Integration, Root-Finding, Linear Systems

Prateek Sharma (prateek@physics.iisc.ernet.in)
Office: D2-08

Numerical Integration

recall that derivative of any function can be calculated analytically, not true of integral!

$$I = \int_a^b f(x)dx$$

can be cast as ODE: $\dfrac{dy}{dx} = f(x) \quad y(a) = 0$

solving for the value $I \equiv y(b)$

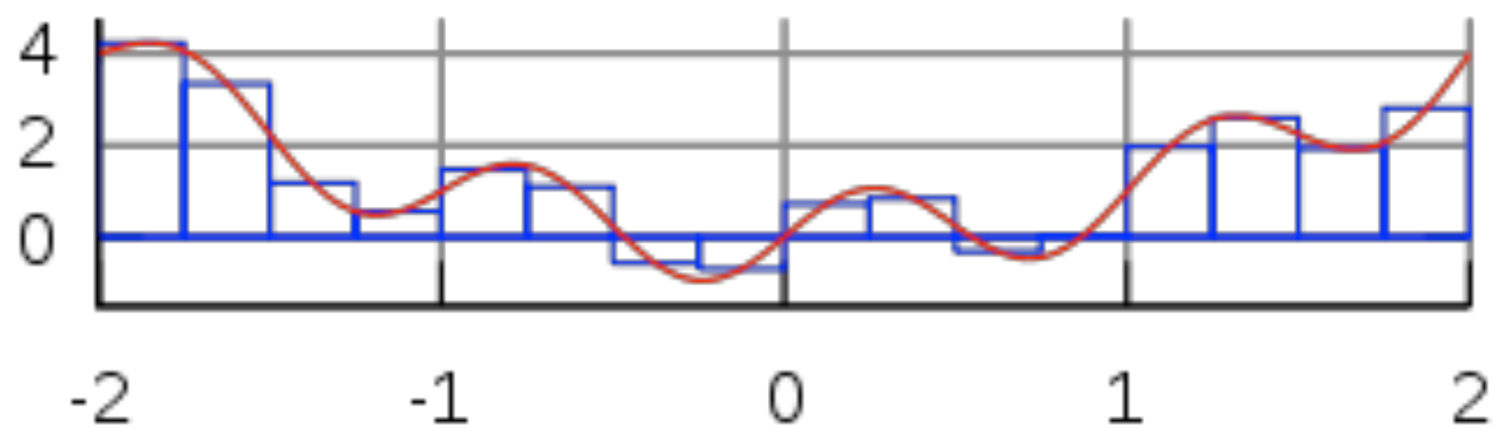
via function approximation: e.g., cubic spline interpolation

some integrals via Fast Fourier Transform (FFT)

