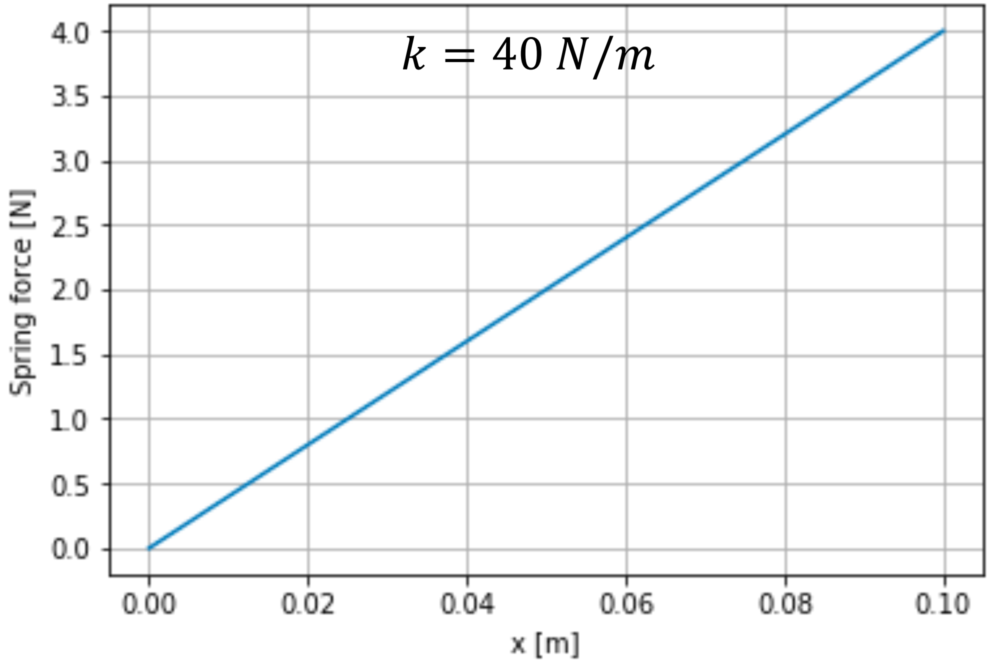
Nonlinear Equations

How can we solve these equations?

- Spring force:

$$F = k\,x$$

What is the displacement when $F = 2\text{N}$?

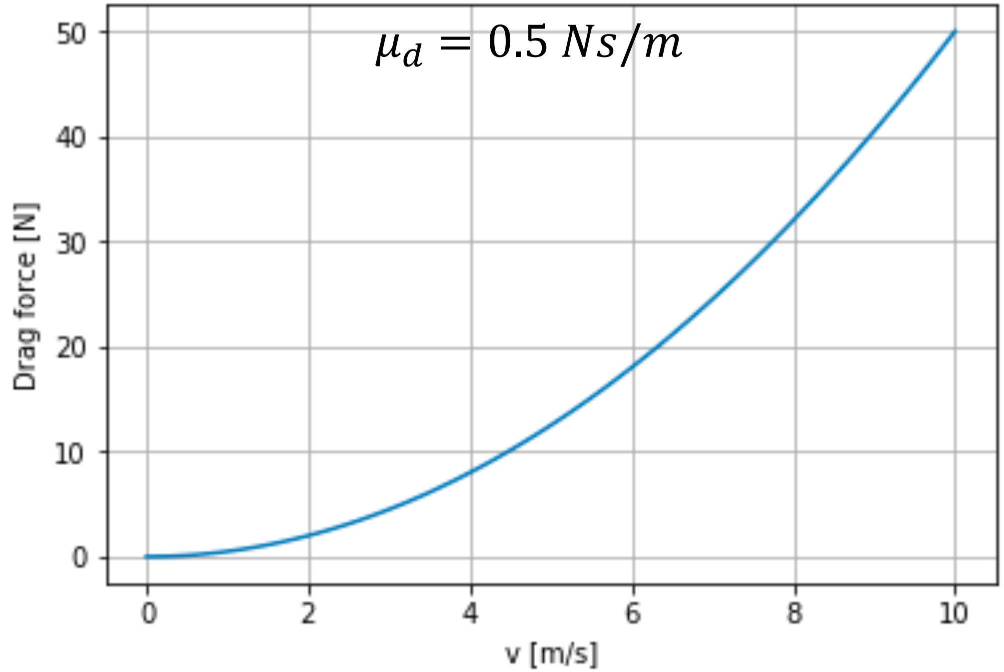


How can we solve these equations?

