

Lecture 7

Systems of Nonlinear Equations

Virginia Tech Department of Computer Science
Fall 2026 • Sumeet Khatri

Recap of Newton's method: solving a non-linear equation

To solve $f(x) = 0$, begin with a current approximation x_n .

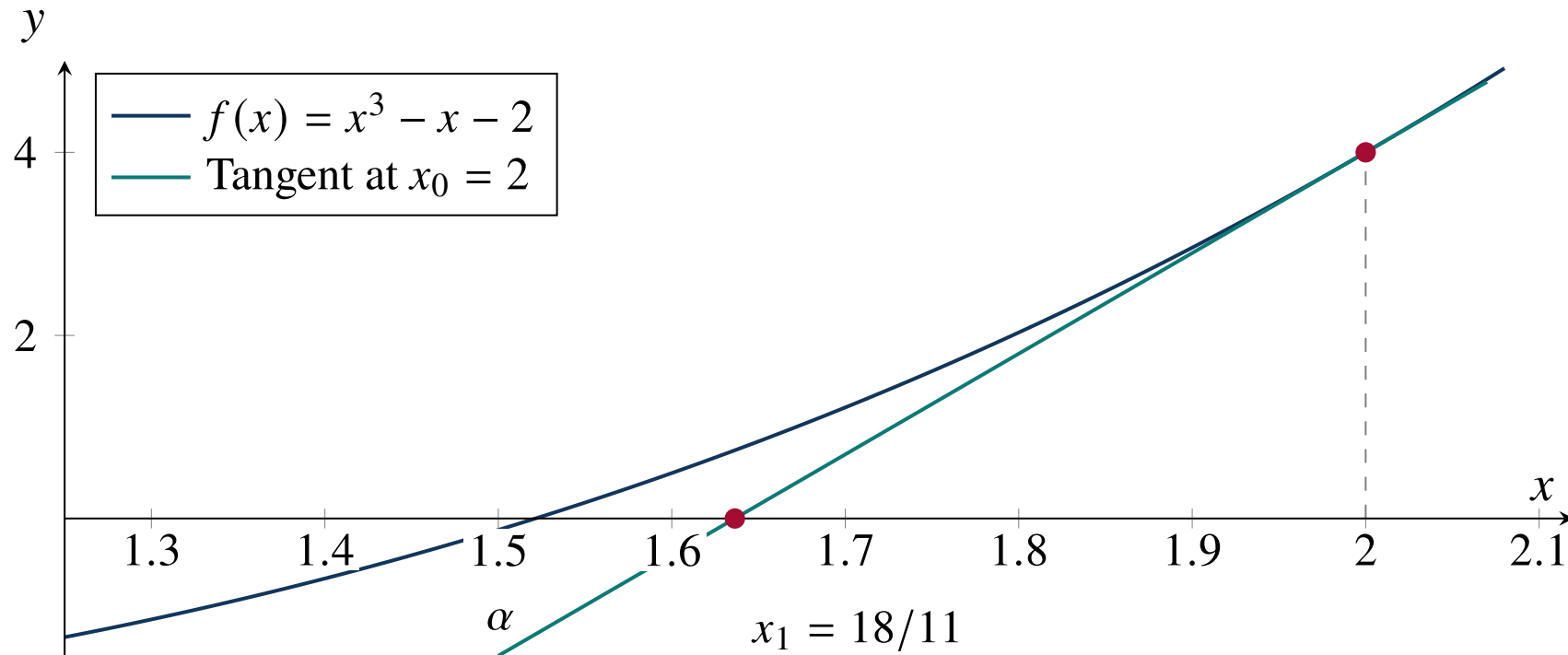
$$f(x_n + s) \approx f(x_n) + f'(x_n)s$$

Set the linear model equal to zero and solve for the correction.

$$f'(x_n)s_n = -f(x_n), \quad x_{n+1} = x_n + s_n$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad f'(x_n) \neq 0$$

