Zeros & Roots

$$f(x) = 0$$

Root Finding

A root-finding algorithm is a numerical method, or algorithm, for finding a value x such that f(x) = 0, for a given function f.

Such an x is called a ***root*** of the function f.

Root Finding

Finding a root of

$$f(x) - g(x) = 0$$

is the same as solving the equation

$$f(x) = g(x)$$

Here, x is called the unknown in the equation. Conversely, any equation can take the canonical form

$$f(x) = 0$$

so equation solving is the same thing as computing (or finding) a root of a function.

Root Finding

Numerical root-finding methods use ***iteration***, producing a sequence of numbers that hopefully ***converge*** towards a limit (the so called "fixed point") which is a root.

The first values of this series are ***initial guesses***. The method computes subsequent values based on the old ones and the function f.

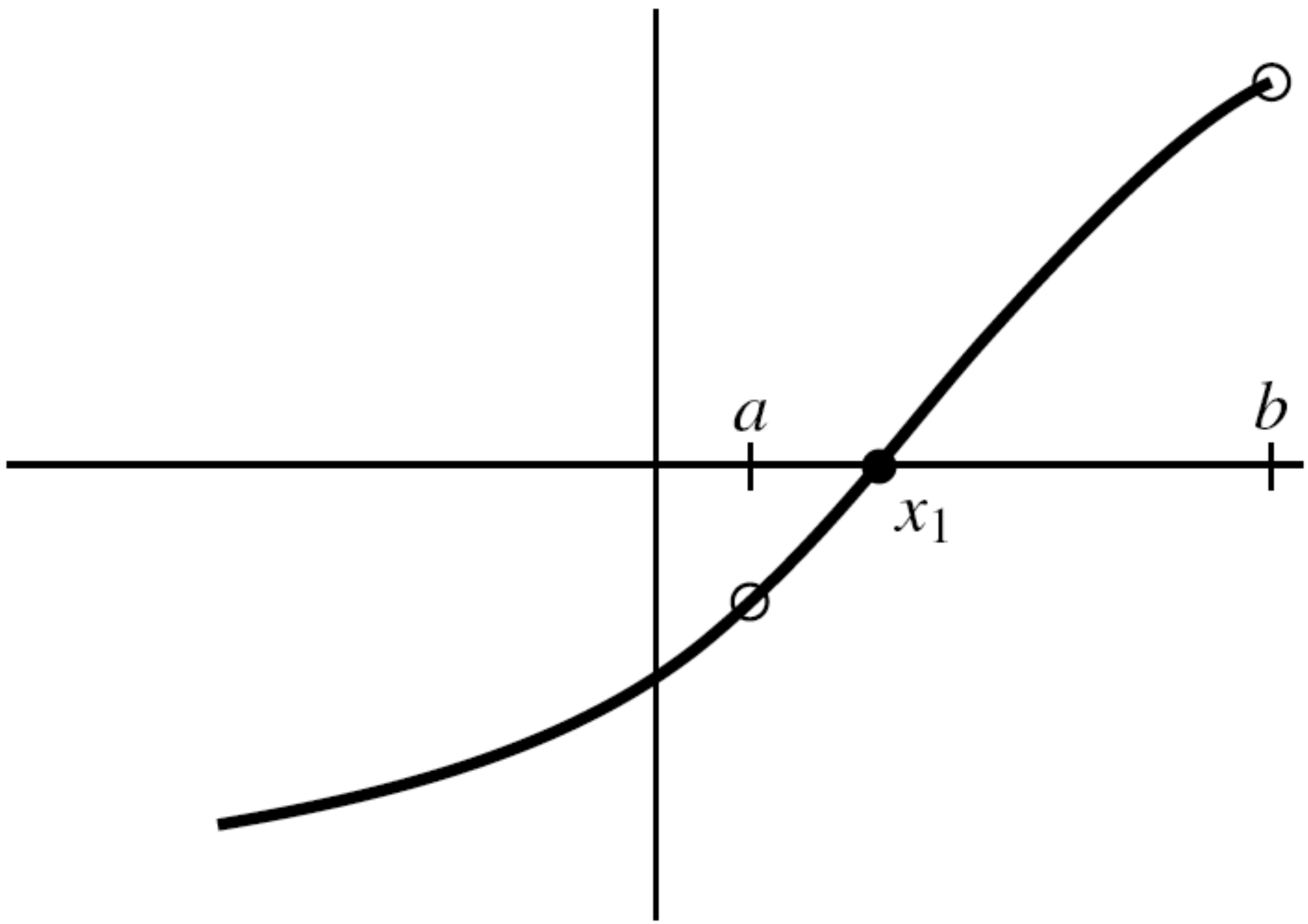
Root Finding

The behavior of root-finding algorithms is studied in ***numerical analysis***.

