6.1 Optimization

Many important applied problems involve finding the best way to accomplish some task. Often this involves finding the maximum or minimum value of some function: the minimum time to make a certain journey, the minimum cost for doing a task, the maximum power that can be generated by a device, and so on. Many of these problems can be solved by finding the appropriate function and then using techniques of calculus to find the maximum or the minimum value required.

Generally such a problem will have the following mathematical form: Find the largest (or smallest) value of \( f(x) \) when \( a \leq x \leq b \). Sometimes \( a \) or \( b \) are infinite, but frequently the real world imposes some constraint on the values that \( x \) may have.

Such a problem differs in two ways from the local maximum and minimum problems we encountered when graphing functions. We are interested only in the function between \( a \) and \( b \), and we want to know the largest or smallest value that \( f(x) \) takes on, not merely values that are the largest or smallest in a small interval. That is, we seek not a local maximum or minimum but a global maximum or minimum.

Any global maximum or minimum must of course be a local maximum or minimum. If we find all possible local extrema, then the global maximum, if it exists, must be the largest of the local maxima and the global minimum, if it exists, must be the smallest of the local minima. We already know where local extrema can occur: only at those points at which \( f'(x) \) is zero or undefined. Actually, there are two additional points at which a maximum or minimum can occur if the endpoints \( a \) and \( b \) are not infinite, namely, at \( a \) and \( b \). We have not previously considered such points because we have not been interested in limiting a function to a small interval. An example may make this clear.

**EXAMPLE 6.1.1**

Find the maximum and minimum values of \( f(x) = x^3 \) on the interval \([-2, 1]\), shown in figure 6.1.1. We compute \( f'(x) = 3x^2 \), which is zero at \( x = 0 \) and is always defined.

Since \( f'(1) = 2 \) we would not normally flag \( x = 1 \) as a point of interest, but it is clear from the graph that when \( f(x) \) is restricted to \([-2, 1]\) there is a local maximum at \( x = 1 \). Likewise we would not normally pay attention to \( x = -2 \), but since we have truncated \( f \) at \(-2\) we have introduced a new local maximum there as well. In a technical sense nothing new is going on here: When we truncate \( f \) we actually create a new function, let’s call it \( g \), that is defined only on the interval \([-2, 1]\). If we try to compute the derivative of this new function we actually find that it does not have a derivative at \(-2 \) or \( 1 \). Why? Because to compute the derivative at \( 1 \) we must compute the limit

\[
\lim_{\Delta x \to 0} \frac{g(1 + \Delta x) - g(1)}{\Delta x}
\]

This limit does not exist because when \( \Delta x > 0 \), \( g(1 + \Delta x) \) is not defined. It is simpler, however, simply to remember that we must always check the endpoints.

So the function \( g \), that is, \( f \) restricted to \([-2, 1]\), has one critical value and two finite endpoints, any of which might be the global maximum or minimum. We could first determine which of these are local maximum or minimum points (or neither); then the largest local maximum must be the global maximum and the smallest local minimum must be the global minimum. It is usually easier, however, to compute the value of \( f \) at every point at which the global maximum or minimum might occur: the largest of these is the global maximum, the smallest is the global minimum.

So we compute \( f(-2) = -4 \), \( f(0) = 0 \), \( f(1) = 1 \). The global maximum is 4 at \( x = -2 \) and the global minimum is 0 at \( x = 0 \).

**EXAMPLE 6.1.2**

Find all local maxima and minima for \( f(x) = x^3 - x \) on the interval \([-2, 2]\), shown in figure 6.1.2. We have

\[
\begin{align*}
\text{Critical points:} & \quad f'(x) = 3x^2 - 1 = 0 \quad \Rightarrow \quad x = \pm \sqrt{\frac{1}{3}} \\
\text{Endpoints:} & \quad f(-2) = -6, \quad f(0) = 0, \quad f(2) = 6
\end{align*}
\]

Since \( f'(x) \) is zero at \( x = \pm \sqrt{\frac{1}{3}} \) and undefined at \( x = 0 \), we compute \( f'(x) \) at \( x = \pm \sqrt{\frac{1}{3}} \). The critical points are \( x = \pm \sqrt{\frac{1}{3}} \). The endpoint is \( x = 0 \).

**EXAMPLE 6.1.3**

Find all local maxima and minima for \( f(x) = x^3 - x \), and determine whether there is a global maximum or minimum on the open interval \((-2, 2)\). In example 6.1.2 we found a local maximum at \((-\sqrt{\frac{1}{3}}, 2\sqrt{\frac{1}{3}})\) and a local minimum at \((\sqrt{\frac{1}{3}}, -2\sqrt{\frac{1}{3}})\). Since the endpoints are not in the interval \((-2, 2)\) they cannot be con-
EXAMPLE 6.1.8  You want to sell a certain number of items in order to maximize your profit. Market research tells you that if you set the price at $1.50, you will be able to sell 5000 items, and for every 10 cents you lower the price below $1.40 you will be able to sell another 1000 items. Suppose that your fixed costs (‘start-up costs’) total $2800, and the per item cost of production (‘marginal cost’) is $0.50. Find the price to set per item in order to maximize profit, and also determine the maximum profit you can get.

