1 Nonlinear Equations & Bisection Method (Section 3.1)

Let's consider the problem of finding zeros.

(Example.) Suppose we have the functions $\sin(x)$ and e^x, and suppose we want to find the values of x such that $\sin(x) - e^x$. Let

$$f(x) = \sin(x) - e^x$$

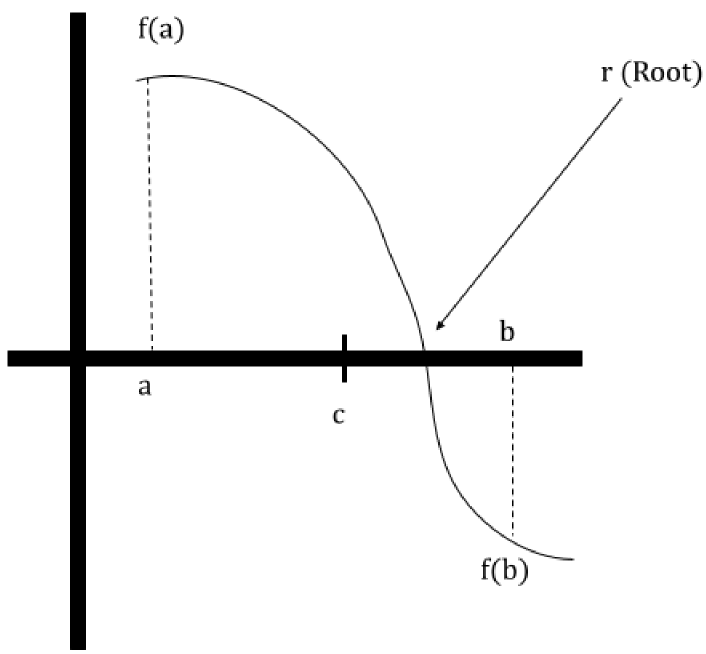
so that we can solve for $\sin(x) - e^x = 0$. We want to seek the "zero" (i.e., the root) of $f(x)$.

1.1 Bisection Method

The bisection method makes the assumption that $f : \mathbb{R} \mapsto \mathbb{R}$ is continuous on an interval $[a, b]$. We want to consider the interval $[a, b]$ so that

$$f(a) \cdot f(b) < 0.$$

This implies that either $f(x)$ or $f(b)$, but not both, are negative (i.e., $\text{sgn}(f(a)) \neq \text{sgn}(f(b))$).



Then, we can solve for

$$c = \frac{b + a}{2} = \frac{a + (b - a)}{2}.$$

c is the midpoint of a and b. Then, we let check

- If $f(a) \cdot f(c) < 0$, then we let $b \leftarrow c$ and start with the new interval $[a, b]$.
- Otherwise, $f(b) \cdot f(c) < 0$ and so we let $a \leftarrow c$ and start with the new interval $[a, b]$.

We keep repeating this until the interval becomes sufficiently small.

1.2 Algorithm Idea and Stopping Conditions

Let

- M be the maximum iterations,

- δ be the interval tolerance ($|b-a|<\delta$), and
- ϵ be the zero tolerance ($|f(c)|<\epsilon$).

The algorithm takes in five inputs: $(M, \delta, \epsilon, a, b)$, where a and b (such that $a \leq b$) represents the interval endpoints (i.e., $[a, b]$).

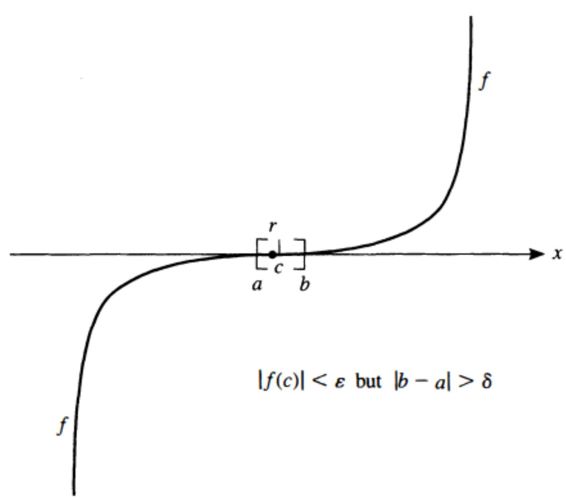
Algorithm 1 Bisection Method

```
function BISECTION(M, δ, ϵ, a, b)
    u ← f(a)
    v ← f(b)
    e = b − a
    if sgn(u) = sgn(v) then
        Error
    end if
    for k ← 1 to M do
        e ← e/2
        c ← a + e
        w ← f(c)
        if |e| < δ or |w| < ϵ then
            Break
        end if
        if sgn(u) = sgn(w) then
            a ← c
            u ← w
        else
            b ← c
            v ← w
        end if
    end for
    return c
end function
```

