

$$f(a)+f'(a)(x-a)+\frac{m}{2}(x-a)^2 \le f(x) \le f(a)+f'(a)(x-a)+\frac{M}{2}(x-a)^2.$$

Now we are ready to tackle the general problem: if we are given upper and lower bounds for the nth derivative $f^{(n)}(x)$ on $[a,b]$, find nth-degree polynomial functions that are upper and lower bounds for $f(x)$. Generalizing the result we obtained for second derivatives, we surmise that the following result holds:

Theorem 4.9. Taylor's inequality. *Suppose that f is an n-times continuously differentiable function on $[a,b]$, and denote by m and M the minimum and maximum, respectively, of $f^{(n)}$ over $[a,b]$; that is,*

$$m \le f^{(n)}(x) \le M, \qquad x \text{ in } [a,b].$$

Then Taylor's inequality

$$f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1}+\frac{m}{n!}(x-a)^n \tag{4.16}$$

$$\le f(x)$$

$$\le f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1}+\frac{M}{n!}(x-a)^n$$

holds for all x in $[a,b]$.

The polynomials on the left and right sides of Taylor's inequality (4.16) are identical up through the next-to-last terms. We call the identical parts Taylor polynomials.

Definition 4.6. If f is n times differentiable at a, the *Taylor polynomials* at a are

$$\begin{aligned}
t_0(x) &= f(a)\\
t_1(x) &= f(a)+f'(a)(x-a)\\
t_2(x) &= f(a)+f'(a)(x-a)+f''(a)\frac{(x-a)^2}{2!}\\
&\cdots\\
t_n(x) &= f(a)+f'(a)(x-a)+f''(a)\frac{(x-a)^2}{2!}+\cdots+f^{(n)}(a)\frac{(x-a)^n}{n!}\\
&\cdots
\end{aligned}$$

Proof of Theorem 4.9. We prove Taylor's inequality for all n inductively, i.e., we first show that it is true for $n=1$, and then we show that if the result is true for any particular number n, then it is true for $n+1$. By the mean value theorem, we know that for some c between a and x,

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

Since f' is continuous on $[a,b]$, it attains a maximum M and minimum m on that interval. Then since $a \le x$, it follows that

$$f(a) + m(x-a) \le f(x) \le f(a) + M(x-a), \qquad (a \le x \le b).$$

Thus the theorem holds for $n=1$. Next, we show that if the result holds for n, then it holds for $n+1$. Assume that Taylor's inequality holds for every function whose nth derivative is bounded on $[a,b]$. If f is $(n+1)$ times continuously differentiable, then there are bounds

$$m \le f^{(n+1)}(x) \le M, \qquad (a \le x \le b).$$

Since $f^{(n+1)}$ is the nth derivative of f', we can apply the inductive hypothesis to the function f' to obtain

$$f'(a) + f''(a)(x-a) + \cdots + \frac{m}{n!}(x-a)^n \le f'(x) \le f'(a) + f''(a)(x-a) + \cdots + \frac{M}{n!}(x-a)^n.$$

The sum on the right is the derivative of

$$t_n(x) + \frac{M}{(n+1)!}(x-a)^{n+1}.$$

Since

$$\left(t_n(x) + \frac{M}{(n+1)!}(x-a)^{n+1}\right)' - f'(x) \ge 0,$$

we see that

$$\left(t_n(x) + \frac{M}{(n+1)!}(x-a)^{n+1}\right) - f(x)$$

is a nondecreasing function on $[a,b]$. At $x=a$, the difference

$$\left(t_n(a) + \frac{M}{(n+1)!}(a-a)^{n+1}\right) - f(a)$$

is zero. So for $x > a$,

$$0 \le \left(t_n(x) + \frac{M}{(n+1)!}(x-a)^{n+1}\right) - f(x).$$

It follows that

$$f(x) \le t_n(x) + \frac{M}{(n+1)!}(x-a)^{n+1},$$

which is the right half of Taylor's inequality. The left half follows in a similar manner. Thus, we have shown that if Taylor's inequality holds for n, it holds for $n+1$. Since the inequality holds for $n = 1$, by induction it must hold for all positive integers. □

Example 4.15. We write Taylor's inequality for $f(x) = \sin x$ on $[0,4]$, where $a = 0$ and $n = 5$. The first four derivatives are

$$f'(x) = \cos x, \;\; f''(x) = -\sin x, \;\; f'''(x) = -\cos x, \;\; f''''(x) = \sin x.$$

Evaluate at $a = 0$: $f(0) = 0,\; f'(0) = 1,\; f''(0) = 0,\; f'''(0) = -1,\; f''''(0) = 0.$ Then the fourth Taylor polynomial $t_4(x) = x - \frac{x^3}{3!}$. We have $f^{(5)}(x) = \cos x$, so $-1 \le f^{(5)}(x) \le 1$ on $[0,4]$. Therefore,

$$x - \frac{x^3}{3!} - \frac{x^5}{5!} \le \sin x \le x - \frac{x^3}{3!} + \frac{x^5}{5!}.$$

Figure 4.15 contains the graphs of the three functions.