The tangent intercept supplies the next iterate



$$f(x) = x^3 - x - 2, \quad x_0 = 2, \quad x_1 = 2 - \frac{4}{11} = \frac{18}{11}$$

Two Newton corrections for the cubic

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

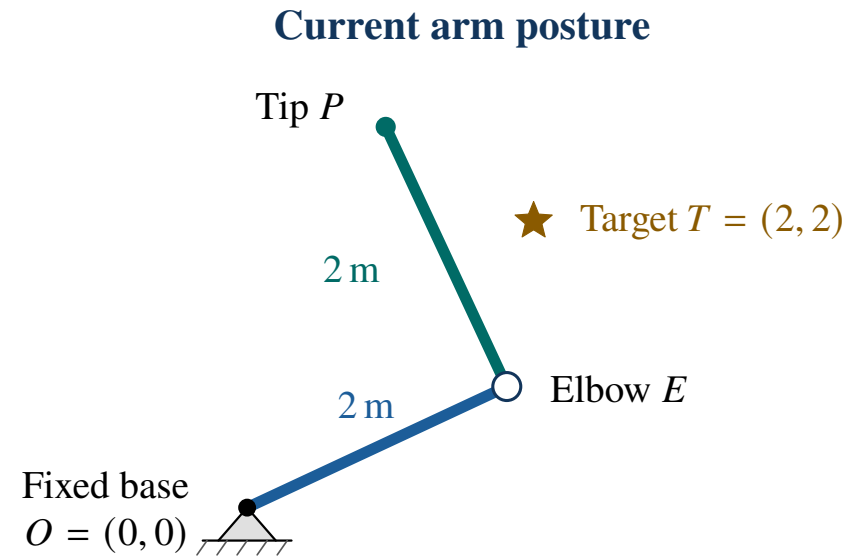
Restart from $x_0 = 1.5$. Then $f(x_0) = -0.125$ and $f'(x_0) = 5.75$.

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

$$x_2 = x_1 - \frac{x_1^3 - x_1 - 2}{3x_1^2 - 1} \approx 1.5213798059647862$$

The error in the approximate root falls from about 3.59×10^{-4} at x_1 to 9.92×10^{-8} at x_2 .

A robot arm reaching a target



The diagram shows the current arm. Move its tip to the target by rotating the two rigid 2 m links.

$$O = (0, 0), \quad T = (2, 2), \quad \text{required elbow: } E = (x, y)$$

A rigid link gives a distance equation

At the desired final position, the tip is at the target. Both links must be 2 m long. Coordinates are in metres.