- Drag force:

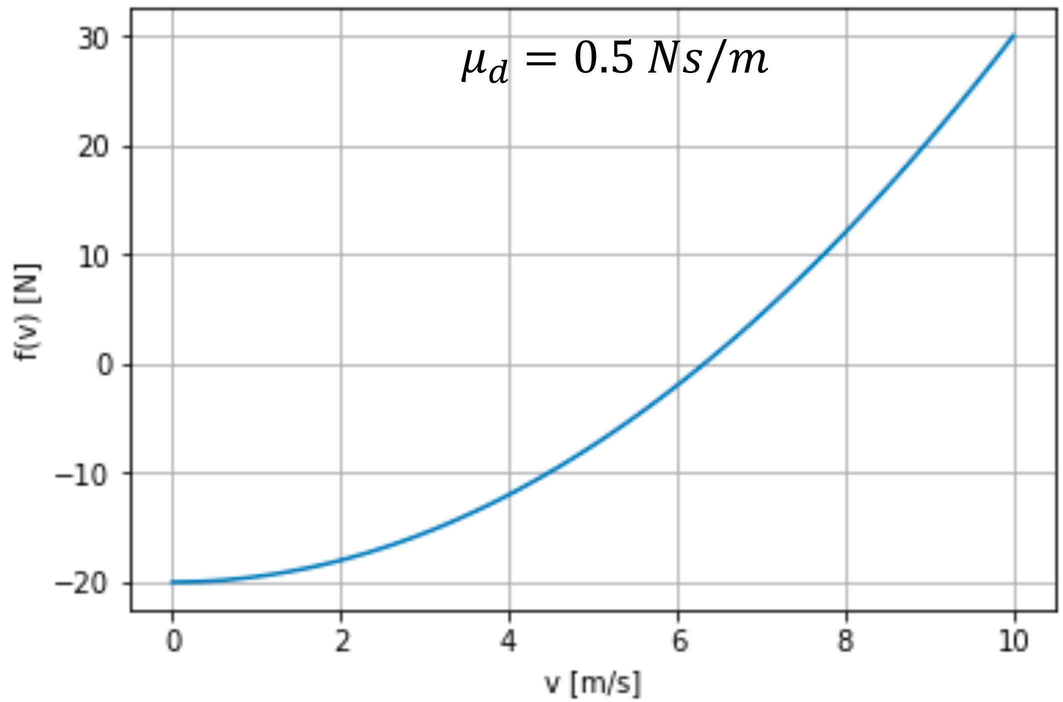
$$F = 0.5\, C_d\, \rho\, A\, v^2 = \mu_d\ v^2$$

What is the velocity when $F = 20\text{N}$?



