Lecture 3

Numerically Solving Equations

Virginia Tech Department of Computer Science
Fall 2026 • Sumeet Khatri

Four methods for locating a root

Bisection

Keep a sign-changing interval and halve its width.

Fixed-point iteration

Rewrite the equation as an iteration that maps a point to its next approximation.

Newton’s method

Use the local tangent line to propose the next point.

Secant method

Estimate the slope from the two most recent function values.

Equations as root-finding problems

A root is an input at which a function equals zero.

$$f(\alpha) = 0, \qquad \alpha \in D \subseteq \mathbb{R}$$

Move both sides of an equation to one side, while retaining its domain.

$$u(x) = v(x) \quad \Longleftrightarrow \quad f(x) = u(x) - v(x) = 0$$

A voltage determined by a nonlinear equation

$$I = a(e^{bV} - 1), \qquad c = dI + V$$

Eliminate the current to obtain one equation for the voltage.

$$F(V) = ad(e^{bV} - 1) + V - c = 0$$

$$F(V) = 14.3(e^{2V} - 1) + V - 12$$

The required voltage is approximately $V = 0.299$. A numerical method supplies the approximation.

Intermediate value theorem

Assume f is continuous on $[a, b]$. Every value between $f(a)$ and $f(b)$ is attained somewhere in that interval.

$$y \text{ between } f(a) \text{ and } f(b) \quad \Longrightarrow \quad f(c) = y \text{ for some } c \in [a, b]$$

The root-finding consequence

$$f(a)f(b) < 0 \quad \Longrightarrow \quad f(\alpha) = 0 \text{ for some } \alpha \in (a, b)$$

Continuity and opposite endpoint signs certify existence. They do not certify uniqueness.

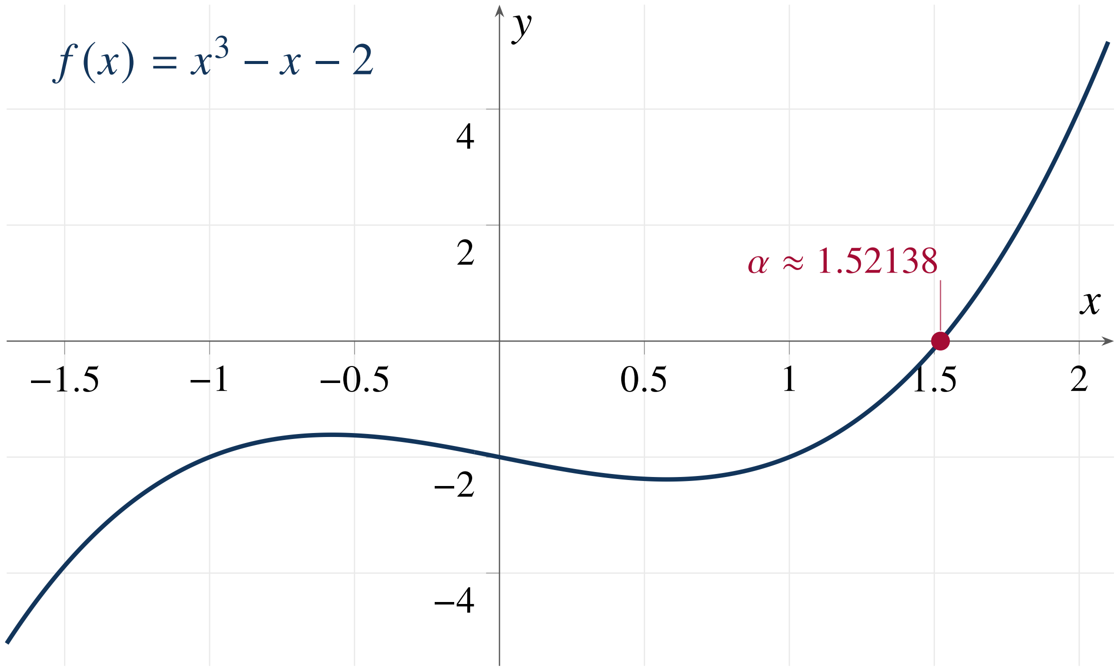
A cubic function

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

Existence of a root

$$f(1) = -2, \qquad f(2) = 4$$

A polynomial is continuous. Opposite signs give a root between the endpoints.



$$\alpha = 1.5213797068045675696040808\ldots$$

This true value is for checking the examples. The algorithms do not receive it, as we don't know the true value in practice!

Method 1: The bisection method

Begin with a continuous function and a interval $[a_0, b_0]$. The initial guess for the root is $c_0 = (a_0 + b_0)/2$. Then narrow down the root by progressively halving the interval.

$$c_n = \frac{a_n + b_n}{2}$$

Retain the left half

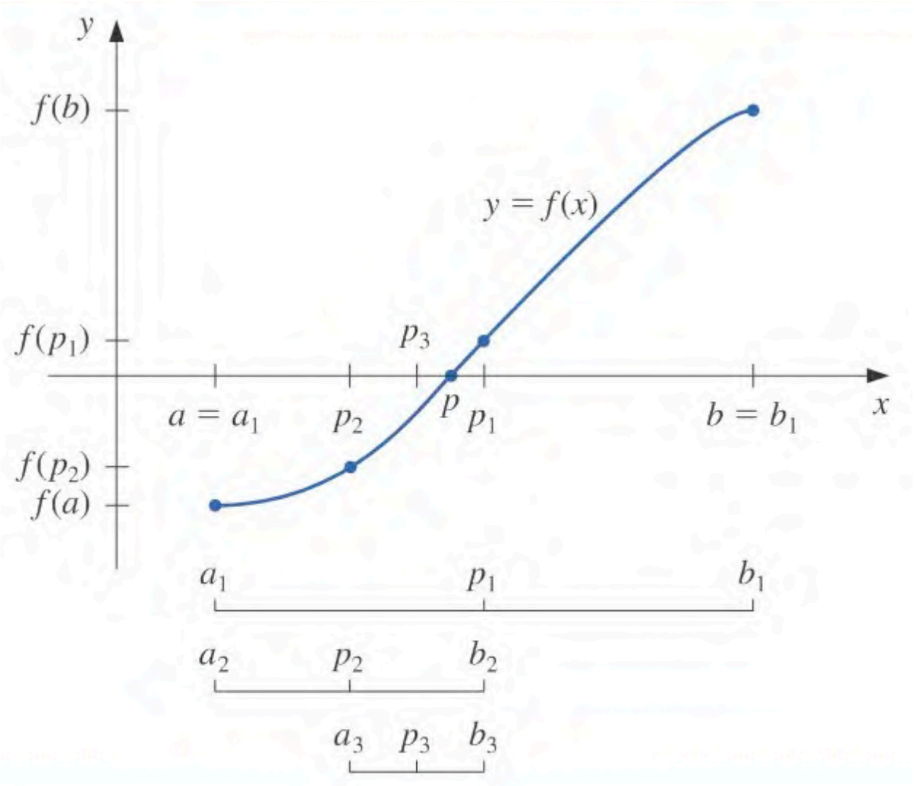
$$f(a_n)f(c_n) < 0$$

$$[a_{n+1}, b_{n+1}] = [a_n, c_n]$$

Retain the right half

$$f(a_n)f(c_n) > 0$$

$$[a_{n+1}, b_{n+1}] = [c_n, b_n]$$



If $f(c_n) = 0$, stop. Otherwise the retained interval has half the previous width.