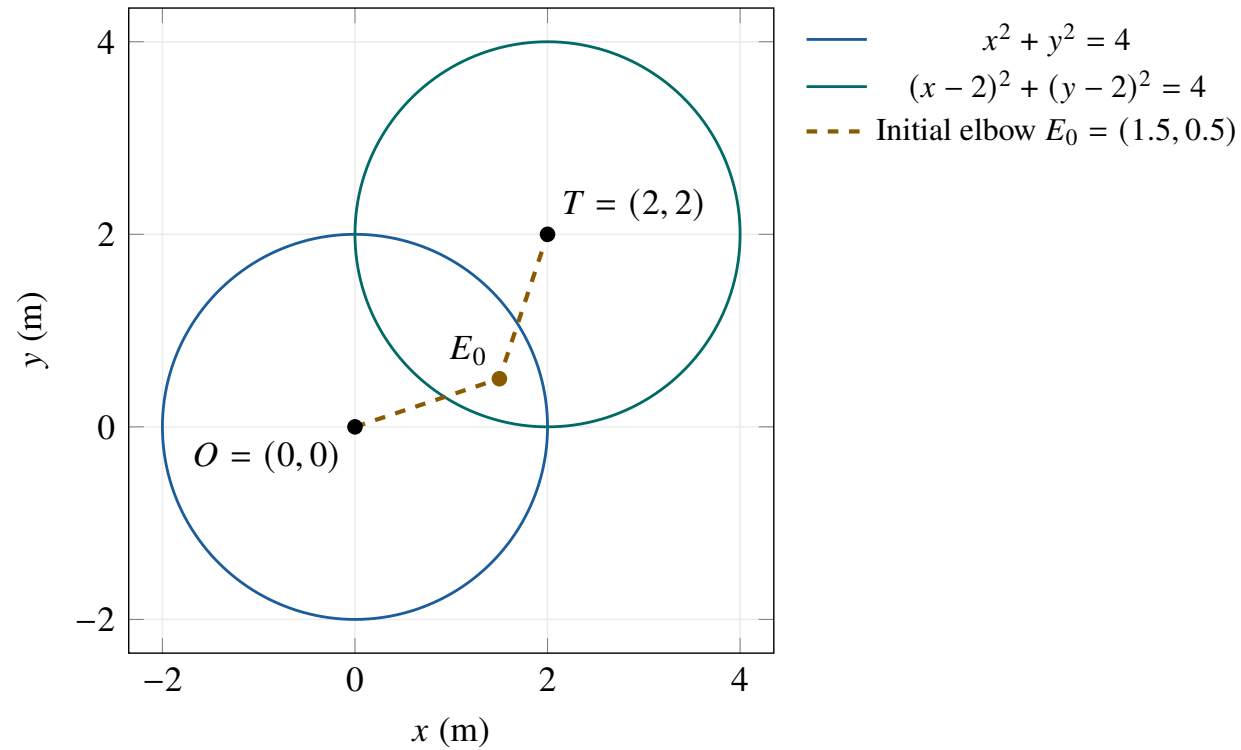
$$\text{Base to elbow: } \sqrt{x^2 + y^2} = 2 \implies x^2 + y^2 = 4$$

$$\text{Elbow to target: } \sqrt{(x - 2)^2 + (y - 2)^2} = 2$$

$$\implies (x - 2)^2 + (y - 2)^2 = 4$$

The elbow position must satisfy both equations at the same time.

Two distance constraints become two circles



General form of a non-linear system

$$f_1(\mathbf{x}) = 0, \quad f_2(\mathbf{x}) = 0, \quad \dots \quad f_n(\mathbf{x}) = 0$$

Collect them all into a vector:

$$F(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Problem: find $\mathbf{x}$ such that all equations are simultaneously satisfied. Such a point $\mathbf{x}_*$ is called a *root* of the system

The Jacobian quantifies changes in the equations

For a single equation with one variable, we took the derivative. Now we find the **Jacobian matrix**.

Rows correspond to the two equations. Columns correspond to the two coordinates.

$$f_1(x, y), \quad f_2(x, y) \quad \Rightarrow \quad J(x, y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ 2(x-2) & 2(y-2) \end{pmatrix}$$

$$J(x, y) \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 2xu + 2yv \\ 2(x-2)u + 2(y-2)v \end{pmatrix}$$

These are the predicted first-order changes in the two squared-length errors.

Iterations of Newton's method for a nonlinear system

At the k -th iteration:

$$\mathbf{x}_k = \begin{pmatrix} x_k \\ y_k \end{pmatrix}, \quad \mathbf{s}_k = \begin{pmatrix} u_k \\ v_k \end{pmatrix}$$

$$F(\mathbf{x}_k + \mathbf{s}) \approx F(\mathbf{x}_k) + J(\mathbf{x}_k)\mathbf{s} \quad \text{(First-order multi-variate Taylor expansion)}$$

Set the predicted new errors equal to zero:

$$J(\mathbf{x}_k)\mathbf{s}_k = -F(\mathbf{x}_k) \Rightarrow \text{linear system!}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{s}_k$$

Solve for the correction, add it, then check the original length equations.

For n equations in n unknowns, F has n components and J is an $n \times n$ matrix.

