Solutions of Equations in One Variable

Simon Fraser University – Surrey Campus

MACM 316 – Spring 2005

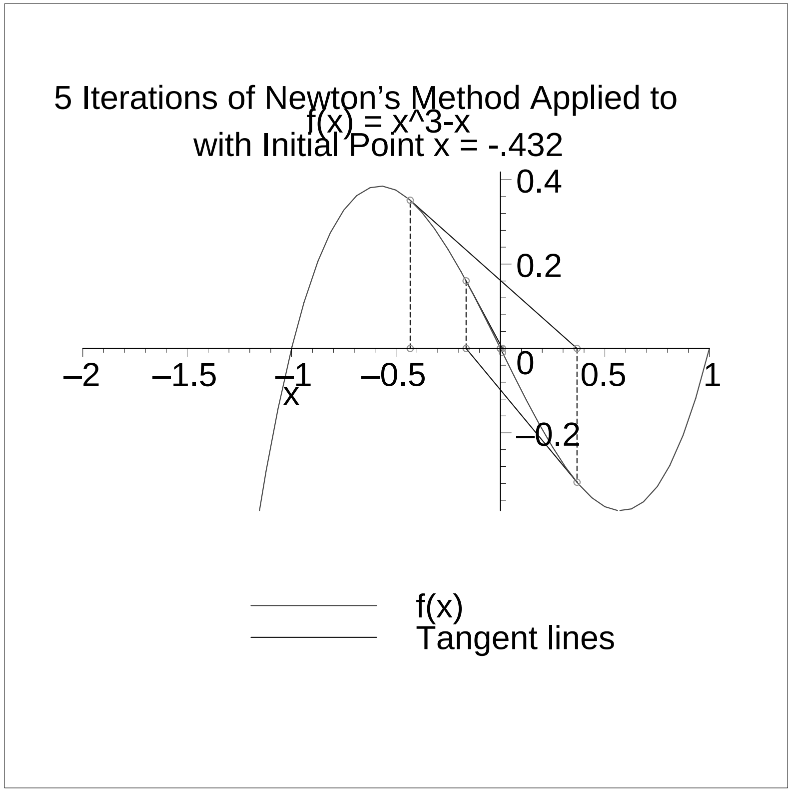
Instructor: Ha Le

Overview

- Iterative Methods
  - The Bisection Method
  - Fixed-Point Iteration
  - Newton-Raphson's Method
  - Secant Method
  - Method of False Position
- Error Analysis for Iterative Methods
- Accelerating Convergence
- Zeros of Polynomials and Müller's Method

The Basic Problem

- Given $f : \mathbb{R} \Longrightarrow \mathbb{R}$, find $x \in \mathbb{R}$ such that $f(x) = 0$.
- In most cases, it is not possible to solve analytically. We will consider *iterative methods* to *approximate* the solution.



The Bisection Method

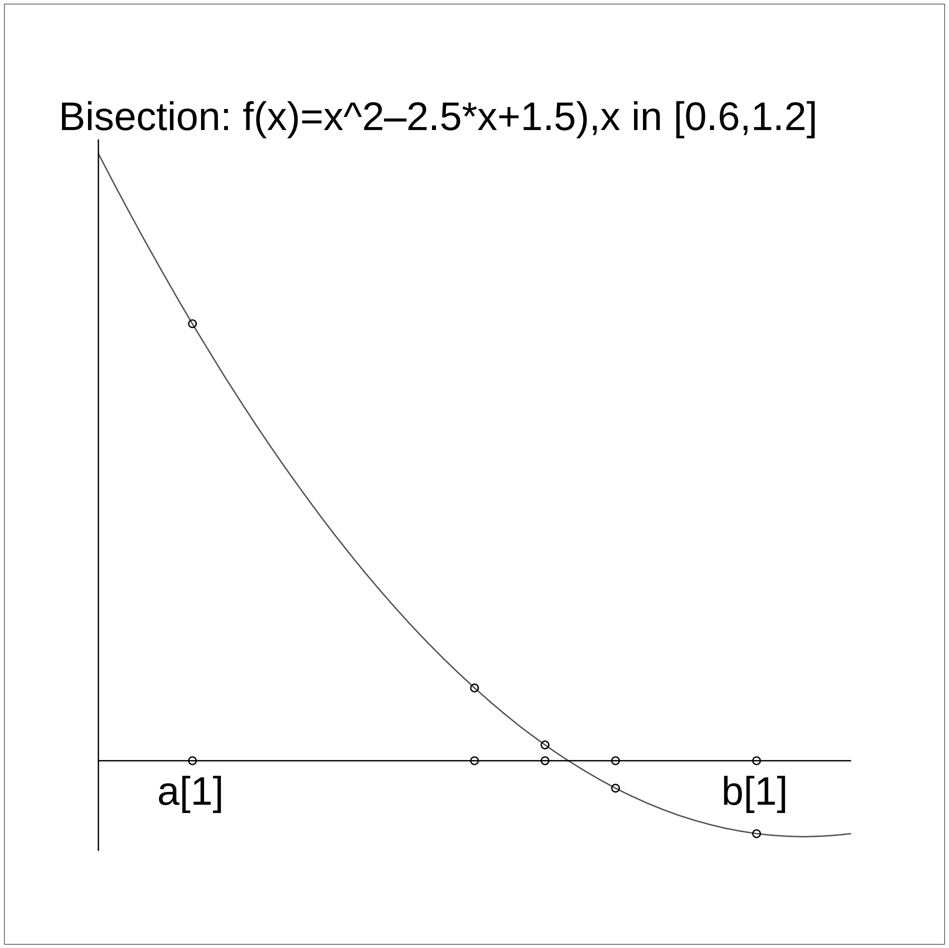
Input. $f \in C[a,b]$ with $f(a) \cdot f(b) < 0$.

