50 Define $f_n(x) = \sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \cdots$ (n terms). Graph f_5 and f_{10} from $-\pi$ to π. Zoom in and describe the *Gibbs phenomenon* at $x = 0$.

On the graphs of 51–56, zoom in to all maxima and minima (3 significant digits). *Estimate inflection points.*

51 $y = 2x^5 - 16x^4 + 5x^3 - 37x^2 + 21x + 683$

52 $y = x^5 - x^4 - \sqrt{3x+1} - 2$

53 $y = x(x-1)(x-2)(x-4)$

54 $y = 7\sin 2x + 5\cos 3x$

55 $y = (x^3 - 2x + 1)/(x^4 - 3x^2 - 15)$, $-3 \leq x \leq 5$

56 $y = x\sin(1/x)$, $0.1 \leq x \leq 1$

57 A 10-digit computer shows $y = 0$ and $dy/dx = .01$ at $x^* = 1$. This root should be correct to about (8 digits) (10 digits) (12 digits). *Hint*: Suppose $y = .01\ (x - 1 + \text{error})$. What errors don't show in 10 digits of y?

58 Which is harder to compute accurately: Maximum point or inflection point? First derivative or second derivative?

3.5 Parabolas, Ellipses, and Hyperbolas

Here is a list of the most important curves in mathematics, so you can tell what is coming. It is not easy to rank the top four:

1. ***straight lines***
2. ***sines and cosines*** (oscillation)
3. ***exponentials*** (growth and decay)
4. ***parabolas, ellipses, and hyperbolas*** (using $1, x, y, x^2, xy, y^2$).

The curves that I wrote last, the Greeks would have written first. It is so natural to go from linear equations to quadratic equations. Straight lines use $1, x, y$. Second degree curves include x^2, xy, y^2. If we go on to x^3 and y^3, the mathematics gets complicated. We now study equations of second degree, and the curves they produce.

It is quite important to see both the ***equations*** and the ***curves***. This section connects two great parts of mathematics—*analysis* of the equation and *geometry* of the curve. Together they produce "***analytic geometry***." You already know about functions and graphs. Even more basic: Numbers correspond to points. We speak about "*the point* (5, 2)." Euclid might not have understood.

Where Euclid drew a 45° line through the origin, Descartes wrote down $y = x$. Analytic geometry has become central to mathematics—we now look at one part of it.

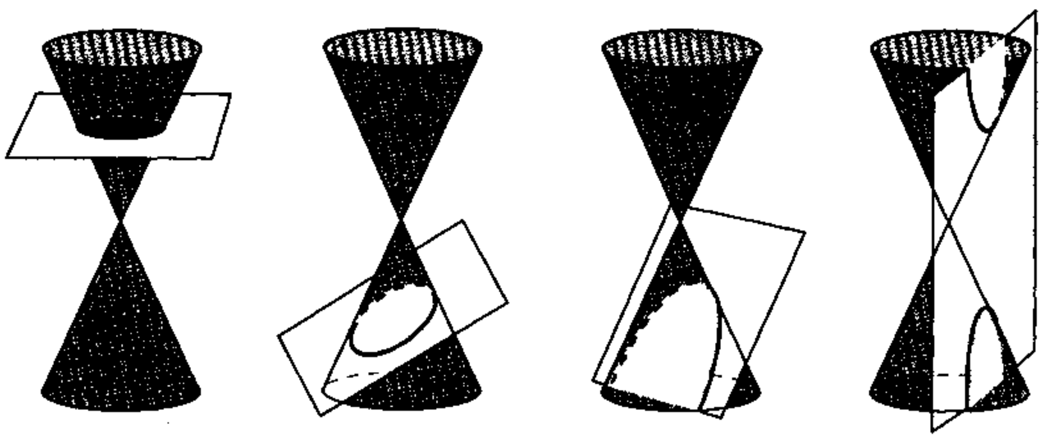


Fig. 3.15 The cutting plane gets steeper: circle to ellipse to parabola to hyperbola.

CONIC SECTIONS

The parabola and ellipse and hyperbola have absolutely remarkable properties. The Greeks discovered that all these curves come from ***slicing a cone by a plane.*** The curves are "conic sections." A level cut gives a ***circle***, and a moderate angle produces an ***ellipse***. A steep cut gives the two pieces of a ***hyperbola*** (Figure 3.15d). At the borderline, when the slicing angle matches the cone angle, the plane carves out a ***parabola***. It has one branch like an ellipse, but it opens to infinity like a hyperbola.

Throughout mathematics, parabolas are on the border between ellipses and hyperbolas.

To repeat: We can slice through cones or we can look for equations. For a cone of light, we see an ellipse on the wall. (The wall cuts into the light cone.) For an equation $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$, we will work to make it simpler. The graph will be centered and rescaled (and rotated if necessary), aiming for an equation like $y = x^2$. Eccentricity and polar coordinates are left for Chapter 9.

THE PARABOLA $y = ax^2 + bx + c$

You knew this function long before calculus. The graph crosses the x axis when $y = 0$. The quadratic formula solves $y = 3x^2 - 4x + 1 = 0$, and so does factoring into $(x-1)(3x-1)$. The crossing points $x = 1$ and $x = \frac{1}{3}$ come from algebra.

The other important point is found by calculus. It is the ***minimum*** point, where $dy/dx = 6x - 4 = 0$. The x coordinate is $\frac{4}{6} = \frac{2}{3}$, halfway between the crossing points. The height is $y_{\min} = -\frac{1}{3}$. This is the ***vertex*** V in Figure 3.16a—at the bottom of the parabola.

A parabola has no asymptotes. The slope $6x - 4$ doesn't approach a constant.

To center the vertex Shift left by $\frac{2}{3}$ and up by $\frac{1}{3}$. So introduce the new variables $X = x - \frac{2}{3}$ and $Y = y + \frac{1}{3}$. Then $x = \frac{2}{3}$ and $y = -\frac{1}{3}$ correspond to $X = Y = 0$—which is the new vertex:

$$y = 3x^2 - 4x + 1 \quad \text{becomes} \quad Y = 3X^2. \tag{1}$$

Check the algebra. $Y = 3X^2$ is the same as $y + \frac{1}{3} = 3(x - \frac{2}{3})^2$. That simplifies to the original equation $y = 3x^2 - 4x + 1$. The second graph shows the centered parabola $Y = 3X^2$, with the vertex moved to the origin.

To zoom in on the vertex Rescale X and Y by the zoom factor a:

$$Y = 3X^2 \quad \text{becomes} \quad \mathbf{y}/a = 3(\mathbf{x}/a)^2.$$

The final equation has $\mathbf{x}$ and $\mathbf{y}$ in boldface. With $a = 3$ we find $\mathbf{y} = \mathbf{x}^2$—the graph is magnified by 3. In two steps we have reached the model parabola opening upward.

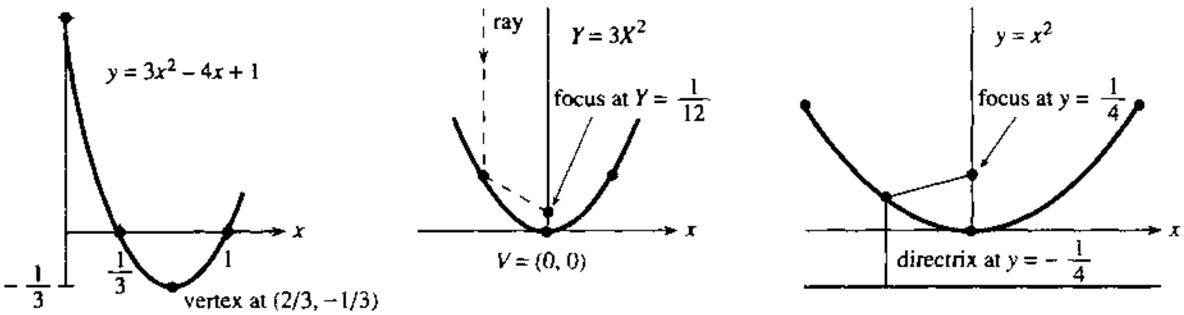


Fig. 3.16 Parabola with minimum at V. Rays reflect to focus. Centered in (b), rescaled in (c).

A parabola has another important point—the *focus*. Its distance from the vertex is called p. The special parabola $\mathbf{y} = \mathbf{x}^2$ has $p = 1/4$, and other parabolas $Y = aX^2$ have $p = 1/4a$. You magnify by a factor a to get $\mathbf{y} = \mathbf{x}^2$. The beautiful property of a parabola is that ***every ray coming straight down is reflected to the focus.***

Problem 2.3.25 located the focus F—here we mention two applications. A solar collector and a TV dish are parabolic. They concentrate sun rays and TV signals onto a point—a heat cell or a receiver collects them at the focus. The 1982 *UMAP Journal* explains how radar and sonar use the same idea. Car headlights turn the idea around, and send the light outward.

Here is a classical fact about parabolas. ***From each point on the curve, the distance to the focus equals the distance to the "directrix."*** The directrix is the line $y = -p$ below the vertex (so the vertex is halfway between focus and directrix). With $p = \frac{1}{4}$, the distance down from any (x, y) is $y + \frac{1}{4}$. Match tbat with tbe distance to the focus at $(0, \frac{1}{4})$— this is the square root below. Out comes the special parabola $y = x^2$:

$$y + \tfrac{1}{4} = \sqrt{x^2 + (y - \tfrac{1}{4})^2} \longrightarrow \text{(square both sides)} \longrightarrow y = x^2. \tag{2}$$

The exercises give practice with all the steps we have taken—center the parabola to $Y = aX^2$, rescale it to $\mathbf{y} = \mathbf{x}^2$, locate the vertex and focus and directrix.

Summary for other parabolas $y = ax^2 + bx + c$ has its vertex where dy/dx is zero. Thus $2ax + b = 0$ and $x = -b/2a$. Shifting across to that point is "completing the square":

$$ax^2 + bx + c \quad \text{equals} \quad a\left(x + \frac{b}{2a}\right)^2 + C. \tag{3}$$

Here $C = c - (b^2/4a)$ is the height of the vertex. The centering transform $X = x + (b/2a)$, $Y = y - C$ produces $Y = aX^2$. It moves the vertex to $(0, 0)$, where it belongs.

For the ellipse and hyperbola, our plan of attack is the same:

1. Center the curve to remove any linear terms Dx and Ey.
2. Locate each focus and discover the reflection property.
3. Rotate to remove Bxy if the equation contains it.

ELLIPSES $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ (CIRCLES HAVE $a = b$)

This equation makes the ellipse symmetric about $(0, 0)$—the center. Changing x to $-x$ or y to $-y$ leaves the same equation. No extra centering or rotation is needed.

The equation also shows that x^2/a^2 and y^2/b^2 cannot exceed one. (They add to one and can't be negative.) Therefore $x^2 \leqslant a^2$, and x stays between $-a$ and a. Similarly y stays between b and $-b$. The ellipse is inside a rectangle.

By solving for y we get a function (or two functions!) of x:

$$\frac{y^2}{b^2} = 1 - \frac{x^2}{a^2} \quad \text{gives} \quad \frac{y}{b} = \pm\sqrt{1 - \frac{x^2}{a^2}} \quad \text{or} \quad y = \pm\frac{b}{a}\sqrt{a^2 - x^2}.$$

The graphs are the top half (+) and bottom half (−) of the ellipse. To draw the ellipse, plot them together. They meet when $y = 0$, at $x = a$ on the far right of Figure 3.17 and at $x = -a$ on the far left. The maximum $y = b$ and minimum $y = -b$ are at the top and bottom of the ellipse, where we bump into the enclosing rectangle.

A circle is a special case of an ellipse, when $a = b$. The circle equation $x^2 + y^2 = r^2$ is the ellipse equation with $a = b = r$. This circle is centered at $(0, 0)$; other circles are centered at $x = h$, $y = k$. The circle is determined by its *radius* r and its *center* (h, k):

$$\textbf{\textit{Equation of circle}}: \ (x-h)^2 + (y-k)^2 = r^2. \tag{4}$$

In words, the distance from (x, y) on the circle to (h, k) at the center is r. The equation has linear terms $-2hx$ and $-2ky$—they disappear when the center is (0, 0).

EXAMPLE 1 Find the circle that has a diameter from (1, 7) to (5, 7).

Solution The center is halfway at (3, 7). So $r = 2$ and $(x-3)^2 + (y-7)^2 = 2^2$.

EXAMPLE 2 Find the center and radius of the circle $x^2 - 6x + y^2 - 14y = -54$.

Solution Complete $x^2 - 6x$ to the square $(x-3)^2$ by adding 9. Complete $y^2 - 14y$ to $(y-7)^2$ by adding 49. Adding 9 and 49 to both sides of the equation leaves $(x-3)^2 + (y-7)^2 = 4$—the same circle as in Example 1.

Quicker Solution Match the given equation with (4). Then $h = 3$, $k = 7$, and $r = 2$:

$$x^2 - 6x + y^2 - 14y = -54 \quad \text{must agree with} \quad x^2 - 2hx + h^2 + y^2 - 2ky + k^2 = r^2.$$

The change to $X = x - h$ and $Y = y - k$ moves the center of the circle from (h, k) to (0, 0). This is equally true for an ellipse:

$$\textit{The ellipse } \frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1 \quad \textit{becomes} \quad \frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1.$$

When we rescale by $x = X/a$ and $y = Y/b$, we get the unit circle $x^2 + y^2 = 1$.

The unit circle has area π. ***The ellipse has area*** πab (proved later in the book). The distance around the circle is 2π. The distance around an ellipse does not rescale—it has no simple formula.

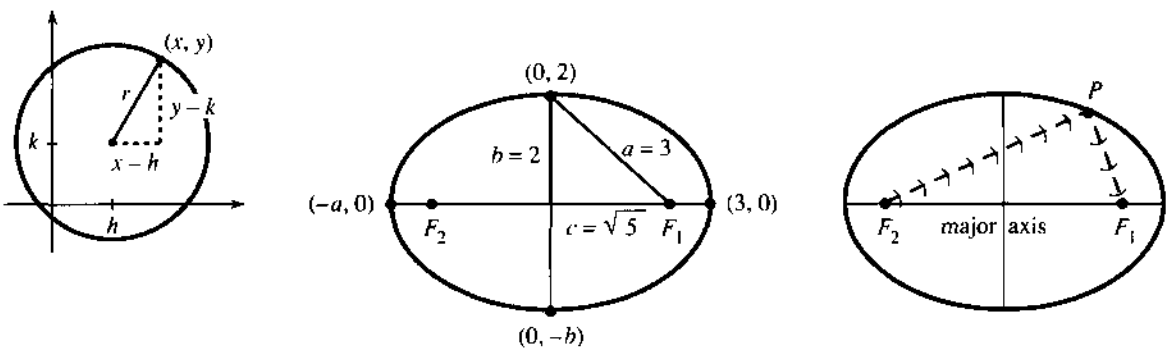


Fig. 3.17 Uncentered circle. Centered ellipse $x^2/3^2 + y^2/2^2 = 1$. *The distance from center to far right is also* $a = 3$. All rays from F_2 reflect to F_1.

Now we leave circles and concentrate on ellipses. They have ***two foci*** (pronounced *fo-sigh*). For a parabola, the second focus is at infinity. For a circle, both foci are at the center. The foci of an ellipse are on its longer axis (its *major* axis), one focus on each side of the center:

$$F_1 \text{ is at } x = c = \sqrt{a^2 - b^2} \qquad \text{and} \qquad F_2 \text{ is at } x = -c.$$

The right triangle in Figure 3.17 *has sides* a, b, c. From the top of the ellipse, the distance to each focus is a. From the endpoint at $x = a$, the distances to the foci are $a + c$ and $a - c$. Adding $(a + c) + (a - c)$ gives $2a$. *As you go around the ellipse, the distance to* F_1 *plus the distance to* F_2 *is constant* (always $2a$).

> **3H** At all points on the ellipse, the sum of distances from the foci is $2a$. This is another equation for the ellipse:
>
> $$\text{from } F_1 \text{ and } F_2 \text{ to } (x, y): \quad \sqrt{(x-c)^2+y^2} + \sqrt{(x+c)^2+y^2} = 2a. \tag{5}$$

To draw an ellipse, tie a string of length $2a$ to the foci. Keep the string taut and your moving pencil will create the ellipse. This description uses a and c—the other form uses a and b (remember $b^2+c^2=a^2$). Problem 24 asks you to simplify equation (5) until you reach $x^2/a^2+y^2/b^2=1$.