Newton's method for a system of equations

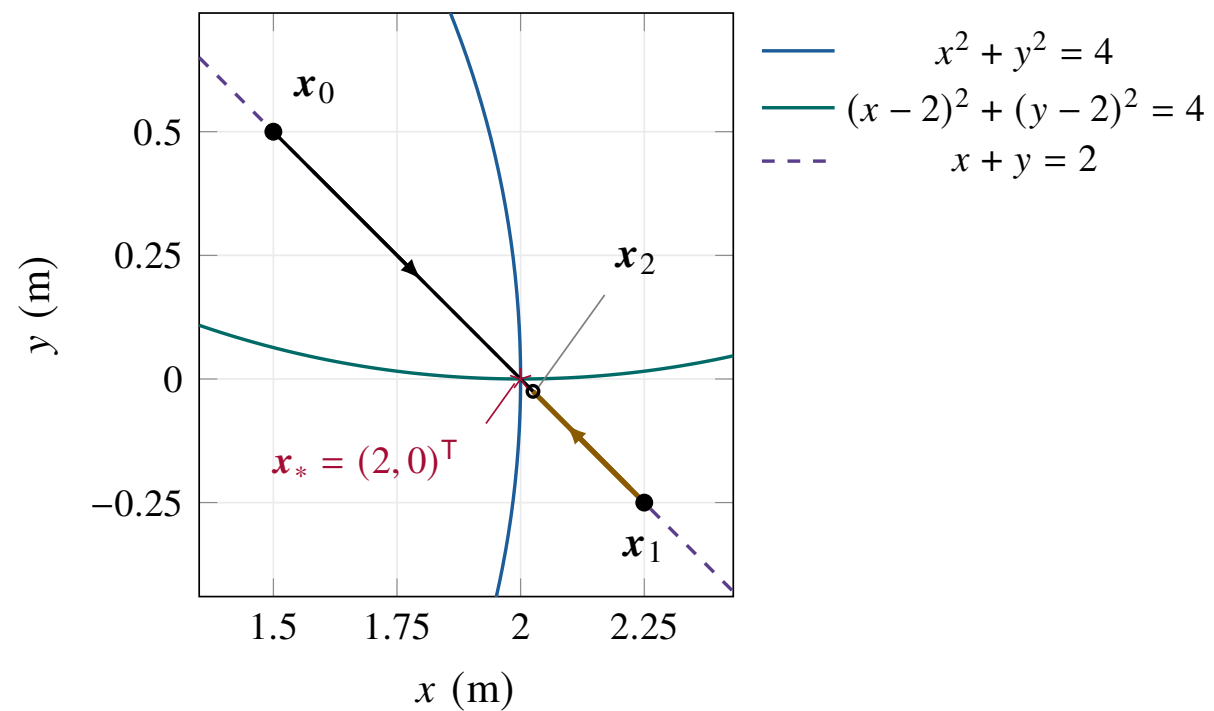
1. Evaluate both equations at the current trial point $\mathbf{x}_k$.
2. Form the Jacobian there and solve for the correction: $J(\mathbf{x}_k)s_k = -F(\mathbf{x}_k)$
3. Add the correction, then check the equations again: $\mathbf{x}_{k+1} = \mathbf{x}_k + s_k \Rightarrow F(\mathbf{x}_{k+1})$
4. Repeat until tolerance and/or iteration limit is reached.

The trial lengths approach the required 2 m

k	x_k (m)	y_k (m)	Both trial link lengths (m)
0	1.500000	0.500000	1.5811
1	2.250000	-0.250000	2.2638
2	2.025000	-0.025000	2.0252
3	2.000305	-0.000305	2.0003
4	2.000000	-4.65×10^{-8}	2.0000

All trial points satisfy $x_k + y_k = 2$, so both trial lengths agree. They approach 2 m; entries are rounded.