The cubic example with bisection

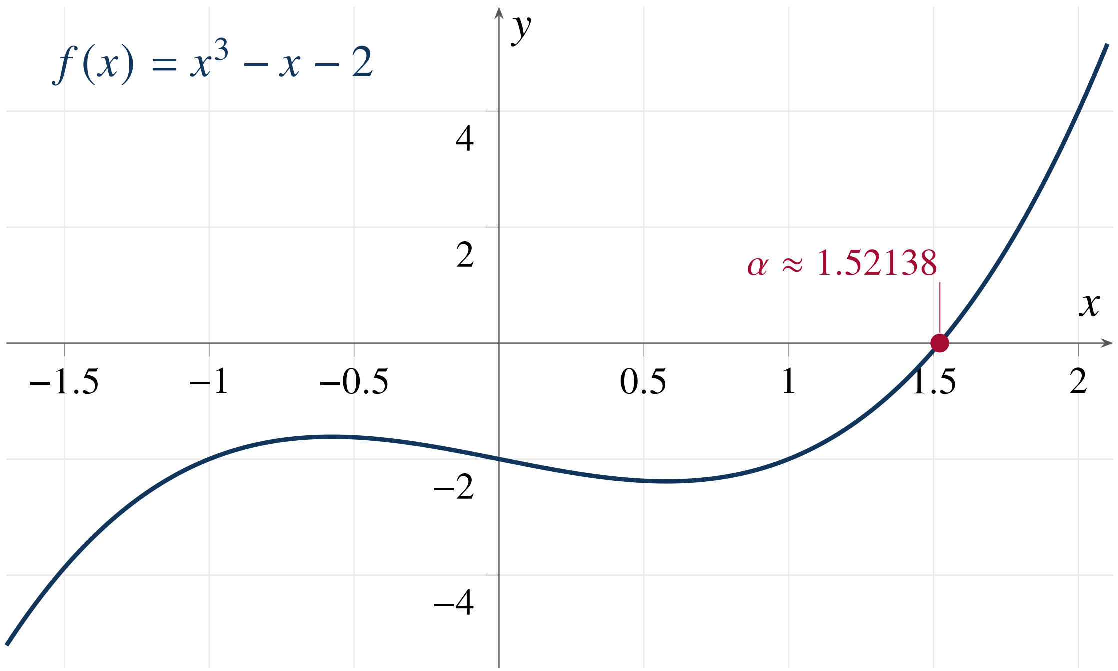
$$[a_0, b_0] = [1, 2], \qquad f(1) = -2, \quad f(2) = 4$$

$$c_0 = \frac{1+2}{2} = 1.5$$

$$f(1.5) = 1.5^3 - 1.5 - 2 = -0.125$$

The midpoint and the left endpoint have the same sign.