The first step is to convert the problem into a function maximization problem. Since we want to maximize profit by setting the price per item, we should look for a function $P(x)$ representing the profit when the price per item is $x$. Profit is revenue minus costs, and revenue is number of items sold times the price per item, so we get $P = nx - 2000 - 0.5nx$. The number of items sold is itself a function of $x$, $n = 5000 + 1000(1.5 - x)/0.10$, because $(1.5 - x)/0.10$ is the number of 10 cents of 10 cents that the price is below $1.50. Now we substitute for $n$ in the profit function:

$$P(x) = (5000 + 1000(1.5 - x)/0.10)x - 2000 - 0.5(5000 + 1000(1.5 - x)/0.10)$$

$$= -1000x^2 + 25000x - 12000$$

We want to know the maximum value of this function when $x$ is between 0 and 1.5. The derivative is $P'(x) = -20000x + 25000$, which is zero when $x = 1.25$. Since $P''(x) = -20000 < 0$, there must be a local maximum at $x = 1.25$, and since this is the only critical value it must be a global maximum as well. (Alternately, we could compute $P(0) = -12000$, $P(1.25) = 3625$, and $P(1.5) = 3000$ and note that $P(1.25)$ is the maximum of these.) Thus the maximum profit is $3625, attained when we set the price at $1.25 and sell 7500 items.

EXAMPLE 6.1.9  Find the largest rectangle (that is, the rectangle with largest area) that fits inside the graph of the parabola $y = x^2$ below the line $y = a$ (is an unspecified constant value), with the top side of the rectangle on the horizontal line $y = a$; see figure 6.1.3.

We want to find the maximum value of some function $A(x)$ representing area. Perhaps the hardest part of this problem is deciding what $x$ should represent. The lower right corner of the rectangle is at $(x, x^2)$, and once this is chosen the rectangle is completely determined. So we can let the $x$ in $A(x)$ be the $x$ of the parabola $f(x) = x^2$. Then the area is $A(x) = (2x)(x^2 - x^2 - 2x^2 + 2x^2)$. We want the maximum value of $A(x)$ when $x$ is in $[0, \sqrt{a}]$. (You might object to allowing $x = 0$ or $x = \sqrt{a}$, since then the “rectangle” has neither a width nor a height, so is not “really” a rectangle. But the problem is somewhat easier if we simply allow such rectangles, which have zero area.)

Setting $0 = A'(x) = -4x^2 + 2x$ we get $x = \sqrt{1/2}$ as the only critical value. Testing this and the two endpoints, we have $A(0) = A(\sqrt{a}) = 0$ and $A(\sqrt{1/2}) = \sqrt{a}/4\sqrt{a}/2$. The maximum area thus occurs when the rectangle has dimensions $\sqrt{a}/\sqrt{3} \times 2\sqrt{a}$.  

EXAMPLE 6.1.10  If you fit the largest possible cone inside a sphere, what fraction of the volume of the sphere is occupied by the cone? (Here by “cone” we mean a right circular cone, i.e., a cone for which the base is perpendicular to the axis of symmetry, and for which the cross-section cut perpendicular to the axis of symmetry at any point is a circle.)

Let $R$ be the radius of the sphere, and let $r$ and $h$ be the base radius and height of the cone inside the sphere. What we want to maximize is the volume of the cone: $\pi r^2 h/3$. Here $R$ is a fixed value, but $r$ and $h$ can vary. Namely, we could choose $r$ to be as large as possible—equal to $R$—by taking the height equal to $R$; or we could make the cone’s height $h$ larger at the expense of making $r$ a little less than $R$. See the cross-section depicted in figure 6.1.4. We have situated the picture in a convenient way relative to the $x$ and $y$ axes, namely, with the center of the sphere at the origin and the vertex of the cone at the far left on the $x$-axis.

Notice that the function we want to maximize, $\pi r^2 h/3$, depends on two variables. This is frequently the case, but often the two variables are related in some way so that “really” there is only one variable. So our next step is to find the relationship and use it to solve for one of the variables in terms of the other, so as to have a function of only one variable to maximize. In this problem, the condition is apparent in the figure: the upper corner of the triangle, whose coordinates are $(h - R, r)$, must be on the circle of radius $R$. That is, $$(h - R)^2 + r^2 = R^2.$$ We can solve for $h$ in terms of $r$ or for $r$ in terms of $h$. Either involves taking a square root, but we notice that the volume function contains $r^2$, not $r$ by itself, so it is easiest to solve for $r^2$ directly: $r^2 = R^2 - (h - R)^2$. Then we substitute the result into $\pi r^2 h/3$.

$$V(h) = \pi r^2 h = \pi \left( R^2 - (h - R)^2 \right) h/3$$

$$= \pi \frac{r^2}{3} + \frac{2}{3} \pi R h$$

We want to maximize $V(h)$ when $h$ is between 0 and $2R$. Now we solve $0 = f'(h) = -\pi R^2 + (4/3)\pi R h$, getting $h = 0$ or $h = 4R/3$. We compute $V(0) = V(2R) = 0$ and $V(4R/3) = (32/81)\pi R^3$. The maximum is the latter, since the volume of the sphere is $(4/3)\pi R^3$, the fraction of the sphere occupied by the cone is

$$\frac{(32/81)\pi R^3}{(4/3)\pi R^3} = \frac{8}{27} \approx 30\%.$$  

EXAMPLE 6.1.11  You are making cylindrical containers to contain a given volume. Suppose that the top and bottom are made of a material that is $N$ times as expensive (cost per unit area) as the material used for the lateral side of the cylinder. Find (in terms of $N$) the ratio of height to base radius of the cylinder that minimizes the cost of making the containers.