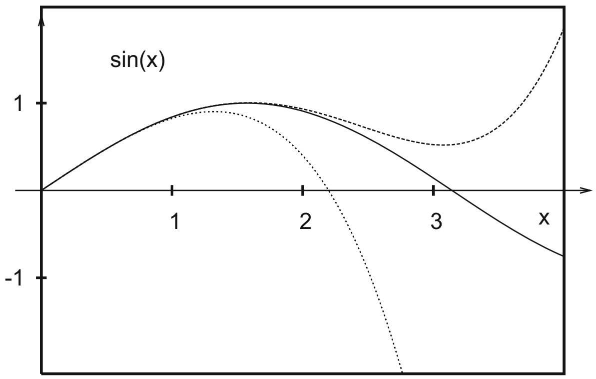

Fig. 4.15 Taylor's inequality for the case in Example 4.15: $x - \frac{x^3}{6} - \frac{x^5}{120} \le \sin x \le x - \frac{x^3}{6} + \frac{x^5}{120}$, where $0 \le x \le 4$

Example 4.16. Let us write Taylor's inequality for $f(x) = \log x$ on the interval $[1,3]$, where $a = 1$ and $n = 4$:

$$f'(x) = \frac{1}{x}, \;\; f''(x) = -\frac{1}{x^2}, \;\; f'''(x) = 2!\frac{1}{x^3}, \;\; f''''(x) = -3!\frac{1}{x^4},$$

$$f(1) = 0, \qquad f'(1) = 1, \qquad f''(1) = -1, \qquad f'''(1) = 2,$$

and since $f''''(x)$ is increasing, $-3! \le f''''(x) \le -3!\frac{1}{3^4}$. According to Taylor's inequality (see Fig. 4.16),

$$(x-1)-\frac{1}{2}(x-1)^2+\frac{1}{3}(x-1)^3-\frac{1}{4}(x-1)^4$$
$$\le \log x \le (x-1)-\frac{1}{2}(x-1)^2+\frac{1}{3}(x-1)^3-\frac{1}{3^4 4}(x-1)^4.$$

Just as we saw in Sect. 4.2, the upper and lower bounds in Taylor's inequality differ inasmuch as one contains the constant m and the other M. So, given $x > a$, there is a number H between m and M such that

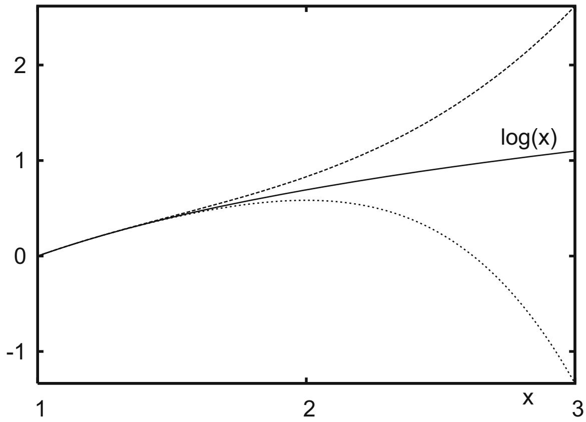

Fig. 4.16 Taylor's inequality for $\log x$ as in Example 4.16

$$f(x) = t_{n-1}(x) + H\frac{(x-a)^n}{n!}.$$

According to the intermediate value theorem, every number H between the minimum m and maximum M of the continuous function $f^{(n)}$ is taken on at some point c between a and b. Now when $x = b$, $f(b) = t_{n-1}(b) + f^{(n)}(c)\frac{(b-a)^n}{n!}$. The difference

$$f(b) - t_{n-1}(b) = f^{(n)}(c)\frac{(b-a)^n}{n!}$$

is called the *remainder*. We express our results with the following theorem.

Theorem 4.10. Taylor's formula with remainder. *Let f be an n-times continuously differentiable function on an interval containing a and b. Then*

$$f(b) = f(a) + f'(a)(b-a) + \cdots + f^{(n-1)}(a)\frac{(b-a)^{n-1}}{(n-1)!} + f^{(n)}(c)\frac{(b-a)^n}{n!}, \tag{4.17}$$

where c lies between a and b.

In the derivation of this theorem we have exploited the fact that $a < b$. It is not hard to show that the theorem remains true if $a > b$. We ask you to do so in Problem 4.50. Here are some applications of Taylor's formula.

Example 4.17. Let $f(x) = x^m$, m a positive integer. Then

$$f^{(k)}(x) = m(m-1)\cdots(m-k+1)x^{m-k}.$$

In particular, $f^{(m)}(x) = m!$, and higher derivatives are 0. Therefore, according to Taylor's formula with $b = 1+y$, $a = 1$, and any $n \geq m$,

$$(1+y)^m = 1 + my + \frac{m(m-1)}{2!}y^2 + \cdots + y^m = \sum_{k=0}^{m}\binom{m}{k}y^k.$$

Example 4.17 is nothing but the binomial expansion, revealed here as a special case of Taylor's formula.

Taylor's inequality

$$t_{n-1}(x) + m\frac{(x-a)^n}{n!} \leq f(x) \leq t_{n-1}(x) + M\frac{(x-a)^n}{n!}$$

is an approximation to f on $[a,b]$. The polynomials on the left- and right-hand sides of Taylor's inequality differ only in the last terms. That difference is due to the variation in the maximum and minimum value of $f^{(n)}$ on $[a,b]$, which leads to the next definition.

Definition 4.7. Denote by C_n the *oscillation* of $f^{(n)}$ on the interval $[a,b]$, i.e.,

$$C_n = M_n - m_n,$$

where M_n is the maximum, m_n the minimum of $f^{(n)}$ over $[a,b]$.