The "whispering gallery" of the United States Senate is an ellipse. If you stand at one focus and speak quietly, you can be heard at the other focus (and nowhere else). Your voice is reflected off the walls to the other focus—following the path of the string. For a parabola the rays come in to the focus from infinity—where the second focus is.

A hospital uses this reflection property to split up kidney stones. The patient sits inside an ellipse with the kidney stone at one focus. At the other focus a *lithotripter* sends out hundreds of small shocks. You get a spinal anesthetic (I mean the patient) and the stones break into tiny pieces.

The most important focus is the Sun. The ellipse is the orbit of the Earth. See Section 12.4 for a terrible printing mistake by the Royal Mint, on England's last pound note. They put the Sun at the center.

Question 1 Why do the whispers (and shock waves) arrive together at the second focus?
Answer Whichever way they go, the distance is $2a$. Exception: straight path is $2c$.

Question 2 Locate the ellipse with equation $4x^2+9y^2=36$.
Answer Divide by 36 to change the constant to 1. Now identify a and b:

$$\frac{x^2}{9}+\frac{y^2}{4}=1 \text{ so } a=\sqrt{9} \text{ and } b=\sqrt{4}. \text{ Foci at } \pm\sqrt{9-4}=\pm\sqrt{5}.$$

Question 3 Shift the center of that ellipse across and down to $x=1$, $y=-5$.
Answer Change x to $x-1$. Change y to $y+5$. The equation becomes $(x-1)^2/9+(y+5)^2/4=1$. In practice we start with this uncentered ellipse and go the other way to center it.

HYPERBOLAS $\frac{y^2}{a^2}-\frac{x^2}{b^2}=1$

Notice the minus sign for a hyperbola. That makes all the difference. Unlike an ellipse, x and y can both be large. The curve goes out to infinity. It is still symmetric, since x can change to $-x$ and y to $-y$.

The center is at (0, 0). Solving for y again yields two functions (+ and −):

$$\frac{y^2}{a^2}-\frac{x^2}{b^2}=1 \quad \text{gives} \quad \frac{y}{a}=\pm\sqrt{1+\frac{x^2}{b^2}} \quad \text{or} \quad y=\pm\frac{a}{b}\sqrt{b^2+x^2}. \tag{6}$$

The hyperbola has two branches that never meet. The upper branch, with a plus sign, has $y \geqslant a$. The ***vertex*** V_1 is at $x=0$, $y=a$—the lowest point on the branch. Much further out, when x is large, the hyperbola climbs up beside its ***sloping asymptotes***:

$$\text{if } \frac{x^2}{b^2}=1000 \quad \text{then} \quad \frac{y^2}{a^2}=1001. \quad \text{So } \frac{y}{a} \text{ is close to } \frac{x}{b} \text{ or } -\frac{x}{b}.$$

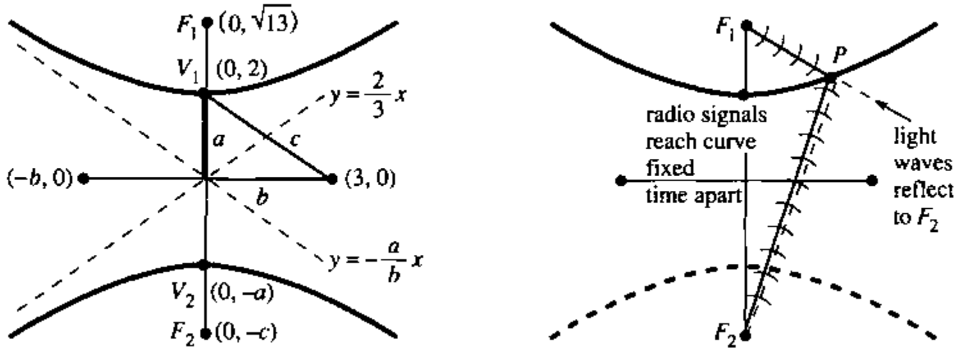


Fig. 3.18 The hyperbola $\frac{1}{4}y^2 - \frac{1}{9}x^2 = 1$ has $a = 2$, $b = 3$, $c = \sqrt{4+9}$. The distances to F_1 and F_2 differ by $2a = 4$.

The asymptotes are the lines $y/a = x/b$ and $y/a = -x/b$. Their slopes are a/b and $-a/b$. You can't miss them in Figure 3.18.

For a hyperbola, the foci are inside the two branches. Their distance from the center is still called c. But now $c = \sqrt{a^2 + b^2}$, which is larger than a and b. The vertex is a distance $c - a$ from one focus and $c + a$ from the other. The *difference* (not the sum) is $(c + a) - (c - a) = 2a$.

All points on the hyperbola have this property: ***The difference between distances to the foci is constantly*** $2a$. A ray coming in to one focus is reflected toward the other. The reflection is on the *outside* of the hyperbola, and the *inside* of the ellipse.

Here is an application to navigation. Radio signals leave two fixed transmitters at the same time. A ship receives the signals a millisecond apart. Where is the ship? *Answer*: It is on a hyperbola with foci at the transmitters. Radio signals travel 186 miles in a millisecond, so $186 = 2a$. This determines the curve. In Long Range Navigation (LORAN) a third transmitter gives another hyperbola. Then the ship is located exactly.

Question 4 How do hyperbolas differ from parabolas, far from the center?
Answer Hyperbolas have asymptotes. Parabolas don't.

The hyperbola has a natural rescaling. The appearance of x/b is a signal to change to X. Similarly y/a becomes Y. Then $Y = 1$ at the vertex, and we have a standard hyperbola:

$$y^2/a^2 - x^2/b^2 = 1 \quad \text{becomes} \quad Y^2 - X^2 = 1.$$

A 90° turn gives $X^2 - Y^2 = 1$—the hyperbola opens to the sides. A 45° turn produces $2XY = 1$. We show below how to recognize $x^2 + xy + y^2 = 1$ as an ellipse and $x^2 + 3xy + y^2 = 1$ as a hyperbola. (They are not circles because of the xy term.) When the xy coefficient increases past 2, $x^2 + y^2$ no longer indicates an ellipse.

Question 5 Locate the hyperbola with equation $9y^2 - 4x^2 = 36$.
Answer Divide by 36. Then $y^2/4 - x^2/9 = 1$. Recognize $a = \sqrt{4}$ and $b = \sqrt{9}$.

Question 6 Locate the uncentered hyperbola $9y^2 - 18y - 4x^2 - 4x = 28$.
Answer Complete $9y^2 - 18y$ to $9(y-1)^2$ by adding 9. Complete $4x^2 + 4x$ to $4(x + \frac{1}{2})^2$ by adding $4(\frac{1}{2})^2 = 1$. The equation is rewritten as $9(y-1)^2 - 4(x + \frac{1}{2})^2 = 28 + 9 - 1$. This is the hyperbola in Question 5 — except its center is $(-\frac{1}{2}, 1)$.

To summarize: Find the center by completing squares. Then read off a and b.

THE GENERAL EQUATION $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$

This equation is of second degree, containing any and all of $1, x, y, x^2, xy, y^2$. A plane is cutting through a cone. ***Is the curve a parabola or ellipse or hyperbola***? Start with the most important case $Ax^2 + Bxy + Cy^2 = 1$.

> **3I** The equation $Ax^2 + Bxy + Cy^2 = 1$ produces a hyperbola if $B^2 > 4AC$ and an ellipse if $B^2 < 4AC$. A parabola has $B^2 = 4AC$.

To recognize the curve, we remove Bxy by *rotating the plane*. This also changes A and C—but the combination $B^2 - 4AC$ is not changed (proof omitted). An example is $2xy = 1$, with $B^2 = 4$. It rotates to $y^2 - x^2 = 1$, with $-4AC = 4$. That positive number 4 signals a hyperbola—since $A = -1$ and $C = 1$ have opposite signs.

Another example is $x^2 + y^2 = 1$. It is a circle (a special ellipse). However we rotate, the equation stays the same. The combination $B^2 - 4AC = 0 - 4 \cdot 1 \cdot 1$ is negative, as predicted for ellipses.

To rotate by an angle α, change x and y to new variables x' and y':

$$\begin{aligned} x &= x' \cos\alpha - y' \sin\alpha \\ y &= x' \sin\alpha + y' \cos\alpha \end{aligned} \quad \text{and} \quad \begin{aligned} x' &= x \cos\alpha + y \sin\alpha \\ y' &= -y \sin\alpha + x \cos\alpha. \end{aligned} \tag{7}$$

Substituting for x and y changes $Ax^2 + Bxy + Cy^2 = 1$ to $A'x'^2 + B'x'y' + C'y'^2 = 1$. The formulas for A', B', C' are painful so I go to the key point:

$$B' \textit{ is zero if the rotation angle } \alpha \textit{ has } \tan 2\alpha = B/(A - C).$$

With $B' = 0$, the curve is easily recognized from $A'x'^2 + C'y'^2 = 1$. It is a hyperbola if A' and C' have opposite signs. Then $B'^2 - 4A'C'$ is positive. The original $B^2 - 4AC$ was also positive, because this special combination stays constant during rotation.

After the xy term is gone, we deal with x and y—by ***centering***. To find the center, complete squares as in Questions 3 and 6. For total perfection, rescale to one of the model equations $\mathbf{y} = \mathbf{x}^2$ or $\mathbf{x}^2 + \mathbf{y}^2 = 1$ or $\mathbf{y}^2 - \mathbf{x}^2 = 1$.

The remaining question is about $F = 0$. What is the graph of $Ax^2 + Bxy + Cy^2 = 0$? The ellipse-hyperbola-parabola have disappeared. But if the Greeks were right, the cone is still cut by a plane. The degenerate case $F = 0$ occurs when the plane cuts ***right through the sharp point of the cone.***

A level cut hits only that one point $(0, 0)$. The equation shrinks to $x^2 + y^2 = 0$, a circle with radius zero. A steep cut gives two lines. The hyperbola becomes $y^2 - x^2 = 0$, leaving only its asymptotes $y = \pm x$. A cut at the exact angle of the cone gives only one line, as in $x^2 = 0$. A ***single point, two lines***, and ***one line*** are very extreme cases of an ellipse, hyperbola, and parabola.

All these "conic sections" come from planes and cones. The beauty of the geometry, which Archimedes saw, is matched by the importance of the equations. Galileo discovered that projectiles go along parabolas (Chapter 12). Kepler discovered that the Earth travels on an ellipse (also Chapter 12). Finally Einstein discovered that light travels on hyperbolas. That is in four dimensions, and not in Chapter 12.

	equation	vertices	foci
P	$y = ax^2 + bx + c$	$\left(-\dfrac{b}{2a}, c - \dfrac{b^2}{4a}\right)$	$\dfrac{1}{4a}$ above vertex, also infinity
E	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1, a > b$	$(a, 0)$ and $(-a, 0)$	$(c, 0)$ and $(-c, 0)$: $c = \sqrt{a^2 - b^2}$
H	$\dfrac{y^2}{a^2} - \dfrac{x^2}{b^2} = 1$	$(0, a)$ and $(0, -a)$	$(0, c)$ and $(0, -c)$: $c = \sqrt{a^2 + b^2}$

3.5 EXERCISES

Read-through questions

The graph of $y = x^2 + 2x + 5$ is a ___a___. Its lowest point (the vertex) is $(x, y) =$ (___b___). Centering by $X = x + 1$ and $Y =$ ___c___ moves the vertex to (0, 0). The equation becomes $Y =$ ___d___. The focus of this centered parabola is ___e___. All rays coming straight down are ___f___ to the focus.

The graph of $x^2 + 4y^2 = 16$ is an ___g___. Dividing by ___h___ leaves $x^2/a^2 + y^2/b^2 = 1$ with $a =$ ___i___ and $b =$ ___j___. The graph lies in the rectangle whose sides are ___k___. The area is $\pi ab =$ ___l___. The foci are at $x = \pm c =$ ___m___. The sum of distances from the foci to a point on this ellipse is always ___n___. If we rescale to $X = x/4$ and $Y = y/2$ the equation becomes ___o___ and the graph becomes a ___p___.

The graph of $y^2 - x^2 = 9$ is a ___q___. Dividing by 9 leaves $y^2/a^2 - x^2/b^2 = 1$ with $a =$ ___r___ and $b =$ ___s___. On the upper branch $y \geqslant$ ___t___. The asymptotes are the lines ___u___. The foci are at $y = \pm c =$ ___v___. The ___w___ of distances from the foci to a point on this hyperbola is ___x___.

All these curves are conic sections—the intersection of a ___y___ and a ___z___. A steep cutting angle yields a ___A___. At the borderline angle we get a ___B___. The general equation is $Ax^2 +$ ___C___ $+ F = 0$. If $D = E = 0$ the center of the graph is at ___D___. The equation $Ax^2 + Bxy + Cy^2 = 1$ gives an ellipse when ___E___. The graph of $4x^2 + 5xy + 6y^2 = 1$ is a ___F___.

1 The vertex of $y = ax^2 + bx + c$ is at $x = -b/2a$. What is special about this x? Show that it gives $y = c - (b^2/4a)$.

2 The parabola $y = 3x^2 - 12x$ has $x_{\min} =$ ________. At this minimum, $3x^2$ is ________ as large as $12x$. Introducing $X = x - 2$ and $Y = y + 12$ centers the equation to ________.

Draw the curves 3–14 by hand or calculator or computer. Locate the vertices and foci.

3 $y = x^2 - 2x - 3$	**4** $y = (x-1)^2$
5 $4y = -x^2$	**6** $4x = y^2$
7 $(x-1)^2 + (y-1)^2 = 1$	**8** $x^2 + 9y^2 = 9$
9 $9x^2 + y^2 = 9$	**10** $x^2/4 - (y-1)^2 = 1$
11 $y^2 - 4x^2 = 1$	**12** $(y-1)^2 - 4x^2 = 1$
13 $y^2 - x^2 = 0$	**14** $xy = 0$

Problems 15–20 are about parabolas, 21–34 are about ellipses, 35–41 are about hyperbolas.

15 Find the parabola $y = ax^2 + bx + c$ that goes through (0, 0) and (1, 1) and (2, 12).

16 $y = x^2 - x$ has vertex at ________. To move the vertex to (0, 0) set $X =$ ________ and $Y =$ ________. Then $Y = X^2$.

17 (a) In equation (2) change $\frac{1}{4}$ to p. Square and simplify.
(b) Locate the focus and directrix of $Y = 3X^2$. Which points are a distance 1 from the directrix and focus?

18 The parabola $y = 9 - x^2$ opens ________ with vertex at ________. Centering by $Y = y - 9$ yields $Y = -x^2$.

19 Find equations for all parabolas which
(a) open to the right with vertex at (0, 0)
(b) open upwards with focus at (0, 0)
(c) open downwards and go through (0, 0) and (1, 0).

20 A projectile is at $x = t$, $y = t - t^2$ at time t. Find dx/dt and dy/dt at the start, the maximum height, and an xy equation for the path.

21 Find the equation of the ellipse with extreme points at $(\pm 2, 0)$ and $(0, \pm 1)$. Then shift the center to (1, 1) and find the new equation.

22 On the ellipse $x^2/a^2 + y^2/b^2 = 1$, solve for y when $x = c = \sqrt{a^2 - b^2}$. This height above the focus will be valuable in proving Kepler's third law.

23 Find equations for the ellipses with these properties:
(a) through (5, 0) with foci at $(\pm 4, 0)$
(b) with sum of distances to (1, 1) and (5, 1) equal to 12
(c) with both foci at (0, 0) and sum of distances $= 2a = 10$.

24 Move a square root to the right side of equation (5) and square both sides. Then isolate the remaining square root and square again. Simplify to reach the equation of an ellipse.

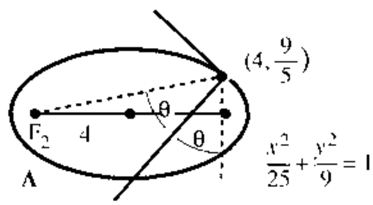
25 Decide between circle-ellipse-parabola-hyperbola, based on the XY equation with $X = x - 1$ and $Y = y + 3$.

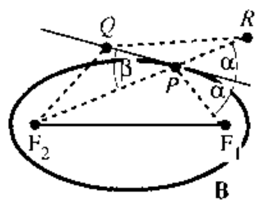
(a) $x^2 - 2x + y^2 + 6y = 6$
(b) $x^2 - 2x - y^2 - 6y = 6$
(c) $x^2 - 2x + 2y^2 + 12y = 6$
(d) $x^2 - 2x - y = 6$.

26 A tilted cylinder has equation $(x - 2y - 2z)^2 + (y - 2x - 2z)^2 = 1$. Show that the water surface at $z = 0$ is an ellipse. What is its equation and what is $B^2 - 4AC$?

