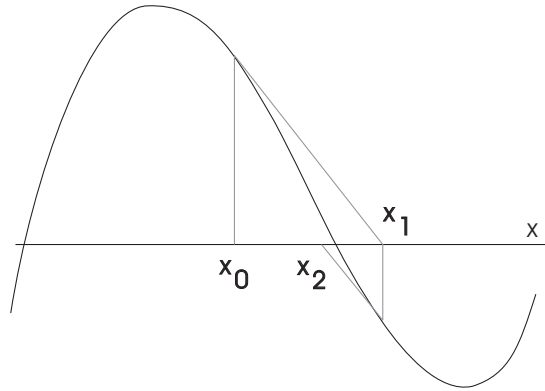
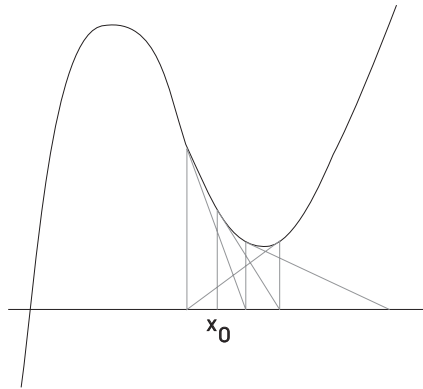


Newton's Method

Newton's method is simple to describe pictorially. To find a root of an equation $f(x) = 0$, start at a point x_0 . Go up to the curve. From there, "slide down" the tangent line till you hit the x -axis. That's x_1 . Now repeat the process. As you can see in the following picture, this can work quite well in "nice" cases.

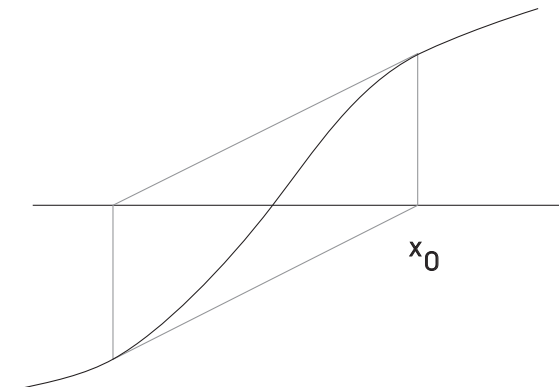


It is possible for things to go wrong. In the next picture, the process seems to be bouncing around aimlessly without getting closer to the root on the left.



In fact, if one of the bounces winds up right below the critical point, the horizontal tangent will shoot you off to infinity!

Here is another situation in which Newton's method fails. This time there's an oscillation of period 2.



When Newton's method *does* work, it works well: Roughly, the number of places of accuracy *doubles* with each step. However, things can also go wrong even more spectacularly than in the bad cases above. Later on, I'll discuss a spectacular result due to Barna (1956, 1961) which illustrates this.

The formula for Newton's method is easy. If your current guess is x_{old} , you want to know where the tangent at x_{old} hits the x -axis. The tangent line has slope $f'(x_{\text{old}})$, and passes through the point $(x_{\text{old}}, f(x_{\text{old}}))$. Its equation is

$$y - f(x_{\text{old}}) = f'(x_{\text{old}}) \cdot (x - x_{\text{old}}).$$

Set $y = 0$ to find the x -intercept. If I call the x -intercept x_{new} , then

$$x_{\text{new}} = x_{\text{old}} - \frac{f(x_{\text{old}})}{f'(x_{\text{old}})}.$$

This is the formula for Newton's method. Note that if you luck out and hit a root, $f(x_{\text{old}}) = 0$ and $x_{\text{new}} = x_{\text{old}}$ — the procedure stabilizes.

Example. Suppose $f(3) = 6$ and $f'(3) = 75$. What is the next value of x if Newton's method is applied to f starting at $x = 3$?

The next value of x is

$$3 - \frac{f(3)}{f'(3)} = 3 - \frac{6}{75} = 2.92. \quad \square$$

Example. Use Newton's method to approximate a solution to $x - \cos x = 0$.

A graph shows that $x = 1$ is a reasonable starting point. I have $f'(x) = 1 + \sin x$, so

$$x_{\text{new}} = x_{\text{old}} - \frac{x_{\text{old}} - \cos x_{\text{old}}}{1 + \sin x_{\text{old}}}.$$

I can iterate the Newton function:

$$(Nf)(x) = x - \frac{x - \cos x}{1 + \sin x}.$$

This might be appropriate if you're using a program.

I'll do things step-by-step so you can see the values of x , $f(x)$, and $f'(x)$:

x	$f(x)$	$f'(x)$
1	0.45970	1.84147
0.75036	0.01892	1.68190
0.73911	4.6×10^{-5}	1.67363
0.739085		

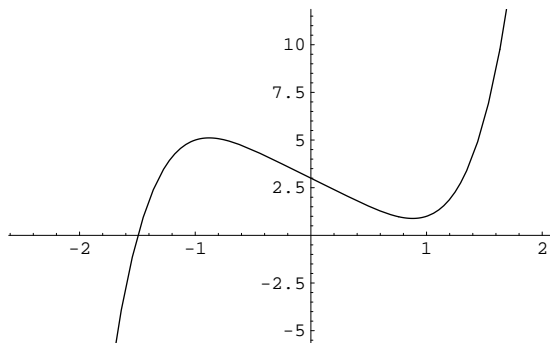
The last result is good to 6 places.

