CHAPTER 4

More Applications of Differentiation

> "In the fall of 1972 President Nixon announced that the rate of increase of inflation was decreasing. This was the first time a sitting president used the third derivative to advance his case for reelection."

Hugo Rossi
Mathematics Is an Edifice, Not a Toolbox, Notices of the AMS, v. 43, Oct. 1996

Introduction

Differential calculus can be used to analyze many kinds of problems and situations that arise in applied disciplines. Calculus has made and will continue to make significant contributions to every field of human endeavour that uses quantitative measurement to further its aims. From economics to physics and from biology to sociology, problems can be found whose solutions can be aided by the use of some calculus.

In this chapter we will examine several kinds of problems to which the techniques we have already learned can be applied. These problems arise both outside and within mathematics. We will deal with the following kinds of problems:

1. Related rates problems, where the rates of change of related quantities are analyzed.
2. Root finding methods, where we try to find numerical solutions of equations.
3. Evaluation of limits.
4. Optimization problems, where a quantity is to be maximized or minimized.
5. Graphing problems, where derivatives are used to illuminate the behaviour of functions.
6. Approximation problems, where complicated functions are approximated by polynomials.

Do not assume that most of the problems we present here are "real-world" problems. Such problems are usually too complex to be treated in a general calculus course. However, the problems we consider, while sometimes artificial, do show how calculus can be applied in concrete situations.

4.1 Related Rates

When two or more quantities that change with time are linked by an equation, that equation can be differentiated with respect to time to produce an equation linking the rates of change of the quantities. Any one of these rates may then be determined when the others, and the values of the quantities themselves, are known. We will consider a couple of examples before formulating a list of procedures for dealing with such problems.

EXAMPLE 1 An aircraft is flying horizontally at a speed of 600 km/h. How fast is the distance between the aircraft and a radio beacon increasing 1 min after the aircraft passes 5 km directly above the beacon?

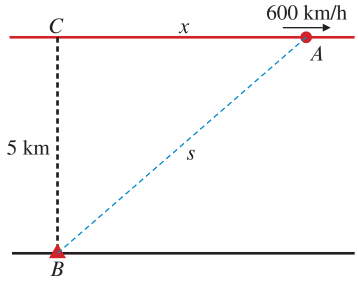


Figure 4.1

Solution A diagram is useful here; see Figure 4.1. Let C be the point on the aircraft's path directly above the beacon B. Let A be the position of the aircraft t min after it is at C, and let x and s be the distances CA and BA, respectively. From the right triangle BCA we have

$$s^2 = x^2 + 5^2.$$

We differentiate this equation implicitly with respect to t to obtain

$$2s\frac{ds}{dt} = 2x\frac{dx}{dt}.$$

We are given that $dx/dt = 600$ km/h $= 10$ km/min. Therefore, $x = 10$ km at time $t = 1$ min. At that time $s = \sqrt{10^2 + 5^2} = 5\sqrt{5}$ km and is increasing at the rate

$$\frac{ds}{dt} = \frac{x}{s}\frac{dx}{dt} = \frac{10}{5\sqrt{5}}(600) = \frac{1,200}{\sqrt{5}} \approx 536.7 \text{ km/h}.$$

One minute after the aircraft passes over the beacon, its distance from the beacon is increasing at about 537 km/h.

EXAMPLE 2 How fast is the area of a rectangle changing if one side is 10 cm long and is increasing at a rate of 2 cm/s and the other side is 8 cm long and is decreasing at a rate of 3 cm/s?

Solution Let the lengths of the sides of the rectangle at time t be x cm and y cm, respectively. Thus, the area at time t is $A = xy$ cm^2. (See Figure 4.2.) We want to know the value of dA/dt when $x = 10$ and $y = 8$, given that $dx/dt = 2$ and $dy/dt = -3$. (Note the negative sign to indicate that y is decreasing.) Since all the quantities in the equation $A = xy$ are functions of time, we can differentiate that equation implicitly with respect to time and obtain

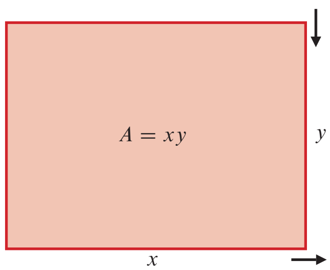


Figure 4.2 Rectangle with sides changing

$$\left.\frac{dA}{dt}\right|_{\substack{x=10\\y=8}} = \left.\left(\frac{dx}{dt}\,y + x\frac{dy}{dt}\right)\right|_{\substack{x=10\\y=8}} = 2(8) + 10(-3) = -14.$$

At the time in question, the area of the rectangle is decreasing at a rate of 14 cm^2/s.

Procedures for Related-Rates Problems

In view of these examples we can formulate a few general procedures for dealing with related-rates problems.

How to solve related-rates problems

1. Read the problem very carefully. Try to understand the relationships between the variable quantities. What is given? What is to be found?
2. Make a sketch if appropriate.
3. Define any symbols you want to use that are not defined in the statement of the problem. Express given and required quantities and rates in terms of these symbols.
4. From a careful reading of the problem or consideration of the sketch, identify one or more equations linking the variable quantities. (You will need as many equations as quantities or rates to be found in the problem.)
5. Differentiate the equation(s) implicitly with respect to time, regarding all variable quantities as functions of time. You can manipulate the equation(s) algebraically before the differentiation is performed (for instance, you could solve for the quantities whose rates are to be found), but it is usually easier to differentiate the equations as they are originally obtained and solve for the desired items later.
6. Substitute any given values for the quantities and their rates, then solve the resulting equation(s) for the unknown quantities and rates.
7. Make a concluding statement answering the question asked. Is your answer reasonable? If not, check back through your solution to see what went wrong.

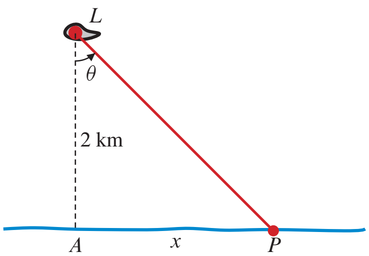


Figure 4.3

EXAMPLE 3 A lighthouse L is located on a small island 2 km from the nearest point A on a long, straight shoreline. If the lighthouse lamp rotates at 3 revolutions per minute, how fast is the illuminated spot P on the shoreline moving along the shoreline when it is 4 km from A?

Solution Referring to Figure 4.3, let x be the distance AP, and let θ be the angle PLA. Then $x = 2\tan\theta$ and

$$\frac{dx}{dt} = 2\sec^2\theta\,\frac{d\theta}{dt}.$$

Now

$$\frac{d\theta}{dt} = (3 \text{ rev/min})(2\pi \text{ radians/rev}) = 6\pi \text{ radians/min}.$$

When $x = 4$, we have $\tan\theta = 2$ and $\sec^2\theta = 1 + \tan^2\theta = 5$. Thus,

$$\frac{dx}{dt} = (2)(5)(6\pi) = 60\pi \approx 188.5.$$

The spot of light is moving along the shoreline at a rate of about 189 km/min when it is 4 km from A.

(Note that it was essential to convert the rate of change of θ from revolutions per minute to radians per minute. If θ were not measured in radians we could not assert that $(d/d\theta)\tan\theta = \sec^2\theta$.)

EXAMPLE 4 A leaky water tank is in the shape of an inverted right circular cone with depth 5 m and top radius 2 m. When the water in the tank is 4 m deep, it is leaking out at a rate of $1/12$ m^3/min. How fast is the water level in the tank dropping at that time?

Solution Let r and h denote the surface radius and depth of water in the tank at time t (both measured in metres). Thus, the volume V (in cubic metres) of water in the tank at time t is

$$V = \frac{1}{3}\pi r^2 h.$$

Using similar triangles (see Figure 4.4), we can find a relationship between r and h:

$$\frac{r}{h} = \frac{2}{5}, \qquad \text{so} \quad r = \frac{2h}{5} \quad \text{and} \quad V = \frac{1}{3}\pi\left(\frac{2h}{5}\right)^2 h = \frac{4\pi}{75}h^3.$$

Differentiating this equation with respect to t, we obtain

$$\frac{dV}{dt} = \frac{4\pi}{25}h^2\frac{dh}{dt}.$$

Since $dV/dt = -1/12$ when $h = 4$, we have

$$\frac{-1}{12} = \frac{4\pi}{25}(4^2)\frac{dh}{dt}, \qquad \text{so} \quad \frac{dh}{dt} = -\frac{25}{768\pi}.$$

When the water in the tank is 4 m deep, its level is dropping at a rate of $25/(768\pi)$ m/min, or about 1.036 cm/min.

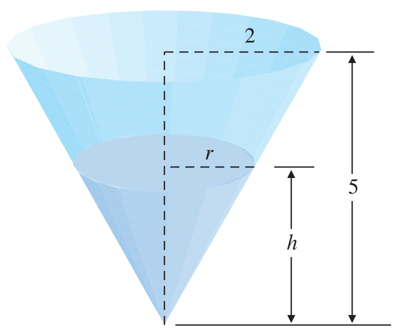


Figure 4.4 The conical tank of Example 4

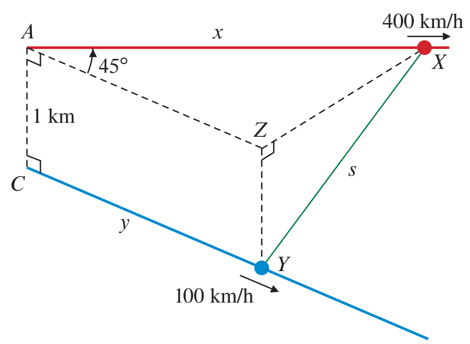


Figure 4.5 Aircraft and car paths in Example 5

EXAMPLE 5 At a certain instant an aircraft flying due east at 400 km/h passes directly over a car travelling due southeast at 100 km/h on a straight, level road. If the aircraft is flying at an altitude of 1 km, how fast is the distance between the aircraft and the car increasing 36 s after the aircraft passes directly over the car?

Solution A good diagram is essential here. See Figure 4.5. Let time t be measured in hours from the time the aircraft was at position A directly above the car at position C. Let X and Y be the positions of the aircraft and the car, respectively, at time t. Let x be the distance AX, y the distance CY, and s the distance XY, all measured in kilometres. Let Z be the point 1 km above Y. Since angle $XAZ = 45°$, the Pythagorean Theorem and Cosine Law yield

$$\begin{aligned} s^2 &= 1 + (ZX)^2 = 1 + x^2 + y^2 - 2xy\cos 45° \\ &= 1 + x^2 + y^2 - \sqrt{2}\,xy. \end{aligned}$$

Thus,

$$\begin{aligned}
2s\frac{ds}{dt} &= 2x\frac{dx}{dt} + 2y\frac{dy}{dt} - \sqrt{2}\frac{dx}{dt}y - \sqrt{2}x\frac{dy}{dt} \\
&= 400(2x - \sqrt{2}y) + 100(2y - \sqrt{2}x),
\end{aligned}$$

since $dx/dt = 400$ and $dy/dt = 100$. When $t = 1/100$ (i.e., 36 s after $t = 0$), we have $x = 4$ and $y = 1$. Hence,

$$\begin{aligned}
s^2 &= 1 + 16 + 1 - 4\sqrt{2} = 18 - 4\sqrt{2} \\
s &\approx 3.5133. \\
\frac{ds}{dt} &= \frac{1}{2s}\big(400(8 - \sqrt{2}) + 100(2 - 4\sqrt{2})\big) \approx 322.86.
\end{aligned}$$

The aircraft and the car are separating at a rate of about 323 km/h after 36 s. (Note that it was necessary to convert 36 s to hours in the solution. In general, all measurements should be in compatible units.)

EXERCISES 4.1

1. Find the rate of change of the area of a square whose side is 8 cm long, if the side length is increasing at 2 cm/min.
2. The area of a square is decreasing at 2 ft^2/s. How fast is the side length changing when it is 8 ft?
3. A pebble dropped into a pond causes a circular ripple to expand outward from the point of impact. How fast is the area enclosed by the ripple increasing when the radius is 20 cm and is increasing at a rate of 4 cm/s?
4. The area of a circle is decreasing at a rate of 2 cm^2/min. How fast is the radius of the circle changing when the area is 100 cm^2?
5. The area of a circle is increasing at $1/3$ km^2/h. Express the rate of change of the radius of the circle as a function of (a) the radius r and (b) the area A of the circle.
6. At a certain instant the length of a rectangle is 16 m and the width is 12 m. The width is increasing at 3 m/s. How fast is the length changing if the area of the rectangle is not changing?
7. Air is being pumped into a spherical balloon. The volume of the balloon is increasing at a rate of 20 cm^3/s when the radius is 30 cm. How fast is the radius increasing at that time? (The volume of a ball of radius r units is $V = \frac{4}{3}\pi r^3$ cubic units.)
8. When the diameter of a ball of ice is 6 cm, it is decreasing at a rate of 0.5 cm/h due to melting of the ice. How fast is the volume of the ice ball decreasing at that time?
9. How fast is the surface area of a cube changing when the volume of the cube is 64 cm^3 and is increasing at 2 cm^3/s?
10. The volume of a right circular cylinder is 60 cm^3 and is increasing at 2 cm^3/min at a time when the radius is 5 cm and is increasing at 1 cm/min. How fast is the height of the cylinder changing at that time?
11. How fast is the volume of a rectangular box changing when the length is 6 cm, the width is 5 cm, and the depth is 4 cm, if the length and depth are both increasing at a rate of 1 cm/s and the width is decreasing at a rate of 2 cm/s?
12. The area of a rectangle is increasing at a rate of 5 m^2/s while the length is increasing at a rate of 10 m/s. If the length is 20 m and the width is 16 m, how fast is the width changing?
13. A point moves on the curve $y = x^2$. How fast is y changing when $x = -2$ and x is decreasing at a rate of 3?
14. A point is moving to the right along the first-quadrant portion of the curve $x^2y^3 = 72$. When the point has coordinates $(3, 2)$, its horizontal velocity is 2 units/s. What is its vertical velocity?
15. The point P moves so that at time t it is at the intersection of the curves $xy = t$ and $y = tx^2$. How fast is the distance of P from the origin changing at time $t = 2$?
16. **(Radar guns)** A police officer is standing near a highway using a radar gun to catch speeders. (See Figure 4.6.) He aims the gun at a car that has just passed his position and, when the gun is pointing at an angle of 45° to the direction of the highway, notes that the distance between the car and the gun is increasing at a rate of 100 km/h. How fast is the car travelling?

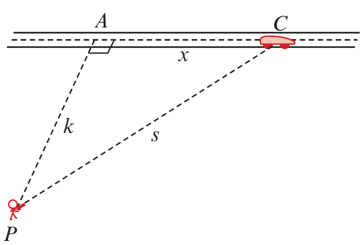


Figure 4.6

17. If the radar gun of Exercise 16 is aimed at a car travelling at 90 km/h along a straight road, what will its reading be when it is aimed making an angle of 30° with the road?

18. The top of a ladder 5 m long rests against a vertical wall. If the base of the ladder is being pulled away from the base of the wall at a rate of 1/3 m/s, how fast is the top of the ladder slipping down the wall when it is 3 m above the base of the wall?

19. A man 2 m tall walks toward a lamppost on level ground at a rate of 0.5 m/s. If the lamp is 5 m high on the post, how fast is the length of the man's shadow decreasing when he is 3 m from the post? How fast is the shadow of his head moving at that time?

20. A woman 6 ft tall is walking at 2 ft/s along a straight path on level ground. There is a lamppost 5 ft to the side of the path. A light 15 ft high on the lamppost casts the woman's shadow on the ground. How fast is the length of her shadow changing when the woman is 12 feet from the point on the path closest to the lamppost?

21. **(Cost of production)** It costs a coal mine owner $\$C$ each day to maintain a production of x tonnes of coal, where $C = 10{,}000 + 3x + x^2/8{,}000$. At what rate is the production increasing when it is 12,000 tonnes and the daily cost is increasing at \$600 per day?

22. **(Distance between ships)** At 1:00 p.m. ship A is 25 km due north of ship B. If ship A is sailing west at a rate of 16 km/h and ship B is sailing south at 20 km/h, at what rate is the distance between the two ships changing at 1:30 p.m?

23. What is the first time after 3:00 p.m. that the hands of a clock are together?

24. **(Tracking a balloon)** A balloon released at point A rises vertically with a constant speed of 5 m/s. Point B is level with and 100 m distant from point A. How fast is the angle of elevation of the balloon at B changing when the balloon is 200 m above A?

25. Sawdust is falling onto a pile at a rate of 1/2 m^3/min. If the pile maintains the shape of a right circular cone with height equal to half the diameter of its base, how fast is the height of the pile increasing when the pile is 3 m high?

26. **(Conical tank)** A water tank is in the shape of an inverted right circular cone with top radius 10 m and depth 8 m. Water is flowing in at a rate of 1/10 m^3/min. How fast is the depth of water in the tank increasing when the water is 4 m deep?

27. **(Leaky tank)** Repeat Exercise 26 with the added assumption that water is leaking out of the bottom of the tank at a rate of $h^3/1{,}000$ m^3/min when the depth of water in the tank is h m. How full can the tank get in this case?

28. **(Another leaky tank)** Water is pouring into a leaky tank at a rate of 10 m^3/h. The tank is a cone with vertex down, 9 m in depth and 6 m in diameter at the top. The surface of water in the tank is rising at a rate of 20 cm/h when the depth is 6 m. How fast is the water leaking out at that time?

29. **(Kite flying)** How fast must you let out line if the kite you are flying is 30 m high, 40 m horizontally away from you, and moving horizontally away from you at a rate of 10 m/min?

30. **(Ferris wheel)** You are on a Ferris wheel of diameter 20 m. It is rotating at 1 revolution per minute. How fast are you rising or falling when you are 6 m horizontally away from the vertical line passing through the centre of the wheel?

31. **(Distance between aircraft)** An aircraft is 144 km east of an airport and is travelling west at 200 km/h. At the same time, a second aircraft at the same altitude is 60 km north of the airport and travelling north at 150 km/h. How fast is the distance between the two aircraft changing?

32. **(Production rate)** If a truck factory employs x workers and has daily operating expenses of $\$y$, it can produce $P = (1/3)x^{0.6}y^{0.4}$ trucks per year. How fast are the daily expenses decreasing when they are \$10,000 and the number of workers is 40, if the number of workers is increasing at 1 per day and production is remaining constant?

33. A lamp is located at point $(3, 0)$ in the xy-plane. An ant is crawling in the first quadrant of the plane and the lamp casts its shadow onto the y-axis. How fast is the ant's shadow moving along the y-axis when the ant is at position $(1, 2)$ and moving so that its x-coordinate is increasing at rate 1/3 units/s and its y-coordinate is decreasing at 1/4 units/s?

34. A straight highway and a straight canal intersect at right angles, the highway crossing over the canal on a bridge 20 m above the water. A boat travelling at 20 km/h passes under the bridge just as a car travelling at 80 km/h passes over it. How fast are the boat and car separating after one minute?

35. **(Filling a trough)** The cross section of a water trough is an equilateral triangle with top edge horizontal. If the trough is 10 m long and 30 cm deep, and if water is flowing in at a rate of 1/4 m^3/min, how fast is the water level rising when the water is 20 cm deep at the deepest?

36. **(Draining a pool)** A rectangular swimming pool is 8 m wide and 20 m long. (See Figure 4.7.) Its bottom is a sloping plane, the depth increasing from 1 m at the shallow end to 3 m at the deep end. Water is draining out of the pool at a rate of 1 m^3/min. How fast is the surface of the water falling when the depth of water at the deep end is (a) 2.5 m? (b) 1 m?

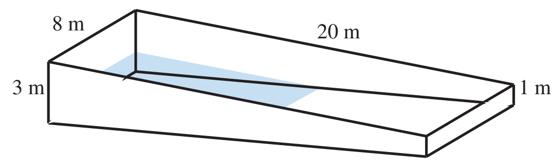


Figure 4.7

37. One end of a 10 m long ladder is on the ground. The ladder is supported partway along its length by resting on top of a 3 m high fence. (See Figure 4.8.) If the bottom of the ladder is 4 m from the base of the fence and is being dragged along the ground away from the fence at a rate of 1/5 m/s, how fast is the free top end of the ladder moving (a) vertically and (b) horizontally?

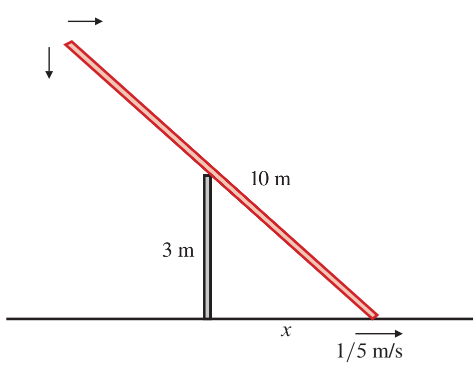


Figure 4.8

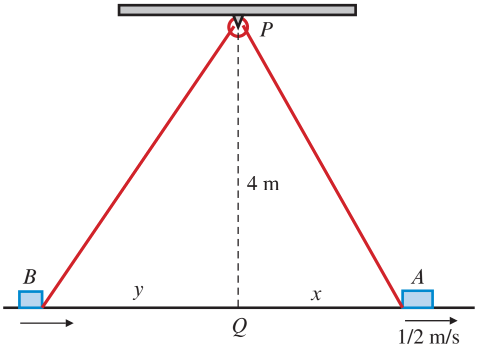


Figure 4.9

38. Two crates, A and B, are on the floor of a warehouse. The crates are joined by a rope 15 m long, each crate being hooked at floor level to an end of the rope. The rope is stretched tight and pulled over a pulley P that is attached to a rafter 4 m above a point Q on the floor directly between the two crates. (See Figure 4.9.) If crate A is 3 m from Q and is being pulled directly away from Q at a rate of 1/2 m/s, how fast is crate B moving toward Q?

39. **(Tracking a rocket)** Shortly after launch, a rocket is 100 km high and 50 km downrange. If it is travelling at 4 km/s at an angle of 30° above the horizontal, how fast is its angle of elevation, as measured at the launch site, changing?

40. **(Shadow of a falling ball)** A lamp is 20 m high on a pole. At time $t = 0$ a ball is dropped from a point level with the lamp and 10 m away from it. The ball falls under gravity (its acceleration is 9.8 m/s^2) until it hits the ground. How fast is the shadow of the ball moving along the ground (a) 1 s after the ball is dropped? (b) just as the ball hits the ground?

41. **(Tracking a rocket)** A rocket blasts off at time $t = 0$ and climbs vertically with acceleration 10 m/s^2. The progress of the rocket is monitored by a tracking station located 2 km horizontally away from the launch pad. How fast is the tracking station antenna rotating upward 10 s after launch?

4.2 Finding Roots of Equations

Finding solutions (roots) of equations is an important mathematical problem to which calculus can make significant contributions. There are only a few general classes of equations of the form $f(x) = 0$ that we can solve exactly. These include **linear equations**:

$$ax + b = 0, \quad (a \neq 0) \qquad \Rightarrow \qquad x = -\frac{b}{a}$$

and **quadratic equations**:

$$ax^2 + bx + c = 0, \quad (a \neq 0) \qquad \Rightarrow \qquad x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Cubic and quartic (3rd- and 4th-degree polynomial) equations can also be solved, but the formulas are very complicated. We usually solve these and most other equations approximately by using numerical methods, often with the aid of a calculator or computer.

In Section 1.4 we discussed the Bisection Method for approximating a root of an equation $f(x) = 0$. That method uses the Intermediate-Value Theorem and depends only on the continuity of f and our ability to find an interval $[x_1, x_2]$ that must contain the root because $f(x_1)$ and $f(x_2)$ have opposite signs. The method is rather slow; it requires between three and four iterations to gain one significant figure of precision in the root being approximated.

If we know that f is more than just continuous, we can devise better (i.e., faster) methods for finding roots of $f(x) = 0$. We study two such methods in this section:

(a) **Fixed-Point Iteration**, which looks for solutions of an equation of the form $x = f(x)$. Such solutions are called **fixed points** of the function f.

(b) **Newton's Method**, which looks for solutions of the equation $f(x) = 0$ as fixed points of the function $g(x) = x - \dfrac{f(x)}{f'(x)}$, that is, points x such that $x = g(x)$. This method is usually very efficient, but it requires that f be differentiable.

Like the Bisection Method, both of these methods require that we have at the outset a rough idea of where a root can be found, and they generate sequences of approximations that get closer and closer to the root.

Discrete Maps and Fixed-Point Iteration

A **discrete map** is an equation of the form

$$x_{n+1} = f(x_n), \qquad \text{for } n = 0,\ 1,\ 2,\ \ldots,$$

which generates a sequence of values $x_1, x_2, x_3, \ldots,$ from a given starting value x_0. In certain circumstances this sequence of numbers will converge to a limit, $r = \lim_{n\to\infty} x_n$, in which case this limit will be a fixed point of f: $r = f(r)$. (A thorough discussion of convergence of sequences can be found in Section 9.1. For our purposes here, an intuitive understanding will suffice: $\lim_{n\to\infty} x_n = r$ if $|x_n - r|$ approaches 0 as $n \to \infty$.)

For certain kinds of functions f, we can solve the equation $f(r) = r$ by starting with an initial guess x_0 and calculating subsequent values of the discrete map until sufficient accuracy is achieved. This is the **Method of Fixed-Point Iteration**. Let us begin by investigating a simple example:

EXAMPLE 1 Find a root of the equation $\cos x = 5x$.

Solution This equation is of the form $f(x) = x$, where $f(x) = \frac{1}{5}\cos x$. Since $\cos x$ is close to 1 for x near 0, we see that $\frac{1}{5}\cos x$ will be close to $\frac{1}{5}$ when $x = \frac{1}{5}$. This suggests that a reasonable first guess at the fixed point is $x_0 = \frac{1}{5} = 0.2$. The values of subsequent approximations

$$x_1 = \frac{1}{5}\cos x_0, \quad x_2 = \frac{1}{5}\cos x_1, \quad x_3 = \frac{1}{5}\cos x_2, \ldots$$

are presented in Table 1. The root is 0.196 164 28 to 8 decimal places.

Table 1.

n	x_n
0	0.2
1	0.196 013 32
2	0.196 170 16
3	0.196 164 05
4	0.196 164 29
5	0.196 164 28
6	0.196 164 28

Why did the method used in Example 1 work? Will it work for any function f? In order to answer these questions, examine the polygonal line in Figure 4.10. Starting at x_0 it goes vertically to the curve $y = f(x)$, the height there being x_1. Then it goes horizontally to the line $y = x$, meeting that line at a point whose x-coordinate must therefore also be x_1. Then the process repeats; the line goes vertically to the curve $y = f(x)$ and horizontally to $y = x$, arriving at $x = x_2$. The line continues in this way, "spiralling" closer and closer to the intersection of $y = f(x)$ and $y = x$. Each value of x_n is closer to the fixed point r than the previous value.

Now consider the function f whose graph appears in Figure 4.11(a). If we try the same method there, starting with x_0, the polygonal line spirals outward, away from the root, and the resulting values x_n will not "converge" to the root as they did in Example 1. To see why the method works for the function in Figure 4.10 but not for the function in Figure 4.11(a), observe the slopes of the two graphs $y = f(x)$ near the fixed point r. Both slopes are negative, but in Figure 4.10 the absolute value of the slope is less than 1 while the absolute value of the slope of f in Figure 4.11(a) is greater than 1. Close consideration of the graphs should convince you that it is this fact that caused the points x_n to get closer to r in Figure 4.10 and farther from r in Figure 4.11(a).

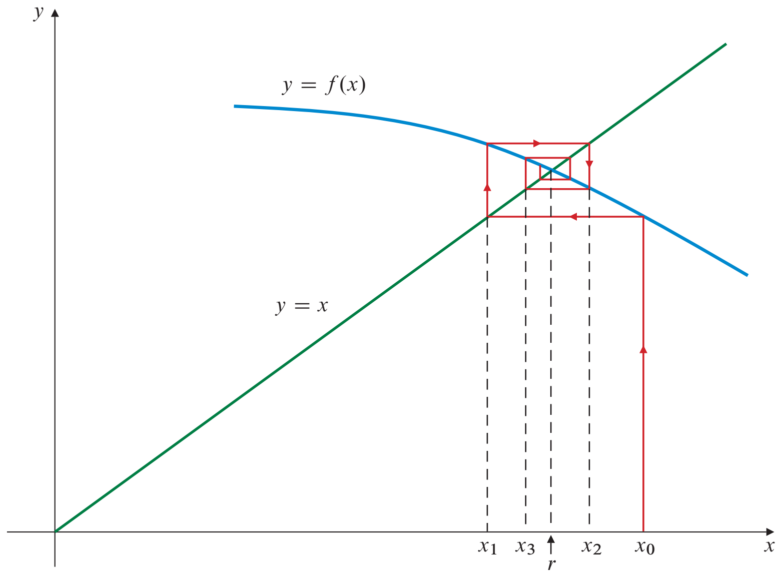


Figure 4.10 Iterations of $x_{n+1} = f(x_n)$ "spiral" toward the fixed point

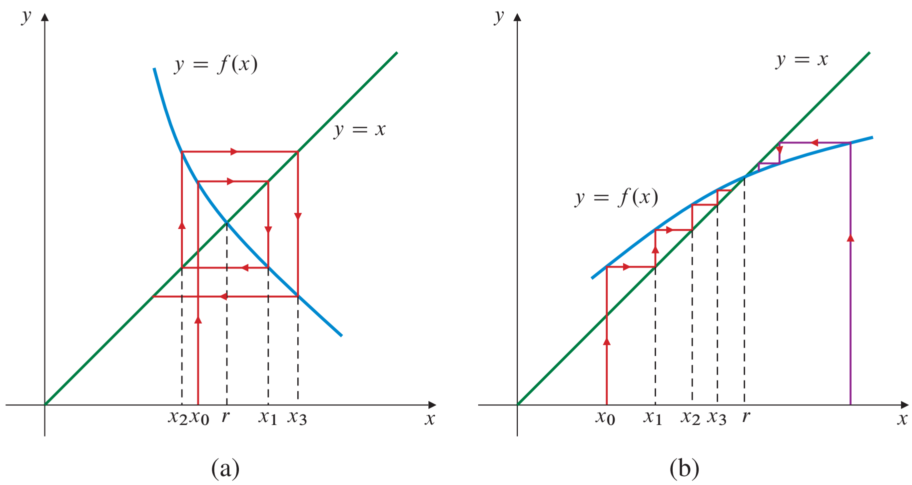


Figure 4.11

(a) A function f for which the iterations $x_{n+1} = f(x_n)$ do not converge

(b) "Staircase" convergence to the fixed point

A third example, Figure 4.11(b), shows that the method can be expected to work for functions whose graphs have positive slope near the fixed point r, provided that the slope is less than 1. In this case the polygonal line forms a "staircase" rather than a "spiral," and the successive approximations x_n increase toward the root if $x_0 < r$ and decrease toward it if $x_0 > r$.

Remark Note that if $|f'(x)| > 1$ near a fixed point r of f, you may still be able to find that fixed point by applying fixed-point iteration to $f^{-1}(x)$. Evidently $f^{-1}(r) = r$ if and only if $r = f(r)$.

The following theorem guarantees that the method of fixed-point iteration will work for a particular class of functions.

THEOREM 1

A fixed-point theorem

Suppose that f is defined on an interval $I = [a, b]$ and satisfies the following two conditions:

(i) $f(x)$ belongs to I whenever x belongs to I and

(ii) there exists a constant K with $0 < K < 1$ such that for every u and v in I,

$$|f(u) - f(v)| \le K|u - v|.$$

Then f has a unique fixed point r in I, that is, $f(r) = r$, and starting with any number x_0 in I, the iterates

$$x_1 = f(x_0), \quad x_2 = f(x_1), \quad \ldots \qquad \text{converge to } r.$$

You are invited to prove this theorem by a method outlined in Exercises 26 and 27 at the end of this section.

EXAMPLE 2 Show that if $0 < k < 1$, then $f(x) = k \cos x$ satisfies the conditions of Theorem 1 on the interval $I = [0, 1]$. Observe that if $k = 1/5$, the fixed point is that calculated in Example 1 above.

Solution Since $0 < k < 1$, f maps I into I. If u and v are in I, then the Mean-Value Theorem says there exists c between u and v such that

$$|f(u) - f(v)| = |(u - v)f'(c)| = k|u - v| \sin c \le k|u - v|.$$

Thus, the conditions of Theorem 1 are satisfied and f has a fixed point r in $[0, 1]$. Of course, even if $k \ge 1$, f may still have a fixed point in I locatable by iteration, provided the slope of f near that point is less than 1.

Newton's Method

We want to find a **root** of the equation $f(x) = 0$, that is, a number r such that $f(r) = 0$. Such a number is also called a **zero** of the function f. If f is differentiable near the root, then tangent lines can be used to produce a sequence of approximations to the root that approaches the root quite quickly. The idea is as follows (see Figure 4.12). Make an initial guess at the root, say $x = x_0$. Draw the tangent line to $y = f(x)$ at $(x_0, f(x_0))$, and find x_1, the x-intercept of this tangent line. Under certain circumstances x_1 will be closer to the root than x_0 was. The process can be repeated over and over to get numbers $x_2, x_3, \ldots$, getting closer and closer to the root r. The number x_{n+1} is the x-intercept of the tangent line to $y = f(x)$ at $(x_n, f(x_n))$.