$$[a_1, b_1] = [1.5, 2]$$



Second midpoint and the next interval

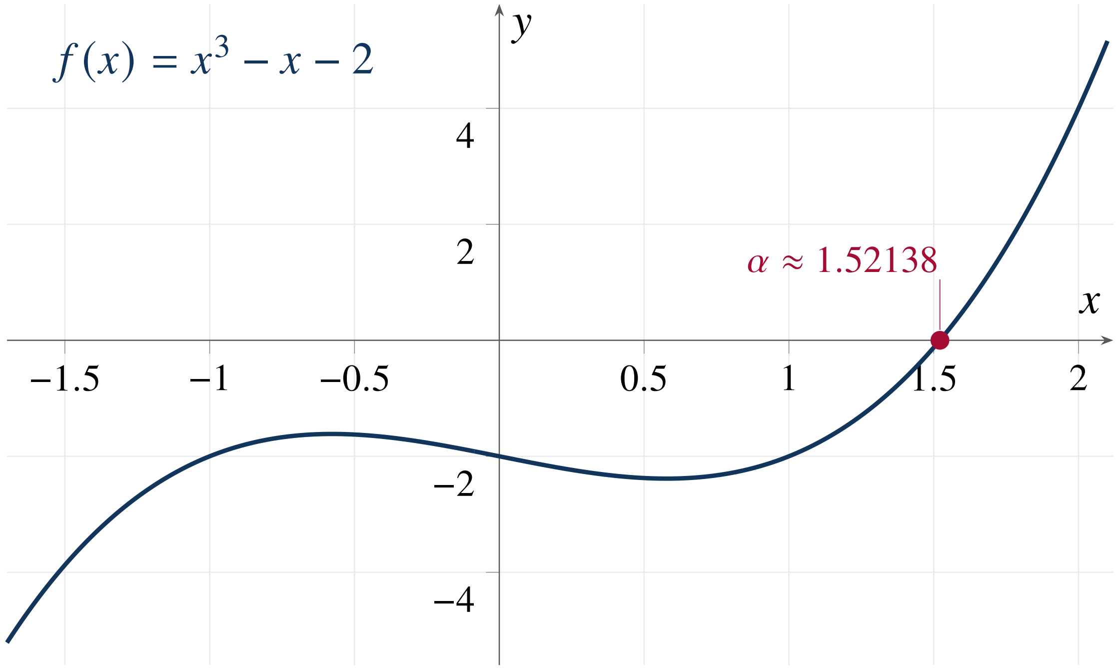
$$[a_1, b_1] = [1.5, 2]$$

$$c_1 = \frac{1.5 + 2}{2} = 1.75$$

$$f(1.75) = 1.75^3 - 1.75 - 2 = 1.609375$$

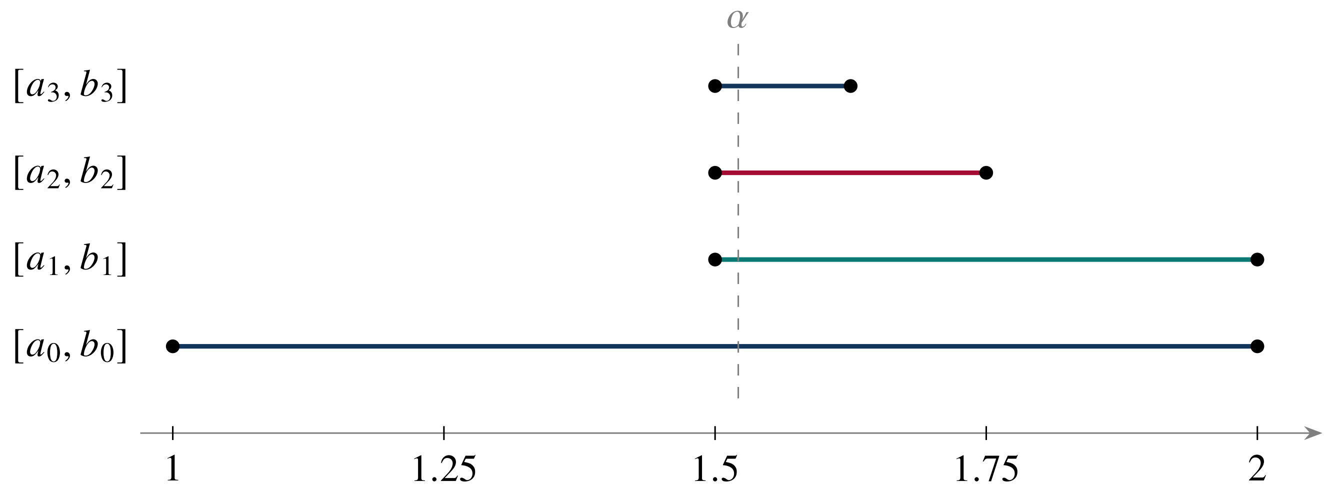
The midpoint and the left endpoint now have opposite signs.

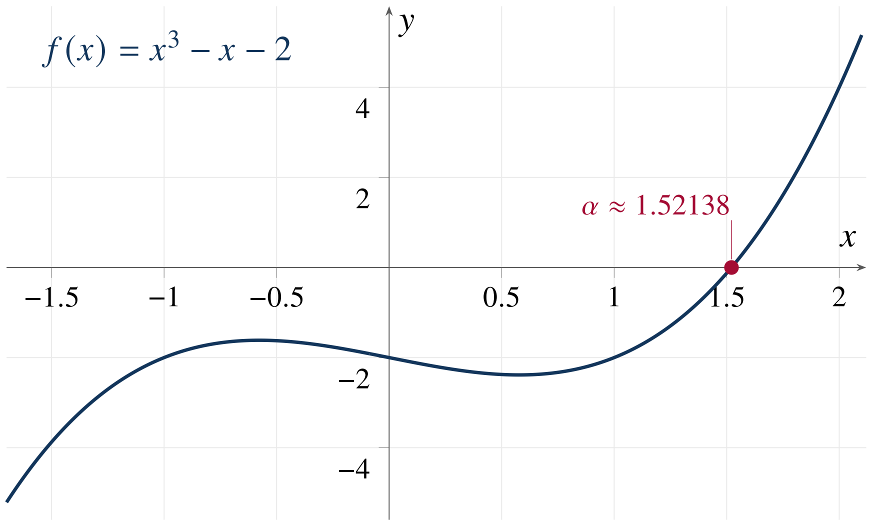
$$[a_2, b_2] = [1.5, 1.75]$$



The first seven bisection iterations

n	a_n	b_n	c_n	$f(c_n)$	$(b_n - a_n) / 2$
0	1.000000	2.000000	1.500000	−0.12500000	0.5000000
1	1.500000	2.000000	1.750000	+1.60937500	0.2500000
2	1.500000	1.750000	1.625000	+0.66601562	0.1250000
3	1.500000	1.625000	1.562500	+0.25219727	0.0625000
4	1.500000	1.562500	1.531250	+0.05911255	0.0312500
5	1.500000	1.531250	1.515625	−0.03405380	0.0156250
6	1.515625	1.531250	1.523438	+0.01225042	0.0078125





Convergence and a guaranteed error bound

Assume f is continuous on $[a_0, b_0]$ and $f(a_0)f(b_0) < 0$.

If no exact zero is encountered, the nested intervals converge to a common point α with $f(\alpha) = 0$.

$$b_n - a_n = \frac{b_0 - a_0}{2^n}$$

$$|c_n - \alpha| \leq \frac{b_0 - a_0}{2^{n+1}}$$

The index $n = 0$ refers to the midpoint of the original interval.

Number of iterations required for a desired accuracy

To guarantee $|c_n - \alpha| \le \varepsilon$, choose

$$n \ge \max\left\{0, \left\lceil \log_2 \frac{b_0 - a_0}{2\varepsilon} \right\rceil\right\}$$

$$b_0 - a_0 = 1, \quad \varepsilon = 10^{-6} \quad \Longrightarrow \quad n \ge 19$$

$$|c_{19} - \alpha| \le 2^{-20} \approx 9.537 \times 10^{-7}$$

Testing every midpoint from c_0 through c_{19} uses 20 midpoint evaluations, plus the two endpoints.

For absolute error 10^{-10}, 33 halvings suffice.

Method 2: The fixed-point method

Find a function g such that $$f(\alpha) = 0 \quad \iff \quad g(\alpha) = \alpha$$

For a nonzero constant λ, one possible choice is

$$g(x) = x - \lambda f(x), \qquad \lambda \neq 0$$

Conversely, a fixed point of g is a root of $f(x) = x - g(x)$.

Different choices of the function g can produce very different iteration behavior.