27 $(4, 9/5)$ is above the focus on the ellipse $x^2/25 + y^2/9 = 1$. Find dy/dx at that point and the equation of the tangent line.

28 (a) Check that the line $xx_0 + yy_0 = r^2$ is tangent to the circle $x^2 + y^2 = r^2$ at (x_0, y_0).
(b) For the ellipse $x^2/a^2 + y^2/b^2 = 1$ show that the tangent equation is $xx_0/a^2 + yy_0/b^2 = 1$. (Check the slope.)





29 The slope of the normal line in Figure **A** is $s = -1/$(slope of tangent) = ________. The slope of the line from F_2 is $S =$ ________. By the reflection property,

$$S = \cot 2\theta = \frac{1}{2}(\cot\theta - \tan\theta) = \frac{1}{2}\left(s - \frac{1}{s}\right).$$

Test your numbers s and S against this equation.

30 Figure **B** proves the reflecting property of an ellipse. R is the mirror image of F_1 in the tangent line; Q is any other point on the line. Deduce steps 2, 3, 4 from 1, 2, 3:

1. $PF_1 + PF_2 < QF_1 + QF_2$ (*left side* $= 2a$, Q *is outside*)
2. $PR + PF_2 < QR + QF_2$
3. P is on the straight line from F_2 to R
4. $\alpha = \beta$: the reflecting property is proved.

31 The ellipse $(x - 3)^2/4 + (y - 1)^2/4 = 1$ is really a ________ with center at ________ and radius ________. Choose X and Y to produce $X^2 + Y^2 = 1$.

32 Compute the area of a square that just fits inside the ellipse $x^2/a^2 + y^2/b^2 = 1$.

33 Rotate the axes of $x^2 + xy + y^2 = 1$ by using equation (7) with $\sin\alpha = \cos\alpha = 1/\sqrt{2}$. The $x'y'$ equation should show an ellipse.

34 What are a, b, c for the Earth's orbit around the sun?

35 Find an equation for the hyperbola with

(a) vertices $(0, \pm 1)$, foci $(0, \pm 2)$
(b) vertices $(0, \pm 3)$, asymptotes $y = \pm 2x$
(c) $(2, 3)$ on the curve, asymptotes $y = \pm x$

36 Find the slope of $y^2 - x^2 = 1$ at (x_0, y_0). Show that $yy_0 - xx_0 = 1$ goes through this point with the right slope (it has to be the tangent line).

37 If the distances from (x, y) to $(8, 0)$ and $(-8, 0)$ differ by 10, what hyperbola contains (x, y)?

38 If a cannon was heard by Napoleon and one second later by the Duke of Wellington, the cannon was somewhere on a ________ with foci at ________.

39 $y^2 - 4y$ is part of $(y - 2)^2 =$ ________ and $2x^2 + 12x$ is part of $2(x + 3)^2 =$ ________. Therefore $y^2 - 4y - 2x^2 - 12x = 0$ gives the hyperbola $(y - 2)^2 - 2(x + 3)^2 =$ ________. Its center is ________ and it opens to the ________.

40 Following Problem 39 turn $y^2 + 2y = x^2 + 10x$ into $Y^2 = X^2 + C$ with X, Y, and C equal to ________.

41 Draw the hyperbola $x^2 - 4y^2 = 1$ and find its foci and asymptotes.

Problems 42–46 are about second-degree curves (conics).

42 For which A, C, F does $Ax^2 + Cy^2 + F = 0$ have no solution (empty graph)?

43 Show that $x^2 + 2xy + y^2 + 2x + 2y + 1 = 0$ is the equation (squared) of a single line.

44 Given any ________ points in the plane, a second-degree curve $Ax^2 + \cdots + F = 0$ goes through those points.

45 (a) When the plane $z = ax + by + c$ meets the cone $z^2 = x^2 + y^2$, eliminate z by squaring the plane equation. Rewrite in the form $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$.
(b) Compute $B^2 - 4AC$ in terms of a and b.
(c) Show that the plane meets the cone in an ellipse if $a^2 + b^2 < 1$ and a hyperbola if $a^2 + b^2 > 1$ (*steeper*).

46 The roots of $ax^2 + bx + c = 0$ also involve the special combination $b^2 - 4ac$. This quadratic equation has two real roots if ________ and no real roots if ________. The roots come together when $b^2 = 4ac$, which is the borderline case like a parabola.

3.6 Iterations $x_{n+1} = F(x_n)$

Iteration means repeating the same function. Suppose the function is $F(x) = \cos x$. Choose any starting value, say $x_0 = 1$. Take its cosine: $x_1 = \cos x_0 = .54$. ***Then take the cosine of*** x_1. That produces $x_2 = \cos .54 = .86$. *The iteration is* $x_{n+1} = \cos x_n$. I am in radian mode on a calculator, pressing "cos" each time. The early numbers are not important, what is important is the output after 12 or 30 or 100 steps:

EXAMPLE 1 $x_{12} = .75, \quad x_{13} = .73, \quad x_{14} = .74, \quad \ldots, \quad x_{29} = .7391, \quad x_{30} = .7391.$

The goal is to explain why the x's approach $x^* = .739085\ldots$.. Every starting value x_0 leads to this same number x^*. ***What is special about*** .7391?

Note on iterations Do $x_1 = \cos x_0$ and $x_2 = \cos x_1$ mean that $x_2 = \cos^2 x_0$? Absolutely not! Iteration creates a new and different function $\cos(\cos x)$. It uses the cos button, not the squaring button. The third step creates $F(F(F(x)))$. As soon as you can, iterate with $x_{n+1} = \frac{1}{2}\cos x_n$. What limit do the x's approach? Is it $\frac{1}{2}(.7931)$?

Let me slow down to understand these questions. ***The central idea is expressed by the equation*** $x_{n+1} = F(x_n)$. Substituting x_0 into F gives x_1. This output x_1 is the input that leads to x_2. In its turn, x_2 is the input and out comes $x_3 = F(x_2)$. This is ***iteration***, and it produces the sequence $x_0, x_1, x_2, \ldots$.

The x's may approach a limit x^*, depending on the function F. Sometimes x^* also depends on the starting value x_0. Sometimes there is *no* limit. Look at a second example, which does not need a calculator.

EXAMPLE 2 $x_{n+1} = F(x_n) = \frac{1}{2}x_n + 4$. Starting from $x_0 = 0$ the sequence is

$$x_1 = \tfrac{1}{2}\cdot 0 + 4 = 4, \quad x_2 = \tfrac{1}{2}\cdot 4 + 4 = 6, \quad x_3 = \tfrac{1}{2}\cdot 6 + 4 = 7, \quad x_4 = \tfrac{1}{2}\cdot 7 + 4 = 7\tfrac{1}{2}, \quad \ldots$$

Those numbers $0, 4, 6, 7, 7\frac{1}{2}, \ldots$ seem to be approaching $x^* = 8$. A computer would convince us. So will mathematics, when we see what is special about 8:

When the x's approach x^*, the limit of $x_{n+1} = \frac{1}{2}x_n + 4$
is $x^* = \frac{1}{2}x^* + 4$. This limiting equation yields $x^* = 8$.

8 is the "steady state" where *input equals output*: $8 = F(8)$. It is the ***fixed point***.

If we start at $x_0 = 8$, the sequence is $8, 8, 8, \ldots$. When we start at $x_0 = 12$, the sequence goes back toward 8:

$$x_1 = \tfrac{1}{2}\cdot 12 + 4 = 10, \qquad x_2 = \tfrac{1}{2}\cdot 10 + 4 = 9, \qquad x_3 = \tfrac{1}{2}\cdot 9 + 4 = 8.5, \qquad \ldots$$

Equation for limit**: **If the iterations $x_{n+1} = F(x_n)$ ***converge to*** x^*, ***then*** $x^* = F(x^*)$.

To repeat: 8 is special because it equals $\frac{1}{2}\cdot 8 + 4$. The number $.7391\ldots$ is special because it equals $\cos .7391\ldots$. ***The graphs of*** $y = x$ ***and*** $y = F(x)$ ***intersect at*** x^*. To explain *why* the x's converge (or why they don't) is the job of calculus.

EXAMPLE 3 $x_{n+1} = x_n^2$ has two fixed points: $0 = 0^2$ and $1 = 1^2$. Here $F(x) = x^2$.

Starting from $x_0 = \frac{1}{2}$ the sequence $\frac{1}{4}, \frac{1}{16}, \frac{1}{256}, \ldots$ goes quickly to $x^* = 0$. The only approaches to $x^* = 1$ are from $x_0 = 1$ (of course) and from $x_0 = -1$. Starting from $x_0 = 2$ we get $4, 16, 256, \ldots$ and *the sequence diverges to* $+\infty$.

Each limit x^* has a "***basin of attraction***." The basin contains all starting points x_0 that lead to x^*. For Examples 1 and 2, every x_0 led to .7391 and 8. The basins were the whole line (that is still to be proved). Example 3 had three basins—the interval $-1 < x_0 < 1$, the two points $x_0 = \pm 1$, and all the rest. The outer basin $|x_0| > 1$ led to $\pm\infty$. I challenge you to find the limits and the basins of attraction (by calculator) for $F(x) = x - \tan x$.

In Example 3, $x^* = 0$ is ***attracting***. Points near x^* move toward x^*. The fixed point $x^* = 1$ is ***repelling***. Points near 1 move away. We now find the rule that decides whether x^* is attracting or repelling. ***The key is the slope dF/dx at x^*.***

> **3J** Start from any x_0 near a fixed point $x^* = F(x^*)$:
>
> x^* is ***attracting*** if $|dF/dx|$ is below 1 at x^*
>
> x^* is ***repelling*** if $|dF/dx|$ is above 1 at x^*.

First I will give a calculus proof. Then comes a picture of convergence, by "*cobwebs*." Both methods throw light on this crucial test for attraction: $|dF/dx| < 1$.

First proof: Subtract $x^* = F(x^*)$ from $x_{n+1} = F(x_n)$. The difference $x_{n+1} - x^*$ is the same as $F(x_n) - F(x^*)$. This is ΔF. ***The basic idea of calculus is that ΔF is close to $F'\Delta x$:***

$$x_{n+1} - x^* = F(x_n) - F(x^*) \approx F'(x^*)(x_n - x^*). \tag{1}$$

The "error" $x_n - x^*$ is multiplied by the slope dF/dx. The next error $x_{n+1} - x^*$ is smaller or larger, based on $|F'| < 1$ or $|F'| > 1$ at x^*. Every step multiplies approximately by $F'(x^*)$. ***Its size controls the speed of convergence.***

In Example 1, $F(x)$ is $\cos x$ and $F'(x)$ is $-\sin x$. There is attraction to .7391 because $|\sin x^*| < 1$. In Example 2, F is $\frac{1}{2}x + 4$ and F' is $\frac{1}{2}$. There is attraction to 8. In Example 3, F is x^2 and F' is $2x$. There is superattraction to $x^* = 0$ (where $F' = 0$). There is repulsion from $x^* = 1$ (where $F' = 2$).

I admit one major difficulty. The approximation in equation (1) only holds *near* x^*. If x_0 is far away, does the sequence still approach x^*? When there are several attracting points, which x^* do we reach? This section starts with good iterations, which solve the equation $x^* = F(x^*)$ or $f(x) = 0$. At the end we discover ***Newton's method***. The next section produces crazy but wonderful iterations, not converging and not blowing up. They lead to "*fractals*" and "*Cantor sets*" and "*chaos*."

The mathematics of iterations is not finished. It may never be finished, but we are converging on the answers. Please choose a function and join in.

THE GRAPH OF AN ITERATION: COBWEBS

The iteration $x_{n+1} = F(x_n)$ involves two graphs at the same time. One is the graph of $y = F(x)$. The other is the graph of $y = x$ (the 45° line). The iteration jumps back and forth between these graphs. It is a very convenient way to see the whole process.

Example 1 was $x_{n+1} = \cos x_n$. Figure 3.19 shows the graph of $\cos x$ and the "*cobweb*." Starting at (x_0, x_0) on the 45° line, the rule is based on $x_1 = F(x_0)$:

From (x_0, x_0) go up or down to (x_0, x_1) ***on the curve***.

From (x_0, x_1) go across to (x_1, x_1) ***on the*** 45° ***line***.

These steps are repeated forever. From x_1 go up to the curve at $F(x_1)$. That height is x_2. Now cross to the 45° line at (x_2, x_2). The iterations are aiming for $(x^*, x^*) = (.7391, .7391)$. This is the *crossing point* of the two graphs $y = F(x)$ and $y = x$.

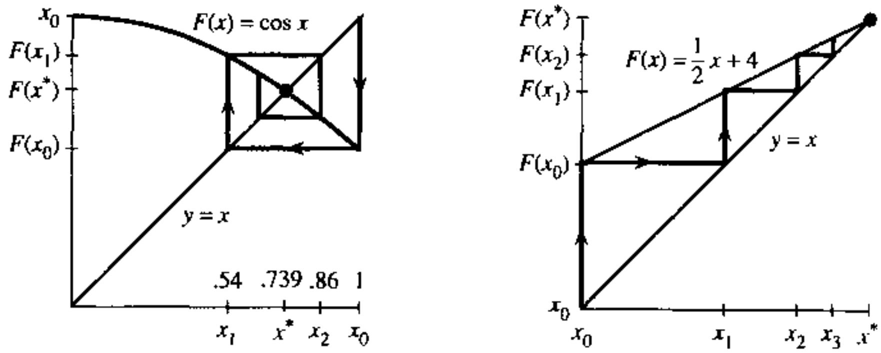


Fig. 3.19 Cobwebs go from (x_0, x_0) to (x_0, x_1) to (x_1, x_1)—line to curve to line.

Example 2 was $x_{n+1} = \frac{1}{2}x_n + 4$. Both graphs are straight lines. The cobweb is one-sided, from (0, 0) to (0, 4) to (4, 4) to (4, 6) to (6, 6). Notice how y changes (vertical line) and then x changes (horizontal line). The slope of $F(x)$ is $\frac{1}{2}$, so the distance to 8 is multiplied by $\frac{1}{2}$ at every step.

Example 3 was $x_{n+1} = x_n^2$. The graph of $y = x^2$ crosses the 45° line at two fixed points: $0^2 = 0$ and $1^2 = 1$. Figure 3.20a starts the iteration close to 1, but it quickly goes away. This fixed point is repelling because $F'(1) = 2$. Distance from $x^* = 1$ is doubled (at the start). One path moves down to $x^* = 0$—which is *superattractive* because $F' = 0$. The path from $x_0 > 1$ diverges to infinity.

EXAMPLE 4 $F(x)$ has two attracting points x^* (a repelling x^* is always between).

Figure 3.20b shows two crossings with slope zero. The iterations and cobwebs converge quickly. In between, the graph of $F(x)$ must cross the 45° line from below. That requires a slope greater than one. Cobwebs diverge from this unstable point, which separates the basins of attraction. The fixed point $x = \pi$ is in a basin by itself!

Note 1 To draw cobwebs on a calculator, graph $y = F(x)$ on top of $y = x$. On a Casio, one way is to plot (x_0, x_0) and give the command `LINE: PLOT X,Y` followed by `EXE`. Now move the cursor vertically to $y = F(x)$ and press `EXE`. Then move horizontally to $y = x$ and press `EXE`. Continue. Each step draws a line.

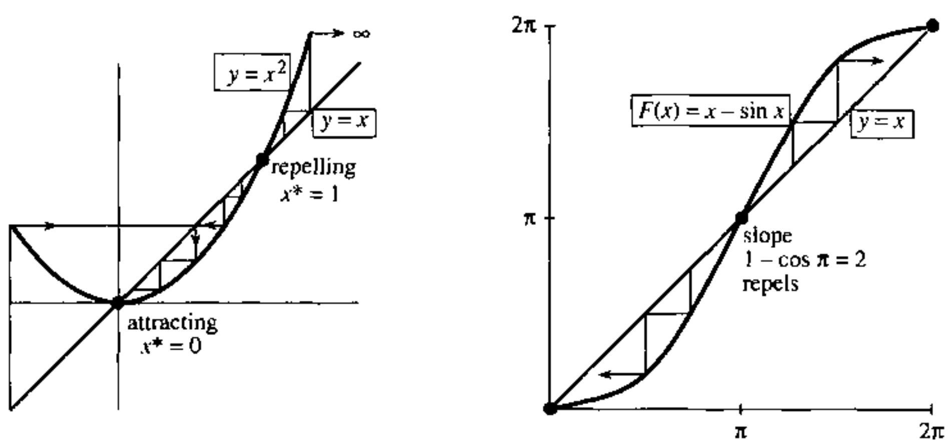


Fig. 3.20 Converging and diverging cobwebs: $F(x) = x^2$ and $F(x) = x - \sin x$.