We derive now a useful variant of Taylor's inequality. Taylor's formula (4.17),

$$f(b) = f(a) + f'(a)(b-a) + \cdots + f^{(n-1)}(a)\frac{(b-a)^{n-1}}{(n-1)!} + f^{(n)}(c)\frac{(b-a)^n}{n!},$$

differs from Taylor's polynomial

$$t_n(b) = f(a) + f'(a)(b-a) + \cdots + f^{(n-1)}(a)\frac{(b-a)^{n-1}}{(n-1)!} + f^{(n)}(a)\frac{(b-a)^n}{n!}$$

in that the last term has $f^{(n)}$ evaluated at c rather than a. Since $f^{(n)}(c)$ and $f^{(n)}(a)$ differ by at most the oscillation C_n, we see that

$$|f(x) - t_n(x)| \le \frac{C_n}{n!}(x-a)^n \le \frac{C_n}{n!}(b-a)^n \qquad \text{for all } x \text{ in } [a,b]. \tag{4.18}$$

Suppose the function f is *infinitely differentiable*, i.e., has derivatives of all orders. Suppose further that

$$\lim_{n\to\infty} \frac{C_n}{n!}(b-a)^n = 0. \tag{4.19}$$

Then as n gets larger and larger, $t_n(x)$ tends to $f(x)$. We can state this result in the following spectacular form.

Theorem 4.11. Taylor's theorem. *Let f be an infinitely differentiable function on an interval $[a,b]$. Denote by C_n the oscillation of $f^{(n)}$, and suppose that $\lim_{n\to\infty} \frac{C_n}{n!}(b-a)^n = 0$. Then f can be represented at every point of $[a,b]$ by the Taylor series*

$$f(x) = \lim_{n\to\infty} t_n(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(a)(x-a)^k,$$

and the Taylor polynomials converge uniformly to f on $[a,b]$. There is an analogous theorem for the interval $[b,a]$ when $b < a$.

Proof. The meaning of the infinite sum on the right is this: Form the nth Taylor polynomials $t_n(x)$ of f at a and take the limit of this sequence of functions as n tends to infinity. Since $|f(x) - t_n(x)| \le \frac{C_n}{n!}(b-a)^n$ and $\lim_{n\to\infty} \frac{C_n}{n!}(b-a)^n = 0$, the sequence $t_n(x)$ tends to $f(x)$ as n tends to infinity. The convergence is uniform on $[a,b]$, because the estimate that we derived for $|f(x) - t_n(x)|$ does not depend on x.

We ask you to verify the proof of the theorem in the case $b < a$ in Problem 4.50. □

4.3a Examples of Taylor Series

The Sine. Let $f(x) = \sin x$ and $a = 0$. The derivatives are

$$f(x) = \sin x,\ \ f'(x) = \cos x,\ \ f''(x) = -\sin x,\ \ f'''(x) = -\cos x,\ \ f''''(x) = \sin x,$$

and so forth. So at 0,

$$f(0) = 0, \quad f'(0) = 1, \quad f''(0) = 0, \quad f'''(0) = -1, \quad f''''(0) = 0, \quad f^{(5)}(0) = 1,$$

etc. The nth Taylor polynomial at $a = 0$ for $\sin x$ is (Fig. 4.17)

$$t_n(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + \sin^{(n)}(0)\frac{x^n}{n!},$$

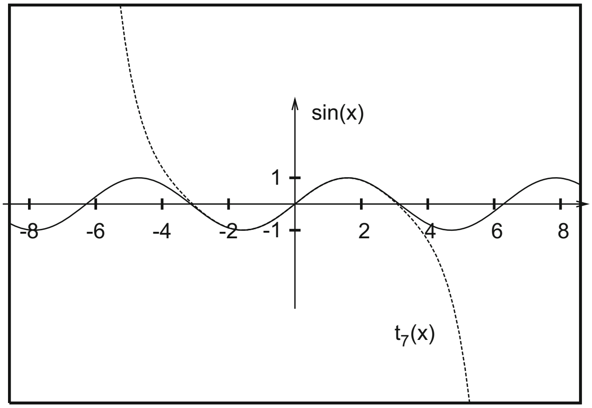

Fig. 4.17 Graphs of $\sin x$ and the Taylor polynomial $t_7(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}$

where the last coefficient is 0, 1, or -1, depending on n. The oscillation C_n is equal to 2, because the sine and cosine have minimum -1 and maximum 1. Then on $[0,b]$,

$$C_n \frac{(b-a)^n}{n!} = 2\frac{b^n}{n!}.$$

For any number b, the terms on the right tend to 0 as n tends to infinity by Example 1.17. Therefore, the Taylor series for $\sin x$ at $a = 0$,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots, \tag{4.20}$$

converges for all x in $[0,b]$ for every b, and therefore on $[0,\infty)$. By a similar argument, $\sin x$ converges for all x in $[b,0]$ when $b < 0$, and therefore on $(-\infty,\infty)$.

The Logarithm. Let $f(x) = \log x$ and $a = 1$. As we saw in Example 4.16, the derivatives follow a pattern,

$$f(x) = \log x,\ \ f'(x) = x^{-1},\ \ f''(x) = -x^{-2},\ \ f'''(x) = 2!x^{-3},\ \ f''''(x) = -3!x^{-4},\ldots.$$

At $a = 1$,

$$f(1) = 0,\ \ f'(1) = 1,\ \ f''(1) = -1,\ \ f'''(1) = 2!,\ \ f''''(1) = -3!,\ldots.$$

The nth Taylor polynomial of $\log x$ at $a = 1$ is

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

In contrast to the case of the sine, the oscillation of $f^{(n)}(x)$ on $[1,b]$ depends on the value of b. Each derivative $f^{(n)}(x) = (-1)^{n-1}(n-1)!x^{-n}$ is monotonic, so

$$C_n = |f^{(n)}(1) - f^{(n)}(b)| = (n-1)!|1-b^{-n}|.$$

On $[1,b]$,

$$|\log(x) - t_n(x)| \le C_n \frac{(b-1)^n}{n!} = |1-b^{-n}|\frac{(b-1)^n}{n}. \tag{4.21}$$

Since $b > 1$, the first factor $|1-b^{-n}|$ tends to 1 as n tends to infinity. How $\frac{(b-1)^n}{n}$ behaves depends on the size of b: for $1 < b \le 2$, $\frac{(b-1)^n}{n}$ tends to 0 as n tends to infinity. If $b > 2$, then $(b-1) > 1$, and we know from the exponential growth theorem, Theorem 2.10, that $\frac{(b-1)^n}{n}$ tends to infinity. Hence the Taylor series converges to $\log x$ in $[1,2]$.