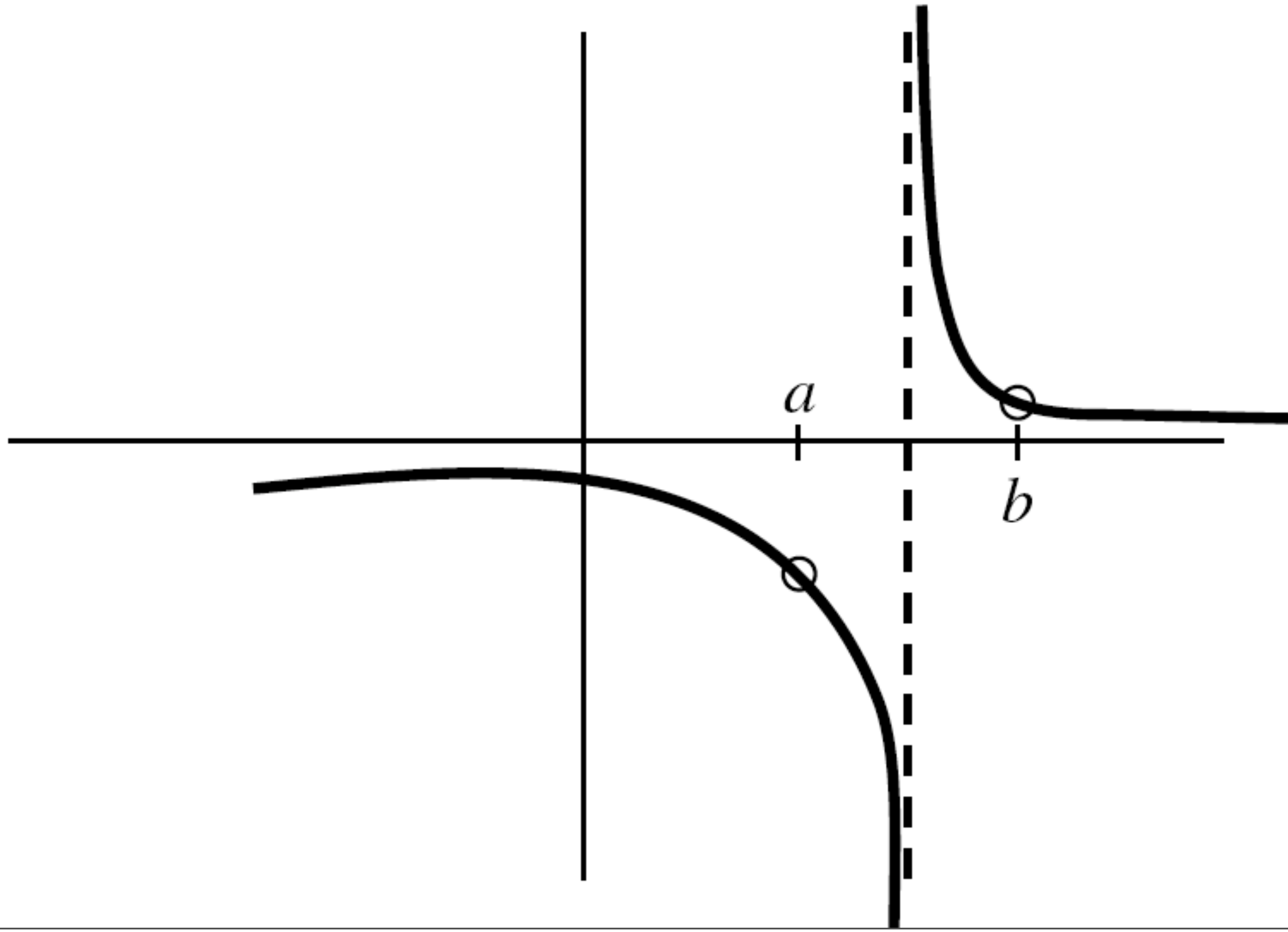
Algorithms perform best when they take advantage of known characteristics of the given function. Thus an algorithm to find isolated real roots of a low-degree polynomial in one variable may bear little resemblance to an algorithm for complex roots of a "black-box" function which is not even known to be differentiable.

Questions include ability to separate close roots, robustness in achieving reliable answers despite inevitable numerical errors, and rate of convergence.

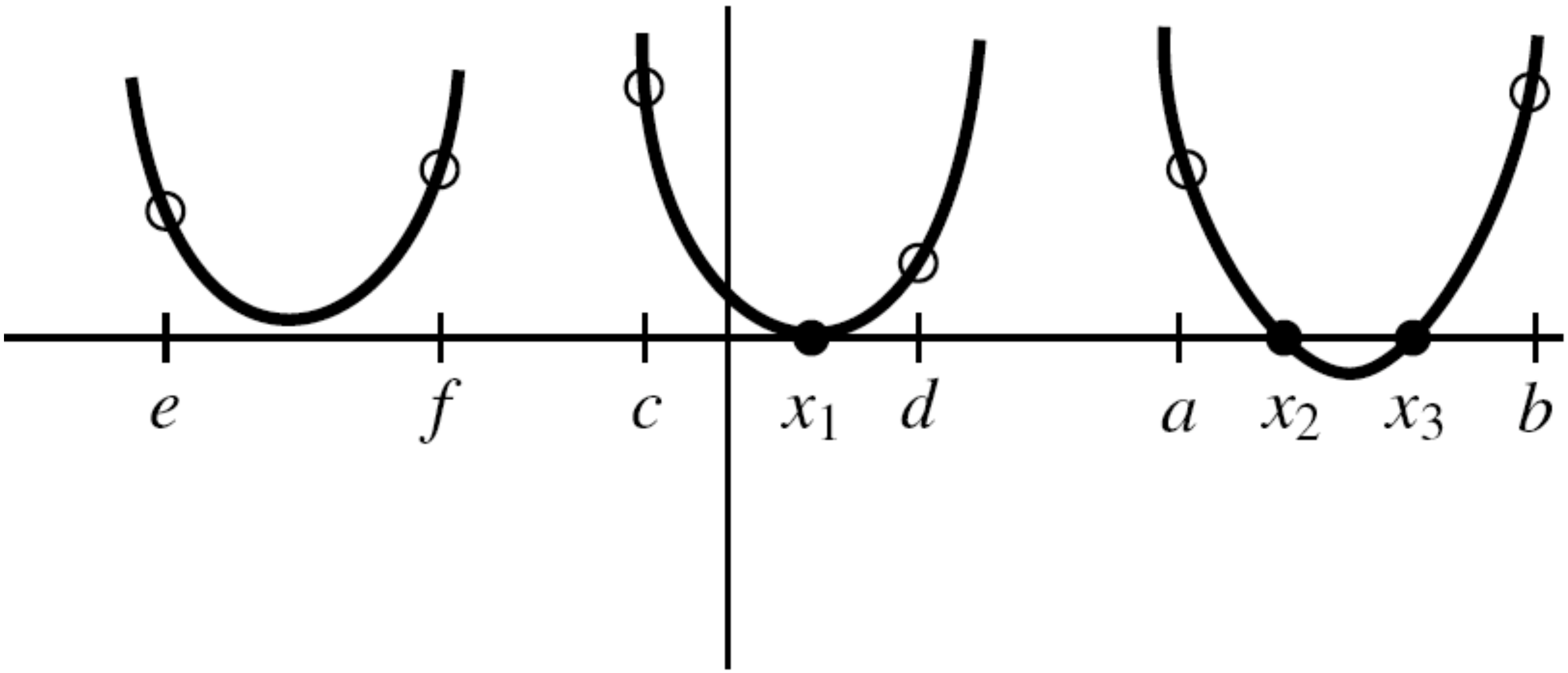
An isolated root bracketed by two points *a* and *b* for which the function has opposite sign