For the TI-81 (and also the Casio) a short program produces a cobweb. Store $F(x)$ in the Y= function slot Y1. Set the range (square window or autoscaling). Run the program and answer the prompt with x_0:

```
PrgmC:COBWEB  :Disp "INITIAL X0"  :Input X  :All-Off
:Y1-On  :"X"→Y4  :Lbl 1  :X→S  :Y1→T  :Line(S,S,S,T)
:Line(S,T,T,T)  :T→X  :Pause  :Goto 1
```

Note 2 The x's approach x^* from one side when $0 < dF/dx < 1$.

Note 3 A basin of attraction can include faraway x_0's (basins can come in infinitely many pieces). This makes the problem interesting. If no fixed points are attracting, see Section 3.7 for "cycles" and "chaos."

THE ITERATION $x_{n+1} = x_n - cf(x_n)$

At this point we offer the reader a choice. One possibility is to jump ahead to the next section on "Newton's Method." That method is an iteration to solve $f(x) = 0$. The function $F(x)$ combines x_n and $f(x_n)$ and $f'(x_n)$ into an optimal formula for x_{n+1}. We will see how quickly Newton's method works (when it works). It is *the* outstanding algorithm to solve equations, and it is totally built on tangent approximations.

The other possibility is to understand (through calculus) a whole family of iterations. This family depends on a number c, which is at our disposal. ***The best choice of c produces Newton's method.*** I emphasize that iteration is by no means a new and peculiar idea. *It is a fundamental technique in scientific computing.*

We start by recognizing that there are many ways to reach $f(x^*) = 0$. (I write x^* for the solution.) A good algorithm may switch to Newton as it gets close. The iterations use $f(x_n)$ to decide on the next point x_{n+1}:

$$x_{n+1} = F(x_n) = x_n - cf(x_n). \tag{2}$$

Notice how $F(x)$ is constructed from $f(x)$—they are different! We move f to the right side and multiply by a "preconditioner" c. *The choice of c* (or c_n, if it changes from step to step) *is absolutely critical.* The starting guess x_0 is also important—but its accuracy is not always under our control.

Suppose the x_n converge to x^*. Then the limit of equation (2) is

$$x^* = x^* - cf(x^*). \tag{3}$$

That gives $f(x^*) = 0$. If the x_n's have a limit, it solves the right equation. It is a fixed point of F (we can assume $c_n \to c \neq 0$ and $f(x_n) \to f(x^*)$). There are two key questions, and both of them are answered by the slope $F'(x^*)$:

1. How quickly does x_n approach x^* (or do the x_n diverge)?
2. What is a good choice of c (or c_n)?

EXAMPLE 5 $f(x) = ax - b$ is zero at $x^* = b/a$. The iteration $x_{n+1} = x_n - c(ax_n - b)$ intends to find b/a without actually dividing. (Early computers could not divide; they used iteration.) Subtracting x^* from both sides leaves an equation for the error:

$$x_{n+1} - x^* = x_n - x^* - c(ax_n - b).$$

Replace b by ax^*. The right side is $(1 - ca)(x_n - x^*)$. This "error equation" is

$$(\text{error})_{n+1} = (1 - ca)(\text{error})_n. \tag{4}$$

At every step the error is multiplied by $(1 - ca)$, which is F'. *The error goes to zero if* $|F'|$ *is less than* 1. The absolute value $|1 - ca|$ decides everything:

$$x_n \textbf{\textit{ converges to }} x^* \textbf{\textit{ if and only if }} -1 < 1 - ca < 1. \tag{5}$$

The perfect choice (if we knew it) is $c = 1/a$, which turns the multiplier $1 - ca$ into zero. Then one iteration gives the exact answer: $x_1 = x_0 - (1/a)(ax_0 - b) = b/a$. That is the horizontal line in Figure 3.21a, converging in one step. But look at the other lines.

This example did not need calculus. Linear equations never do. The key idea is that *close to* x^* *the nonlinear equation* $f(x) = 0$ *is nearly linear.* We apply the tangent approximation. You are seeing how calculus is used, in a problem that doesn't start by asking for a derivative.

THE BEST CHOICE OF c

The immediate goal is to study the errors $x_n - x^*$. They go quickly to zero, if the multiplier is small. To understand $x_{n+1} = x_n - cf(x_n)$, subtract the equation $x^* = x^* - cf(x^*)$:

$$x_{n+1} - x^* = x_n - x^* - c(f(x_n) - f(x^*)). \tag{6}$$

Now calculus enters. ***When you see a difference of*** f's ***think of*** df/dx. Replace $f(x_n) - f(x^*)$ by $A(x_n - x^*)$, where A stands for the slope df/dx at x^*:

$$x_{n+1} - x^* \approx (1 - cA)(x_n - x^*). \tag{7}$$

This is the ***error equation***. The new error at step $n + 1$ is approximately the old error multiplied by $m = 1 - cA$. This corresponds to $m = 1 - ca$ in the linear example. We keep returning to the basic test $|m| = |F'(x^*)| < 1$:

> **3K** Starting near x^*, the errors $x_n - x^*$ go to zero if the multiplier has $|m| < 1$. The perfect choice is $c = 1/A = 1/f'(x^*)$. Then $m = 1 - cA = 0$.

There is only one difficulty: *We don't know* x^*. Therefore we don't know the perfect c. It depends on the slope $A = f'(x^*)$ at the unknown solution. However we can come close, by using the slope at x_n:

$$\textit{Choose } c_n = 1/f'(x_n). \textit{ Then } x_{n+1} = x_n - f(x_n)/f'(x_n) = F(x_n).$$

This is Newton's method. The multiplier $m = 1 - cA$ is as near to zero as we can make it. By building df/dx into $F(x)$, Newton speeded up the convergence of the iteration.

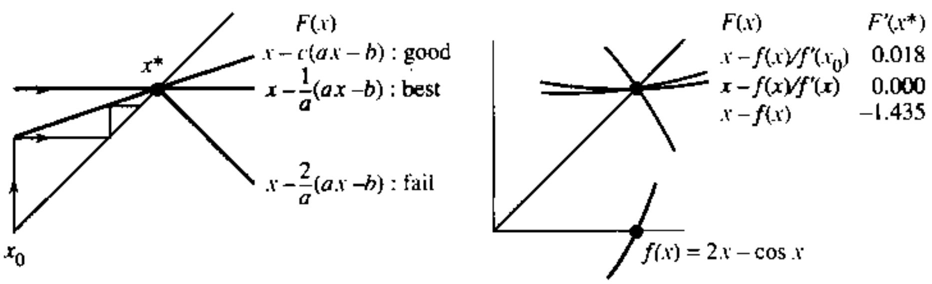


Fig. 3.21 The error multiplier is $m = 1 - cf'(x^*)$. Newton has $c = 1/f'(x_n)$ and $m \to 0$.

EXAMPLE 6 Solve $f(x) = 2x - \cos x = 0$ with different iterations (different c's).

The line $y = 2x$ crosses the cosine curve somewhere near $x = \frac{1}{2}$. The intersection point where $2x^* = \cos x^*$ has no simple formula. We start from $x_0 = \frac{1}{2}$ and iterate $x_{n+1} = x_n - c(2x_n - \cos x_n)$ with *three different choices* of c.

Take $c = 1$ or $c = 1/f'(x_0)$ or update c by Newton's rule $c_n = 1/f'(x_n)$:

$x_0 = .50$	$c = 1$	$c = 1/f'(x_0)$	$c_n = 1/f'(x_n)$
$x_1 =$	.38	.45063	.45062669
$x_2 =$	.55	.45019	.45018365
$x_3 =$	.30	.45018	.45018361...

The column with $c = 1$ is diverging (repelled from x^*). The second column shows convergence (attracted to x^*). The third column (Newton's method) approaches x^* so quickly that .4501836 *and seven more digits* are exact for x_3.

How does this convergence match the prediction? Note that $f'(x) = 2 + \sin x$ so $A = 2.435$. Look to see whether the actual errors $x_n - x^*$, going down each column, are multiplied by the predicted m below that column:

	$c = 1$	$c = 1/(2 + \sin \frac{1}{2})$	$c_n = 1/(2 + \sin x_n)$
$x_0 - x^* =$	0.05	$4.98 \cdot 10^{-2}$	$4.98 \cdot 10^{-2}$
$x_1 - x^* =$	-0.07	$4.43 \cdot 10^{-4}$	$4.43 \cdot 10^{-4}$
$x_2 - x^* =$	0.10	$7.88 \cdot 10^{-6}$	$3.63 \cdot 10^{-8}$
$x_3 - x^* =$	-0.15	$1.41 \cdot 10^{-7}$	$2.78 \cdot 10^{-16}$
multiplier	$m = -1.4$	$m = .018$	$m \to 0$ (Newton)

The first column shows a multiplier below -1. The errors grow at every step. Because m is negative the errors change sign—the cobweb goes outward.

The second column shows convergence with $m = .018$. It takes one genuine Newton step, then c is fixed. After n steps the error is closely proportional to $m^n = (.018)^n$—that is "***linear convergence***" with a good multiplier.

The third column shows the "***quadratic convergence***" of Newton's method. Multiplying the error by m is more attractive than ever, because $m \to 0$. In fact m itself is proportional to the error, so ***at each step the error is squared.*** Problem 3.8.31 will show that $(\text{error})_{n+1} \leq M(\text{error})_n^2$. This squaring carries us from 10^{-2} to 10^{-4} to 10^{-8} to "machine ε" in three steps. The number of correct digits is doubled at every step as Newton converges.

Note 1 The choice $c = 1$ produces $x_{n+1} = x_n - f(x_n)$. This is "successive substitution." The equation $f(x) = 0$ is rewritten as $x = x - f(x)$, and each x_n is substituted back to produce x_{n+1}. Iteration with $c = 1$ does not always fail!

Note 2 Newton's method is successive substitution for f/f', not f. Then $m \approx 0$.

Note 3 Edwards and Penney happened to choose the same example $2x = \cos x$. But they cleverly wrote it as $x_{n+1} = \frac{1}{2} \cos x_n$, which has $|F'| = |\frac{1}{2} \sin x| < 1$. This iteration fits into our family with $c = \frac{1}{2}$, and it succeeds. We asked earlier if its limit is $\frac{1}{2}(.7391)$. *No*, it is $x^* = .450\ldots$.

Note 4 The choice $c = 1/f'(x_0)$ is "***modified Newton.***" After one step of Newton's method, c is fixed. The steps are quicker, because they don't require a new $f'(x_n)$. But we need more steps. Millions of dollars are spent on Newton's method, so speed is important. In all its forms, $f(x) = 0$ is the central problem of computing.

3.6 EXERCISES

Read-through questions

$x_{n+1} = x_n^3$ describes an __a__. After one step $x_1 =$ __b__. After two steps $x_2 = F(x_1) =$ __c__. If it happens that input = output, or $x^* =$ __d__, then x^* is a __e__ point. $F = x^3$ has __f__ fixed points, at $x^* =$ __g__. Starting near a fixed point, the x_n will converge to it if __h__ < 1. That is because $x_{n+1} - x^* = F(x_n) - F(x^*) \approx$ __i__. The point is called __j__. The x_n are repelled if __k__. For $F = x^3$ the fixed points have $F' =$ __l__. The cobweb goes from (x_0, x_0) to (,) to (,) and converges to $(x^*, x^*) =$ __m__. This is an intersection of $y = x^3$ and $y =$ __n__, and it is super-attracting because __o__.

$f(x) = 0$ can be solved iteratively by $x_{n+1} = x_n - cf(x_n)$, in which case $F'(x^*) =$ __p__. Subtracting $x^* = x^* - cf(x^*)$, the error equation is $x_{n+1} - x^* \approx m($__q__$)$. The multiplier is $m =$ __r__. The errors approach zero if __s__. The choice $c_n =$ __t__ produces Newton's method. The choice $c = 1$ is "successive __u__" and $c =$ __v__ is modified Newton. Convergence to x^* is __w__ certain.

We have three ways to study iterations $x_{n+1} = F(x_n)$: (1) compute $x_1, x_2, \ldots$ from different x_0 (2) find the fixed points x^* and test $|dF/dx| < 1$ (3) draw cobwebs.

In Problems 1–8 start from $x_0 = .6$ and $x_0 = 2$. Compute $x_1, x_2, \ldots$ to test convergence:

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

2 $x_{n+1} = 2x_n(1 - x_n)$

3 $x_{n+1} = \sqrt{x_n}$

4 $x_{n+1} = 1/\sqrt{x_n}$

5 $x_{n+1} = 3x_n(1 - x_n)$

6 $x_{n+1} = x_n^2 + x_n - 2$

7 $x_{n+1} = \frac{1}{2}x_n - 1$

8 $x_{n+1} = |x_n|$

9 Check dF/dx at all fixed points in Problems 1–6. Are they attracting or repelling?

10 From $x_0 = -1$ compute the sequence $x_{n+1} = -x_n^3$. Draw the cobweb with its "cycle." Two steps produce $x_{n+2} = x_n^9$, which has the fixed points _______.

11 Draw the cobwebs for $x_{n+1} = \frac{1}{2}x_n - 1$ and $x_{n+1} = 1 - \frac{1}{2}x_n$ starting from $x_0 = 2$. Rule: Cobwebs are two-sided when dF/dx is _______.

12 Draw the cobweb for $x_{n+1} = x_n^2 - 1$ starting from the periodic point $x_0 = 0$. Another periodic point is _______. Start nearby at $x_0 = .1$ to see if the iterations are attracted to $0, -1, 0, -1, \ldots$.

Solve equations 13–16 within 1% by iteration.

13 $x = \cos \frac{1}{2}x$

14 $x = \cos^2 x$

15 $x = \cos \sqrt{x}$

16 $x = 2x - 1$ (??)

17 For which numbers a does $x_{n+1} = a(x_n - x_n^2)$ converge to $x^* = 0$?

18 For which numbers a does $x_{n+1} = a(x_n - x_n^2)$ converge to $x^* = (a-1)/a$?

19 Iterate $x_{n+1} = 4(x_n - x_n^2)$ to see chaos. Why don't the x_n approach $x^* = \frac{3}{4}$?

20 One fixed point of $F(x) = x^2 - \frac{1}{2}$ is attracting, the other is repelling. By experiment or cobwebs, find the basin of x_0's that go to the attractor.

21 (important) Find the fixed point for $F(x) = ax + s$. When is it attracting?

22 What happens in the linear case $x_{n+1} = ax_n + 4$ when $a = 1$ and when $a = -1$?

23 Starting with \$1000, you spend half your money each year and a rich but foolish aunt gives you a new \$1000. What is your steady state balance x^*? What is x^* if you start with a million dollars?

24 The US national debt was once \$1 trillion. Inflation reduces its real value by 5% each year (so multiply by $a = .95$), but overspending adds another \$100 billion. What is the steady state debt x^*?

25 $x_{n+1} = b/x_n$ has the fixed point $x^* = \sqrt{b}$. Show that $|dF/dx| = 1$ at that point—what is the sequence starting from x_0?

26 Show that both fixed points of $x_{n+1} = x_n^2 + x_n - 3$ are repelling. What do the iterations do?

27 A \$5 calculator takes square roots but not cube roots. Explain why $x_{n+1} = \sqrt{2/x_n}$ converges to $\sqrt[3]{2}$.

28 Start the cobwebs for $x_{n+1} = \sin x_n$ and $x_{n+1} = \tan x_n$. In both cases $dF/dx = 1$ at $x^* = 0$. (a) Do the iterations converge? (b) Propose a theory based on F'' for cases when $F' = 1$.

Solve $f(x) = 0$ in 29–32 by the iteration $x_{n+1} = x_n - cf(x_n)$, to find a c that succeeds and a c that fails.

29 $f(x) = x^2 - 4$

30 $f(x) = x^2 - 4x + 3$

31 $f(x) = (x-2)^9 - 1$

32 $f(x) = (1-x)^{-1} - 3$

33 Newton's method computes a new $c = 1/f'(x_n)$ at each step. Write out the iteration formulas for $f(x) = x^3 - 2 = 0$ and $f(x) = \sin x - \frac{1}{2} = 0$.

34 Apply Problem 33 to find the first six decimals of $\sqrt[3]{2}$ and $\pi/6$.

35 By experiment find each x^* and its basin of attraction, when Newton's method is applied to $f(x) = x^2 - 5x + 4$.

36 Test Newton's method on $x^2 - 1 = 0$, starting far out at $x_0 = 10^6$. At first the error is reduced by about $m = \frac{1}{2}$. Near $x^* = 1$ the multiplier approaches $m = 0$.