Monte-Carlo integration for multidimensions

Simple Formulae

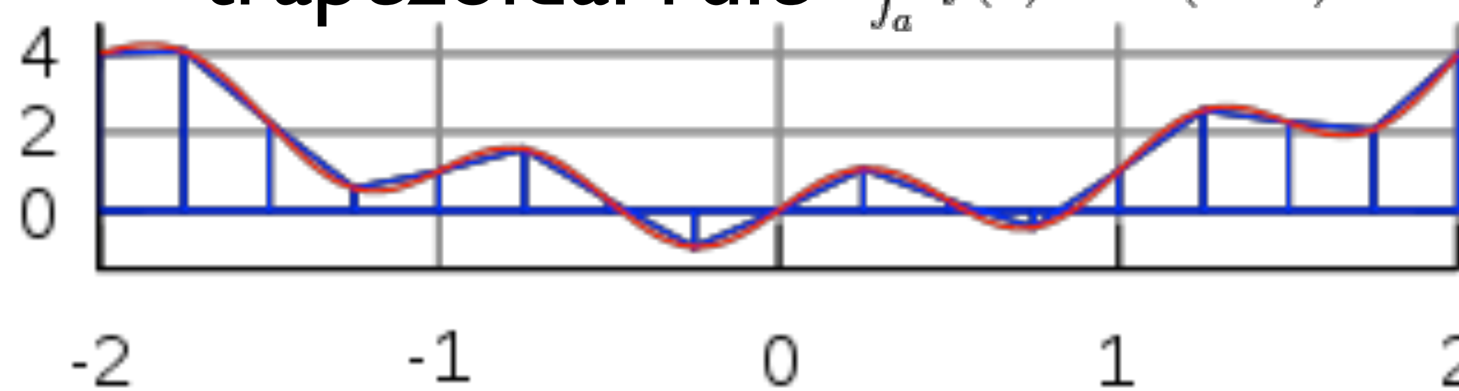
mid-point/rectangle rule



$$\int_a^b f(x)\,dx \approx (b-a)\, f\left(\frac{a+b}{2}\right).$$

$O(h^3)$

trapezoidal rule $\int_a^b f(x)\,dx \approx (b-a)\,\frac{f(a)+f(b)}{2}.$



possible to construct higher order w. higher deg. polynomial

can apply these rules at much smaller intervals and obtain accurate integrals

Aim: minimum function evaluations & highest accuracy

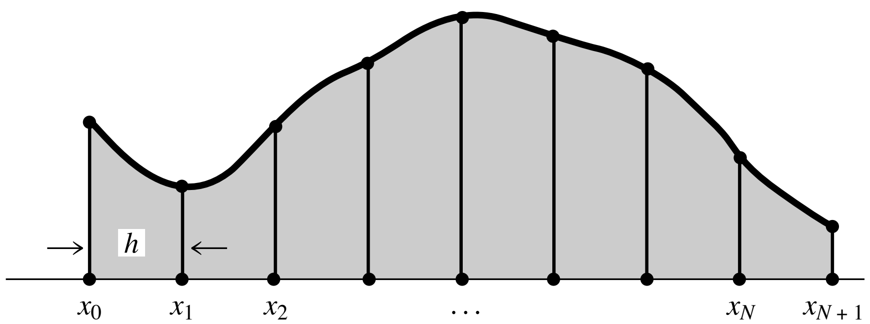
Quadrature

an approx. of definite integral as a wtd. sum of fn. values at specified points

equally spaced abscissa:

extended Trapezoidal rule, exact for linear polynomial

$$\int_{x_1}^{x_N} f(x)dx = h\left[\frac{1}{2}f_1 + f_2 + f_3 + \cdots + f_{N-1} + \frac{1}{2}f_N\right] + O\left(\frac{(b-a)^3 f''}{N^2}\right)$$



$$\int_{x_1}^{x_N} f(x)dx = h\left[\frac{1}{3}f_1 + \frac{4}{3}f_2 + \frac{2}{3}f_3 + \frac{4}{3}f_4 + \cdots + \frac{2}{3}f_{N-2} + \frac{4}{3}f_{N-1} + \frac{1}{3}f_N\right] + O\left(\frac{1}{N^4}\right)$$

extended Simpson's rule, exact for cubic

Richardson Extrapolation

both mid-point & trapezoidal rule are 2nd order accurate

similar formula for Trapezoidal

$$\int_a^b f(x)dx \quad = \quad h\sum_{i=1}^{n} f(m_i) + c_2h^2 + c_4h^4 + \cdots$$

$$h = \tfrac{b-a}{n}$$

remember Richardson extrapolation for smooth functions:

$$M(f) \quad = \quad N(f,h) + c_ph^p + c_qh^q + \cdots$$

$$M(f) \quad = \quad N(f,2h) + c_p(2h)^p + c_q(2h)^q + \cdots$$

$$c_ph^p \quad = \quad \frac{N(f,h) - N(f,2h)}{2^p - 1} - \tilde{c}_qh^q - \cdots$$ leading order error w. $$\tilde{c}_q \quad = \quad c_q\frac{2^q - 1}{2^p - 1}$$

reduces error to $O(h^q)$

Romberg Integration

$$R_{k,1} = R_{k,0} + \frac{R_{k,0} - R_{k-1,0}}{3} \qquad R_{k,2} = R_{k,1} + \frac{R_{k,1} - R_{k-1,1}}{15}$$

$$R_{k,m} = R_{k,m-1} + \frac{R_{k,m-1} - R_{k-1,m-1}}{2^{2m} - 1} \qquad \int_0^1 e^x dx$$

$R_{k,0}$ n=2^k	$R_{k,1}$	$R_{k,2}$	$R_{k,3}$	$R_{k,4}$
1.6487212707001282				
1.7005127166502081	1.7177765319669014			
1.7138152797710871	1.7182494674780466	1.7182809965121231		
1.7171636649956870	1.7182797934038869	1.7182818151322763	1.7182818281262471	
1.7180021920526605	1.7182817010716516	1.7182818282495025	1.7182818284577124	1.7182818284590122
1.7182119133838587	1.7182818204942580	1.7182818284557650	1.7182818284590391	1.7182818284590442
1.7182643493168632	1.7182818279611980	1.7182818284589940	1.7182818284590453	1.7182818284590453

→ cheap, more accurate

↓ expensive

Aitken extrapolation:
can determine leading order error by comparing N(f,h), N(f,2h) and N(f,4h)

Improper Integrals

integrable singularity at a known x, $\int_0^1 \frac{\sin x}{x}dx \quad \int_0^1 x^{-1/2}dx \quad \int_{-\infty}^{\infty}$

quadrature formula must not evaluate function at the integrable singularity!

$$\int_{x_1}^{x_N} f(x)dx = h[f_{3/2} + f_{5/2} + f_{7/2} + \cdots + f_{N-3/2} + f_{N-1/2}] \quad + O\left(\frac{1}{N^2}\right)$$

extended mid-point rule

apply w. Romberg method

$$\int_a^b f(x)dx = \int_{1/b}^{1/a} \frac{1}{t^2} f\left(\frac{1}{t}\right) dt \qquad ab > 0$$

change of variable w. finite range
having intuitive feel of \`f' very useful

Gaussian quadrature:

conveniently spaced abscissa & smooth enough function can give double the order of accuracy compared to e.g., Trapezoidal. We won't cover it.