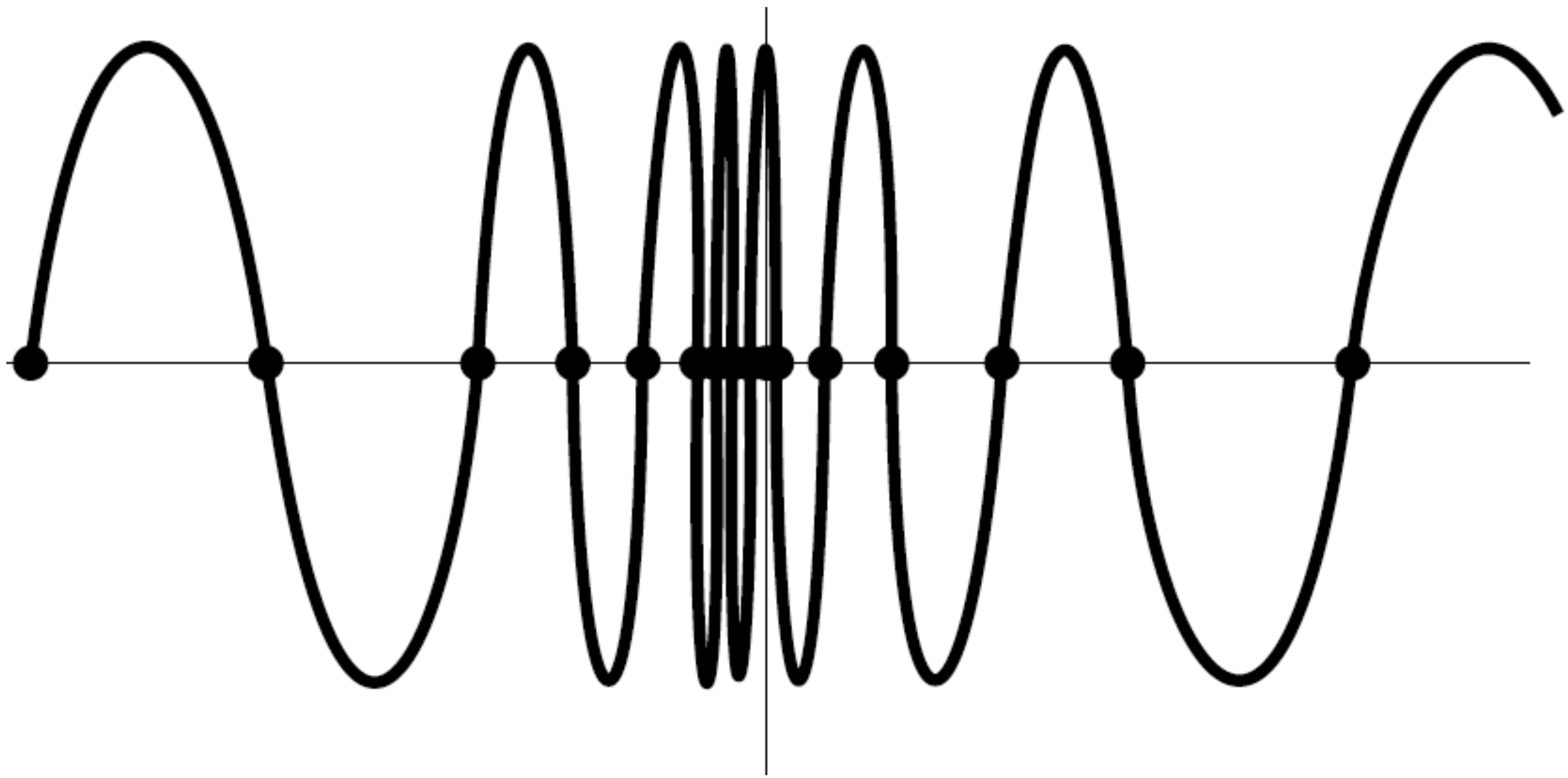
The function has opposite signs at points *a* and *b*, but the points bracket a *singularity* <u>not</u> a *root*