37 Find the multiplier m at each fixed point of $x_{n+1} = x_n - c(x_n^2 - x_n)$. Predict the convergence for different c (to which x^*?).

38 Make a table of iterations for $c = 1$ and $c = 1/f'(x_0)$ and $c = 1/f'(x_n)$, when $f(x) = x^2 - \frac{1}{2}$ and $x_0 = 1$.

39 In the iteration for $x^2 - 2 = 0$, find dF/dx at x^*:

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

(b) Newton's iteration has $F(x) = x - f(x)/f'(x)$. Show that $F' = 0$ when $f(x) = 0$. *The multiplier for Newton is* $m = 0$.

40 What are the solutions of $f(x) = x^2 + 2 = 0$ and why is Newton's method sure to fail? But carry out the iteration to see whether $x_n \to \infty$.

41 ***Computer project*** $F(x) = x - \tan x$ has fixed points where $\tan x^* = 0$. So x^* is any multiple of π. From $x_0 = 2.0$ and 1.8 and 1.9, which multiple do you reach? Test points in $1.7 < x_0 < 1.9$ to find basins of attraction to π, 2π, 3π, 4π.

Between any two basins there are basins for *every* multiple of π. And more basins between these (***a fractal***). Mark them on the line from 0 to π. Magnify the picture around $x_0 = 1.9$ (in color?).

42 Graph $\cos x$ and $\cos(\cos x)$ and $\cos(\cos(\cos x))$. Also $(\cos)^8 x$. What are these graphs approaching?

43 Graph $\sin x$ and $\sin(\sin x)$ and $(\sin)^8 x$. What are these graphs approaching? Why so slow?

3.7 Newton's Method (and Chaos)

The equation to be solved is $f(x) = 0$. Its solution x^* is the point where the graph crosses the x axis. Figure 3.22 shows x^* and a starting guess x_0. Our goal is to come as close as possible to x^*, *based on the information* $f(x_0)$ *and* $f'(x_0)$.

Section 3.6 reached Newton's formula for x_1 (the next guess). We now do that directly.

What do we see at x_0? The graph has height $f(x_0)$ and slope $f'(x_0)$. We know where we are, and which direction the curve is going. We don't know if the curve bends (we don't have f''). The best plan is ***to follow the tangent line***, which uses all the information we have.

Newton replaces $f(x)$ by its linear approximation (= tangent approximation):

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0). \tag{1}$$

We want the left side to be zero. The best we can do is to make the right side zero! The tangent line crosses the axis at x_1, while the curve crosses at x^*. The new guess x_1 comes from $f(x_0) + f'(x_0)(x_1 - x_0) = 0$. Dividing by $f'(x_0)$ and solving for x_1, this is step 1 of Newton's method:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \tag{2}$$

At this new point, compute $f(x_1)$ and $f'(x_1)$—the height and slope at x_1. They give a new tangent line, which crosses at x_2. ***At every step we want*** $f(x_{n+1}) = 0$ ***and we settle for*** $f(x_n) + f'(x_n)(x_{n+1} - x_n) = 0$. After dividing by $f'(x_n)$, the formula for x_{n+1} is Newton's method.

3L The tangent line from x_n crosses the axis at x_{n+1}:

$$\textbf{\textit{Newton's method}} \qquad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{3}$$

Usually this iteration $x_{n+1} = F(x_n)$ converges quickly to x^*.

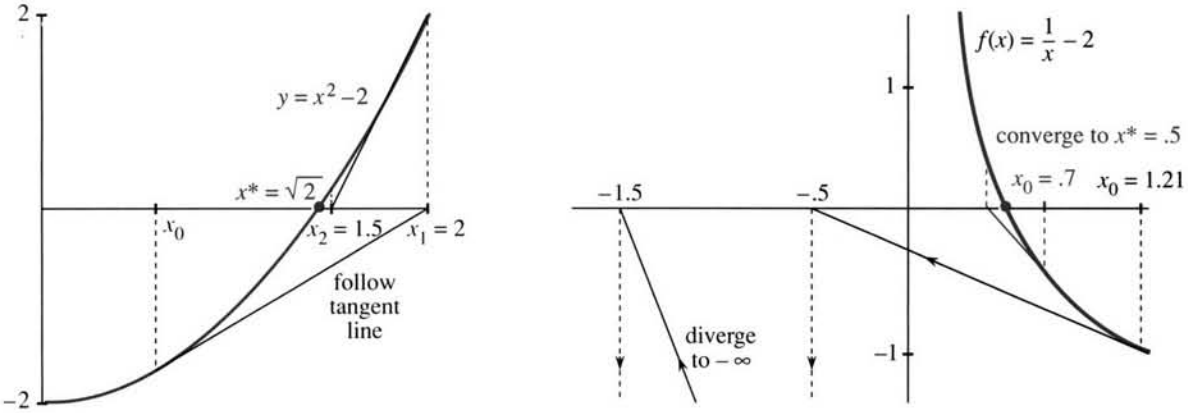


Fig. 3.22 Newton's method along tangent lines from x_0 to x_1 to x_2.

Linear approximation involves three numbers. They are Δx (across) and Δf (up) and the slope $f'(x)$. If we know two of those numbers, we can estimate the third. It is remarkable to realize that calculus has now used all three calculations—they are the key to this subject:

1. *Estimate the slope $f'(x)$ from $\Delta f/\Delta x$* (Section 2.1)
2. *Estimate the change Δf from $f'(x)\,\Delta x$* (Section 3.1)
3. *Estimate the change Δx from $\Delta f/f'(x)$* (Newton's method)

The desired Δf is $-f(x_n)$. Formula (3) is exactly $\Delta x = -f(x_n)/f'(x_n)$.

EXAMPLE 1 (Square roots) $f(x) = x^2 - b$ is zero at $x^* = \sqrt{b}$ and also at $-\sqrt{b}$. Newton's method is a quick way to find square roots—probably built into your calculator. The slope is $f'(x_n) = 2x_n$, and formula (3) for the new guess becomes

$$x_{n+1} = x_n - \frac{x_n^2 - b}{2x_n} = x_n - \frac{1}{2}x_n + \frac{b}{2x_n}. \tag{4}$$

This simplifies to $x_{n+1} = \frac{1}{2}(x_n + b/x_n)$. ***Guess the square root, divide into*** b, ***and average the two numbers***. The ancient Babylonians had this same idea, without knowing functions or slopes. They iterated $x_{n+1} = F(x_n)$:

$$F(x) = \frac{1}{2}\left(x + \frac{b}{x}\right) \quad \text{and} \quad F'(x) = \frac{1}{2}\left(1 - \frac{b}{x^2}\right). \tag{5}$$

The Babylonians did exactly the right thing. The slope F' is zero *at the solution*, when $x^2 = b$. That makes Newton's method converge at high speed. The convergence test is $|F'(x^*)| < 1$. Newton achieves $F'(x^*) = 0$—which is *superconvergence*.

To find $\sqrt{4}$, start the iteration $x_{n+1} = \frac{1}{2}(x_n + 4/x_n)$ at $x_0 = 1$. Then $x_1 = \frac{1}{2}(1+4)$:

$$x_1 = 2.5 \quad x_2 = 2.05 \quad x_3 = 2.0006 \quad x_4 = 2.000000009.$$

The wrong decimal is twice as far out at each step. ***The error is squared.*** Subtracting $x^* = 2$ from both sides of $x_{n+1} = F(x_n)$ gives an *error equation* which displays that square:

$$x_{n+1} - 2 = \frac{1}{2}\left(x_n + \frac{4}{x_n}\right) - 2 = \frac{1}{2x_n}(x_n - 2)^2. \tag{6}$$

This is $(error)_{n+1} \approx \frac{1}{4}(error)_n^2$. It explains the speed of Newton's method.

Remark **1** You can't start this iteration at $x_0 = 0$. The first step computes 4/0 and blows up. Figure 3.22a shows why—the tangent line at zero is horizontal. It will never cross the axis.

Remark **2** Starting at $x_0 = -1$, Newton converges to $-\sqrt{2}$ instead of $+\sqrt{2}$. That is the other x^*. Often it is difficult to predict which x^* Newton's method will choose. Around every solution is a "basin of attraction," but other parts of the basin may be far away. Numerical experiments are needed, with many starts x_0. Finding basins of attraction was one of the problems that led to fractals.

EXAMPLE 2 *Solve* $\frac{1}{x} - a = 0$ *to find* $x^* = \frac{1}{a}$ *without dividing by a.*

Here $f(x) = (1/x) - a$. Newton uses $f'(x) = -1/x^2$. Surprisingly, we don't divide:

$$x_{n+1} = x_n - \frac{(1/x_n) - a}{-1/x_n^2} = x_n + x_n - ax_n^2. \tag{7}$$

Do these iterations converge? I will take $a = 2$ and aim for $x^* = \frac{1}{2}$. Subtracting $\frac{1}{2}$ from both sides of (7) changes the iteration into the error equation:

$$x_{n+1} = 2x_n - 2x_n^2 \quad \text{becomes} \quad x_{n+1} - \tfrac{1}{2} = -2(x_n - \tfrac{1}{2})^2. \tag{8}$$

At each step the error is squared. This is terrific if (and only if) you are close to $x^* = \frac{1}{2}$. Otherwise squaring a large error and multiplying by -2 is not good:

$$x_0 = .70 \quad x_1 = .42 \quad x_2 = .487 \quad x_3 = .4997 \quad x_4 = .4999998$$

$$x_0 = 1.21 \quad x_1 = -.5 \quad x_2 = -1.5 \quad x_3 = -7.5 \quad x_4 = -127.5$$

The algebra in Problem 18 confirms those experiments. There is fast convergence if $0 < x_0 < 1$. There is divergence if x_0 is negative or $x_0 > 1$. The tangent line goes to a negative x_1. After that Figure 3.22 shows a long trip backwards.

In the previous section we drew $F(x)$. The iteration $x_{n+1} = F(x_n)$ converged to the 45° line, where $x^* = F(x^*)$. In this section we are drawing $f(x)$. Now x^* is the point on the axis where $f(x^*) = 0$.

To repeat: It is $f(x^*) = 0$ that we aim for. But it is the slope $F'(x^*)$ that decides whether we get there. Example 2 has $F(x) = 2x - 2x^2$. The fixed points are $x^* = \frac{1}{2}$ (our solution) and $x^* = 0$ (not attractive). The slopes $F'(x^*)$ are zero (typical Newton) and 2 (typical repeller). *The key to Newton's method is $F' = 0$ at the solution*:

$$\textit{The slope of} \quad F(x) = x - \frac{f(x)}{f'(x)} \quad \textit{is} \quad \frac{f(x)f''(x)}{(f'(x))^2}. \quad \textit{Then } F'(x) = 0 \textit{ when } f(x) = 0.$$

The examples $x^2 = b$ and $1/x = a$ show fast convergence or failure. In Chapter 13, and in reality, Newton's method solves much harder equations. Here I am going to choose a third example that came from pure curiosity about what might happen. The results are absolutely amazing. The equation is $x^2 = -1$.

EXAMPLE 3 ***What happens to Newton's method if you ask it to solve $f(x) = x^2 + 1 = 0$?***

The only solutions are the imaginary numbers $x^* = i$ and $x^* = -i$. There is no real square root of -1. Newton's method might as well give up. But it has no way to know that! The tangent line still crosses the axis at a new point x_{n+1}, even if the curve $y = x^2 + 1$ never crosses. Equation (5) still gives the iteration for $b = -1$:

$$x_{n+1} = \frac{1}{2}\left(x_n - \frac{1}{x_n}\right) = F(x_n). \tag{9}$$

The x's cannot approach i or $-i$ (nothing is imaginary). So what do they do?

The starting guess $x_0 = 1$ is interesting. It is followed by $x_1 = 0$. Then x_2 divides by zero and blows up. I expected other sequences to go to infinity. But the experiments showed something different (and mystifying). When x_n is large, x_{n+1} is less than half as large. After $x_n = 10$ comes $x_{n+1} = \frac{1}{2}(10 - \frac{1}{10}) = 4.95$. After much indecision and a long wait, a number near zero eventually appears. Then the next guess divides by that small number and goes far out again. This reminded me of "chaos."

It is tempting to retreat to ordinary examples, where Newton's method is a big success. By trying exercises from the book or equations of your own, you will see that the fast convergence to $\sqrt{4}$ is very typical. The function can be much more complicated than $x^2 - 4$ (in practice it certainly is). The iteration for $2x = \cos x$ was in the previous section, and the error was squared at every step. If Newton's method starts close to x^*, its convergence is overwhelming. That has to be the main point of this section: ***Follow the tangent line.***

Instead of those good functions, may I stay with this strange example $x^2 + 1 = 0$? It is not so predictable, and maybe not so important, but somehow it is more interesting. There is no real solution x^*, and Newton's method $x_{n+1} = \frac{1}{2}(x_n - 1/x_n)$ bounces around. We will now discover x_n.

A FORMULA FOR x_n

The key is an exercise from trigonometry books. Most of those problems just give practice with sines and cosines, but this one exactly fits $\frac{1}{2}(x_n - 1/x_n)$:

$$\frac{1}{2}\left(\frac{\cos\theta}{\sin\theta} - \frac{\sin\theta}{\cos\theta}\right) = \frac{\cos 2\theta}{\sin 2\theta} \qquad \text{or} \qquad \frac{1}{2}\left(\cot\theta - \frac{1}{\cot\theta}\right) = \cot 2\theta$$

In the left equation, the common denominator is $2\sin\theta\cos\theta$ (which is $\sin 2\theta$). The numerator is $\cos^2\theta - \sin^2\theta$ (which is $\cos 2\theta$). Replace cosine/sine by cotangent, and the identity says this:

If $\quad x_0 = \cot\theta \quad$ then $\quad x_1 = \cot 2\theta. \qquad$ Then $\quad x_2 = \cot 4\theta. \qquad$ Then $\quad x_n = \cot 2^n\theta.$

This is the formula. *Our points are on the cotangent curve.* Figure 3.23 starts from $x_0 = 2 = \cot\theta$, and every iteration doubles the angle.

Example A The sequence $x_0 = 1, x_1 = 0, x_2 = \infty$ matches the cotangents of $\pi/4$, $\pi/2$, and π. This sequence blows up because x_2 has a division by $x_1 = 0$.

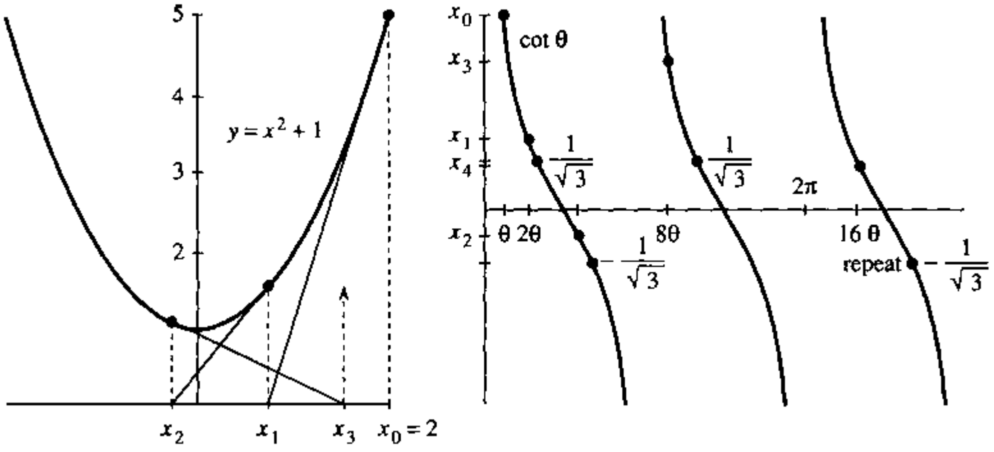


Fig. 3.23 Newton's method for $x^2+1=0$. Iteration gives $x_n = \cot 2^n\theta$.

Example B The sequence $1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3}$ matches the cotangents of $\pi/3$, $2\pi/3$, and $4\pi/3$. This sequence *cycles forever* because $x_0 = x_2 = x_4 = \ldots$.

Example C Start with a large x_0 (a small θ). Then x_1 is about half as large (at 2θ). Eventually one of the angles $4\theta, 8\theta, \ldots$ hits on a large cotangent, and the x's go far out again. *This is typical.* Examples A and B were special, when θ/π was $\frac{1}{4}$ or $\frac{1}{3}$.