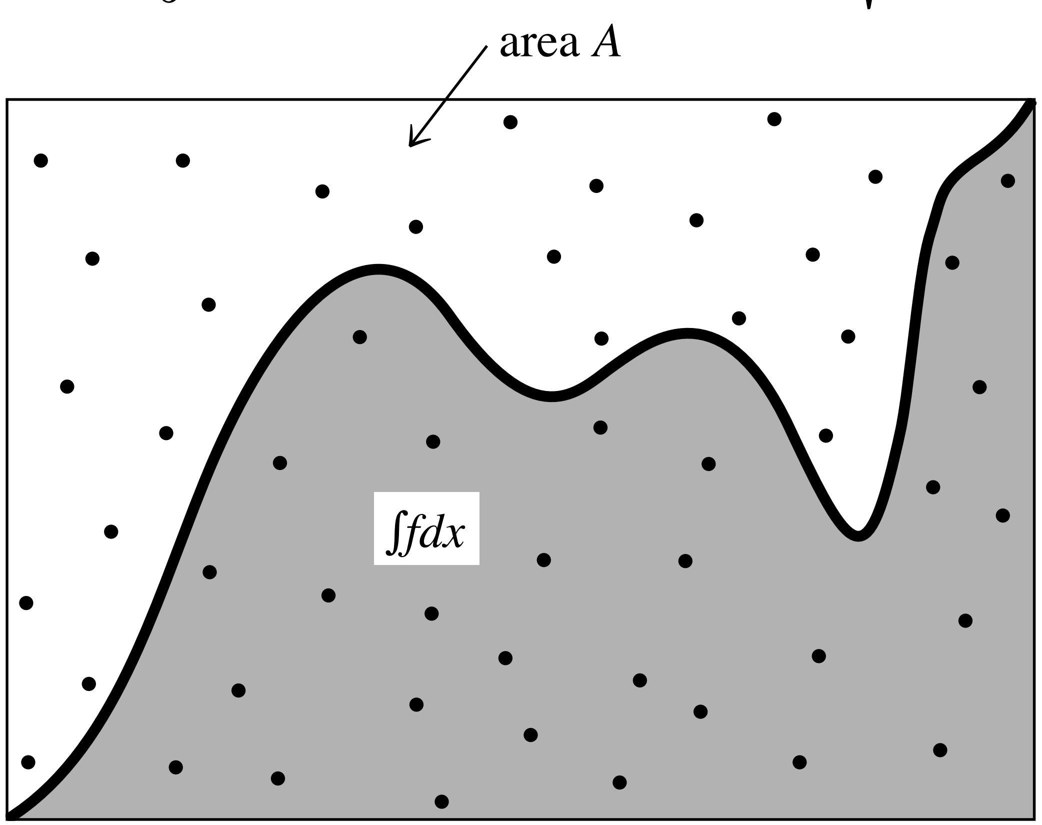
Multidimensional Integn.

expensive: N^3 in 3D, where N fn. evals. needed in 1D; complicated boundaries

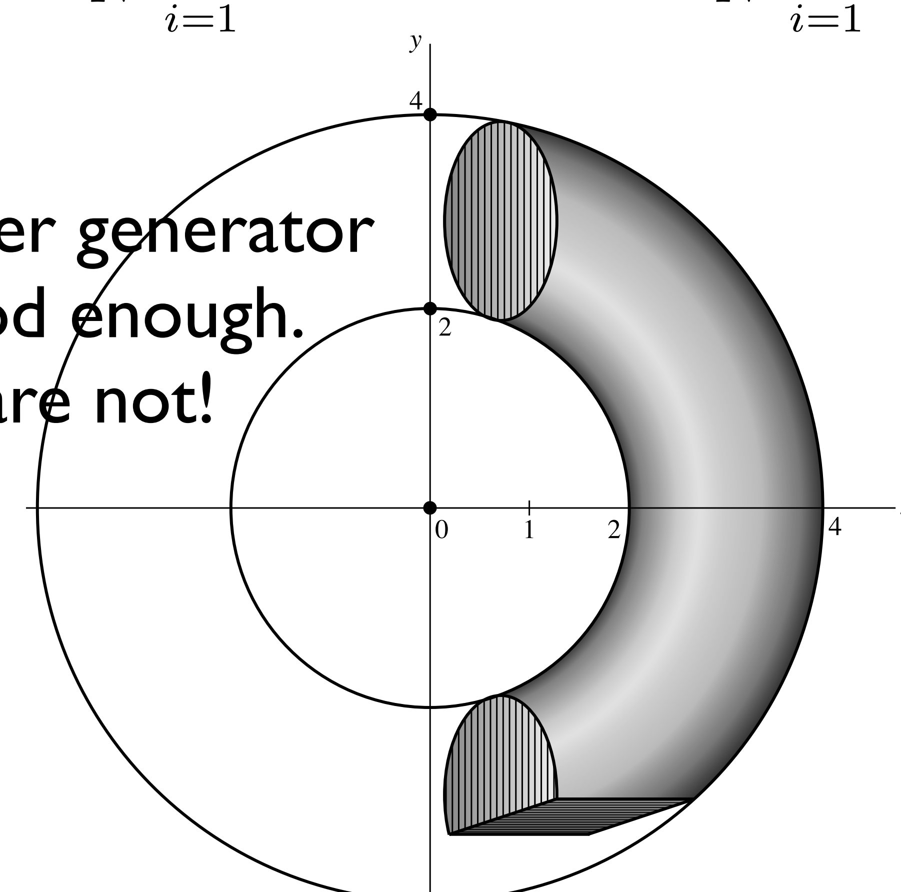
try to reduce dimension

if complicated boundary & integrand is not strongly peaked in isolated regions & low accuracy is fine => MonteCarlo integration

$$\int f\, dV \approx V \langle f \rangle \pm V \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{N}} \qquad \langle f \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} f(x_i) \qquad \langle f^2 \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} f^2(x_i)$$



random number generator must be good enough. defaults are not!