Fixed points found algebraically

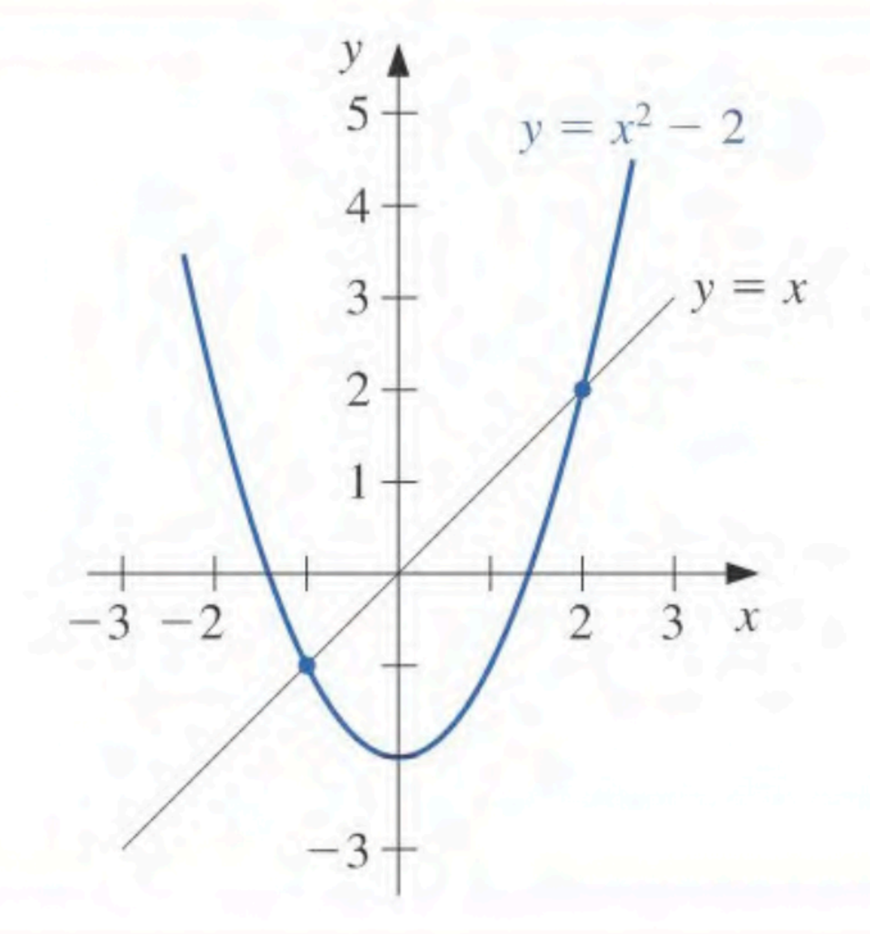
$$g(x) = x^2 - 2$$

$$\alpha = g(\alpha) \quad \iff \quad \alpha^2 - \alpha - 2 = 0$$

$$(\alpha + 1)(\alpha - 2) = 0$$

$$\alpha = -1 \quad \text{or} \quad \alpha = 2$$

These are the intersections of $y = g(x)$ and $y = x$.



Existence and uniqueness on an interval

Existence

If g is continuous on $[a, b]$ and maps every point of $[a, b]$ back into $[a, b]$, then it has a fixed point there

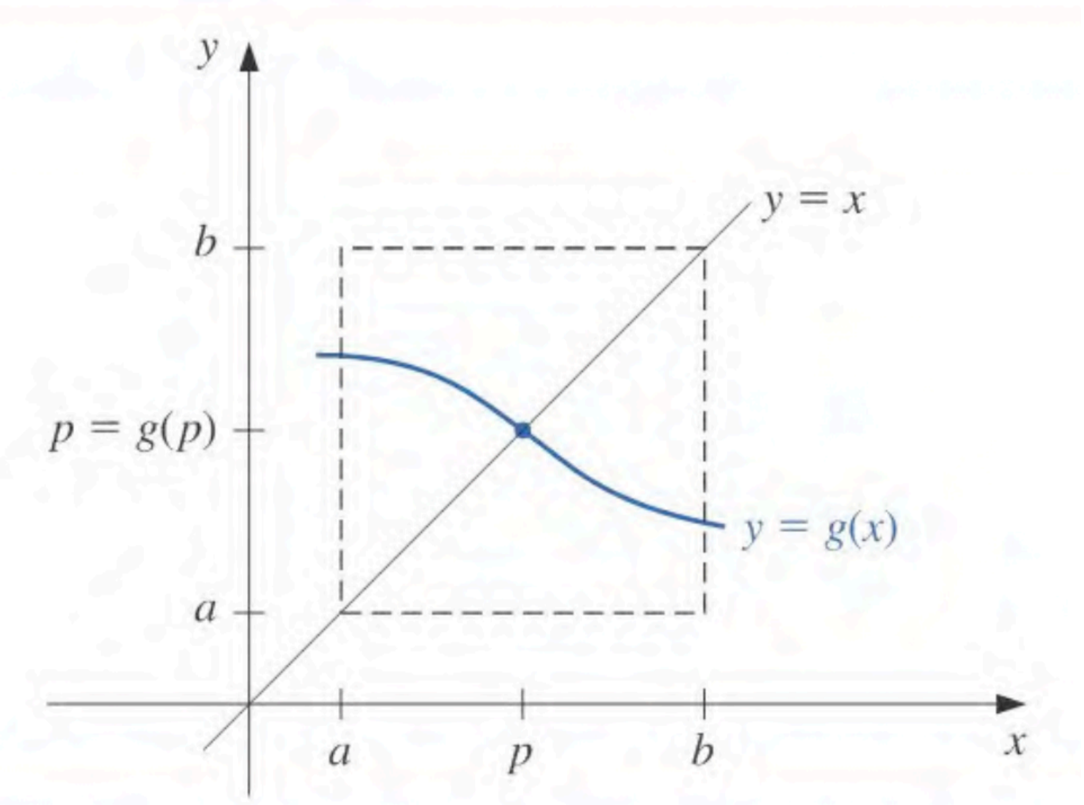
$$g([a, b]) \subseteq [a, b]$$

Uniqueness

If, in addition, g' exists on (a, b) and satisfies

$$|\, g'(x) \,| \leq k < 1 \qquad (a < x < b)$$

then the fixed point in the interval is unique.



The fixed point of an exponential function

$$g(x) = 3^{-x}, \qquad x \in [0, 1]$$

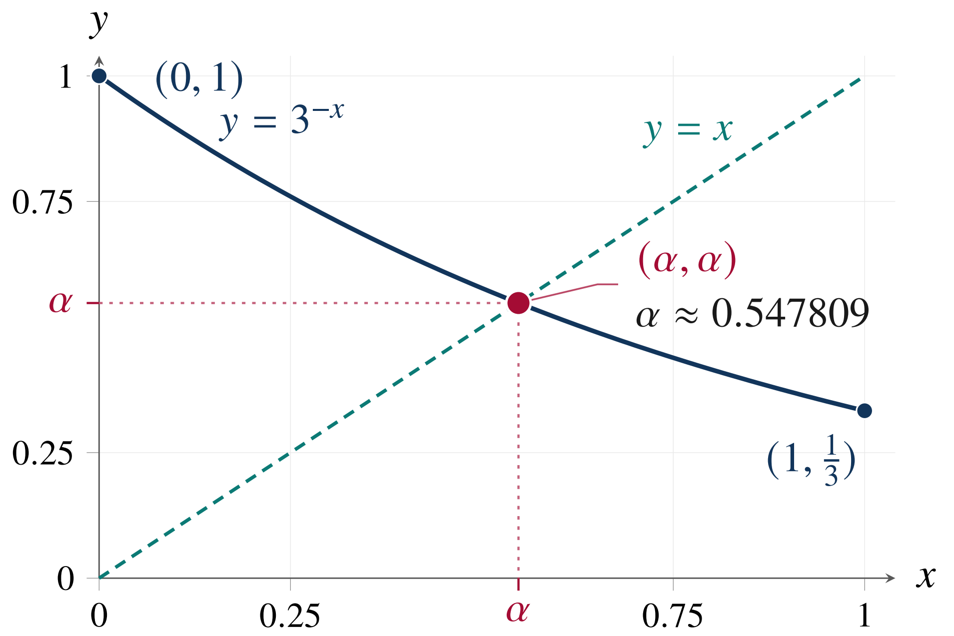
$$g([0, 1]) = [1 / 3, 1] \subseteq [0, 1]$$