Let us first choose letters to represent various things: $h$ for the height, $r$ for the base radius, $V$ for the volume of the cylinder, and $c$ for the cost per unit area of the lateral side of the cylinder; $V$ and $c$ are constants, $h$ and $r$ are variables. Now we can write the cost of materials:

$$c(2\pi rh) + N(2\pi r^2).$$

Again we have two variables; the relationship is provided by the fixed volume of the cylinder: $V = \pi r^2 h$. We use this relationship to eliminate $h$ (we could eliminate $r$, but it’s a little easier if we eliminate $h$, which appears in only one place in the above formula for cost).

The result is

$$f(r) = 2\pi r V/\pi r^2 + 2Nc\pi r^2 = \frac{2V}{r} + 2Nc\pi r^2.$$  

We want to know the minimum value of this function when $r$ is in $(0, \infty)$. We now set $0 = f'(r) = -2V/r^2 + 4Ncr$, giving $r = \sqrt{V/4Nc}$. Since $f''(r) = 4V/r^3 + 4Nc$ is positive when $r$ is positive, there is a local minimum at the critical value, and hence a global minimum since there is only one critical value.

Finally, since $h = V/(\pi r^2)$,

$$\frac{h}{r} = \frac{V}{\pi r^2} = \frac{V}{(V/(2\pi r^2))} = 2N,$$

so the minimum cost occurs when the height $h$ is $2N$ times the radius. If, for example, there is no difference in the cost of materials, the height is twice the radius (or the height is equal to the diameter).
EXAMPLE 6.1.12 Suppose you want to reach a point $A$ that is located across the sand from a nearby road (see figure 6.1.5). Suppose that the road is straight, and $b$ is the distance from $A$ to the closest point $C$ on the road. Let $v$ be your speed on the road, and let $w$, which is less than $v$, be your speed on the sand. Right now you are at the point $B$, which is a distance $a$ from $C$. At what point $B$ should you turn off the road and head across the sand in order to minimize your travel time to $A$?

Let $x$ be the distance short of $C$ where you turn off, i.e., the distance from $B$ to $C$. We want to minimize the total travel time. Recall that when traveling at constant velocity, time is distance divided by velocity.

You travel the distance $Bf$ at speed $v$, and then the distance $fC$ at speed $w$. Since $Bf = a - x$ and, by the Pythagorean theorem, $fC = \sqrt{x^2 + b^2}$, the total time for the trip is

$$t(x) = \frac{a-x}{v} + \frac{x}{w\sqrt{x^2 + b^2}}$$

We want to find the minimum value of $t$ when $x$ is between 0 and $a$. As usual we set $f(x) = 0$ and solve for $x$:

$$0 = f'(x) = \frac{1}{v} - \frac{x}{w\sqrt{x^2 + b^2}} x \left(\frac{x}{w\sqrt{x^2 + b^2}}\right)$$

Notice that $a$ does not appear in the last expression, but $a$ is not irrelevant, since we are interested only in critical values that are in $[0, a]$, and $wb/\sqrt{v^2 - w^2}$ is either in this interval or not. If it is, we can use the second derivative to test it:

$$f''(x) = \frac{2v^2 - x^2}{w^2(x^2 + b^2)^{3/2}}$$

Since this is always positive there is a local minimum at the critical point, and so it is a global minimum as well.

If the critical value is not in $[0, a]$ it is larger than $a$. In this case the minimum must occur at one of the endpoints. We can compute

$$f(0) = \frac{a}{v} \quad f(a) = \frac{b}{w}$$

You want to make cylindrical containers to hold 1 liter (1000 cubic centimeters) using the least amount of construction material. The side is made from a rectangular piece of material, and this can be done with no material wasted. However, the top and bottom are cut from squares of side $2r$, so that $2r^2 = h^2$ of material is needed (rather than $2\pi r^2$, which is the total area of the top and bottom). Find the dimensions of the container using the least amount of material, and also find the ratio of height to radius for this container.

15. You want to make cylindrical containers of a given volume $V$ using the least amount of construction material. The side is made from a rectangular piece of material, and this can be done with no material wasted. However, the top and bottom are cut from squares of side $2r$, so that $2r^2 = h^2$ of material is needed (rather than $2\pi r^2$, which is the total area of the top and bottom). Find the optimal ratio of height to radius.

16. Given a right circular cone, you put an upside-down cone inside it so that its vertex is at the center of the base of the larger cone and its base is parallel to the base of the larger cone. If you choose the upside-down cone to have the largest possible volume, what fraction of the volume of the larger cone does it occupy? (Let $h$ and $R$ be the height and base radius of the larger cone, and let $h$ and $r$ be the height and base radius of the smaller cone. Hint: Use similar triangles to get an equation relating $R$ and $r$.)

17. In example 6.1.12, what happens if $w \geq v$ (i.e., your speed on sand is at least your speed on the road)?

18. A container holding a fixed volume is being made in the shape of a cylinder with a hemispherical top. (The hemispherical top has the same radius as the cylinder.) Find the ratio of height to radius of the cylinder which minimizes the cost of the container if (a) the cost per unit area of the top is twice as great as the cost per unit area of the side, and the container is made with no bottom; (b) the same as in (a), except that the container is made with a circular bottom, for which the cost per unit area is 1.5 times the cost per unit area of the side.