$$f(v) = \mu_d \ v^2 - F = 0$$

Find the root (zero) of the nonlinear equation $f(v)$

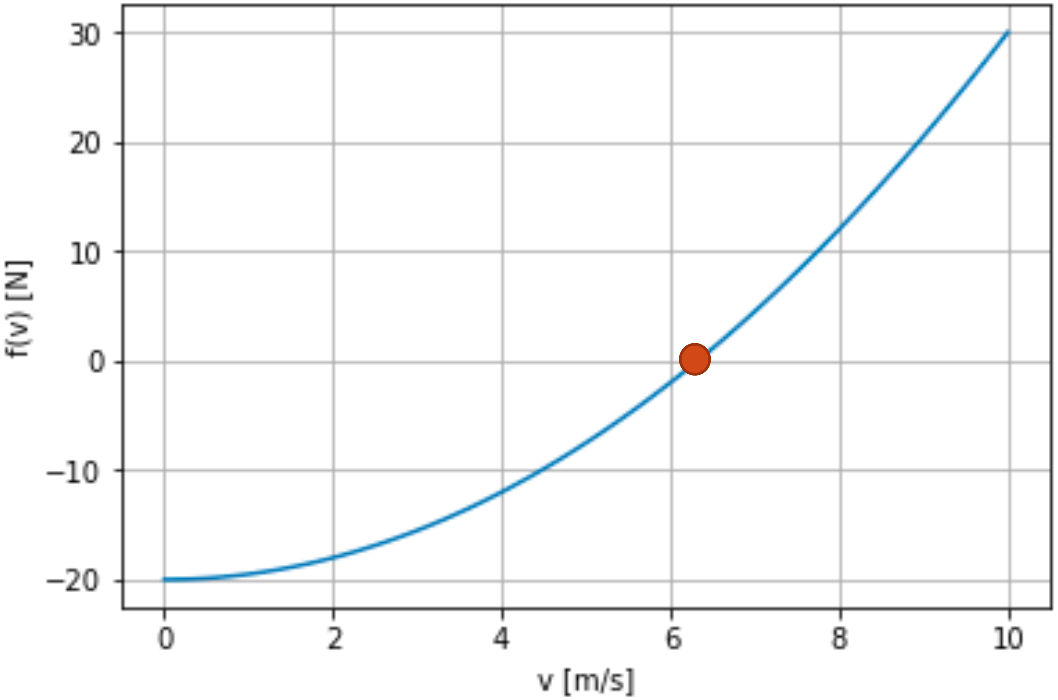


Nonlinear Equations in 1D

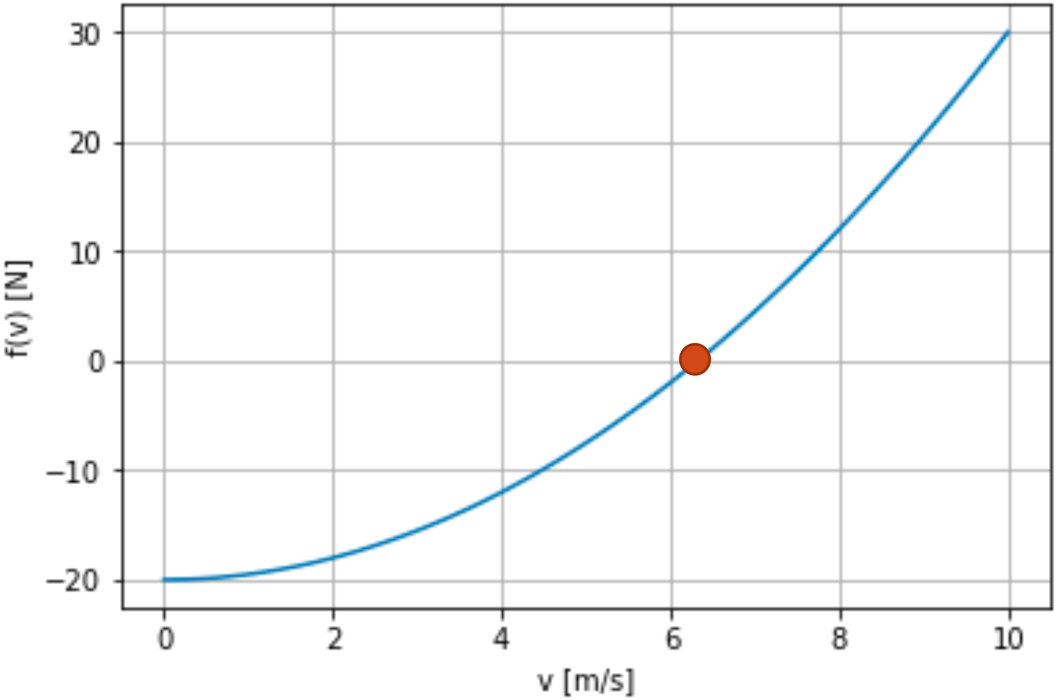
Goal: Solve $f(x) = 0$ for $f\colon \mathcal{R} \to \mathcal{R}$

Often called **Root Finding**

Bisection method



Bisection method



Convergence

An iterative method **converges with rate** r if:

$$\lim_{k\to\infty} \frac{||e_{k+1}||}{||e_k||^r} = C, \qquad 0 < C < \infty \qquad r = 1\text{: linear convergence}$$

Linear convergence gains a constant number of accurate digits each step (and $C < 1$ matters!)

For example: Power Iteration

Convergence

An iterative method **converges with rate** r if:

$$\lim_{k\to\infty} \frac{||e_{k+1}||}{||e_k||^r} = C, \qquad 0 < C < \infty$$

$r = 1$: linear convergence
$r > 1$: superlinear convergence
$r = 2$: quadratic convergence

Linear convergence gains a constant number of accurate digits each step (and $C < 1$ matters!)

