4.7 Newton's Method

HISTORICAL BIOGRAPHY

Niels Henrik Abel
(1802–1829)

One of the basic problems of mathematics is solving equations. Using the quadratic root formula, we know how to find a point (solution) where $x^2 - 3x + 2 = 0$. There are more complicated formulas to solve cubic or quartic equations (polynomials of degree 3 or 4), but the Norwegian mathematician Niels Abel showed that no simple formulas exist to solve polynomials of degree equal to five. There is also no simple formula for solving equations like $\sin x = x^2$, which involve transcendental functions as well as polynomials or other algebraic functions.

In this section we study a numerical method, called *Newton's method* or the *Newton–Raphson method*, which is a technique to approximate the solution to an equation $f(x) = 0$. Essentially it uses tangent lines in place of the graph of $y = f(x)$ near the points where f is zero. (A value of x where f is zero is a *root* of the function f and a *solution* of the equation $f(x) = 0$.)

Procedure for Newton's Method

The goal of Newton's method for estimating a solution of an equation $f(x) = 0$ is to produce a sequence of approximations that approach the solution. We pick the first number x_0 of the sequence. Then, under favorable circumstances, the method does the rest by moving step by step toward a point where the graph of f crosses the x-axis (Figure 4.43). At each

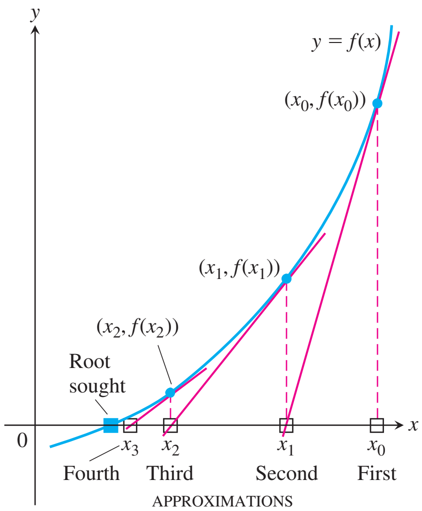

FIGURE 4.43 Newton's method starts with an initial guess x_0 and (under favorable circumstances) improves the guess one step at a time.

step the method approximates a zero of f with a zero of one of its linearizations. Here is how it works.

The initial estimate, x_0, may be found by graphing or just plain guessing. The method then uses the tangent to the curve $y = f(x)$ at $(x_0, f(x_0))$ to approximate the curve, calling the point x_1 where the tangent meets the x-axis (Figure 4.43). The number x_1 is usually a better approximation to the solution than is x_0. The point x_2 where the tangent to the curve at $(x_1, f(x_1))$ crosses the x-axis is the next approximation in the sequence. We continue on, using each approximation to generate the next, until we are close enough to the root to stop.

We can derive a formula for generating the successive approximations in the following way. Given the approximation x_n, the point-slope equation for the tangent to the curve at $(x_n, f(x_n))$ is

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

We can find where it crosses the x-axis by setting $y = 0$ (Figure 4.44).

$$0 = f(x_n) + f'(x_n)(x - x_n)$$

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

$$x = x_n - \frac{f(x_n)}{f'(x_n)} \qquad \text{If } f'(x_n) \neq 0$$

This value of x is the next approximation x_{n+1}. Here is a summary of Newton's method.

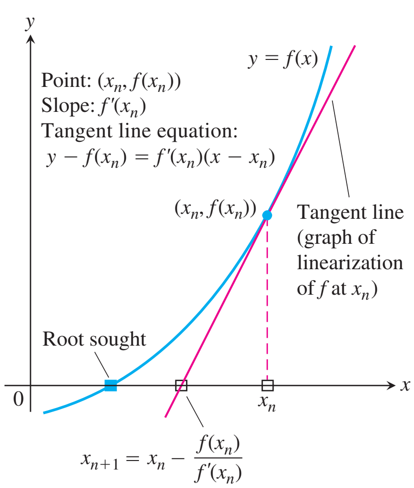

FIGURE 4.44 The geometry of the successive steps of Newton's method. From x_n we go up to the curve and follow the tangent line down to find x_{n+1}.