19. A piece of cardboard is 1 meter by 1/2 meter. A square is to be cut from each corner and the sides folded up to make an open-top box. What are the dimensions of the box with maximum possible volume?
6.2 Related Rates

Suppose we have two variables $x$ and $y$ (in most problems the letters will be different, but for now let’s use $x$ and $y$) which are both changing with time. A “related rates” problem is a problem in which we know one of the rates of change at a given instant—in, say, $dx/dt$ for $x$, and $dy/dt$ for $y$. We are interested in the time at which $\frac{dy}{dx}$ is rising after 1 sec; (b) the angular speed of the rope in deg/sec after 1 sec.

Was interested in the time at which $x = 4$; at this time we know that $4^2 + 9 = y^2$, so $y = 5$. Putting together all the information we get

\[
2(\frac{4}{(100)} - \frac{2(3)})
\]

Thus, $y = 400$ mph.

**EXAMPLE 6.2.3** You are flying a spherical balloon at the rate of 7 cm$^3$/sec. How fast is its radius increasing when the radius is 4 cm? Here the variables are the radius $r$ and the volume $V$. We know $\frac{dv}{dt}$, and we want $dr/dt$. The two variables are related by means of the equation $V = 4\pi r^3/3$. Taking the derivative of both sides gives $\frac{dv}{dt} = 4\pi r^2 \cdot \frac{dr}{dt}$. We now substitute the values we know at the instant in question: $T = 4\pi r^2$, so $r = 7/(4\pi)$ cm/sec.

**EXAMPLE 6.2.4** Water is poured into a conical container at the rate of 10 cm$^3$/sec. The cone points directly down, and it has a height of 30 cm and a base radius of 10 cm; see figure 6.2.2. How fast is the water level rising when the water is 4 cm deep (at its deepest point)? The cone points directly down, and it has a height of 30 cm and a base radius of 10 cm; see figure 6.2.2. Without stopping to do any calculus, you are interested in the time at which $\frac{dy}{dx}$ is rising after 1 sec; (b) the angular speed of the rope in deg/sec after 1 sec.

The water forms a conical shape within the big cone; its height and base radius and volume are all increasing as water is poured into the container. This means that we actually have three things varying with time: the water level $h$ (the height of the cone of water), the radius $r$ of the circular top surface of water (the base radius of the cone of water), and the volume of water $V$. The volume of a cone is given by $V = \pi r^2 h/3$. We know $\frac{dv}{dt}$, and we want $\frac{dh}{dt}$. At first something seems to be wrong: we have a third variable $r$ whose rate we don’t know.

But the dimensions of the cone of water must have the same proportions as those of the container. That is, because of similar triangles, $\frac{r}{h} = \frac{10}{30}$ so $r = h/3$. Now we can eliminate $r$ from the problem entirely: $V = \pi (h/3)^2 h/3 = \pi h^3/27$. We take the derivative of both sides and plug in $h = 4$ and $\frac{dv}{dt} = 10$, obtaining $10 = (3\pi \cdot 4^2/27)(\frac{dh}{dt})$. Thus, $\frac{dh}{dt} = 90/(16\pi)$ cm/sec.

**EXAMPLE 6.2.5** A swing consists of a board at the end of a 10 ft long rope. Think of the board as a point $P$ at the end of the rope, and let $Q$ be the point of attachment at the other end. Suppose that the swing is directly below you at time $t = 0$, and is being pushed by someone who walks at 6 ft/sec from left to right. Find (a) how fast the swing is rising after 1 sec; (b) the angular speed of the rope in deg/sec after 1 sec.

We start out by asking: What is the geometric quantity whose rate of change we know, and what is the geometric quantity whose rate of change we’re being asked about? Note that the person pushing the swing is moving horizontally at a rate we know. In other words, the horizontal coordinate of $P$ is increasing at 6 ft/sec. In the xy-plane let us make the convenient choice of putting the origin at the location of $P$ at time $t = 0$, i.e., a distance 10 directly below the point of attachment. Then the rate we know is $\frac{dx}{dt}$, and in part (a) the rate we want is $\frac{dy}{dx}$ (the rate at which $P$ is rising). In part (b) the rate we want is $\theta = \frac{dy}{dx}$, where $\theta$ stands for the angle in radians through which the swing has swung from the vertical. (Actually, since we want our answer in deg/sec, at the end we must convert $\frac{dy}{dx}$ from rad/sec by multiplying by $180/\pi$.)

(a) From the diagram we see that we have a right triangle whose legs are $x = 10$ and $y$, and whose hypotenuse is 10. Hence $x^2 + (10 - y)^2 = 100$. Taking the derivative of both sides we obtain: $2x - 2(10 - y) - y = 0$. We now look at what we know after 1 second, namely $x = 6$ (because $x$ started at 0 and has been increasing at the rate of 6 ft/sec for 1 sec), $y = 2$ (because we get $10 - y = 8$ from the Pythagorean theorem applied to the triangle with hypotenuse 10 and leg 6), and $x = 6$. Putting in these values gives us $2 \cdot 6 - 2 \cdot 2 \cdot y = 0$, from which we can easily solve for $y = 4.5$ ft/sec.

(b) Here our two variables are $x$ and $\theta$, so we want to use the same right triangle as in part (a), but this time relate $\theta$ to $x$. Since the hypotenuse is constant (equal to 10), the best way to do this is to use the sine: $\sin \theta = x/10$. Taking derivatives we obtain

\[
\frac{dy}{dx} = \frac{x}{10}
\]
Car A is driving north along the first road, and car B is driving east along the second road. The angle between the roads is 90°. Car A is traveling at 60 m/sec, and car B is traveling at 80 m/sec. At the instant in question (t = 1 sec), when we have a right triangle with sides 6–8–10, \( \cos \theta = 8/10 \), \( \cos \theta = 0.8 \), and \( \dot{x} = 6 \). Thus \( (8/10)\dot{t} = 6/10, i.e., \( \dot{t} = 6/8 = 3/4 \text{ rad/sec} \), or approximately 43 deg/sec.