The numerical path of the trial elbow



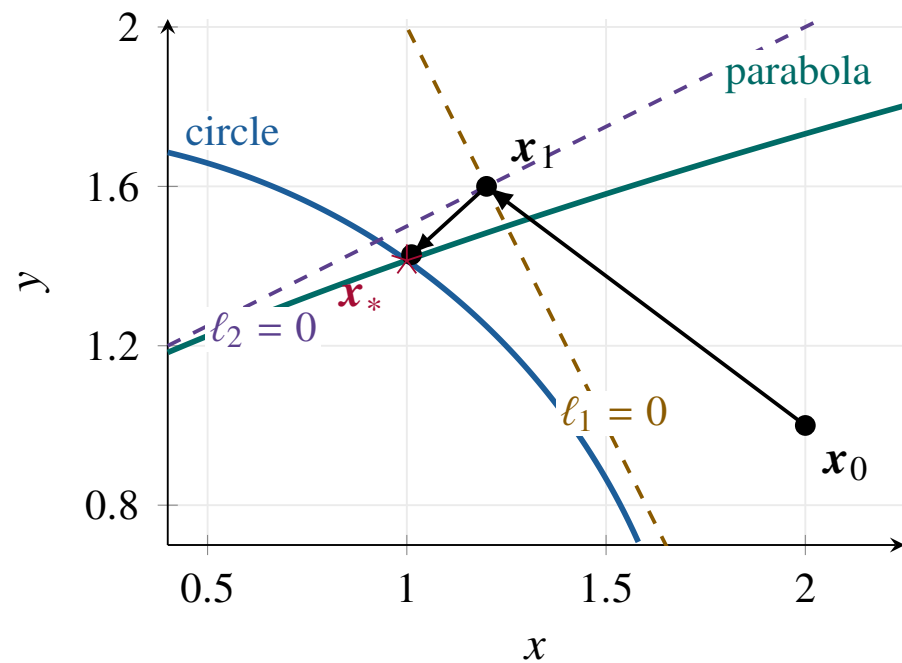
A different system, the same recipe

$$F(x, y) = \begin{pmatrix} x^2 + y^2 - 3 \\ x - y^2 + 1 \end{pmatrix}, \quad J(x, y) = \begin{pmatrix} 2x & 2y \\ 1 & -2y \end{pmatrix}$$

$$\mathbf{x}_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 4 & 2 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$$

$$\mathbf{s}_0 = \begin{pmatrix} -4/5 \\ 3/5 \end{pmatrix}, \quad \mathbf{x}_1 = \begin{pmatrix} 6/5 \\ 8/5 \end{pmatrix}$$

The linearized equations meet at the next point



Compare residual and coordinate error

$$\mathbf{x}_* = (1, \sqrt{2})^\top, \quad e_k = \|\mathbf{x}_k - \mathbf{x}_*\|_2$$

k	(x_k, y_k)	$\ F(\mathbf{x}_k)\ _2$	e_k
0	(2,1)	2.8284	1.0824
1	(1.2,1.6)	1.0628	2.7298×10^{-1}
2	(1.011765,1.428676)	7.1123×10^{-2}	1.8644×10^{-2}
3	(1.000046,1.414303)	4.0122×10^{-4}	1.0028×10^{-4}
4	(1.000000,1.414214)	1.2826×10^{-8}	3.1402×10^{-9}