Quadratic convergence doubles the number of accurate digits in each step (however it only starts making sense once $||e_k||$ is small (and C does not matter much)

Convergence

- The bisection method does not estimate x_k, the approximation of the desired root x. It instead finds an interval smaller than a given tolerance that contains the root.

Example:

Consider the nonlinear equation

$$f(x) = 0.5x^2 - 2$$

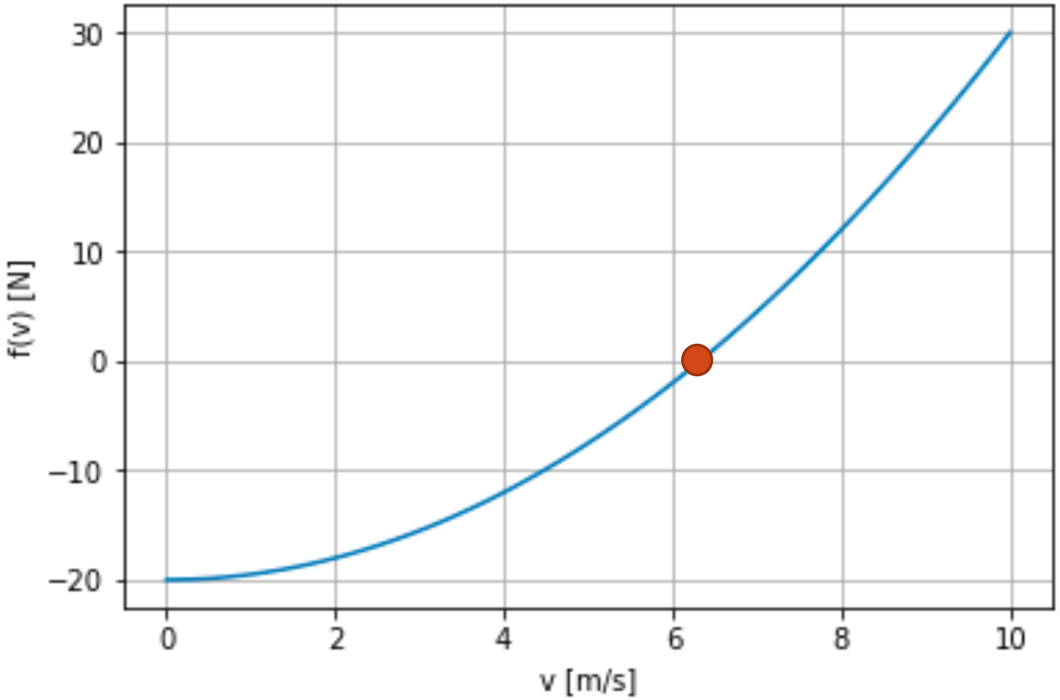
and solving $\mathrm{f(x)} = 0$ using the Bisection Method. For each of the initial intervals below, how many iterations are required to ensure the root is accurate within 2^{-4}?

A) $[-10, -1.8]$

B) $[-3, -2.1]$

C) $[-4, 1.9]$

Bisection method



Algorithm:

1. Take two points, a and b, on each side of the root such that $f(a)$ and $f(b)$ have opposite signs.

2. Calculate the midpoint $m = \frac{a+b}{2}$

3. Evaluate $f(m)$ and use m to replace either a or b, keeping the signs of the endpoints opposite.

Bisection Method - summary

- The function must be continuous with a root in the interval $[a, b]$
- Requires only one function evaluations for each iteration!
  - The first iteration requires two function evaluations.
- Given the initial internal $[a, b]$, the length of the interval after k iterations is $\frac{b-a}{2^k}$
- Has linear convergence