What we have here is ***chaos***. The x's can't converge. They are strongly repelled by all points. They are also extremely sensitive to the value of θ. After ten steps θ is multiplied by $2^{10} = 1024$. The starting angles 60° and 61° look close, but now they are different by 1024°. If that were a multiple of 180°, the cotangents would still be close. In fact the x_{10}'s are 0.6 and 14.

This chaos in mathematics is also seen in nature. The most familiar example is the weather, which is much more delicate than you might think. The headline "Forecasting Pushed Too Far" appeared in *Science* (1989). The article said that the snowballing of small errors destroys the forecast after six days. We can't follow the weather equations for a month—the flight of a plane can change everything. This is a revolutionary idea, that a simple rule can lead to answers that are too sensitive to compute.

We are accustomed to complicated formulas (or no formulas). We are not accustomed to innocent-looking formulas like $\cot 2^n\theta$, which are absolutely hopeless after 100 steps.

CHAOS FROM A PARABOLA

Now I get to tell you about new mathematics. First I will change the iteration $x_{n+1} = \frac{1}{2}(x_n - 1/x_n)$ into one that is even simpler. By switching from x to $z = 1/(1+x^2)$, each new z turns out to involve only the old z and z^2:

$$z_{n+1} = 4z_n - 4z_n^2. \tag{10}$$

This is the most famous quadratic iteration in the world. There are books about it, and Problem 28 shows where it comes from. Our formula for x_n leads to z_n:

$$z_n = \frac{1}{1+x_n^2} = \frac{1}{1+(\cot 2^n\theta)^2} = (\sin 2^n\theta)^2. \tag{11}$$

The sine is just as unpredictable as the cotangent, when $2^n\theta$ gets large. The new thing is to locate this quadratic as the last member (when $a=4$) of the family

$$z_{n+1} = az_n - az_n^2, \quad 0 \leqslant a \leqslant 4. \tag{12}$$

Example 2 happened to be the middle member $a=2$, converging to $\frac{1}{2}$. I would like to give a brief and very optional report on this iteration, for different a's.

The general principle is to start with a number z_0 between 0 and 1, and compute $z_1, z_2, z_3, \ldots$. It is fascinating to watch the behavior change as a increases. ***You can see it on your own computer.*** Here we describe some things to look for. All numbers stay between 0 and 1 and they may approach a limit. That happens when a is small:

for $0 \leqslant a \leqslant 1$ the z_n approach $z^* = 0$
for $1 \leqslant a \leqslant 3$ the z_n approach $z^* = (a-1)/a$

Those limit points are the solutions of $z = F(z)$. They are the fixed points where $z^* = az^* - a(z^*)^2$. But remember the test for approaching a limit: *The slope at z^* cannot be larger than one.* Here $F = az - az^2$ has $F' = a - 2az$. It is easy to check $|F'| \leqslant 1$ at the limits predicted above. The hard problem—sometimes impossible—is to predict what happens above $a=3$. Our case is $a=4$.

The** z's **cannot approach a limit when $|F'(z^*)| > 1$. Something has to happen, and there are at least three possibilities:

The z_n's can cycle or fill the whole interval (0, 1) *or approach a Cantor set.*

I start with a random number z_0, take 100 steps, and write down steps 101 to 105:

	$a=3.4$	$a=3.5$	$a=3.8$	$a=4.0$
$z_{101}=$	.842	.875	.336	.169
$z_{102}=$	.452	.383	.848	.562
$z_{103}=$	.842	.827	.491	.985
$z_{104}=$	.452	.501	.950	.060
$z_{105}=$	.842	.875	.182	.225

The first column is converging to a "2-cycle." It alternates between $x = .842$ and $y = .452$. Those satisfy $y = F(x)$ and $x = F(y) = F(F(x))$. If we look at a *double step* when $a = 3.4$, x and y are fixed points of the double iteration $z_{n+2} = F(F(z_n))$. When a increases past 3.45, this cycle becomes unstable.

At that point the period doubles from 2 ***to*** 4. With $a = 3.5$ you see a "4-cycle" in the table—it repeats after four steps. The sequence bounces from .875 to .383 to .827 to .501 *and back to* .875. This cycle must be attractive or we would not see it. But it also becomes unstable as a increases. Next comes an 8-cycle, which is stable in a little window (you could compute it) around $a = 3.55$. ***The cycles are stable for shorter and shorter intervals of*** a's. Those stability windows are reduced by the *Feigenbaum shrinking factor* 4.6692.... Cycles of length 16 and 32 and 64 can be seen in physical experiments, but they are all unstable before $a = 3.57$. What happens then?

The new and unexpected behavior is between 3.57 and 4. Down each line of Figure 3.24, the computer has plotted the values of z_{1001} to z_{2000}—omitting the first thousand points to let a stable period (or chaos) become established. No points appeared in the big white wedge. I don't know why. In the window for period 3, you

The period 2, 4, ... is the number of z's in a cycle.

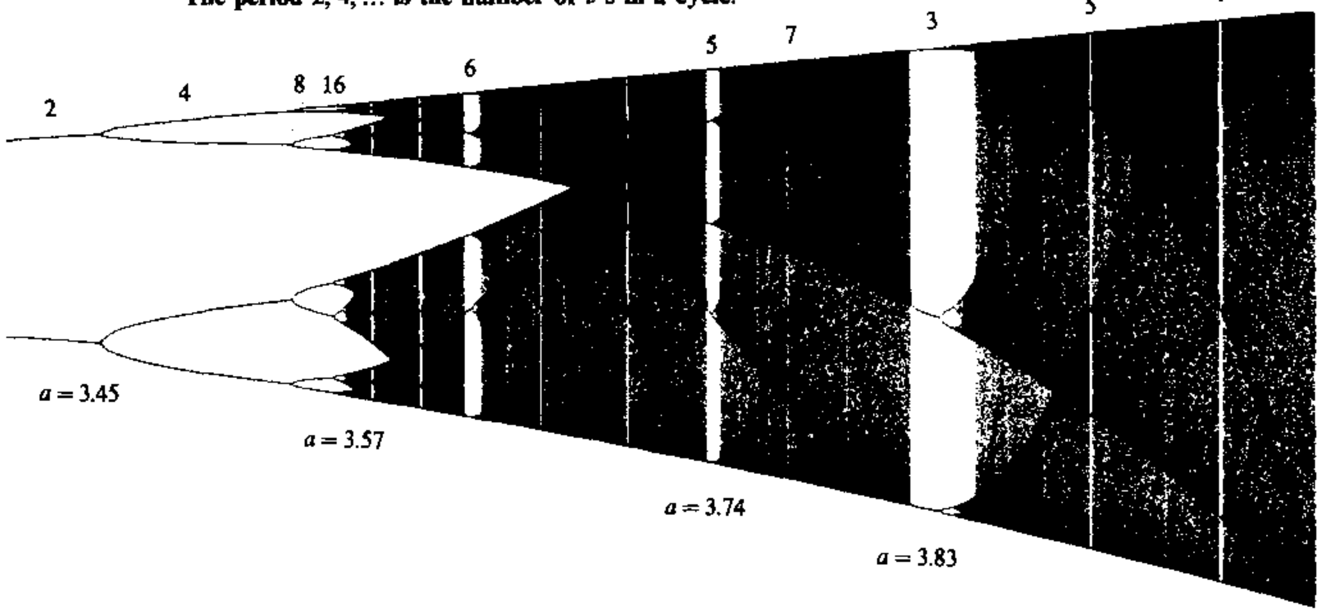


Fig. 3.24 Period doubling and chaos from iterating $F(z)$ (stolen by special permission from *Introduction to Applied Mathematics* by Gilbert Strang, Wellesley-Cambridge Press).

see only three z's. Period 3 is followed by 6, 12, 24, ... There is ***period doubling*** at the end of every window (including all the windows that are too small to see). You can reproduce this figure by iterating $z_{n+1} = az_n - az_n^2$ from any z_0 and plotting the results.

CANTOR SETS AND FRACTALS

I can't tell what happens at $a = 3.8$. There may be a stable cycle of some long period. The z's may come close to every point between 0 and 1. A third possibility is to approach a very thin limit set, which looks like the famous ***Cantor set***:

> To construct the Cantor set, divide [0, 1] into three pieces and remove the open interval $(\frac{1}{3}, \frac{2}{3})$. Then remove $(\frac{1}{9}, \frac{2}{9})$ and $(\frac{7}{9}, \frac{8}{9})$ from what remains. At each step ***take out the middle thirds***. The points that are left form the Cantor set.

All the endpoints $\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \ldots$ are in the set. So is $\frac{3}{4}$ (Problem 42). Nevertheless the lengths of the removed intervals add to 1 and the Cantor set has "measure zero." What is especially striking is its ***self-similarity***: *Between* 0 *and* $\frac{1}{3}$ *you see the same Cantor set three times smaller*. From 0 to $\frac{1}{9}$ the Cantor set is there again, scaled down by 9. Every section, when blown up, copies the larger picture.

Fractals That self-similarity is typical of a *fractal*. There is an infinite sequence of scales. A mathematical snowflake starts with a triangle and adds a bump in the middle of each side. At every step the bumps lengthen the sides by 4/3. The final boundary is self-similar, like an infinitely long coastline.

The word "fractal" comes from ***fractional dimension***. The snowflake boundary has dimension larger than 1 and smaller than 2. The Cantor set has dimension larger than 0 and smaller than 1. Covering an ordinary line segment with circles of radius r would take c/r circles. For fractals it takes c/r^D circles—and D is the dimension.

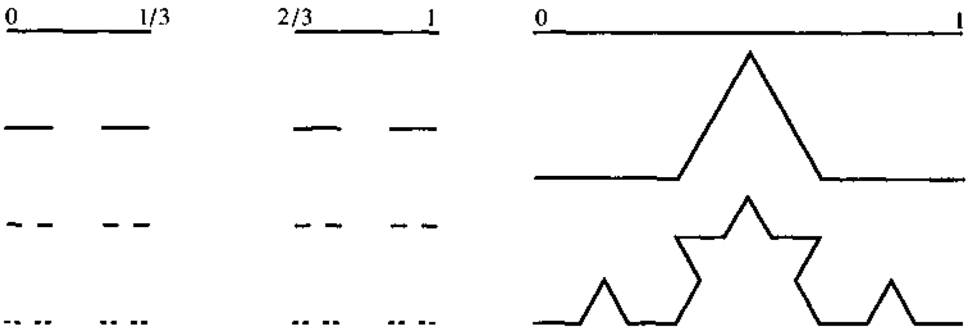


Fig. 3.25 Cantor set (middle thirds removed). Fractal snowflake (infinite boundary).

Our iteration $z_{n+1} = 4z_n - 4z_n^2$ has $a = 4$, at the end of Figure 3.24. The sequence $z_0, z_1, \ldots$ goes everywhere and nowhere. Its behavior is chaotic, and statistical tests find no pattern. For all practical purposes the numbers are random.

Think what this means in an experiment (or the stock market). If simple rules produce chaos, there is *absolutely no way* to predict the results. No measurement can ever be sufficiently accurate. The newspapers report that Pluto's orbit is chaotic—even though it obeys the law of gravity. The motion is totally unpredictable over long times. I don't know what that does for astronomy (or astrology).

The most readable book on this subject is Gleick's best-seller *Chaos: Making a New Science*. The most dazzling books are *The Beauty of Fractals* and *The Science of Fractal Images*, in which Peitgen and Richter and Saupe show photographs that have been in art museums around the world. The most original books are Mandelbrot's *Fractals* and *Fractal Geometry*. Our cover has a fractal from Figure 13.11.

We return to friendlier problems in which calculus is not helpless.

NEWTON'S METHOD VS. SECANT METHOD: CALCULATOR PROGRAMS

The hard part of Newton's method is to find df/dx. We need it for the slope of the tangent line. But calculus can approximate by $\Delta f/\Delta x$—using the values of $f(x)$ already computed at x_n and x_{n-1}.

The ***secant method*** follows the secant line instead of the tangent line:

$$\textbf{\textit{Secant}}: \qquad x_{n+1} = x_n - \frac{f(x_n)}{(\Delta f/\Delta x)_n} \quad \text{where} \quad \left(\frac{\Delta f}{\Delta x}\right)_n = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}. \qquad (13)$$

The secant line connects the two latest points on the graph of $f(x)$. Its equation is $y - f(x_n) = (\Delta f/\Delta x)(x - x_n)$. Set $y = 0$ to find equation (13) for the new $x = x_{n+1}$, where the line crosses the axis.

Prediction: *Three* secant steps are about as good as *two* Newton steps. Both should give four times as many correct decimals: $(error) \to (error)^4$. Probably the secant method is also chaotic for $x^2 + 1 = 0$.

These Newton and secant programs are for the TI-81. Place the formula for $f(x)$ in slot `Y1` and the formula for $f'(x)$ in slot `Y2` on the `Y` = function edit screen. Answer the prompt with the initial x_0 = `X0`. The programs pause to display each approximation x_n, the value $f(x_n)$, and the difference $x_n - x_{n-1}$. Press `ENTER` to continue or press `ON` and select item `2: Quit` to break. If $f(x_n) = 0$, the programs display `ROOT AT` and the root x_n.

```
PrgmN:NEWTON  :Disp "ENTER FOR MORE"
:Disp "X0="   :Disp "ON 2 TO BREAK"
:Input X      :Disp " "
:X→S          :Disp "XN FXN XN-XNM1"
:Y1→Y         :Disp X
:Lbl 1        :Disp Y
:X-Y/Y2→X     :Disp D
:X-S→D        :Pause
:X→S          :If Y≠0
:Y1→Y         :Goto 1
              :Disp "ROOT AT"
              :Disp X
```

```
PrgmS:SECANT      :Y→T
:Disp "X0="       :Y1→Y
:Input X          :Disp "ENTER FOR MORE"
:X→S              :Disp "XN FXN XN-XNM1"
:Y1→T             :Disp X
:Disp "X1="       :Disp Y
:Input X          :Disp D
:Y1→Y             :Pause
:Lbl 1            :If Y≠0
:X-S→D            :Goto 1
:X→S              :Disp "ROOT AT"
:X-YD/(Y-T)→X     :Disp X
```

3.7 EXERCISES

Read-through questions

When $f(x)=0$ is linearized to $f(x_n)+f'(x_n)(x-x_n)=0$, the solution $x =$ __a__ is Newton's x_{n+1}. The __b__ to the curve crosses the axis at x_{n+1}, while the __c__ crosses at x^*. The errors at x_n and x_{n+1} are normally related by $(\text{error})_{n+1} \approx M$ __d__. This is __e__ convergence. The number of correct decimals __f__ at every step.

For $f(x)=x^2-b$, Newton's iteration is $x_{n+1} =$ __g__. The x_n converge to __h__ if $x_0>0$ and to __i__ if $x_0<0$. For $f(x)=x^2+1$, the iteration becomes $x_{n+1} =$ __j__. This cannot converge to __k__. Instead it leads to chaos. Changing to $z=1/(x^2+1)$ yields the parabolic iteration $z_{n+1} =$ __l__.

For $a \leqslant 3$, $z_{n+1}=az_n-az_n^2$ converges to a single __m__. After $a=3$ the limit is a 2-cycle, which means __n__. Later the limit is a Cantor set, which is a one-dimensional example of a __o__. The Cantor set is self-__p__.

1 To solve $f(x)=x^3-b=0$, what iteration comes from Newton's method?

2 For $f(x)=(x-1)/(x+1)$ Newton's formula is $x_{n+1}=F(x_n)=$ ________. Solve $x^*=F(x^*)$ and find $F'(x^*)$. What limit do the x_n's approach?

3 I believe that Newton only applied his method in public to one equation $x^3-2x-5=0$. Raphson carried the idea forward but got partial credit at best. After two steps from $x_0=2$, how many decimals in $x^*=2.09455148$ are correct?

4 Show that Newton's method for $f(x)=x^{1/3}$ gives the strange formula $x_{n+1}=-2x_n$. Draw a graph to show the iterations.

5 Find x_1 if (a) $f(x_0)=0$; (b) $f'(x_0)=0$.

6 Graph $f(x)=x^3-3x-1$ and estimate its roots x^*. Run Newton's method starting from 0, 1, $-\frac{1}{2}$, and 1.1. Experiment to decide which x_0 converge to which root.

7 Solve $x^2-6x+5=0$ by Newton's method with $x_0=2.5$ and 3. Draw a graph to show which x_0 lead to which root.

8 If $f(x)$ is increasing and concave up ($f'>0$ and $f''>0$) show by a graph that Newton's method converges. From which side?

Solve 9–17 to four decimal places by Newton's method with a computer or calculator. Choose any x_0 except x^*.

9 $x^2-10=0$

10 $x^4-100=0$ (faster or slower than Problem 9?)