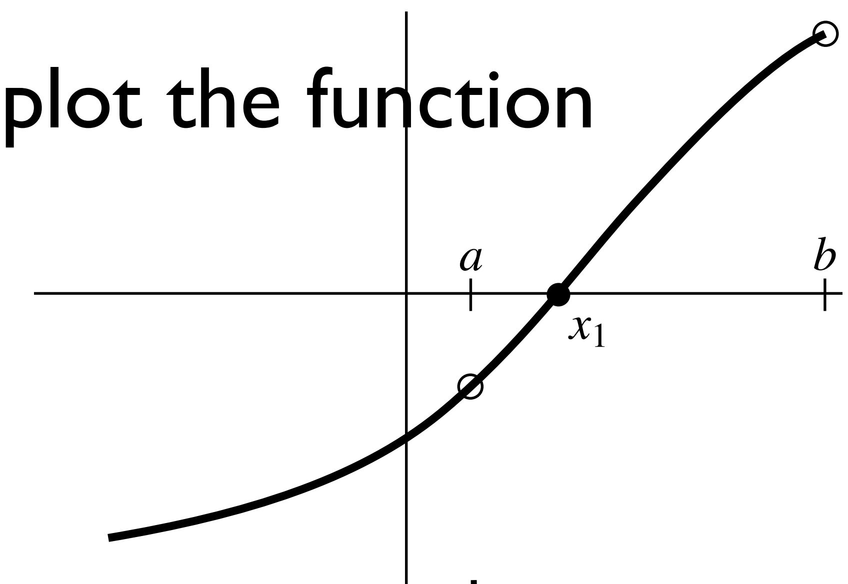
Root-Finding

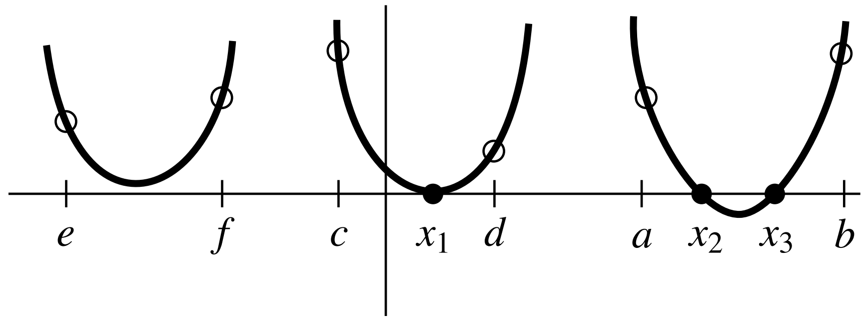
solving f(x)=0, multi-D in general

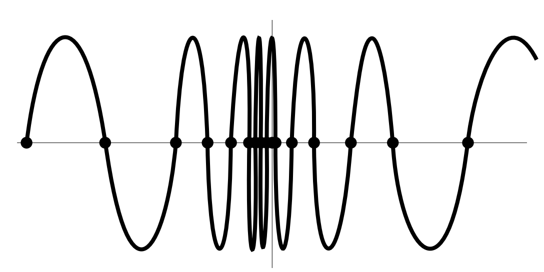
finding roots in multi-D very challenging, bracketing works in 1-D
except for linear systems root-finding is iterative (having a good initial guess crucial)
may not converge or worse converge to a wrong root

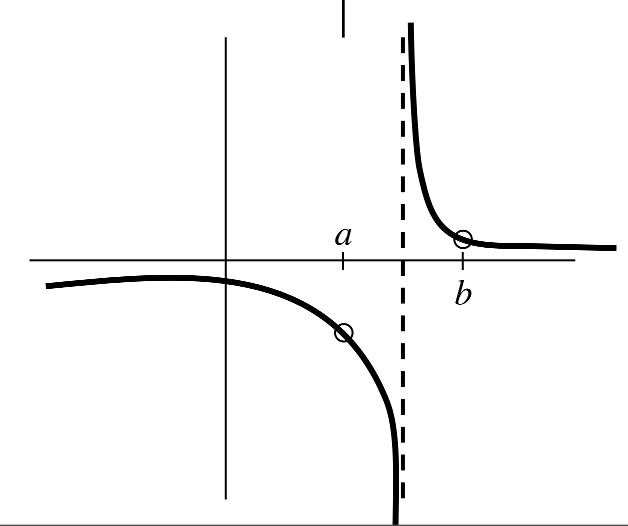
feel for what the fn. looks like! always best to plot the function

bracketing the root: function changes
sign in a given interval; $f_a f_b<0$
don’t let the guess go out of best bracket