We have seen that sometimes there are apparently more than two variables that change with time, but in reality these are just two, as the others can be expressed in terms of just two. But sometimes there really are several variables that change with time; as long as you know the rates of change of all but one of them you can find the rate of change of the remaining one. As in the case when there are just two variables, take the derivative of both sides of the equation relating all of the variables, and then substitute all of the known values and solve for the unknown rate.

**EXAMPLE 6.2.6** A road running north to south crosses a road going east to west at the point \( P \). Car A is driving north along the first road, and car B is driving east along the second road. At a particular time \( t \) car A is 10 kilometers to the north of \( P \) and traveling at 80 km/hr, while car B is 15 kilometers to the east of \( P \) and traveling at 108 km/hr. How fast is the distance between the two cars changing?

**Figure 6.2.4 Cars moving apart.**

\[ \frac{dx}{dt} = \frac{d}{dt}(x^2 + y^2) = 2x \frac{dx}{dt} + 2y \frac{dy}{dt} \]

At the instant in question (when \( x = 10 \) and \( y = 15 \)), we have:

\[ \frac{dx}{dt} = \frac{d}{dt}(x^2 + y^2) = 2x \frac{dx}{dt} + 2y \frac{dy}{dt} = 2(10)(80) + 2(15)(108) = 2600 \text{ km/hr} \]

Therefore, the distance between the two cars is increasing at 2600 km/hr.

**EXERCISES 6.2.**

1. A woman 5 ft tall walks at the rate of 3.5 ft/sec away from a streetlight that is 12 ft above the ground. At what rate is the tip of her shadow moving? At what rate is her shadow lengthening?

2. A man 1.8 meters tall walks at the rate of 1 meter per second toward a streetlight that is 4 meters above the ground. At what rate is the tip of his shadow moving? At what rate is his shadow shortening?

3. A police helicopter is flying at 150 mph at a constant altitude of 0.5 mile above a straight road. The pilot uses radar to determine that an oncoming car is at a distance of exactly 1 mile from the helicopter, and that this distance is decreasing at 180 mph. Find the speed of the car.

4. A police helicopter is flying at 200 kilometers per hour at a constant altitude of 1 km above a straight road. The pilot uses radar to determine that an oncoming car is at a distance of exactly 2 kilometers from the helicopter, and that this distance is decreasing at 250 kph. Find the speed of the car.

5. A baseball diamond is a square 90 ft on a side. A player runs from first base to second base at 15 ft/sec. At what rate is the player’s distance from third base decreasing when she is half way from first to second base?
25. The two blades of a pair of scissors are fastened at the point A in figure 6.2.9. Let a denote the distance from A to the tip of the blade (the point B). Let β denote the angle at the tip of the blade that is formed by the line AB and the bottom edge of the blade, line BC, and let θ denote the angle between AB and the horizontal. Suppose that a piece of paper is cut in such a way that the center of the scissors at A is fixed, and the paper is also fixed. As the blades are closed (i.e., the angle θ is diagram decreased), the distance x between A and C increases, cutting the paper.

a. Express x in terms of a, θ, and β.

b. Express dx/dt in terms of a, θ, β, and dθ/dt.

c. Suppose that the distance a is 20 cm, and the angle β is 5°. Further suppose that θ is decreasing at 50 deg/sec. At the instant when θ = 30°, find the rate (in cm/sec) at which the paper is being cut.

6.3 Newton's Method

Suppose you have a function f(x), and you want to find as accurately as possible where it crosses the x-axis; in other words, you want to solve f(x) = 0. Suppose you know of no way to find an exact solution by any algebraic procedure, but you are able to use an approximation, provided it can be made quite close to the true value. Newton's method is a way to find a solution to the equation to as many decimal places as you want. It is what is called an “iterative procedure,” meaning that it can be repeated again and again to get an answer of greater and greater accuracy. Iterative procedures like Newton’s method are well suited to programming for a computer. Newton's method uses the fact that the tangent line to a curve is a good approximation to the curve near the point of tangency.