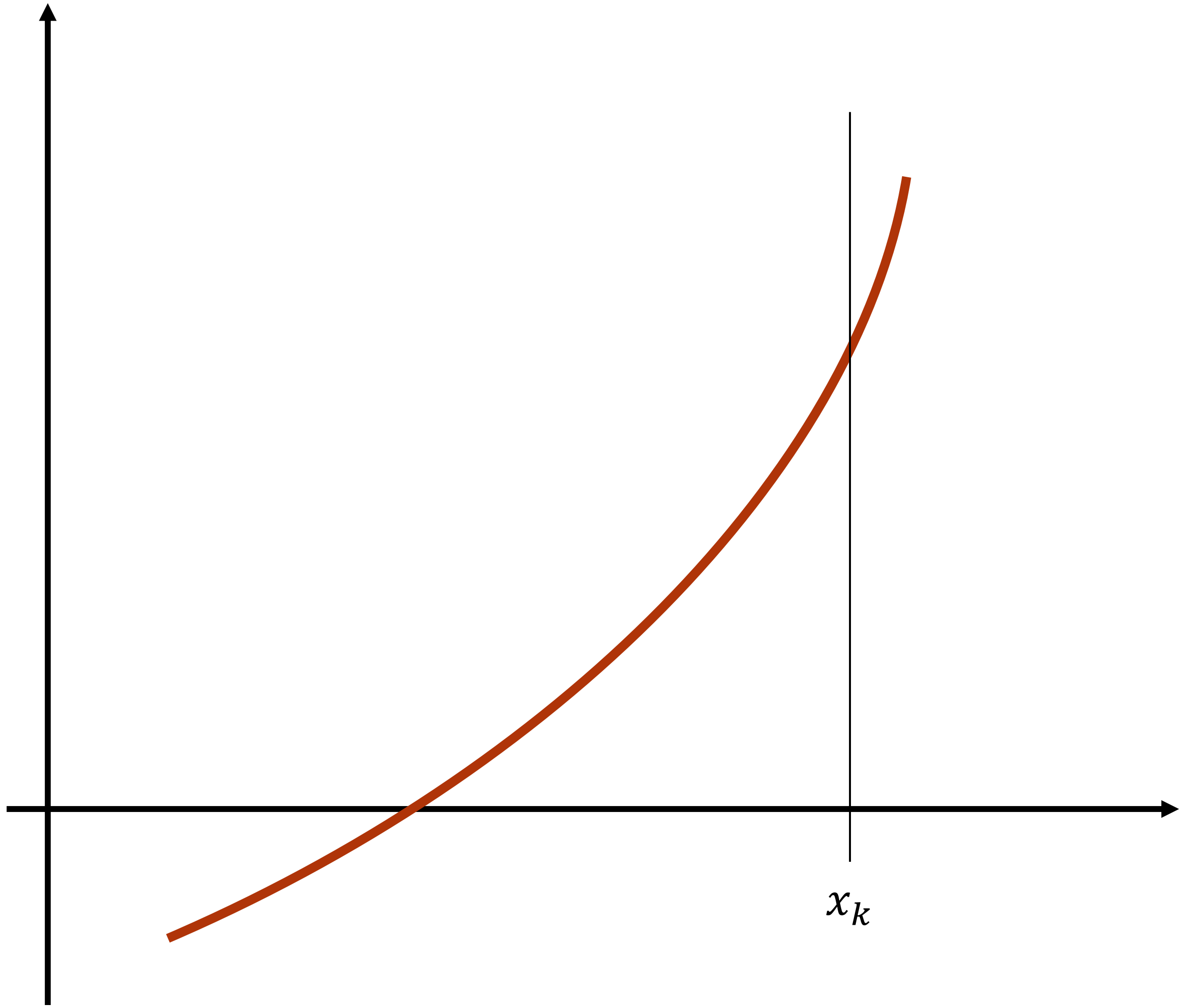
Newton's method

- Recall we want to solve $f(x) = 0$ for $f: \mathcal{R} \to \mathcal{R}$
- The Taylor expansion:

$$f(x_k + h) \approx f(x_k) + f'(x_k)h$$

gives a linear approximation for the nonlinear function f near x_k.

Newton's method



Example

Consider solving the nonlinear equation

$$5 = 2.0\, e^{x} + x^{2}$$

What is the result of applying **one iteration** of Newton's method for solving nonlinear equations with initial starting guess $x_0 = 0$, i.e. what is x_1?