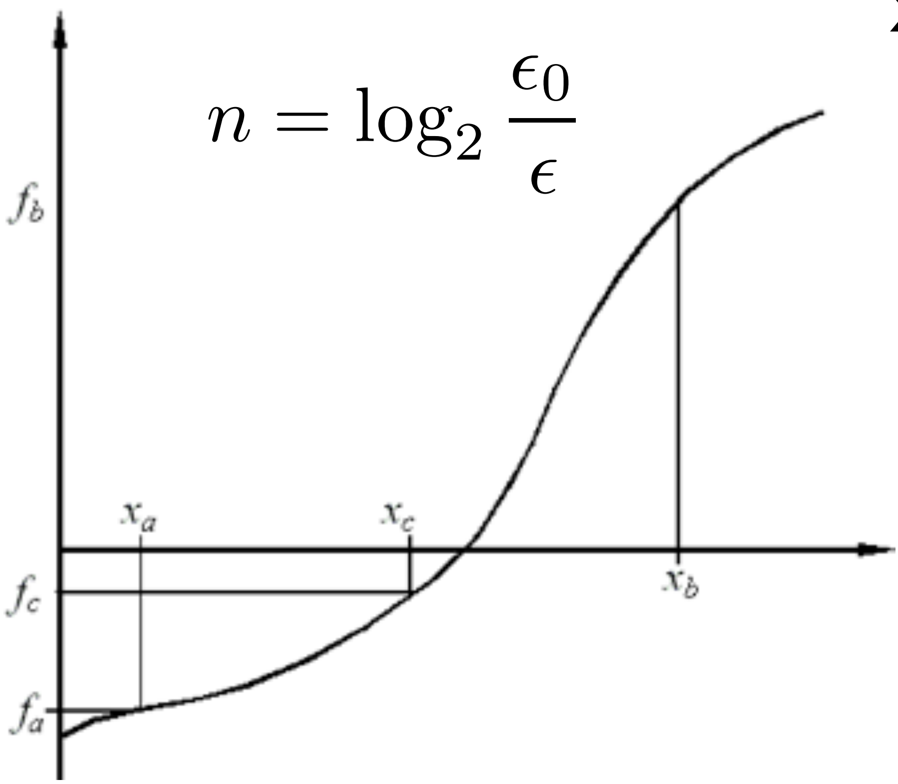


Bisection, Secant Methods

convergence criterion: fractional error in x_0: e.g., $|dx/x_0|<10^{-6}$

$$x^{n+1}=x^n-f^n(x^n-x^{n-1})/(f^n-f^{n-1})$$
$$=(x^{n-1}f^n-x^nf^{n-1})/(f^n-f^{n-1})$$

$$n = \log_2 \frac{\epsilon_0}{\epsilon}$$

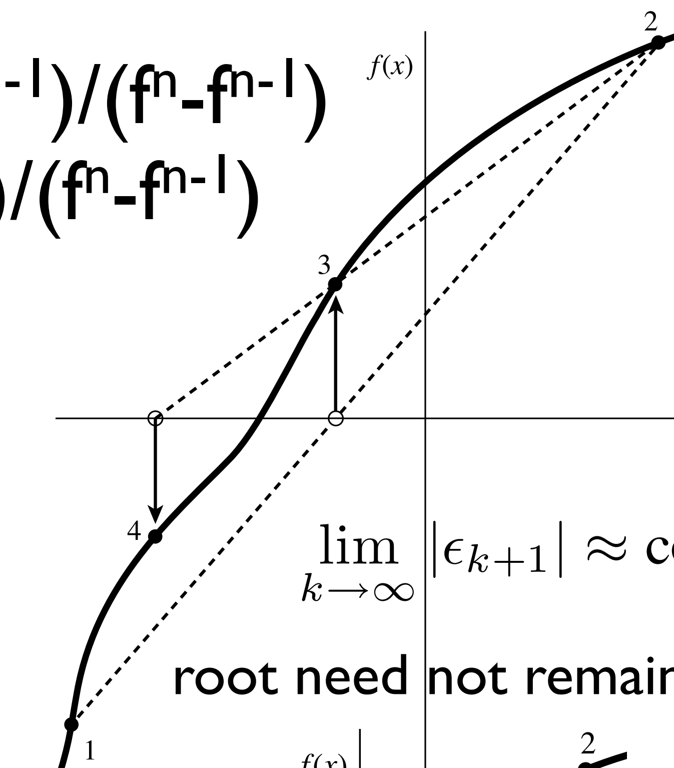


-bracket the root, $f_a f_b<0$

-halve the interval $c=(a+b)/2$

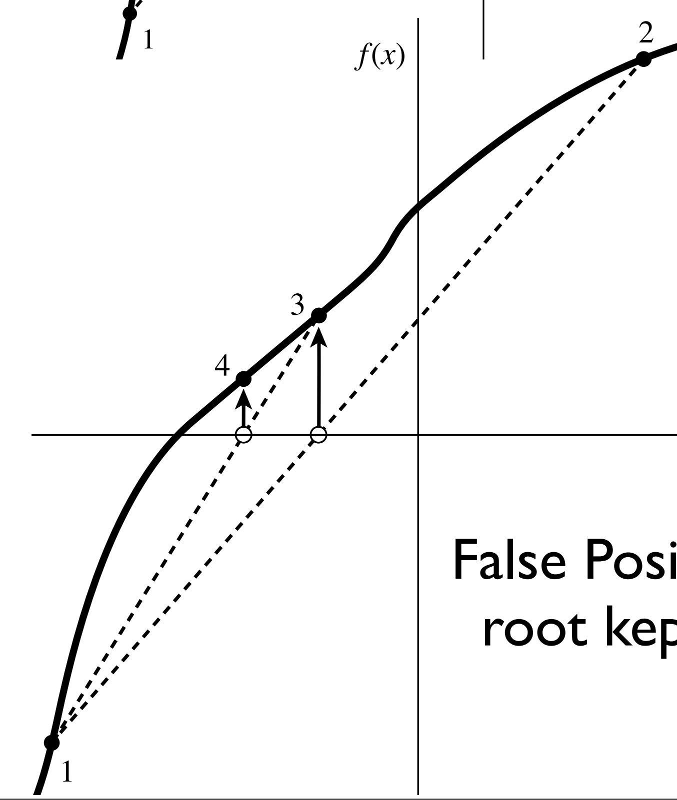
$f_a f_c<0$ or $f_b f_c<0$?