Procedure for Newton's Method

1. Guess a first approximation to a solution of the equation $f(x) = 0$. A graph of $y = f(x)$ may help.
2. Use the first approximation to get a second, the second to get a third, and so on, using the formula

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad \text{if } f'(x_n) \neq 0 \tag{1}$$

Applying Newton's Method

Applications of Newton's method generally involve many numerical computations, making them well suited for computers or calculators. Nevertheless, even when the calculations are done by hand (which may be very tedious), they give a powerful way to find solutions of equations.

In our first example, we find decimal approximations to $\sqrt{2}$ by estimating the positive root of the equation $f(x) = x^2 - 2 = 0$.

Video You Try It

EXAMPLE 1 Finding the Square Root of 2

Find the positive root of the equation

$$f(x) = x^2 - 2 = 0.$$

Solution With $f(x) = x^2 - 2$ and $f'(x) = 2x$, Equation (1) becomes

$$\begin{aligned} x_{n+1} &= x_n - \frac{x_n{}^2 - 2}{2x_n} \\ &= x_n - \frac{x_n}{2} + \frac{1}{x_n} \\ &= \frac{x_n}{2} + \frac{1}{x_n}. \end{aligned}$$

The equation

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

enables us to go from each approximation to the next with just a few keystrokes. With the starting value $x_0 = 1$, we get the results in the first column of the following table. (To five decimal places, $\sqrt{2} = 1.41421$.)

	Error	Number of correct digits
$x_0 = 1$	-0.41421	1
$x_1 = 1.5$	0.08579	1
$x_2 = 1.41667$	0.00246	3
$x_3 = 1.41422$	0.00001	5

■

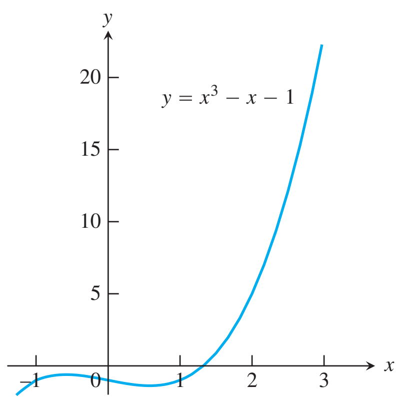

FIGURE 4.45 The graph of $f(x) = x^3 - x - 1$ crosses the x-axis once; this is the root we want to find (Example 2).

Newton's method is the method used by most calculators to calculate roots because it converges so fast (more about this later). If the arithmetic in the table in Example 1 had been carried to 13 decimal places instead of 5, then going one step further would have given $\sqrt{2}$ correctly to more than 10 decimal places.

EXAMPLE 2 Using Newton's Method

Find the x-coordinate of the point where the curve $y = x^3 - x$ crosses the horizontal line $y = 1$.

Solution The curve crosses the line when $x^3 - x = 1$ or $x^3 - x - 1 = 0$. When does $f(x) = x^3 - x - 1$ equal zero? Since $f(1) = -1$ and $f(2) = 5$, we know by the Intermediate Value Theorem there is a root in the interval $(1, 2)$ (Figure 4.45).

We apply Newton's method to f with the starting value $x_0 = 1$. The results are displayed in Table 4.1 and Figure 4.46.

At $n = 5$, we come to the result $x_6 = x_5 = 1.3247\ 17957$. When $x_{n+1} = x_n$, Equation (1) shows that $f(x_n) = 0$. We have found a solution of $f(x) = 0$ to nine decimals. ■

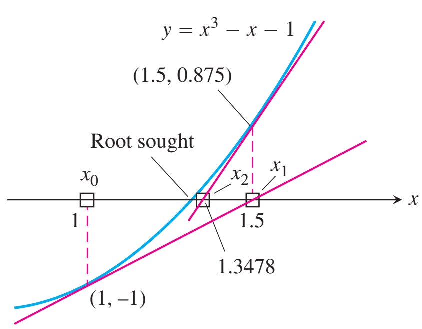

FIGURE 4.46 The first three x-values in Table 4.1 (four decimal places).

In Figure 4.47 we have indicated that the process in Example 2 might have started at the point $B_0(3, 23)$ on the curve, with $x_0 = 3$. Point B_0 is quite far from the x-axis, but the tangent at B_0 crosses the x-axis at about $(2.12, 0)$, so x_1 is still an improvement over x_0. If we use Equation (1) repeatedly as before, with $f(x) = x^3 - x - 1$ and $f'(x) = 3x^2 - 1$, we confirm the nine-place solution $x_7 = x_6 = 1.3247\ 17957$ in seven steps.

