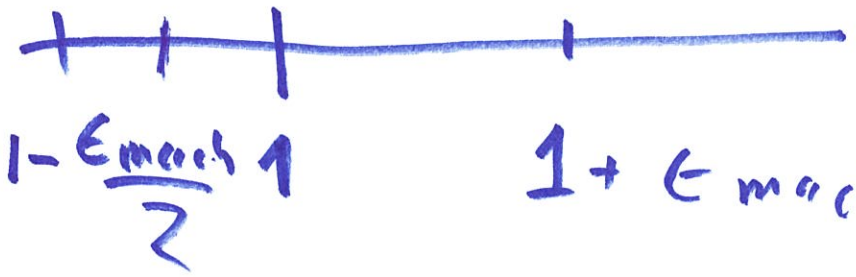
A

$$f(x) = \frac{1-(1-x)}{x}$$

$x = \epsilon_{mach}/4$



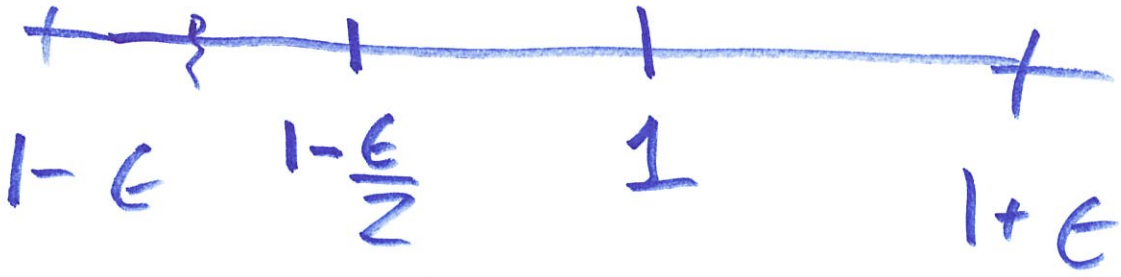
$1-x \approx 1$; $f(x) = \frac{1-1}{1} = 0$

$x = \frac{\epsilon_{mach}}{4}\,1.0001$;

$1-x \approx 1-\frac{\epsilon_{mach}}{2}$; $1-(1-x) = \frac{\epsilon_{mach}}{4}$

$f(x) = 2$

$X = \epsilon_{mach} \cdot 3/4$



$1 - X = 1 - \epsilon/2$

$1 - (1-\epsilon) = \epsilon/2$.

$f(x) = \dfrac{\epsilon}{2 \cdot 3/4\,\epsilon} = \dfrac{4}{6} = \dfrac{2}{3} = 0.66.$

$X = \epsilon_{mach} \cdot 3/4 \cdot 1.001$

$1 - X = 1 - \epsilon$; $f(x) = \dfrac{\epsilon}{3/4\,\epsilon} = 4/3 = 1.3$

C

Chopping $\left| \frac{x - \text{float}(x)}{x} \right| < \epsilon_{mach}$

round to nearest

$$\left| \frac{x - fl(x)}{x} \right| \leq \frac{\epsilon_{mach}}{2}$$

$\pm$ shift mantissa to match exponent

$-$ Cancellation

$N = 3$ $x = 1.32$ ← trailing

$y = 1.31$ ← leading

$x - y = 0.01$

① X, y D

Single floating pointing

L = -1 U = 1 N = 3 β = 2

write #, see the distances

Double N = 5

② $S_1 = \sum \frac{1}{1+e}$

$$S_1 = \sum_{k=1}^{1000} \left(\frac{e^k}{1+e^k} \right) \frac{e^{-k}}{e^{-k}}$$

$$= \sum_{k=1}^{1000} \frac{1}{e^{-k}+1} \approx 992$$

$$\sqrt{a} - \sqrt{b} = \frac{(\sqrt{a} - \sqrt{b})(\sqrt{a} + \sqrt{b})}{\sqrt{a} + \sqrt{b}}$$