There is not necessarily a sign change in the function near a double root



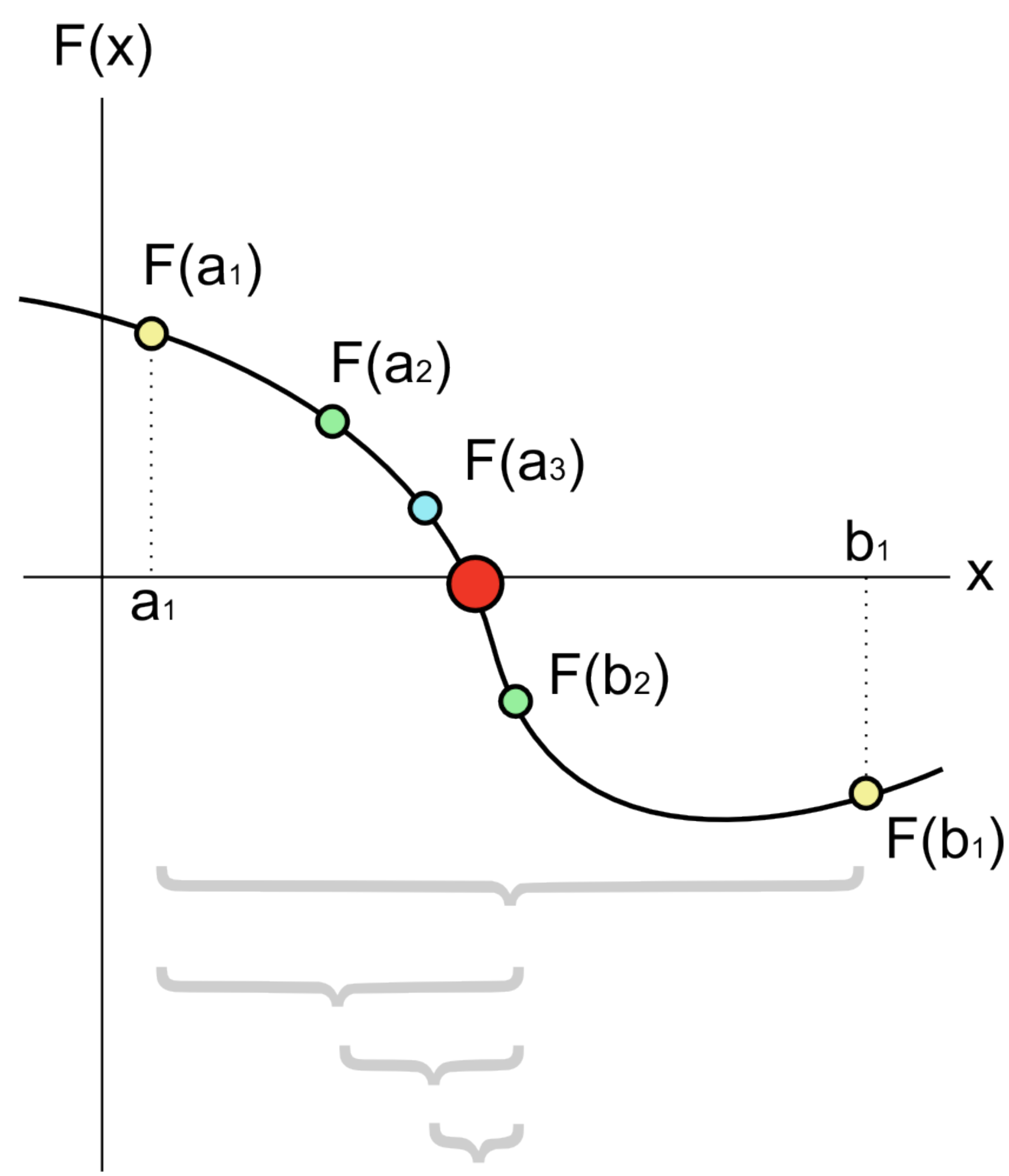
A function with many roots



Bisection Method

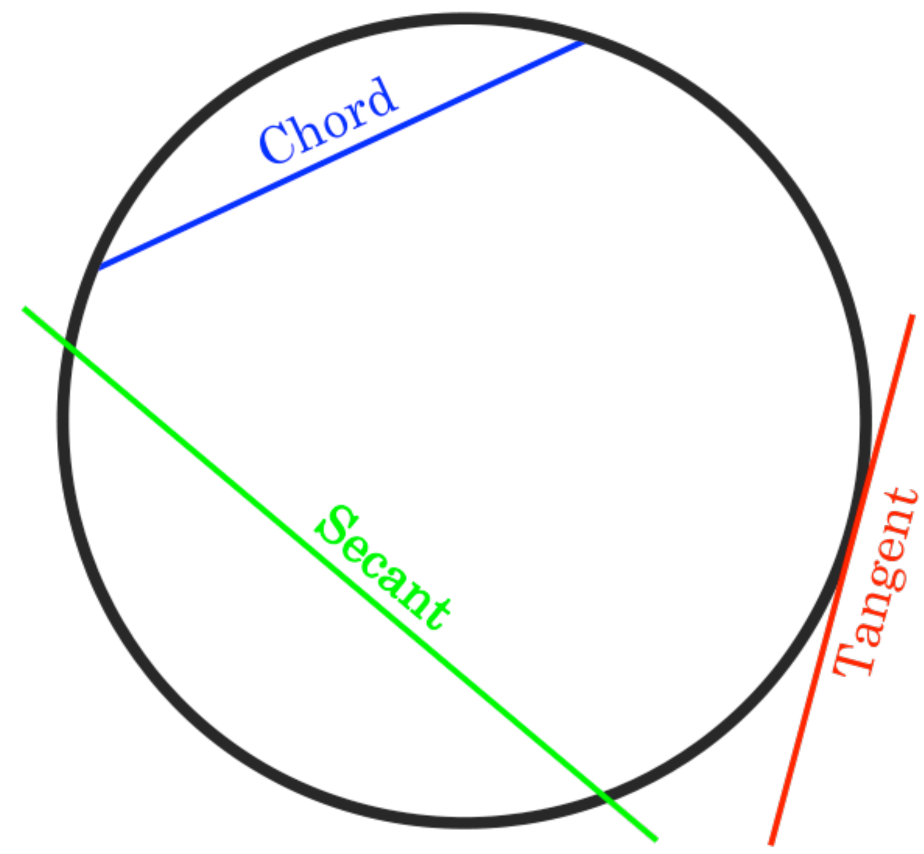
The bisection repeatedly divides an interval in half and then selects the subinterval in which a root exists. It is a very ***simple*** and ***robust*** method, but it is also rather ***slow***.

A few steps of the bisection method applied over the starting range $[a_1;b_1]$. The bigger red dot is the root of the function.



Secant line

A secant line of a curve is a line that (locally) intersects two points on the curve. The word secant comes from the Latin *secare*, meaning ***to cut***.



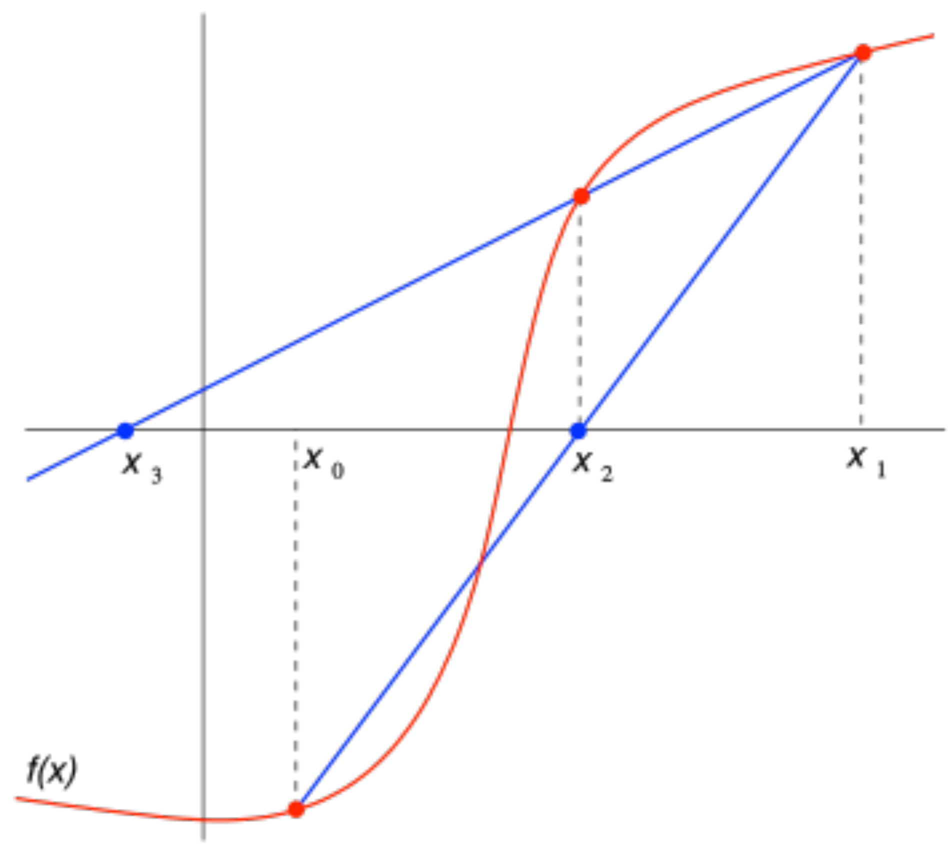
Secant Method

The secant method uses a succession of roots of secant lines to better approximate a root of a function f.