There must be a $p \in \mathbb{R}$ in (a,b) with $f(p) = 0$ by the

Theorem 1 *Intermediate Value Theorem (IVT). If $f \in C[a,b]$ and K is any number between $f(a)$ and $f(b)$, then there exists a number c in (a,b) for which $f(c) = K$.*

Algorithm description.

- $a_1 := a$, $b_1 := b$, $p_1 := 1/2(a_1 + b_1)$;
- if $f(p_1) = 0$ then we are done ($p = p_1$).
- if $f(p_1)f(a_1) < 0$, then $\exists p \in (a_1, p_1)$ s.t. $f(p) = 0$. Set $a_2 := a_1$, and $b_2 := p_1$. Otherwise, $f(p_1)f(b_1) < 0$. Set $a_2 := p_1$ and $b_2 := b_1$.
- Reapply the process to $[a_2, b_2]$.


Bisection: f(x)=x^2–2.5*x+1.5),x in [0.6,1.2]
a[1]
b[1]

The Bisection Method: Stopping Criteria

Once appropriate stopping criteria are satisfied, we set the midpoint of the interval equal to the estimate for the root.

Possible stopping criteria.

(1) $\dfrac{1}{2}(b_n - a_n) < \text{TOL}$ or $|p_n - p_{n-1}| < \text{TOL}$.

- GOOD: ensures that the returned root value p_n is within TOL of the exact value p; easy error analysis.
- BAD: does not ensure that $f(p_n)$ is small; an absolute rather than a relative error.

(2) $\dfrac{|p_n - p_{n-1}|}{|p_n|} < \text{TOL}, \; p_n \neq 0.$

Usually preferred over (1) if nothing is known about f or p.

(3) $|f(p_n)| < \text{TOL}$.

Ensures that $f(p_n)$ is small, but p_n may differ significantly from the true root p.

(4) We can also carry out a fixed number of iterations N. This is closely related to (1).

- The best stopping criteria will depend on what is known about f and p and on the type of problem.

- It is often useful to use the relative error test (2) with a fixed, maximum number of steps (4).

The Bisection Method: Fine Points

- $p_i = a_i + \dfrac{b_i - a_i}{2}$ is preferred over $p_i = \dfrac{a_i + b_i}{2}$.

It is usually best in bisection to add a small correction to a previous approximation. Could otherwise lead to $p \notin [a, b]$.

- To avoid underflows and overflows, it is sometimes preferred to use $\text{sign}(f(a_i)) \cdot \text{sign}(f(b_i)) > 0$ rather than $f(a_i) \cdot f(b_i) > 0$.

- We want to choose the initial interval as small as possible to minimize the number of iterations.

- The bisection method in general is slow to converge. However, the method always converges (by the intermediate value theorem) which makes it an excellent choice to start other methods, or to use when other methods fail.

Accuracy of Bisection Method

Theorem 2 *Suppose that $f \in C[a,b]$ and $f(a) \cdot f(b) < 0$. The Bisection method generates a sequence $\{p_n\}_{n=1}^{\infty}$ approximating a zero p of f with*

$$|p_n - p| \le \frac{b-a}{2^n}, \qquad \textit{when } n \ge 1.$$

Proof.

$$\forall n \ge 1, \; b_n - a_n = \frac{1}{2^{n-1}}(b-a) \quad \text{and} \quad p \in (a_n, b_n).$$

Since $p_n = \dfrac{1}{2}(a_n + b_n), \; \forall n \ge 1,$

$$|p_n - p| \le \frac{1}{2}(b_n - a_n) = \frac{b-a}{2^n}.$$

$\blacksquare$

Convergence of Bisection Method

Recall. Suppose $\{\beta_n\}_{n=1}^{\infty}$ is a sequence known to converge to zero, and $\{\alpha_n\}_{n=1}^{\infty}$ converges to a number α. If a positive constant K exists with

$$|\alpha_n - \alpha| \leq K|\beta_n| \quad \text{for large } n,$$

then we say that $\{\alpha_n\}_{n=1}^{\infty}$ converges to α with *rate of convergence* $O(\beta_n)$. It is indicated by writing $\alpha_n = \alpha + O(\beta_n)$.