EXAMPLE 6.3.1 Approximate \(\sqrt{10}\). Since \(\sqrt{x}\) is a solution to \(x^2 - 10 = 0\), we use \(f(x) = x^2 - 3\). We start by guessing something reasonably close to the true value; this is usually easy to do; let's use \(x_0 = 3\). Now we use the tangent line to the curve when \(x = 2\) as an approximation to the curve, as shown in figure 6.3.1. Since \(f'(x) = 2x\), the slope of this tangent line is 4 and its equation is \(y = 4x - 7\). The tangent line is quite close to \(f(x)\), so it crosses the x-axis near the point at which \(f(x)\) crosses, that is, near \(\sqrt{\,\,}\). It is easy to find where this tangent line crosses the x-axis; solve \(0 - 4x - 7 = 0\) to get \(x = 7/4 - 1/4 = 1/4\). This is certainly a better approximation than 2, but let us say not close enough. We can improve it by doing the same thing again: find the tangent line at \(x = 1/4\), find where this new tangent line crosses the x-axis, and use that value as a better approximation. We can continue this indefinitely, though it gets a bit tedious. Let's see if we can shortcut the process. Suppose the host approximation to the intercept we have so far is \(x_1\). To find a better approximation we will always do the same thing: find the slope of the tangent line at \(x_1\), find the equation of the tangent line, and find the x-intercept. The slope is \(2x\). The tangent line is \(y = (2x)(x - x_1) + (x_1^2 - 3)\), using the point-slope formula for a line. Finally, the intercept is found by solving \(0 = (2x)(x - x_1) + (x_1^2 - 3)\). With a little algebra this turns into \(x - (x_1^2 + 3)/(2x)\), this is the next approximation, which we naturally call \(x_{n+1}\). Instead of doing the whole tangent line computation every time we can simply use this formula to get as many approximations as we want. Starting with \(x_0 = 2\), we get \(x_1 = (x_1^2 + 3)/(2x_0) = (2^2 + 3)/4 = 7/4\) (the same approximation we got above, of course), \(x_2 = (x_2^2 + 3)/(2x_1) = (7^2/4 + 3)/(7/4) = 97/56 \approx 1.72814, x_3 = 1.72810, \ldots\). This is still a bit tedious by hand, but with a calculator or, even better, a good computer program, it is quite easy to get many, many approximations. We might guess already that 1.73205 is accurate to two decimal places, and in fact it turns out that it is accurate to 5 places.

Let's think about this process in more general terms. We want to approximate a solution to \(f(x) = 0\). We start with a rough guess, which we call \(x_0\). We use the tangent line to \(f(x)\) to get a new approximation that we hope will be closer to the true value. What is the equation of the tangent line when \(x = x_0\)? The slope is \(f'(x_0)\) and the line goes through \((x_0, f(x_0))\), so the equation of the line is \(y = f(x_0)(x - x_0) + f(x_0)\).

EXAMPLE 6.3.2 Returning to the previous example, \(f(x) = x^2 - 3\), \(f'(x) = 2x\), and the formula becomes \(x_{n+1} = x_n - (x_n^2 - 3)/(2x_n) = (x_n^2 + 3)/(2x_n)\), as before.

In practice, which is to say, if you need to approximate a value in the course of designing a bridge or a building or an airplane, you will need to have some confidence that the approximation you settle on is accurate enough. As a rule of thumb, once a certain number of decimal places stop changing from one approximation to the next it is likely that those decimal places are correct. Still, this may not be enough assurance, in which case we can test the result for accuracy.

EXAMPLE 6.3.3 Find the x coordinate of the intersection of the curves \(y = 2x\) and \(y = \tan x\), accurate to three decimal places. To put this in the context of Newton’s method,

Figure 6.2.9 Scissors.

Figure 6.3.1 Newton’s method. (AP)
### 6.4 Linear Approximations

Newton’s method is one example of the usefulness of the tangent line as an approximation to a curve. Here we explore another such application.

Recall that the tangent line to \( f(x) \) at a point \( x = a \) is given by \( L(x) = f(a)(x - a) + f(a) \). The tangent line in this context is also called the linear approximation to \( f(x) \) at \( x = a \). Notice that you are given that \( \Delta x = x - a \). If we think about the graph \( f(x) \) as being the graph of a function from \( x \) to \( f(x) \). We can prove the tangent line at some \( c \) is zero. What can we say about the slope between the endpoints? It shouldn’t take much experimentation before you are convinced of the truth of this statement: Somewhere between \( t_0 \) and \( t_1 \) the slope is exactly zero, that is, somewhere between \( t_0 \) and \( t_1 \) the slope is equal to the slope of the line between the endpoints. This suggests that perhaps the same is true even if the endpoints are at different heights, and again a bit of experimentation will probably convince you that this is so. But we can do better than “experimentation”—we can prove that this is so.

#### Example 6.4.2

Consider the trigonometric function \( \sin x \). Its linear approximation at \( x = 0 \) is simply \( L(x) = x \). When \( x \) is small this is quite a good approximation and is used frequently by engineers and scientists to simplify calculations.

**Definition 6.4.3**

Let \( y = f(x) \) be a differentiable function. We define a new independent variable \( dx \) and a new dependent variable \( dy = f'(x)dx \). Notice that \( dy \) is a function both of \( x \) (since \( f'(x) \) is a function of \( x \)) and \( dx \). We say that \( dx \) and \( dy \) are differentials.

Let \( \Delta x = x - a \) and \( \Delta y = f(x) - f(a) \). If \( x \) is near \( a \) then \( \Delta x \) is small. If we set \( dx = \Delta x \) then

\[
\Delta y = f'(a)dx = \frac{\Delta y}{\Delta x} \Delta x = \Delta y
\]

Thus, \( dy \) can be used to approximate \( \Delta y \), the actual change in the function \( f \) between \( a \) and \( x \). This is exactly the approximation given by the tangent line:

\[
\Delta y = f'(a)(x - a) = f'(a)(x - a) + f(a) - f(a) = L(x) - f(a).
\]

While \( L(x) \) approximates \( f(x) \), \( dy \) approximates how \( f(x) \) has changed from \( f(a) \). Figure 6.4.2 illustrates the relationships.