We show by another method in Example 7.35 that $|\log x - t_n(x)|$ also tends to 0 for $0 < x \le 1$. Given that future result, we have

$$\log x = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(x-1)^n}{n} \qquad (0 < x \le 2).$$

Figure 4.18 shows part of the graphs of $\log x$ and $t_4(x)$.

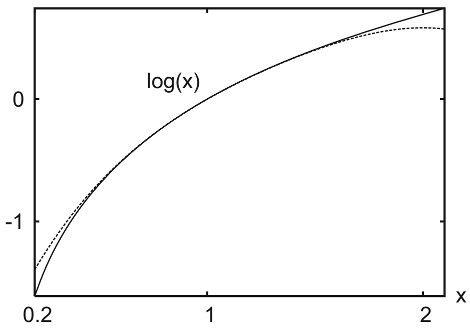

Fig. 4.18 Graphs of $\log x$ and the Taylor polynomial $t_4(x) = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4}$

Remark. According to the ratio test, using

$$\lim_{n\to\infty} \left| \frac{\frac{(x-1)^{n+1}}{n+1}}{\frac{(x-1)^n}{n}} \right| = |x-1| < 1,$$

we see that the Taylor series for the logarithm, $\lim_{n\to\infty} t_n(x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(x-1)^n}{n}$, converges uniformly on every closed interval in $(0,2)$. Checking the endpoints, we see that the power series converges at $x = 2$ (alternating series theorem) and diverges at $x = 0$ (harmonic series). But this does not tell us that $t_n(x)$ converges to $\log x$ for $0 < x \le 2$. To show that a Taylor series $\sum_{n=0}^{\infty} f^{(n)}(a)\frac{(x-a)^n}{n!}$ converges to the function f from which it is derived, it is necessary to show that $|f(x) - t_n(x)|$ tends to 0 as n tends to infinity. Examining the oscillation is one way to do this. Another way is to examine the behavior of the remainders

$$|f(b) - t_n(b)| = \left| f^{(n+1)}(c)\frac{(b-a)^{n+1}}{n+1!} \right|$$

from Taylor's formula, as we do in the next example. After we study integration in Chap. 7, we will have an integral formula for the remainder that gives another way to estimate the remainder.

The Exponential Function. Let $f(x) = e^x$ and $a = 0$. Since $f^{(n)}(x) = e^x$, it follows that

$$f(0) = 1, \;\; f'(0) = 1, \;\; f''(0) = 1, \;\; f'''(0) = 1, \ldots,$$

and the nth Taylor polynomial for e^x at $a = 0$ is

$$t_n(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}.$$

We want to show that $\lim_{n\to\infty} |e^x - t_n(x)| = 0$ for all x. According to Taylor's formula,

$$|e^x - t_n(x)| = \left| e^c \frac{x^{n+1}}{n+1!} \right|$$

for some c between 0 and x. Suppose x is in $[-b, b]$. Then

$$|e^x - t_n(x)| \le e^b \frac{b^{n+1}}{n+1!}.$$

We saw in Example 1.17 that $\lim_{n\to\infty} \frac{b^n}{n!} = 0$ for every number b. Hence $e^b \frac{b^{n+1}}{n+1!}$ tends to 0. The Taylor series

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

converges to e^x for all x in $[-b, b]$. Since b is arbitrary, the series converges to e^x for all x. Figure 4.19 shows graphs of e^x, $t_3(x)$, and $t_4(x)$.

The Binomial Series. Here we point out that the binomial expansions as in Example 4.17 have a generalization to every real exponent. We prove the following.

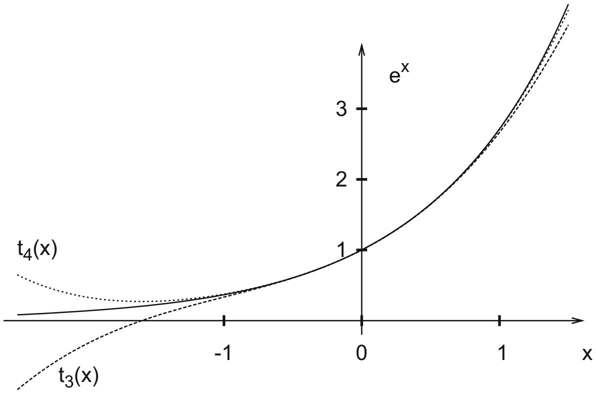

Fig. 4.19 Graphs of e^x with Taylor polynomials $t_3(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}$ and $t_4(x) = t_3(x) + \frac{x^4}{4!}$ on $[-2.5, 1.5]$

Theorem 4.12. The binomial theorem. *If ℓ is any number and $|y| < 1$, then*

$$(1+y)^\ell = \sum_{k=0}^{\infty} \binom{\ell}{k} y^k,$$

where the binomial coefficients are defined by

$$\binom{\ell}{0} = 1, \qquad \binom{\ell}{k} = \frac{\ell(\ell-1)\cdots(\ell-k+1)}{k!} \quad (k > 0).$$