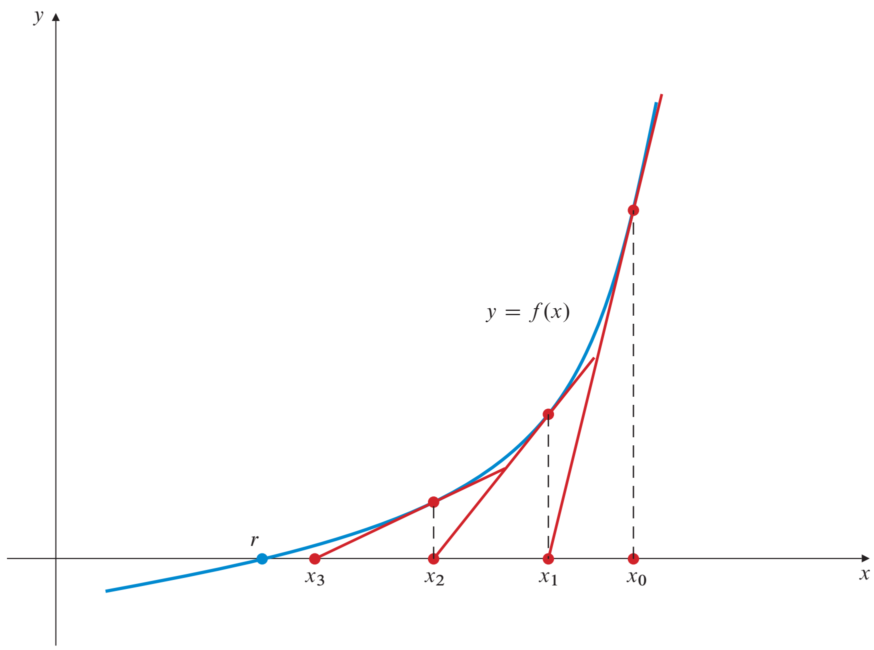


Figure 4.12

The tangent line to $y = f(x)$ at $x = x_0$ has equation

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Since the point $(x_1, 0)$ lies on this line, we have $0 = f(x_0) + f'(x_0)(x_1 - x_0)$. Hence,

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Similar formulas produce x_2 from x_1, then x_3 from x_2, and so on. The formula producing x_{n+1} from x_n is the discrete map $x_{n+1} = g(x_n)$, where $g(x) = x - \dfrac{f(x)}{f'(x)}$. That is,

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

which is known as the **Newton's Method formula**. If r is a fixed point of g then $f(r) = 0$ and r is a zero of f. We usually use a calculator or computer to calculate the successive approximations x_1, x_2, x_3, $\ldots$, and observe whether these numbers appear to converge to a limit. Convergence will not occur if the graph of f has a horizontal or vertical tangent at any of the numbers in the sequence. However, if $\lim_{n\to\infty} x_n = r$ exists, and if f/f' is continuous near r, then r must be a zero of f. This method is known as **Newton's Method** or **The Newton-Raphson Method**. Since Newton's Method is just a special case of fixed-point iteration applied to the function $g(x)$ defined above, the general properties of fixed-point iteration apply to Newton's Method as well.

EXAMPLE 3 Use Newton's Method to find the only real root of the equation $x^3 - x - 1 = 0$ correct to 10 decimal places.

Solution We have $f(x) = x^3 - x - 1$ and $f'(x) = 3x^2 - 1$. Since f is continuous and since $f(1) = -1$ and $f(2) = 5$, the equation has a root in the interval $[1, 2]$. Figure 4.13 shows that the equation has only one root to the right of $x = 0$. Let us make the initial guess $x_0 = 1.5$. The Newton's Method formula here is

$$x_{n+1} = x_n - \frac{x_n^3 - x_n - 1}{3x_n^2 - 1} = \frac{2x_n^3 + 1}{3x_n^2 - 1},$$

so that, for example, the approximation x_1 is given by

$$x_1 = \frac{2(1.5)^3 + 1}{3(1.5)^2 - 1} \approx 1.347\,826\ldots.$$

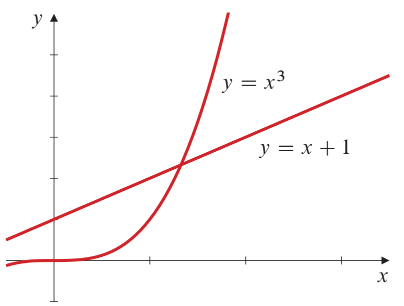


Figure 4.13 The graphs of x^3 and $x + 1$ meet only once to the right of $x = 0$, and that meeting is between 1 and 2

The values of $x_1, x_2, x_3, \ldots$ are given in Table 2.

Table 2.

n	x_n	$f(x_n)$
0	1.5	0.875 000 000 000 $\cdots$
1	1.347 826 086 96 $\cdots$	0.100 682 173 091 $\cdots$
2	1.325 200 398 95 $\cdots$	0.002 058 361 917 $\cdots$
3	1.324 718 174 00 $\cdots$	0.000 000 924 378 $\cdots$
4	1.324 717 957 24 $\cdots$	0.000 000 000 000 $\cdots$
5	1.324 717 957 24 $\cdots$	

The values in Table 2 were obtained with a scientific calculator. Evidently $r = 1.324\,717\,957\,2$ correctly rounded to 10 decimal places.

Observe the behaviour of the numbers x_n. By the third iteration, x_3, we have apparently achieved a precision of 6 decimal places, and by x_4 over 10 decimal places. It is characteristic of Newton's Method that when you begin to get close to the root the convergence can be very rapid. Compare these results with those obtained for the same equation by the Bisection Method in Example 12 of Section 1.4; there we achieved only 3 decimal place precision after 11 iterations.

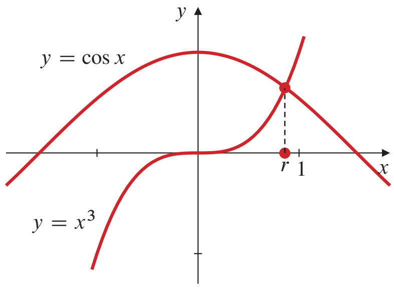


Figure 4.14 Solving $x^3 = \cos x$

EXAMPLE 4 Solve the equation $x^3 = \cos x$ to 11 decimal places.

Solution We are looking for the x-coordinate r of the intersection of the curves $y = x^3$ and $y = \cos x$. From Figure 4.14 it appears that the curves intersect slightly to the left of $x = 1$. Let us start with the guess $x_0 = 0.8$. If $f(x) = x^3 - \cos x$, then $f'(x) = 3x^2 + \sin x$. The Newton's Method formula for this function is

$$x_{n+1} = x_n - \frac{x_n^3 - \cos x_n}{3x_n^2 + \sin x_n} = \frac{2x_n^3 + x_n \sin x_n + \cos x_n}{3x_n^2 + \sin x_n}.$$

The approximations $x_1, x_2, \ldots$ are given in Table 3.

Table 3.

n	x_n	$f(x_n)$
0	0.8	−0.184 706 709 347···
1	0.870 034 801 135···	0.013 782 078 762···
2	0.865 494 102 425···	0.000 006 038 051···
3	0.865 474 033 493···	0.000 000 001 176···
4	0.865 474 033 102···	0.000 000 000 000···
5	0.865 474 033 102···	

The two curves intersect at $x = 0.865\,474\,033\,10$, rounded to 11 decimal places.

Remark Example 4 shows how useful a sketch can be for determining an initial guess x_0. Even a rough sketch of the graph of $y = f(x)$ can show you how many roots the equation $f(x) = 0$ has and approximately where they are. Usually, the closer the initial approximation is to the actual root, the smaller the number of iterations needed to achieve the desired precision. Similarly, for an equation of the form $g(x) = h(x)$, making a sketch of the graphs of g and h (on the same set of axes) can suggest starting approximations for any intersection points. In either case, you can then apply Newton's Method to improve the approximations.

Remark When using Newton's Method to solve an equation that is of the form $g(x) = h(x)$ (such as the one in Example 4), we must rewrite the equation in the form $f(x) = 0$ and apply Newton's Method to f. Usually we just use $f(x) = g(x) - h(x)$, although $f(x) = \big(g(x)/h(x)\big) - 1$ is also a possibility.

Remark If your calculator is programmable, you should learn how to program the Newton's Method formula for a given equation so that generating new iterations requires pressing only a few buttons. If your calculator has graphing capabilities, you can use them to locate a good initial guess.

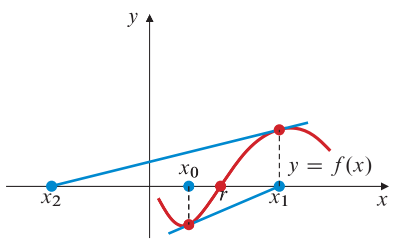


Figure 4.15 Here the Newton's Method iterations do not converge to the root

Newton's Method does not always work as well as it does in the preceding examples. If the first derivative f' is very small near the root, or if the second derivative f'' is very large near the root, a single iteration of the formula can take us from quite close to the root to quite far away. Figure 4.15 illustrates this possibility. (Also see Exercises 21 and 22 at the end of this section.)

Before you try to use Newton's Method to find a real root of a funcion f, you should make sure that a real root actually exists. If you use the method starting with a real initial guess, but the function has no real root nearby, the successive "approximations" can exhibit strange behaviour. The following example illustrates this for a very simple function.

EXAMPLE 5 Consider the function $f(x) = 1 + x^2$. Clearly f has no real roots though it does have complex roots $x = \pm i$. The Newton's Method formula for f is

$$x_{n+1} = x_n - \frac{1+x_n^2}{2x_n} = \frac{x_n^2-1}{2x_n}.$$

If we start with a real guess $x_0 = 2$, iterate this formula 20,000 times, and plot the resulting points (n, x_n), we obtain Figure 4.16, which was done using a Maple procedure. It is clear from this plot that not only do the iterations not converge (as one might otherwise expect), but they do not diverge to ∞ or $-\infty$, and they are not periodic either. This phenomenon is known as **chaos**.

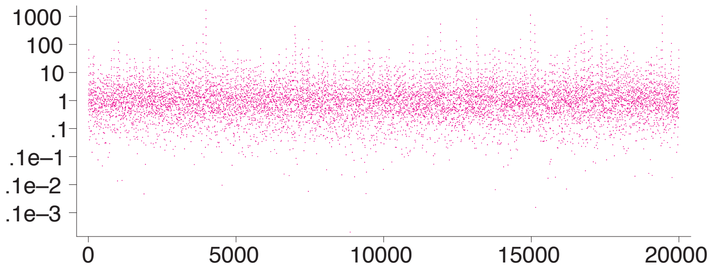


Figure 4.16 Plot of 20,000 points (n, x_n) for Example 5

A definitive characteristic of this phenomenon is sensitivity to initial conditions. To demonstrate this sensitivity in the case at hand we make a change of variables. Let

$$y_n = \frac{1}{1+x_n^2},$$

then the Newton's Method formula for f becomes

$$y_{n+1} = 4y_n(1-y_n),$$

(see Exercise 24), which is a special case of a discrete map called the **logistic map**. It represents one of the best-known and simplest examples of chaos. If, for example, $y_n = \sin^2(u_n)$, for $n = 0, 1, 2, \ldots$, then it follows (see Exercise 25 below) that $u_n = 2^n u_0$. Unless u_0 is a rational multiple of π, it follows that two different choices of u_0 will lead to differences in the resulting values of u_n that grow exponentially with n. In Exercise 25 it is shown that this sensitivity is carried through to the first order in x_n.

Remark The above example does not imply that Newton's Method cannot be used to find complex roots; the formula simply cannot escape from the real line if a real initial guess is used. To accomodate a complex initial guess, $z_0 = a_0 + ib_0$, we can substitute, $z_n = a_n + ib_n$ into the complex version of Newton's Method formula $z_{n+1} = \dfrac{z_n^2-1}{2z_n}$ (see Appendix I for a discussion of complex arithmetic) to get the following coupled equations:

$$a_{n+1} = \frac{a_n^3 + a_n(b_n^2-1)}{2(a_n^2+b_n^2)}$$

$$b_{n+1} = \frac{b_n^3 + b_n(a_n^2+1)}{2(a_n^2+b_n^2)}.$$

With initial guess $z_0 = 1 + i$, the next six members of the sequence of complex numbers (in 14-figure precision) become

$$z_1 = 0.250\,000\,000\,000\,00 + i\,0.750\,000\,000\,000\,00$$
$$z_2 = -0.075\,000\,000\,000\,00 + i\,0.975\,000\,000\,000\,00$$
$$z_3 = 0.001\,715\,686\,274\,51 + i\,0.997\,303\,921\,568\,63$$
$$z_4 = -0.000\,004\,641\,846\,27 + i\,1.000\,002\,160\,490\,67$$
$$z_5 = -0.000\,000\,000\,010\,03 + i\,0.999\,999\,999\,991\,56$$
$$z_6 = 0.000\,000\,000\,000\,00 + i\,1.000\,000\,000\,000\,00$$

converging to the root $+i$. For an initial guess, $1-i$, the resulting sequence converges as rapidly to the root $-i$. Note that for the real initial guess $z_0 = 0 + i0$, neither a_1 nor b_1 is defined, so the process fails. This corresponds to the fact that $1 + x^2$ has a horizontal tangent $y = 1$ at $(0, 1)$, and this tangent has no finite x-intercept.

The following theorem gives sufficient conditions for the Newton approximations to converge to a root r of the equation $f(x) = 0$ if the initial guess x_0 is sufficiently close to that root.

THEOREM 2

Error bounds for Newton's Method

Suppose that f, f', and f'' are continuous on an interval I containing x_n, x_{n+1}, and a root $x = r$ of $f(x) = 0$. Suppose also that there exist constants $K > 0$ and $L > 0$ such that for all x in I we have

(i) $|f''(x)| \le K$ and

(ii) $|f'(x)| \ge L$.

Then

(a) $|x_{n+1} - r| \le \dfrac{K}{2L}|x_{n+1} - x_n|^2$ and

(b) $|x_{n+1} - r| \le \dfrac{K}{2L}|x_n - r|^2$.

Conditions (i) and (ii) assert that near r the slope of $y = f(x)$ is not too small in size and does not change too rapidly. If $K/(2L) < 1$, the theorem shows that x_n converges quickly to r once n becomes large enough that $|x_n - r| < 1$.

The proof of Theorem 2 depends on the Mean-Value Theorem. We will not give it since the theorem is of little practical use. In practice, we calculate successive approximations using Newton's formula and observe whether they seem to converge to a limit. If they do, and if the values of f at these approximations approach 0, we can be confident that we have located a root.

"Solve" Routines

Many of the more advanced models of scientific calculators and most computer-based mathematics software have built-in routines for solving general equations numerically or, in a few cases, symbolically. These "Solve" routines assume continuity of the left and right sides of the given equations and often require the user to specify an interval in which to search for the root or an initial guess at the value of the root, or both. Typically the calculator or computer software also has graphing capabilities, and you are expected to use them to get an idea of how many roots the equation has and roughly where they are located before invoking the solving routines. It may also be possible to specify a *tolerance* on the difference of the two sides of the equation. For instance, if we want a solution to the equation $f(x) = 0$, it may be more important to us to be sure that an approximate solution $\hat{x}$ satisfies $|f(\hat{x})| < 0.0001$ than it is to be sure that $\hat{x}$ is within any particular distance of the actual root.

The methods used by the solve routines vary from one calculator or software package to another and are frequently very sophisticated, making use of numerical differentiation and other techniques to find roots very quickly, even when the search interval is large. If you have an advanced scientific calculator and/or computer software with similar capabilities, it is well worth your while to read the manuals that describe how to make effective use of your hardware/software for solving equations. Applications of mathematics to solving "real-world" problems frequently require finding approximate solutions of equations that are intractable by exact methods.

EXERCISES 4.2

Use fixed-point iteration to solve the equations in Exercises 1–6. Obtain 5 decimal place precision.

1. $2x = e^{-x}$, start with $x_0 = 0.3$

2. $1 + \frac{1}{4}\sin x = x$

3. $\cos \dfrac{x}{3} = x$

4. $(x+9)^{1/3} = x$

5. $\dfrac{1}{2+x^2} = x$

6. Solve $x^3 + 10x - 10 = 0$ by rewriting it in the form $1 - \frac{1}{10}x^3 = x$.

In Exercises 7–16, use Newton's Method to solve the given equations to the precision permitted by your calculator.

7. Find $\sqrt{2}$ by solving $x^2 - 2 = 0$.

8. Find $\sqrt{3}$ by solving $x^2 - 3 = 0$.

9. Find the root of $x^3 + 2x - 1 = 0$ between 0 and 1.

10. Find the root of $x^3 + 2x^2 - 2 = 0$ between 0 and 1.

11. Find the two roots of $x^4 - 8x^2 - x + 16 = 0$ in $[1, 3]$.

12. Find the three roots of $x^3 + 3x^2 - 1 = 0$ in $[-3, 1]$.

13. Solve $\sin x = 1 - x$. A sketch can help you make a guess x_0.

14. Solve $\cos x = x^2$. How many roots are there?

15. How many roots does the equation $\tan x = x$ have? Find the one between $\pi/2$ and $3\pi/2$.

16. Solve $\dfrac{1}{1+x^2} = \sqrt{x}$ by rewriting it $(1+x^2)\sqrt{x} - 1 = 0$.

17. If your calculator has a built-in Solve routine, or if you use computer software with such a routine, use it to solve the equations in Exercises 7–16.

Find the maximum and minimum values of the functions in Exercises 18–19.

18. $\dfrac{\sin x}{1+x^2}$

19. $\dfrac{\cos x}{1+x^2}$

20. Let $f(x) = x^2$. The equation $f(x) = 0$ clearly has solution $x = 0$. Find the Newton's Method iterations x_1, x_2, and x_3, starting with $x_0 = 1$.

(a) What is x_n?

(b) How many iterations are needed to find the root with error less than 0.0001 in absolute value?

(c) How many iterations are needed to get an approximation x_n for which $|f(x_n)| < 0.0001$?

(d) Why do the Newton's Method iterations converge more slowly here than in the examples done in this section?

21. **(Oscillation)** Apply Newton's Method to

$$f(x) = \begin{cases} \sqrt{x} & \text{if } x \ge 0, \\ \sqrt{-x} & \text{if } x < 0, \end{cases}$$

starting with the initial guess $x_0 = a > 0$. Calculate x_1 and x_2. What happens? (Make a sketch.) If you ever observed this behaviour when you were using Newton's Method to find a root of an equation, what would you do next?

22. **(Divergent oscillations)** Apply Newton's Method to $f(x) = x^{1/3}$ with $x_0 = 1$. Calculate x_1, x_2, x_3, and x_4. What is happening? Find a formula for x_n.

23. **(Convergent oscillations)** Apply Newton's Method to find $f(x) = x^{2/3}$ with $x_0 = 1$. Calculate x_1, x_2, x_3, and x_4. What is happening? Find a formula for x_n.

24. Verify that the Newton's Method map for $1 + x^2$, namely $x_{n+1} = x_n - \dfrac{1+x_n^2}{2x_n}$, transforms into the logistic map $y_{n+1} = 4y_n(1 - y_n)$ under the transformation $y_n = \dfrac{1}{1+x_n^2}$.

25. Sensitivity to initial conditions is regarded as a definitive property of chaos. If the initial values of two sequences differ, and the differences between the two sequences tends to grow exponentially, the map is said to be sensitive to initial values. Growing exponentially in this sense does not require that each sequence grow exponentially on its own. In fact, for chaos the growth should only be exponential in the differential. Moreover, the growth only needs to be exponential for large n.

a) Show that the logistic map is sensitive to initial conditions by making the substitution $y_j = \sin^2 u_j$ and taking the differential, given that u_0 is not an integral multiple of π.

b) Use part (a) to show that the Newton's Method map for $1 + x^2$ is also sensitive to initial conditions. Make the reasonable assumption, based on Figure 4.16, that the iterates neither converge nor diverge.

Exercises 26–27 constitute a proof of Theorem 1.

26. Condition (ii) of Theorem 1 implies that f is continuous on $I = [a, b]$. Use condition (i) to show that f has a unique fixed point r on I. *Hint:* Apply the Intermediate-Value Theorem to $g(x) = f(x) - x$ on $[a, b]$.

27. Use condition (ii) of Theorem 1 and mathematical induction to show that $|x_n - r| \le K^n|x_0 - r|$. Since $0 < K < 1$, we know that $K^n \to 0$ as $n \to \infty$. This shows that $\lim_{n\to\infty} x_n = r$.

4.3 Indeterminate Forms

In Section 2.5 we showed that

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

We could not readily see this by substituting $x = 0$ into the function $(\sin x)/x$ because both $\sin x$ and x are zero at $x = 0$. We call $(\sin x)/x$ an **indeterminate form** of type $[0/0]$ at $x = 0$. The limit of such an indeterminate form can be any number. For instance, each of the quotients kx/x, x/x^3, and x^3/x^2 is an indeterminate form of type [0/0] at $x = 0$, but

$$\lim_{x\to 0}\frac{kx}{x} = k, \qquad \lim_{x\to 0}\frac{x}{x^3} = \infty, \qquad \lim_{x\to 0}\frac{x^3}{x^2} = 0.$$

There are other types of indeterminate forms. Table 4 lists them together with an example of each type.

Table 4. Types of indeterminate forms

Type	Example
$[0/0]$	$\lim_{x\to 0}\frac{\sin x}{x}$
$[\infty/\infty]$	$\lim_{x\to 0}\frac{\ln(1/x^2)}{\cot(x^2)}$
$[0\cdot\infty]$	$\lim_{x\to 0+} x\ln\frac{1}{x}$
$[\infty-\infty]$	$\lim_{x\to(\pi/2)-}\left(\tan x - \frac{1}{\pi-2x}\right)$
$[0^0]$	$\lim_{x\to 0+} x^x$
$[\infty^0]$	$\lim_{x\to(\pi/2)-}(\tan x)^{\cos x}$
$[1^\infty]$	$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x$

Indeterminate forms of type [0/0] are the most common. You can evaluate many indeterminate forms of type [0/0] with simple algebra, typically by cancelling common factors. Examples can be found in Sections 1.2 and 1.3. We will now develop another method called **l'Hôpital's Rules**[1] for evaluating limits of indeterminate forms of the types [0/0] and [∞/∞]. The other types of indeterminate forms can usually be reduced to one of these two by algebraic manipulation and the taking of logarithms. In Section 4.10 we will discover yet another method for evaluating limits of type [0/0].

l'Hôpital's Rules

THEOREM 3

The first l'Hôpital Rule

Suppose the functions f and g are differentiable on the interval (a, b), and $g'(x) \neq 0$ there. Suppose also that

(i) $\lim_{x\to a+} f(x) = \lim_{x\to a+} g(x) = 0$ and

(ii) $\lim_{x\to a+}\frac{f'(x)}{g'(x)} = L$ (where L is finite or ∞ or $-\infty$).

Then

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

Similar results hold if every occurrence of $\lim_{x\to a+}$ is replaced by $\lim_{x\to b-}$ or even $\lim_{x\to c}$ where $a < c < b$. The cases $a = -\infty$ and $b = \infty$ are also allowed.

[1] The Marquis de l'Hôpital (1661–1704), for whom these rules are named, published the first textbook on calculus. The circumflex (^) did not come into use in the French language until after the French Revolution. The Marquis would have written his name "l'Hospital."

PROOF We prove the case involving $\lim_{x\to a+}$ for finite a. Define

$$F(x) = \begin{cases} f(x) & \text{if } a < x < b \\ 0 & \text{if } x = a \end{cases} \qquad \text{and} \qquad G(x) = \begin{cases} g(x) & \text{if } a < x < b \\ 0 & \text{if } x = a \end{cases}$$

Then F and G are continuous on the interval $[a, x]$ and differentiable on the interval (a, x) for every x in (a, b). By the Generalized Mean-Value Theorem (Theorem 16 of Section 2.8) there exists a number c in (a, 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 $a < c < x$, if $x \to a+$, then necessarily $c \to a+$, so we have

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

The case involving $\lim_{x\to b-}$ for finite b is proved similarly. The cases where $a = -\infty$ or $b = \infty$ follow from the cases already considered via the change of variable $x = 1/t$:

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{t\to 0+} \frac{f\left(\dfrac{1}{t}\right)}{g\left(\dfrac{1}{t}\right)} = \lim_{t\to 0+} \frac{f'\left(\dfrac{1}{t}\right)\left(\dfrac{-1}{t^2}\right)}{g'\left(\dfrac{1}{t}\right)\left(\dfrac{-1}{t^2}\right)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = L.$$

EXAMPLE 1 Evaluate $\displaystyle\lim_{x\to 1} \frac{\ln x}{x^2 - 1}$.

Solution We have

$$\lim_{x\to 1} \frac{\ln x}{x^2 - 1} \quad \left[\frac{0}{0}\right]$$
$$= \lim_{x\to 1} \frac{1/x}{2x} = \lim_{x\to 1} \frac{1}{2x^2} = \frac{1}{2}.$$

BEWARE! Note that in applying l'Hôpital's Rule we calculate the quotient of the derivatives, *not* the derivative of the quotient.

This example illustrates how calculations based on l'Hôpital's Rule are carried out. Having identified the limit as that of a $[0/0]$ indeterminate form, we replace it by the limit of the quotient of derivatives; the existence of this latter limit will justify the equality. It is possible that the limit of the quotient of derivatives may still be indeterminate, in which case a second application of l'Hôpital's Rule can be made. Such applications may be strung out until a limit can finally be extracted, which then justifies all the previous applications of the rule.

EXAMPLE 2 Evaluate $\displaystyle\lim_{x\to 0} \frac{2\sin x - \sin(2x)}{2e^x - 2 - 2x - x^2}$.

Solution We have (using l'Hôpital's Rule three times)

$$\begin{aligned}
&\lim_{x\to 0} \frac{2\sin x - \sin(2x)}{2e^x - 2 - 2x - x^2} && \left[\frac{0}{0}\right] \\
&= \lim_{x\to 0} \frac{2\cos x - 2\cos(2x)}{2e^x - 2 - 2x} && \text{cancel the 2s} \\
&= \lim_{x\to 0} \frac{\cos x - \cos(2x)}{e^x - 1 - x} && \text{still } \left[\frac{0}{0}\right] \\
&= \lim_{x\to 0} \frac{-\sin x + 2\sin(2x)}{e^x - 1} && \text{still } \left[\frac{0}{0}\right] \\
&= \lim_{x\to 0} \frac{-\cos x + 4\cos(2x)}{e^x} = \frac{-1 + 4}{1} = 3.
\end{aligned}$$

EXAMPLE 3 Evaluate (a) $\displaystyle\lim_{x\to(\pi/2)-}\frac{2x-\pi}{\cos^2 x}$ and (b) $\displaystyle\lim_{x\to 1+}\frac{x}{\ln x}$.

Solution

(a) $$\lim_{x\to(\pi/2)-}\frac{2x-\pi}{\cos^2 x}\qquad\left[\frac{0}{0}\right]$$
$$=\lim_{x\to(\pi/2)-}\frac{2}{-2\sin x\cos x}=-\infty$$

(b) l'Hôpital's Rule cannot be used to evaluate $\lim_{x\to 1+} x/(\ln x)$ because this is not an indeterminate form. The denominator approaches 0 as $x\to 1+$, but the numerator does not approach 0. Since $\ln x>0$ for $x>1$, we have, directly,

$$\lim_{x\to 1+}\frac{x}{\ln x}=\infty.$$

(Had we tried to apply l'Hôpital's Rule, we would have been led to the erroneous answer $\lim_{x\to 1+}(1/(1/x))=1$.)

BEWARE! Do not use l'Hôpital's Rule to evaluate a limit that is not indeterminate.

EXAMPLE 4 Evaluate $\displaystyle\lim_{x\to 0+}\left(\frac{1}{x}-\frac{1}{\sin x}\right)$.

Solution The indeterminate form here is of type $[\infty-\infty]$, to which l'Hôpital's Rule cannot be applied. However, it becomes $[0/0]$ after we combine the fractions into one fraction:

$$\begin{aligned}
&\lim_{x\to 0+}\left(\frac{1}{x}-\frac{1}{\sin x}\right) && [\infty-\infty]\\
&=\lim_{x\to 0+}\frac{\sin x-x}{x\sin x} && \left[\frac{0}{0}\right]\\
&=\lim_{x\to 0+}\frac{\cos x-1}{\sin x+x\cos x} && \left[\frac{0}{0}\right]\\
&=\lim_{x\to 0+}\frac{-\sin x}{2\cos x-x\sin x}=\frac{-0}{2}=0.
\end{aligned}$$

A version of l'Hôpital's Rule also holds for indeterminate forms of the type $[\infty/\infty]$.

THEOREM 4

The second l'Hôpital Rule

Suppose that f and g are differentiable on the interval (a,b) and that $g'(x)\neq 0$ there. Suppose also that

(i) $\displaystyle\lim_{x\to a+} g(x)=\pm\infty$ and

(ii) $\displaystyle\lim_{x\to a+}\frac{f'(x)}{g'(x)}=L$ (where L is finite, or ∞ or $-\infty$).

Then

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

Again, similar results hold for $\lim_{x\to b-}$ and for $\lim_{x\to c}$, and the cases $a=-\infty$ and $b=\infty$ are allowed.

The proof of the second l'Hôpital Rule is technically rather more difficult than that of the first Rule and we will not give it here. A sketch of the proof is outlined in Exercise 35 at the end of this section.

Remark Do *not* try to use l'Hôpital's Rules to evaluate limits that are not indeterminate of type $[0/0]$ or $[\infty/\infty]$; such attempts will almost always lead to false conclusions, as observed in Example 3(b) above. (Strictly speaking, the second l'Hôpital Rule can be applied to the form $[a/\infty]$, but there is no point to doing so if a is not infinite, since the limit is obviously 0 in that case.)

Remark No conclusion about $\lim f(x)/g(x)$ can be made using either l'Hôpital Rule if $\lim f'(x)/g'(x)$ does not exist. Other techniques might still be used. For example, $\lim_{x\to 0} (x^2 \sin(1/x))/\sin(x) = 0$ by the Squeeze Theorem even though $\lim_{x\to 0} (2x\sin(1/x) - \cos(1/x))/\cos(x)$ does not exist.

EXAMPLE 5 Evaluate (a) $\displaystyle\lim_{x\to\infty} \frac{x^2}{e^x}$ and (b) $\displaystyle\lim_{x\to 0+} x^a \ln x$, where $a > 0$.

Solution Both of these limits are covered by Theorem 5 in Section 3.4. We do them here by l'Hôpital's Rule.

(a)
$$\begin{aligned}
&\lim_{x\to\infty} \frac{x^2}{e^x} \qquad \left[\frac{\infty}{\infty}\right] \\
&= \lim_{x\to\infty} \frac{2x}{e^x} \qquad \text{still } \left[\frac{\infty}{\infty}\right] \\
&= \lim_{x\to\infty} \frac{2}{e^x} = 0.
\end{aligned}$$

Similarly, one can show that $\lim_{x\to\infty} x^n/e^x = 0$ for any positive integer n by repeated applications of l'Hôpital's Rule.

(b)
$$\begin{aligned}
&\lim_{x\to 0+} x^a \ln x \quad (a > 0) \qquad [0\cdot(-\infty)] \\
&= \lim_{x\to 0+} \frac{\ln x}{x^{-a}} \qquad \left[\frac{-\infty}{\infty}\right] \\
&= \lim_{x\to 0+} \frac{1/x}{-a x^{-a-1}} = \lim_{x\to 0+} \frac{x^a}{-a} = 0.
\end{aligned}$$

The easiest way to deal with indeterminate forms of types $[0^0]$, $[\infty^0]$, and $[1^\infty]$ is to take logarithms of the expressions involved. Here are two examples.

EXAMPLE 6 Evaluate $\displaystyle\lim_{x\to 0+} x^x$.

Solution This indeterminate form is of type $[0^0]$. Let $y = x^x$. Then

$$\lim_{x\to 0+} \ln y = \lim_{x\to 0+} x \ln x = 0,$$

by Example 5(b). Hence, $\displaystyle\lim_{x\to 0} x^x = \lim_{x\to 0+} y = e^0 = 1$.

EXAMPLE 7 Evaluate $\displaystyle\lim_{x\to\infty} \left(1 + \sin\frac{3}{x}\right)^x$.

Solution This indeterminate form is of type 1^∞. Let $y = \left(1 + \sin\dfrac{3}{x}\right)^x$. Then, taking ln of both sides,

$$\begin{aligned}
\lim_{x\to\infty} \ln y &= \lim_{x\to\infty} x \ln\left(1+\sin\frac{3}{x}\right) \qquad [\infty\cdot 0] \\
&= \lim_{x\to\infty} \frac{\ln\left(1+\sin\dfrac{3}{x}\right)}{\dfrac{1}{x}} \qquad \left[\frac{0}{0}\right] \\
&= \lim_{x\to\infty} \frac{\dfrac{1}{1+\sin\dfrac{3}{x}}\left(\cos\dfrac{3}{x}\right)\left(-\dfrac{3}{x^2}\right)}{-\dfrac{1}{x^2}} = \lim_{x\to\infty} \frac{3\cos\dfrac{3}{x}}{1+\sin\dfrac{3}{x}} = 3.
\end{aligned}$$

Hence, $\displaystyle\lim_{x\to\infty}\left(1+\sin\frac{3}{x}\right)^x = e^3.$

EXERCISES 4.3

Evaluate the limits in Exercises 1–32.

1. $\displaystyle\lim_{x\to 0}\frac{3x}{\tan 4x}$

2. $\displaystyle\lim_{x\to 2}\frac{\ln(2x-3)}{x^2-4}$

3. $\displaystyle\lim_{x\to 0}\frac{\sin ax}{\sin bx}$

4. $\displaystyle\lim_{x\to 0}\frac{1-\cos ax}{1-\cos bx}$

5. $\displaystyle\lim_{x\to 0}\frac{\sin^{-1}x}{\tan^{-1}x}$

6. $\displaystyle\lim_{x\to 1}\frac{x^{1/3}-1}{x^{2/3}-1}$

7. $\displaystyle\lim_{x\to 0} x\cot x$

8. $\displaystyle\lim_{x\to 0}\frac{1-\cos x}{\ln(1+x^2)}$

9. $\displaystyle\lim_{t\to \pi}\frac{\sin^2 t}{t-\pi}$

10. $\displaystyle\lim_{x\to 0}\frac{10^x-e^x}{x}$

11. $\displaystyle\lim_{x\to \pi/2}\frac{\cos 3x}{\pi-2x}$

12. $\displaystyle\lim_{x\to 1}\frac{\ln(ex)-1}{\sin \pi x}$

13. $\displaystyle\lim_{x\to \infty} x\,\sin\frac{1}{x}$

14. $\displaystyle\lim_{x\to 0}\frac{x-\sin x}{x^3}$

15. $\displaystyle\lim_{x\to 0}\frac{x-\sin x}{x-\tan x}$

16. $\displaystyle\lim_{x\to 0}\frac{2-x^2-2\cos x}{x^4}$

17. $\displaystyle\lim_{x\to 0+}\frac{\sin^2 x}{\tan x-x}$

18. $\displaystyle\lim_{r\to \pi/2}\frac{\ln\sin r}{\cos r}$

19. $\displaystyle\lim_{t\to \pi/2}\frac{\sin t}{t}$

20. $\displaystyle\lim_{x\to 1-}\frac{\arccos x}{x-1}$

21. $\displaystyle\lim_{x\to \infty} x(2\tan^{-1}x-\pi)$

22. $\displaystyle\lim_{t\to (\pi/2)-}(\sec t-\tan t)$

23. $\displaystyle\lim_{t\to 0}\left(\frac{1}{t}-\frac{1}{te^{at}}\right)$

24. $\displaystyle\lim_{x\to 0+} x^{\sqrt{x}}$

25. $\displaystyle\lim_{x\to 0+}(\csc x)^{\sin^2 x}$

26. $\displaystyle\lim_{x\to 1+}\left(\frac{x}{x-1}-\frac{1}{\ln x}\right)$

27. $\displaystyle\lim_{t\to 0}\frac{3\sin t-\sin 3t}{3\tan t-\tan 3t}$

28. $\displaystyle\lim_{x\to 0}\left(\frac{\sin x}{x}\right)^{1/x^2}$

29. $\displaystyle\lim_{t\to 0}(\cos 2t)^{1/t^2}$

30. $\displaystyle\lim_{x\to 0+}\frac{\csc x}{\ln x}$

31. $\displaystyle\lim_{x\to 1-}\frac{\ln\sin \pi x}{\csc \pi x}$

32. $\displaystyle\lim_{x\to 0}(1+\tan x)^{1/x}$

33. **(A Newton quotient for the second derivative)** Evaluate $\lim_{h\to 0}\dfrac{f(x+h)-2f(x)+f(x-h)}{h^2}$ if f is a twice differentiable function.