1.3 Tolerances

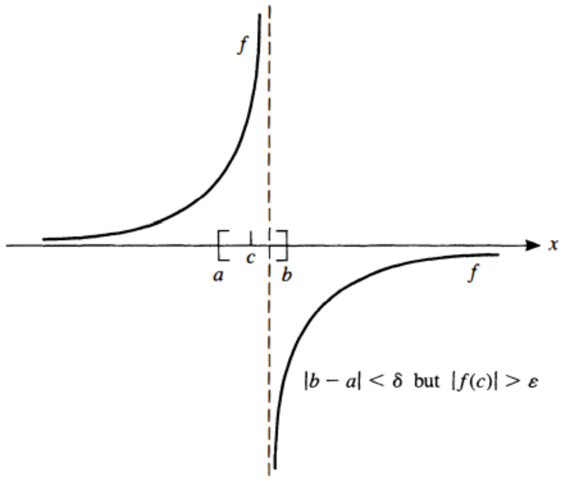
We introduces the tolerances, δ and ϵ, for robustness.

- We might have $|f(c) < \epsilon$ but $|b-a|>\delta$.



Notice how the graph is flat near the zero. This corresponds to a multiple root, which means the bisection method could have difficulty determining this zero to a high precision.

- We also have $|b-a| < \delta$ and $|f(c)| > \epsilon$.



The curve here, which is in the interval $[a, b]$, is not continuous.

Remark: From the first point, if we have a function with a double root, then approximation may become less precise or outright impossible. For example, consider $f(x) = (x-1)^2$, which has a double root of $x = 1$. Because $\forall x \in \mathbb{R}, f(x) \geq 0$, the bisection method cannot be used here.

1.4 Error Analysis

Let $I_i = [a_i, b_i]$ be an interval. Then, we essentially have a sequence of intervals, $I_0, I_1, I_2, I_3, \ldots$, where $a_0 \leq a_1 \leq a_2 \leq \ldots \leq b_0$ and $b_0 \geq b_1 \geq b_2 \geq \ldots \geq a_0$.

```
a_m+1       b_m+1
 |---X------|----------|
a_m         c_m         b_m
```

For $m \geq 0$, We know that $b_{m+1} - a_{m+1} = \frac{1}{2}(b_m - a_m)$, so applying it repeatedly gives us

$$\begin{aligned} b_m - a_m &= \frac{1}{2}(b_{m-1} - a_{m-1}) \\ &= \left(\frac{1}{2}\right)^m (b_0 - a_0) \end{aligned}.$$

The sequence, b_m and a_m, are monotonic and convergent. Because we're constantly dividing the interval by 2, we know that

$$\lim_{m\mapsto\infty}(b_m - a_m) = \lim_{m\mapsto\infty}\left(\frac{1}{2}\right)^m (b_0 - a_0) = 0.$$

We also know that, as a and b are the left and right endpoints of this interval, both a and b will converge to the root of the equation, r, i.e.,

$$\lim_{m\mapsto\infty} a_m = \lim_{m\mapsto\infty} b_m = r.$$

Then,

$$\lim_{m\mapsto\infty} f(a_m)f(b_m) \leq 0.$$

This yields

$$f(r)f(r) \leq 0 \implies f(r) = 0.$$

Therefore,

```
 |-------r--|----------|
a_m        c_m        b_m
```

In this sense, if at a certain stage in the process we have the interval I_n and the process is now stopped, the root is certain to lie in this interval. However, the best estimate of the root at this stage is not a_n or b_n, but the midpoint of the interval, c_n. In this sense, the error is then bounded by

$$|r - c_m| \leq \frac{1}{2}|b_m - a_m|.$$

This gives us the following theorem.

Theorem 1.1: Bisection Method

If $I_i = [a_i, b_i]$ is an interval and $I_0, I_1, \ldots, I_n$ denote the intervals in the bisection method, then the limits $\lim_{n\mapsto\infty} a_n$ and $\lim_{n\mapsto\infty} b_n$ exist, are equal, and represent a zero of f. If $r = \lim_{n\mapsto\infty} c_n$ and $c_n = \frac{1}{2}(a_n + b_n)$, then

$$|r - c_n| \leq \frac{1}{2}|b_n - a_n| = 2^{-(n+1)}|b_0 - a_0|.$$

(Exercise.) Suppose $[a_0, b_0] = [50, 63]$. How many steps needs to be done using the bisection method to compute a root with relative accuracy of 10^{-12}?

We want to seek $|r - c_m|/|r| \leq 10^{-12}$. This means that

$$|r - c_m|/50 \leq 10^{-12}.$$

(Note that we asked for the <u>relative</u> error.) The sufficient condition is

$$|r - c_m|/50 \leq \left(\frac{1}{2}\right)^{m+1}(b_0 - a_0)/50 \leq 10^{-12}.$$

From the theorem, we have

$$2^{-(n+1)}(13/50) \leq 10^{-12}.$$

From there, we can solve for n.