A) -2
B) 0.75
C) -1.5
D) 1.5
E) 3.0

Newton's Method - summary

- Must be started with initial guess close enough to root (convergence is only local). Otherwise it may not converge at all.
- Requires function and first derivative evaluation at each iteration (think about two function evaluations)
- Typically has quadratic convergence

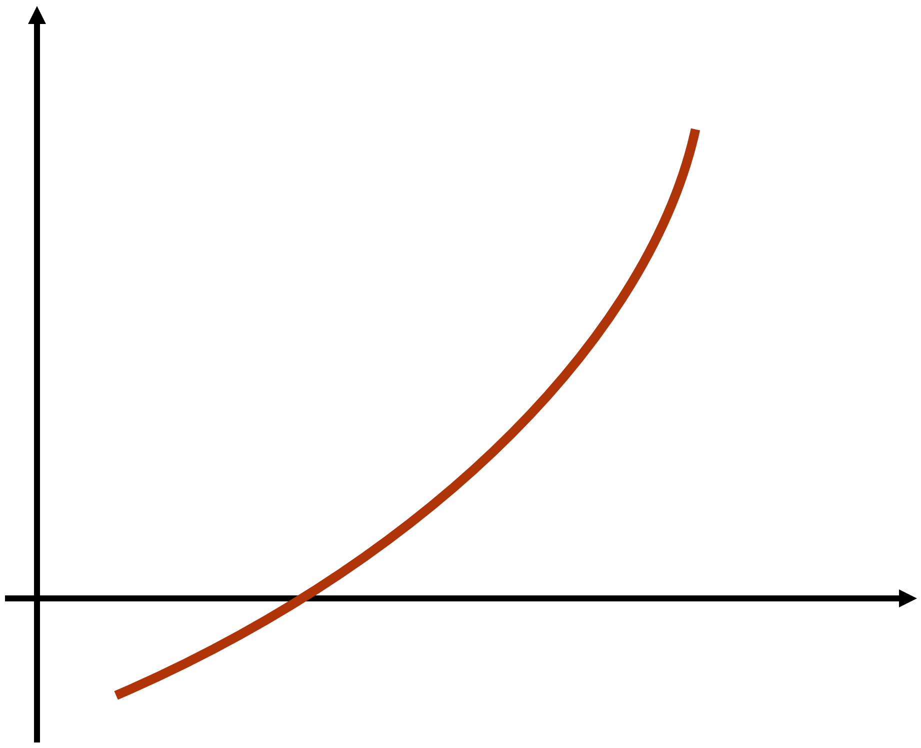
$$\lim_{k\to\infty} \frac{||e_{k+1}||}{||e_k||^2} = C, \qquad 0 < C < \infty$$

- What can we do when the derivative evaluation is too costly (or difficult to evaluate)?

Secant method

Also derived from Taylor expansion, but instead of using $f'(x_k)$, it approximates the tangent with the secant line:

$$x_{k+1} = x_k - f(x_k)/f'(x_k)$$



Secant Method - summary

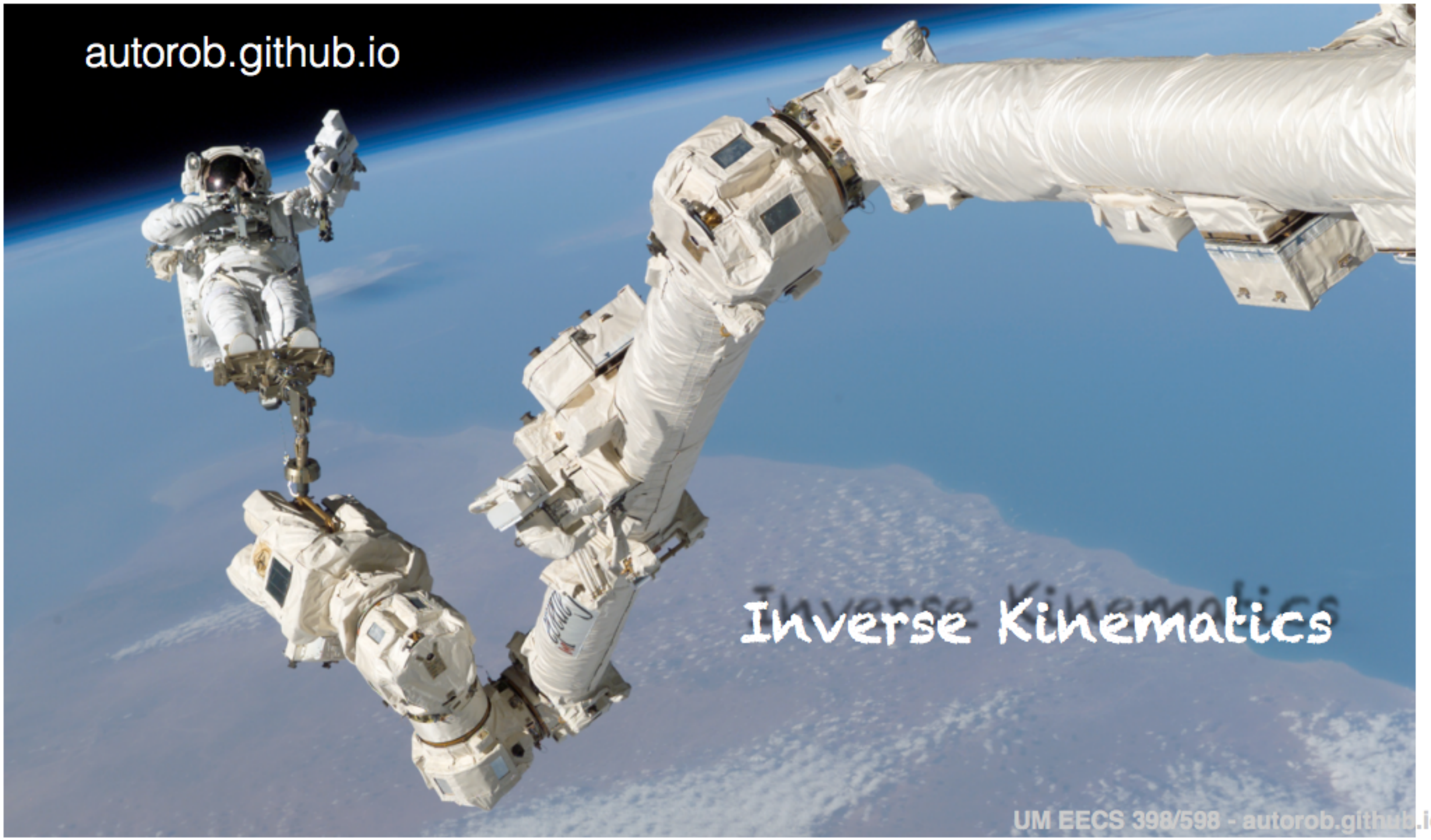
- Still local convergence
- Requires only one function evaluation per iteration (only the first iteration requires two function evaluations)
- Needs two starting guesses
- Has slower convergence than Newton's Method – superlinear convergence

$$\lim_{k\to\infty} \frac{||e_{k+1}||}{||e_k||^r} = C, \qquad 1 < r < 2$$

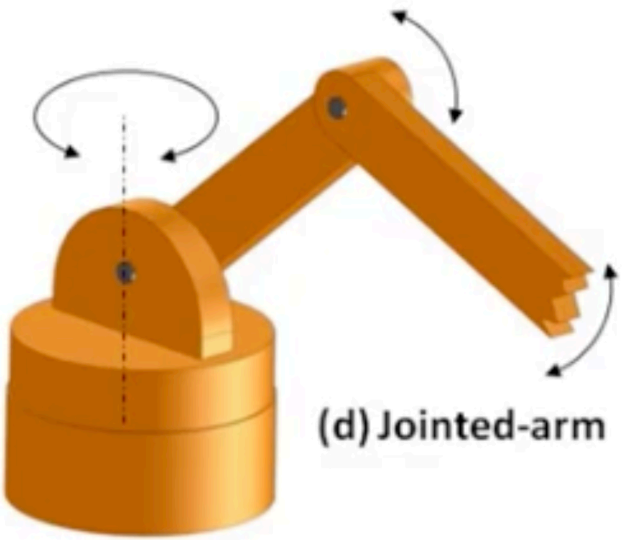
1D methods for root finding:

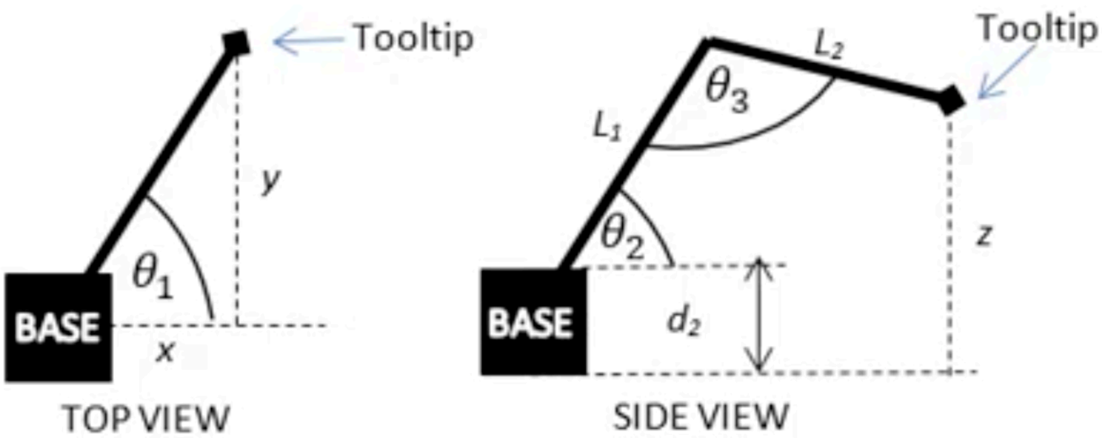
Method	Update	Convergence	Cost
Bisection	Check signs of $f(a)$ and $f(b)$ $$t_k = \frac{\lvert b-a\rvert}{2^k}$$	Linear ($r = 1$ and $c = 0.5$)	One function evaluation per iteration, no need to compute derivatives
Secant	$$x_{k+1} = x_k + h$$ $$h = -f(x_k)/dfa$$ $$dfa = \frac{f(x_k) - f(x_{k-1})}{(x_k - x_{k-1})}$$	Superlinear ($r = 1.618$), local convergence properties, convergence depends on the initial guess	One function evaluation per iteration (two evaluations for the initial guesses only), no need to compute derivatives
Newton	$$x_{k+1} = x_k + h$$ $$h = -f(x_k)/f'(x_k)$$	Quadratic ($r = 2$), local convergence properties, convergence depends on the initial guess	Two function evaluations per iteration, requires first order derivatives

Nonlinear system of equations



Robotic arms





https://www.youtube.com/watch?v=NRgNDlVtmz0 (Robotic arm 1)
https://www.youtube.com/watch?v=9DqRkLQ5Sv8 (Robotic arm 2)
https://www.youtube.com/watch?v=DZ_ocmY8xEI (Blender)

Inverse Kinematics

Nonlinear system of equations

Goal: Solve $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ for $\boldsymbol{f}: \mathcal{R}^n \rightarrow \mathcal{R}^n$

Newton's method

Approximate the nonlinear function $\boldsymbol{f(x)}$ by a linear function using Taylor expansion:

Newton's method

Algorithm:

Convergence:

- Typically has quadratic convergence
- Drawback: Still only locally convergent

Cost:

- Main cost associated with computing the Jacobian matrix and solving the Newton step.

Example

Consider solving the nonlinear system of equations

$$2 = 2y + x$$
$$4 = x^2 + 4y^2$$

What is the result of applying one iteration of Newton's method with the following initial guess?

$$\boldsymbol{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Newton's method

$\boldsymbol{x}_0 = initial\ guess$

$For\ k = 1,2, \dots$

Evaluate $\mathbf{J} = \boldsymbol{J}(\boldsymbol{x}_k)$

Evaluate $\boldsymbol{f}(\boldsymbol{x}_k)$

Factorization of Jacobian (for example $\mathbf{LU} = \boldsymbol{J}$)

Solve using factorized J (for example $\mathbf{LU}\ \boldsymbol{s}_k = -\boldsymbol{f}(\boldsymbol{x}_k)$)

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

Newton's method - summary

- Typically quadratic convergence (local convergence)

- Computing the Jacobian matrix requires the equivalent of n^2 function evaluations for a dense problem (where every function of $\boldsymbol{f}(\boldsymbol{x})$ depends on every component of $\boldsymbol{x}$).

- Computation of the Jacobian may be cheaper if the matrix is sparse.

- The cost of calculating the step $\boldsymbol{s}$ is $O(n^3)$ for a dense Jacobian matrix (Factorization + Solve)

- If the same Jacobian matrix $\boldsymbol{J}(\boldsymbol{x}_k)$ is reused for several consecutive iterations, the convergence rate will suffer accordingly (trade-off between cost per iteration and number of iterations needed for convergence)

Inverse Kinematics