$$= \frac{a - b}{\sqrt{a} + \sqrt{b}}$$

$$\sin\left(\pi \cdot \left(k^{10} + \frac{1}{k}\right)\right)$$

$$\sin(x + 2\pi) = \sin x$$

$$\sin(x + 2\pi \cdot n) = \sin x$$

n is integer

$$\sin(x + \pi) = -\sin(x)$$

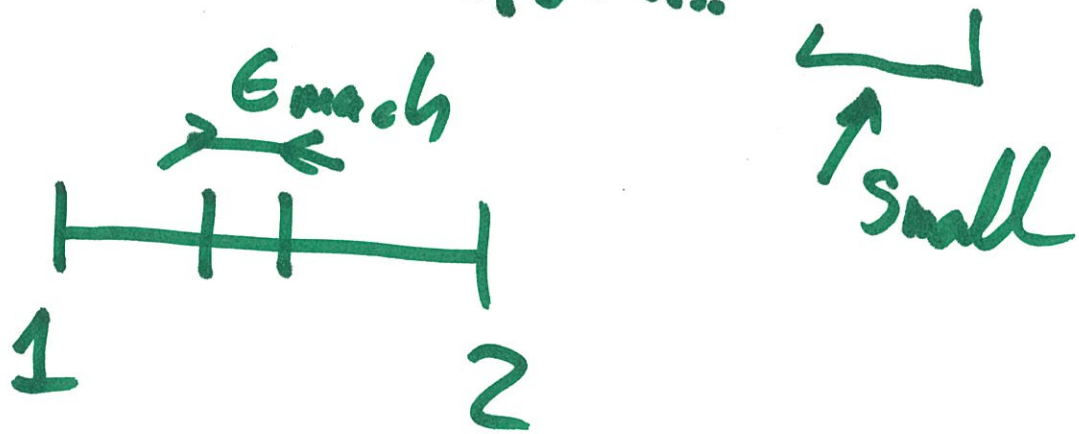
F

$$f(x) = \frac{e^x - 1}{x}$$

$$S2 = \sum_{k=1}^{\infty} 1/k^2$$

Analytical answer: $\pi^2/6$

$1 + 1/4 + 1/9 + 1/16 + \ldots$



$$S = \sum_{k=1000}^{1}$$

Sum, from large to small

$K = K_{max}$

$K = K_{max} + 1 \ldots$

What is the smallest positive # representable in Matlab (Double precision)

IEEE

$realmin = \beta^L = 2^{-1022}$

subnormals

$x_{smallest} = 0.\underbrace{00\ldots01}_{51} \cdot \beta^L =$

$= 2^{-1022} \cdot 2^{-52}$

1
10
11
100
101
110
111
1000

$52 = 32 + 16 + 4$

$= 110100$

$x = \underbrace{1111\ldots1}_{53} =$

$= 1.\underbrace{111}_{52} \cdot 10^{110100}$

$x_2 = x_1 + 1 = 1.000\ldots \cdot 10^{110101}$

$x_3 = \underbrace{11111}_{56\text{ '1's}}$

$$X = \pm \left(d_1 + \frac{d_2}{\beta} + \frac{d_3}{\beta^2} + \dots + \frac{d_N}{\beta^{N-1}} \right) \beta^E$$

$$= \pm \left(d_1 \cdot \beta^{N-1} + d_2 \cdot \beta^{N-2} + d_3 \cdot \beta^{N-3} \right.$$

$$\left. + \dots + d_{N-1} \cdot \beta + d_N \right) \cdot \frac{\beta^E}{\beta^{N-1}}$$

Find x^* s.t.

$f(x^*) = \emptyset$

$x^6 + 5x^5 + 2x^4 + 3x^3 + 27x^2 + x + 5 = \emptyset$

$X = \cos x$

g

$\cos g$

$\cos(\cos(g))$

$X = \cos X$

Guess x_0 as a possible solution to $f(x) = 0$

iterations

$$x_{k+1} = g(x_k)$$

↳ iterative function

$$x_1 = g(x_0)$$

$$x_2 = g(x_1) = g(g(x_0))$$

$$x_3 = g(x_2) = g(g(g(x_0))) \ldots$$

Hope

x_n – "final" results

How fast the correct solution is obtained?

convergence rate

$f(x) = 0$;