34. If f has a continuous third derivative, evaluate

$$\lim_{h\to 0}\frac{f(x+3h)-3f(x+h)+3f(x-h)-f(x-3h)}{h^3}.$$

35. **(Proof of the second l'Hôpital Rule)** Fill in the details of the following outline of a proof of the second l'Hôpital Rule (Theorem 4) for the case where a and L are both finite. Let $a<x<t<b$ and show that there exists c in (x,t) such that

$$\frac{f(x)-f(t)}{g(x)-g(t)} = \frac{f'(c)}{g'(c)}.$$

Now juggle the above equation algebraically into the form

$$\frac{f(x)}{g(x)}-L = \frac{f'(c)}{g'(c)}-L+\frac{1}{g(x)}\left(f(t)-g(t)\frac{f'(c)}{g'(c)}\right).$$

It follows that

$$\begin{aligned}
&\left|\frac{f(x)}{g(x)}-L\right| \\
&\le \left|\frac{f'(c)}{g'(c)}-L\right| + \frac{1}{|g(x)|}\left(|f(t)|+|g(t)|\left|\frac{f'(c)}{g'(c)}\right|\right).
\end{aligned}$$

Now show that the right side of the above inequality can be made as small as you wish (say, less than a positive number ϵ) by choosing first t and then x close enough to a. Remember, you are given that $\lim_{c\to a+}\big(f'(c)/g'(c)\big) = L$ and $\lim_{x\to a+}|g(x)| = \infty$.

4.4 Extreme Values

The first derivative of a function is a source of much useful information about the behaviour of the function. As we have already seen, the sign of f' tells us whether f is increasing or decreasing. In this section we use this information to find maximum and minimum values of functions. In Section 4.8 we will put the techniques developed here to use solving problems that require finding maximum and minimum values.

Maximum and Minimum Values

Recall (from Section 1.4) that a function has a maximum value at x_0 if $f(x) \le f(x_0)$ for all x in the domain of f. The maximum value is $f(x_0)$. To be more precise, we should call such a maximum value an *absolute* or *global* maximum because it is the largest value that f attains anywhere on its entire domain.

DEFINITION 1

Absolute extreme values

Function f has an **absolute maximum value** $f(x_0)$ at the point x_0 in its domain if $f(x) \le f(x_0)$ holds for every x in the domain of f.
Similarly, f has an **absolute minimum value** $f(x_1)$ at the point x_1 in its domain if $f(x) \ge f(x_1)$ holds for every x in the domain of f.

A function can have at most one absolute maximum or minimum value, although this value can be assumed at many points. For example, $f(x) = \sin x$ has absolute maximum value 1 occurring at every point of the form $x = (\pi/2) + 2n\pi$, where n is an integer, and an absolute minimum value -1 at every point of the form $x = -(\pi/2) + 2n\pi$. A function need not have any absolute extreme values. The function $f(x) = 1/x$ becomes arbitrarily large as x approaches 0 from the right, so has no finite absolute maximum. (Remember, ∞ is not a number and is not a value of f.) It doesn't have an absolute minimum either. Even a bounded function may not have an absolute maximum or minimum value. The function $g(x) = x$ with domain specified to be the *open* interval $(0, 1)$ has neither; the range of g is also the interval $(0, 1)$, and there is no largest or smallest number in this interval. Of course, if the domain of g (and therefore also its range) were extended to be the *closed* interval $[0, 1]$, then g would have both a maximum value, 1, and a minimum value, 0.

Maximum and minimum values of a function are collectively referred to as **extreme values**. The following theorem is a restatement (and slight generalization) of Theorem 8 of Section 1.4. It will prove very useful in some circumstances when we want to find extreme values.

THEOREM 5

Existence of extreme values

If the domain of the function f is a *closed, finite interval* or a union of finitely many such intervals, and if f is *continuous* on that domain, then f must have an absolute maximum value and an absolute minimum value.

Consider the graph $y = f(x)$ shown in Figure 4.17. Evidently the absolute maximum value of f is $f(x_2)$, and the absolute minimum value is $f(x_3)$. In addition to these extreme values, f has several other "local" maximum and minimum values corresponding to points on the graph that are higher or lower than neighbouring points. Observe that f has *local maximum values* at a, x_2, x_4, and x_6 and local minimum values at x_1, x_3, x_5, and b. The absolute maximum is the highest of the local maxima; the absolute minimum is the lowest of the local minima.

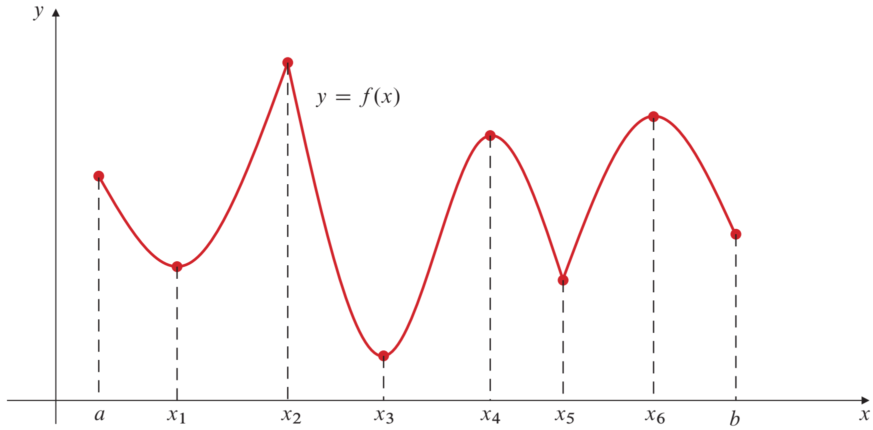


Figure 4.17 Local extreme values

DEFINITION 2

Local extreme values

Function f has a **local maximum value (loc max)** $f(x_0)$ at the point x_0 in its domain provided there exists a number $h > 0$ such that $f(x) \le f(x_0)$ whenever x is in the domain of f and $|x - x_0| < h$.

Similarly, f has a **local minimum value (loc min)** $f(x_1)$ at the point x_1 in its domain provided there exists a number $h > 0$ such that $f(x) \ge f(x_1)$ whenever x is in the domain of f and $|x - x_1| < h$.

Thus, f has a local maximum (or minimum) value at x if it has an absolute maximum (or minimum) value at x when its domain is restricted to points sufficiently near x. Geometrically, the graph of f is at least as high (or low) at x as it is at nearby points.

Critical Points, Singular Points, and Endpoints

Figure 4.17 suggests that a function $f(x)$ can have local extreme values only at points x of three special types:

(i) **critical points** of f (points x in $\mathcal{D}(f)$ where $f'(x) = 0$),

(ii) **singular points** of f (points x in $\mathcal{D}(f)$ where $f'(x)$ is not defined), and

(iii) **endpoints** of the domain of f (points in $\mathcal{D}(f)$ that do not belong to any open interval contained in $\mathcal{D}(f)$).

In Figure 4.17, x_1, x_3, x_4, and x_6 are critical points, x_2 and x_5 are singular points, and a and b are endpoints.

THEOREM 6

Locating extreme values

If the function f is defined on an interval I and has a local maximum (or local minimum) value at point $x = x_0$ in I, then x_0 must be either a critical point of f, a singular point of f, or an endpoint of I.

PROOF Suppose that f has a local maximum value at x_0 and that x_0 is neither an endpoint of the domain of f nor a singular point of f. Then for some $h > 0$, $f(x)$ is defined on the open interval $(x_0 - h, x_0 + h)$ and has an absolute maximum (for that interval) at x_0. Also, $f'(x_0)$ exists. By Theorem 14 of Section 2.8, $f'(x_0) = 0$. The proof for the case where f has a local minimum value at x_0 is similar. ∎

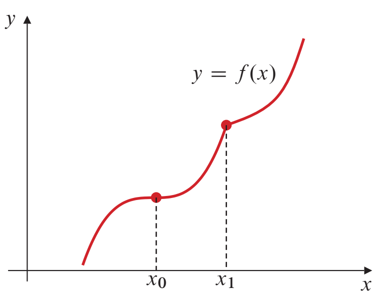


Figure 4.18 A function need not have extreme values at a critical point or a singular point

Although a function cannot have extreme values anywhere other than at endpoints, critical points, and singular points, it need not have extreme values at such points. Figure 4.18 shows the graph of a function with a critical point x_0 and a singular point x_1 at neither of which it has an extreme value. It is more difficult to draw the graph of a function whose domain has an endpoint at which the function fails to have an extreme value. See Exercise 49 at the end of this section for an example of such a function.

Finding Absolute Extreme Values

If a function f is defined on a closed interval or a union of finitely many closed intervals, Theorem 5 assures us that f must have an absolute maximum value and an absolute minimum value. Theorem 6 tells us how to find them. We need only check the values of f at any critical points, singular points, and endpoints.

EXAMPLE 1 Find the maximum and minimum values of the function $g(x) = x^3 - 3x^2 - 9x + 2$ on the interval $-2 \le x \le 2$.

Solution Since g is a polynomial, it can have no singular points. For critical points, we calculate

$$\begin{aligned} g'(x) = 3x^2 - 6x - 9 &= 3(x^2 - 2x - 3) \\ &= 3(x+1)(x-3) \\ &= 0 \qquad \text{if} \quad x = -1 \text{ or } x = 3. \end{aligned}$$

However, $x = 3$ is not in the domain of g, so we can ignore it. We need to consider only the values of g at the critical point $x = -1$ and at the endpoints $x = -2$ and $x = 2$:

$$g(-2) = 0, \qquad g(-1) = 7, \qquad g(2) = -20.$$

The maximum value of $g(x)$ on $-2 \le x \le 2$ is 7, at the critical point $x = -1$, and the minimum value is -20, at the endpoint $x = 2$. See Figure 4.19.

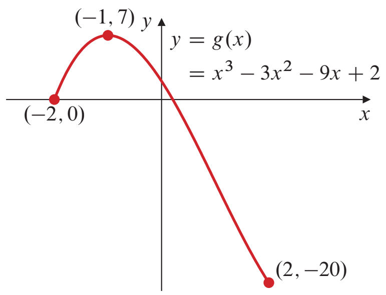


Figure 4.19 g has maximum and minimum values 7 and -20, respectively

EXAMPLE 2 Find the maximum and minimum values of $h(x) = 3x^{2/3} - 2x$ on the interval $[-1, 1]$.

Solution The derivative of h is

$$h'(x) = 3\left(\frac{2}{3}\right)x^{-1/3} - 2 = 2(x^{-1/3} - 1).$$

Note that $x^{-1/3}$ is not defined at the point $x = 0$ in $\mathcal{D}(h)$, so $x = 0$ is a singular point of h. Also, h has a critical point where $x^{-1/3} = 1$, that is, at $x = 1$, which also happens to be an endpoint of the domain of h. We must therefore examine the values of h at the points $x = 0$ and $x = 1$, as well as at the other endpoint $x = -1$. We have

$$h(-1) = 5, \qquad h(0) = 0, \qquad h(1) = 1.$$

The function h has maximum value 5 at the endpoint -1 and minimum value 0 at the singular point $x = 0$. See Figure 4.20.

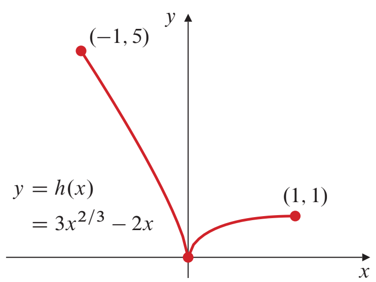


Figure 4.20 h has absolute minimum value 0 at a singular point

The First Derivative Test

Most functions you will encounter in elementary calculus have nonzero derivatives everywhere on their domains except possibly at a finite number of critical points, singular points, and endpoints of their domains. On intervals between these points the derivative exists and is not zero, so the function is either increasing or decreasing there. If f is continuous and increases to the left of x_0 and decreases to the right, then it must have a local maximum value at x_0. The following theorem collects several results of this type together.

THEOREM 7

The First Derivative Test

PART I. Testing interior critical points and singular points.

Suppose that f is continuous at x_0, and x_0 is not an endpoint of the domain of f.

(a) If there exists an open interval (a, b) containing x_0 such that $f'(x) > 0$ on (a, x_0) and $f'(x) < 0$ on (x_0, b), then f has a local maximum value at x_0.

(b) If there exists an open interval (a, b) containing x_0 such that $f'(x) < 0$ on (a, x_0) and $f'(x) > 0$ on (x_0, b), then f has a local minimum value at x_0.

PART II. Testing endpoints of the domain.

Suppose a is a left endpoint of the domain of f and f is right continuous at a.

(c) If $f'(x) > 0$ on some interval (a, b), then f has a local minimum value at a.

(d) If $f'(x) < 0$ on some interval (a, b), then f has a local maximum value at a.

Suppose b is a right endpoint of the domain of f and f is left continuous at b.

(e) If $f'(x) > 0$ on some interval (a, b), then f has a local maximum value at b.

(f) If $f'(x) < 0$ on some interval (a, b), then f has a local minimum value at b.

Remark If f' is positive (or negative) on *both* sides of a critical or singular point, then f has neither a maximum nor a minimum value at that point.

EXAMPLE 3 Find the local and absolute extreme values of $f(x) = x^4 - 2x^2 - 3$ on the interval $[-2, 2]$. Sketch the graph of f.

Solution We begin by calculating and factoring the derivative $f'(x)$:

$$f'(x) = 4x^3 - 4x = 4x(x^2 - 1) = 4x(x - 1)(x + 1).$$

The critical points are 0, -1, and 1. The corresponding values are $f(0) = -3$, $f(-1) = f(1) = -4$. There are no singular points. The values of f at the endpoints -2 and 2 are $f(-2) = f(2) = 5$. The factored form of $f'(x)$ is also convenient for determining the sign of $f'(x)$ on intervals between these endpoints and critical points. Where an odd number of the factors of $f'(x)$ are negative, $f'(x)$ will itself be negative; where an even number of factors are negative, $f'(x)$ will be positive. We summarize the positive/negative properties of $f'(x)$ and the implied increasing/decreasing behaviour of $f(x)$ in chart form:

	EP		CP		CP		CP		EP
x	-2		-1		0		1		2
f'		$-$	0	$+$	0	$-$	0	$+$	
f	max	↘	min	↗	max	↘	min	↗	max

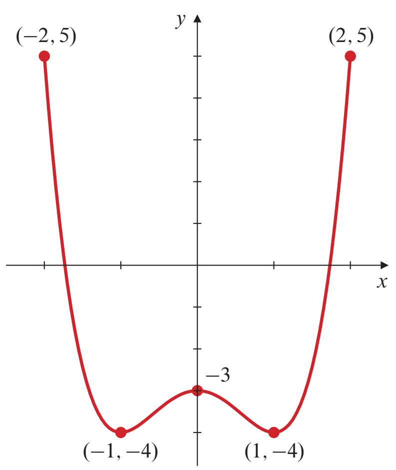


Figure 4.21 The graph $y = x^4 - 2x^2 - 3$

Note how the sloping arrows indicate visually the appropriate classification of the endpoints (EP) and critical points (CP) as determined by the First Derivative Test. We will make extensive use of such charts in future sections. The graph of f is shown in Figure 4.21. Since the domain is a closed, finite interval, f must have absolute maximum and minimum values. These are 5 (at ± 2) and -4 (at ± 1).

EXAMPLE 4 Find and classify the local and absolute extreme values of the function $f(x) = x - x^{2/3}$ with domain $[-1, 2]$. Sketch the graph of f.

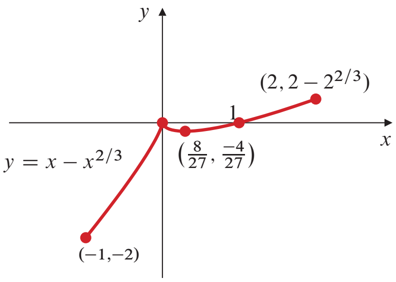


Figure 4.22 The graph for Example 4

Solution $f'(x) = 1 - \frac{2}{3}x^{-1/3} = \left(x^{1/3} - \frac{2}{3}\right)/x^{1/3}$. There is a singular point, $x = 0$, and a critical point, $x = 8/27$. The endpoints are $x = -1$ and $x = 2$. The values of f at these points are $f(-1) = -2$, $f(0) = 0$, $f(8/27) = -4/27$, and $f(2) = 2 - 2^{2/3} \approx 0.4126$ (see Figure 4.22). Another interesting point on the graph is the x-intercept at $x = 1$. Information from f' is summarized in the chart:

	EP		SP		CP		EP
x	-1		0		$8/27$		2
f'		$+$	undef	$-$	0	$+$	
f	min	↗	max	↘	min	↗	max

There are two local minima and two local maxima. The absolute maximum of f is $2 - 2^{2/3}$ at $x = 2$; the absolute minimum is -2 at $x = -1$.

Functions Not Defined on Closed, Finite Intervals

If the function f is not defined on a closed, finite interval, then Theorem 5 cannot be used to guarantee the existence of maximum and minimum values for f. Of course, f may still have such extreme values. In many applied situations we will want to find extreme values of functions defined on infinite and/or open intervals. The following theorem adapts Theorem 5 to cover some such situations.

THEOREM 8

Existence of extreme values on open intervals

If f is continuous on the open interval (a, b), and if

$$\lim_{x\to a+} f(x) = L \qquad \text{and} \qquad \lim_{x\to b-} f(x) = M,$$

then the following conclusions hold:

(i) If $f(u) > L$ and $f(u) > M$ for some u in (a, b), then f has an absolute maximum value on (a, b).

(ii) If $f(v) < L$ and $f(v) < M$ for some v in (a, b), then f has an absolute minimum value on (a, b).

In this theorem a may be $-\infty$, in which case $\lim_{x\to a+}$ should be replaced with $\lim_{x\to-\infty}$, and b may be ∞, in which case $\lim_{x\to b-}$ should be replaced with $\lim_{x\to\infty}$. Also, either or both of L and M may be either ∞ or $-\infty$.

PROOF We prove part (i); the proof of (ii) is similar. We are given that there is a number u in (a, b) such that $f(u) > L$ and $f(u) > M$. Here, L and M may be finite numbers or $-\infty$. Since $\lim_{x\to a+} f(x) = L$, there must exist a number x_1 in (a, u) such that

$$f(x) < f(u) \qquad \text{whenever} \quad a < x < x_1.$$

Similarly, there must exist a number x_2 in (u, b) such that

$$f(x) < f(u) \qquad \text{whenever} \quad x_2 < x < b.$$

(See Figure 4.23.) Thus, $f(x) < f(u)$ at all points of (a, b) that are not in the closed, finite subinterval $[x_1, x_2]$. By Theorem 5, the function f, being continuous on $[x_1, x_2]$, must have an absolute maximum value on that interval, say at the point w. Since u belongs to $[x_1, x_2]$, we must have $f(w) \geq f(u)$, so $f(w)$ is the maximum value of $f(x)$ for all of (a, b).

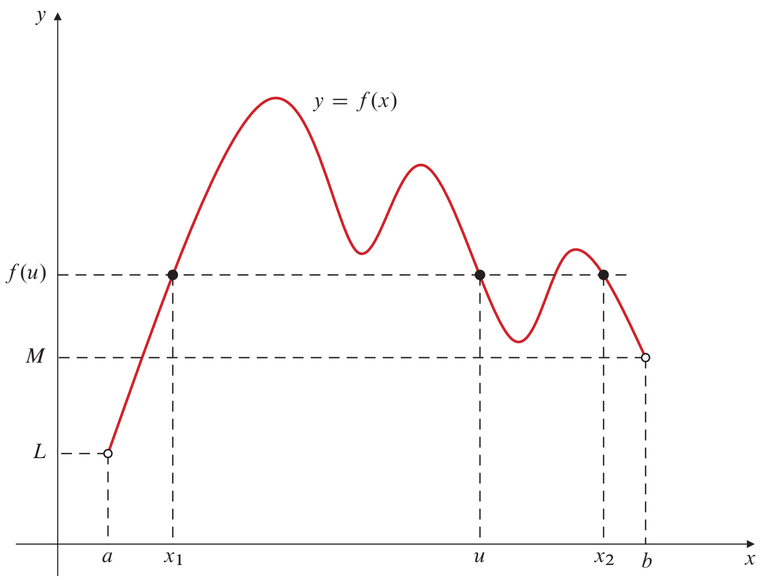


Figure 4.23

Theorem 6 still tells us where to look for extreme values. There are no endpoints to consider in an open interval, but we must still look at the values of the function at any critical points or singular points in the interval.

EXAMPLE 5 Show that $f(x) = x + (4/x)$ has an absolute minimum value on the interval $(0, \infty)$, and find that minimum value.

Solution We have

$$\lim_{x \to 0+} f(x) = \infty \qquad \text{and} \qquad \lim_{x \to \infty} f(x) = \infty.$$

Since $f(1) = 5 < \infty$, Theorem 8 guarantees that f must have an absolute minimum value at some point in $(0, \infty)$. To find the minimum value we must check the values of f at any critical points or singular points in the interval. We have

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

which equals 0 only at $x = 2$ and $x = -2$. Since f has domain $(0, \infty)$, it has no singular points and only one critical point, namely, $x = 2$, where f has the value $f(2) = 4$. This must be the minimum value of f on $(0, \infty)$. (See Figure 4.24.)

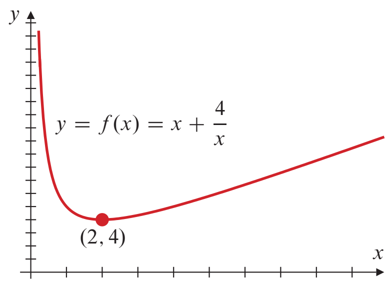


Figure 4.24 f has minimum value 4 at $x = 2$

EXAMPLE 6 Let $f(x) = x\,e^{-x^2}$. Find and classify the critical points of f, evaluate $\lim_{x \to \pm\infty} f(x)$, and use these results to help you sketch the graph of f.

Solution $f'(x) = e^{-x^2}(1 - 2x^2) = 0$ only if $1 - 2x^2 = 0$ since the exponential is always positive. Thus, the critical points are $\pm\frac{1}{\sqrt{2}}$. We have $f\left(\pm\frac{1}{\sqrt{2}}\right) = \pm\frac{1}{\sqrt{2e}}$. f' is positive (or negative) when $1 - 2x^2$ is positive (or negative). We summarize the intervals where f is increasing and decreasing in chart form:

		CP		CP	
x		$-1/\sqrt{2}$		$1/\sqrt{2}$	
f'	$-$	0	$+$	0	$-$
f	$\searrow$	min	$\nearrow$	max	$\searrow$

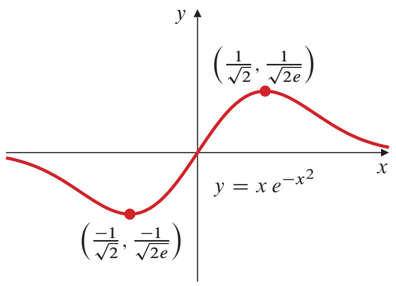


Figure 4.25 The graph for Example 6

Note that $f(0) = 0$ and that f is an odd function ($f(-x) = -f(x)$), so the graph is symmetric about the origin. Also,

$$\lim_{x\to\pm\infty} x\,e^{-x^2} = \left(\lim_{x\to\pm\infty} \frac{1}{x}\right)\left(\lim_{x\to\pm\infty} \frac{x^2}{e^{x^2}}\right) = 0 \times 0 = 0$$

because $\lim_{x\to\pm\infty} x^2 e^{-x^2} = \lim_{u\to\infty} u\,e^{-u} = 0$ by Theorem 5 of Section 3.4. Since $f(x)$ is positive at $x = 1/\sqrt{2}$ and is negative at $x = -1/\sqrt{2}$, f must have absolute maximum and minimum values by Theorem 8. These values can only be the values $\pm 1/\sqrt{2e}$ at the two critical points. The graph is shown in Figure 4.25. The x-axis is an asymptote as $x \to \pm\infty$.

EXERCISES 4.4

In Exercises 1–17, determine whether the given function has any local or absolute extreme values, and find those values if possible.

1. $f(x) = x + 2$ on $[-1, 1]$

2. $f(x) = x + 2$ on $(-\infty, 0]$

3. $f(x) = x + 2$ on $[-1, 1)$

4. $f(x) = x^2 - 1$

5. $f(x) = x^2 - 1$ on $[-2, 3]$

6. $f(x) = x^2 - 1$ on $(2, 3)$

7. $f(x) = x^3 + x - 4$ on $[a, b]$

8. $f(x) = x^3 + x - 4$ on (a, b)

9. $f(x) = x^5 + x^3 + 2x$ on $(a, b]$

10. $f(x) = \dfrac{1}{x-1}$

11. $f(x) = \dfrac{1}{x-1}$ on $(0, 1)$

12. $f(x) = \dfrac{1}{x-1}$ on $[2, 3]$

13. $f(x) = |x - 1|$ on $[-2, 2]$

14. $|x^2 - x - 2|$ on $[-3, 3]$

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

16. $f(x) = (x + 2)^{2/3}$

17. $f(x) = (x - 2)^{1/3}$

In Exercises 18–40, locate and classify all local extreme values of the given function. Determine whether any of these extreme values are absolute. Sketch the graph of the function.

18. $f(x) = x^2 + 2x$

19. $f(x) = x^3 - 3x - 2$

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

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

22. $f(x) = x^2(x - 1)^2$

23. $f(x) = x(x^2 - 1)^2$

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

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

26. $f(x) = \dfrac{x}{\sqrt{x^4+1}}$

27. $f(x) = x\sqrt{2 - x^2}$

28. $f(x) = x + \sin x$

29. $f(x) = x - 2\sin x$

30. $f(x) = x - 2\tan^{-1} x$

31. $f(x) = 2x - \sin^{-1} x$

32. $f(x) = e^{-x^2/2}$

33. $f(x) = x\,2^{-x}$

34. $f(x) = x^2 e^{-x^2}$

35. $f(x) = \dfrac{\ln x}{x}$

36. $f(x) = |x + 1|$

37. $f(x) = |x^2 - 1|$

38. $f(x) = \sin|x|$

39. $f(x) = |\sin x|$

40. $f(x) = (x - 1)^{2/3} - (x + 1)^{2/3}$

In Exercises 41–46, determine whether the given function has absolute maximum or absolute minimum values. Justify your answers. Find the extreme values if you can.

41. $\dfrac{x}{\sqrt{x^2+1}}$

42. $\dfrac{x}{\sqrt{x^4+1}}$

43. $x\sqrt{4 - x^2}$

44. $\dfrac{x^2}{\sqrt{4-x^2}}$

45. $\dfrac{1}{x\sin x}$ on $(0, \pi)$

46. $\dfrac{\sin x}{x}$

47. If a function has an absolute maximum value, must it have any local maximum values? If a function has a local maximum value, must it have an absolute maximum value? Give reasons for your answers.

48. If the function f has an absolute maximum value and $g(x) = |f(x)|$, must g have an absolute maximum value? Justify your answer.

49. **(A function with no max or min at an endpoint)** Let

$$f(x) = \begin{cases} x\sin\dfrac{1}{x} & \text{if } x > 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Show that f is continuous on $[0, \infty)$ and differentiable on $(0, \infty)$ but that it has neither a local maximum nor a local minimum value at the endpoint $x = 0$.

4.5 Concavity and Inflections

Like the first derivative, the second derivative of a function also provides useful information about the behaviour of the function and the shape of its graph: it determines whether the graph is *bending upward* (i.e., has increasing slope) or *bending downward* (i.e., has decreasing slope) as we move along the graph toward the right.

DEFINITION 3

We say that the function f is **concave up** on an open interval I if it is differentiable there and the derivative f' is an increasing function on I. Similarly, f is **concave down** on I if f' exists and is decreasing on I.

The terms "concave up" and "concave down" are used to describe the graph of the function as well as the function itself.

Note that concavity is defined only for differentiable functions, and even for those, only on intervals on which their derivatives are not constant. According to the above definition, a function is neither concave up nor concave down on an interval where its graph is a straight line segment. We say the function has no concavity on such an interval. We also say a function has opposite concavity on two intervals if it is concave up on one interval and concave down on the other.

The function f whose graph is shown in Figure 4.26 is concave up on the interval (a, b) and concave down on the interval (b, c).

Some geometric observations can be made about concavity:

(i) If f is concave up on an interval, then, on that interval, the graph of f lies above its tangents, and chords joining points on the graph lie above the graph.

(ii) If f is concave down on an interval, then, on that interval, the graph of f lies below its tangents, and chords to the graph lie below the graph.

(iii) If the graph of f has a tangent at a point, and if the concavity of f is opposite on opposite sides of that point, then the graph crosses its tangent at that point. (This occurs at the point $\big(b, f(b)\big)$ in Figure 4.26. Such a point is called an *inflection point* of the graph of f.)

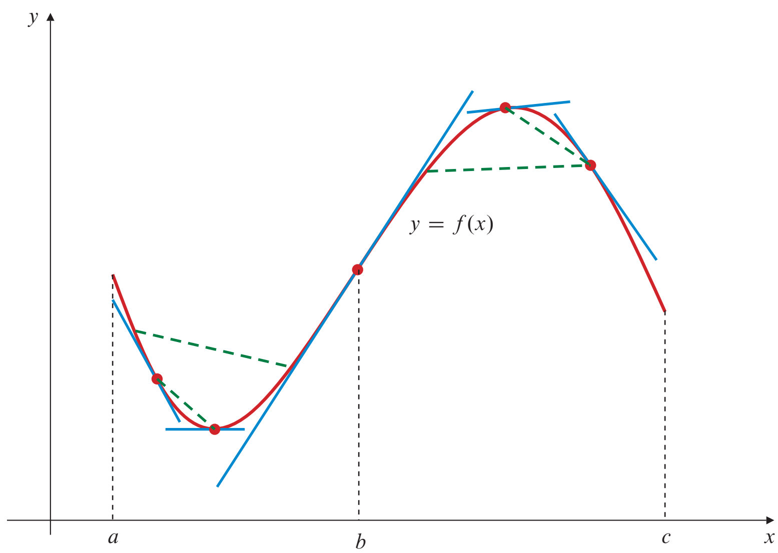


Figure 4.26 f is concave up on (a, b) and concave down on (b, c)

DEFINITION 4

Inflection points

We say that the point $\big(x_0, f(x_0)\big)$ *is* an **inflection point** of the curve $y = f(x)$ (or that the function f *has* an **inflection point** at x_0) if the following two conditions are satisfied:

(a) the graph of $y = f(x)$ has a tangent line at $x = x_0$, and

(b) the concavity of f is opposite on opposite sides of x_0.

Note that (a) implies that either f is differentiable at x_0 or its graph has a vertical tangent line there, and (b) implies that the graph crosses its tangent line at x_0. An inflection point of a function f is a point on the graph of a function, rather than a point in its domain like a critical point or a singular point. A function may or may not have an inflection point at a critical point or singular point. In general, a point P

is an inflection point (or simply *an inflection*) of a curve C (which is not necessarily the graph of a function) if C has a tangent at P and arcs of C extending in opposite directions from P are on opposite sides of that tangent line.

Figures 4.27–4.29 illustrate some situations involving critical and singular points and inflections.

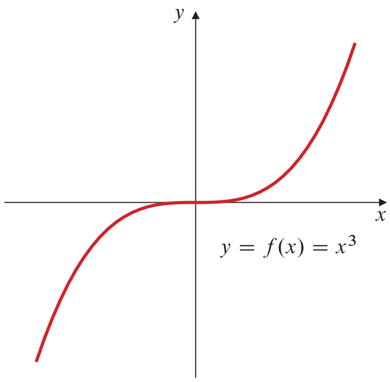


Figure 4.27 $x = 0$ is a critical point of $f(x) = x^3$, and f has an inflection point there

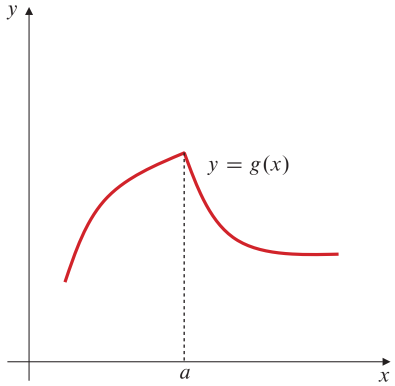


Figure 4.28 The concavity of g is opposite on opposite sides of the singular point a, but its graph has no tangent and therefore no inflection point there

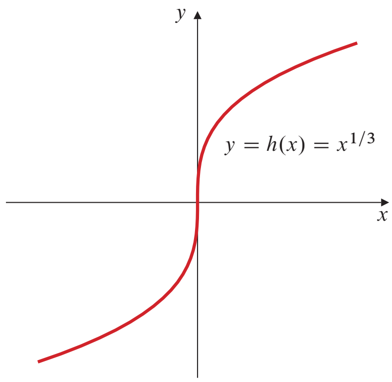


Figure 4.29 This graph of h has an inflection point at the origin even though $x = 0$ is a singular point of h

If a function f has a second derivative f'', the sign of that second derivative tells us whether the first derivative f' is increasing or decreasing and hence determines the concavity of f.