Proof. Let $f(y) = (1+y)^\ell$. The nth derivative of f is

$$f^{(n)}(y) = \ell(\ell-1)\cdots(\ell-n+1)(1+y)^{\ell-n}. \tag{4.22}$$

If $|y| < 1$, the power series

$$g(y) = \sum_{k=0}^{\infty} \binom{\ell}{k} y^k = \sum_{k=0}^{\infty} \frac{\ell(\ell-1)\cdots(\ell-k+1)}{k!} y^k$$

converges by the ratio test, since

$$\lim_{k\to\infty} \left| \frac{\binom{\ell}{k+1} y^{k+1}}{\binom{\ell}{k} y^k} \right| = \lim_{k\to\infty} \frac{\frac{\ell(\ell-1)\cdots(\ell-k)}{(k+1)!} y^{k+1}}{\frac{\ell(\ell-1)\cdots(\ell-k+1)}{k!} y^k} = \lim_{k\to\infty} \left| \frac{\ell-k}{k+1} \right| |y| = |y|.$$

We want to show that $g(y) = (1+y)^\ell$ for $|y| < 1$. According to Theorem 3.17, we can differentiate $g(y)$ term by term to get

$$g'(y) = \sum_{k=0}^{\infty} \frac{\ell(\ell-1)\cdots(\ell-k+1)}{k!} Ly^{k-1} = \sum_{k=0}^{\infty} \binom{\ell}{k-1} y^{k-1}.$$

Multiply $g(y)$ by y and add $g'(y)$ to get

$$(1+y)g'(y) = \sum_{k=0}^{\infty} \left((k+1)\binom{\ell}{k+1} + k\binom{\ell}{k} \right) y^k = \sum_{k=0}^{\infty} \ell \binom{\ell}{k} y^k = \ell g(y).$$

Now let us examine

$$\frac{\mathrm{d}}{\mathrm{d}y} \frac{g(y)}{(1+y)^\ell} = \frac{(1+y)^\ell g'(y) - g(y)\ell(1+y)^{\ell-1}}{\left((1+y)^\ell\right)^2}$$

$$= \frac{\ell(1+y)^{\ell-1}}{(1+y)^{2\ell}} \big((1+y)g'(y) - \ell g(y)\big) = 0.$$

Therefore, $\dfrac{g(y)}{(1+y)^\ell}$ is constant. But at $y = 0$, we know that $\dfrac{g(0)}{(1)^\ell} = 1$. Therefore, the power series $g(y)$ equals $(1+y)^\ell$ for $|y| < 1$. □

This generalization of the binomial theorem to noninteger exponents was derived by Newton. This shows that Newton was familiar with Taylor's theorem, even though Taylor's book appeared 50 years after Newton's.

Problems

4.42. Find the Taylor polynomials $t_2(x)$ and $t_3(x)$ for $f(x) = 1 + x + x^2 + x^3 + x^4$ in powers of x.

4.43. Find the Taylor series for $\cos x$ in powers of x. For what values of x does the series converge to $\cos x$?

4.44. Find the Taylor series for $\cos(3x)$ in powers of x. For what values of x does the series converge to $\cos(3x)$?

4.45. Compare Taylor polynomials t_3 and t_4 for $\sin x$ in powers of x. Give the best estimate you can of $|\sin x - t_3(x)|$.

4.46. Find the Taylor polynomial of degree 4 and estimate the remainder in

$$\tan^{-1} x = x - \frac{x^3}{3} + (\text{remainder}) \quad x \text{ in } [-\frac{1}{2}, \frac{1}{2}].$$

4.47. Find the Taylor series for $\cosh x$ in powers of x. Use the Taylor remainder formula to show that the series converges uniformly to $\cosh x$ on every interval $[-b, b]$.

4.48. Find the Taylor series for $\sinh(2x)$ in powers of x. Use the Taylor remainder formula to show that the series converges uniformly to $\sinh(2x)$ on every interval $[-b,b]$.

4.49. Find the Taylor series for $\cos x$ about $a=\pi/3$, i.e. in powers of $(x-\pi/3)$.

4.50. Prove the validity of Taylor's formula with remainder, Eq. (4.17), when $b<a$. *Hint*: Consider the function $g(x)=f(a+b-x)$ over the interval $[b,a]$.

4.51. Find the binomial coefficients $b_k=\binom{\frac{1}{2}}{k}$ through b_3 in

$$\sqrt{1+y}=b_0+b_1y+b_2y^2+\cdots.$$

4.52. Consider the function $f(x)=\sqrt{x}$ on the interval $1\le x\le 1+d$. Find values of d small enough that $t_2(x)$, the second-degree Taylor polynomial at $x=1$, approximates $f(x)$ on $[1,1+d]$ with an error of at most

(a) .1
(b) .01
(c) .001

4.53. Answer the question posed in Problem 4.52 for the third-degree Taylor polynomial t_3 in place of t_2.

4.54. Let s be a function with the following properties:

(a) s has derivatives of all orders.
(b) All derivatives of s lie between -1 and 1.
(c) $s^{(j)}(0)=\begin{cases}0, & j \text{ even.}\\ (-1)^{(j-1)/2}, & j \text{ odd.}\end{cases}$

Determine a value of n so large that the nth-degree Taylor polynomial $t_n(x)$ approximates $s(x)$ with an error less than 10^{-3} on the interval $[-1,1]$. Determine the value of $s(0.7854)$ with an error less than 10^{-3}. What is the standard name for this function?

4.55. Let c be a function that has properties (a) and (b) of Problem 4.54 and satisfies

$$c^{(j)}(0)=\begin{cases}(-1)^{j/2}, & j \text{ even,}\\ 0, & j \text{ odd.}\end{cases}$$

Using a Taylor polynomial of appropriate degree, determine the value of $c(0.7854)$ with an error less than 10^{-3}. What is the standard name for this function?