It turns out that for the equation $x = \cos x$, there is an easy way to find the solution iteratively on your calculator. Take the starting number — say $x = 1$ — and repeatedly take the cosine. On some calculators, you need to take the cosine of the *last* result explicitly; on others, you can just keep pushing the cosine button. The process will converge — more slowly than Newton's method — to the solution above. \square

Example. Solve $x^5 - 3x + 3 = 0$.

Note that there is no “general quintic formula”, so without using very complicated mathematical functions, I must approximate a root.

First, I graph $f(x) = x^5 - 3x + 3$.



There appears to be a single root around $x = -1.5$.

The derivative is $f'(x) = 5x^4 - 3$. I can iterate the Newton function

$$N(f)(x) = x - \frac{x^5 - 3x + 3}{5x^4 - 3}.$$

This might be appropriate if you’re using a program.

I’ll set up a table and compute the values step-by-step. Here is the table:

x	$f(x)$	$f'(x)$
-1.5	-0.09375	22.31250
-1.49580	-5.94160×10^{-4}	22.03008
-1.49577	$-.243437 \times 10^{-8}$	22.02827

$x \approx -1.49577$ is correct to 5 places.

Notice that Newton’s method converges very rapidly (when it works). As I noted earlier, the number of correct places will roughly double with each iteration. \square

Example. (A square root algorithm) Use Newton’s method to approximate $\sqrt{3}$.

The positive solution to $x^2 - 3 = 0$ is $\sqrt{3}$, so I’ll apply Newton’s method to $f(x) = x^2 - 3$. I have $f'(x) = 2x$, so

$$x_{\text{new}} = x_{\text{old}} - \frac{x_{\text{old}}^2 - 3}{2x_{\text{old}}} = \frac{1}{2} \left(x_{\text{old}} + \frac{3}{x_{\text{old}}} \right).$$

If you iterate this function with a good starting point, you approximate $\sqrt{3}$. For example, starting at $x = 1$, the first few iterates are

$$1, \quad 2, \quad 1.75, \quad 1.73214, \quad 1.73205.$$

The last one is already good to 5 places.

In general, for $c > 0$, \sqrt{c} can be approximated by iterating

$$x_{\text{new}} = \frac{1}{2} \left(x_{\text{old}} + \frac{c}{x_{\text{old}}} \right).$$

This algorithm is often used in square root routines for computers. \square

There is a spectacular result due to Barna ([1], [2]) which demonstrates how badly Newton's method can fail. I'll state a special case.

Take a polynomial $f(x)$ of degree > 3 . In fact, take one of degree 28, with 28 distinct roots. For example, you could use

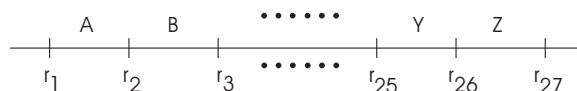
$$f(x) = (x - 1)(x - 2) \cdots (x - 28).$$

Choose f so that:

- (a) Between two roots of $f'(x)$ there is a root of $f(x)$.
- (b) Between two roots of $f''(x)$ there is a root of $f'(x)$.
- (c) All the roots of f' are inside the interval determined by the biggest and smallest roots of f .
- (d) All the roots of f'' are inside the interval determined by the biggest and smallest roots of f' .

These conditions simply say that f is a rather generic polynomial; they rule out things such as multiple roots for f' and f'' .

For $f(x) = (x - 1)(x - 2) \cdots (x - 28)$, you can see that f has 28 roots, and f' has 27. I've labelled the roots of f' as r_1, r_2, \dots . Here's a schematic picture of the intervals:



I've also labelled the intervals between the r 's with the letters A, B, C, \dots .

When you do Newton's method, if you hit a root r of f' , the tangent is horizontal and you get shot out to infinity. Otherwise, you bounce around between the intervals labelled with letters.

Barna's result says that if you take *any* string of letters, there is a starting point for Newton's method which "spells out" that string. That is, the successive iterates land in the intervals labelled by the letters for the string.

So if you took the string "STROMBOLI", the first iterate would be in the S interval, the second in the T interval, the third in the R interval, and so on.

You could take a string which spelled out an encyclopedia — or a history of your life! — and Barna's result says there's a starting point such that Newton's method would spell the string.

As you might expect, there's a catch. The catch is that while the result guarantees that such a starting point *exists*, it doesn't give any way of *finding* it. Maybe that's a good thing!

-
- [1] B. Barna, Über die divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln algebraischer Gleichungen II, *Publicationes Mathematicae, Debrecen*, 4(1956), 384–397.
 - [2] B. Barna, Über die divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln algebraischer Gleichungen III, *Publicationes Mathematicae, Debrecen*, 4(1956), 193–207.
 - [3] Donald G. Saari and John B. Urenko, Newton's method, circle maps, and chaotic motion, *American Mathematical Monthly*, 91(1)(1984), 3–17.