-again halve w. bracketing until desired accuracy in x_0



$$\lim_{k\to\infty} |\epsilon_{k+1}| \approx \text{const} \times |\epsilon_k|^{1.618}$$

root need not remain bracketed

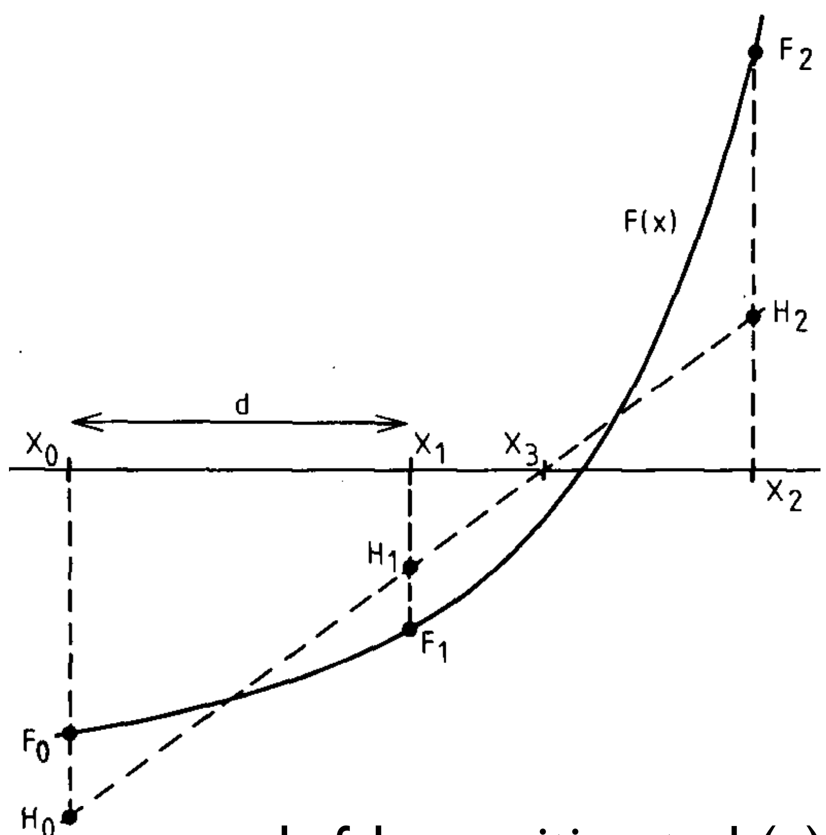


False Position Method: root kept bracketed

Ridder's Method

instead of $f(x)=0$; solve $h(x)=f(x)e^{mx}=0$; m chosen s. t.
$h_0-2h_1+h_2=0$ (i.e., center-pt. lies in middle)
$x_1=(x_0+x_2)/2$;
$x_0-x_2=x_2-x_1=d$

$$e^{md} = \frac{F_1 - \mathrm{sign}(F_0)\cdot\sqrt{W}}{F_2}, \qquad \text{with} \quad W = F_1^2 - F_0F_2.$$

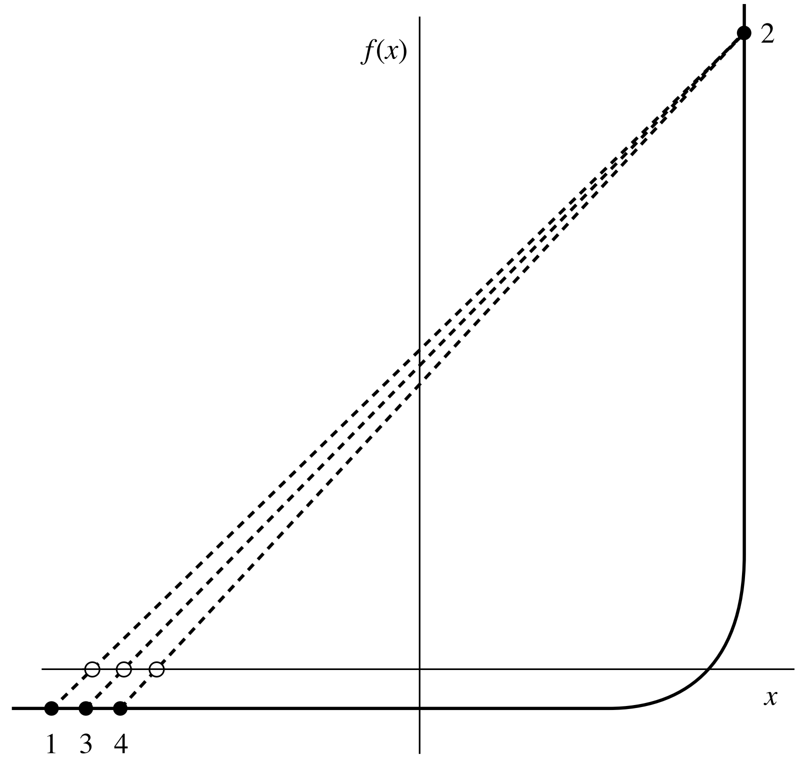


apply false position to h(x):