4.56. Explain why there is no power series $|t|=\sum_{n=0}^{\infty}a_nt^n$.

4.57. In this problem, we rediscover Taylor's theorem for the case of polynomials. Let f be a polynomial of degree n, and a any constant, and set

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

(a) Evaluate $g(0)$ and $g(-a)$.
(b) Evaluate $g'(x)$ and simplify your result as much as possible.
(c) Conclude from part (b) the somewhat curious result that g is a constant function.
(d) Use the result of part (c) to express the relation between the values you found in part (a).

4.4 Approximating Derivatives

By definition, $f'(x)$ requires that we use a limiting process. That may be fine for human beings, but not for computers. The difference quotient

$$f_h(x) = \frac{f(x+h) - f(x)}{h}$$

can be computed once you know f, x, and h. But how good would such an approximation be? Let us look at some examples.

Example 4.18. If $f(x) = x^2$, then

$$f_h(x) = \frac{f(x+h) - f(x)}{h} = \frac{(x+h)^2 - x^2}{h} = \frac{x^2 + 2xh + h^2 - x^2}{h} = 2x + h.$$

For this function, replacing $f'(x)$ by $f_h(x)$ would lead to an error of only h. If we are willing to accept an error of say 0.00001, then we could approximate $f'(x)$ by $f_{0.00001}(x)$. Figure 4.20 shows the case in which $h = 0.07$.

Example 4.19. If $f(x) = x^3$, then

$$f_h(x) = \frac{f(x+h) - f(x)}{h} = \frac{(x+h)^3 - x^3}{h}$$

$$= \frac{x^3 + 3x^2h + 3xh^2 + h^3 - x^3}{h} = 3x^2 + 3xh + h^2.$$

For this function, replacing $f'(x)$ by $f_h(x)$ would introduce an error of $3xh + h^2$, an amount that depends on both h and x. This is illustrated in Fig. 4.21.

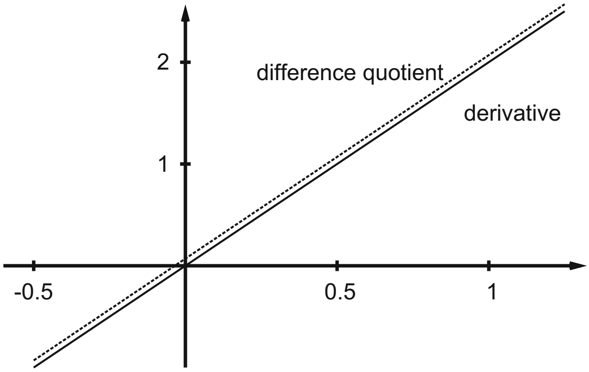

Fig. 4.20 Derivative $f'(x)$ and difference quotient $f_h(x) = \dfrac{f(x+h)-f(x)}{h}$ for $f(x) = x^2$ using $h = 0.07$

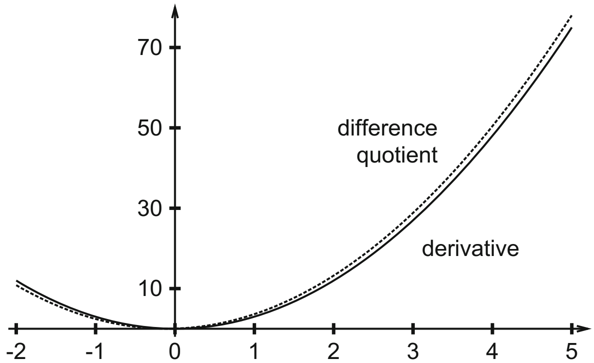

Fig. 4.21 Derivative $f'(x)$ and difference quotient $f_h(x) = \frac{f(x+h)-f(x)}{h}$ for $f(x) = x^3$ using $h = 0.2$

These examples lead to the concept of uniform differentiability.

Definition 4.8. A function f defined on an interval is called *uniformly differentiable* if given a tolerance $\varepsilon > 0$, there is a δ such that

$$\text{if } |h| < \delta, \text{ then } \left| \frac{f(x+h)-f(x)}{h} - f'(x) \right| < \varepsilon \text{ for every } x.$$

Example 4.20. The linear function $f(x) = mx + b$ is uniformly differentiable: the derivative is the constant function m. The difference quotient is given by

$$f_h(x) = \frac{m(x+h) + b - (mx+b)}{h} = m.$$

So it not only tends to $f'(x)$, but is equal to $f'(x)$ for all h and all x.

Example 4.21. We have shown in Sect. 3.3a that the exponential function e^x is differentiable at every x. We will now show that e^x is uniformly differentiable on each interval $[-c,c]$. We have

$$\frac{e^{x+h}-e^x}{h}-e^x = e^x\left(\frac{e^h-1}{h}-1\right).$$

Therefore, for every x in $[-c,c]$, this quantity is at most e^c times $\left(\frac{e^h-1}{h}-1\right)$, which does not depend on x, and which tends to zero as h tends to zero.

We claim the following theorem.