TABLE 4.1 The result of applying Newton's method to $f(x) = x^3 - x - 1$ with $x_0 = 1$

n	x_n	$f(x_n)$	$f'(x_n)$	$x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$
0	1	−1	2	1.5
1	1.5	0.875	5.75	1.3478 26087
2	1.3478 26087	0.1006 82173	4.4499 05482	1.3252 00399
3	1.3252 00399	0.0020 58362	4.2684 68292	1.3247 18174
4	1.3247 18174	0.0000 00924	4.2646 34722	1.3247 17957
5	1.3247 17957	−1.8672E-13	4.2646 32999	1.3247 17957

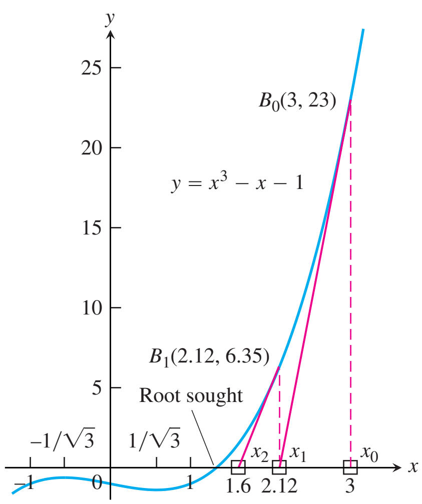

FIGURE 4.47 Any starting value x_0 to the right of $x = 1/\sqrt{3}$ will lead to the root.

The curve in Figure 4.47 has a local maximum at $x = -1/\sqrt{3}$ and a local minimum at $x = 1/\sqrt{3}$. We would not expect good results from Newton's method if we were to start with x_0 between these points, but we can start any place to the right of $x = 1/\sqrt{3}$ and get the answer. It would not be very clever to do so, but we could even begin far to the right of B_0, for example with $x_0 = 10$. It takes a bit longer, but the process still converges to the same answer as before.

Convergence of Newton's Method

In practice, Newton's method usually converges with impressive speed, but this is not guaranteed. One way to test convergence is to begin by graphing the function to estimate a good starting value for x_0. You can test that you are getting closer to a zero of the function by evaluating $|f(x_n)|$ and check that the method is converging by evaluating $|x_n - x_{n+1}|$.

Theory does provide some help. A theorem from advanced calculus says that if

$$\left|\frac{f(x)f''(x)}{(f'(x))^2}\right| < 1 \tag{2}$$

for all x in an interval about a root r, then the method will converge to r for any starting value x_0 in that interval. Note that this condition is satisfied if the graph of f is not too horizontal near where it crosses the x-axis.

Newton's method always converges if, between r and x_0, the graph of f is concave up when $f(x_0) > 0$ and concave down when $f(x_0) < 0$. (See Figure 4.48.) In most cases, the speed of the convergence to the root r is expressed by the advanced calculus formula

$$\underbrace{|x_{n+1} - r|}_{\text{error } e_{n+1}} \le \frac{\max|f''|}{2\min|f'|}|x_n - r|^2 = \text{constant} \cdot \underbrace{|x_n - r|}_{\text{error } e_n}{}^2, \tag{3}$$

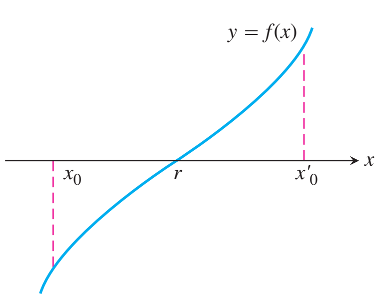

FIGURE 4.48 Newton's method will converge to r from either starting point.

where max and min refer to the maximum and minimum values in an interval surrounding r. The formula says that the error in step $n + 1$ is no greater than a constant times the square of the error in step n. This may not seem like much, but think of what it says. If the constant is less than or equal to 1 and $|x_n - r| < 10^{-3}$, then $|x_{n+1} - r| < 10^{-6}$. *In a single step*, the method moves from three decimal places of accuracy to six, and the number of decimals of accuracy continues to double with each successive step.