The figure shows the first two iterations of the secant method. The red curve shows the function f and the blue lines are the secants.

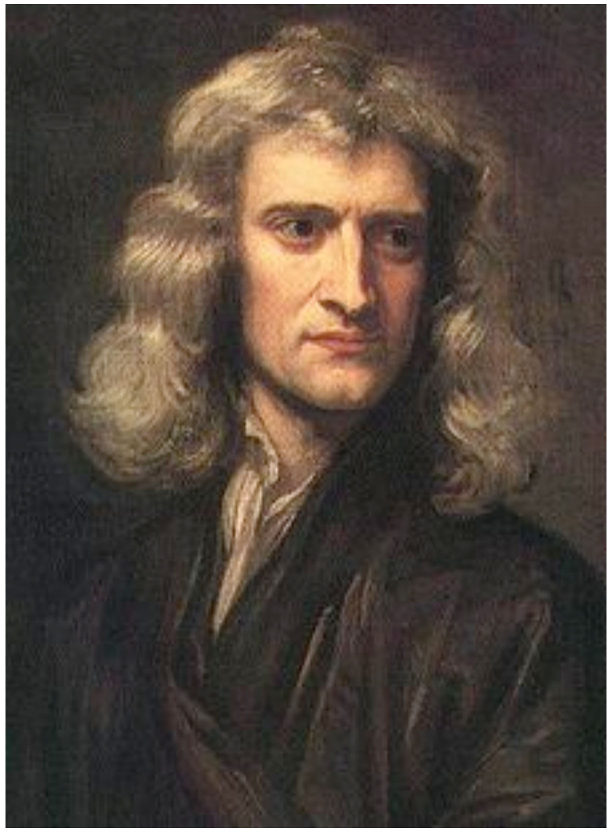
The function is assumed to be approximately linear in local region of interest. The next improvement in the root is taken as the point where the approximating line crosses the axis.

This is repeated until we reach a sufficiently high level of precision (a sufficiently small difference between x_n and x_{n-1}.



$$x_n = x_{n-1} - f(x_{n-1})\frac{x_{n-1} - x_{n-2}}{f(x_{n-1}) - f(x_{n-2})}$$

Newton-Raphson Method



Probably the most celebrated root finding method is the Newton-Raphson method.

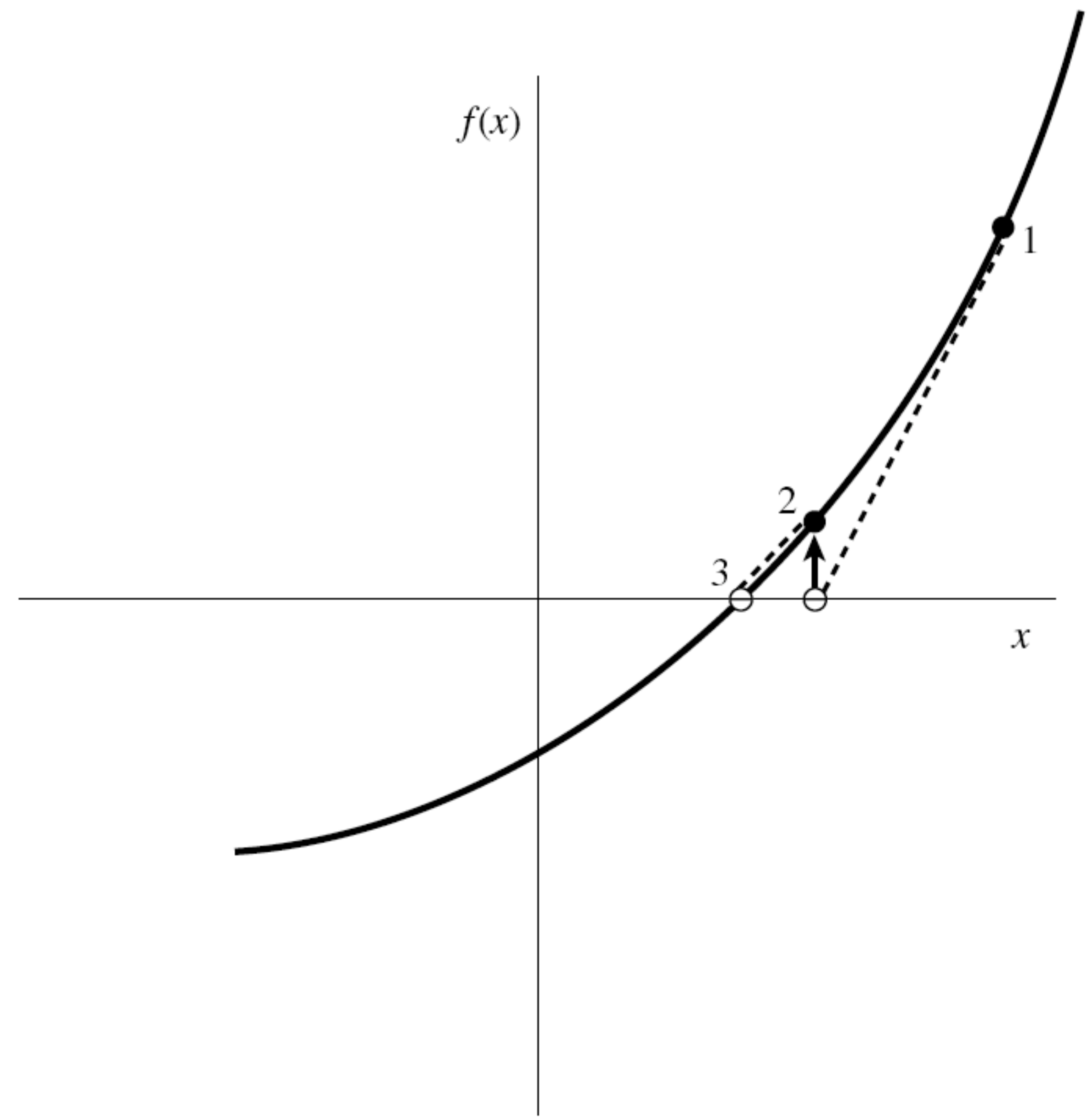
Given a function f(x) and its derivative f'(x), we begin with a first guess x_0 . A better approximation x_1 is

$$x_1 = x_o - \frac{f(x_o)}{f'(x_o)}$$

Newton-Raphson Method

This method is distinguished from the previous method in that it requires the evaluation of both the function and its derivative at arbitrary points.