THEOREM 9

Concavity and the second derivative

(a) If $f''(x) > 0$ on interval I, then f is concave up on I.
(b) If $f''(x) < 0$ on interval I, then f is concave down on I.
(c) If f has an inflection point at x_0 and $f''(x_0)$ exists, then $f''(x_0) = 0$.

PROOF Parts (a) and (b) follow from applying Theorem 12 of Section 2.8 to the derivative f' of f. If f has an inflection point at x_0 and $f''(x_0)$ exists, then f must be differentiable in an open interval containing x_0. Since f' is increasing on one side of x_0 and decreasing on the other side, it must have a local maximum or minimum value at x_0. By Theorem 6, $f''(x_0) = 0$.

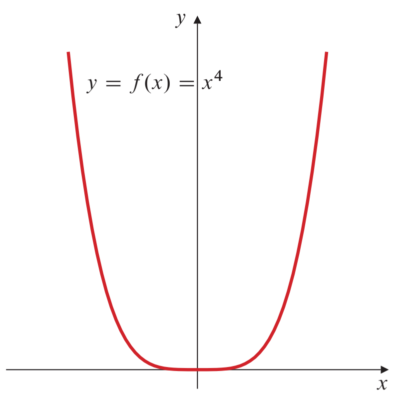


Figure 4.30 $f''(0) = 0$, but f does not have an inflection point at 0

Theorem 9 tells us that to find (the x-coordinates of) inflection points of a twice differentiable function f, we need only look at points where $f''(x) = 0$. However, not every such point has to be an inflection point. For example, $f(x) = x^4$, whose graph is shown in Figure 4.30, does not have an inflection point at $x = 0$ even though $f''(0) = 12x^2|_{x=0} = 0$. In fact, x^4 is concave up on every interval.

EXAMPLE 1 Determine the intervals of concavity of $f(x) = x^6 - 10x^4$ and the inflection points of its graph.

Solution We have

$$f'(x) = 6x^5 - 40x^3,$$
$$f''(x) = 30x^4 - 120x^2 = 30x^2(x-2)(x+2).$$

Having factored $f''(x)$ in this manner, we can see that it vanishes only at $x = -2$, $x = 0$, and $x = 2$. On the intervals $(-\infty, -2)$ and $(2, \infty)$, $f''(x) > 0$, so f is

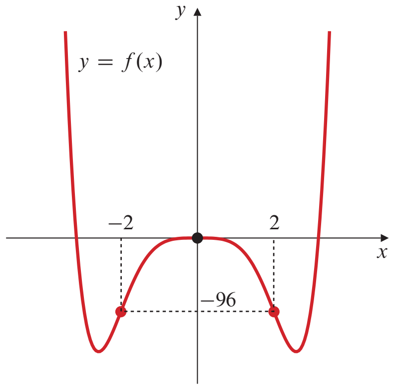


Figure 4.31 The graph of $f(x) = x^6 - 10x^4$

concave up. On $(-2, 0)$ and $(0, 2)$, $f''(x) < 0$, so f is concave down. $f''(x)$ changes sign as we pass through -2 and 2. Since $f(\pm 2) = -96$, the graph of f has inflection points at $(\pm 2, -96)$. However, $f''(x)$ does not change sign at $x = 0$, since $x^2 > 0$ for both positive and negative x. Thus, there is no inflection point at 0. As was the case for the first derivative, information about the sign of $f''(x)$ and the consequent concavity of f can be conveniently conveyed in a chart:

x		-2		0		2	
f''	$+$	0	$-$	0	$-$	0	$+$
f	$\smile$	infl	$\frown$		$\frown$	infl	$\smile$

The graph of f is sketched in Figure 4.31.

EXAMPLE 2 Determine the intervals of increase and decrease, the local extreme values, and the concavity of $f(x) = x^4 - 2x^3 + 1$. Use the information to sketch the graph of f.

Solution

$$f'(x) = 4x^3 - 6x^2 = 2x^2(2x - 3) = 0 \quad \text{at } x = 0 \text{ and } x = 3/2,$$
$$f''(x) = 12x^2 - 12x = 12x(x - 1) = 0 \quad \text{at } x = 0 \text{ and } x = 1.$$

The behaviour of f is summarized in the following chart:

		CP				CP	
x		0		1		$3/2$	
f'	$-$	0	$-$		$-$	0	$+$
f''	$+$	0	$-$	0	$+$		$+$
f	$\searrow$		$\searrow$		$\searrow$	min	$\nearrow$
	$\smile$	infl	$\frown$	infl	$\smile$		$\smile$

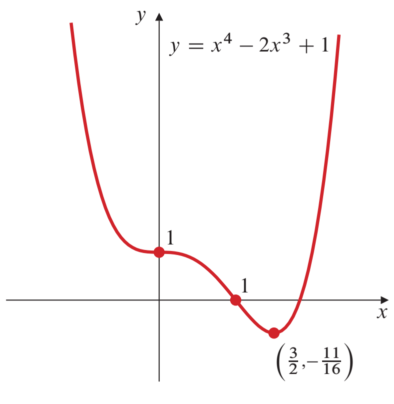


Figure 4.32 The function of Example 2

Note that f has an inflection at the critical point $x = 0$. We calculate the values of f at the "interesting values of x" in the charts:

$$f(0) = 1, \qquad f(1) = 0, \qquad f\left(\tfrac{3}{2}\right) = -\tfrac{11}{16}.$$

The graph of f is sketched in Figure 4.32.

The Second Derivative Test

A function f will have a local maximum (or minimum) value at a critical point if its graph is concave down (or up) in an interval containing that point. In fact, we can often use the value of the second derivative at the critical point to determine whether the function has a local maximum or a local minimum value there.

THEOREM 10

The Second Derivative Test

(a) If $f'(x_0) = 0$ and $f''(x_0) < 0$, then f has a local maximum value at x_0.
(b) If $f'(x_0) = 0$ and $f''(x_0) > 0$, then f has a local minimum value at x_0.
(c) If $f'(x_0) = 0$ and $f''(x_0) = 0$, no conclusion can be drawn; f may have a local maximum at x_0 or a local minimum, or it may have an inflection point instead.

PROOF Suppose that $f'(x_0) = 0$ and $f''(x_0) < 0$. Since

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

it follows that $f'(x_0+h) < 0$ for all sufficiently small positive h, and $f'(x_0+h) > 0$ for all sufficiently small negative h. By the first derivative test (Theorem 7), f must have a local maximum value at x_0. The proof of the local minimum case is similar.

The functions $f(x) = x^4$ (Figure 4.30), $f(x) = -x^4$, and $f(x) = x^3$ (Figure 4.27) all satisfy $f'(0) = 0$ and $f''(0) = 0$. But x^4 has a minimum value at $x = 0$, $-x^4$ has a maximum value at $x = 0$, and x^3 has neither a maximum nor a minimum value at $x = 0$ but has an inflection there. Therefore, we cannot make any conclusion about the nature of a critical point based on knowing that $f''(x) = 0$ there.

EXAMPLE 3 Find and classify the critical points of $f(x) = x^2e^{-x}$.

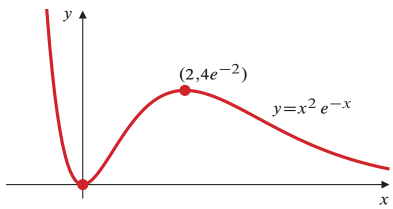


Figure 4.33 The critical points of $f(x) = x^2e^{-x}$

Solution We begin by calculating the first two derivatives of f:

$$f'(x) = (2x - x^2)e^{-x} = x(2-x)e^{-x} = 0 \quad \text{at } x = 0 \text{ and } x = 2,$$
$$f''(x) = (2 - 4x + x^2)e^{-x}$$
$$f''(0) = 2 > 0, \qquad f''(2) = -2e^{-2} < 0.$$

Thus, f has a local minimum value at $x = 0$ and a local maximum value at $x = 2$. See Figure 4.33.

For many functions the second derivative is more complicated to calculate than the first derivative, so the First Derivative Test is likely to be of more use in classifying critical points than is the Second Derivative Test. Also note that the First Derivative Test can classify local extreme values that occur at endpoints and singular points as well as at critical points.

It is possible to generalize the Second Derivative Test to obtain a higher derivative test to deal with some situations where the second derivative is zero at a critical point. (See Exercise 40 at the end of this section.)

EXERCISES 4.5

In Exercises 1–22, determine the intervals of constant concavity of the given function, and locate any inflection points.

1. $f(x) = \sqrt{x}$

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

3. $f(x) = x^2 + 2x + 3$

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

5. $f(x) = 10x^3 - 3x^5$

6. $f(x) = 10x^3 + 3x^5$

7. $f(x) = (3 - x^2)^2$

8. $f(x) = (2 + 2x - x^2)^2$

9. $f(x) = (x^2 - 4)^3$

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

11. $f(x) = \sin x$

12. $f(x) = \cos 3x$

13. $f(x) = x + \sin 2x$

14. $f(x) = x - 2\sin x$

15. $f(x) = \tan^{-1} x$

16. $f(x) = x\,e^x$

17. $f(x) = e^{-x^2}$

18. $f(x) = \dfrac{\ln(x^2)}{x}$

19. $f(x) = \ln(1 + x^2)$

20. $f(x) = (\ln x)^2$

21. $f(x) = \dfrac{x^3}{3} - 4x^2 + 12x - \dfrac{25}{3}$

22. $f(x) = (x-1)^{1/3} + (x+1)^{1/3}$

23. Discuss the concavity of the linear function $f(x) = ax + b$. Does it have any inflections?

Classify the critical points of the functions in Exercises 24–35 using the Second Derivative Test whenever possible.

24. $f(x) = 3x^3 - 36x - 3$

25. $f(x) = x(x-2)^2 + 1$

26. $f(x) = x + \dfrac{4}{x}$

27. $f(x) = x^3 + \dfrac{1}{x}$

28. $f(x) = \dfrac{x}{2^x}$

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

30. $f(x) = xe^x$

31. $f(x) = x\ln x$

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

33. $f(x) = (x^2-4)^3$

34. $f(x) = (x^2 - 3)e^x$ **35.** $f(x) = x^2 e^{-2x^2}$

36. Let $f(x) = x^2$ if $x \ge 0$ and $f(x) = -x^2$ if $x < 0$. Is 0 a critical point of f? Does f have an inflection point there? Is $f''(0) = 0$? If a function has a nonvertical tangent line at an inflection point, does the second derivative of the function necessarily vanish at that point?

37. Verify that if f is concave up on an interval, then its graph lies above its tangent lines on that interval. *Hint:* Suppose f is concave up on an open interval containing x_0. Let $h(x) = f(x) - f(x_0) - f'(x_0)(x - x_0)$. Show that h has a local minimum value at x_0 and hence that $h(x) \ge 0$ on the interval. Show that $h(x) > 0$ if $x \ne x_0$.

38. Verify that the graph $y = f(x)$ crosses its tangent line at an inflection point. *Hint:* Consider separately the cases where the tangent line is vertical and nonvertical.

39. Let $f_n(x) = x^n$ and $g_n(x) = -x^n$, $(n = 2, 3, 4, \ldots)$. Determine whether each function has a local maximum, a local minimum, or an inflection point at $x = 0$.

40. **(Higher Derivative Test)** Use your conclusions from Exercise 39 to suggest a generalization of the Second Derivative Test that applies when

$$f'(x_0) = f''(x_0) = \ldots = f^{(k-1)}(x_0) = 0, \quad f^{(k)}(x_0) \ne 0,$$

for some $k \ge 2$.

41. This problem shows that no test based solely on the signs of derivatives at x_0 can determine whether every function with a critical point at x_0 has a local maximum or minimum or an inflection point there. Let

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

Prove the following:

(a) $\lim_{x\to 0} x^{-n} f(x) = 0$ for $n = 0, 1, 2, 3, \ldots$.

(b) $\lim_{x\to 0} P(1/x) f(x) = 0$ for every polynomial P.

(c) For $x \ne 0$, $f^{(k)}(x) = P_k(1/x) f(x) (k = 1, 2, 3, \ldots)$, where P_k is a polynomial.

(d) $f^{(k)}(0)$ exists and equals 0 for $k = 1, 2, 3, \ldots$.

(e) f has a local minimum at $x = 0$; $-f$ has a local maximum at $x = 0$.

(f) If $g(x) = x f(x)$, then $g^{(k)}(0) = 0$ for every positive integer k and g has an inflection point at $x = 0$.

42. A function may have neither a local maximum nor a local minimum nor an inflection at a critical point. Show this by considering the following function:

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

Show that $f'(0) = f(0) = 0$, so the x-axis is tangent to the graph of f at $x = 0$; but $f'(x)$ is not continuous at $x = 0$, so $f''(0)$ does not exist. Show that the concavity of f is not constant on any interval with endpoint 0.

4.6 Sketching the Graph of a Function

When sketching the graph $y = f(x)$ of a function f, we have three sources of useful information:

(i) **the function f itself**, from which we determine the coordinates of some points on the graph, the symmetry of the graph, and any asymptotes;

(ii) **the first derivative, f'**, from which we determine the intervals of increase and decrease and the location of any local extreme values; and

(iii) **the second derivative, f''**, from which we determine the concavity and inflection points, and sometimes extreme values.

Items (ii) and (iii) were explored in the previous two sections. In this section we consider what we can learn from the function itself about the shape of its graph, and then we illustrate the entire sketching procedure with several examples using all three sources of information.

We could sketch a graph by plotting the coordinates of many points on it and joining them by a suitably smooth curve. This is what computer software and graphics calculators do. When carried out by hand (without a computer or calculator), this simplistic approach is at best tedious and at worst can fail to reveal the most interesting aspects of the graph (singular points, extreme values, and so on). We could also compute the slope at each of the plotted points and, by drawing short line segments through these points with the appropriate slopes, ensure that the sketched graph passes through each plotted point with the correct slope. A more efficient procedure is to obtain the coordinates of only a few points and use qualitative information from the function and its first and second derivatives to determine the *shape* of the graph between these points.

Besides critical and singular points and inflections, a graph may have other "interesting" points. The **intercepts** (points at which the graph intersects the coordinate axes) are usually among these. When sketching any graph it is wise to try to find all such intercepts, that is, all points with coordinates $(x, 0)$ and $(0, y)$ that lie on the graph. Of course, not every graph will have such points, and even when they do exist it may not always be possible to compute them exactly. Whenever a graph is made up of several disconnected pieces (called **components**), the coordinates of *at least one point on each component* must be obtained. It can sometimes be useful to determine the slopes at those points too. Vertical asymptotes (discussed below) usually break the graph of a function into components.

Realizing that a given function possesses some symmetry can aid greatly in obtaining a good sketch of its graph. In Section P.4 we discussed odd and even functions and observed that odd functions have graphs that are symmetric about the origin, while even functions have graphs that are symmetric about the y-axis, as shown in Figure 4.34. These are the symmetries you are most likely to notice, but functions can have other symmetries. For example, the graph of $2 + (x - 1)^2$ will certainly be symmetric about the line $x = 1$, and the graph of $2 + (x - 3)^3$ is symmetric about the point $(3, 2)$.

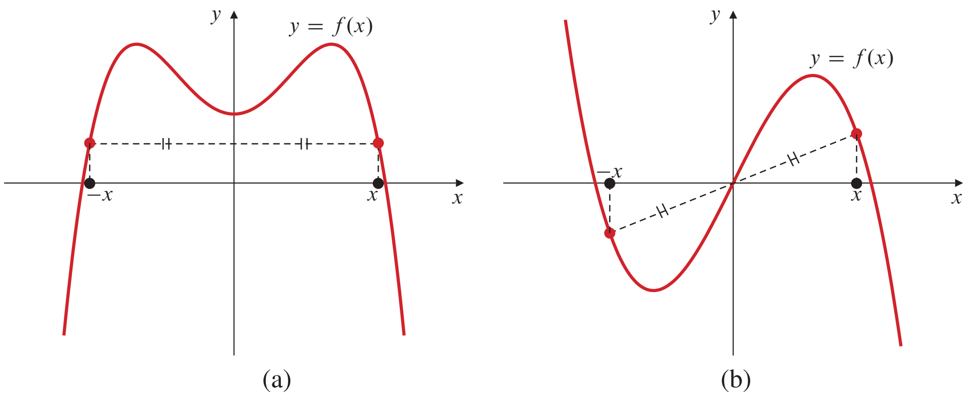


Figure 4.34

(a) The graph of an even function is symmetric about the y-axis

(b) The graph of an odd function is symmetric about the origin

Asymptotes

Some of the curves we have sketched in previous sections have had **asymptotes**, that is, straight lines to which the curve draws arbitrarily close as it recedes to infinite distance from the origin. Asymptotes are of three types: vertical, horizontal, and oblique.

DEFINITION 5

The graph of $y = f(x)$ has a **vertical asymptote** at $x = a$ if

either $\lim_{x\to a-} f(x) = \pm\infty$ or $\lim_{x\to a+} f(x) = \pm\infty$, or both.

This situation tends to arise when $f(x)$ is a quotient of two expressions and the denominator is zero at $x = a$.

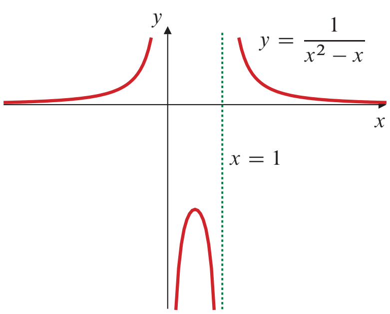


Figure 4.35

EXAMPLE 1 Find the vertical asymptotes of $f(x) = \dfrac{1}{x^2 - x}$. How does the graph approach these asymptotes?

Solution The denominator $x^2 - x = x(x - 1)$ approaches 0 as x approaches 0 or 1, so f has vertical asymptotes at $x = 0$ and $x = 1$ (Figure 4.35). Since $x(x - 1)$ is positive on $(-\infty, 0)$ and on $(1, \infty)$ and is negative on $(0, 1)$, we have

$$\lim_{x\to 0-} \frac{1}{x^2 - x} = \infty, \qquad \lim_{x\to 1-} \frac{1}{x^2 - x} = -\infty,$$

$$\lim_{x\to 0+} \frac{1}{x^2 - x} = -\infty, \qquad \lim_{x\to 1+} \frac{1}{x^2 - x} = \infty.$$

DEFINITION 6

The graph of $y = f(x)$ has a **horizontal asymptote** $y = L$ if

$$\text{either} \quad \lim_{x\to\infty} f(x) = L \quad \text{or} \quad \lim_{x\to-\infty} f(x) = L, \quad \text{or both.}$$

EXAMPLE 2 Find the horizontal asymptotes of

$$\text{(a) } f(x) = \frac{1}{x^2 - x} \quad \text{and} \quad \text{(b) } g(x) = \frac{x^4 + x^2}{x^4 + 1}.$$

Solution

(a) The function f has horizontal asymptote $y = 0$ (Figure 4.35) since

$$\lim_{x\to\pm\infty} \frac{1}{x^2 - x} = \lim_{x\to\pm\infty} \frac{1/x^2}{1 - (1/x)} = \frac{0}{1} = 0.$$

(b) The function g has horizontal asymptote $y = 1$ (Figure 4.36) since

$$\lim_{x\to\pm\infty} \frac{x^4 + x^2}{x^4 + 1} = \lim_{x\to\pm\infty} \frac{1 + (1/x^2)}{1 + (1/x^4)} = \frac{1}{1} = 1.$$

Observe that the graph of g crosses its asymptote twice. (There is a popular misconception among students that curves cannot cross their asymptotes. Exercise 41 below gives an example of a curve that crosses its asymptote infinitely often.)

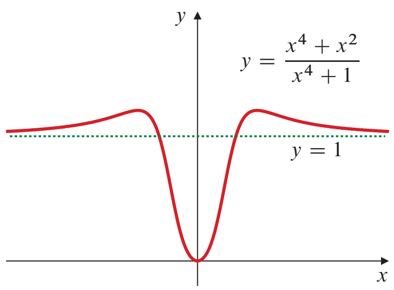


Figure 4.36

The horizontal asymptotes of both functions f and g in Example 2 are **two-sided**, which means that the graphs approach the asymptotes as x approaches both infinity and negative infinity. The function $\tan^{-1} x$ has two **one-sided** asymptotes, $y = \pi/2$ (as $x \to \infty$) and $y = -(\pi/2)$ (as $x \to -\infty$). See Figure 4.37.

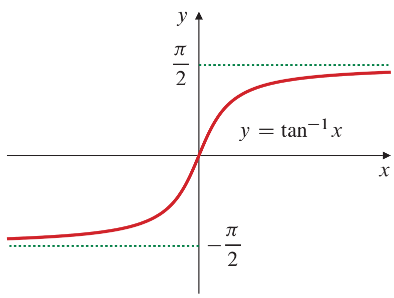


Figure 4.37 One-sided horizontal asymptotes

It can also happen that the graph of a function f approaches a nonhorizontal straight line as x approaches ∞ or $-\infty$ (or both). Such a line is called an *oblique asymptote* of the graph.

DEFINITION 7

The straight line $y = ax + b$ (where $a \neq 0$) is an **oblique asymptote** of the graph of $y = f(x)$ if

$$\text{either} \quad \lim_{x\to-\infty} \big(f(x) - (ax+b)\big) = 0 \quad \text{or} \quad \lim_{x\to\infty} \big(f(x) - (ax+b)\big) = 0,$$

or both.

EXAMPLE 3 Consider the function $f(x) = \dfrac{x^2 + 1}{x} = x + \dfrac{1}{x}$, whose graph is shown in Figure 4.38(a). The straight line $y = x$ is a *two-sided* oblique asymptote of the graph of f because

$$\lim_{x\to\pm\infty} \big(f(x) - x\big) = \lim_{x\to\pm\infty} \frac{1}{x} = 0.$$

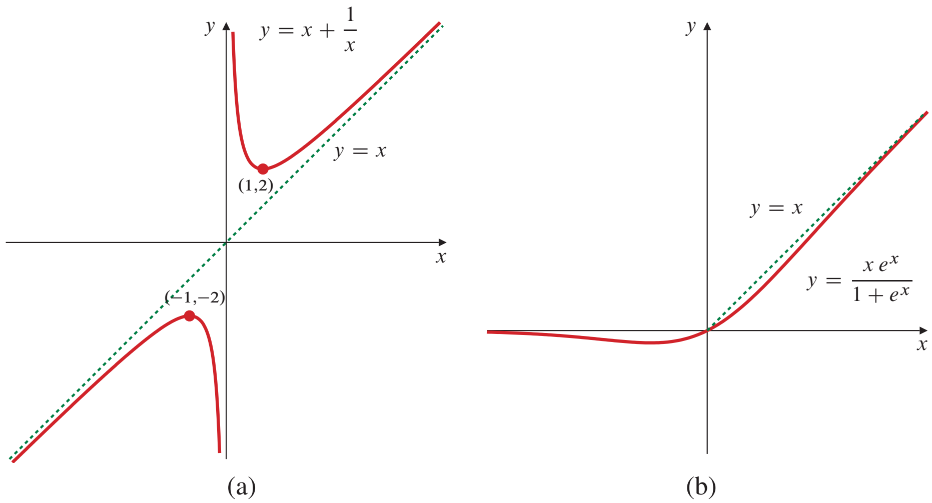


Figure 4.38

(a) The graph of $y = f(x)$ has a two-sided oblique asymptote, $y = x$

(b) This graph has a horizontal asymptote at the left and an oblique asymptote at the right

EXAMPLE 4 The graph of $y = \dfrac{x\,e^x}{1 + e^x}$, shown in Figure 4.38(b), has a horizontal asymptote $y = 0$ at the left and an oblique asymptote $y = x$ at the right:

$$\lim_{x\to-\infty} \frac{x\,e^x}{1 + e^x} = \frac{0}{1} = 0 \qquad \text{and}$$

$$\lim_{x\to\infty} \left(\frac{x\,e^x}{1 + e^x} - x \right) = \lim_{x\to\infty} \frac{x(e^x - 1 - e^x)}{1 + e^x} = \lim_{x\to\infty} \frac{-x}{1 + e^x} = 0.$$

Recall that a **rational function** is a function of the form $f(x) = P(x)/Q(x)$, where P and Q are polynomials. Following observations made in Sections P.6, 1.2, and 1.3, we can be quite specific about the asymptotes of a rational function.

Asymptotes of a rational function

Suppose that $f(x) = \dfrac{P_m(x)}{Q_n(x)}$, where P_m and Q_n are polynomials of degree m and n, respectively. Suppose also that P_m and Q_n have no common linear factors. Then

(a) The graph of f has a vertical asymptote at every position x such that $Q_n(x) = 0$.

(b) The graph of f has a two-sided horizontal asymptote $y = 0$ if $m < n$.

(c) The graph of f has a two-sided horizontal asymptote $y = L$, $(L \neq 0)$ if $m = n$. L is the quotient of the coefficients of the highest degree terms in P_m and Q_n.

(d) The graph of f has a two-sided oblique asymptote if $m = n + 1$. This asymptote can be found by dividing Q_n into P_m to obtain a linear quotient, $ax + b$, and remainder, R, a polynomial of degree at most $n - 1$. That is,

$$f(x) = ax + b + \frac{R(x)}{Q_n(x)}.$$

The oblique asymptote is $y = ax + b$.

(e) The graph of f has no horizontal or oblique asymptotes if $m > n + 1$.

EXAMPLE 5 Find the oblique asymptote of $y = \dfrac{x^3}{x^2 + x + 1}$.

Solution We can either obtain the quotient by long division:

$$
\begin{array}{r|l}
\multicolumn{1}{r}{} & x \;-\; 1 \\ \cline{2-2}
x^2+x+1 & x^3 \\
\multicolumn{1}{r}{} & x^3 + x^2 + x \\ \cline{2-2}
\multicolumn{1}{r}{} & \quad -x^2 - x \\
\multicolumn{1}{r}{} & \quad -x^2 - x - 1 \\ \cline{2-2}
\multicolumn{1}{r}{} & \hfill 1
\end{array}
\qquad\qquad
\frac{x^3}{x^2+x+1} = x - 1 + \frac{1}{x^2+x+1}
$$

or we can obtain the same result by short division:

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

In any event, we see that the oblique asymptote has equation $y = x - 1$.

Examples of Formal Curve Sketching

Here is a checklist of things to consider when you are asked to make a careful sketch of the graph of $y = f(x)$. It will, of course, not always be possible to obtain every item of information mentioned in the list.

Checklist for curve sketching

1. Calculate $f'(x)$ and $f''(x)$, and express the results in factored form.
2. Examine $f(x)$ to determine its domain and the following items:
   (a) Any vertical asymptotes. (Look for zeros of denominators.)
   (b) Any horizontal or oblique asymptotes. (Consider $\lim_{x\to\pm\infty} f(x)$.)
   (c) Any obvious symmetry. (Is f even or odd?)
   (d) Any easily calculated intercepts (points with coordinates $(x, 0)$ or $(0, y)$) or endpoints or other "obvious" points. You will add to this list when you know any critical points, singular points, and inflection points. Eventually you should make sure you know the coordinates of at least one point on every component of the graph.
3. Examine $f'(x)$ for the following:
   (a) Any critical points.
   (b) Any points where f' is not defined. (These will include singular points, endpoints of the domain of f, and vertical asymptotes.)
   (c) Intervals on which f' is positive or negative. It's a good idea to convey this information in the form of a chart such as those used in the examples. Conclusions about where f is increasing and decreasing and classification of some critical and singular points as local maxima and minima can also be indicated on the chart.
4. Examine $f''(x)$ for the following:
   (a) Points where $f''(x) = 0$.
   (b) Points where $f''(x)$ is undefined. (These will include singular points, endpoints, vertical asymptotes, and possibly other points as well, where f' is defined but f'' isn't.)
   (c) Intervals where f'' is positive or negative and where f is therefore concave up or down. Use a chart.
   (d) Any inflection points.

When you have obtained as much of this information as possible, make a careful sketch that reflects *everything* you have learned about the function. Consider where best to place the axes and what scale to use on each so the "interesting features" of the graph show up most clearly. Be alert for seeming inconsistencies in the information—that is a strong suggestion you may have made an error somewhere. For example, if you have determined that $f(x) \to \infty$ as x approaches the vertical asymptote $x = a$ from the right, and also that f is decreasing and concave down on the interval (a, b), then you have very likely made an error. (Try to sketch such a situation to see why.)

EXAMPLE 6 Sketch the graph of $y = \dfrac{x^2 + 2x + 4}{2x}$.

Solution It is useful to rewrite the function y in the form

$$y = \frac{x}{2} + 1 + \frac{2}{x},$$

since this form not only shows clearly that $y = (x/2) + 1$ is an oblique asymptote, but also makes it easier to calculate the derivatives

$$y' = \frac{1}{2} - \frac{2}{x^2} = \frac{x^2 - 4}{2x^2}, \qquad y'' = \frac{4}{x^3}.$$

From y: Domain: all x except 0. Vertical asymptote: $x = 0$,
Oblique asymptote: $y = \dfrac{x}{2} + 1$, $\quad y - \left(\dfrac{x}{2} + 1\right) = \dfrac{2}{x} \to 0$ as $x \to \pm\infty$.
Symmetry: none obvious (y is neither odd nor even).
Intercepts: none. $x^2 + 2x + 4 = (x + 1)^2 + 3 \ge 3$ for all x, and y is not defined at $x = 0$.

From y': Critical points: $x = \pm 2$; points $(-2, -1)$ and $(2, 3)$.
y' not defined at $x = 0$ (vertical asymptote).

From y'': $y'' = 0$ nowhere; y'' undefined at $x = 0$.

		CP		ASY		CP	
x		-2		0		2	
y'	$+$	0	$-$	undef	$-$	0	$+$
y''	$-$		$-$	undef	$+$		$+$
y	$\nearrow$	max	$\searrow$	undef	$\searrow$	min	$\nearrow$
	$\frown$		$\frown$		$\smile$		$\smile$

The graph is shown in Figure 4.39.

EXAMPLE 7 Sketch the graph of $f(x) = \dfrac{x^2 - 1}{x^2 - 4}$.

Solution We have

$$f'(x) = \frac{-6x}{(x^2 - 4)^2}, \qquad f''(x) = \frac{6(3x^2 + 4)}{(x^2 - 4)^3}.$$

From f: Domain: all x except ± 2. Vertical asymptotes: $x = -2$ and $x = 2$.
Horizontal asymptote: $y = 1$ (as $x \to \pm\infty$).
Symmetry: about the y-axis (y is even).
Intercepts: $(0, 1/4)$, $(-1, 0)$, and $(1, 0)$.
Other points: $(-3, 8/5)$, $(3, 8/5)$. (The two vertical asymptotes divide the graph into three components; we need points on each. The outer components require points with $|x| > 2$.)

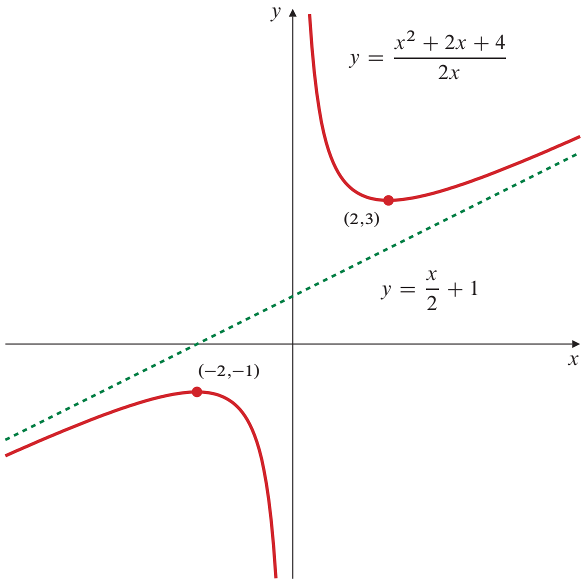


Figure 4.39

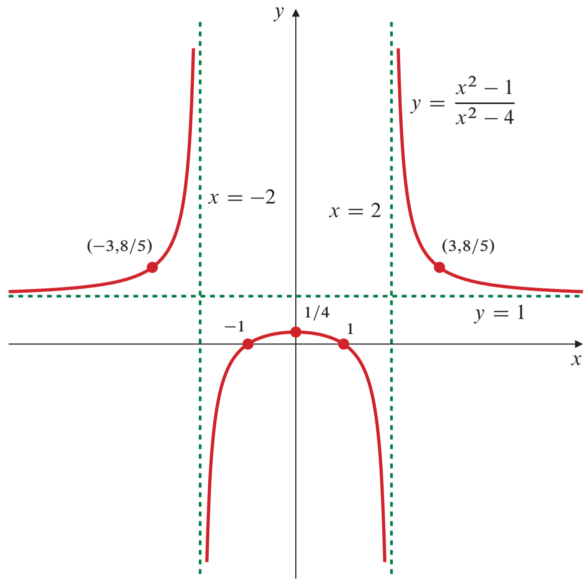


Figure 4.40

From f': Critical point: $x = 0$; f' not defined at $x = 2$ or $x = -2$.

From f'': $f''(x) = 0$ nowhere; f'' not defined at $x = 2$ or $x = -2$.

		ASY		CP		ASY	
x		-2		0		2	
f'	$+$	undef	$+$	0	$-$	undef	$-$
f''	$+$	undef	$-$		$-$	undef	$+$
f	$\nearrow$ $\smile$	undef	$\nearrow$ $\frown$	max	$\searrow$ $\frown$	undef	$\searrow$ $\smile$

The graph is shown in Figure 4.40.

EXAMPLE 8 Sketch the graph of $y = xe^{-x^2/2}$.

Solution We have $y' = (1 - x^2)e^{-x^2/2}$, $\quad y'' = x(x^2 - 3)e^{-x^2/2}$.

From y: Domain: all x.
Horizontal asymptote: $y = 0$. Note that if $t = x^2/2$, then $|xe^{-x^2/2}| = \sqrt{2t}\, e^{-t} \to 0$ as $t \to \infty$ (hence as $x \to \pm\infty$).
Symmetry: about the origin (y is odd). Intercepts: $(0, 0)$.

From y': Critical points: $x = \pm 1$; points $(\pm 1, \pm 1/\sqrt{e}) \approx (\pm 1, \pm 0.61)$.