Existence follows. The derivative test for uniqueness does not apply on the whole interval.

$$g'(x) = -(\log 3)3^{-x}, \qquad |g'(0)| = \log 3 > 1$$

Yet $g(x) - x$ is strictly decreasing, so it can have at most one zero.

$$\alpha = 3^{-\alpha}, \qquad \alpha \approx 0.547808621654$$



Fixed-point iterations

$$x_{n+1} = g(x_n), \qquad n = 0, 1, 2, \dots$$

Basic iteration

Choose x_0. Compute the next point, compare it with the previous point, and continue up to a finite iteration limit.

When $|x_n - x_{n-1}| \le \varepsilon$, the usual step test returns the point just computed, x_n.

A small step becomes a certified root error only when an appropriate bound is available.

$$x_n \to \alpha \text{ and } g \text{ continuous} \quad \Longrightarrow \quad g(\alpha) = \alpha$$

Several maps share the same fixed point

$$f(x) = x^3 - x - 2 = 0$$

$$g_1(x) = (x + 2)^{1/3}$$

$$g_2(x) = x^3 - 2$$

$$g_3(x) = \sqrt{1 + 2 / x} \quad (x > 0)$$

The root is the same. The derivative of each map controls its local behavior.

The first updates from the left endpoint

$$x_0 = 1$$

$$x_1 = (x_0 + 2)^{1/3} = \sqrt[3]{3} \approx 1.44224957030741$$

$$x_2 = (x_1 + 2)^{1/3} \approx 1.50989744933236$$

$$x_3 = (x_2 + 2)^{1/3} \approx 1.51972430499164$$

Use the stored values in each new update. Round only the displayed numbers.

Early iterates and a computable error bound

n	x_n	$\|x_n - \alpha\|$	$\frac{q}{1-q}\|x_n - x_{n-1}\|$
0	1.00000000000000	5.21×10^{-1}	–
1	1.44224957030741	7.91×10^{-2}	8.44×10^{-2}
2	1.50989744933236	1.15×10^{-2}	1.29×10^{-2}
3	1.51972430499164	1.66×10^{-3}	1.88×10^{-3}
4	1.52114126906273	2.38×10^{-4}	2.70×10^{-4}
5	1.52134536775190	3.43×10^{-5}	3.89×10^{-5}

The true error is shown for comparison. The final column can be computed without knowing the root.

A contraction keeps points close

A map $g: I \to I$ is a contraction if some constant $0 \le q < 1$ satisfies

$$|g(x) - g(y)| \le q\,|x - y| \qquad \text{for all } x, y \in I$$

Two separate checks

Self-mapping: $g(I) \subseteq I$.

Distance reduction: a derivative bound $|g'(x)| \le q < 1$ is sufficient on an interval.

Contraction theorem on an interval

Let $I = [a, b]$ be a nonempty closed interval. Assume $g(I) \subseteq I$ and that g is a contraction with constant $0 \leq q < 1$.

Conclusion

There is exactly one fixed point $\alpha \in I$. Every starting point $x_0 \in I$ generates iterates that stay in I and converge to α.

$$|x_n - \alpha| \leq q^n |x_0 - \alpha|$$

If $q = 0$, the map is constant and one update reaches the fixed point.

Verifying a contraction for the cubic

$$g_1(x) = (x+2)^{1/3}, \qquad I = [1,2]$$

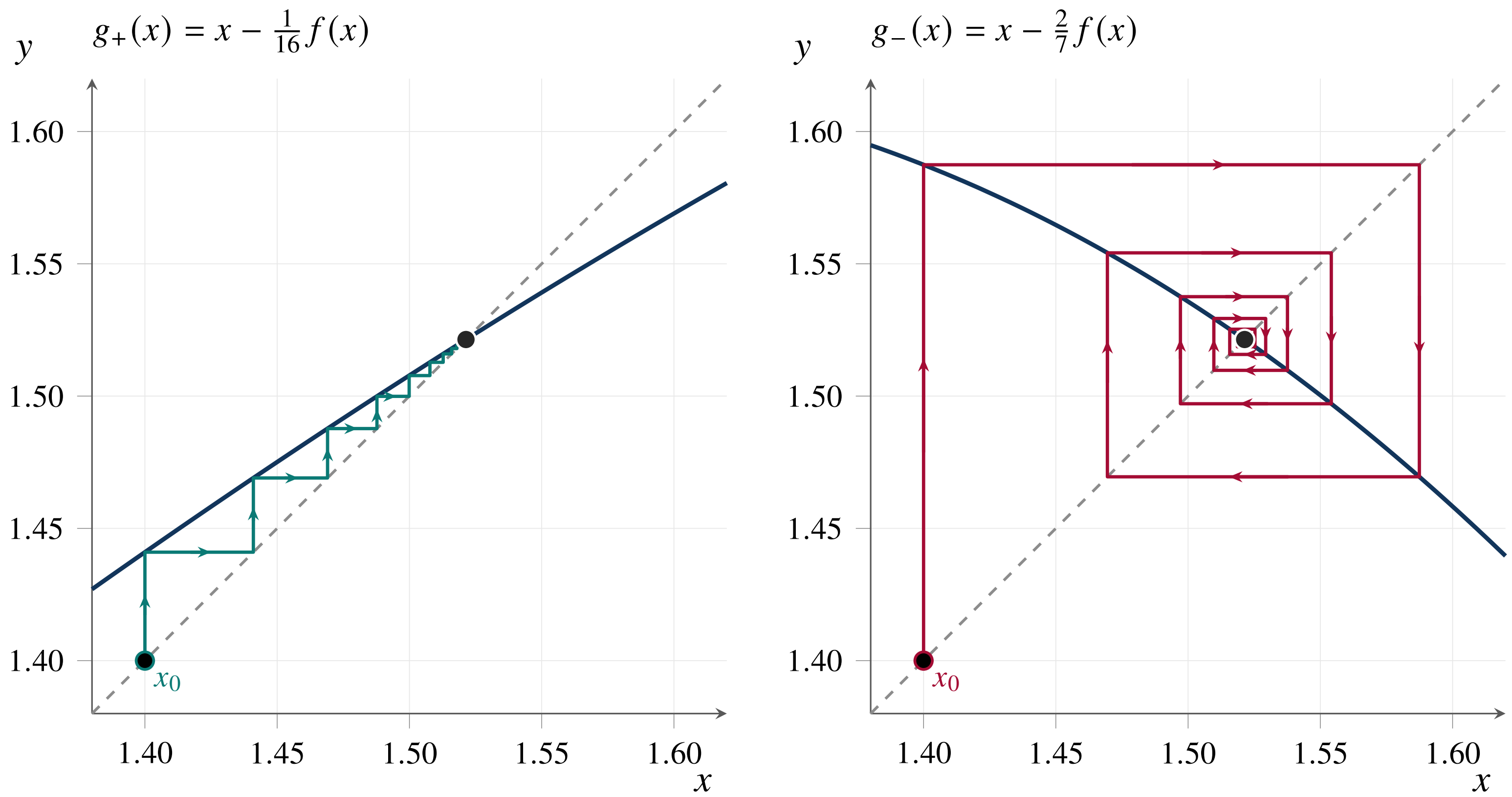
$$g_1(I) = [\sqrt[3]{3}, \sqrt[3]{4}] \subset [1,2]$$