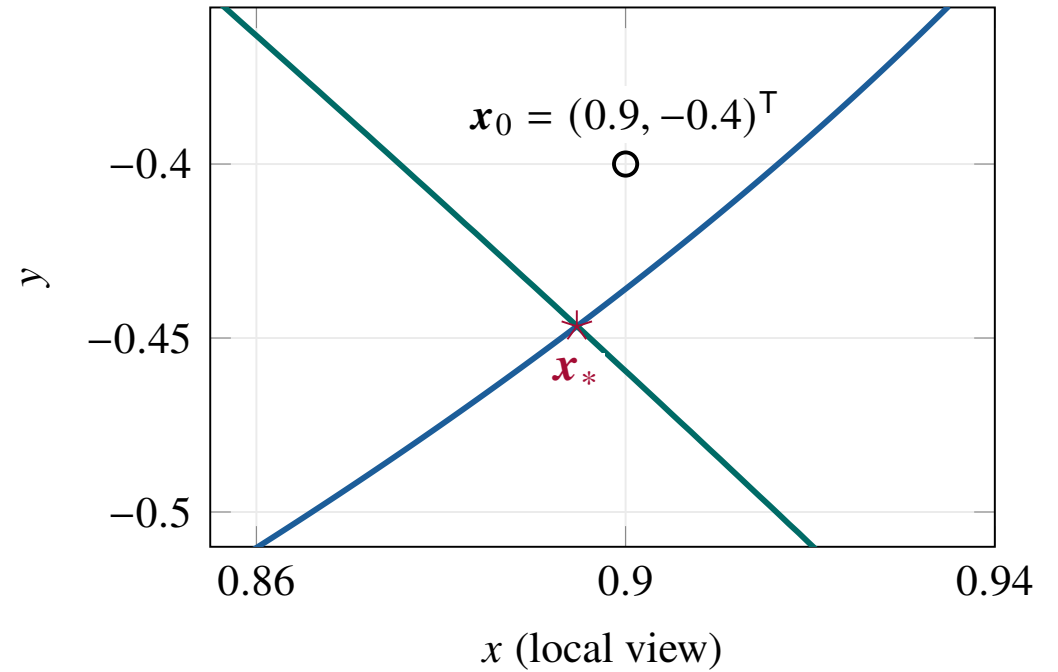
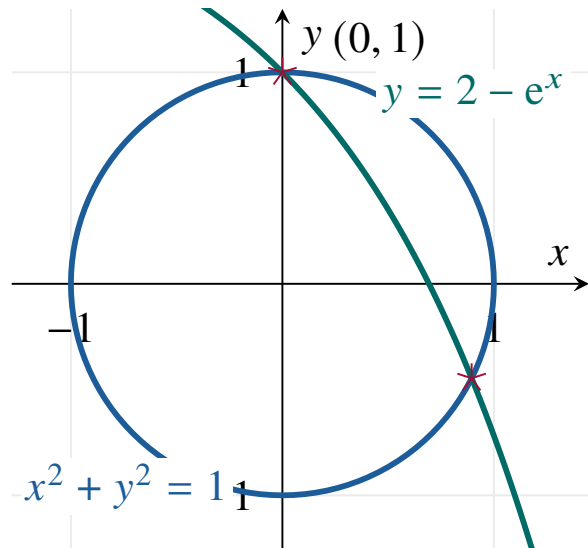
Newton's methods can also handle exponentials

$$F(x, y) = \begin{pmatrix} e^x + y - 2 \\ x^2 + y^2 - 1 \end{pmatrix}, \quad J(x, y) = \begin{pmatrix} e^x & 1 \\ 2x & 2y \end{pmatrix}$$

$$\begin{pmatrix} 2.4596031112 & 1 \\ 1.8 & -0.8 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \approx \begin{pmatrix} -0.0596031112 \\ 0.03 \end{pmatrix}$$

$$\mathbf{x}_0 = \begin{pmatrix} 0.9 \\ -0.4 \end{pmatrix} \longrightarrow \mathbf{x}_1 \approx \begin{pmatrix} 0.8953067996 \\ -0.4480597009 \end{pmatrix}$$

One initial guess selects one intersection



Recompute, correct, and check

k	x_k	y_k	$\ F(\mathbf{x}_k)\ _2$
0	0.9000000000	-0.4000000000	6.6727×10^{-2}
1	0.8953067996	-0.4480597009	2.3319×10^{-3}
2	0.8947154930	-0.4466391766	2.4059×10^{-6}
3	0.8947148013	-0.4466379119	2.1588×10^{-12}

Derivative information is enough for Newton; the equations need not be polynomial.

Why the later steps are so effective

Near a simple root, sufficiently smooth equations and a sufficiently close start give quadratic convergence.

$$e_k = \|\mathbf{x}_k - \mathbf{x}_*\|_2$$

$$\text{Linear: } e_{k+1} \leq qe_k, \quad 0 < q < 1,$$

$$\text{Quadratic: } e_{k+1} \leq Ce_k^2.$$

In the quadratic regime, the number of correct digits roughly doubles per step.