From y'': $y'' = 0$ at $x = 0$ and $x = \pm\sqrt{3}$;
points $(0, 0)$, $(\pm\sqrt{3}, \pm\sqrt{3}e^{-3/2}) \approx (\pm 1.73, \pm 0.39)$.

				CP				CP			
x		$-\sqrt{3}$		-1		0		1		$\sqrt{3}$	
y'	$-$		$-$	0	$+$		$+$	0	$-$		$-$
y''	$-$	0	$+$		$+$	0	$-$		$-$	0	$+$
y	$\searrow$		$\searrow$	min	$\nearrow$		$\nearrow$	max	$\searrow$		$\searrow$
	$\frown$	infl	$\smile$		$\smile$	infl	$\frown$		$\frown$	infl	$\smile$

The graph is shown in Figure 4.41.

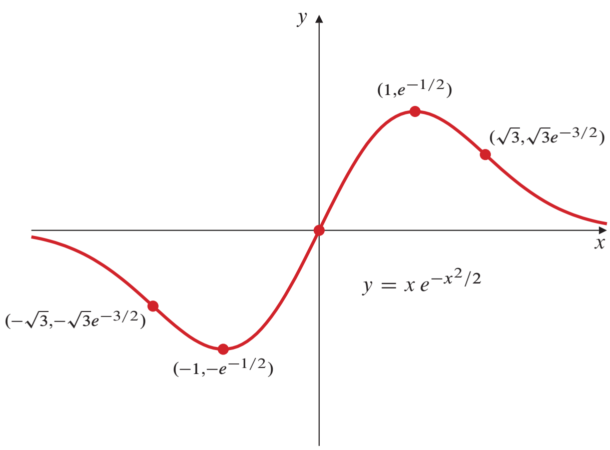


Figure 4.41

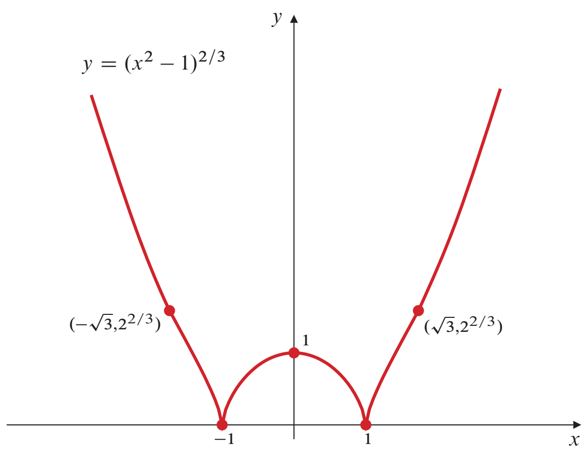


Figure 4.42

EXAMPLE 9 Sketch the graph of $f(x) = (x^2 - 1)^{2/3}$. (See Figure 4.42.)

Solution $f'(x) = \dfrac{4}{3}\,\dfrac{x}{(x^2 - 1)^{1/3}}$, $\qquad f''(x) = \dfrac{4}{9}\,\dfrac{x^2 - 3}{(x^2 - 1)^{4/3}}$.

From f: Domain: all x.
Asymptotes: none. ($f(x)$ grows like $x^{4/3}$ as $x \to \pm\infty$.)
Symmetry: about the y-axis (f is an even function).
Intercepts: $(\pm 1, 0)$, $(0, 1)$.

From f': Critical points: $x = 0$; singular points: $x = \pm 1$.

From f'': $f''(x) = 0$ at $x = \pm\sqrt{3}$; points $(\pm\sqrt{3}, 2^{2/3}) \approx (\pm 1.73, 1.59)$;
$f''(x)$ not defined at $x = \pm 1$.

				SP		CP		SP			
x		$-\sqrt{3}$		-1		0		1		$\sqrt{3}$	
f'	$-$		$-$	undef	$+$	0	$-$	undef	$+$		$+$
f''	$+$	0	$-$	undef	$-$		$-$	undef	$-$	0	$+$
f	$\searrow$		$\searrow$	min	$\nearrow$	max	$\searrow$	min	$\nearrow$		$\nearrow$
	$\smile$	infl	$\frown$		$\frown$		$\frown$		$\frown$	infl	$\smile$

EXERCISES 4.6

1. Figure 4.43 shows the graphs of a function f, its two derivatives f' and f'', and another function g. Which graph corresponds to each function?

2. List, for each function graphed in Figure 4.43, such information that you can determine (approximately) by inspecting the graph (e.g., symmetry, asymptotes, intercepts, intervals of increase and decrease, critical and singular points, local maxima and minima, intervals of constant concavity, inflection points).

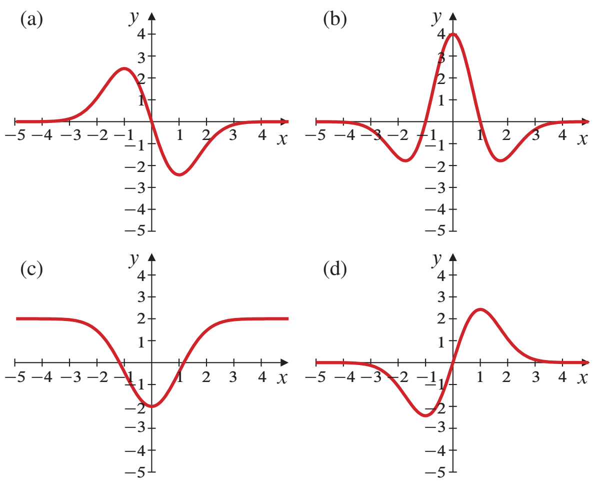


Figure 4.43

3. Figure 4.44 shows the graphs of four functions:

$$f(x) = \frac{x}{1-x^2}, \qquad g(x) = \frac{x^3}{1-x^4},$$

$$h(x) = \frac{x^3 - x}{\sqrt{x^6+1}}, \qquad k(x) = \frac{x^3}{\sqrt{|x^4-1|}}.$$

Which graph corresponds to each function?

4. Repeat Exercise 2 for the graphs in Figure 4.44.

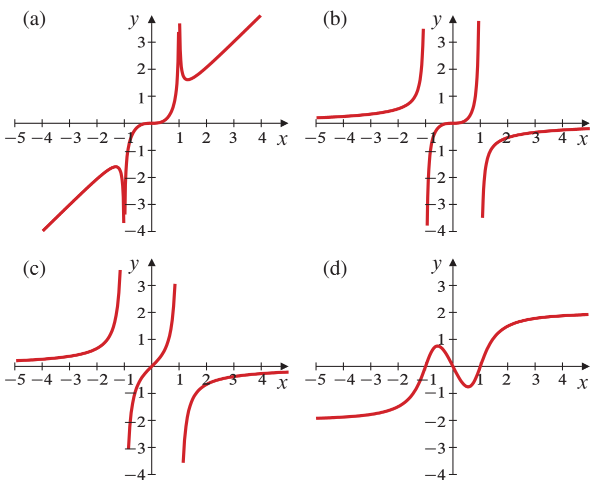


Figure 4.44

In Exercises 5–6, sketch the graph of a function that has the given properties. Identify any critical points, singular points, local maxima and minima, and inflection points. Assume that f is continuous and its derivatives exist everywhere unless the contrary is implied or explicitly stated.

5. $f(0) = 1$, $f(\pm 1) = 0$, $f(2) = 1$, $\lim_{x\to\infty} f(x) = 2$, $\lim_{x\to-\infty} f(x) = -1$, $f'(x) > 0$ on $(-\infty, 0)$ and on $(1, \infty)$, $f'(x) < 0$ on $(0, 1)$, $f''(x) > 0$ on $(-\infty, 0)$ and on $(0, 2)$, and $f''(x) < 0$ on $(2, \infty)$.

6. $f(-1) = 0$, $f(0) = 2$, $f(1) = 1$, $f(2) = 0$, $f(3) = 1$, $\lim_{x\to\pm\infty}(f(x) + 1 - x) = 0$, $f'(x) > 0$ on $(-\infty, -1)$, $(-1, 0)$ and $(2, \infty)$, $f'(x) < 0$ on $(0, 2)$, $\lim_{x\to-1} f'(x) = \infty$, $f''(x) > 0$ on $(-\infty, -1)$ and on $(1, 3)$, and $f''(x) < 0$ on $(-1, 1)$ and on $(3, \infty)$.

In Exercises 7–39, sketch the graphs of the given functions, making use of any suitable information you can obtain from the function and its first and second derivatives.

7. $y = (x^2 - 1)^3$

8. $y = x(x^2 - 1)^2$

9. $y = \dfrac{2-x}{x}$

10. $y = \dfrac{x-1}{x+1}$

11. $y = \dfrac{x^3}{1+x}$

12. $y = \dfrac{1}{4+x^2}$

13. $y = \dfrac{1}{2-x^2}$

14. $y = \dfrac{x}{x^2-1}$

15. $y = \dfrac{x^2}{x^2-1}$

16. $y = \dfrac{x^3}{x^2-1}$

17. $y = \dfrac{x^3}{x^2+1}$

18. $y = \dfrac{x^2}{x^2+1}$

19. $y = \dfrac{x^2-4}{x+1}$

20. $y = \dfrac{x^2-2}{x^2-1}$

21. $y = \dfrac{x^3-4x}{x^2-1}$

22. $y = \dfrac{x^2-1}{x^2}$

23. $y = \dfrac{x^5}{(x^2-1)^2}$

24. $y = \dfrac{(2-x)^2}{x^3}$

25. $y = \dfrac{1}{x^3-4x}$

26. $y = \dfrac{x}{x^2+x-2}$

27. $y = \dfrac{x^3-3x^2+1}{x^3}$

28. $y = x + \sin x$

29. $y = x + 2\sin x$

30. $y = e^{-x^2}$

31. $y = xe^x$

32. $y = e^{-x}\sin x, \ (x \ge 0)$

33. $y = x^2 e^{-x^2}$

34. $y = x^2 e^x$

35. $y = \dfrac{\ln x}{x}, \ (x > 0)$

36. $y = \dfrac{\ln x}{x^2}, \ (x > 0)$

37. $y = \dfrac{1}{\sqrt{4-x^2}}$

38. $y = \dfrac{x}{\sqrt{x^2+1}}$

39. $y = (x^2 - 1)^{1/3}$

! 40. What is $\lim_{x\to 0+} x\ln x$? $\lim_{x\to 0} x\ln|x|$? If $f(x) = x\ln|x|$ for $x \ne 0$, is it possible to define $f(0)$ in such a way that f is continuous on the whole real line? Sketch the graph of f.

41. What straight line is an asymptote of the curve $y = \dfrac{\sin x}{1+x^2}$? At what points does the curve cross this asymptote?

4.7 Graphing with Computers

The techniques for sketching, developed in the previous section, are useful for graphs of functions that are simple enough to allow you to calculate and analyze their derivatives. They are also essential for testing the validity of graphs produced by computers or calculators, which can be inaccurate or misleading for a variety of reasons, including the case of numerical monsters introduced in previous chapters. In practice, it is often easiest to first produce a graph using a computer or graphing calculator, but many times this will not turn out to be the last step. (We will use the term "computer" for both computers and calculators.) For many simple functions this can be a quick and painless activity, but sometimes functions have properties that complicate the process. Knowledge of the function, from techniques like those above, is important to guide you on what the next steps must be.

The Maple command[1] for viewing the graph of the function from Example 6 of Section 4.6, together with its oblique asymptote, is a straightforward example of plotting; we ask Maple to plot both $(x^2 + 2x + 4)/(2x)$ and $1 + (x/2)$:

```
> plot({(x^2+2*x+4)/(2*x), 1+(x/2)}, x=-6..6, y=-7..7);
```

This command sets the window $-6 \le x \le 6$ and $-7 \le y \le 7$. Why that window? To get a plot that characterizes the function, knowledge of its vertical asymptote at $x = 0$ is essential. (If $x - 10$ were substituted for x in the expression, the given window would no longer produce a reasonable graph of the key features of the function. The new function would be better viewed on the interval $4 \le x \le 16$.) If the range $[-7, 7]$ were not specified, the computer would plot all of the points where it evaluates the function, including those very close to the vertical asymptote where the function is very large. The resulting plot would compress all of the features of the graph onto the x-axis. Even the asymptote would look like a horizontal line in that scaling. You might even miss the vertical asymptote, which is squeezed into the y-axis.

Getting Maple to plot the curve in Example 9 of Section 4.6 is a bit trickier. Because Maple doesn't deal well with fractional powers of negative numbers, even when they have positive real values, we must actually plot $|x^2 - 1|^{2/3}$ or $((x^2 - 1)^2)^{1/3}$. Otherwise, the part of the graph between -1 and 1 will be missing. Either of the plot commands

```
> plot((abs(x^2-1))^(2/3), x=-4..4, y=-1..5);
> plot(((x^2-1)^2)^(1/3), x=-4..4, y=-1..5);
```

will produce the desired graph. In order to ensure a complete plot with all of the features of the function present, the graph of the simple expression should be viewed critically, and not taken at face value.

Numerical Monsters and Computer Graphing

The next obvious problem is that of false features and false behaviours. Functions that are mathematically well-behaved can still be computationally poorly behaved, leading to false features on graphs, as we have already seen.

EXAMPLE 1 Consider the function $f(x) = e^x \ln(1 + e^{-x})$, which has suitably simplified derivative

$$f'(x) = e^x\, g(x), \qquad \text{where} \qquad g(x) = \ln(1 + e^{-x}) - \frac{1}{e^x + 1}.$$

In turn, the derivative of $g(x)$ simplifies to

$$g'(x) = -\frac{1}{(e^x + 1)^2},$$

[1] Although we focus on Maple to illustrate the issues of graphing with computers, the issues presented are general ones, pertaining to all software and computers.

which is negative for all x, so g is decreasing. Since $g(0) = \ln 2 - 1/2 > 0$ and $\lim_{x\to\infty} g(x) = 0$, it follows that $g(x) > 0$ and decreasing for all x. Thus, $f'(x)$ is positive, and $f(x)$ is an increasing function for all x. Furthermore, l'Hôpital's Rules show that

$$\lim_{x\to\infty} f(x) = 1 \qquad \textit{and} \qquad \lim_{x\to-\infty} f(x) = 0.$$

This gives us a pretty full picture of how the function f behaves. It grows with increasing x from 0 at $-\infty$, crosses the y-axis at $\ln 2$, and finally approaches 1 asymptotically from below as x increases toward ∞.

Now let's plot the graph of f using the Maple command

```
> plot(exp(x)*ln(1+(1/exp(x))), x=-20..45, style=point,
 symbol=point, numpoints=1500);
```

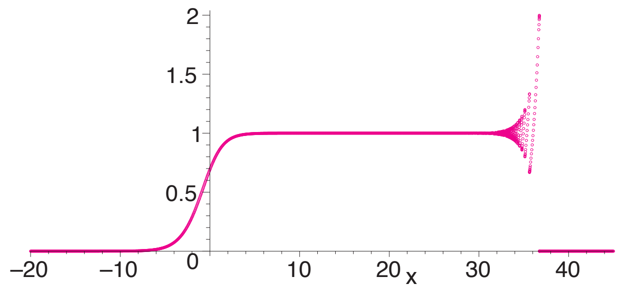


Figure 4.45 A faulty computer plot of $y = e^x \ln(1 + e^{-x})$

The result is shown in Figure 4.45. Clearly something is wrong. From $x = -20$ to about $x = 30$, the graph behaves in accordance with the mathematical analysis. However, for larger values of x, peculiarities emerge that sharply disagree with the analysis. The calculus of this chapter tells us that the function is increasing with no horizontal tangents, but the computer suggests that it decreases in some places. The calculus tells us that the function rises asymptotically to 1, but the computer suggests that the function starts to oscillate and ultimately becomes 0 at about $x = 36$.

This is another numerical monster. What a computer does can simply be wrong. In this case, it is significantly so. In practical applications an erroneous value of 0 instead of 1 could, for example, be a factor in a product, and that would change everything dramatically. If the mathematics were not known in this case, how could we even know that the computer is wrong? Another computer cannot be used to check it, as the problem is one that all computers share. Another program cannot be used because all software must use the special floating-point arithmetic that is subject to the roundoff errors responsible for the problem. Figure 4.45 is not particular to Maple. This monster, or one much like it, can be created in nearly any software package.

Floating-Point Representation of Numbers in Computers

It is necessary that you know mathematics in order to use computers correctly and effectively. It is equally necessary to understand why *all* computers fail to fully capture the mathematics. As indicated previously, the reason is that no computer can represent all numbers. Computer designers artfully attempt to minimize the effects of this by making the number of representable numbers as large as possible. But, speaking in terms of physics, a finite-sized machine can only represent a finite number of numbers. Having only a finite number of numbers leads to numbers sufficiently small, compared to 1, that the computer simply discards them in a sum. When digits are lost in this manner, the resulting error is known as **roundoff error**.

In many cases the finiteness shows up in the use of floating-point numbers and a set of corresponding arithmetic rules that approximate correct arithmetic. These approximate rules and approximate representations are not unique by any means. For example, the software package *Derive* uses so-called **slash arithmetic**, which works with a representation of numbers as continued fractions instead of decimals. This has certain advantages and disadvantages, but, in the end, finiteness forces truncation just the same.

The term "roundoff" implies that there is some kind of mitigation procedure or **rounding** done to reduce error once the smallest digits have been discarded. There are a number of different kinds of rounding practices. The various options can be quite intricate, but they all begin with the aim to slightly reduce error as a result of truncation. The truncation is the source of error, not the rounding, despite the terminology that seems to suggest otherwise. The entire process of truncation and rounding have come to be termed "roundoff," although the details of the error mitigation are immaterial for the purposes of this discussion. Rounding is beyond the scope of this section and will not be considered further.

Historically, the term "decimal" implies base ten, but the idea works the same in any base. In particular, in any base, multiplying by the base to an integral power simply shifts the position of the "decimal point." Thus, multiplying or dividing by the base is known as a **shift operation**. The term "floating-point" signifies this shifting of the point to the left or right. The general technical term for the decimal point is **radix point**. Specifically for base two, the point is sometimes called the **binary point**. However, we will use the term **decimal point** or just **decimal** for all bases, as the etymological purity is not worth having several names for one small symbol.

While computers, for the most part, work in base two, they can be and have been built in other bases. For example, there have been base-three computers, and many computers group numbers so that they work as if they were built in base eight (**octal**) or base sixteen (**hexadecimal**). (If you are feeling old, quote your age in hexadecimal. For example, $48 = 3 \times 16$ or 30 in hexadecimal. If you are feeling too young, use octal.)

In a normal binary computer, floating-point numbers approximate the mathematical *real numbers*. Several **bytes** of memory (frequently 8 bytes) are allocated for each floating-point number. Each byte consists of eight **bits**, each of which has two (physical) states and can thus store one of the two base-two digits "0" or "1," as it is the equivalent of a switch being either *off* or *on*.

Thus, an eight-byte allocation for a floating-point number can store 64 bits of data. The computer uses something similar to scientific notation, which is often used to express numbers in base ten. However, the convention is to place the decimal immediately to the left of all significant figures. For example, the computer convention would call for the base-ten number 284,070,000 to be represented as 0.28407×10^9. Here 0.28407 is called the **mantissa**, and it has 5 significant base-ten digits following the decimal point, the 2 being the most significant and the 7 the least significant digit. The 9 in the factor 10^9 is called the **exponent**, which defines the number of shift operations needed to locate the correct position of the decimal point of the actual number.

The computer only needs to represent the mantissa and the exponent, each with its appropriate sign. The base is set by the architecture and so is not stored. Neither is the decimal point nor the leading zero in the mantissa stored. These are all just implied. If the floating-point number has 64 bits, two are used for the two signs, leaving 62 bits for significant digits in the mantissa and the exponent.

As an example of base two (i.e., binary) representation, the number

$$101.011 = 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 + 0 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3}$$

stands for the base-ten number 4 + 1 + (1/4) + (1/8) = 43/8. On a computer the stored bits would be +101011 for the mantissa and +11 for the exponent. Thus, the base-two floating-point form is 0.101011×2^3, with mantissa 0.101011 and exponent 3. Note that we are representing the exponent in base ten (3), and not base two (11), because that is more convenient for counting shift operations.

While the base-two representation of two is 10, we will continue, for convenience, to write two as 2 when using it as the base for base-two representations. After all, any base b is represented by 10 with respect to itself as base. So, if we chose to write the number above as 0.101011×10^{11}, the numeral could as well denote a number in any base. However, for us people normally thinking in base ten, 0.101011×2^3 clearly indicates that the base is two and the decimal point is shifted 3 digits to the right of the most significant digit in the mantissa.

Now consider $x = 0.101 \times 2^{-10} = 0.0000000000101$, the base-two floating-point number whose value as a base-ten fraction is $x = 5/8192$. The only significant base-two digits are the 101 in the mantissa. Now add x to 1; the result is

$$1 + x = 0.10000000000101 \times 2^1,$$

which has mantissa 0.10000000000101 and exponent 1. The mantissa now has 14 significant base-two digits; all the zeros between the first and last 1s are significant. If your computer or calculator software only allocates, say, 12 bits for mantissas, then it would be unable to represent $1 + x$. It would have to throw away the two least significant base-two digits and save the number as

$$1 + x = 0.100000000001 \times 2^1 = \frac{2{,}049}{2{,}048},$$

thus creating a roundoff error of 1/8,192. Even worse, if only ten base-two digits were used to store mantissas, the computer would store $1 + x = 0.1000000000 \times 2^1$ (i.e., it would not be able to distinguish $1 + x$ from 1). Of course, calculators and computer software use many more than ten or twelve base-two digits to represent mantissas of floating-point numbers, but the number of digits used is certainly finite, and so the problem of roundoff will always occur for sufficiently small floating-point numbers x.

Machine Epsilon and Its Effect on Figure 4.45

The smallest number x for which the computer recognizes that $1 + x$ is greater than 1 is called **machine epsilon** (denoted ϵ) for that computer. The computer does not return 1 when evaluating $1 + \epsilon$, but for all positive numbers x smaller than ϵ, the computer simply returns 1 when asked to evaluate $1 + x$, because the computer only keeps a finite number of (normally base two) digits.

When using computer algebra packages like Maple, the number of digits can be increased in the software. Thus, the number of numbers that the computer can represent can be extended beyond what is native to the processor's hardware, by stringing together bits to make available larger numbers of digits for a single number. The Maple command for this is "`Digits`," which defaults to 10 (decimal digits). However, the computer remains finite in size, so there will always be an effective value for ϵ, no matter how the software is set. A hardware value for ϵ is not uniform for all devices either. Thus, for any device you may be using (calculator or computer), the value of machine epsilon may not be immediately obvious. To anticipate where a computer may be wrong, you need the value of machine epsilon, and you need to understand where the function may run afoul of it. We will outline a simple way to determine this below.

In the case of the function f in Example 1, it is clear where the computer discards digits in a sum. The factor $\ln(1 + e^{-x})$ decreases as x increases, but for sufficiently large x a computer must discard the exponential in the sum because it is too small to show up in the digits allotted for 1. When the exponential term decreases below the value of ϵ, the computer will return 1 for the argument of the natural logarithm, and the factor will be determined by the computer to be 0. Thus, f will be represented as 0 instead of nearly 1.

Of course, pathological behaviour begins to happen before the exponential e^{-x} decreases to below ϵ. When the exponential is small enough, all change with x happens in the smaller digits. The sum forces them to be discarded by the computer, so the change is discarded with it. That means for finite intervals the larger digits from the decreasing exponential term do not change until the smaller changes accrue. In the case of f, this means it behaves like an increasing exponential times a constant between corrections of the larger digits. This is confirmed in Figure 4.46, which is a close-up of the pathological region given by adjusting the interval of the plot command.

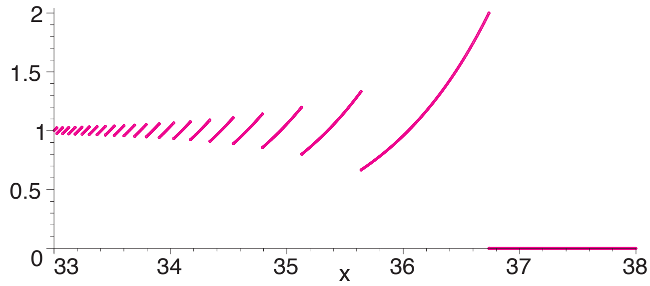


Figure 4.46 Part of the graph of f from Example 1 over the interval $[33, 38]$

Determining Machine Epsilon

A small alteration in the function f of Example 1 provides an easy way to determine the value of machine epsilon. As computers store and process data in base-two form, it is useful to use instead of f the function $h(x) = 2^x \ln(1 + 2^{-x})$. The Maple plot command

```
> plot(2^x*ln(1+1/2^x), x=50..55, style=line,
   thickness=5, xtickmarks=[50,51,52,53,54]);
```

produces the graph in Figure 4.47. The graph drops to 0 at $x = 53$. Thus, 2^{-53} is the next number below ϵ that the computer can represent. Because the first nonzero digit in a base-two number is 1, the next largest number must be up to twice as large. But because all higher digits are discarded, the effect is to have simply a change in the exponent of the number, a shift operation. A single shift operation larger than 2^{-53} is 2^{-52}, so $\epsilon = 2^{-52}$ in the settings for this plot.

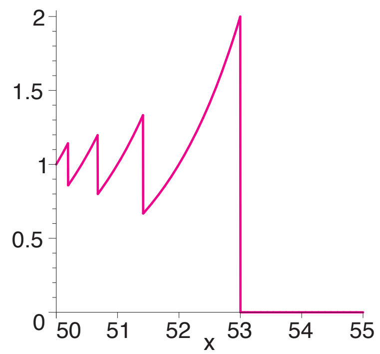


Figure 4.47 This indicates that machine epsilon is $\epsilon = 2^{-52}$

From this we can predict when f will drop to zero in Figure 4.45 and Figure 4.46. It will be when $\epsilon/2 = 2^{-53} = e^{-x}$, or approximately $x = 36.74$. While this seems to give us a complete command of the effect for most computers, there is much more going on with computer error that depends on specific algorithms. While significant error erupts when ϵ is reached in a sum with 1, other sources of error are in play well before that for smaller values of x.

It is interesting to look at some of the complex and structured patterns of error in a close-up of what should be a single curve well before the catastrophic drop to zero. Figure 4.48 is produced by the plot instruction

```
> plot(exp(x)*ln(1+1/exp(x)), x = 29.5 .. 30,
   style = point, symbol = point, numpoints = 3000);
```

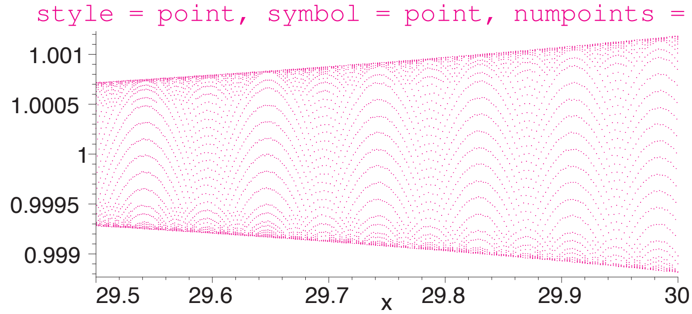


Figure 4.48 Illusions of computation

In this figure there are many fascinating and beautiful patterns created, which are completely spurious. In this region the exponential curves are collapsed together, forming what seems like a single region contained within an expanding envelope. The beautiful patterns make it easy to forget that the mathematically correct curve would appear as a single horizontal line at height 1. The patterns here are created by Maple's selection of points at which to evaluate the function and their placement in the plot. If you change the plot window, try to zoom in on them, or change the numbers of points or the interval; they will change too, or disappear. They are completely illusive and spurious features. Computers can't be trusted blindly. You can trust mathematics.

EXERCISES 4.7

1. Use Maple to get a plot instruction that plots an exponential function through one of the stripes in Figure 4.46. You can use the cursor position in the Maple display to read off the approximate coordinates of the lower left endpoint on one of the stripes.

2. Why should the expression $h(x) - \sqrt{h(x)^2}$ not be expected to be exactly zero, especially for large $h(x)$, when evaluated on a computer?

3. Consider Figure 4.49. It is the result of the plot instruction

```
> plot([ln(2^x-sqrt(2^(2*x)-1)),
-ln(2^x+sqrt(2^(2*x)-1))], x=0..50,
 y=-30..10, style=line, symbol=point,
thickness=[1,4],
color=[magenta, grey], numpoints=8000);
```

The grey line is a plot of $f(x) = -\ln(2^x + \sqrt{2^{2x} - 1})$. The coloured line is a plot of $g(x) = \ln(2^x - \sqrt{2^{2x} - 1})$.

(a) Show that $g(x) = f(x)$.

(b) Why do the graphs of f and g behave differently?

(c) Estimate a value of x beyond which the plots of f and g will behave differently. Assume machine epsilon is $\epsilon = 2^{-52}$.

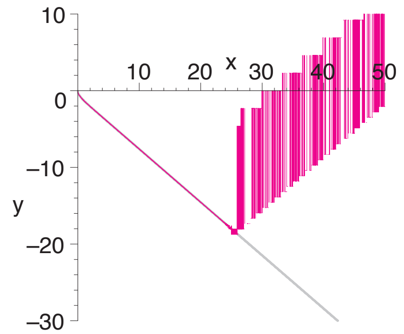


Figure 4.49

4. If you use a graphing calculator or other mathematical graphing software, try to determine machine epsilon for it.

In Exercises 5–6 assume that a computer uses 64 bits (binary digits) of memory to store a floating-point number, and that of these 64 bits 52 are used for the mantissa and one each for the signs of the mantissa and the exponent.

5. To the nearest power of 10, what is the smallest positive number that can be represented in floating-point form by the computer?

6. To the nearest power of 10, what is the largest positive number that can be represented in floating-point form by the computer?

4.8 Extreme-Value Problems

In this section we solve various word problems that, when translated into mathematical terms, require the finding of a maximum or minimum value of a function of one variable. Such problems can range from simple to very complex and difficult; they can be phrased in terminology appropriate to some other discipline, or they can be already partially translated into a more mathematical context. We have already encountered a few such problems in earlier chapters.

Let us consider a couple of examples before attempting to formulate any general principles for dealing with such problems.

EXAMPLE 1 A rectangular animal enclosure is to be constructed having one side along an existing long wall and the other three sides fenced. If 100 m of fence are available, what is the largest possible area for the enclosure?

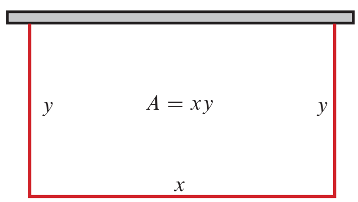


Figure 4.50

Solution This problem, like many others, is essentially a geometric one. A sketch should be made at the outset, as we have done in Figure 4.50. Let the length and width of the enclosure be x and y m, respectively, and let its area be A m^2. Thus $A = xy$. Since the total length of the fence is 100 m, we must have $x + 2y = 100$. A appears to be a function of two variables, x and y, but these variables are not independent; they are related by the *constraint* $x + 2y = 100$. This constraint equation can be solved for one variable in terms of the other, and A can therefore be written as a function of only one variable:

$$x = 100 - 2y,$$
$$A = A(y) = (100 - 2y)y = 100y - 2y^2.$$

Evidently, we require $y \geq 0$ and $y \leq 50$ (i.e., $x \geq 0$) in order that the area make sense. (It would otherwise be negative.) Thus, we must maximize the function $A(y)$ on the interval $[0, 50]$. Being continuous on this closed, finite interval, A must have a maximum value, by Theorem 5. Clearly, $A(0) = A(50) = 0$ and $A(y) > 0$ for $0 < y < 50$. Hence, the maximum cannot occur at an endpoint. Since A has no singular points, the maximum must occur at a critical point. To find any critical points, we set

$$0 = A'(y) = 100 - 4y.$$

Therefore, $y = 25$. Since A must have a maximum value and there is only one possible point where it can be, the maximum must occur at $y = 25$. The greatest possible area for the enclosure is therefore $A(25) = 1{,}250$ m^2.

EXAMPLE 2 A lighthouse L is located on a small island 5 km north of a point A on a straight east-west shoreline. A cable is to be laid from L to point B on the shoreline 10 km east of A. The cable will be laid through the water in a straight line from L to a point C on the shoreline between A and B, and from there to B along the shoreline. (See Figure 4.51.) The part of the cable lying in the water costs \$5,000/km, and the part along the shoreline costs \$3,000/km.

(a) Where should C be chosen to minimize the total cost of the cable?

(b) Where should C be chosen if B is only 3 km from A?

Solution

(a) Let C be x km from A toward B. Thus $0 \leq x \leq 10$. The length of LC is $\sqrt{25 + x^2}$ km, and the length of CB is $10 - x$ km, as illustrated in Figure 4.51. Hence, the total cost of the cable is \$$T$, where

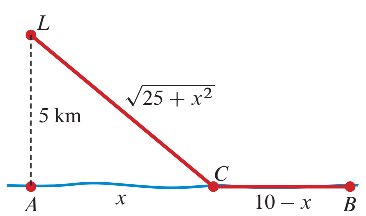


Figure 4.51

$$T = T(x) = 5{,}000\sqrt{25 + x^2} + 3{,}000(10 - x), \qquad (0 \leq x \leq 10).$$

T is continuous on the closed, finite interval $[0, 10]$, so it has a minimum value that may occur at one of the endpoints $x = 0$ or $x = 10$ or at a critical point in the interval $(0, 10)$. (T has no singular points.) To find any critical points, we set

$$0 = \frac{dT}{dx} = \frac{5{,}000x}{\sqrt{25 + x^2}} - 3{,}000.$$

$$\begin{aligned} \text{Thus,} \qquad 5{,}000x &= 3{,}000\sqrt{25 + x^2} \\ 25x^2 &= 9(25 + x^2) \\ 16x^2 &= 225 \\ x^2 &= \frac{225}{16} = \frac{15^2}{4^2}. \end{aligned}$$

This equation has two solutions, but only one, $x = 15/4 = 3.75$, lies in the interval $(0, 10)$. Since $T(0) = 55{,}000$, $T(15/4) = 50{,}000$, and $T(10) \approx 55{,}902$, the critical point 3.75 evidently provides the minimum value for $T(x)$. For minimal cost, C should be 3.75 km from A.