$$0 < g_1'(x) = \frac{1}{3(x+2)^{2/3}} \leq \frac{1}{3 \cdot 3^{2/3}} = q$$

$$q \approx 0.16025 < 1$$

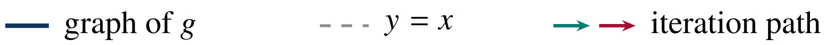
Every starting point in this interval converges to the cubic root.

Two other fixed-point maps for the cubic example $f(x) = x^3 - x - 2$



Local slope: $g'_+(\alpha) \approx 0.629$

Local slope: $g'_-(\alpha) \approx -0.698$



Both maps start at $x_0 = 1.4$ and are contractions on $[1.4, 1.6]$.

Method 3: Newton's method

A special case of the fixed-point method!

Recall: find a function g such that $f(\alpha) = 0 \iff g(\alpha) = \alpha$

For a nonzero constant λ, one possible choice is

$$g(x) = x - \lambda f(x), \qquad \lambda \neq 0$$

Now set $\lambda = \dfrac{1}{f'(x)}$:

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

Pseudocode

Input: f, f', a starting point, tolerances, and a maximum iteration count.

1. Evaluate and store $y_n = f(x_n)$. Return x_n if the value is zero.

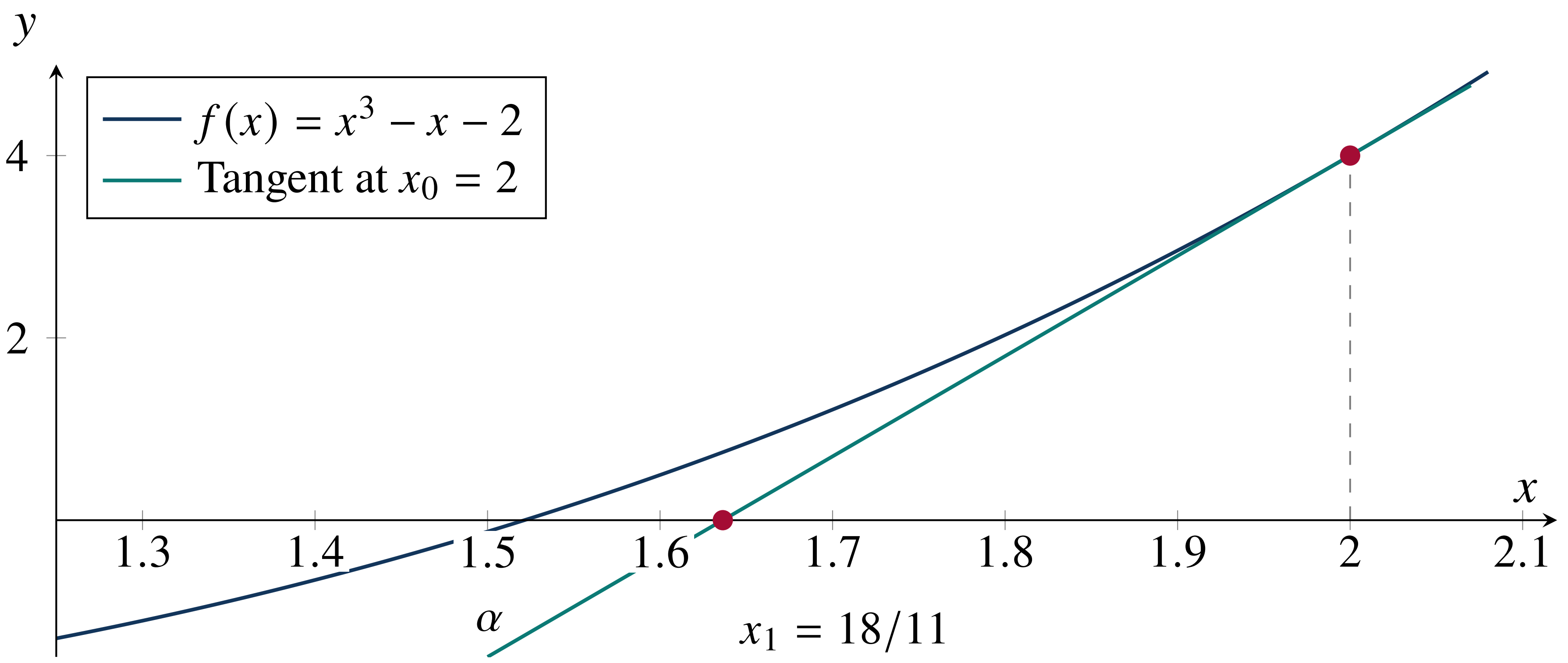
2. Evaluate $d_n = f'(x_n)$. Report failure if $d_n = 0$.

$$x_{n+1} = x_n - y_n \,/\, d_n, \qquad y_{n+1} = f(x_{n+1})$$

3. Return the new point if it meets the stopping conditions.

4. Otherwise continue. Report failure on an undefined or nonfinite value, stagnation, or the iteration limit.

Return to the cubic example!



$$x_0 = 2, \quad f(2) = 4, \quad f'(2) = 11, \qquad x_1 = 2 - \frac{4}{11} = \frac{18}{11}$$

Newton's method iterates for the cubic example

$$f(x) = x^3 - x - 2, \qquad f'(x) = 3x^2 - 1$$

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

$$x_0 = 1.5, \qquad f(x_0) = -0.125, \qquad f'(x_0) = 5.75$$

$$x_1 = 1.5 + \frac{0.125}{5.75} = 1.5217391304347826\ldots$$

The next tangent correction

$$x_1 = 1.5217391304347826\ldots$$

$$f(x_1) \approx 2.136927755 \times 10^{-3}, \qquad f'(x_1) \approx 5.947069943$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} \approx 1.5213798059647862$$

$$|x_1 - \alpha| \approx 3.59 \times 10^{-4}, \qquad |x_2 - \alpha| \approx 9.92 \times 10^{-8}$$

One more correction reduces the error by several orders of magnitude.