Geometrically it consists of extending a tangent line until it crosses zero.



Newton-Raphson Method

Algebraically the method derives from the familiar Taylor series expansion of a function

$$f(x+\delta x) \approx f(x) + f'(x)\delta + \frac{f''(x)}{2}\delta^2 + \cdots$$

for small enough δ and well behaved functions the terms beyond linear are small hence f(x+δ)=0 implies

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

Newton-Raphson Method

The method readily generalizes to more than one dimension.

Newton's method can often converge remarkably quickly, especially if the iteration begins "sufficiently near" the desired root.

Unfortunately, when iteration begins far from the desired root, Newton's method can easily lead an unwary user astray with little warning.

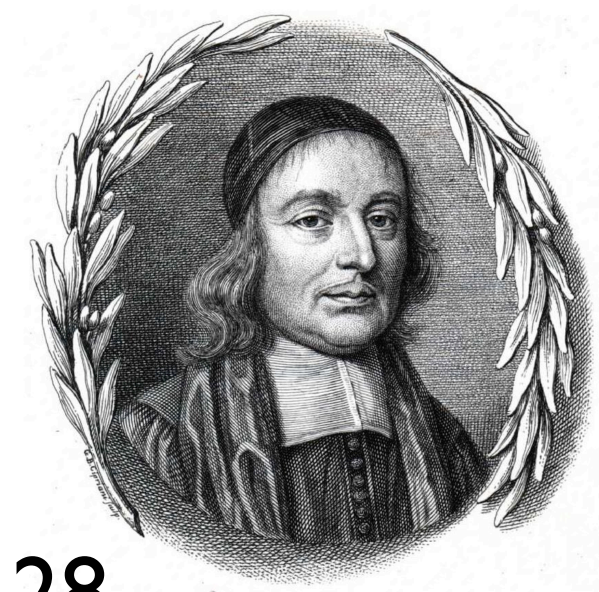
Thus, good implementations of the method embed it in a routine that also detects and perhaps overcomes possible convergence failures.

More Methods

Check out more methods in Numerical Recipes Chapter 9:

- Secant Method, False Position Method, and Ridders' Method
- Van Wijngaarden-Dekker-Brent Method

John Wallis



John Wallis (November 23, 1616 - October 28, 1703) was an English mathematician who is given partial credit for the development of modern calculus. Between 1643 and 1689 he served as chief cryptographer for Parliament and, later, the royal court. He is also credited with introducing the symbol ∞ for infinity.

The polynomial used by Wallis when he first presented Newton's method to the French Academy was

$$x^3-2x-5$$

Solution of x^3-2x-5

There are three roots of Wallis' function, one real root, between x = 2 and x = 3, and a pair of complex conjugate roots:

Solution of x^3-2x-5

There are three roots of Wallis' function, one real root, between x = 2 and x = 3, and a pair of complex conjugate roots:

$$\begin{pmatrix} \dfrac{2}{3\left(\frac{\sqrt{108}\,\sqrt{643}}{108}+\frac{5}{2}\right)^{\frac{1}{3}}} + \left(\dfrac{\sqrt{108}\,\sqrt{643}}{108}+\dfrac{5}{2}\right)^{\frac{1}{3}} \\ -\dfrac{1}{3\left(\frac{\sqrt{108}\,\sqrt{643}}{108}+\frac{5}{2}\right)^{\frac{1}{3}}} - \dfrac{\left(\frac{\sqrt{108}\,\sqrt{643}}{108}+\frac{5}{2}\right)^{\frac{1}{3}}}{2} - \dfrac{\sqrt{3}\left(\frac{2}{3\left(\frac{\sqrt{108}\,\sqrt{643}}{108}+\frac{5}{2}\right)^{\frac{1}{3}}} - \left(\frac{\sqrt{108}\,\sqrt{643}}{108}+\frac{5}{2}\right)^{\frac{1}{3}}\right)\mathrm{i}}{2} \\ -\dfrac{1}{3\left(\frac{\sqrt{108}\,\sqrt{643}}{108}+\frac{5}{2}\right)^{\frac{1}{3}}} - \dfrac{\left(\frac{\sqrt{108}\,\sqrt{643}}{108}+\frac{5}{2}\right)^{\frac{1}{3}}}{2} + \dfrac{\sqrt{3}\left(\frac{2}{3\left(\frac{\sqrt{108}\,\sqrt{643}}{108}+\frac{5}{2}\right)^{\frac{1}{3}}} - \left(\frac{\sqrt{108}\,\sqrt{643}}{108}+\frac{5}{2}\right)^{\frac{1}{3}}\right)\mathrm{i}}{2} \end{pmatrix}$$