- Since $|p_n - p| \leq \dfrac{b-a}{2^n}$, the sequence $\{p_n\}_{n=1}^{\infty}$ converges to p with rate of convergence $O\left(\dfrac{1}{2^n}\right)$. That is, $p_n = p + O\left(\dfrac{1}{2^n}\right)$.

The Bisection Method: an Example

- Find a bound N for the number of iterations needed to approximate a solution to the equation $x^3 + 4x^2 - 10 = 0$ on the interval $[1, 2]$ to an accuracy of 10^{-3}.
- $|p_N - p| \leq \dfrac{b-a}{2^N} = 2^{-N} < 10^{-3}$.

$\Longrightarrow \log_{10}\, 2^{-N} < \log_{10}\, 10^{-3}$

$\Longrightarrow -N\, \log_{10} 2 < -3$

$\Longrightarrow N > \dfrac{3}{\log_{10} 2} \approx 9.96$

$\Longrightarrow N \geq 10.$

Fixed Point Iteration

Definition 1 *A number p is a* fixed point *for a given function g if $g(p) = p$.*

Example 1 For $g(x) = x^2 - 2$, $-2 \le x \le 3$, $x = -1$ and $x = 2$ are two fixed points for g.

Remark 1 Fixed-point problems and root finding problems are *equivalent.*

Let $g(x) - x = f(x)$.

$$\begin{aligned} f \text{ has a root } p &\implies g \text{ has a fixed point } p; \\ g \text{ has a fixed point } p &\implies f \text{ has a root } p. \end{aligned}$$

Remark 2 For a given $f(x)$, there might be many choices for $g(x)$.

Example 2 For $f(x) = x^3 + 4x^2 - 10 = 0$, one can verify that the fixed point of each $g_i(x)$ is a solution to $f(x) = 0$.

(a) $x = g_1(x) = x - x^3 - 4x^2 + 10;$

(b) $x = g_2(x) = \left(\frac{10}{x} - 4x\right)^{1/2};$

(c) $x = g_3(x) = \frac{1}{2}(10 - x^3)^{1/2};$

(d) $x = g_4(x) = \left(\frac{10}{4 + x}\right)^{1/2};$

(e) $x = g_5(x) = x - \frac{x^3 + 4x^2 - 10}{3x^2 + 8x}.$

Fixed Points: Existence & Uniqueness

Theorem 3 (a) *If $g \in C[a,b]$ and $g(x) \in [a,b]$ for all $x \in [a,b]$, then g has a fixed point in $[a,b]$;*

(b) *If, in addition, $g'(x)$ exists on (a,b) and a positive $k < 1$ exists with $|g'(x)| \le k$, $\forall x \in (a,b)$, then the fixed point in $[a,b]$ is unique.*

Proof. (a)

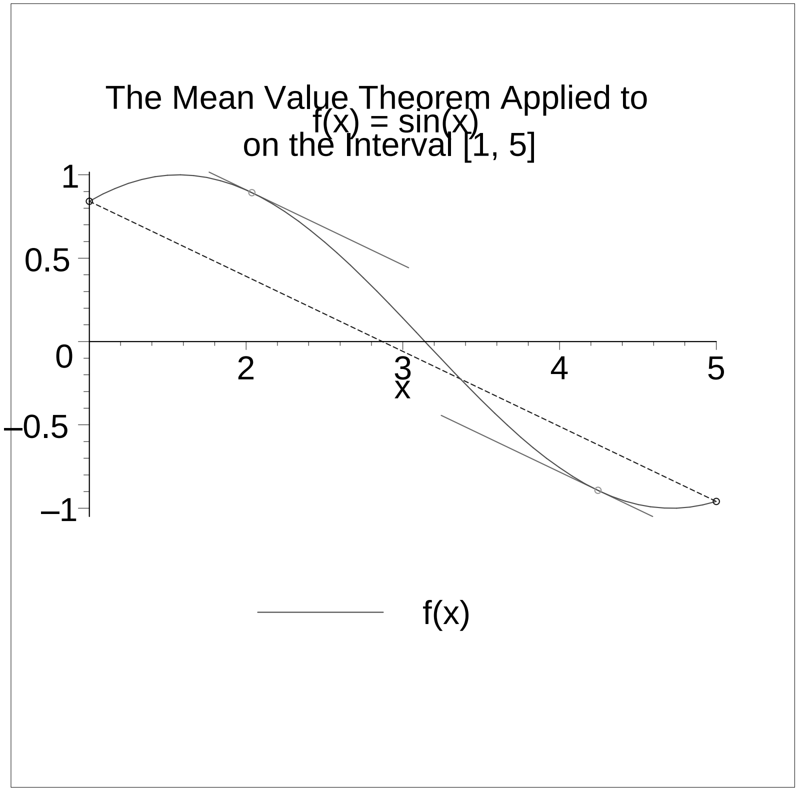
- If $g(a) = a$ or $g(b) = b$, then g has a fixed point at an end point;
- Otherwise, $g(a) > a$ and $g(b) < b$. Define $h(x) = g(x) - x$.