Theorem 4.13. *If f is uniformly differentiable on $[a,b]$, then f' is uniformly continuous on $[a,b]$.*

The proof is outlined in Problem 4.62 at the end of this section. The theorem has a converse, whose significance is that we may easily approximate derivatives:

Theorem 4.14. *If f' is uniformly continuous on $[a,b]$, then f is uniformly differentiable on $[a,b]$.*

Proof. We need to prove that $f'(x)$ and the difference quotient $f_h(x)$ differ by an amount that is small independently of x. More precisely, consider

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

According to the mean value theorem (Theorem 4.1), the difference quotient equals $f'(c)$ at some point c between x and $x+h$. So x and c differ by less than h, and we can rewrite the previous expression as

$$\frac{f(x+h)-f(x)}{h}-f'(x) = f'(c)-f'(x).$$

Since f' is uniformly continuous on $[a,b]$, given any tolerance ε, there is a precision δ such that if x and c are in $[a,b]$ and differ by less than δ, then $f'(c)$ and $f'(x)$ differ by less than ε. This proves that f is uniformly differentiable. □

Many of the functions we work with, such as polynomials, sine, cosine, exponential, and logarithm, have continuous derivatives, and are therefore uniformly differentiable on closed intervals. We give an example in Problem 4.63 of a differentiable function f for which f' is not continuous, and thus according to Theorem 4.13, f is not uniformly differentiable.

A Word of Caution about actually approximating $f'(x)$ by $f_h(x)$. When we asked how good an approximation $f_h(x) = \dfrac{f(x+h) - f(x)}{h}$ is to $f'(x)$, we assumed that we could perform the subtraction accurately. But when we subtract decimal approximations of numbers that are very close, we get very few digits of the difference correctly. Figure 4.22 shows the result of a computer program that attempted to calculate the difference quotient minus the derivative,

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

for $f(x) = x^2$ and $x = 1$. We know by algebra that

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

so the graph should be a straight line. But we see something quite different in Fig. 4.22.

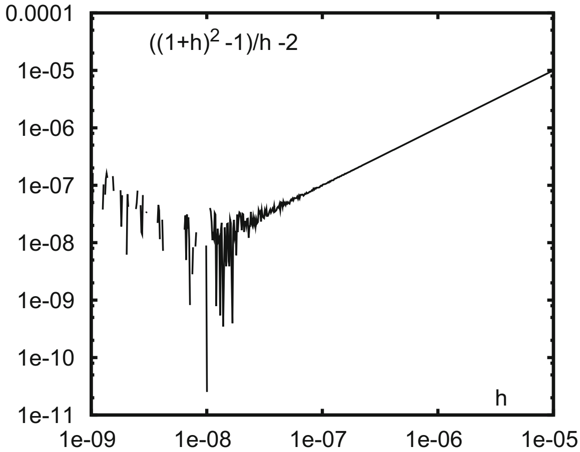

Fig. 4.22 The result of computing the difference quotient minus the derivative $\frac{(1+h)^2 - 1^2}{h} - 2$ is plotted for $10^{-9} \le h \le 10^{-5}$

Approximate Derivatives and Data. In experimental settings, functions are represented through tables of data, rather than by a formula. How can we compute f' when only tabular data are known? We give an application of the linear approximation theorem, Theorem 4.5, to the problem of approximating derivatives.

The difference quotient $\dfrac{f(x+h) - f(x)}{h}$ is asymmetric in the sense that it favors one side of the point x. By the linear approximation theorem,

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(c_1)h^2 \quad \text{and} \quad f(x-h) = f(x) - f'(x)h + \frac{1}{2}f''(c_2)h^2,$$

where c_1 lies between x and $x+h$, and c_2 lies between x and $x-h$. Subtract and divide by $2h$ to get an estimate for the *symmetric* difference quotient

$$\frac{f(x+h) - f(x-h)}{2h} = f'(x) + \frac{1}{4}\big(f''(c_1) - f''(c_2)\big)h. \tag{4.23}$$

See Fig. 4.23, which illustrates a symmetric difference quotient. Both c_1 and c_2 differ by less than h from x. If f'' is continuous, then for small h, both $f''(c_1)$ and $f''(c_2)$ differ little from $f''(x)$. Thus we deduce that the symmetric difference quotient differs from $f'(x)$ by an amount sh, where $s = \frac{1}{4}\big(f''(c_1) - f''(c_2)\big)$ is small when h is small.

But we saw that the one-sided difference quotient differs from $f'(x)$ by $\frac{1}{2}f''(c_1)h$. We conclude that for twice differentiable functions and for small h, *the symmetric difference quotient is a better approximation of the derivative at x than the one-sided quotient.*

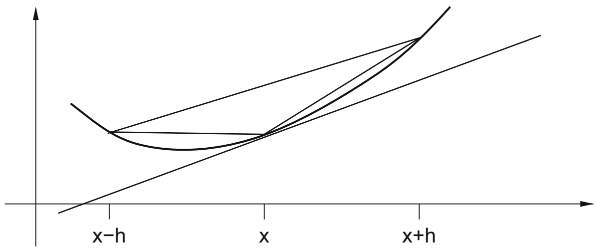

Fig. 4.23 The symmetric difference quotient is a better approximation to $f'(x)$ than one-sided quotients

Let us look at an example. Table 4.1 contains data for a function at eleven equidistant points between 0 and 1. Note that if we take $x+h = x_2$ and $x-h = x_1$ in Eq. (4.23), it becomes

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'\left(\frac{x_2 + x_1}{2}\right) + \frac{1}{4}\big(f''(c_1) - f''(c_2)\big)\frac{x_2 - x_1}{2}.$$

Table 4.1 shows estimates for f' at the midpoints $0.05, 0.15, 0.25, \ldots, 0.85, 0.95$ of the intervals using

$$f'\big(\frac{x_1 + x_2}{2}\big) \approx \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

For example,

$$f'(0.35) = f'\big(\frac{0.3 + 0.4}{2}\big) \approx \frac{f(0.4) - f(0.3)}{0.4 - 0.3} = \frac{0.38941 - 0.29552}{0.1} = 0.9389$$