Find root symbolically

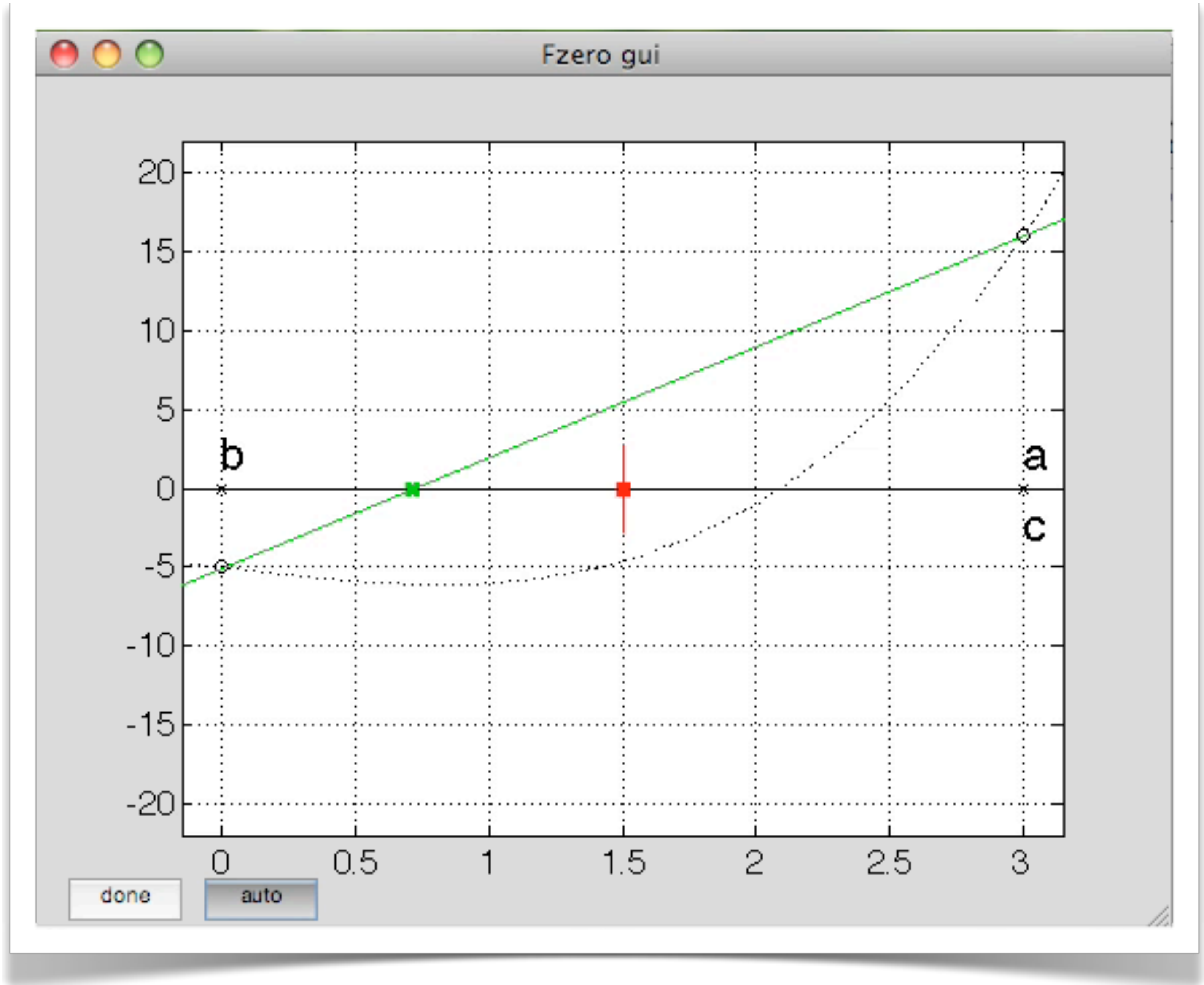
```
syms x
solve(x^3 - 2*x -5)
```

Use fzerogui to find root

```
F = @(x) x.^3-2*x-5;
[out1, out2] = fzerogui(F,[0 3])
```