(b) If B is 3 km from A, the corresponding total cost function is

$$T(x) = 5{,}000\sqrt{25 + x^2} + 3{,}000(3 - x), \qquad (0 \le x \le 3),$$

which differs from the total cost function $T(x)$ of part (a) only in the added constant (9,000 rather than 30,000). It therefore has the same critical point, $x = 15/4 = 3.75$, which does not lie in the interval $(0, 3)$. Since $T(0) = 34{,}000$ and $T(3) \approx 29{,}155$, in this case we should choose $x = 3$. To minimize the total cost, the cable should go straight from L to B.

Procedure for Solving Extreme-Value Problems

Based on our experience with the examples above, we can formulate a checklist of steps involved in solving optimization problems.

Solving extreme-value problems

1. Read the problem very carefully, perhaps more than once. You must understand clearly what is given and what must be found.
2. Make a diagram if appropriate. Many problems have a geometric component, and a good diagram can often be an essential part of the solution process.
3. Define any symbols you wish to use that are not already specified in the statement of the problem.
4. Express the quantity Q to be maximized or minimized as a function of one or more variables.
5. If Q depends on n variables, where $n > 1$, find $n - 1$ equations (constraints) linking these variables. (If this cannot be done, the problem cannot be solved by single-variable techniques.)
6. Use the constraints to eliminate variables and hence express Q as a function of only one variable. Determine the interval(s) in which this variable must lie for the problem to make sense. Alternatively, regard the constraints as implicitly defining $n - 1$ of the variables, and hence Q, as functions of the remaining variable.
7. Find the required extreme value of the function Q using the techniques of Section 4.4. Remember to consider any critical points, singular points, and endpoints. Make sure to give a convincing argument that your extreme value is the one being sought; for example, if you are looking for a maximum, the value you have found should not be a minimum.
8. Make a concluding statement answering the question asked. Is your answer for the question *reasonable*? If not, check back through the solution to see what went wrong.

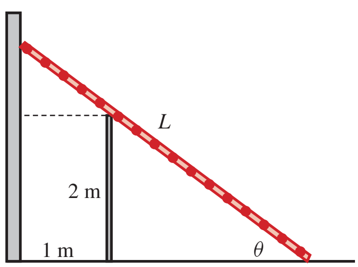


Figure 4.52

EXAMPLE 3 Find the length of the shortest ladder that can extend from a vertical wall, over a fence 2 m high located 1 m away from the wall, to a point on the ground outside the fence.

Solution Let θ be the angle of inclination of the ladder, as shown in Figure 4.52. Using the two right-angled triangles in the figure, we obtain the length L of the ladder as a function of θ:

$$L = L(\theta) = \frac{1}{\cos\theta} + \frac{2}{\sin\theta},$$

where $0 < \theta < \pi/2$. Since

$$\lim_{\theta \to (\pi/2)-} L(\theta) = \infty \quad \text{and} \quad \lim_{\theta \to 0+} L(\theta) = \infty,$$

$L(\theta)$ must have a minimum value on $(0, \pi/2)$, occurring at a critical point. (L has no singular points in $(0, \pi/2)$.) To find any critical points, we set

$$0 = L'(\theta) = \frac{\sin\theta}{\cos^2\theta} - \frac{2\cos\theta}{\sin^2\theta} = \frac{\sin^3\theta - 2\cos^3\theta}{\cos^2\theta\,\sin^2\theta}.$$

Any critical point satisfies $\sin^3\theta = 2\cos^3\theta$, or, equivalently, $\tan^3\theta = 2$. We don't need to solve this equation for $\theta = \tan^{-1}(2^{1/3})$ since it is really the corresponding value of $L(\theta)$ that we want. Observe that

$$\sec^2\theta = 1 + \tan^2\theta = 1 + 2^{2/3}.$$

It follows that

$$\cos\theta = \frac{1}{(1 + 2^{2/3})^{1/2}} \quad \text{and} \quad \sin\theta = \tan\theta\cos\theta = \frac{2^{1/3}}{(1 + 2^{2/3})^{1/2}}.$$

Therefore, the minimal value of $L(\theta)$ is

$$\frac{1}{\cos\theta} + \frac{2}{\sin\theta} = (1 + 2^{2/3})^{1/2} + 2\,\frac{(1 + 2^{2/3})^{1/2}}{2^{1/3}} = \left(1 + 2^{2/3}\right)^{3/2} \approx 4.16.$$

The shortest ladder that can extend from the wall over the fence to the ground outside is about 4.16 m long.

EXAMPLE 4 Find the most economical shape of a cylindrical tin can.

Solution This problem is stated in a rather vague way. We must consider what is meant by "most economical" and even "shape." Without further information, we can take one of two points of view:

(i) the volume of the tin can is to be regarded as given, and we must choose the dimensions to minimize the total surface area, or

(ii) the total surface area is given (we can use just so much metal), and we must choose the dimensions to maximize the volume.

We will discuss other possible interpretations later. Since a cylinder is determined by its radius and height (Figure 4.53), its shape is determined by the ratio radius/height. Let r, h, S, and V denote, respectively, the radius, height, total surface area, and volume of the can. The volume of a cylinder is the base area times the height:

$$V = \pi r^2 h.$$

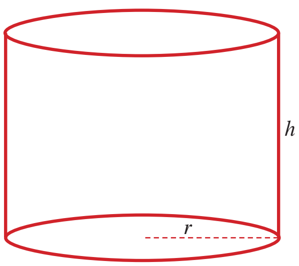


Figure 4.53

The surface of the can is made up of the cylindrical wall and circular disks for the top and bottom. The disks each have area πr^2, and the cylindrical wall is really just a rolled-up rectangle with base $2\pi r$ (the circumference of the can) and height h. Therefore, the total surface area of the can is

$$S = 2\pi r h + 2\pi r^2.$$

Let us use interpretation (i): V is a given constant, and S is to be minimized. We can use the equation for V to eliminate one of the two variables r and h on which S depends. Say we solve for $h = V/(\pi r^2)$ and substitute into the equation for S to obtain S as a function of r alone:

$$S = S(r) = 2\pi r\,\frac{V}{\pi r^2} + 2\pi r^2 = \frac{2V}{r} + 2\pi r^2 \qquad (0 < r < \infty).$$

Evidently, $\lim_{r\to 0+} S(r) = \infty$ and $\lim_{r\to\infty} S(r) = \infty$. Being differentiable and therefore continuous on $(0, \infty)$, $S(r)$ must have a minimum value, and it must occur at a critical point. To find any critical points,

$$0 = S'(r) = -\frac{2V}{r^2} + 4\pi r,$$
$$r^3 = \frac{2V}{4\pi} = \frac{1}{2\pi}\pi r^2 h = \frac{1}{2} r^2 h.$$

Thus, $h = 2r$ at the critical point of S. Under interpretation (i), the most economical can is shaped so that its height equals the diameter of its base. You are encouraged to show that interpretation (ii) leads to the same conclusion.

Remark A different approach to the problem in Example 4 shows directly that interpretations (i) and (ii) must give the same solution. Again, we start from the two equations

$$V = \pi r^2 h \qquad \text{and} \qquad S = 2\pi r h + 2\pi r^2.$$

If we regard h as a function of r and differentiate implicitly, we obtain

$$\frac{dV}{dr} = 2\pi r h + \pi r^2 \frac{dh}{dr},$$
$$\frac{dS}{dr} = 2\pi h + 2\pi r \frac{dh}{dr} + 4\pi r.$$

Under interpretation (i), V is constant and we want a critical point of S; under interpretation (ii), S is constant and we want a critical point of V. In *either* case, $dV/dr = 0$ and $dS/dr = 0$. Hence, both interpretations yield

$$2\pi r h + \pi r^2 \frac{dh}{dr} = 0 \qquad \text{and} \qquad 2\pi h + 4\pi r + 2\pi r \frac{dh}{dr} = 0.$$

If we divide the first equation by πr^2 and the second equation by $2\pi r$ and subtract to eliminate dh/dr, we again get $h = 2r$.

Remark **Modifying Example 4** Given the sparse information provided in the statement of the problem in Example 4, interpretations (i) and (ii) are the best we can do. The problem could be made more meaningful economically (from the point of view, say, of a tin can manufacturer) if more elements were brought into it. For example:

(a) Most cans use thicker material for the cylindrical wall than for the top and bottom disks. If the cylindrical wall material costs \$$A$ per unit area and the material for the top and bottom costs \$$B$ per unit area, we might prefer to minimize the total cost of materials for a can of given volume. What is the optimal shape if $A = 2B$?

(b) Large numbers of cans are to be manufactured. The material is probably being cut out of sheets of metal. The cylindrical walls are made by bending up rectangles, and rectangles can be cut from the sheet with little or no waste. There will, however, always be a proportion of material wasted when the disks are cut out. The exact proportion will depend on how the disks are arranged; two possible arrangements are shown in Figure 4.54. What is the optimal shape of the can if a square packing of disks is used? A hexagonal packing? Any such modification of the original problem will alter the optimal shape to some extent. In "real-world" problems, many factors may have to be taken into account to come up with a "best" strategy.

(c) The problem makes no provision for costs of manufacturing the can other than the cost of sheet metal. There may also be costs for joining the opposite edges of the rectangle to make the cylinder and for joining the top and bottom disks to the cylinder. These costs may be proportional to the lengths of the joins.

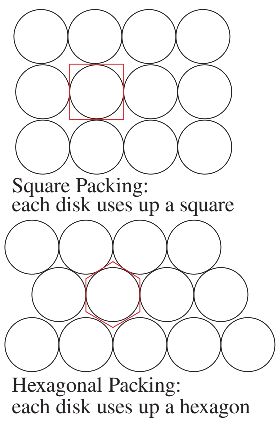


Figure 4.54 Square and hexagonal packing of disks in a plane

In most of the examples above, the maximum or minimum value being sought occurred at a critical point. Our final example is one where this is not the case.

EXAMPLE 5 A man can run twice as fast as he can swim. He is standing at point A on the edge of a circular swimming pool 40 m in diameter, and he wishes to get to the diametrically opposite point B as quickly as possible. He can run around the edge to point C, then swim directly from C to B. Where should C be chosen to minimize the total time taken to get from A to B?

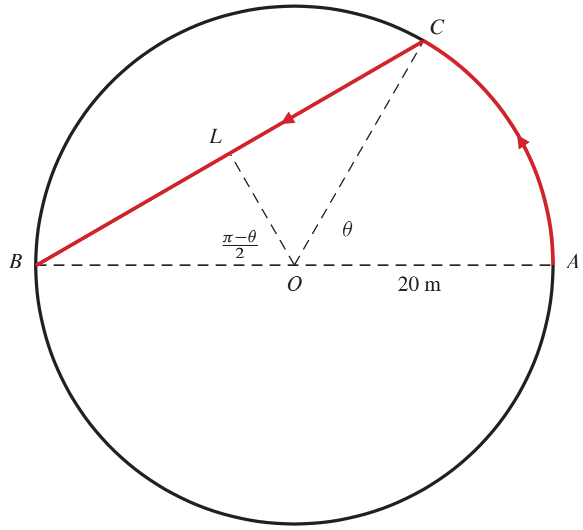


Figure 4.55 Running and swimming to get from A to B

Solution It is convenient to describe the position of C in terms of the angle AOC, where O is the centre of the pool. (See Figure 4.55.) Let θ denote this angle. Clearly, $0 \le \theta \le \pi$. (If $\theta = 0$, the man swims the whole way; if $\theta = \pi$, he runs the whole way.) The radius of the pool is 20 m, so arc $AC = 20\theta$. Since angle $BOC = \pi - \theta$, we have angle $BOL = (\pi - \theta)/2$ and chord $BC = 2BL = 40\sin\big((\pi - \theta)/2\big)$.

Suppose the man swims at a rate k m/s and therefore runs at a rate $2k$ m/s. If t is the total time he takes to get from A to B, then

$$\begin{aligned} t = t(\theta) &= \text{time running} + \text{time swimming} \\ &= \frac{20\theta}{2k} + \frac{40}{k}\sin\frac{\pi - \theta}{2}. \end{aligned}$$

(We are assuming that no time is wasted in jumping into the water at C.) The domain of t is $[0, \pi]$ and t has no singular points. Since t is continuous on a closed, finite interval, it must have a minimum value, and that value must occur at a critical point or an endpoint. For critical points,

$$0 = t'(\theta) = \frac{10}{k} - \frac{20}{k}\cos\frac{\pi - \theta}{2}.$$

Thus,

$$\cos\frac{\pi - \theta}{2} = \frac{1}{2}, \qquad \frac{\pi - \theta}{2} = \frac{\pi}{3}, \qquad \theta = \frac{\pi}{3}.$$

This is the only critical value of θ lying in the interval $[0, \pi]$. We have

$$t\left(\frac{\pi}{3}\right) = \frac{10\pi}{3k} + \frac{40}{k}\sin\frac{\pi}{3} = \frac{10}{k}\left(\frac{\pi}{3} + \frac{4\sqrt{3}}{2}\right) \approx \frac{45.11}{k}.$$

We must also look at the endpoints $\theta = 0$ and $\theta = \pi$:

$$t(0) = \frac{40}{k}, \qquad t(\pi) = \frac{10\pi}{k} \approx \frac{31.4}{k}.$$

Evidently, $t(\pi)$ is the least of these three times. To get from A to B as quickly as possible, the man should run the entire distance.

Remark This problem shows how important it is to check every candidate point to see whether it gives a maximum or minimum. Here, the critical point $\theta = \pi/3$ yielded the *worst* possible strategy: running one-third of the way around and then swimming the remainder would take the greatest time, not the least.

EXERCISES 4.8

1. Two positive numbers have sum 7. What is the largest possible value for their product?

2. Two positive numbers have product 8. What is the smallest possible value for their sum?

3. Two nonnegative numbers have sum 60. What are the numbers if the product of one of them and the square of the other is maximal?

4. Two numbers have sum 16. What are the numbers if the product of the cube of one and the fifth power of the other is as large as possible?

5. The sum of two nonnegative numbers is 10. What is the smallest value of the sum of the cube of one number and the square of the other?

6. Two nonnegative numbers have sum n. What is the smallest possible value for the sum of their squares?

7. Among all rectangles of given area, show that the square has the least perimeter.

8. Among all rectangles of given perimeter, show that the square has the greatest area.

9. Among all isosceles triangles of given perimeter, show that the equilateral triangle has the greatest area.

10. Find the largest possible area for an isosceles triangle if the length of each of its two equal sides is 10 m.

11. Find the area of the largest rectangle that can be inscribed in a semicircle of radius R if one side of the rectangle lies along the diameter of the semicircle.

12. Find the largest possible perimeter of a rectangle inscribed in a semicircle of radius R if one side of the rectangle lies along the diameter of the semicircle. (It is interesting that the rectangle with the largest perimeter has a different shape than the one with the largest area, obtained in Exercise 11.)

13. A rectangle with sides parallel to the coordinate axes is inscribed in the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Find the largest possible area for this rectangle.

14. Let ABC be a triangle right-angled at C and having area S. Find the maximum area of a rectangle inscribed in the triangle if (a) one corner of the rectangle lies at C, or (b) one side of the rectangle lies along the hypotenuse, AB.

15. Find the maximum area of an isosceles triangle whose equal sides are 10 cm in length. Use half the length of the third side of the triangle as the variable in terms of which to express the area of the triangle.

16. Repeat Exercise 15, but use instead the angle between the equal sides of the triangle as the variable in terms of which to express the area of the triangle. Which solution is easier?

17. **(Designing a billboard)** A billboard is to be made with 100 m^2 of printed area and with margins of 2 m at the top and bottom and 4 m on each side. Find the outside dimensions of the billboard if its total area is to be a minimum.

18. **(Designing a box)** A box is to be made from a rectangular sheet of cardboard 70 cm by 150 cm by cutting equal squares out of the four corners and bending up the resulting four flaps to make the sides of the box. (The box has no top.) What is the largest possible volume of the box?

19. **(Using rebates to maximize profit)** An automobile manufacturer sells 2,000 cars per month, at an average profit of \$1,000 per car. Market research indicates that for each \$50 of factory rebate the manufacturer offers to buyers it can expect to sell 200 more cars each month. How much of a rebate should it offer to maximize its monthly profit?

20. **(Maximizing rental profit)** All 80 rooms in a motel will be rented each night if the manager charges \$40 or less per room. If he charges \$$(40 + x)$ per room, then $2x$ rooms will remain vacant. If each rented room costs the manager \$10 per day and each unrented room \$2 per day in overhead, how much should the manager charge per room to maximize his daily profit?

21. **(Minimizing travel time)** You are in a dune buggy in the desert 12 km due south of the nearest point A on a straight east-west road. You wish to get to point B on the road 10 km east of A. If your dune buggy can average 15 km/h travelling over the desert and 39 km/h travelling on the road, toward what point on the road should you head in order to minimize your travel time to B?

22. Repeat Exercise 21, but assume that B is only 4 km from A.

23. **(Flying with least energy)** At the altitude of airliners, winds can typically blow at a speed of about 100 knots (nautical miles per hour) from the west toward the east. A westward-flying passenger jet from London, England, on its way to Toronto, flies directly against this wind for 3,000 nautical miles. The energy per unit time expended by the airliner is proportional to v^3, where v is the speed of the airliner relative to the air. This reflects the power required to push aside the air exerting ram pressure proportional to v^2. What speed uses the least energy on this trip? Estimate the time it would take to fly this route at the resulting optimal speed. Is this a typical speed at which airliners travel? Explain.

24. **(Energy for a round trip)** In the preceding problem we found that an airliner flying against the wind at speed v with respect to the air consumes the least energy over a flight if it travels at $v = 3u/2$, where u is the speed of the headwind with respect to the ground. Assume the power (energy per unit time) required to push aside the air is kv^3.

    (a) Write the general expression for energy consumed over a trip of distance ℓ flying with an airspeed v into a headwind of speed u. Also write the general expression for energy used on the return journey along the same path with airspeed w aided by a tailwind of speed u.

(b) Show that the energy consumed in the return journey is a strictly increasing function of w. What is the least energy consumed in the return journey if the airliner must have a minimum airspeed of s (known as "stall speed") to stay aloft?

(c) What is the least energy consumed in the round trip if $u > 2s/3$? What is the energy consumed when $u < 2s/3$?

25. A one-metre length of stiff wire is cut into two pieces. One piece is bent into a circle, the other piece into a square. Find the length of the part used for the square if the sum of the areas of the circle and the square is (a) maximum and (b) minimum.

26. Find the area of the largest rectangle that can be drawn so that each of its sides passes through a different vertex of a rectangle having sides a and b.

27. What is the length of the shortest line segment having one end on the x-axis, the other end on the y-axis, and passing through the point $(9, \sqrt{3})$?

28. **(Getting around a corner)** Find the length of the longest beam that can be carried horizontally around the corner from a hallway of width a m to a hallway of width b m. (See Figure 4.56; assume the beam has no width.)

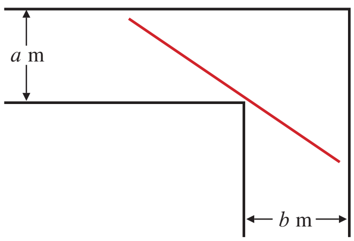


Figure 4.56

29. If the height of both hallways in Exercise 28 is c m, and if the beam need not be carried horizontally, how long can it be and still get around the corner? *Hint:* You can use the result of the previous exercise to do this one easily.

30. The fence in Example 3 is demolished and a new fence is built 2 m away from the wall. How high can the fence be if a 6 m ladder must be able to extend from the wall, over the fence, to the ground outside?

31. Find the shortest distance from the origin to the curve $x^2 y^4 = 1$.

32. Find the shortest distance from the point $(8, 1)$ to the curve $y = 1 + x^{3/2}$.

33. Find the dimensions of the largest right-circular cylinder that can be inscribed in a sphere of radius R.

34. Find the dimensions of the circular cylinder of greatest volume that can be inscribed in a cone of base radius R and height H if the base of the cylinder lies in the base of the cone.

35. A box with square base and no top has a volume of 4 m^3. Find the dimensions of the most economical box.

36. **(Folding a pyramid)** A pyramid with a square base and four faces, each in the shape of an isosceles triangle, is made by cutting away four triangles from a 2 ft square piece of cardboard (as shown in Figure 4.57) and bending up the resulting triangles to form the walls of the pyramid. What is the largest volume the pyramid can have? *Hint:* The volume of a pyramid having base area A and height h measured perpendicular to the base is $V = \frac{1}{3}Ah$.

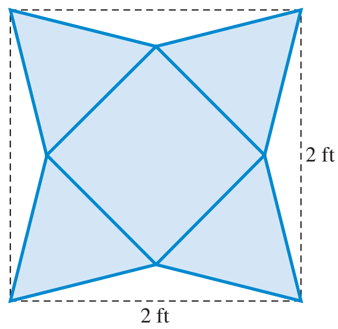


Figure 4.57

37. **(Getting the most light)** A window has perimeter 10 m and is in the shape of a rectangle with the top edge replaced by a semicircle. Find the dimensions of the rectangle if the window admits the greatest amount of light.

38. **(Fuel tank design)** A fuel tank is made of a cylindrical part capped by hemispheres at each end. If the hemispheres are twice as expensive per unit area as the cylindrical wall, and if the volume of the tank is V, find the radius and height of the cylindrical part to minimize the total cost. The surface area of a sphere of radius r is $4\pi r^2$; its volume is $\frac{4}{3}\pi r^3$.

39. **(Reflection of light)** Light travels in such a way that it requires the minimum possible time to get from one point to another. A ray of light from C reflects off a plane mirror AB at X and then passes through D. (See Figure 4.58.) Show that the rays CX and XD make equal angles with the normal to AB at X. (*Remark:* You may wish to give a proof based on elementary geometry without using any calculus, or you can minimize the travel time on CXD.)

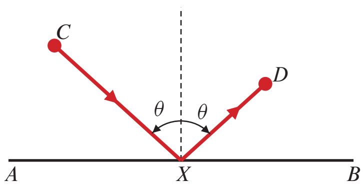


Figure 4.58

40. **(Snell's Law)** If light travels with speed v_1 in one medium and speed v_2 in a second medium, and if the two media are separated by a plane interface, show that a ray of light passing from point A in one medium to point B in the other is bent at the interface in such a way that

$$\frac{\sin i}{\sin r} = \frac{v_1}{v_2},$$

where i and r are the angles of incidence and refraction, as is shown in Figure 4.59. This is known as Snell's Law. Deduce it from the least-time principle stated in Exercise 39.

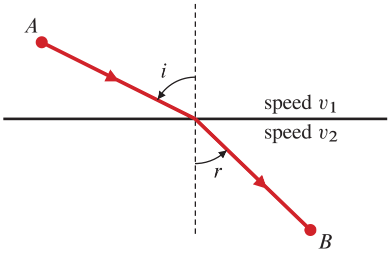


Figure 4.59

41. **(Cutting the stiffest beam)** The stiffness of a wooden beam of rectangular cross section is proportional to the product of the width and the cube of the depth of the cross section. Find the width and depth of the stiffest beam that can be cut out of a circular log of radius R.

42. Find the equation of the straight line of maximum slope tangent to the curve $y = 1 + 2x - x^3$.

43. A quantity Q grows according to the differential equation

$$\frac{dQ}{dt} = kQ^3(L - Q)^5,$$

where k and L are positive constants. How large is Q when it is growing most rapidly?

44. Find the smallest possible volume of a right-circular cone that can contain a sphere of radius R. (The volume of a cone of base radius r and height h is $\frac{1}{3}\pi r^2 h$.)

45. **(Ferry loading)** A ferry runs between the mainland and the island of Dedlos. The ferry has a maximum capacity of 1,000 cars, but loading near capacity is very time consuming. It is found that the number of cars that can be loaded in t hours is

$$f(t) = 1{,}000\,\frac{t}{e^{-t} + t}.$$

(Note that $\lim_{t\to\infty} f(t) = 1{,}000$, as expected.) Further, it is found that it takes $x/1{,}000$ hours to unload x cars. The sailing time to or from the island is 1 hour. Assume there are always more cars waiting for each sailing than can be loaded. How many cars should be loaded on the ferry for each sailing to maximize the average movement of cars back and forth to the island? (You will need to use a graphing calculator or computer software like Maple's `fsolve` routine to find the appropriate critical point.)

46. **(The best view of a mural)** How far back from a mural should one stand to view it best if the mural is 10 ft high and the bottom of it is 2 ft above eye level? (See Figure 4.60.)

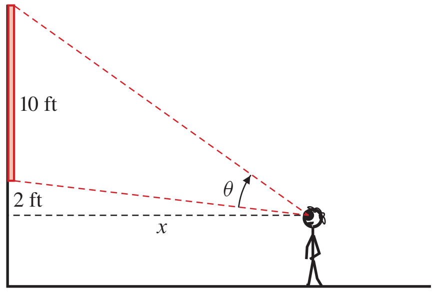


Figure 4.60

47. **(Improving the enclosure of Example 1)** An enclosure is to be constructed having part of its boundary along an existing straight wall. The other part of the boundary is to be fenced in the shape of an arc of a circle. If 100 m of fencing is available, what is the area of the largest possible enclosure? Into what fraction of a circle is the fence bent?

48. **(Designing a Dixie cup)** A sector is cut out of a circular disk of radius R, and the remaining part of the disk is bent up so that the two edges join and a cone is formed (see Figure 4.61). What is the largest possible volume for the cone?

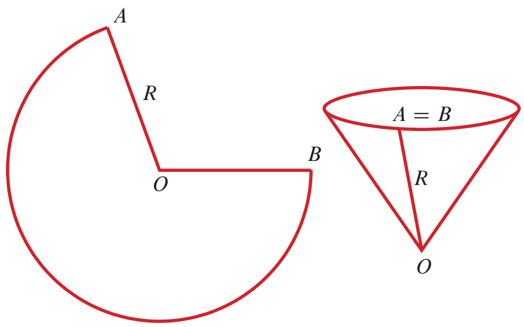


Figure 4.61

49. **(Minimize the fold)** One corner of a strip of paper a cm wide is folded up so that it lies along the opposite edge. (See Figure 4.62.) Find the least possible length for the fold line.

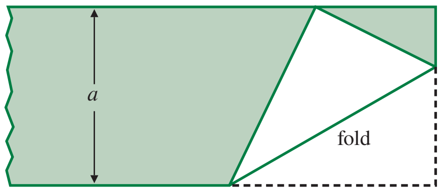


Figure 4.62

4.9 Linear Approximations

Many problems in applied mathematics are too difficult to be solved exactly—that is why we resort to using computers, even though in many cases they may only give approximate answers. However, not all approximation is done with machines. Linear approximation can be a very effective way to estimate values or test the plausibility of

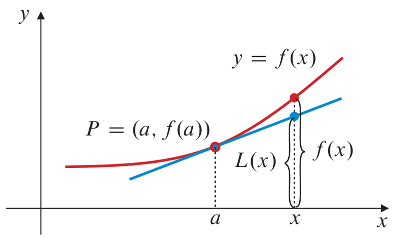


Figure 4.63 The linearization of function f about a

numbers given by a computer. In Section 2.7 we observed how differentials could be used to approximate (changes in) the values of functions between nearby points. In this section we reconsider such approximations in a more formal way and obtain estimates for the size of the errors encountered when such "linear" approximations are made.

The tangent to the graph $y = f(x)$ at $x = a$ describes the behaviour of that graph near the point $P = (a, f(a))$ better than any other straight line through P, because it goes through P in the same direction as the curve $y = f(x)$. (See Figure 4.63.) We exploit this fact by using the height to the tangent line to calculate approximate values of $f(x)$ for values of x near a. The tangent line has equation $y = f(a) + f'(a)(x - a)$. We call the right side of this equation the linearization of f about a (or the linearization of $f(x)$ about $x = a$).

DEFINITION 8

The **linearization** of the function f about a is the function L defined by

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

We say that $f(x) \approx L(x) = f(a) + f'(a)(x - a)$ provides **linear approximations** for values of f near a.

EXAMPLE 1 Find linearizations of (a) $f(x) = \sqrt{1 + x}$ about $x = 0$ and (b) $g(t) = 1/t$ about $t = 1/2$.

Solution

(a) We have $f(0) = 1$ and, since $f'(x) = 1/(2\sqrt{1+x})$, $f'(0) = 1/2$. The linearization of f about 0 is

$$L(x) = 1 + \frac{1}{2}(x - 0) = 1 + \frac{x}{2}.$$

(b) We have $g(1/2) = 2$ and, since $g'(t) = -1/t^2$, $g'(1/2) = -4$. The linearization of $g(t)$ about $t = 1/2$ is

$$L(t) = 2 - 4\left(t - \frac{1}{2}\right) = 4 - 4t.$$

Approximating Values of Functions

We have already made use of linearization in Section 2.7, where it was disguised as the formula

$$\Delta y \approx \frac{dy}{dx}\,\Delta x$$

and used to approximate a small change $\Delta y = f(a + \Delta x) - f(a)$ in the values of function f corresponding to the small change in the argument of the function from a to $a + \Delta x$. This is just the linear approximation

$$f(a + \Delta x) \approx L(a + \Delta x) = f(a) + f'(a)\Delta x.$$

EXAMPLE 2 A ball of ice melts so that its radius decreases from 5 cm to 4.92 cm. By approximately how much does the volume of the ball decrease?

Solution The volume V of a ball of radius r is $V = \frac{4}{3}\pi r^3$, so that $dV/dr = 4\pi r^2$ and $L(r + \Delta r) = V(r) + 4\pi r^2\,\Delta r$. Thus,

$$\Delta V \approx L(r + \Delta r) = 4\pi r^2\,\Delta r.$$

For $r = 5$ and $\Delta r = -0.08$, we have

$$\Delta V \approx 4\pi(5^2)(-0.08) = -8\pi \approx -25.13.$$

The volume of the ball decreases by about 25 cm^3.

The following example illustrates the use of linearization to find an approximate value of a function near a point where the values of the function and its derivative are known.

EXAMPLE 3 Use the linearization for $\sqrt{x}$ about $x = 25$ to find an approximate value for $\sqrt{26}$.

Solution If $f(x) = \sqrt{x}$, then $f'(x) = 1/(2\sqrt{x})$. Since we know that $f(25) = 5$ and $f'(25) = 1/10$, the linearization of $f(x)$ about $x = 25$ is

$$L(x) = 5 + \frac{1}{10}(x - 25).$$

Putting $x = 26$, we get

$$\sqrt{26} = f(26) \approx L(26) = 5 + \frac{1}{10}(26 - 25) = 5.1.$$

If we use the square root function on a calculator we can obtain the "true value" of $\sqrt{26}$ (actually, just another approximation, although presumably a better one): $\sqrt{26} = 5.099\,019\,5\ldots$, but if we have such a calculator we don't need the approximation in the first place. Approximations are useful when there is no easy way to obtain the true value. However, if we don't know the true value, we would at least like to have some way of determining how good the approximation must be; that is, we want an *estimate for the error*. After all, *any number* is an approximation to $\sqrt{26}$, but the error may be unacceptably large; for instance, the size of the error in the approximation $\sqrt{26} \approx 1{,}000{,}000$ is greater than 999,994.

Error Analysis

In any approximation, the **error** is defined by

$$\text{error} = \text{true value} - \text{approximate value}.$$

If the linearization of f about a is used to approximate $f(x)$ near $x = a$, that is,

$$f(x) \approx L(x) = f(a) + f'(a)(x - a),$$

then the error $E(x)$ in this approximation is

$$E(x) = f(x) - L(x) = f(x) - f(a) - f'(a)(x - a).$$

It is the vertical distance at x between the graph of f and the tangent line to that graph at $x = a$, as shown in Figure 4.64. Observe that if x is "near" a, then $E(x)$ is small compared to the horizontal distance between x and a.

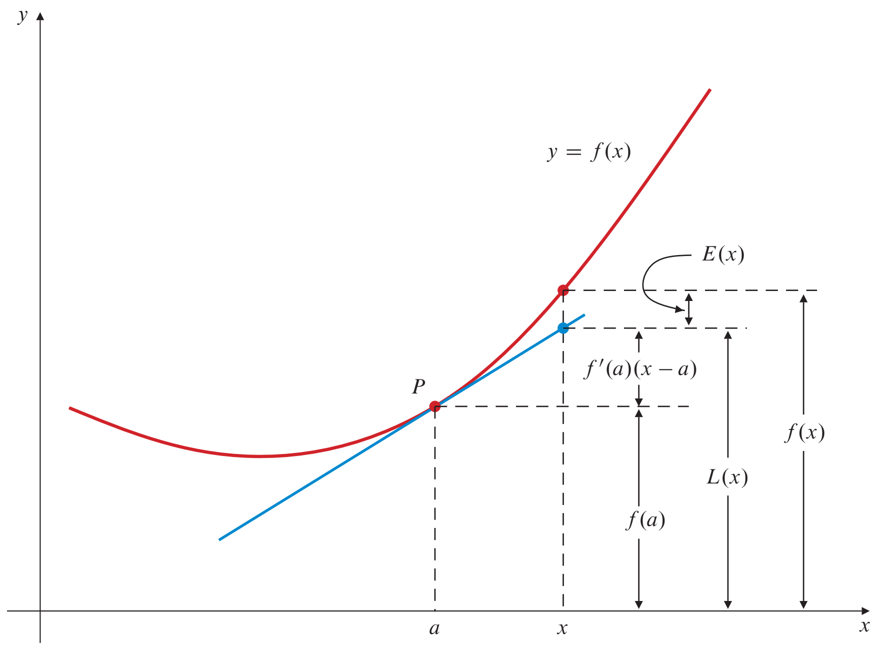


Figure 4.64 $f(x)$ and its linearization $L(x)$ about $x = a$. $E(x)$ is the error in the approximation $f(x) \approx L(x)$

The following theorem and its corollaries give us a way to estimate this error if we know bounds for the *second derivative* of f.