11 $x^2-x=0$ (which x_0 to which root?)

12 $x^3-x=0$ (which x_0 to which root?)

13 $x+5\cos x=0$ (this has three roots)

14 $x+\tan x=0$ (find two roots) (are there more?)

15 $1/(1-x)=2$

16 $1+x+x^2+x^3+x^4=2$

17 $x^3+(x+1)^3=10^3$

18 (a) Show that $x_{n+1}=2x_n-2x_n^2$ in Example 2 is the same as $(1-2x_{n+1})=(1-2x_n)^2$.
(b) Prove divergence if $|1-2x_0|>1$. Prove convergence if $|1-2x_0|<1$ or $0<x_0<1$.

19 With $a=3$ in Example 2, experiment with the Newton iteration $x_{n+1}=2x_n-3x_n^2$ to decide which x_0 lead to $x^*=\frac{1}{3}$.

20 Rewrite $x_{n+1}=2x_n-ax_n^2$ as $(1-ax_{n+1})=(1-ax_n)^2$. For which x_0 does the sequence $1-ax_n$ approach zero (so $x_n\to 1/a$)?

21 What is Newton's method to find the kth root of 7? Calculate $\sqrt[7]{7}$ to 7 places.

22 Find all solutions of $x^3=4x-1$ (5 decimals).

Problems 23–29 are about $x^2+1=0$ and chaos.

23 For $\theta=\pi/16$ when does $x_n=\cot 2^n\theta$ blow up? For $\theta=\pi/7$ when does $\cot 2^n\theta=\cot\theta$? (The angles $2^n\theta$ and θ differ by a multiple of π.)

24 For $\theta=\pi/9$ follow the sequence until $x_n=x_0$.

25 For $\theta=1$, x_n never returns to $x_0=\cot 1$. The angles 2^n and 1 never differ by a multiple of π because ______.

26 If z_0 equals $\sin^2\theta$, show that $z_1=4z_0-4z_0^2$ equals $\sin^2 2\theta$.

27 If $y=x^2+1$, each new y is

$$y_{n+1}=x_{n+1}^2+1=\frac{1}{4}\left(x_n-\frac{1}{x_n}\right)^2+1.$$

Show that this equals $y_n^2/4(y_n-1)$.

28 Turn Problem 27 upside down, $1/y_{n+1}=4(y_n-1)/y_n^2$, to find the quadratic iteration (10) for $z_n=1/y_n=1/(1+x_n^2)$.

29 If $F(z)=4z-4z^2$ what is $F(F(z))$? How many solutions to $z=F(F(z))$? How many are not solutions to $z=F(z)$?

30 Apply Newton's method to $x^3-.64x-.36=0$ to find the basin of attraction for $x^*=1$. Also find a pair of points for which $y=F(z)$ and $z=F(y)$. In this example Newton does not always find a root.

31 Newton's method solves $x/(1-x)=0$ by $x_{n+1}=$ ______. From which x_0 does it converge? The distance to $x^*=0$ is exactly squared.

Problems 33–41 are about competitors of Newton.

32 At a double root, Newton only converges linearly. What is the iteration to solve $x^2=0$?

33 To speed up Newton's method, find the step Δx from $f(x_n)+\Delta x f'(x_n)+\frac{1}{2}(\Delta x)^2 f''(x_n)=0$. Test on $f(x)=x^2-1$ from $x_0=0$ and explain.

34 Halley's method uses $f_n+\Delta x f_n'+\frac{1}{2}\Delta x(-f_n/f_n')f_n''=0$. For $f(x)=x^2-1$ and $x_0=1+\varepsilon$, show that $x_1=1+O(\varepsilon^3)$—which is *cubic* convergence.

35 Apply the secant method to $f(x)=x^2-4=0$, starting from $x_0=1$ and $x_1=2.5$. Find $\Delta f/\Delta x$ and the next point x_2 by hand. Newton uses $f'(x_1)=5$ to reach $x_2=2.05$. Which is closer to $x^*=2$?

36 Draw a graph of $f(x)=x^2-4$ to show the secant line in Problem 35 and the point x_2 where it crosses the axis.

Bisection method **If $f(x)$ changes sign between x_0 and x_1, find its sign at the midpoint $x_2=\frac{1}{2}(x_0+x_1)$. Decide whether $f(x)$ changes sign between x_0 and x_2 or x_2 and x_1. Repeat on that half-length (bisected) interval.** Continue. **Switch to a faster method when the interval is small enough.**

37 $f(x)=x^2-4$ is negative at $x=1$, positive at $x=2.5$, and negative at the midpoint $x=1.75$. So x^* lies in what interval? Take a second step to cut the interval in half again.

38 Write a code for the bisection method. At each step print out an interval that contains x^*. The inputs are x_0 and x_1; the code calls $f(x)$. Stop if $f(x_0)$ and $f(x_1)$ have the same sign.

39 Three bisection steps reduce the interval by what factor? Starting from $x_0=0$ and $x_1=8$, take three steps for $f(x)=x^2-10$.

40 A direct method is to *zoom in* where the graph crosses the axis. Solve $10x^3-8.3x^2+2.295x-.21141=0$ by several zooms.

41 If the zoom factor is 10, then the number of correct decimals ______ for every zoom. Compare with Newton.

42 The number $\frac{3}{4}$ equals $\frac{2}{3}(1+\frac{1}{9}+\frac{1}{81}+\cdots)$. Show that it is in the Cantor set. It survives when middle thirds are removed.

43 The solution to $f(x)=(x-1.9)/(x-2.0)=0$ is $x^*=1.9$. Try Newton's method from $x_0=1.5$, 2.1, and 1.95. Extra credit: Which x_0's give convergence?

44 Apply the secant method to solve $\cos x=0$ from $x_0=.308$.

45 Try Newton's method on $\cos x=0$ from $x_0=.308$. If $\cot x_0$ is exactly π, show that $x_1=x_0+\pi$ (and $x_2=x_1+\pi$). From $x_0=.308169071$ does Newton's method ever stop?

46 Use the Newton and secant programs to solve $x^3-10x^2+22x+6=0$ from $x_0=2$ and 1.39.

47 Newton's method for $\sin x=0$ is $x_{n+1}=x_n-\tan x_n$. Graph $\sin x$ and three iterations from $x_0=2$ and $x_0=1.8$. Predict the result for $x_0=1.9$ and test. This leads to the *computer project* in Problem 3.6.41, which finds fractals.

48 Graph $Y_1(x)=3.4(x-x^2)$ and $Y_2(x)=Y_1(Y_1(x))$ in the square window $(0,0)\leqslant(x,y)\leqslant(1,1)$. Then graph $Y_3(x)=Y_2(Y_1(x))$ and $Y_4,\ldots,Y_9$. The cycle is from .842 to .452.

49 Repeat Problem 48 with 3.4 changed to 2 or 3.5 or 4.

3.8 The Mean Value Theorem and l'Hôpital's Rule

Now comes one of the cornerstones of calculus: the *Mean Value Theorem.* It connects the local picture (slope at a point) to the global picture (average slope across an interval). In other words it relates df/dx to $\Delta f/\Delta x$. Calculus depends on this connec-

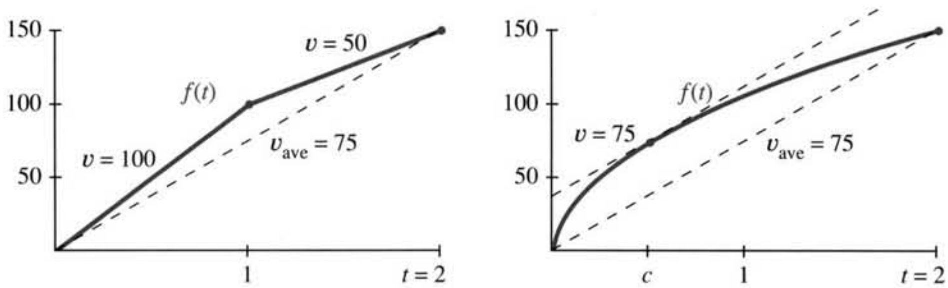


Fig. 3.26 (a) v jumps over v_{average}. (b) v equals v_{average}.

tion, which we saw first for velocities. If the average velocity is 75, is there a moment when the instantaneous velocity is 75?

Without more information, the answer to that question is *no*. The velocity could be 100 and then 50—averaging 75 but never equal to 75. If we allow a jump in velocity, it can jump right over its average. At that moment the velocity does not exist. (The distance function in Figure 3.26a has no derivative at $x = 1$.) We will take away this cheap escape by requiring a derivative at all points inside the interval.

In Figure 3.26b the distance increases by 150 when t increases by 2. There is a derivative df/dt at all interior points (but an infinite slope at $t = 0$). The average velocity is

$$\frac{\Delta f}{\Delta t} = \frac{f(2) - f(0)}{2 - 0} = \frac{150}{2} = 75.$$

The conclusion of the theorem is that $df/dt = 75$ ***at some point inside the interval.*** There is at least one point where $f'(c) = 75$.

This is not a constructive theorem. The value of c is not known. We don't find c, we just claim (with proof) that such a point exists.

3M ***Mean Value Theorem*** Suppose $f(x)$ is continuous in the closed interval $a \leqslant x \leqslant b$ and has a derivative everywhere in the open interval $a < x < b$. Then

$$\frac{f(b) - f(a)}{b - a} = f'(c) \text{ at some point } a < c < b. \tag{1}$$

The left side is the average slope $\Delta f/\Delta x$. It equals df/dx at c. The notation for a closed interval [with endpoints] is $[a, b]$. For an open interval (without endpoints) we write (a, b). Thus f' is defined in (a, b), and f remains continuous at a and b. A derivative is allowed at those endpoints too—but the theorem doesn't require it.

The proof is based on a special case—when $f(a) = 0$ and $f(b) = 0$. *Suppose the function starts at zero and returns to zero.* The average slope or velocity is zero. We have to prove that $f'(c) = 0$ at a point in between. This special case (keeping the assumptions on $f(x)$) is called *Rolle's theorem.*

Geometrically, if f goes away from zero and comes back, then $f' = 0$ *at the turn.*

3N ***Rolle's theorem*** Suppose $f(a) = f(b) = 0$ (zero at the ends). Then $f'(c) = 0$ at some point with $a < c < b$.

Proof At a point inside the interval where $f(x)$ reaches its maximum or minimum, df/dx must be zero. That is an acceptable point c. Figure 3.27a shows the difference between $f = 0$ (assumed at a and b) and $f' = 0$ (proved at c).

Small problem: The maximum could be reached at the ends a and b, if $f(x)<0$ in between. At those endpoints df/dx might not be zero. But in that case the ***minimum*** is reached at an interior point c, which is equally acceptable. The key to our proof is that ***a continuous function on*** $[a, b]$ ***reaches its maximum and minimum***. This is the *Extreme Value Theorem*.†

It is ironic that Rolle himself did not believe the logic behind calculus. He may not have believed his own theorem! Probably he didn't know what it meant—the language of "evanescent quantities" (Newton) and "infinitesimals" (Leibniz) was exciting but frustrating. Limits were close but never reached. Curves had infinitely many flat sides. Rolle didn't accept that reasoning, and what was really serious, he didn't accept the conclusions. The Académie des Sciences had to stop his battles (he fought against ordinary mathematicians, not Newton and Leibniz). So he went back to number theory, but his special case when $f(a)=f(b)=0$ leads directly to the big one.

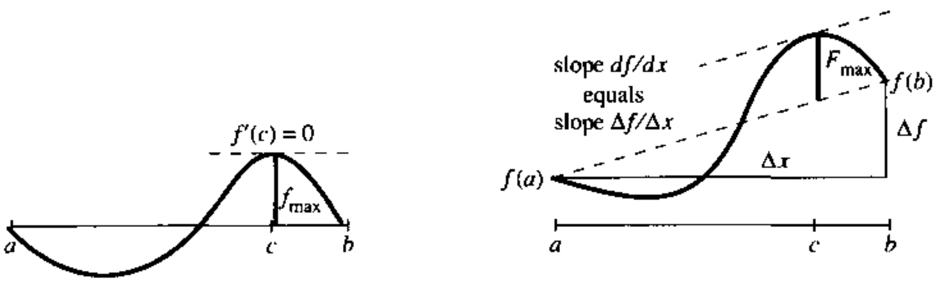


Fig. 3.27 Rolle's theorem is when $f(a)=f(b)=0$ in the Mean Value Theorem.

Proof of the Mean Value Theorem We are looking for a point where df/dx equals $\Delta f/\Delta x$. The idea is *to tilt the graph back to Rolle's special case* (when Δf was zero). In Figure 3.27b, the distance $F(x)$ between the curve and the dotted secant line comes from subtraction:

$$F(x)=f(x)-\left[f(a)+\frac{\Delta f}{\Delta x}(x-a)\right]. \tag{2}$$

At a and b, this distance is $F(a)=F(b)=0$. *Rolle's theorem applies to* $F(x)$. There is an interior point where $F'(c)=0$. At that point take the derivative of equation (2): $0=f'(c)-(\Delta f/\Delta x)$. The desired point c is found, proving the theorem.

EXAMPLE 1 The function $f(x)=\sqrt{x}$ goes from zero at $x=0$ to ten at $x=100$. Its average slope is $\Delta f/\Delta x=10/100$. The derivative $f'(x)=1/2\sqrt{x}$ exists in the open interval $(0, 100)$, even though it blows up at the end $x=0$. By the Mean Value Theorem there must be a point where $10/100=f'(c)=1/2\sqrt{c}$. That point is $c=25$.

The truth is that nobody cares about the exact value of c. Its existence is what matters. Notice how it affects the linear approximation $f(x)\approx f(a)+f'(a)(x-a)$, which was basic to this chapter. Close becomes exact ($\approx$ becomes $=$) when f' is computed at c instead of a:

†If $f(x)$ doesn't reach its maximum M, then $1/(M-f(x))$ would be continuous but also approach infinity. Essential fact: ***A continuous function on*** $[a, b]$ ***cannot approach infinity.***

3O The derivative at c gives an exact prediction of $f(x)$:

$$f(x) = f(a) + f'(c)(x - a). \tag{3}$$

The Mean Value Theorem is rewritten here as $\Delta f = f'(c)\Delta x$. Now $a < c < x$.

EXAMPLE 2 The function $f(x) = \sin x$ starts from $f(0) = 0$. The linear prediction (tangent line) uses the slope $\cos 0 = 1$. The exact prediction uses the slope $\cos c$ at an unknown point between 0 and x:

$$(approximate)\ \sin x \approx x \qquad (exact)\ \sin x = (\cos c)x. \tag{4}$$

The approximation is useful, because everything is computed at $x = a = 0$. The exact formula is interesting, because $\cos c \leqslant 1$ proves again that $\sin x \leqslant x$. The slope is below 1, so the sine graph stays below the 45° line.

EXAMPLE 3 ***If $f'(c) = 0$ at all points in an interval then $f(x)$ is constant.***

Proof When f' is everywhere zero, the theorem gives $\Delta f = 0$. Every pair of points has $f(b) = f(a)$. The graph is a horizontal line. That deceptively simple case is a key to the Fundamental Theorem of Calculus.

Most applications of $\Delta f = f'(c)\Delta x$ do not end up with a number. They end up with another theorem (like this one). The goal is to connect derivatives (local) to differences (global). But the next application—***l'Hôpital's Rule***—manages to produce a number out of 0/0.

L'HÔPITAL'S RULE

When $f(x)$ and $g(x)$ both approach zero, what happens to their ratio $f(x)/g(x)$?

$$\frac{f(x)}{g(x)} = \frac{x^2}{x} \quad \text{or} \quad \frac{\sin x}{x} \quad \text{or} \quad \frac{x - \sin x}{1 - \cos x} \quad \text{all become} \quad \frac{0}{0} \quad \text{at } x = 0.$$

Since 0/0 is meaningless, we cannot work separately with $f(x)$ and $g(x)$. This is a "*race toward zero*," in which two functions become small while their ratio might do anything. The problem is to find the limit of $f(x)/g(x)$.

One such limit is already studied. *It is the derivative!* $\Delta f/\Delta x$ automatically builds in a race toward zero, whose limit is df/dx:

$$\begin{matrix} f(x) - f(a) \to 0 \\ x - a \to 0 \end{matrix} \quad \text{but} \quad \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = f'(a). \tag{5}$$

The idea of l'Hôpital is to use f'/g' to handle f/g. The derivative is the special case $g(x) = x - a$, with $g' = 1$. The Rule is followed by examples and proofs.