This is a local result; the initial guess and numerical precision still matter.

(A *simple root* $\mathbf{x}_*$ is when the Jacobian $J(\mathbf{x}_*)$ at that point is a non-singular matrix.)

A nonlinear system that simplifies exactly

A **Gröbner basis** can expose the solution through simpler polynomial equations. Here is one we can calculate by hand.

$$f_1 = x + y - 2 = 0, \quad f_2 = x^2 + xy - 2 = 0$$

$$\begin{aligned} f_2 - xf_1 &= (x^2 + xy - 2) - (x^2 + xy - 2x) \\ &= 2x - 2. \end{aligned}$$

$$x = 1, \quad y = 2 - x = 1$$

This is similar to row operations in linear systems!

A simple Gröbner basis

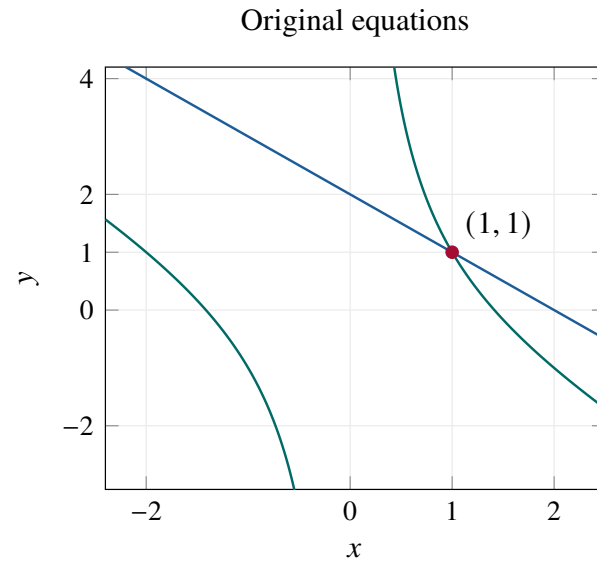
$$G = \{g_1, g_2\} = \{x - 1, y - 1\}$$

$$g_1 = \frac{1}{2}(f_2 - xf_1), \quad g_2 = f_1 - g_1$$

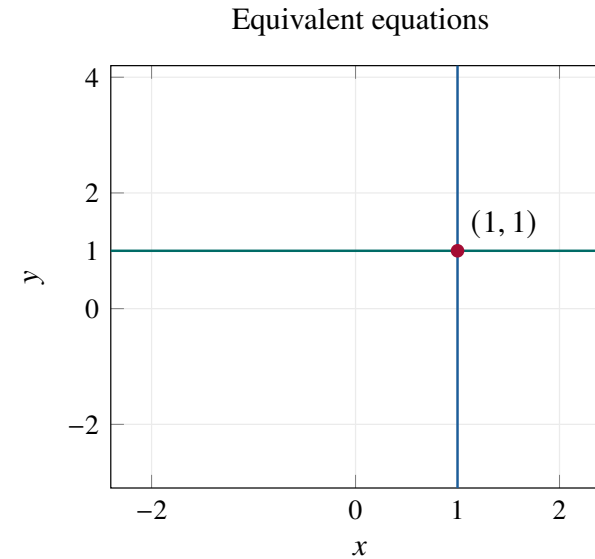
$$f_1 = g_1 + g_2, \quad f_2 = (x + 2)g_1 + xg_2$$

Both systems imply each other. With this basis, polynomial simplification is substitution at the single point (1,1).

The same solution in two descriptions



— $x + y = 2$
— $x^2 + xy = 2$



— $x - 1 = 0$
— $y - 1 = 0$

The original curves and the simpler lines share the single solution (1,1).