THEOREM 11

An error formula for linearization

If $f''(t)$ exists for all t in an interval containing a and x, then there exists some point s between a and x such that the error $E(x) = f(x) - L(x)$ in the linear approximation $f(x) \approx L(x) = f(a) + f'(a)(x - a)$ satisfies

$$E(x) = \frac{f''(s)}{2}(x-a)^2.$$

PROOF Let us assume that $x > a$. (The proof for $x < a$ is similar.) Since

$$E(t) = f(t) - f(a) - f'(a)(t-a),$$

we have $E'(t) = f'(t) - f'(a)$. We apply the Generalized Mean-Value Theorem (Theorem 16 of Section 2.8) to the two functions $E(t)$ and $(t-a)^2$ on $[a, x]$. Noting that $E(a) = 0$, we obtain a number u in (a, x) such that

$$\frac{E(x)}{(x-a)^2} = \frac{E(x) - E(a)}{(x-a)^2 - (a-a)^2} = \frac{E'(u)}{2(u-a)} = \frac{f'(u) - f'(a)}{2(u-a)} = \frac{1}{2}f''(s)$$

for some s in (a, u); the latter expression is a consequence of applying the Mean-Value Theorem again, this time to f' on $[a, u]$. Thus,

$$E(x) = \frac{f''(s)}{2}(x-a)^2$$

as claimed.

The following three corollaries are immediate consequences of Theorem 11.

Corollary A. If $f''(t)$ has constant sign (i.e., is always positive or always negative) between a and x, then the error $E(x)$ in the linear approximation $f(x) \approx L(x)$ in the Theorem has that same sign; if $f''(t) > 0$ between a and x, then $f(x) > L(x)$; if $f''(t) < 0$ between a and x, then $f(x) < L(x)$.

Corollary B. If $|f''(t)| < K$ for all t between a and x (where K is a constant), then $|E(x)| < (K/2)(x-a)^2$.

Corollary C. If $f''(t)$ satisfies $M < f''(t) < N$ for all t between a and x (where M and N are constants), then

$$L(x) + \frac{M}{2}(x-a)^2 < f(x) < L(x) + \frac{N}{2}(x-a)^2.$$

If M and N have the same sign, a better approximation to $f(x)$ is given by the midpoint of this interval containing $f(x)$:

$$f(x) \approx L(x) + \frac{M+N}{4}(x-a)^2.$$

For this approximation the error is less than half the length of the interval:

$$|\text{Error}| < \frac{N-M}{4}(x-a)^2.$$

EXAMPLE 4 Determine the sign and estimate the size of the error in the approximation $\sqrt{26} \approx 5.1$ obtained in Example 3. Use these to give a small interval that you can be sure contains $\sqrt{26}$.

Solution For $f(t) = t^{1/2}$, we have

$$f'(t) = \frac{1}{2}t^{-1/2} \qquad \text{and} \qquad f''(t) = -\frac{1}{4}t^{-3/2}.$$

For $25 < t < 26$, we have $f''(t) < 0$, so $\sqrt{26} = f(26) < L(26) = 5.1$. Also, $t^{3/2} > 25^{3/2} = 125$, so $|f''(t)| < (1/4)(1/125) = 1/500$ and

$$|E(26)| < \frac{1}{2} \times \frac{1}{500} \times (26-25)^2 = \frac{1}{1{,}000} = 0.001.$$

Therefore, $f(26) > L(26) - 0.001 = 5.099$, and $\sqrt{26}$ is in the interval $(5.099, 5.1)$.

Remark We can use Corollary C of Theorem 11 and the fact that $\sqrt{26} < 5.1$ to find a better (i.e., smaller) interval containing $\sqrt{26}$ as follows. If $25 < t < 26$, then $125 = 25^{3/2} < t^{3/2} < 26^{3/2} < 5.1^3$. Thus,

$$M = -\frac{1}{4 \times 125} < f''(t) < -\frac{1}{4 \times 5.1^3} = N$$

$$\sqrt{26} \approx L(26) + \frac{M+N}{4} = 5.1 - \frac{1}{4}\left(\frac{1}{4 \times 125} + \frac{1}{4 \times 5.1^3}\right) \approx 5.099\,028\,8$$

$$|\text{Error}| < \frac{N-M}{4} = \frac{1}{16}\left(-\frac{1}{5.1^3} + \frac{1}{125}\right) \approx 0.000\,028\,8.$$

Thus, $\sqrt{26}$ lies in the interval $(5.099\,00, 5.099\,06)$.

EXAMPLE 5 Use a suitable linearization to find an approximate value for $\cos 36° = \cos(\pi/5)$. Is the true value greater than or less than your approximation? Estimate the size of the error, and give an interval that you can be sure contains $\cos 36°$.

Solution Let $f(t) = \cos t$, so that $f'(t) = -\sin t$ and $f''(t) = -\cos t$. The value of a nearest to $36°$ for which we know $\cos a$ is $a = 30° = \pi/6$, so we use the linearization about that point:

$$L(x) = \cos\frac{\pi}{6} - \sin\frac{\pi}{6}\left(x - \frac{\pi}{6}\right) = \frac{\sqrt{3}}{2} - \frac{1}{2}\left(x - \frac{\pi}{6}\right).$$

Since $(\pi/5) - (\pi/6) = \pi/30$, our approximation is

$$\cos 36° = \cos\frac{\pi}{5} \approx L\left(\frac{\pi}{5}\right) = \frac{\sqrt{3}}{2} - \frac{1}{2}\left(\frac{\pi}{30}\right) \approx 0.813\,67.$$

If $(\pi/6) < t < (\pi/5)$, then $f''(t) < 0$ and $|f''(t)| < \cos(\pi/6) = \sqrt{3}/2$. Therefore, $\cos 36° < 0.813\,67$ and

$$|E(36°)| < \frac{\sqrt{3}}{4}\left(\frac{\pi}{30}\right)^2 < 0.004\,75.$$

Thus, $0.813\,67 - 0.004\,75 < \cos 36° < 0.813\,67$, so $\cos 36°$ lies in the interval $(0.808\,92, 0.813\,67)$.

Remark The error in the linearization of $f(x)$ about $x = a$ can be interpreted in terms of differentials (see Section 2.7 and the beginning of this section) as follows: if $\Delta x = dx = x - a$, then the change in $f(x)$ as we pass from $x = a$ to $x = a + \Delta x$ is $f(a + \Delta x) - f(a) = \Delta y$, and the corresponding change in the linearization $L(x)$ is $f'(a)(x - a) = f'(a)\,dx$, which is just the value at $x = a$ of the differential $dy = f'(x)\,dx$. Thus,

$$E(x) = \Delta y - dy.$$

The error $E(x)$ is small compared with Δx as Δx approaches 0, as seen in Figure 4.64. In fact,

$$\lim_{\Delta x \to 0} \frac{\Delta y - dy}{\Delta x} = \lim_{\Delta x \to 0}\left(\frac{\Delta y}{\Delta x} - \frac{dy}{dx}\right) = \frac{dy}{dx} - \frac{dy}{dx} = 0.$$

If $|f''(t)| \le K$ (constant) near $t = a$, a stronger assertion can be made:

$$\left|\frac{\Delta y - dy}{(\Delta x)^2}\right| = \left|\frac{E(x)}{(\Delta x)^2}\right| \le \frac{K}{2}, \qquad \text{so} \qquad |\Delta y - dy| \le \frac{K}{2}(\Delta x)^2.$$

EXERCISES 4.9

In Exercises 1–10, find the linearization of the given function about the given point.

1. x^2 about $x = 3$
2. x^{-3} about $x = 2$
3. $\sqrt{4 - x}$ about $x = 0$
4. $\sqrt{3 + x^2}$ about $x = 1$
5. $1/(1 + x)^2$ about $x = 2$
6. $1/\sqrt{x}$ about $x = 4$
7. $\sin x$ about $x = \pi$
8. $\cos(2x)$ about $x = \pi/3$
9. $\sin^2 x$ about $x = \pi/6$
10. $\tan x$ about $x = \pi/4$
11. By approximately how much does the area of a square increase if its side length increases from 10 cm to 10.4 cm?
12. By about how much must the edge length of a cube decrease from 20 cm to reduce the volume of the cube by 12 cm^3?
13. A spacecraft orbits the earth at a distance of 4,100 miles from the centre of the earth. By about how much will the circumference of its orbit decrease if the radius decreases by 10 miles?
14. **(Acceleration of gravity)** The acceleration a of gravity at an altitude of h miles above the surface of the earth is given by

$$a = g\left(\frac{R}{R + h}\right)^2,$$

where $g \approx 32$ ft/s^2 is the acceleration at the surface of the earth, and $R \approx 3,960$ miles is the radius of the earth. By about what percentage will a decrease if h increases from 0 to 10 miles?

In Exercises 15–22, use a suitable linearization to approximate the indicated value. Determine the sign of the error and estimate its size. Use this information to specify an interval you can be sure contains the value.

15. $\sqrt{50}$

16. $\sqrt{47}$

17. $\sqrt[4]{85}$

18. $\dfrac{1}{2.003}$

19. $\cos 46°$

20. $\sin \dfrac{\pi}{5}$

21. $\sin(3.14)$

22. $\sin 33°$

Use Corollary C of Theorem 11 in the manner suggested in the remark following Example 4 to find better intervals and better approximations to the values in Exercises 23–26.

23. $\sqrt{50}$ as first approximated in Exercise 15.

24. $\sqrt{47}$ as first approximated in Exercise 16.

25. $\cos 36°$ as first approximated in Example 5.

26. $\sin 33°$ as first approximated in Exercise 22.

27. If $f(2) = 4$, $f'(2) = -1$, and $0 \le f''(x) \le 1/x$ for $x > 0$, find the smallest interval you can be sure contains $f(3)$.

28. If $f(2) = 4$, $f'(2) = -1$, and $\dfrac{1}{2x} \le f''(x) \le \dfrac{1}{x}$ for $2 \le x \le 3$, find the best approximation you can for $f(3)$.

29. If $g(2) = 1$, $g'(2) = 2$, and $|g''(x)| < 1 + (x-2)^2$ for all $x > 0$, find the best approximation you can for $g(1.8)$. How large can the error be?

30. Show that the linearization of $\sin\theta$ at $\theta = 0$ is $L(\theta) = \theta$. How large can the percentage error in the approximation $\sin\theta \approx \theta$ be if $|\theta|$ is less than $17°$?

31. A spherical balloon is inflated so that its radius increases from 20.00 cm to 20.20 cm in 1 min. By approximately how much has its volume increased in that minute?

4.10 Taylor Polynomials

The linearization of a function $f(x)$ about $x = a$, namely, the linear function

$$P_1(x) = L(x) = f(a) + f'(a)(x-a),$$

describes the behaviour of f near a better than any other polynomial of degree 1 because both P_1 and f have the same value and the same derivative at a:

$$P_1(a) = f(a) \qquad \text{and} \qquad P_1'(a) = f'(a).$$

(We are now using the symbol P_1 instead of L to stress the fact that the linearization is a polynomial of degree at most 1.)

We can obtain even better approximations to $f(x)$ by using quadratic or higher-degree polynomials and matching more derivatives at $x = a$. For example, if f is twice differentiable near a, then the polynomial

$$P_2(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2$$

satisfies $P_2(a) = f(a)$, $P_2'(a) = f'(a)$, and $P_2''(a) = f''(a)$ and describes the behaviour of f near a better than any other polynomial of degree at most 2.

In general, if $f^{(n)}(x)$ exists in an open interval containing $x = a$, then the polynomial

$$\begin{aligned} P_n(x) = f(a) &+ \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 \\ &+ \frac{f'''(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n \end{aligned}$$

matches f and its first n derivatives at $x = a$,

$$P_n(a) = f(a), \quad P_n'(a) = f'(a), \quad \ldots, \quad P_n^{(n)}(a) = f^{(n)}(a),$$

and so describes $f(x)$ near $x = a$ better than any other polynomial of degree at most n. P_n is called the ***n*th-order Taylor polynomial for *f* about *a***. (Taylor polynomials about 0 are usually called **Maclaurin** polynomials.) The 0th-order Taylor polynomial for f about a is just the constant function $P_0(x) = f(a)$. The nth-order Taylor polynomial for f about a is sometimes called the nth-*degree* Taylor polynomial, but its degree will actually be less than n if $f^{(n)}(a) = 0$.

EXAMPLE 1 Find the following Taylor polynomials:

(a) $P_2(x)$ for $f(x) = \sqrt{x}$ about $x = 25$.
(b) $P_3(x)$ for $g(x) = \ln x$ about $x = e$.

Solution (a) $f'(x) = (1/2)x^{-1/2}$, $f''(x) = -(1/4)x^{-3/2}$. Thus,

$$\begin{aligned}P_2(x) &= f(25) + f'(25)(x-25) + \frac{f''(25)}{2!}(x-25)^2\\ &= 5 + \frac{1}{10}(x-25) - \frac{1}{1{,}000}(x-25)^2.\end{aligned}$$

(b) $g'(x) = \dfrac{1}{x}$, $g''(x) = -\dfrac{1}{x^2}$, $g'''(x) = \dfrac{2}{x^3}$. Thus,

$$\begin{aligned}P_3(x) &= g(e) + g'(e)(x-e) + \frac{g''(e)}{2!}(x-e)^2 + \frac{g'''(e)}{3!}(x-e)^3\\ &= 1 + \frac{1}{e}(x-e) - \frac{1}{2e^2}(x-e)^2 + \frac{1}{3e^3}(x-e)^3.\end{aligned}$$

EXAMPLE 2 Find the nth-order Maclaurin polynomial $P_n(x)$ for e^x. Use $P_0(1)$, $P_1(1)$, $P_2(1)$, $\ldots$ to calculate approximate values for $e = e^1$. Stop when you think you have 3 decimal places correct.

Solution Since every derivative of e^x is e^x and so is 1 at $x = 0$, the nth-order Maclaurin polynomial for e^x (i.e., Taylor polynomial at $x = 0$) is

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

Thus, we have for $x = 1$, adding one more term at each step:

$$\begin{aligned}P_0(1) &= 1\\ P_1(1) &= P_0(1) + \frac{1}{1!} = 1 + 1 = 2\\ P_2(1) &= P_1(1) + \frac{1}{2!} = 2 + \frac{1}{2} = 2.5\\ P_3(1) &= P_2(1) + \frac{1}{3!} = 2.5 + \frac{1}{6} = 2.6666\\ P_4(1) &= P_3(1) + \frac{1}{4!} = 2.6666 + \frac{1}{24} = 2.7083\\ P_5(1) &= P_4(1) + \frac{1}{5!} = 2.7083 + \frac{1}{120} = 2.7166\\ P_6(1) &= P_5(1) + \frac{1}{6!} = 2.7166 + \frac{1}{720} = 2.7180\\ P_7(1) &= P_6(1) + \frac{1}{7!} = 2.7180 + \frac{1}{5{,}040} = 2.7182.\end{aligned}$$

It appears that $e \approx 2.718$ to 3 decimal places. We will verify in Example 5 below that $P_7(1)$ does indeed give this much precision. The graphs of e^x and its first four Maclaurin polynomials are shown in Figure 4.65.

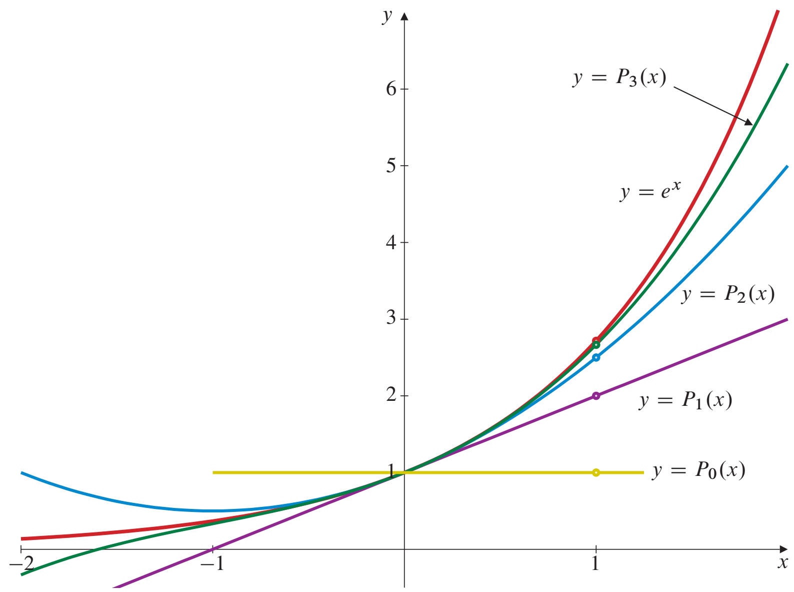


Figure 4.65 Some Maclaurin polynomials for e^x

EXAMPLE 3 Find Maclaurin polynomials $P_1(x)$, $P_2(x)$, $P_3(x)$, and $P_4(x)$ for $f(x) = \sin x$. Then write the general Maclaurin polynomials $P_{2n-1}(x)$ and $P_{2n}(x)$ for that function.

Solution We have $f'(x) = \cos x$, $f''(x) = -\sin x$, $f'''(x) = -\cos x$, and $f^{(4)}(x) = \sin x = f(x)$, so the pattern repeats for higher derivatives. Since

$$\begin{array}{llll} f(0) = 0, & f''(0) = 0, & f^{(4)}(0) = 0, & f^{(6)}(0) = 0, \dots \\ f'(0) = 1, & f'''(0) = -1, & f^{(5)}(0) = 1, & f^{(7)}(0) = -1, \dots \end{array}$$

we have

$$\begin{aligned} P_1(x) &= 0 + x = x \\ P_2(x) &= x + \frac{0}{2!}x^2 = x = P_1(x) \\ P_3(x) &= x - \frac{1}{3!}x^3 = x - \frac{x^3}{3!} \\ P_4(x) &= x - \frac{1}{3!}x^3 + \frac{0}{4!}x^4 = x - \frac{x^3}{3!} = P_3(x). \end{aligned}$$

In general, $f^{(2n-1)}(0) = (-1)^{n-1}$ and $f^{(2n)}(0) = 0$, so

$$P_{2n-1}(x) = P_{2n}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!}.$$

Taylor's Formula

The following theorem provides a formula for the error in a Taylor approximation $f(x) \approx P_n(x)$ similar to that provided for linear approximation by Theorem 11.

THEOREM 12

Taylor's Theorem

If the $(n+1)$st-order derivative, $f^{(n+1)}(t)$, exists for all t in an interval containing a and x, and if $P_n(x)$ is the nth-order Taylor polynomial for f about a, that is,

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

then the error $E_n(x) = f(x) - P_n(x)$ in the approximation $f(x) \approx P_n(x)$ is given by

$$E_n(x) = \frac{f^{(n+1)}(s)}{(n+1)!}(x-a)^{n+1},$$

where s is some number between a and x. The resulting formula

$$\begin{aligned} f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n \\ + \frac{f^{(n+1)}(s)}{(n+1)!}(x-a)^{n+1}, \quad \text{for some } s \text{ between } a \text{ and } x, \end{aligned}$$

is called **Taylor's formula with Lagrange remainder**; the Lagrange remainder term is the explicit formula given above for $E_n(x)$.

Note that the error term (Lagrange remainder) in Taylor's formula looks just like the next term in the Taylor polynomial would look if we continued the Taylor polynomial to include one more term (of degree $n+1$) EXCEPT that the derivative $f^{(n+1)}$ is not evaluated at a but rather at some (generally unknown) point s between a and x. This makes it easy to remember Taylor's formula.

PROOF Observe that the case $n = 0$ of Taylor's formula, namely,

$$f(x) = P_0(x) + E_0(x) = f(a) + \frac{f'(s)}{1!}(x-a),$$

is just the Mean-Value Theorem

$$\frac{f(x)-f(a)}{x-a} = f'(s) \quad \text{for some } s \text{ between } a \text{ and } x.$$

Also note that the case $n = 1$ is just the error formula for linearization given in Theorem 11.

We will complete the proof for higher n using mathematical induction. (See the proof of Theorem 2 in Section 2.3.) Suppose, therefore, that we have proved the case $n = k-1$, where $k \geq 2$ is an integer. Thus, we are assuming that if f is *any* function whose kth derivative exists on an interval containing a and x, then

$$E_{k-1}(x) = \frac{f^{(k)}(s)}{k!}(x-a)^k,$$

where s is some number between a and x. Let us consider the next higher case: $n = k$. As in the proof of Theorem 11, we assume $x > a$ (the case $x < a$ is similar) and apply the Generalized Mean-Value Theorem to the functions $E_k(t)$ and $(t-a)^{k+1}$ on $[a, x]$. Since $E_k(a) = 0$, we obtain a number u in (a, x) such that

$$\frac{E_k(x)}{(x-a)^{k+1}} = \frac{E_k(x) - E_k(a)}{(x-a)^{k+1} - (a-a)^{k+1}} = \frac{E_k'(u)}{(k+1)(u-a)^k}.$$

Now

$$\begin{aligned} E_k'(u) &= \frac{d}{dt}\left(f(t) - f(a) - f'(a)(t-a) - \frac{f''(a)}{2!}(t-a)^2 \right. \\ &\qquad \left. - \cdots - \frac{f^{(k)}(a)}{k!}(t-a)^k \right)\Bigg|_{t=u} \\ &= f'(u) - f'(a) - f''(a)(u-a) - \cdots - \frac{f^{(k)}(a)}{(k-1)!}(u-a)^{k-1}. \end{aligned}$$

This last expression is just $E_{k-1}(u)$ for the function f' instead of f. By the induction assumption it is equal to

$$\frac{(f')^{(k)}(s)}{k!}(u-a)^k = \frac{f^{(k+1)}(s)}{k!}(u-a)^k$$

for some s between a and u. Therefore,

$$E_k(x) = \frac{f^{(k+1)}(s)}{(k+1)!}(x-a)^{k+1}.$$

We have shown that the case $n = k$ of Taylor's Theorem is true if the case $n = k-1$ is true, and the inductive proof is complete.

Remark For any value of x for which $\lim_{n\to\infty} E_n(x) = 0$, we can ensure that the Taylor approximation $f(x) \approx P_n(x)$ is as close as we want by choosing n large enough.

EXAMPLE 4 Use the 2nd-order Taylor polynomial for $\sqrt{x}$ about $x = 25$ found in Example 1(a) to approximate $\sqrt{26}$. Estimate the size of the error, and specify an interval that you can be sure contains $\sqrt{26}$.

Solution In Example 1(a) we calculated $f''(x) = -(1/4)x^{-3/2}$ and obtained the Taylor polynomial

$$P_2(x) = 5 + \frac{1}{10}(x-25) - \frac{1}{1{,}000}(x-25)^2.$$

The required approximation is

$$\sqrt{26} = f(26) \approx P_2(26) = 5 + \frac{1}{10}(26-25) - \frac{1}{1{,}000}(26-25)^2 = 5.099.$$

Now $f'''(x) = (3/8)x^{-5/2}$. For $25 < s < 26$, we have

$$|f'''(s)| \le \frac{3}{8}\,\frac{1}{25^{5/2}} = \frac{3}{8\times 3{,}125} = \frac{3}{25{,}000}.$$

Thus, the error in the approximation satisfies

$$|E_2(26)| \le \frac{3}{25{,}000\times 6}(26-25)^3 = \frac{1}{50{,}000} = 0.000\,02.$$

Therefore, $\sqrt{26}$ lies in the interval $(5.098\,98, 5.099\,02)$.

EXAMPLE 5 Use Taylor's Theorem to confirm that the Maclaurin polynomial $P_7(x)$ for e^x is sufficient to give e correct to 3 decimal places as claimed in Example 2.

Solution The error in the approximation $e^x \approx P_n(x)$ satisfies

$$E_n(x) = \frac{e^s}{(n+1)!}x^{n+1}, \quad \text{for some } s \text{ between } 0 \text{ and } x.$$

If $x = 1$, then $0 < s < 1$, so $e^s < e < 3$ and $0 < E_n(1) < 3/(n+1)!$. To get an approximation for $e = e^1$ correct to 3 decimal places, we need to have $E_n(1) < 0.0005$. Since $3/(8!) = 3/40{,}320 \approx 0.000\,074$, but $3/(7!) = 3/5{,}040 \approx 0.000\,59$, we can be sure $n = 7$ will do, but we cannot be sure $n = 6$ will do:

$$e \approx 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + \frac{1}{7!} \approx 2.7183 \approx 2.718$$

to 3 decimal places.

Big-O Notation

DEFINITION 9

We write $f(x) = O\big(u(x)\big)$ as $x \to a$ (read this "$f(x)$ is big-Oh of $u(x)$ as x approaches a") provided that

$$|f(x)| \le K|u(x)|$$

holds for some constant K on some open interval containing $x = a$.

Similarly, $f(x) = g(x) + O\big(u(x)\big)$ as $x \to a$ if $f(x) - g(x) = O\big(u(x)\big)$ as $x \to a$, that is, if

$$|f(x) - g(x)| \le K|u(x)| \quad \text{near } a.$$

For example, $\sin x = O(x)$ as $x \to 0$ because $|\sin x| \le |x|$ near 0.

The following properties of big-O notation follow from the definition:

(i) If $f(x) = O\big(u(x)\big)$ as $x \to a$, then $Cf(x) = O\big(u(x)\big)$ as $x \to a$ for any value of the constant C.

(ii) If $f(x) = O\big(u(x)\big)$ as $x \to a$ and $g(x) = O\big(u(x)\big)$ as $x \to a$, then $f(x) \pm g(x) = O\big(u(x)\big)$ as $x \to a$.

(iii) If $f(x) = O\big((x-a)^k u(x)\big)$ as $x \to a$, then $f(x)/(x-a)^k = O\big(u(x)\big)$ as $x \to a$ for any constant k.

Taylor's Theorem says that if $f^{(n+1)}(t)$ exists on an interval containing a and x, and if P_n is the nth-order Taylor polynomial for f at a, then, as $x \to a$,

$$f(x) = P_n(x) + O\big((x-a)^{n+1}\big).$$

This is a statement about how rapidly the graph of the Taylor polynomial $P_n(x)$ approaches that of $f(x)$ as $x \to a$; the vertical distance between the graphs decreases as fast as $|x - a|^{n+1}$. The following theorem shows that the Taylor polynomial $P_n(x)$ is the *only* polynomial of degree at most n whose graph approximates the graph of $f(x)$ that rapidly.

THEOREM 13

If $f(x) = Q_n(x) + O\big((x-a)^{n+1}\big)$ as $x \to a$, where Q_n is a polynomial of degree at most n, then $Q_n(x) = P_n(x)$, that is, Q_n is the Taylor polynomial for $f(x)$ at $x = a$.

PROOF Let P_n be the Taylor polynomial, then properties (i) and (ii) of big-O imply that $R_n(x) = Q_n(x) - P_n(x) = O\big((x-a)^{n+1}\big)$ as $x \to a$. We want to show that $R_n(x)$ is identically zero so that $Q_n(x) = P_n(x)$ for all x. By replacing x with $a + (x - a)$ and expanding powers, we can write $R_n(x)$ in the form

$$R_n(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + \cdots + c_n(x-a)^n.$$

If $R_n(x)$ is not identically zero, then there is a smallest coefficient c_k ($k \le n$), such that $c_k \ne 0$, but $c_j = 0$ for $0 \le j \le k - 1$. Thus,

$$R_n(x) = (x-a)^k\big(c_k + c_{k+1}(x-a) + \cdots + c_n(x-a)^{n-k}\big).$$

Therefore, $\lim_{x\to a} R_n(x)/(x-a)^k = c_k \ne 0$. However, by property (iii) above we have $R_n(x)/(x-a)^k = O\big((x-a)^{n+1-k}\big)$. Since $n + 1 - k > 0$, this says $R_n(x)/(x-a)^k \to 0$ as $x \to a$. This contradiction shows that $R_n(x)$ must be identically zero. Therefore, $Q_n(x) = P_n(x)$ for all x.

Table 5 lists Taylor formulas about 0 (Maclaurin formulas) for some elementary functions, with error terms expressed using big-O notation.

Table 5. Some Maclaurin Formulas with Errors in Big-O Form

	As $x \to 0$:
(a)	$e^x = 1 + x + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \cdots + \dfrac{x^n}{n!} + O\big(x^{n+1}\big)$
(b)	$\cos x = 1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \cdots + (-1)^n \dfrac{x^{2n}}{(2n)!} + O\big(x^{2n+2}\big)$
(c)	$\sin x = x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \cdots + (-1)^n \dfrac{x^{2n+1}}{(2n+1)!} + O\big(x^{2n+3}\big)$
(d)	$\dfrac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots + x^n + O\big(x^{n+1}\big)$
(e)	$\ln(1+x) = x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \cdots + (-1)^{n-1} \dfrac{x^n}{n} + O\big(x^{n+1}\big)$
(f)	$\tan^{-1} x = x - \dfrac{x^3}{3} + \dfrac{x^5}{5} - \cdots + (-1)^n \dfrac{x^{2n+1}}{2n+1} + O\big(x^{2n+3}\big)$

It is worthwhile remembering these. The first three can be established easily by using Taylor's formula with Lagrange remainder; the other three would require much more effort to verify for general n. In Section 9.6 we will return to the subject of Taylor and Maclaurin polynomials in relation to Taylor and Maclaurin series. At that time we will have access to much more powerful machinery to establish such results. The need to calculate high-order derivatives can make the use of Taylor's formula difficult for all but the simplest functions.

The real importance of Theorem 13 is that it enables us to obtain Taylor polynomials for new functions by combining others already known; as long as the error term is of higher degree than the order of the polynomial obtained, the polynomial must be the Taylor polynomial. We illustrate this with a few examples.

EXAMPLE 6 Find the Maclaurin polynomial of order $2n$ for $\cosh x$.

Solution Write the Taylor formula for e^x at $x = 0$ (from Table 5) with n replaced by $2n+1$, and then rewrite that with x replaced by $-x$. We get

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{2n}}{(2n)!} + \frac{x^{2n+1}}{(2n+1)!} + O\big(x^{2n+2}\big), \\ e^{-x} &= 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots + \frac{x^{2n}}{(2n)!} - \frac{x^{2n+1}}{(2n+1)!} + O\big(x^{2n+2}\big) \end{aligned}$$

as $x \to 0$. Now average these two to get

$$\cosh x = \frac{e^x + e^{-x}}{2} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2n}}{(2n)!} + O\big(x^{2n+2}\big)$$

as $x \to 0$. By Theorem 13 the Maclaurin polynomial $P_{2n}(x)$ for $\cosh x$ is

$$P_{2n}(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2n}}{(2n)!}.$$

EXAMPLE 7 Obtain the Taylor polynomial of order 3 for e^{2x} about $x = 1$ from the corresponding Maclaurin polynomial for e^x (from Table 5).

Solution Writing $x = 1 + (x-1)$, we have

$$\begin{aligned} e^{2x} &= e^{2+2(x-1)} = e^2 e^{2(x-1)} \\ &= e^2\left[1 + 2(x-1) + \frac{2^2(x-1)^2}{2!} + \frac{2^3(x-1)^3}{3!} + O\big((x-1)^4\big)\right] \end{aligned}$$

as $x \to 1$. By Theorem 13 the Taylor polynomial $P_3(x)$ for e^{2x} at $x = 1$ must be

$$P_3(x) = e^2 + 2e^2(x-1) + 2e^2(x-1)^2 + \frac{4e^2}{3}(x-1)^3.$$

EXAMPLE 8 Use the Taylor formula for $\ln(1+x)$ (from Table 5) to find the Taylor polynomial $P_3(x)$ for $\ln x$ about $x = e$. (This provides an alternative to using the definition of Taylor polynomial as was done to solve the same problem in Example 1(b).)

Solution We have $x = e + (x-e) = e(1+t)$ where $t = (x-e)/e$. As $x \to e$ we have $t \to 0$, so

$$\begin{aligned} \ln x &= \ln e + \ln(1+t) = \ln e + t - \frac{t^2}{2} + \frac{t^3}{3} + O(t^4) \\ &= 1 + \frac{x-e}{e} - \frac{1}{2}\left(\frac{x-e}{e}\right)^2 + \frac{1}{3}\left(\frac{x-e}{e}\right)^3 + O\big((x-e)^4\big). \end{aligned}$$

Therefore, by Theorem 13,

$$P_3(x) = 1 + \frac{x-e}{e} - \frac{1}{2}\left(\frac{x-e}{e}\right)^2 + \frac{1}{3}\left(\frac{x-e}{e}\right)^3.$$

Evaluating Limits of Indeterminate Forms

Taylor and Maclaurin polynomials provide us with another method for evaluating limits of indeterminate forms of type $[0/0]$. For some such limits this method can be considerably easier than using l'Hôpital's Rule.

EXAMPLE 9 Evaluate $\displaystyle\lim_{x\to 0} \frac{2\sin x - \sin(2x)}{2e^x - 2 - 2x - x^2}$.

Solution Both the numerator and denominator approach 0 as $x \to 0$. Let us replace the trigonometric and exponential functions with their degree-3 Maclaurin polynomials plus error terms written in big-O notation:

$$\begin{aligned} &\lim_{x\to 0} \frac{2\sin x - \sin(2x)}{2e^x - 2 - 2x - x^2} \\ &\quad = \lim_{x\to 0} \frac{2\left(x - \frac{x^3}{3!} + O(x^5)\right) - \left(2x - \frac{2^3x^3}{3!} + O(x^5)\right)}{2\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + O(x^4)\right) - 2 - 2x - x^2} \\ &\quad = \lim_{x\to 0} \frac{-\frac{x^3}{3} + \frac{4x^3}{3} + O(x^5)}{\frac{x^3}{3} + O(x^4)} \\ &\quad = \lim_{x\to 0} \frac{1 + O(x^2)}{\frac{1}{3} + O(x)} = \frac{1}{\frac{1}{3}} = 3. \end{aligned}$$

Observe how we used the properties of big-O as listed in this section. We needed to use Maclaurin polynomials of degree at least 3 because all lower degree terms cancelled out in the numerator and the denominator.

EXAMPLE 10 Evaluate $\lim\limits_{x\to 1} \dfrac{\ln x}{x^2-1}$.

Solution This is also of type [0/0]. We begin by substituting $x = 1 + t$. Note that $x \to 1$ corresponds to $t \to 0$. We can use a known Maclaurin polynomial for $\ln(1+t)$. For this limit even the degree 1 polynomial $P_1(t) = t$ with error $O(t^2)$ will do.

$$\lim_{x\to 1} \frac{\ln x}{x^2-1} = \lim_{t\to 0} \frac{\ln(1+t)}{(1+t)^2-1} = \lim_{t\to 0} \frac{\ln(1+t)}{2t+t^2}$$
$$= \lim_{t\to 0} \frac{t+O(t^2)}{2t+t^2} = \lim_{t\to 0} \frac{1+O(t)}{2+t} = \frac{1}{2}.$$

EXERCISES 4.10

Find the indicated Taylor polynomials for the functions in Exercises 1–8 by using the definition of Taylor polynomial.