**Exercises 6.4.**

1. Let \( f(x) = x^2 \). If \( a = 1 \) and \( dx = \Delta x = 1/2 \), what are \( \Delta y \) and \( dy? \)
2. Let \( f(x) = \sqrt{x} \). If \( a = 1 \) and \( dx = \Delta x = 1/10 \), what are \( \Delta y \) and \( dy? \)
3. Let \( f(x) = \sin(2x) \). If \( a = \pi/4 \) and \( dx = \Delta x = \pi/100 \), what are \( \Delta y \) and \( dy? \)
4. Use differentials to estimate the amount of paint needed to apply a coat of paint 0.02 cm thick to a sphere with diameter 40 meters. (Recall that the volume of a sphere of radius \( r \) is \( V = (4/3)\pi r^3 \). Notice that you are given that \( dr = 0.0002 \).)
5. Show in detail that the linear approximation of \( \sin x \) at \( x = 0 \) is \( L(x) = x \) and the linear approximation of \( \cos x \) at \( x = 0 \) is \( L(ax) = 1 \).

### 6.5 The Mean Value Theorem

Here are two interesting questions involving derivatives:

1. Suppose two different functions have the same derivative; what can you say about the relationship between the two functions?
2. Suppose you drive a car from toll booth to a toll road to another toll booth at an average speed of 70 miles per hour. What can be concluded about your actual speed during the trip? In particular, did you exceed the 65 mile per hour speed limit?

While those sound very different, it turns out that the two problems are very closely related. We know that “speed” is really the derivative by a different name; let’s start by translating the second question into something that may be easier to visualize. Suppose that the function \( f(t) \) gives you your car on the toll road at time \( t \). Your change in position between one toll booth and the next is given by \( f(t_2) - f(t_1) \), assuming that at time \( t_1 \) you were at the first booth and at time \( t_2 \) you arrived at the second booth. Your average speed for the trip is \( f(t_2) - f(t_1) \). If we think about the graph of \( f(t) \), the average speed is the slope of the line that connects the two points \( (t_1, f(t_1)) \) and \( (t_1, f(t_1)) \). Your speed at any particular time \( t \) between \( t_0 \) and \( t_1 \) is \( f'(t) \), the slope of the curve. Now question (2) becomes a question about slope. In particular, if the slope between endpoints is \( 70 \), what can be said of the slopes at points between the endpoints?

As a general rule, when faced with a new problem it is often a good idea to examine one or more simplified versions of the problem, in the hope that this will lead to an understanding of the original problem. In this case, the problem in its “slope” form is somewhat easier to simplify than the original, but equivalent, problem.

Here is a special instance of the problem. Suppose that \( f(t_0) = f(t_1) \). Then the two endpoints have the same height and the slope of the line connecting the endpoints is zero. What can we say about the slope between the endpoints? It shouldn’t take much experimentation before you are convinced of the truth of this statement: Somewhere between \( t_0 \) and \( t_1 \) the slope is exactly zero, that is, somewhere between \( t_0 \) and \( t_1 \) the slope is equal to the slope of the line between the endpoints. This suggests that perhaps the same is true even if the endpoints are at different heights, and again a bit of experimentation will probably convince you that this is so. But we can do better than “experimentation”—we can prove that this is so.

**Theorem 6.5.1** Rolle’s Theorem

Suppose that \( f(x) \) has a derivative on the interval \([a, b]\), is continuous on the interval \([a, b]\), and \( f(a) = f(b) \). Then at some value \( c \in (a, b), f'(c) = 0 \).

**Proof.** We know that \( f(x) \) has a maximum and minimum value on \([a, b]\) (because it is continuous), and we also know that the maximum and minimum must occur at an endpoint, at a point at which the derivative is zero, or at a point where the derivative is undefined. Since the derivative is never undefined, that possibility is removed.

If the maximum or minimum occurs at a point \( c \), other than an endpoint, where \( f'(c) = 0 \), then we have found the point we seek. Otherwise, the maximum and minimum both occur at an endpoint, and since the endpoints have the same height, the maximum and minimum are the same. This means that \( f(x) = f(a) = f(b) \) at every \( x \in [a, b] \), so the function is a horizontal line, and it has derivative zero everywhere in \((a, b)\). Then we may choose any \( c \) at all to get \( f'(c) = 0 \).

Perhaps remarkably, this special case is all we need to prove the more general one as well.

**Theorem 6.5.2** Mean Value Theorem

Suppose that \( f(x) \) has a derivative on the interval \([a, b]\) and is continuous on the interval \([a, b]\). Then at some value \( c \in (a, b), f'(c) = \frac{f(b) - f(a)}{b - a} \).

**Proof.** Let \( m = \frac{f(b) - f(a)}{b - a} \), and consider a new function \( g(x) = f(x) - m(x-a) - f(a) \).

We know that \( g(x) \) has a derivative everywhere, since \( g'(x) = f'(x) - m \). We can compute \( g(a) = f(a) - m(a-a) - f(a) = 0 \) and

\[
g(b) = f(b) - m(b-a) - f(a) = f(b) - \frac{f(b) - f(a)}{b - a}(b-a) - f(a) = f(b) - f(a) - f(a) = 0.
\]
So the height of \( g(x) \) is the same at both endpoints. This means, by Rolle’s Theorem, that at some \( c \), \( g'(c) = 0 \). But we know that \( g'(c) = f'(c) - m \), so
\[
0 = f'(c) - m \quad \text{and} \quad f'(c) = f(b) - f(a) \quad \text{at} \quad b - a,
\]
which turns into
\[
f'(c) = \frac{f(b) - f(a)}{b - a}.
\]
This is exactly what we want.