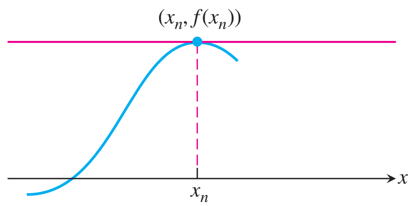

FIGURE 4.49 If $f'(x_n) = 0$, there is no intersection point to define x_{n+1}.

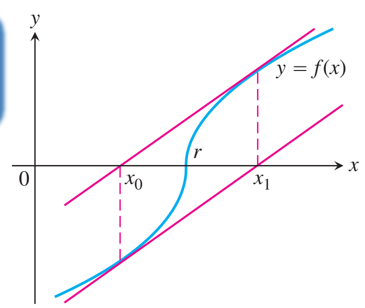

FIGURE 4.50 Newton's method fails to converge. You go from x_0 to x_1 and back to x_0, never getting any closer to r.

But Things Can Go Wrong

Newton's method stops if $f'(x_n) = 0$ (Figure 4.49). In that case, try a new starting point. Of course, f and f' may have the same root. To detect whether this is so, you could first find the solutions of $f'(x) = 0$ and check f at those values, or you could graph f and f' together.

Newton's method does not always converge. For instance, if

$$f(x) = \begin{cases} -\sqrt{r - x}, & x < r \\ \sqrt{x - r}, & x \geq r, \end{cases}$$

the graph will be like the one in Figure 4.50. If we begin with $x_0 = r - h$, we get $x_1 = r + h$, and successive approximations go back and forth between these two values. No amount of iteration brings us closer to the root than our first guess.

If Newton's method does converge, it converges to a root. Be careful, however. There are situations in which the method appears to converge but there is no root there. Fortunately, such situations are rare.

When Newton's method converges to a root, it may not be the root you have in mind. Figure 4.51 shows two ways this can happen.

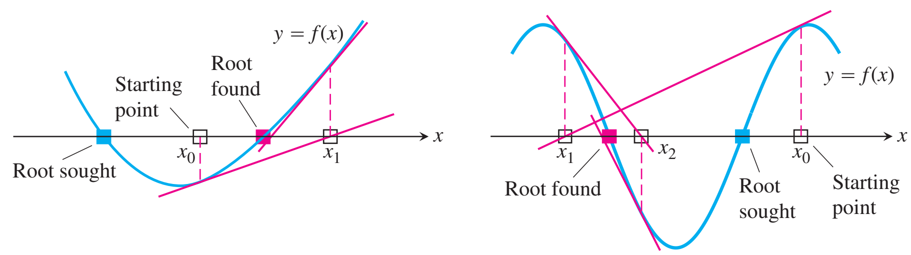

FIGURE 4.51 If you start too far away, Newton's method may miss the root you want.

Fractal Basins and Newton's Method

The process of finding roots by Newton's method can be uncertain in the sense that for some equations, the final outcome can be extremely sensitive to the starting value's location.

The equation $4x^4 - 4x^2 = 0$ is a case in point (Figure 4.52a). Starting values in the blue zone on the x-axis lead to root A. Starting values in the black lead to root B, and starting values in the red zone lead to root C. The points $\pm\sqrt{2}/2$ give horizontal tangents. The points $\pm\sqrt{21}/7$ "cycle," each leading to the other, and back (Figure 4.52b).

The interval between $\sqrt{21}/7$ and $\sqrt{2}/2$ contains infinitely many open intervals of points leading to root A, alternating with intervals of points leading to root C (Figure 4.52c). The boundary points separating consecutive intervals (there are infinitely many) do not lead to roots, but cycle back and forth from one to another. Moreover, as we select points that approach $\sqrt{21}/7$ from the right, it becomes increasingly difficult to distinguish which lead to root A and which to root C. On the same side of $\sqrt{21}/7$, we find arbitrarily close together points whose ultimate destinations are far apart.

If we think of the roots as "attractors" of other points, the coloring in Figure 4.52 shows the intervals of the points they attract (the "intervals of attraction"). You might think that points between roots A and B would be attracted to either A or B, but, as we see, that is not the case. Between A and B there are infinitely many intervals of points attracted to C. Similarly between B and C lie infinitely many intervals of points attracted to A.