x^* is an exact solution

$f(x^*) = 0$

x_0 ; $\epsilon_0 = |x_0 - x^*|$

$x_1 = g(x_0)$; $\epsilon_1 = |x_1 - x^*|$

$x_2 = g(x_1)$; $\epsilon_2 = |x_2 - x^*|$

$$\epsilon_{k+1} = C \cdot \epsilon_k^r$$

C is a constant; r = convergence rate

Bisection

$r = 1 \Rightarrow$ linear convergence

$$\epsilon_{k+1} = C \cdot \epsilon_k$$

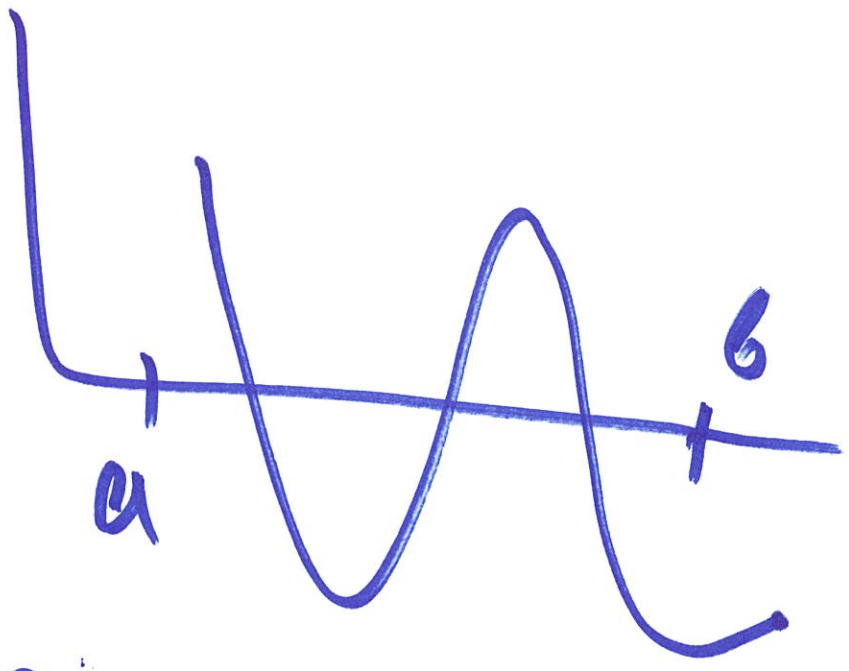
$0 < C < 1$

Newton method

$r = 2$ quadratic

$1 < r < 2$ - superlinear

Cubic convergence (HW)



Positive: it will ALWAYS find a zero

$$\epsilon_{k+1} \leq \frac{\epsilon_k}{2}; \quad r = 1$$

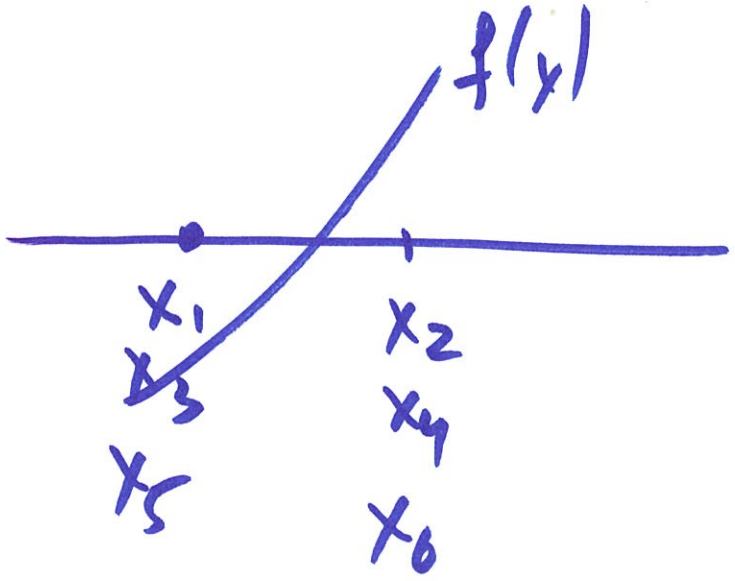
Linear convergence

One bit of a solution at a time

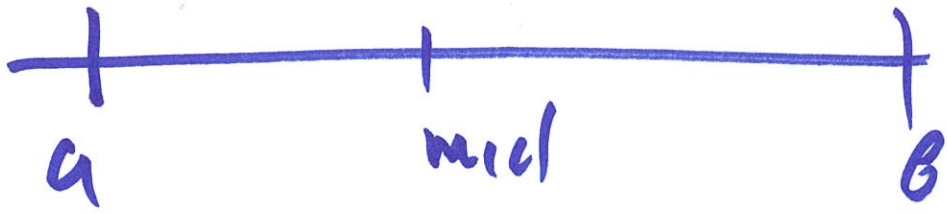
Q

Stopping criteria

1. Stop if $|f(x_k)| <$ Tolerance
2. $|x_{k+1} - x_k| <$ tolerance ✓
3. Max # iterations



R



while $|a-b| <$ Tolerance

$$m = \frac{a+b}{2}$$

if $f(a) \cdot f(m) < 0$

then $b = m$

else $a = m$

end if

end while

B [a, b] :=

{

m = $\frac{a+b}{2}$; IF |f(m)| < Tolerance Return m

~~IF f(m) < tolerance~~

IF f(a) · f(m) < 0

then B[a, m]

else B[b, m]

end if

}