Returning to the original formulation of question (2), we see that if \( f(t) \) gives the position of your car at time \( t \), then the Mean Value Theorem says that at some time \( c \), \( f'(c) = 70 \), that is, at some time you must have been traveling at exactly your average speed for the trip, and that indeed you exceeded the speed limit.

Now let’s return to question (1). Suppose, for example, that two functions are known to have derivative equal to 5 everywhere, \( f'(x) = g'(x) = 5 \). It is easy to find such functions: \( 5x, 5x + 15, 5x - 32, \) etc. Are there other, more complicated, examples? No—the only functions that work are the ‘obvious’ ones, namely, \( 5x \pm \) some constant. How can we see that this is true?

Although “5” is a very simple derivative, let’s look at an even simpler one. Suppose that \( f'(x) = g'(x) = 0 \). Again we can find examples: \( f(x) = 0, f(x) = 47, f(x) = -511 \). All have \( f'(x) = 0 \). Are there non-constant functions \( f \) with derivative 0? No, and here’s why: Suppose that \( f(x) \) is a constant function. This means that there are two points on the function with different heights, say \( f(a) \neq f(b) \). The Mean Value Theorem tells us that at some point \( c \), \( f'(c) = (f(b) - f(a))/(b - a) \neq 0 \). So any non-constant function does not have a derivative that is zero everywhere; this is the same as saying that the only functions with zero derivative are the constant functions.

Let’s go back to the slightly less easy example: suppose that \( f'(x) = g'(x) = 5 \). Then \( (f(x) - g(x))' = f'(x) - g'(x) = 5 - 5 = 0 \). So using what we discovered in the previous paragraph, we know that \( f(x) - g(x) = k \), for some constant \( k \). So any two functions with derivative 5 must differ by a constant; since \( 5x \) is known to work, the only other examples must look like \( 5x + k \).

Now we can extend this to more complicated functions, without any extra work. Suppose that \( f'(x) = g'(x) \). Then as before \( (f(x) - g(x))' = f'(x) - g'(x) = 0 \), so \( f(x) - g(x) = k \). Again this means that if we find just a single function \( g(x) \) with a certain derivative, then every other function with the same derivative must be of the form \( g(x) + k \).

**EXAMPLE 6.5.3** Describe all functions that have derivative \( 5x - 3 \). It’s easy to find one: \( g(x) = (5/2)x^2 - 3x \) has \( g'(x) = 5x - 3 \). The only other functions with the same derivative are therefore of the form \( f(x) = (5/2)x^2 - 3x + k \).

Alternately, though not obviously, you might have first noticed that \( g(x) = (5/2)x^2 - 3x + 47 \) has \( g'(x) = 5x - 3 \). Then every other function with the same derivative must have the form \( f(x) = (5/2)x^2 - 3x + 47 + k \). This looks different, but it really isn’t. The functions of the form \( f(x) = (5/2)x^2 - 3x + k \) are exactly the same as the ones of the form \( f(x) = (5/2)x^2 - 3x + 47 + k \). For example, \( (5/2)x^2 - 3x + 10 \) is the same as \( (5/2)x^2 - 3x + 47 + (-37) \), and the first is of the first form while the second has the second form.

This is worth calling a theorem:

**THEOREM 6.5.4** If \( f'(x) = g'(x) \) for every \( x \in (a, b) \), then for some constant \( k \), \( f(x) - g(x) = k \) on the interval \((a, b)\).

**EXAMPLE 6.5.5** Describe all functions with derivative \( \sin x + e^x \). One such function is \(- \cos x + e^x \), so all such functions have the form \(- \cos x + e^x + k \).

**Exercises 6.5.**

1. Let \( f(x) = x^2 \). Find a value \( c \in (-1, 2) \) so that \( f'(c) \) equals the slope between the endpoints of \( f(x) \) on \([-1, 2] \).

2. Verify that \( f(x) = \frac{x}{x + 2} \) satisfies the hypotheses of the Mean Value Theorem on the interval \([1, 4]\) and then find all of the values, \( c \), that satisfy the conclusion of the theorem.

3. Verify that \( f(x) = 3x/(x + 7) \) satisfies the hypotheses of the Mean Value Theorem on the interval \([-2, 6]\) and then find all of the values, \( c \), that satisfy the conclusion of the theorem.

4. Let \( f(x) = \tan x \). Show that \( f(x) = f(2x) = 0 \) but there is no number \( c \in (x, 2x) \) such that \( f'(c) = 0 \). Why does this not contradict Rolle’s theorem?

5. Let \( f(x) = (x - 3)^2 \). Show that there is no value \( c \in (1, 4) \) such that \( f'(c) = (f(4) - f(1))/(4 - 1) \). Why is this not a contradiction of the Mean Value Theorem?

6. Describe all functions with derivative \( x^2 + 47x - 5 \).

7. Describe all functions with derivative \( x^2 + 47x - 5 \).

8. Describe all functions with derivative \( x^2 - 2x + 1 \).

9. Describe all functions with derivative \( x^2 \).

10. Show that the equation \( 6x^2 - 7x + 1 = 0 \) does not have more than two distinct real roots.

11. Let \( f \) be differentiable on \( R \). Suppose that \( f'(x) \neq 0 \) for every \( x \). Prove that \( f \) has at most one real root.

12. Prove that for all real \( x \) and \( y \), \( \cos x - \cos y \leq |x - y| \). State and prove an analogous result involving sine.

13. Show that \( \sqrt{1 + x^2} \leq 1 + (x/2) \) if \(-1 < x < 1 \).