We encounter an even more dramatic example of such behavior when we apply Newton's method to solve the complex-number equation $z^6 - 1 = 0$. It has six solutions: 1, -1, and the four numbers $\pm(1/2) \pm \left(\sqrt{3}/2\right)i$. As Figure 4.53 suggests, each of the

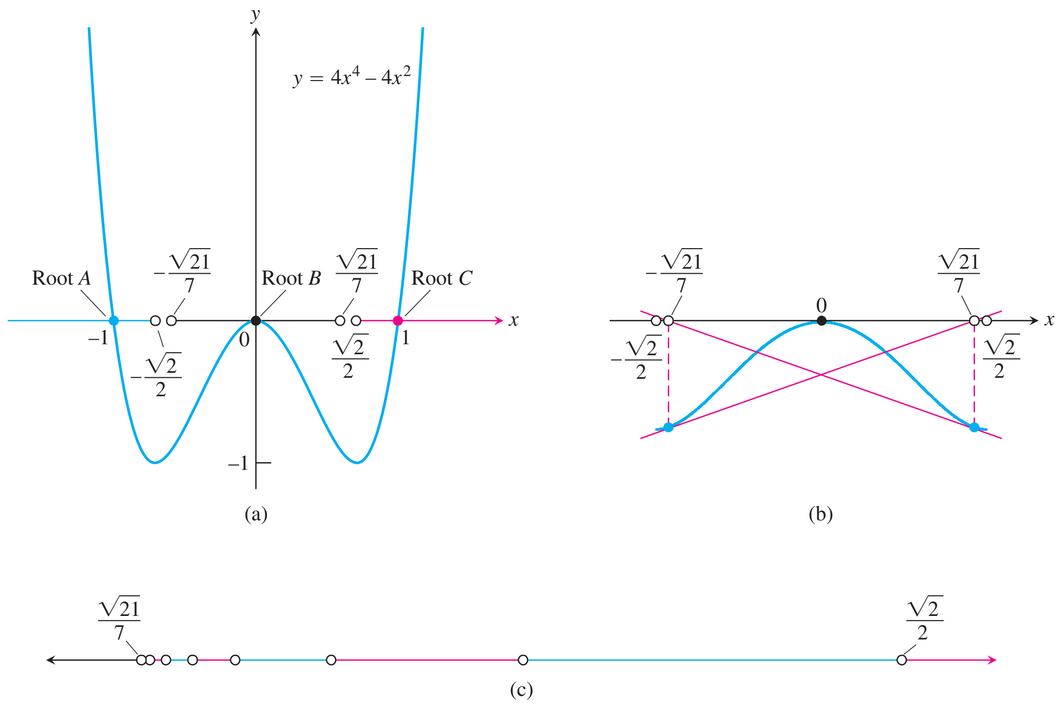

FIGURE 4.52 (a) Starting values in $\left(-\infty, -\sqrt{2}/2\right)$, $\left(-\sqrt{21}/7, \sqrt{21}/7\right)$, and $\left(\sqrt{2}/2, \infty\right)$ lead respectively to roots A, B, and C. (b) The values $x = \pm\sqrt{21}/7$ lead only to each other. (c) Between $\sqrt{21}/7$ and $\sqrt{2}/2$, there are infinitely many open intervals of points attracted to A alternating with open intervals of points attracted to C. This behavior is mirrored in the interval $\left(-\sqrt{2}/2, -\sqrt{21}/7\right)$.

FIGURE 4.53 This computer-generated initial value portrait uses color to show where different points in the complex plane end up when they are used as starting values in applying Newton's method to solve the equation $z^6 - 1 = 0$. Red points go to 1, green points to $(1/2) + \left(\sqrt{3}/2\right)i$, dark blue points to $(-1/2) + \left(\sqrt{3}/2\right)i$, and so on. Starting values that generate sequences that do not arrive within 0.1 unit of a root after 32 steps are colored black.

six roots has infinitely many "basins" of attraction in the complex plane (Appendix 5). Starting points in red basins are attracted to the root 1, those in the green basin to the root $(1/2) + \left(\sqrt{3}/2\right)i$, and so on. Each basin has a boundary whose complicated pattern repeats without end under successive magnifications. These basins are called **fractal basins**.