Eliminating one variable from symmetric sums

$$x + y + z = 6, \quad xy + xz + yz = 11, \quad xyz = 6$$

Using a series of eliminations, we arrive at

$$g_1 = x + y + z - 6 = 0$$

$$g_2 = y^2 + yz + z^2 - 6y - 6z + 11 = 0$$

$$g_3 = z^3 - 6z^2 + 11z - 6 = 0$$

This is like a triangular system: we can solve for z in the final equation, then back substitute!

A triangular system

$$g_1(x, y, z) = x + y + z - 6 = 0$$

Recover the remaining coordinate: $x = 6 - y - z$.

↑ substitute each compatible pair (y, z)

$$g_2(y, z) = y^2 + yz + z^2 - 6y - 6z + 11 = 0$$

For each chosen z , solve this equation for y .

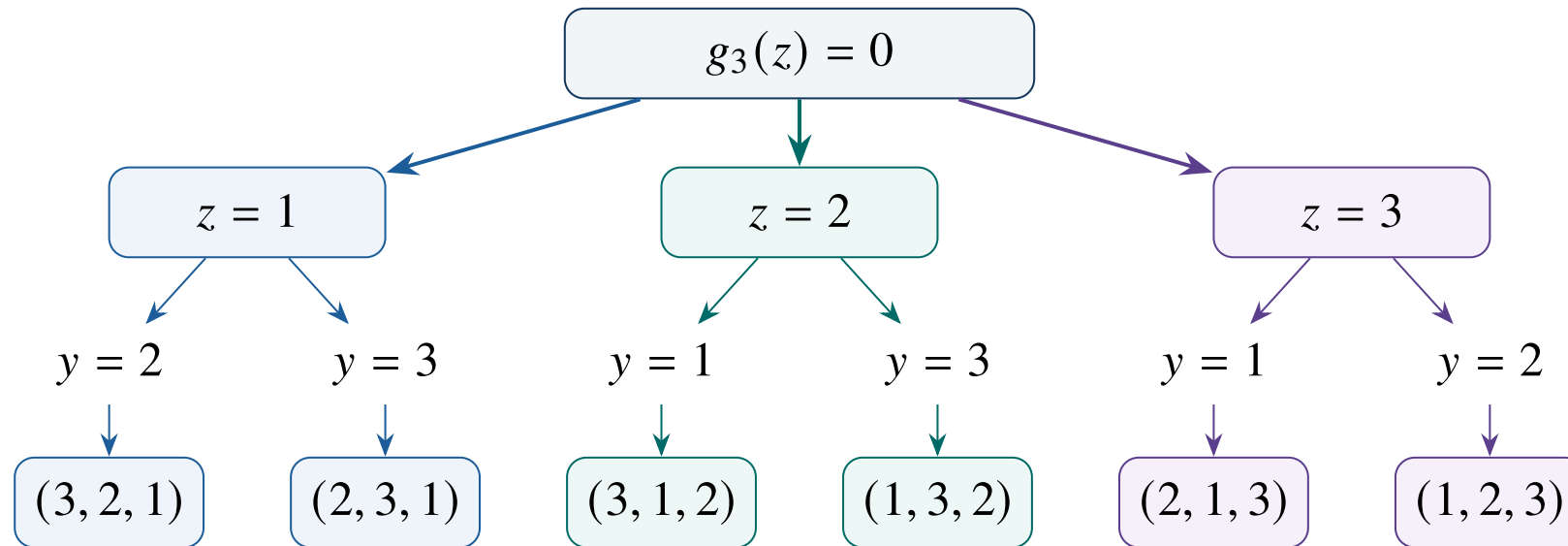
↑ substitute each possible z

$$g_3(z) = (z - 1)(z - 2)(z - 3) = 0$$

Start here: $z \in \{1, 2, 3\}$.

$$q = yz(6 - y - z) - 6, \quad q + zg_2 = z^3 - 6z^2 + 11z - 6$$

Back-substitution gives six solutions



Each leaf is an ordered triple (x, y, z) , with $x = 6 - y - z$.

$$z^3 - 6z^2 + 11z - 6 = (z - 1)(z - 2)(z - 3)$$