$$x_3 = \frac{x_1H_2 - x_2H_1}{H_2 - H_1} = x_1 - \frac{d}{H_2/H_1 - 1}$$

$$x_3 = x_1 + d\cdot\frac{F_1/F_0}{\sqrt{(F_1/F_0)^2 - F_2/F_0}}$$

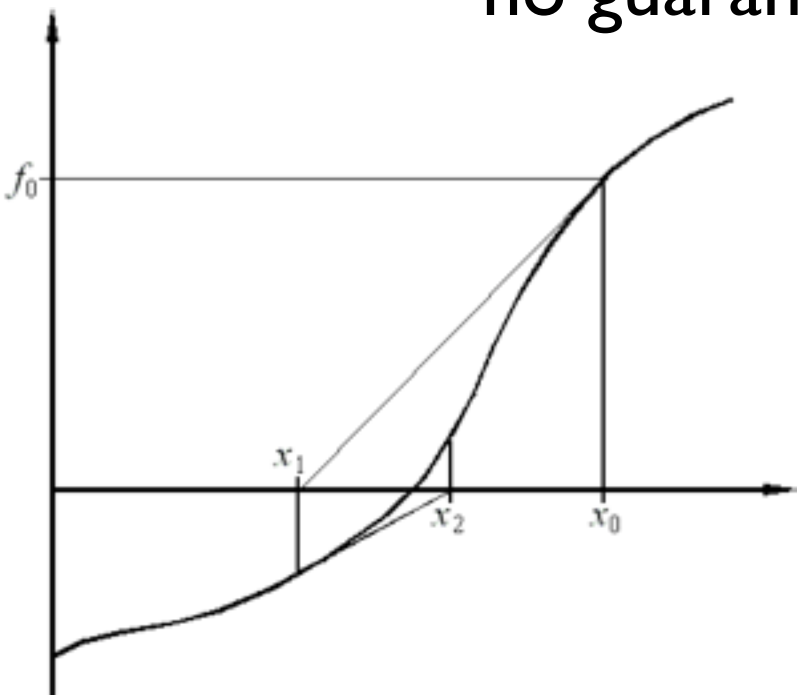
secant method/false position may be too slow!