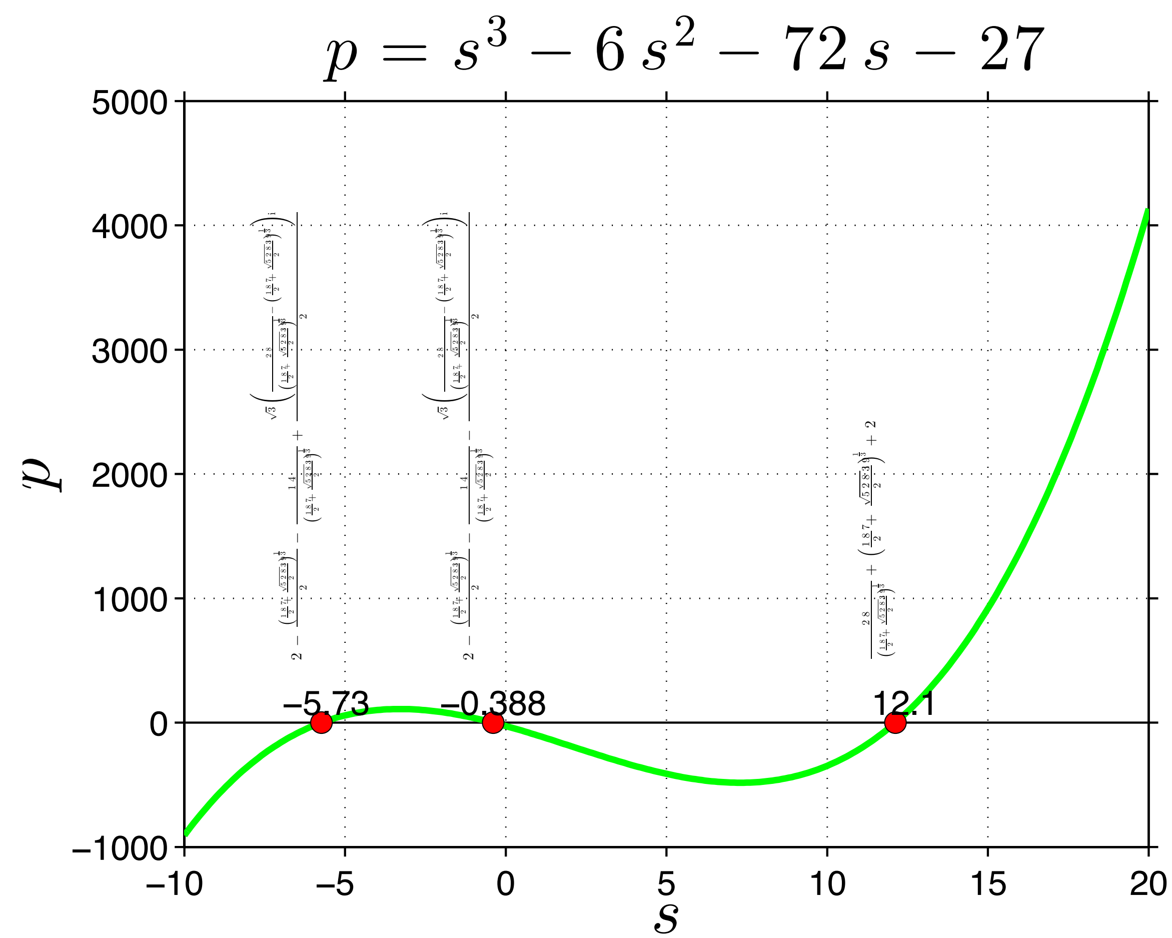
fzero

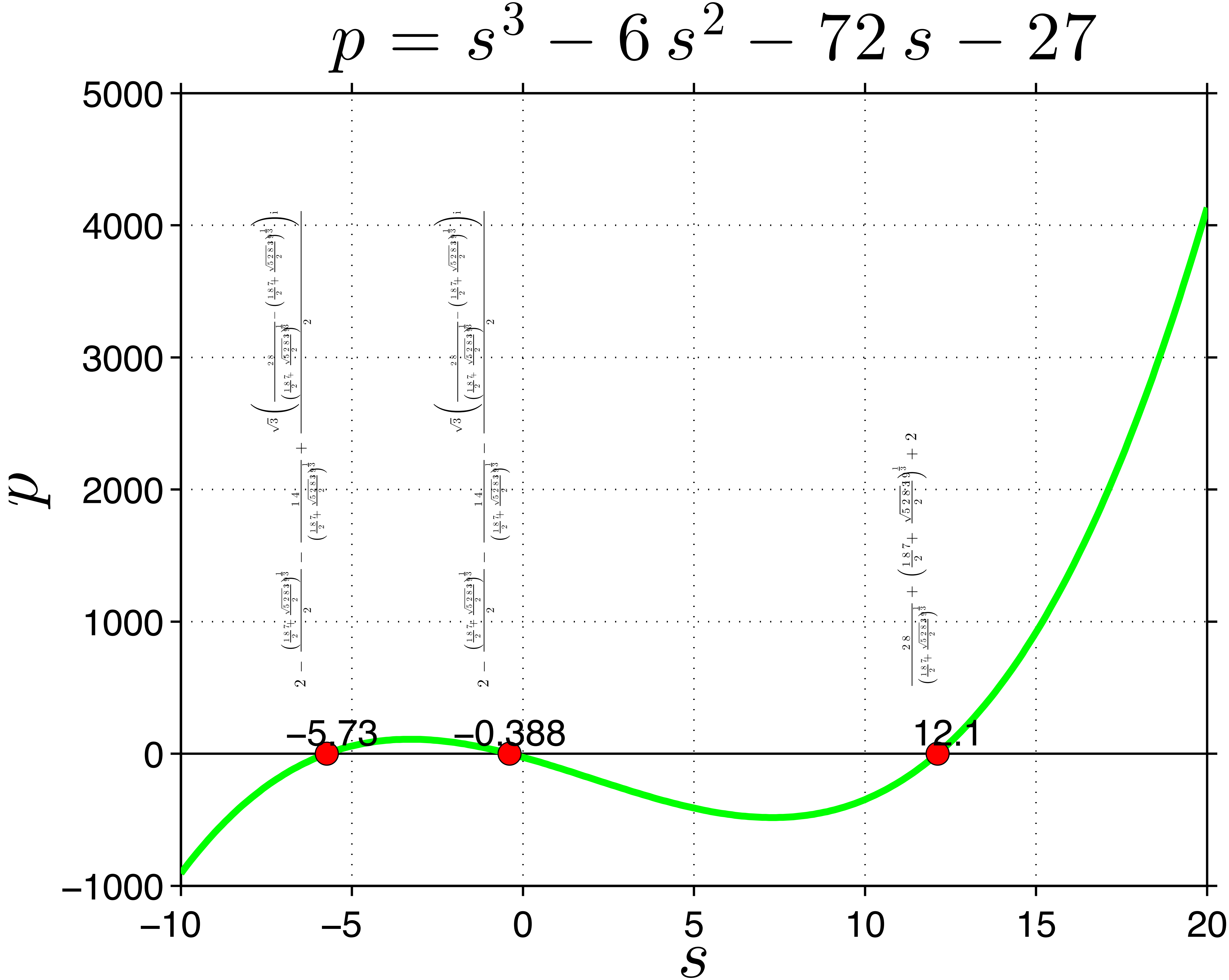
A Matlab root finding algorithm is called fzero. It has several features beyond the basic algorithms. A preamble takes a single starting guess and searches for an interval with a sign change. The values returned by the function f(x) are checked for infinities, NaNs, and complex numbers. Default tolerances can be changed. Additional output, including a count of function evaluations, can be requested.



fzero and roots

1. In Matlab help browser type "`find roots`" in the search results field. Implement the examples shown
2. do `doc fzero` and do the examples.
3. Use the roots function to find and plot the roots of the polynomial as shown to the right: $s^3 - 6s^2 - 72s - 27$





$p = s^3 - 6\,s^2 - 72\,s - 27$
5000
4000
3000
2000
1000
0
−1000
−10
−5
0
5
10
15
20
s
p
−5.73
−0.388
12.1

Examples

Use `fzerogui` to try to find a zero of each of the following functions in the given interval. Do you see any interesting or unusual behavior?

$x^3 - 2x - 5$	$[0, 3]$
$\sin x$	$[1, 4]$
$x^3 - 0.001$	$[-1, 1]$
$\log(x + 2/3)$	$[0, 1]$
$\text{sign}(x-2)\sqrt{\lvert x-2 \rvert}$	$[1, 4]$
$\text{atan}(x) - \pi/3$	$[0, 5]$
$1/(x - \pi)$	$[0, 5]$

Examples

Here is a little footnote to the history of numerical methods. The polynomial

$$x^3 - 2x - 5$$

was used by Wallis when he first presented Newton's method to the French Academy. It has one real root, between $x = 2$ and $x = 3$, and a pair of complex conjugate roots.
(a) Use the Symbolic Toolbox to find symbolic expressions for the three roots. Warning: The results are not pretty. Convert the expressions to numerical values.
(b) Use the `roots` function in MATLAB to find numerical values for all three roots.
(c) Use `fzerotx` to find the real root.
(d) Use Newton's method starting with a complex initial value to find a complex root.
(e) Can bisection be used to find the complex root? Why or why not?