Table 4.1 Data in the $f(x)$ column, and approximate derivatives for the unknown function

x	$f(x)$	$\approx f'(x)$	$\approx f''(x)$	$\approx f'''(x)$	$\approx f''''(x)$
0	0				
0.05		0.9983			
0.1	0.09983		−0.100		
0.15		0.9883		−0.97	
0.2	0.19866		−0.197		( )
0.25		0.9686		−1	
0.3	0.29552		−0.297		( )
0.35		0.9389		−0.91	
0.4	0.38941		−0.388		( )
0.45		0.9001		−0.91	
0.5	0.47942		−0.479		( )
0.55		0.8522		( )	
0.6	0.56464		−0.565		( )
0.65		0.7957		( )	
0.7	0.64421		−0.643		( )
0.75		0.7314		( )	
0.8	0.71735		−0.717		( )
0.85		0.6597		( )	
0.9	0.78332		−0.782		
0.95		0.5815			
1	0.84147				

and

$$f'(0.45) = f'\Big(\frac{0.4+0.5}{2}\Big) \approx \frac{f(0.5)-f(0.4)}{0.5-0.4} = \frac{0.47942-0.38941}{0.1} = 0.9001.$$

Now using our estimates for $f'(0.35)$ and $f'(0.45)$ we can repeat the process to find estimates for f'' at $0.1, 0.2, \ldots, 0.9$. For example,

$$f''(0.4) \approx \frac{f'(0.45)-f'(0.35)}{0.45-0.35} \approx \frac{0.9001-0.9389}{0.1} = -0.388.$$

In Problem 4.60, we ask you to complete the table.

Problems

4.58. Consider the symmetric difference quotient graphed in Fig. 4.24. Use Taylor's theorem with remainder to show that the difference

$$\left|\frac{\sin(x+0.1)-\sin(x-0.1)}{0.2} - \cos x\right|$$

is less than 0.002 for all x. This is why the cosine appears to have been graphed in the figure.

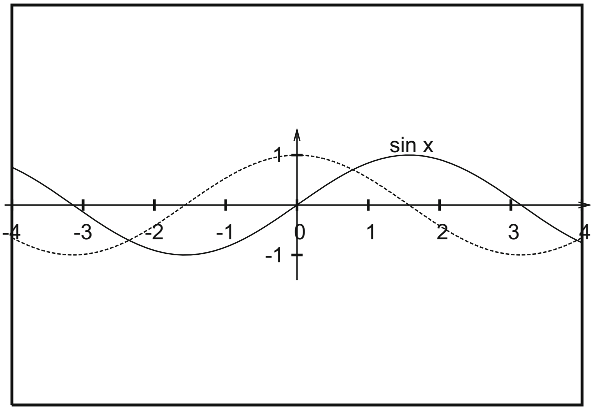

Fig. 4.24 The sine and a symmetric difference quotient $\dfrac{\sin(x+0.1)-\sin(x-0.1)}{0.2}$

4.59. Evaluate the one-sided difference quotient $\dfrac{f(x+h)-f(x)}{h}$ and the symmetric difference quotient $\dfrac{f(x+h)-f(x-h)}{2h}$ for the cases $f(x)=x^2$ and $f(x)=x^3$. If $x=10$ and $h=0.1$, by how much do these quotients differ from the derivative?

4.60.

(a) Use the approximation $f'''\left(\dfrac{x_1+x_2}{2}\right)\approx\dfrac{f''(x_2)-f''(x_1)}{x_2-x_1}$ to find estimates for f''' at the points 0.55, 0.65, 0.75, 0.85, which were left blank in Table 4.1.

(b) Use the approximation $f''''\left(\dfrac{x_1+x_2}{2}\right)\approx\dfrac{f'''(x_2)-f'''(x_1)}{x_2-x_1}$ to find estimates for f'''' at $x=0.2,0.3,\ldots,0.7,0.8$.

4.61. Here we explore how to use approximate derivatives in a somewhat different way: to detect an isolated error in the tabulation of a smooth function. Suppose that in tabulating the data column that shows values of $f(x)$ in Table 4.1, a small error was made that interchanged two digits, $f(0.4)=0.38914$ instead of $f(0.4)=0.38941$.

(a) Recompute the table.
(b) Plot graphs of f, f', f'', f''' and f'''' for Table 4.1, and again for the recomputed table. What do you notice?

4.62. Assume that f is uniformly differentiable on $[a,b]$. Show that f' is continuous on $[a,b]$ by carrying out or justifying each of the following steps.

(a) Write down the definition of uniformly differentiable on $[a,b]$.
(b) Given any tolerance ε, explain why there is a positive integer n such that if $|h|<\frac{1}{n}$, then

$$\left|\frac{f\left(x+\frac{1}{n}\right)-f(x)}{\frac{1}{n}}-f'(x)\right|<\varepsilon$$

for every x in $[a,b]$.

(c) Define a sequence of continuous functions $f_n(x) = \dfrac{f\left(x+\frac{1}{n}\right) - f(x)}{\frac{1}{n}}$. Explain why f_n is continuous on $[a,b]$ and show that f_n converges uniformly to f' on $[a,b]$. Conclude that f' is continuous on $[a,b]$.

4.63. Define $f(x) = x^2 \sin\left(\dfrac{1}{x}\right)$ and $f(0) = 0$.

(a) Find $f'(x)$ for values of $x \neq 0$.
(b) Verify that f is also differentiable at $x = 0$, and evaluate $f'(0)$, by considering the difference quotient $\dfrac{f(h) - f(0)}{h}$.
(c) Verify that f' is not continuous at 0 by showing that $f'(x)$ does not tend to $f'(0)$ as x tends to zero.
(d) Use Theorem 4.13 to argue that f is not uniformly differentiable.