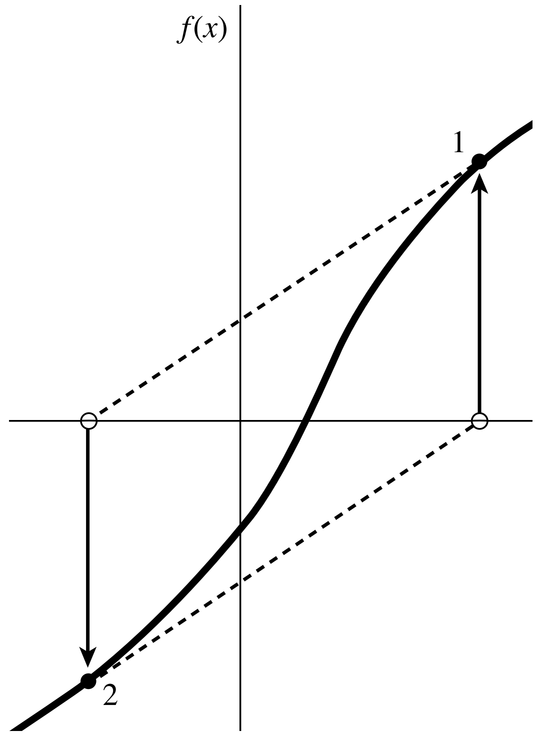


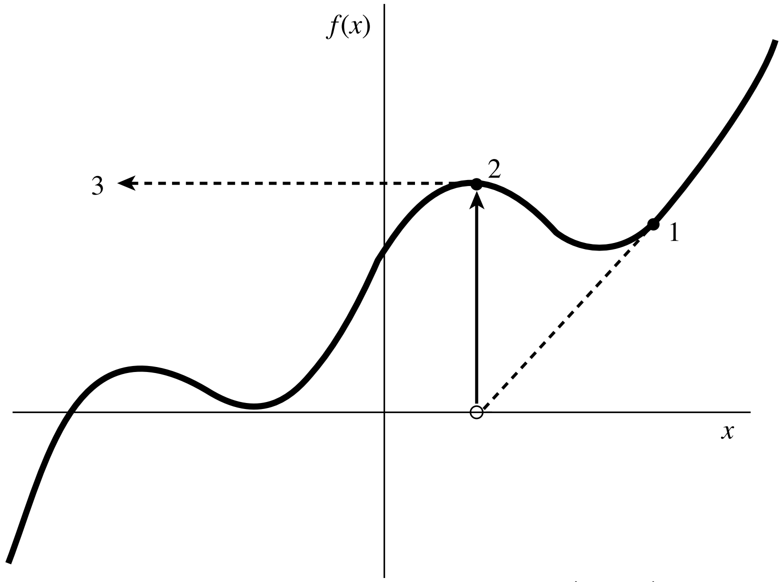
apply FP in comb. w. bisection

Newton/Newton-Raphson

no guaranteed convergence, useful when f' known







$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

$$\epsilon_{i+1} = -\epsilon_i^2 \frac{f''(x)}{2f'(x)}.$$

very rapid convergence; use w. bisection when high accuracy needed

can calculate f' numerically $f' = [f(x_i + \varepsilon) - f(x_i)]/\varepsilon$

Nonlinear Systems of Eqs.

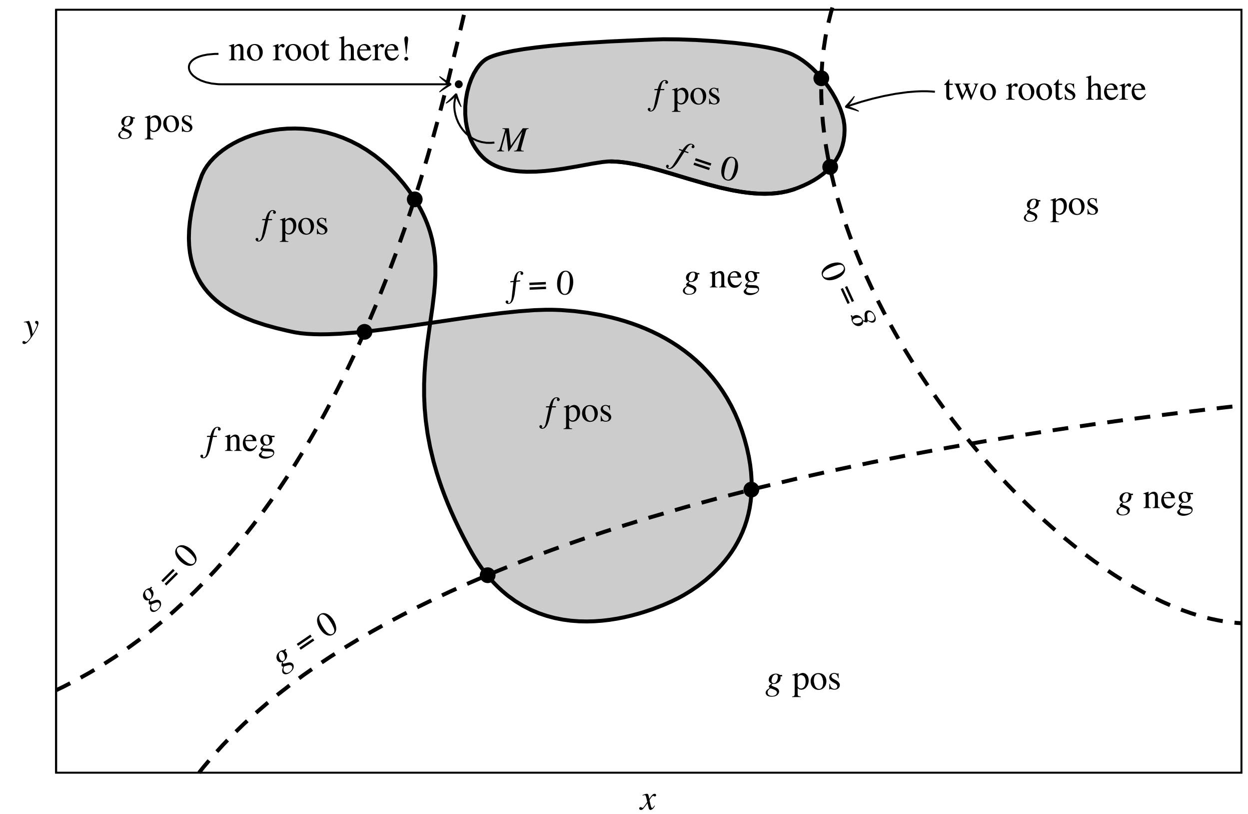
We make an extreme, but wholly defensible, statement: There are *no* good, general methods for solving systems of more than one nonlinear equation. Furthermore,

-Numerical Recipes

$$f(x, y) = 0$$

$$g(x, y) = 0$$

must have intuition about how functions look like!



Newton for multi-D

$$F_i(x_1, x_2, \ldots, x_N) = 0 \qquad i = 1, 2, \ldots, N.$$

$$F_i(\mathbf{x} + \delta\mathbf{x}) = F_i(\mathbf{x}) + \sum_{j=1}^{N} \frac{\partial F_i}{\partial x_j} \delta x_j + O(\delta\mathbf{x}^2). \qquad J_{ij} \equiv \frac{\partial F_i}{\partial x_j}$$

$$\mathbf{J} \cdot \delta\mathbf{x} = -\mathbf{F}$$ matrix equation; linear system

$$\mathbf{x}_{\text{new}} = \mathbf{x}_{\text{old}} + \delta\mathbf{x}$$

step only if |F| is smaller; otherwise try a new initial guess

other methods available too, see NR.