1. for e^{-x} about $x = 0$, order 4.

2. for $\cos x$ about $x = \pi/4$, order 3.

3. for $\ln x$ about $x = 2$, order 4.

4. for $\sec x$ about $x = 0$, order 3.

5. for $\sqrt{x}$ about $x = 4$, order 3.

6. for $1/(1-x)$ about $x = 0$, order n.

7. for $1/(2+x)$ about $x = 1$, order n.

8. for $\sin(2x)$ about $x = \pi/2$, order $2n-1$.

In Exercises 9–14, use second order Taylor polynomials $P_2(x)$ for the given function about the point specified to approximate the indicated value. Estimate the error, and write the smallest interval you can be sure contains the value.

9. $f(x) = x^{1/3}$ about 8; approximate $9^{1/3}$.

10. $f(x) = \sqrt{x}$ about 64; approximate $\sqrt{61}$.

11. $f(x) = \dfrac{1}{x}$ about 1; approximate $\dfrac{1}{1.02}$.

12. $f(x) = \tan^{-1} x$ about 1; approximate $\tan^{-1}(0.97)$.

13. $f(x) = e^x$ about 0; approximate $e^{-0.5}$.

14. $f(x) = \sin x$ about $\pi/4$; approximate $\sin(47°)$.

In Exercises 15–20, write the indicated case of Taylor's formula for the given function. What is the Lagrange remainder in each case?

15. $f(x) = \sin x,\ a = 0,\ n = 7$

16. $f(x) = \cos x,\ a = 0,\ n = 6$

17. $f(x) = \sin x,\ a = \pi/4,\ n = 4$

18. $f(x) = \dfrac{1}{1-x},\ a = 0,\ n = 6$

19. $f(x) = \ln x,\ a = 1, n = 6$

20. $f(x) = \tan x,\ a = 0,\ n = 3$

Find the requested Taylor polynomials in Exercises 21–26 by using known Taylor or Maclaurin polynomials and changing variables as in Examples 6–8.

21. $P_3(x)$ for e^{3x} about $x = -1$.

22. $P_8(x)$ for e^{-x^2} about $x = 0$.

23. $P_4(x)$ for $\sin^2 x$ about $x = 0$. *Hint:* $\sin^2 x = \dfrac{1-\cos(2x)}{2}$.

24. $P_5(x)$ for $\sin x$ about $x = \pi$.

25. $P_6(x)$ for $1/(1+2x^2)$ about $x = 0$

26. $P_8(x)$ for $\cos(3x-\pi)$ about $x = 0$.

27. Find all Maclaurin polynomials $P_n(x)$ for $f(x) = x^3$.

28. Find all Taylor polynomials $P_n(x)$ for $f(x) = x^3$ at $x = 1$.

29. Find the Maclaurin polynomial $P_{2n+1}(x)$ for $\sinh x$ by suitably combining polynomials for e^x and e^{-x}.

30. By suitably combining Maclaurin polynomials for $\ln(1+x)$ and $\ln(1-x)$, find the Maclaurin polynomial of order $2n+1$ for $\tanh^{-1}(x) = \dfrac{1}{2}\ln\left(\dfrac{1+x}{1-x}\right)$.

31. Write Taylor's formula for $f(x) = e^{-x}$ with $a = 0$, and use it to calculate $1/e$ to 5 decimal places. (You may use a calculator but not the e^x function on it.)

! **32.** Write the general form of Taylor's formula for $f(x) = \sin x$ at $x = 0$ with Lagrange remainder. How large need n be taken to ensure that the corresponding Taylor polynomial approximation will give the sine of 1 radian correct to 5 decimal places?

33. What is the best order 2 approximation to the function $f(x) = (x-1)^2$ at $x = 0$? What is the error in this approximation? Now answer the same questions for $g(x) = x^3 + 2x^2 + 3x + 4$. Can the constant $1/6 = 1/3!$, in the error formula for the degree 2 approximation, be improved (i.e., made smaller)?

34. By factoring $1 - x^{n+1}$ (or by long division), show that

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

Next, show that if $|x| \le K < 1$, then

$$\left|\frac{x^{n+1}}{1-x}\right| \le \frac{1}{1-K}|x^{n+1}|.$$

This implies that $x^{n+1}/(1-x) = O(x^{n+1})$ as $x \to 0$ and confirms formula (d) of Table 5. What does Theorem 13 then say about the nth-order Maclaurin polynomial for $1/(1-x)$?

35. By differentiating identity (*) in Exercise 34 and then replacing n with $n+1$, show that

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

Then use Theorem 13 to determine the nth-order Maclaurin polynomial for $1/(1-x)^2$.

4.11 Roundoff Error, Truncation Error, and Computers

In Section 4.7 we introduced the idea of **roundoff error**, while in Sections 4.9 and 4.10 we discussed the result of approximating a function by its Taylor polynomials. The resulting error here is known as **truncation error.** This conventional terminology may be a bit confusing at first because rounding off is itself a kind of truncation of the digital representation of a number. However in numerical analysis "truncation" is reserved for discarding higher order terms, typically represented by big-O, often leaving a Taylor polynomial.

Truncation error is a crucial source of error in using computers to do mathematical operations. In computation with computers, many of the mathematical functions and structures being investigated are approximated by polynomials in order to make it possible for computers to manipulate them. However, the other source of error, roundoff, is ubiquitous, so it is inevitable that mathematics on computers has to involve consideration of both sources of error. These sources can sometimes be treated independently, but in other circumstances they can interact with each other in fascinating ways. In this section we look at some of these fascinating interactions in the form of numerical monsters using Maple. Of course, as stated previously, the issues concern all calculation on computers and not Maple in particular.

Taylor Polynomials in Maple

In much of the following discussion we will be examining the function $\sin x$. Let us begin by defining the Maple expression `s := sin(x)` to denote this function. The Maple input

```
> u := taylor(s, x=0, 5);
```

produces the Taylor polynomial of degree 4 about $x = 0$ (i.e., a Maclaurin polynomial) for $\sin(x)$ together with a big-O term of order x^5:

$$u := x - \frac{1}{6}x^3 + O(x^5)$$

The presence of the big-O term means that u is an actual representation of $\sin x$; there is no error involved. If we want to get an actual Taylor polynomial, we need to convert the expression for u to drop off the big-O term. Since the coefficient of x^4 is zero, let us call the resulting polynomial P_3:

```
> P3 := convert(u, polynom);
```

$$P3 := x - \frac{1}{6}x^3$$

Unlike u, P_3 is not an exact representation of $\sin x$; it is only an approximation. The discarded term $O(x^5) = s - P_3 = u - P_3$ is the error in this approximation. On the basis of the discussion in the previous section, this truncation error can be expected to be quite small for x close to 0, a fact that is confirmed by the Maple plot in Figure 4.66(a). The behaviour is much as expected. $\sin x$ behaves like the cubic polynomial near 0 (so the difference is nearly 0), while farther from 0 the cubic term dominates the expression.

```
> plot(s-P3, x=-1..1, style=point,
   symbol=point, numpoints=1000);
```

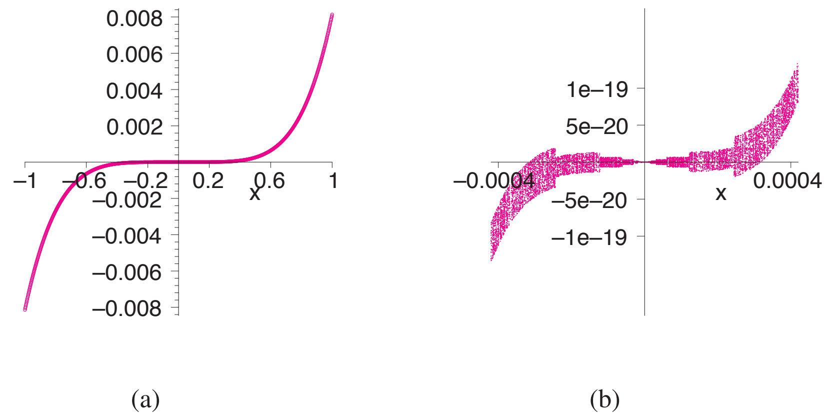


Figure 4.66 The error $\sin x - P_3(x)$ over (a) the interval $[-1, 1]$, and (b) the interval $[-4.2 \times 10^{-4}, 4.2 \times 10^{-4}]$

The limiting behaviour near 0 can be explored by changing the plot window. If the Maple plot instruction is revised to

```
> plot(s-P3, x=-0.42e-3..0.42e-3, style=point,
   symbol=point, numpoints=1000);
```

the plot in Figure 4.66(b) results. What is this structure? Clearly the distances from the x-axis are very small, and one can see the cubic-like behaviour. But why are the points not distributed along a single curve, filling out a jagged arrow-like structure instead? This is another numerical monster connected to roundoff error, as we can see if we plot $\sin(x) - P_3(x)$ together with the functions $\pm(\epsilon/2)\sin(x)$ and $\pm(\epsilon/4)\sin x$, where $\epsilon = 2^{-52}$ is machine epsilon, as calculated in Section 4.7.

```
> eps := evalf(2^(-52)):
> plot([s-P3, -eps*s/2, eps*s/2, -eps*s/4, eps*s/4],
  x=-0.1e-3,0.1e-3, colour=[magenta,grey,grey,black,black],
  style=point, symbol=point, numpoints=1000);
```

The result is in Figure 4.67. The black and grey envelope curves (which appear like straight lines since the plot window is so close to the origin) link the structure of the plot to machine epsilon; the seemingly random points are not as random as they first seemed.

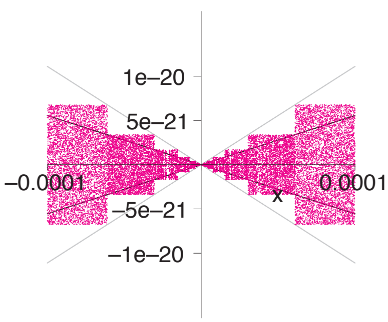


Figure 4.67 Examining the structure of the Maple plot of $\sin x - P_3(x)$ for x in $[-0.0001, 0.0001]$. Note the relationship to the envelope curves $y = \pm(\epsilon/4)\sin x$ (black), and $y = \pm(\epsilon/2)\sin x$ (grey)

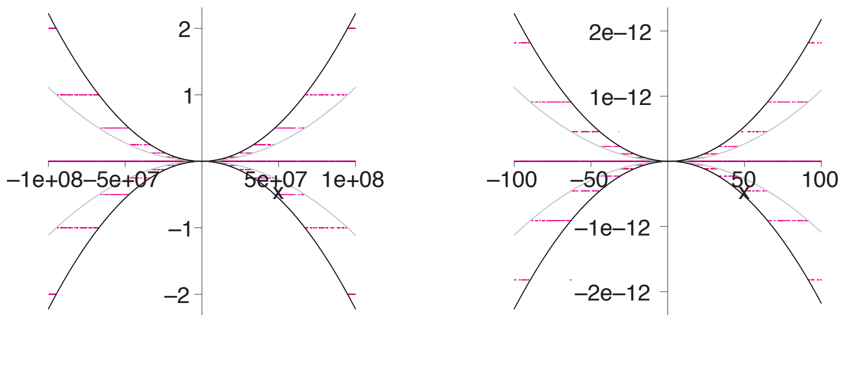
Moreover, this structure is distinctive to Maple. Other software packages, such as Matlab, produce a somewhat different, but still spurious, structure for the same plotting window. Try some others. If different software produces different behaviour under the same instructions, it is certain that some type of computational error is involved. Software-dependent behaviour is one sure sign of computational error.

A distinctive aspect of this monster is that for a large plot window, the truncation error dominates, while near zero, where the truncation error approaches zero, the roundoff error dominates. This is a common relationship between truncation error and roundoff error. However, the roundoff error shows up for plot windows near zero, while the truncation error is dominant over wide ranges of plot windows. Is this always true for truncation error? No—as the next monster shows.

Persistent Roundoff Error

The trade-off between truncation error and roundoff error is distinctive, but one should not get the impression that roundoff error only matters in extreme limiting cases in certain plot windows. Consider, for example, the function $f(x) = x^2 - 2x + 1 - (x - 1)^2$. It is identically 0, not just 0 in the limiting case $x = 0$. However, the computer evaluates the two mathematically equivalent parts of the function f differently, leaving different errors from rounding off the true values of the numbers inserted into the expression. The difference of the result is then not exactly 0. A plot of $f(x)$ on the interval $[-10^8, 10^8]$ is produced by the Maple command

```
> plot([eps*(x-1)^2,eps*(x-1)^2/2,-eps*(x-1)^2,
  -eps*(x-1)^2/2,(x^2-2*x+1)-(x-1)^2],
  x=-1e8..1e8,numpoints=1500,style=point,symbol=point,
  color=[black,grey,black,grey,magenta],
  tickmarks=[[-1e8,-5e7,5e7,1e8],[-2,-1,1,2]]);
```



(a) (b)

Figure 4.68 The values of $x^2 - 2x + 1 - (x-1)^2$ (colour) lie between the parabolas $\pm\epsilon(x-1)^2$ (black) and $\pm\epsilon(x-1)^2/2$ (grey) for (a) $-10^8 \le x \le 10^8$, and (b) $-100 \le x \le 100$

It is shown in Figure 4.68(a). The spurious values of $f(x)$ seem like rungs on a ladder. Note that these false nonzero values of $f(x)$ (colour) are not small compared to 1. This is because the window is so wide. But the error is clearly due to roundoff as the grey and black envelope curves are proportional to machine epsilon. This plot is largely independent of the width of the window chosen. Figure 4.68(b) is the same plot with a window one million times narrower. Except for a change of scale, it is virtually identical to the plot in Figure 4.68(a). This behaviour is quite different from the numerical monster involving Taylor polynomials encountered above.

Truncation, Roundoff, and Computer Algebra

One of the more modern developments in computer mathematics is the computer's ability to deal with mathematics symbolically. This important capability is known as "computer algebra." For example, Maple can generate Taylor expansions of very high order. This might appear to make the issue of error less important. If one can generate exact Taylor polynomials of very high order, how could error remain an issue?

To see how the finiteness of computers intrudes on our calculations in this case too, let us consider the Taylor (Maclaurin) polynomial of degree 99 for $\sin x$:

```
> v := taylor(s, x=0, 100): P99 := convert(v, polynom):
```

It is good to suppress the output here; each command produces screensfull of output. Figure 4.69 shows the result of the Maple plot command

```
> plot([P99,s],x=35..39,y=-3..3,colour=[magenta,black],
  style=point,symbol=point,numpoints=500,
  xtickmarks=[36,37,38,39]);
```

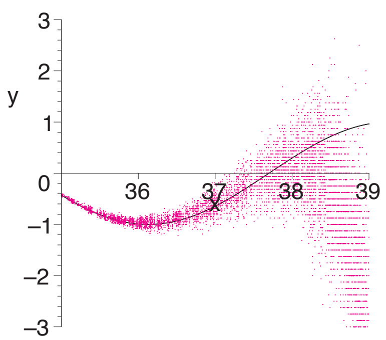


Figure 4.69 The coloured cloud results from Maple's attempt to evaluate the polynomial $P_{99}(x)$ at 500 values of x between 35 and 39

The black curve is the graph of the sine function, and the colour tornado-like cloud is the plot of $P_{99}(x)$ that Maple produces. For plotting, the polynomial must be evaluated at specific values of x. The algorithm cannot employ the large rational expressions for coefficients and high powers of input values. In order to place the result into an actual pixel on the computer screen, the value of the polynomial must be converted to a floating-point number. Then, with the adding and subtracting of 100 terms involving rounded powers, roundoff error returns despite the exact polynomial that we began with.

Of course, there are often tactics to fix these types of problems, but the only way to know what the problems are that need fixing is to understand the mathematics in the first place. But this also means that careful calculations on computers constitute a full field of modern research, requiring considerable mathematical knowledge.

EXERCISES 4.11

1. Use Maple to repeat the plots of Figure 4.68, except using the mathematically equivalent function $(x-1)^2 - (x^2 - 2x + 1)$. Does the result look the same? Is the result surprising?

2. Use Maple to graph $f - P_4(x)$ where $f(x) = \cos x$ and $P_4(x)$ is the 4th degree Taylor polynomial of f about $x = 0$. Use the interval $[-10^{-3}/2, 10^{-3}/2]$ for the plot and plot 1000 points. On the same plot, graph $\pm\epsilon f/2$ and $\pm\epsilon f/4$, where ϵ is machine epsilon. How does the result differ from what is expected mathematically?

! 3. If a real number x is represented on a computer, it is replaced by a floating-point number $F(x)$; x is said to be "floated" by the function F. Show that the relative error in floating for a base-two machine satisfies

$$|\text{error}| = |x - F(x)| \le \epsilon |x|,$$

where $\epsilon = 2^{-t}$ and t is the number of base-two digits (bits) in the floating-point number.

! 4. Consider two different but mathematically equivalent expressions, having the value C after evaluation. On a computer, with each step in the evaluation of each of the expressions, roundoff error is introduced as digits are discarded and rounded according to various rules. In subsequent steps, resulting error is added or subtracted according to the details of the expression producing a final error that depends in detail on the expression, the particular software package, the operating system, and the machine hardware. Computer errors are not equivalent for the two expressions, even when the expressions are mathematically equivalent.

   (a) If we suppose that the computer satisfactorily evaluates the expressions for many input values within an interval, all to within machine precision, why might we expect the difference of these expressions on a computer to have an error contained within an interval $[-\epsilon C, \epsilon C]$?

   (b) Is it possible for exceptional values of the error to lie outside that interval in some cases? Why?

   (c) Is it possible for the error to be much smaller than the interval indicates? Why?

CHAPTER REVIEW

Key Ideas

- **What do the following words, phrases, and statements mean?**
  - ◇ critical point of f
  - ◇ singular point of f
  - ◇ inflection point of f
  - ◇ f has absolute maximum value M
  - ◇ f has a local minimum value at $x = c$
  - ◇ vertical asymptote
  - ◇ horizontal asymptote
  - ◇ oblique asymptote
  - ◇ machine epsilon
  - ◇ the linearization of $f(x)$ about $x = a$
  - ◇ the Taylor polynomial of degree n of $f(x)$ about $x = a$
  - ◇ Taylor's formula with Lagrange remainder
  - ◇ $f(x) = O\big((x-a)^n\big)$ as $x \to a$
  - ◇ a root of $f(x) = 0$
  - ◇ a fixed point of $f(x)$
  - ◇ an indeterminate form
  - ◇ l'Hôpital's Rules
- **Describe how to estimate the error in a linear (tangent line) approximation to the value of a function.**
- **Describe how to find a root of an equation $f(x) = 0$ by using Newton's Method. When will this method work well?**

Review Exercises

1. If the radius r of a ball is increasing at a rate of 2 percent per minute, how fast is the volume V of the ball increasing?

2. **(Gravitational attraction)** The gravitational attraction of the earth on a mass m at distance r from the centre of the earth is a continuous function of r for $r \ge 0$, given by

$$F = \begin{cases} \dfrac{mgR^2}{r^2} & \text{if } r \ge R \\ mkr & \text{if } 0 \le r < R, \end{cases}$$

where R is the radius of the earth, and g is the acceleration due to gravity at the surface of the earth.

   (a) Find the constant k in terms of g and R.

   (b) F decreases as m moves away from the surface of the earth, either upward or downward. Show that F decreases as r increases from R at twice the rate at which F decreases as r decreases from R.

3. **(Resistors in parallel)** Two variable resistors R_1 and R_2 are connected in parallel so that their combined resistance R is given by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}.$$

At an instant when $R_1 = 250$ ohms and $R_2 = 1{,}000$ ohms, R_1 is increasing at a rate of 100 ohms/min. How fast must R_2 be changing at that moment (a) to keep R constant? and (b) to enable R to increase at a rate of 10 ohms/min?

4. **(Gas law)** The volume V (in m^3), pressure P (in kilopascals, kPa), and temperature T (in kelvin, K) for a sample of a certain gas satisfy the equation $pV = 5.0T$.
   (a) How rapidly does the pressure increase if the temperature is 400 K and increasing at 4 K/min while the gas is kept confined in a volume of 2.0 m^3?
   (b) How rapidly does the pressure decrease if the volume is 2 m^3 and increases at 0.05 m^3/min while the temperature is kept constant at 400 K?

5. **(The size of a print run)** It costs a publisher \$10,000 to set up the presses for a print run of a book and \$8 to cover the material costs for each book printed. In addition, machinery servicing, labour, and warehousing add another $\$6.25 \times 10^{-7} x^2$ to the cost of each book if x copies are manufactured during the printing. How many copies should the publisher print in order to minimize the average cost per book?

6. **(Maximizing profit)** A bicycle wholesaler must pay the manufacturer \$75 for each bicycle. Market research tells the wholesaler that if she charges her customers \$$x$ per bicycle, she can expect to sell $N(x) = 4.5 \times 10^6/x^2$ of them. What price should she charge to maximize her profit, and how many bicycles should she order from the manufacturer?

7. Find the largest possible volume of a right-circular cone that can be inscribed in a sphere of radius R.

8. **(Minimizing production costs)** The cost \$$C(x)$ of production in a factory varies with the amount x of product manufactured. The cost may rise sharply with x when x is small, and more slowly for larger values of x because of economies of scale. However, if x becomes too large, the resources of the factory can be overtaxed, and the cost can begin to rise quickly again. Figure 4.70 shows the graph of a typical such cost function $C(x)$.

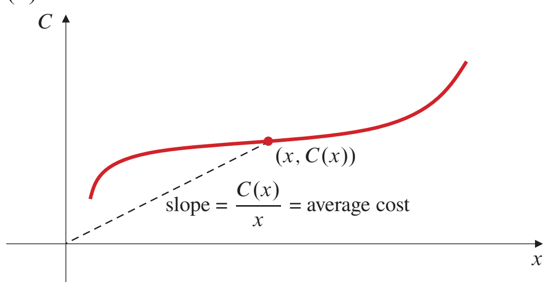


Figure 4.70

If x units are manufactured, the average cost per unit is \$$C(x)/x$, which is the slope of the line from the origin to the point $(x, C(x))$ on the graph.
   (a) If it is desired to choose x to minimize this average cost per unit (as would be the case if all units produced could be sold for the same price), show that x should be chosen to make the average cost equal to the marginal cost:
   $$\frac{C(x)}{x} = C'(x).$$
   (b) Interpret the conclusion of (a) geometrically in the figure.
   (c) If the average cost equals the marginal cost for some x, does x necessarily minimize the average cost?

9. **(Box design)** Four squares are cut out of a rectangle of cardboard 50 cm by 80 cm, as shown in Figure 4.71, and the remaining piece is folded into a closed, rectangular box, with two extra flaps tucked in. What is the largest possible volume for such a box?

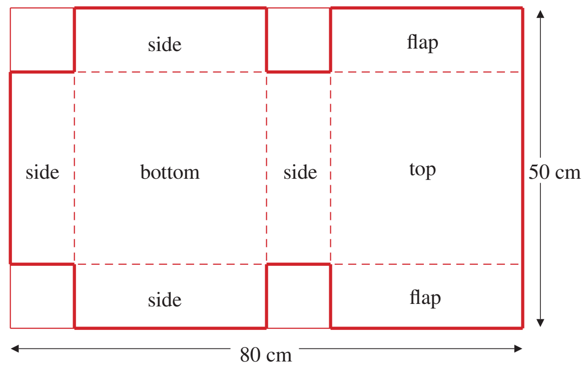


Figure 4.71

10. **(Yield from an orchard)** A certain orchard has 60 trees and produces an average of 800 apples per tree per year. If the density of trees is increased, the yield per tree drops; for each additional tree planted, the average yield per tree is reduced by 10 apples per year. How many more trees should be planted to maximize the total annual yield of apples from the orchard?

11. **(Rotation of a tracking antenna)** What is the maximum rate at which the antenna in Exercise 41 of Section 4.1 must be able to turn in order to track the rocket during its entire vertical ascent?

12. An oval table has its outer edge in the shape of the curve $x^2 + y^4 = 1/8$, where x and y are measured in metres. What is the width of the narrowest hallway in which the table can be turned horizontally through 180°?

13. A hollow iron ball whose shell is 2 cm thick weighs half as much as it would if it were solid iron throughout. What is the radius of the ball?

14. **(Range of a cannon fired from a hill)** A cannon ball is fired with a speed of 200 ft/s at an angle of 45° above the horizontal from the top of a hill whose height at a horizontal distance x ft from the top is $y = 1{,}000/(1 + (x/500)^2)$ ft above sea level. How far does the cannon ball travel horizontally before striking the ground?

15. **(Linear approximation for a pendulum)** Because $\sin\theta \approx \theta$ for small values of $|\theta|$, the nonlinear equation of motion of a simple pendulum
$$\frac{d^2\theta}{dt^2} = -\frac{g}{L}\sin\theta,$$
which determines the displacement angle $\theta(t)$ away from the vertical at time t for a simple pendulum, is frequently approximated by the simpler linear equation
$$\frac{d^2\theta}{dt^2} = -\frac{g}{L}\theta$$
when the maximum displacement of the pendulum is not large. What is the percentage error in the right side of the equation if $|\theta|$ does not exceed 20°?

16. Find the Taylor polynomial of degree 6 for $\sin^2 x$ about $x = 0$ and use it to help you evaluate
$$\lim_{x\to 0} \frac{3\sin^2 x - 3x^2 + x^4}{x^6}.$$

17. Use a second-order Taylor polynomial for $\tan^{-1} x$ about $x = 1$ to find an approximate value for $\tan^{-1}(1.1)$. Estimate the size of the error by using Taylor's formula.

18. The line $2y = 10x - 19$ is tangent to $y = f(x)$ at $x = 2$. If an initial approximation $x_0 = 2$ is made for a root of $f(x) = 0$ and Newton's Method is applied once, what will be the new approximation that results?

19. Find all solutions of the equation $\cos x = (x - 1)^2$ to 10 decimal places.

20. Find the shortest distance from the point $(2, 0)$ to the curve $y = \ln x$.

21. A car is travelling at night along a level, curved road whose equation is $y = e^x$. At a certain instant its headlights illuminate a signpost located at the point $(1, 1)$. Where is the car at that instant?

Challenging Problems

1. (Growth of a crystal) A single cubical salt crystal is growing in a beaker of salt solution. The crystal's volume V increases at a rate proportional to its surface area and to the amount by which its volume is less than a limiting volume V_0:

$$\frac{dV}{dt} = kx^2(V_0 - V),$$

where x is the edge length of the crystal at time t.

(a) Using $V = x^3$, transform the equation above to one that gives the rate of change dx/dt of the edge length x in terms of x.

(b) Show that the growth rate of the edge of the crystal decreases with time but remains positive as long as $x < x_0 = V_0^{1/3}$.

(c) Find the volume of the crystal when its edge length is growing at half the rate it was initially.

2. (A review of calculus!) You are in a tank (the military variety) moving down the y-axis toward the origin. At time $t = 0$ you are 4 km from the origin, and 10 min later you are 2 km from the origin. Your speed is decreasing; it is proportional to your distance from the origin. You know that an enemy tank is waiting somewhere on the positive x-axis, but there is a high wall along the curve $xy = 1$ (all distances in kilometres) preventing you from seeing just where it is. How fast must your gun turret be capable of turning to maximize your chances of surviving the encounter?

3. (The economics of blood testing) Suppose that it is necessary to perform a blood test on a large number N of individuals to detect the presence of a virus. If each test costs \$$C$, then the total cost of the testing program is \$$NC$. If the proportion of people in the population who have the virus is not large, this cost can be greatly reduced by adopting the following strategy. Divide the N samples of blood into N/x groups of x samples each. Pool the blood in each group to make a single sample for that group and test it. If it tests negative, no further testing is necessary for individuals in that group. If the group sample tests positive, test all the individuals in that group.

Suppose that the fraction of individuals in the population infected with the virus is p, so the fraction uninfected is $q = 1 - p$. The probability that a given individual is unaffected is q, so the probability that all x individuals in a group are unaffected is q^x. Therefore, the probability that a pooled sample is infected is $1 - q^x$. Each group requires one test, and the infected groups require an extra x tests. Therefore, the expected total number of tests to be performed is

$$T = \frac{N}{x} + \frac{N}{x}(1 - q^x)x = N\left(\frac{1}{x} + 1 - q^x\right).$$

For example, if $p = 0.01$, so that $q = 0.99$ and $x = 20$, then the expected number of tests required is $T = 0.23N$, a reduction of over 75%. But maybe we can do better by making a different choice for x.

(a) For $q = 0.99$, find the number x of samples in a group that minimizes T (i.e., solve $dT/dx = 0$). Show that the minimizing value of x satisfies

$$x = \frac{(0.99)^{-x/2}}{\sqrt{-\ln(0.99)}}.$$

(b) Use the technique of fixed-point iteration (see Section 4.2) to solve the equation in (a) for x. Start with $x = 20$, say.

4. (Measuring variations in g) The period P of a pendulum of length L is given by

$$P = 2\pi\sqrt{L/g},$$

where g is the acceleration of gravity.

(a) Assuming that L remains fixed, show that a 1% increase in g results in approximately a 0.5% decrease in the period P. (Variations in the period of a pendulum can be used to detect small variations in g from place to place on the earth's surface.)

(b) For fixed g, what percentage change in L will produce a 1% increase in P?

5. (Torricelli's Law) The rate at which a tank drains is proportional to the square root of the depth of liquid in the tank above the level of the drain: if $V(t)$ is the volume of liquid in the tank at time t, and $y(t)$ is the height of the surface of the liquid above the drain, then $dV/dt = -k\sqrt{y}$, where k is a constant depending on the size of the drain. For a cylindrical tank with constant cross-sectional area A with drain at the bottom:

(a) Verify that the depth $y(t)$ of liquid in the tank at time t satisfies $dy/dt = -(k/A)\sqrt{y}$.

(b) Verify that if the depth of liquid in the tank at $t = 0$ is y_0, then the depth at subsequent times during the draining process is $y = \left(\sqrt{y_0} - \frac{kt}{2A}\right)^2$.

(c) If the tank drains completely in time T, express the depth $y(t)$ at time t in terms of y_0 and T.

(d) In terms of T, how long does it take for half the liquid in the tank to drain out?

6. If a conical tank with top radius R and depth H drains according to Torricelli's Law and empties in time T, show that the depth of liquid in the tank at time t $(0 < t < T)$ is

$$y = y_0\left(1 - \frac{t}{T}\right)^{2/5},$$

where y_0 is the depth at $t = 0$.

7. Find the largest possible area of a right-angled triangle whose perimeter is P.

8. Find a tangent to the graph of $y = x^3 + ax^2 + bx + c$ that is not parallel to any other tangent.

9. (Branching angles for electric wires and pipes)

(a) The resistance offered by a wire to the flow of electric current through it is proportional to its length and inversely proportional to its cross-sectional area. Thus, the resistance R of a wire of length L and radius r is $R = kL/r^2$, where k is a positive constant. A long straight wire of length L and radius r_1 extends from A to B. A second straight wire of smaller radius r_2 is to be connected between a point P on AB and a point C at distance h from B such that CB is perpendicular to AB. (See Figure 4.72.) Find the value of the angle $\theta = \angle BPC$ that minimizes the total resistance of the path APC, that is, the resistance of AP plus the resistance of PC.

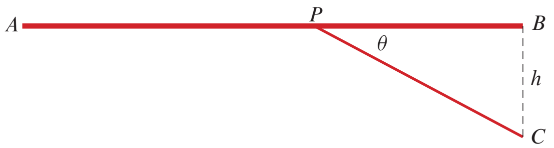


Figure 4.72

(b) The resistance of a pipe (e.g., a blood vessel) to the flow of liquid through it is, by Poiseuille's Law, proportional to its length and inversely proportional to the *fourth power* of its radius: $R = kL/r^4$. If the situation in part (a) represents pipes instead of wires, find the value of θ that minimizes the total resistance of the path APC. How does your answer relate to the answer for part (a)? Could you have predicted this relationship?

10. (The range of a spurt) A cylindrical water tank sitting on a horizontal table has a small hole located on its vertical wall at height h above the bottom of the tank. Water escapes from the tank horizontally through the hole and then curves down under the influence of gravity to strike the table at a distance R from the base of the tank, as shown in Figure 4.73. (We ignore air resistance.) Torricelli's Law implies that the speed v at which water escapes through the hole is proportional to the square root of the depth of the hole below the surface of the water: if the depth of water in the tank at time t is $y(t) > h$, then $v = k\sqrt{y - h}$, where the constant k depends on the size of the hole.

(a) Find the range R in terms of v and h.

(b) For a given depth y of water in the tank, how high should the hole be to maximize R?

(c) Suppose that the depth of water in the tank at time $t = 0$ is y_0, that the range R of the spurt is R_0 at that time, and that the water level drops to the height h of the hole in T minutes. Find, as a function of t, the range R of the water that escaped through the hole at time t.

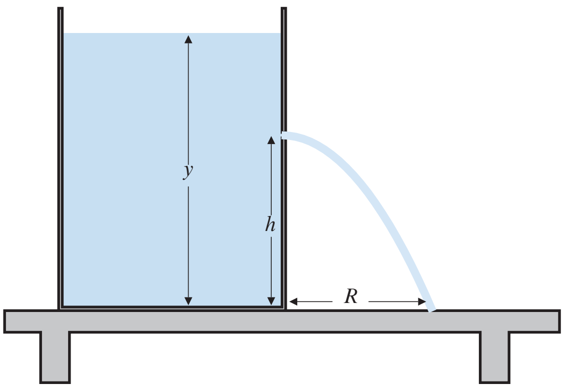


Figure 4.73

11. (Designing a dustpan) Equal squares are cut out of two adjacent corners of a square of sheet metal having sides of length 25 cm. The three resulting flaps are bent up, as shown in Figure 4.74, to form the sides of a dustpan. Find the maximum volume of a dustpan made in this way.

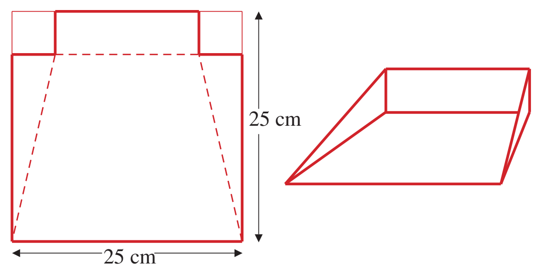


Figure 4.74