3P ***l'Hôpital's Rule*** Suppose $f(x)$ and $g(x)$ both approach zero as $x \to a$. Then $f(x)/g(x)$ approaches the same limit as $f'(x)/g'(x)$, if that second limit exists:

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}. \qquad \text{Normally this limit is } \frac{f'(a)}{g'(a)}. \tag{6}$$

This is not the quotient rule! The derivatives of $f(x)$ and $g(x)$ are taken separately. Geometrically, l'Hôpital is saying that *when functions go to zero their slopes control their size*. An easy case is $f = 6(x - a)$ and $g = 2(x - a)$. The ratio f/g is exactly 6/2,

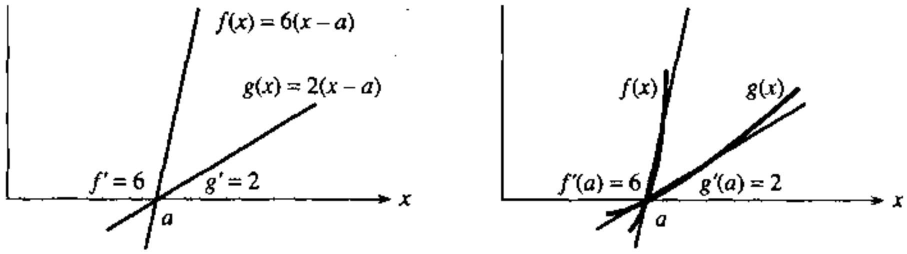


Fig. 3.28 (a) $\frac{f(x)}{g(x)}$ is exactly $\frac{f'(a)}{g'(a)}=3$. (b) $\frac{f(x)}{g(x)}$ approaches $\frac{f'(a)}{g'(a)}=3$.

the ratio of their slopes. Figure 3.28 shows these straight lines dropping to zero, controlled by 6 and 2.

The next figure shows the same limit 6/2, when the curves are ***tangent*** to the lines. That picture is the key to l'Hôpital's rule.

Generally the limit of f/g can be a finite number L or $+\infty$ or $-\infty$. (Also the limit point $x=a$ can represent a finite number or $+\infty$ or $-\infty$. We keep it finite.) The one absolute requirement is that $f(x)$ and $g(x)$ must separately approach zero—we insist on 0/0. Otherwise there is no reason why equation (6) should be true. With $f(x)=x$ and $g(x)=x-1$, ***don't*** use l'Hôpital:

$$\frac{f(x)}{g(x)}\to\frac{a}{a-1} \quad \text{but} \quad \frac{f'(x)}{g'(x)}=\frac{1}{1}.$$

Ordinary ratios approach $\lim f(x)$ divided by $\lim g(x)$. l'Hôpital enters only for 0/0.

EXAMPLE 4 (an old friend) $\lim_{x\to 0}\frac{1-\cos x}{x}$ equals $\lim_{x\to 0}\frac{\sin x}{1}$. This equals zero.

EXAMPLE 5 $\frac{f}{g}=\frac{\tan x}{\sin x}$ leads to $\frac{f'}{g'}=\frac{\sec^2 x}{\cos x}$. At $x=0$ the limit is $\frac{1}{1}$.

EXAMPLE 6 $\frac{f}{g}=\frac{x-\sin x}{1-\cos x}$ leads to $\frac{f'}{g'}=\frac{1-\cos x}{\sin x}$. At $x=0$ this is still $\frac{0}{0}$.

Solution ***Apply the Rule to*** f'/g'. It has the same limit as f''/g'':

$$\text{if } \frac{f}{g}\to\frac{0}{0} \text{ and } \frac{f'}{g'}\to\frac{0}{0} \text{ then compute } \frac{f''(x)}{g''(x)}=\frac{\sin x}{\cos x}\to\frac{0}{1}=0.$$

The reason behind l'Hôpital's Rule is that the following fractions are the same:

$$\frac{f(x)}{g(x)}=\frac{f(x)-f(a)}{x-a}\Big/\frac{g(x)-g(a)}{x-a}. \tag{7}$$

That is just algebra; the limit hasn't happened yet. The factors $x-a$ cancel, and the numbers $f(a)$ and $g(a)$ are zero by assumption. Now take the limit on the right side of (7) as x approaches a.

What normally happens is that one part approaches f' at $x=a$. The other part approaches $g'(a)$. We hope $g'(a)$ is not zero. In this case we can divide one limit by the other limit. That gives the "normal" answer

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\text{limit of }(7)=\frac{f'(a)}{g'(a)}. \tag{8}$$

This is also l'Hôpital's answer. When $f'(x)\to f'(a)$ and separately $g'(x)\to g'(a)$, his overall limit is $f'(a)/g'(a)$. He published this rule in the first textbook ever written on differential calculus. (That was in 1696—the limit was actually discovered by his teacher Bernoulli.) Three hundred years later we apply his name to other cases permitted in (6), when f'/g' might approach a limit even if the separate parts do not.

To prove this more general form of l'Hôpital's Rule, we need a more general Mean Value Theorem. *I regard the discussion below as optional in a calculus course* (but required in a calculus book). The important idea already came in equation (8).

Remark** **The basic "indeterminate" is $\infty-\infty$. If $f(x)$ and $g(x)$ approach infinity, anything is possible for $f(x)-g(x)$. We could have x^2-x or $x-x^2$ or $(x+2)-x$. Their limits are ∞ and $-\infty$ and 2.

At the next level are $0/0$ and ∞/∞ and $0\cdot\infty$. To find the limit in these cases, try l'Hôpital's Rule. See Problem 24 when $f(x)/g(x)$ approaches ∞/∞. When $f(x)\to 0$ and $g(x)\to\infty$, apply the $0/0$ rule to $f(x)/(1/g(x))$.

The next level has 0^0 and 1^∞ and ∞^0. Those come from limits of $f(x)^{g(x)}$. If $f(x)$ approaches 0, 1, or ∞ while $g(x)$ approaches 0, ∞, or 0, we need more information. A really curious example is $x^{1/\ln x}$, which shows all three possibilities 0^0 and 1^∞ and ∞^0. This function is actually a constant! It equals e.

To go back down a level, take logarithms. Then $g(x)\ln f(x)$ returns to $0/0$ and $0\cdot\infty$ and l'Hôpital's Rule. But logarithms and e have to wait for Chapter 6.

THE GENERALIZED MEAN VALUE THEOREM

The MVT can be extended to *two functions*. The extension is due to Cauchy, who cleared up the whole idea of limits. You will recognize the special case $g=x$ as the ordinary Mean Value Theorem.

> **3Q** ***Generalized* MVT** If $f(x)$ and $g(x)$ are continuous on $[a,b]$ and differentiable on (a,b), there is a point $a<c<b$ where
>
> $$[f(b)-f(a)]g'(c)=[g(b)-g(a)]f'(c). \tag{9}$$

The proof comes by constructing a new function that has $F(a)=F(b)$:

$$F(x)=[f(b)-f(a)]g(x)-[g(b)-g(a)]f(x).$$

The ordinary Mean Value Theorem leads to $F'(c)=0$—which is equation (9).

Application 1 (Proof of l'Hôpital's Rule) The rule deals with $f(a)/g(a)=0/0$. Inserting those zeros into equation (9) leaves $f(b)g'(c)=g(b)f'(c)$. Therefore

$$\frac{f(b)}{g(b)}=\frac{f'(c)}{g'(c)}. \tag{10}$$

As b approaches a, so does c. The point c is squeezed between a and b. The limit of equation (10) as $b\to a$ and $c\to a$ is l'Hôpital's Rule.

Application 2 (Error in linear approximation) Section 3.2 stated that the distance between a curve and its tangent line grows like $(x-a)^2$. Now we can prove this, and find out more. Linear approximation is

$$f(x)=f(a)+f'(a)(x-a)+\textit{error } e(x). \tag{11}$$

The pattern suggests an error involving $f''(x)$ and $(x-a)^2$. The key example $f=x^2$ shows the need for a factor $\frac{1}{2}$ (to cancel $f''=2$). ***The error in linear approximation is***

$$e(x)=\tfrac{1}{2}f''(c)(x-a)^2 \quad \text{with} \quad a<c<x. \tag{12}$$

Key idea Compare the error $e(x)$ to $(x-a)^2$. Both are zero at $x=a$:

$$e=f(x)-f(a)-f'(a)(x-a) \qquad e'=f'(x)-f'(a) \qquad e''=f''(x)$$
$$g=(x-a)^2 \qquad g'=2(x-a) \qquad g''=2$$

The Generalized Mean Value Theorem finds a point C between a and x where $e(x)/g(x)=e'(C)/g'(C)$. This is equation (10) with different letters. After checking $e'(a)=g'(a)=0$, apply the same theorem to $e'(x)$ and $g'(x)$. It produces a point c between a and C—certainly between a and x—where

$$\frac{e'(C)}{g'(C)}=\frac{e''(c)}{g''(c)} \quad \text{and therefore} \quad \frac{e(x)}{g(x)}=\frac{e''(c)}{g''(c)}.$$

With $g=(x-a)^2$ and $g''=2$ and $e''=f''$, the equation on the right is $e(x)=\frac{1}{2}f''(c)(x-a)^2$. The error formula is proved. A very good approximation is $\frac{1}{2}f''(a)(x-a)^2$.

EXAMPLE 7 $f(x)=\sqrt{x}$ near $a=100$: $\sqrt{102}\approx 10+\left(\frac{1}{20}\right)2+\frac{1}{2}\left(\frac{-1}{4000}\right)2^2.$

That last term predicts $e=-.0005$. The actual error is $\sqrt{102}-10.1=-.000496$.

3.8 EXERCISES

Read-through questions

The Mean Value Theorem equates the average slope $\Delta f/\Delta x$ over an __a__ $[a,b]$ to the slope df/dx at an unknown __b__. The statement is __c__. It requires $f(x)$ to be __d__ on the __e__ interval $[a,b]$, with a __f__ on the open interval (a,b). Rolle's theorem is the special case when $f(a)=f(b)=0$, and the point c satisfies __g__. The proof chooses c as the point where f reaches its __h__.

Consequences of the Mean Value Theorem include: If $f'(x)=0$ everywhere in an interval then $f(x)=$ __i__. The prediction $f(x)=f(a)+$ __j__ $(x-a)$ is exact for some c between a and x. The quadratic prediction $f(x)=f(a)+f'(a)(x-a)+$ __k__ $(x-a)^2$ is exact for another c. The error in $f(a)+f'(a)(x-a)$ is less than $\frac{1}{2}M(x-a)^2$ where M is the maximum of __l__.

A chief consequence is l'Hôpital's Rule, which applies when $f(x)$ and $g(x)\to$ __m__ as $x\to a$. In that case the limit of $f(x)/g(x)$ equals the limit of __n__, provided this limit exists. Normally this limit is $f'(a)/g'(a)$. If this is also 0/0, go on to the limit of __o__.

Find all points $0<c<2$ where $f(2)-f(0)=f'(c)(2-0)$.

1 $f(x)=x^3$

2 $f(x)=\sin \pi x$

3 $f(x)=\tan 2\pi x$

4 $f(x)=1+x+x^2$

5 $f(x)=(x-1)^{10}$

6 $f(x)=(x-1)^9$

In 7–10 show that no point c yields $f(1)-f(-1)=f'(c)(2)$. Explain why the Mean Value Theorem fails to apply.

7 $f(x)=|x-\frac{1}{2}|$

8 $f(x)=$ unit step function

9 $f(x)=|x|^{1/2}$

10 $f(x)=1/x^2$

11 Show that $\sec^2 x$ and $\tan^2 x$ have the same derivative, and draw a conclusion about $f(x)=\sec^2 x-\tan^2 x$.

12 Show that $\csc^2 x$ and $\cot^2 x$ have the same derivative and find $f(x)=\csc^2 x-\cot^2 x$.

Evaluate the limits in 13–22 by l'Hôpital's Rule.

13 $\lim_{x\to 3}\frac{x^2-9}{x-3}$

14 $\lim_{x\to 3}\frac{x^2-9}{x+3}$

15 $\lim_{x\to 0} \frac{(1+x)^{-2}-1}{x}$ **16** $\lim_{x\to 0} \frac{\sqrt{1-\cos x}}{x}$

17 $\lim_{x\to \pi} \frac{x-\pi}{\sin x}$ **18** $\lim_{x\to 1} \frac{x-1}{\sin x}$

19 $\lim_{x\to 0} \frac{(1+x)^n-1}{x}$ **20** $\lim_{x\to 0} \frac{(1+x)^n-1-nx}{x^2}$

21 $\lim_{x\to 0} \frac{\sin x-\tan x}{x^3}$ **22** $\lim_{x\to 0} \frac{\sqrt{1+x}-\sqrt{1-x}}{x}$

23 For $f=x^2-4$ and $g=x+2$, the ratio f'/g' approaches 4 as $x\to 2$. What is the limit of $f(x)/g(x)$? What goes wrong in l'Hôpital's Rule?

24 ***l'Hôpital's Rule still holds for*** $f(x)/g(x)\to\infty/\infty$: L ***is***

$$\lim \frac{f(x)}{g(x)} = \lim \frac{1/g(x)}{1/f(x)} = \lim \frac{g'(x)/g^2(x)}{f'(x)/f^2(x)} = L^2 \lim \frac{g'(x)}{f'(x)}.$$

Then L equals $\lim\,[f'(x)/g'(x)]$ if this limit exists. Where did we use the rule for 0/0? What other limit rule was used?

25 Compute $\lim_{x\to 0} \frac{1+(1/x)}{1-(1/x)}$. **26** Compute $\lim_{x\to\infty} \frac{x^2+x}{2x^2}$.

27 Compute $\lim_{x\to\infty} \frac{x+\cos x}{x+\sin x}$ by common sense. Show that l'Hôpital gives no answer.

28 Compute $\lim_{x\to 0} \frac{\csc x}{\cot x}$ by common sense or trickery.

29 The Mean Value Theorem applied to $f(x)=x^3$ guarantees that some number c between 1 and 4 has a certain property. Say what the property is and find c.

30 If $|df/dx|\leqslant 1$ at all points, prove this fact:

$$|f(x)-f(y)|\leqslant|x-y| \text{ at all } x \text{ and } y.$$

31 The error in Newton's method is squared at each step: $|x_{n+1}-x^*|\leqslant M|x_n-x^*|^2$. The proof starts from $0=f(x^*)=f(x_n)+f'(x_n)(x^*-x_n)+\frac{1}{2}f''(c)(x^*-x_n)^2$. Divide by $f'(x_n)$, recognize x_{n+1}, and estimate M.

32 (Rolle's theorem backward) Suppose $f'(c)=0$. Are there necessarily two points around c where $f(a)=f(b)$?

33 Suppose $f(0)=0$. If $f(x)/x$ has a limit as $x\to 0$, that limit is better known to us as ________. L'Hôpital's Rule looks instead at the limit of ________.

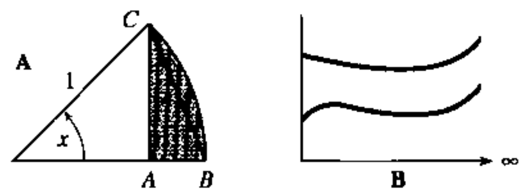
Conclusion from l'Hôpital: The limit of $f'(x)$, if it exists, agrees with $f'(0)$. Thus $f'(x)$ cannot have a "removable ________."

34 It is possible that $f'(x)/g'(x)$ has *no* limit but $f(x)/g(x)\to L$. This is why l'Hôpital included an "if."

(a) Find L as $x\to 0$ when $f(x)=x^2\cos(1/x)$ and $g(x)=x$. Remember that cosines are below 1.

(b) From the formula $f'(x)=\sin(1/x)+2x\cos(1/x)$ show that f'/g' has no limit as $x\to 0$.

35 Stein's calculus book asks for the limiting ratio of $f(x)=$ triangular area ABC to $g(x)=$ curved area ABC. (a) Guess the limit of f/g as the angle x goes to zero. (b) Explain why $f(x)$ is $\frac{1}{2}(\sin x-\sin x\cos x)$ and $g(x)$ is $\frac{1}{2}(x-\sin x\cos x)$. (c) Compute the true limit of $f(x)/g(x)$.



36 If you drive 3000 miles from New York to L.A. in 100 hours (sleeping and eating and going backwards are allowed) then at some moment your speed is ________.

37 As $x\to\infty$ l'Hôpital's Rule still applies. The limit of $f(x)/g(x)$ equals the limit of $f'(x)/g'(x)$, if that limit exists. What is the limit as the graphs become parallel in Figure B?

38 Prove that $f(x)$ is increasing when its slope is positive: *If* $f'(c)>0$ *at all points* c, *then* $f(b)>f(a)$ *at all pairs of points* $b>a$.