Rapid convergence in the iteration table

n	x_n	$\lvert f(x_n) \rvert$	$\lvert x_n - \alpha \rvert$
0	1.5000000000000000	1.25×10^{-1}	2.14×10^{-2}
1	1.5217391304347826	2.14×10^{-3}	3.59×10^{-4}
2	1.5213798059647862	5.89×10^{-7}	9.92×10^{-8}
3	1.5213797068045751	4.49×10^{-14}	7.55×10^{-15}
4	1.5213797068045676	2.60×10^{-28}	4.38×10^{-29}

Linear and quadratic convergence

For a sequence $x_n \to \alpha$, set $e_n = |x_n - \alpha|$.

$$\lim_{n\to\infty} \frac{e_{n+1}}{e_n^p} = \lambda > 0$$

Linear

$$p = 1, \qquad 0 < \lambda < 1$$

Eventually, each step reduces the error by approximately a fixed factor.

Quadratic

$$p = 2, \qquad \lambda > 0$$

Once the error is small, the next error is proportional to its square.

Local convergence: hypotheses

Let $f \in C^2$ near $I = [\alpha - r, \alpha + r]$, with $f(\alpha) = 0$.

$$|f'(x)| \geq m > 0, \qquad |f''(x)| \leq M \qquad (x \in I)$$

$$K = \frac{M}{2m}, \qquad 0 < \delta \leq r, \qquad K\delta \leq \frac{1}{2}$$

$$|x_0 - \alpha| \leq \delta$$

The derivative stays away from zero, the curvature is bounded, and the starting point is sufficiently close.

Local convergence: conclusions

Under the preceding hypotheses, every Newton step is well defined, stays within the chosen neighborhood, and converges to the root.

$$|e_{n+1}| \le K|e_n|^2, \qquad e_n = x_n - \alpha$$

If $f''(\alpha) \neq 0$ and there is no exact termination, the exact order is two:

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^2} = \left|\frac{f''(\alpha)}{2f'(\alpha)}\right|$$

For the cubic, the constant is approximately 0.768. If $M = 0$, one step is exact.

A zero derivative away from a root

$$f(x) = x^3 - 1, \qquad f'(x) = 3x^2$$

$$x_0 = 0, \qquad f(0) = -1, \qquad f'(0) = 0$$

The Newton denominator vanishes, although the starting point is not a root.

A derivative that is merely very small can instead produce a very large step.

An exact two-cycle

$$f(x) = x^3 - 2x + 2, \qquad f'(x) = 3x^2 - 2$$

$$N(0) = 0 - \frac{2}{-2} = 1$$

$$N(1) = 1 - \frac{1}{1} = 0$$

$$x_0 = 0, \quad x_1 = 1, \quad x_2 = 0, \quad x_3 = 1, \ldots$$

Smoothness and nonzero derivatives at the iterates do not guarantee convergence from an arbitrary start.

Method 4: Secant method

The function's derivative may be expensive to compute or unavailable. Estimate the slope from the two latest points.

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$

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

Distinct starting points are required. Their function values must also differ.

Pseudocode

1. Require distinct x_0, x_1. Store $y_0 = f(x_0)$ and $y_1 = f(x_1)$; return either point if it is a root.
2. Set $d_n = y_n - y_{n-1}$. Report failure if $d_n = 0$.

$$x_{n+1} = x_n - y_n \frac{x_n - x_{n-1}}{d_n}, \qquad y_{n+1} = f(x_{n+1})$$

3. Apply the same step and residual tests as for Newton. Retain the newest pair if the tests fail.
4. Report failure for undefined or nonfinite arithmetic, stagnation, or the iteration limit.

Only one new function evaluation is needed per update after initialization.

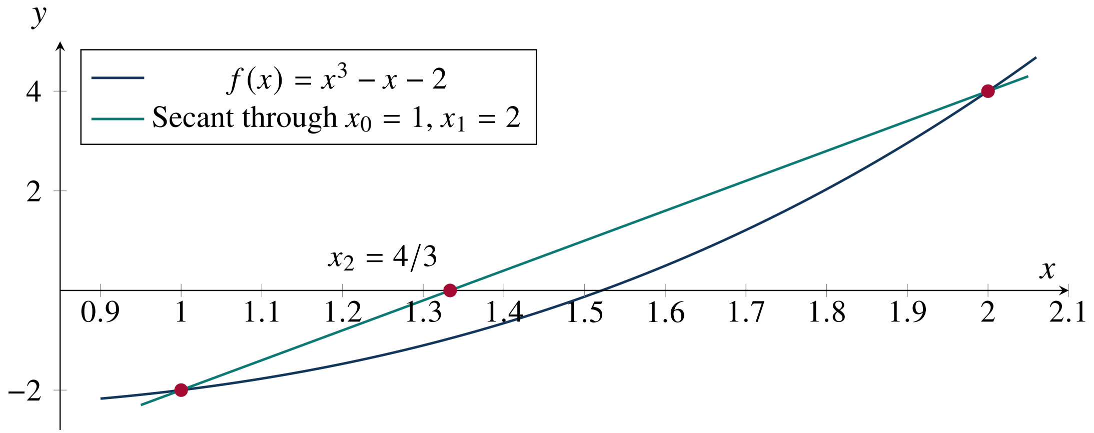
First secant update for the cubic

$$x_0 = 1, \qquad x_1 = 2, \qquad f(x_0) = -2, \quad f(x_1) = 4$$

$$d_1 = \frac{4 - (-2)}{2 - 1} = 6$$

$$x_2 = 2 - \frac{4}{6} = \frac{4}{3} \approx 1.33333333333333$$

The new point is the horizontal-axis intercept of the line through the two sampled points.



Use the newest pair for the next step

$$x_1 = 2, \qquad x_2 = \frac{4}{3}, \qquad f(x_2) = -\frac{26}{27}$$

$$x_3 = \frac{4}{3} - \left(-\frac{26}{27}\right)\frac{\frac{4}{3} - 2}{-\frac{26}{27} - 4}$$

$$x_3 = \frac{98}{67} \approx 1.46268656716418$$

The next step will use x_2 and x_3, whose function values have the same sign.

The early secant iterates

n	x_n	$\lvert f(x_n)\rvert$	$\lvert x_n - \alpha\rvert$
0	1.00000000000000	2.00×10^{0}	5.21×10^{-1}
1	2.00000000000000	4.00×10^{0}	4.79×10^{-1}
2	1.33333333333333	9.63×10^{-1}	1.88×10^{-1}
3	1.46268656716418	3.33×10^{-1}	5.87×10^{-2}
4	1.53116943214120	5.86×10^{-2}	9.79×10^{-3}
5	1.52092642051528	2.69×10^{-3}	4.53×10^{-4}

The new secant slope changes as the latest points move toward the root.

Later iterates gain accuracy rapidly

n	x_n	$\lvert f(x_n)\rvert$	$\lvert x_n - \alpha\rvert$
6	1.52137631666974	2.02×10^{-5}	3.39×10^{-6}
7	1.52137970798487	7.02×10^{-9}	1.18×10^{-9}
8	1.52137970680456	1.83×10^{-14}	3.07×10^{-15}
9	1.52137970680457	1.66×10^{-23}	2.78×10^{-24}

Local convergence: hypotheses

Let $f \in C^2$ near $I = [\alpha - r, \alpha + r]$, where $f(\alpha) = 0$.

$$|f'(x)| \geq m > 0, \qquad |f''(x)| \leq M \quad (x \in I)$$

$$K = \frac{M}{2m}, \qquad 0 < \delta \leq r, \qquad K\delta \leq \frac{1}{2}$$

$$x_0, x_1 \in J = [\alpha - \delta, \alpha + \delta], \qquad x_0 \neq x_1$$

Both starting points must be sufficiently close. They need not bracket the root.

Local convergence: conclusions

The method stays in J and remains well defined until it reaches α, or it generates an infinite sequence converging to α.

$$E_{n+1} \le K E_n E_{n-1}, \qquad n \ge 1$$

$$E_n \le 2^{-(n-1)} E_1 \quad (n \ge 1), \qquad \frac{E_{n+1}}{E_n} \to 0$$

The golden-ratio order of convergence

Under the local convergence hypotheses, additionally assume $f''(\alpha) \neq 0$ and no exact termination.

$$A = \left| \frac{f''(\alpha)}{2f'(\alpha)} \right| > 0, \qquad \varphi = \frac{1+\sqrt{5}}{2} \approx 1.618$$

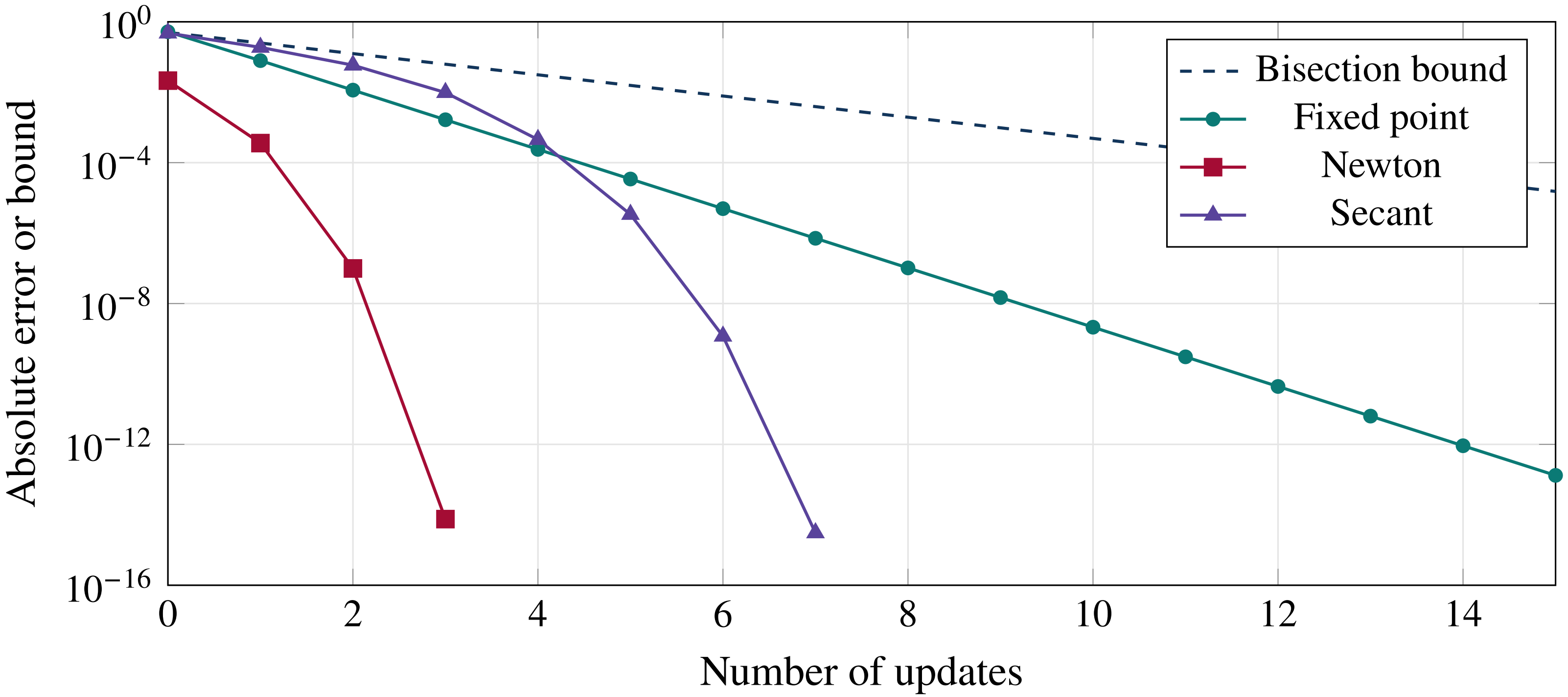
$$\lim_{n\to\infty} \frac{E_{n+1}}{E_n^{\varphi}} = A^{1/\varphi}$$

This order lies between linear and quadratic convergence — sometimes referred to as *superlinear*.

Comparing the root-finding methods

Method	Starting information	Convergence
Bisection	A sign-changing interval	The error bound halves
Fixed point	A map and an initial point	Usually linear
Newton	A function, derivative, and initial point	Quadratic in the generic case
Secant	A function and two initial points	Superlinear, golden-ratio order

Comparing the methods for the cubic example



Bisection shows an error bound. The other curves show actual errors. Values below 10^{-16} are omitted.