$$\left.\begin{array}{l} h \text{ is continuous on } [a,b] \\ h(a) = g(a) - a > 0 \\ h(b) = g(b) - b < 0 \end{array}\right\} \stackrel{\text{IVT}}{\Longrightarrow} \exists p \in (a,b) \text{ s.t. } h(p) = 0.$$

Since $0 = h(p) = g(p) - p$, p is a fixed point of g. ∎

In order to prove part (b), we need

Theorem 4 *Mean Value Theorem (MVT). If $f \in C[a,b]$ and f is differentiable on (a,b), then a number c in (a,b) exists with* $f'(c) = \dfrac{f(b) - f(a)}{b - a}$.



Proof. (b) Suppose that p and q are two different fixed points on $[a, b]$, by MVT, $\exists\, \xi \in (p, q) \in [a, b]$ s.t $\dfrac{g(p) - g(q)}{p - q} = g'(\xi)$. Thus

$$\begin{aligned} |p - q| &= |g(p) - g(q)| &= |g'(\xi)||p - q| \\ &\leq k|p - q| &< |p - q|. \end{aligned}$$

A contradiction. Hence $p = q$, i.e., the fixed point on $[a, b]$ is unique. ∎

Fixed Point Iteration: Main Idea

We want to approximate the fixed point of a function $g(x)$.

- choose an initial approximation p_0;
- generate a sequence $\{p_n\}_{n=0}^{\infty}$ such that $p_n = g(p_{n-1})$, $n \geq 1$.

IF the sequence converges to p and g is continuous

$$p = \lim_{n\to\infty} p_n = \lim_{n\to\infty} g(p_{n-1}) = g\left(\lim_{n\to\infty} p_{n-1}\right) = g(p).$$

Fixed Points: Convergence

The following theorem and its corollary give us some clues concerning the paths we should pursue and, perhaps more importantly, some we should reject.

Theorem 5 *Fixed-Point Theorem (FPT). Let $g \in C[a,b]$ be such that $g(x) \in [a,b]$, for all x in $[a,b]$. Suppose, in addition, that g' exists on (a,b) and that a constant $0 < k < 1$ exists with*

$$|g'(x)| \leq k, \text{ for all } x \in (a,b).$$

Then, for any number p_0 in $[a,b]$, the sequence defined by

$$p_n = g(p_{n-1}), \ \ n \geq 1,$$

converges to the unique fixed point p in $[a,b]$.

Proof.

$$\left.\begin{array}{l} p_n = g(p_{n-1}) \\ g : [a,b] \rightarrow [a,b] \\ p_0 \in [a,b] \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \{p_n\}_{n=0}^{\infty} \text{ is defined} \quad \forall n \geq 0 \\ p_n \in [a,b] \quad \forall n. \end{array}\right.$$

$$\begin{aligned} |p_n - p| &= |g(p_{n-1}) - g(p)| \overset{\text{MVT}}{=} |g'(\xi_n)||p_{n-1} - p| \qquad (1) \\ &\leq k|p_{n-1} - p| \end{aligned}$$

where $\xi_n \in (a,b)$. Applying (1) inductively gives

$$|p_n - p| \leq k|p_{n-1} - p| \leq k^2|p_{n-2} - p| \leq \cdots \leq k^n|p_0 - p|.$$

Since $0 < k < 1$, $\lim_{k\to\infty} k^n = 0$, and

$$\lim_{k\to\infty} |p_n - p| \leq \lim_{k\to\infty} k^n|p_0 - p| = 0.$$

Hence, $\{p_n\}_{n=0}^{\infty}$ converges to p. ∎

Corollary 1 *If g satisfies the hypotheses of Theorem 5, then bounds for the error involved in using p_n to approximate p are given by*

$$|p_n - p| \leq k^n \max\{p_0 - a, b - p_0\}$$

and

$$|p_n - p| \leq \frac{k^n}{1-k}|p_1 - p_0|, \quad \textit{for all } \; n \geq 1.$$

Newton-Raphson Method

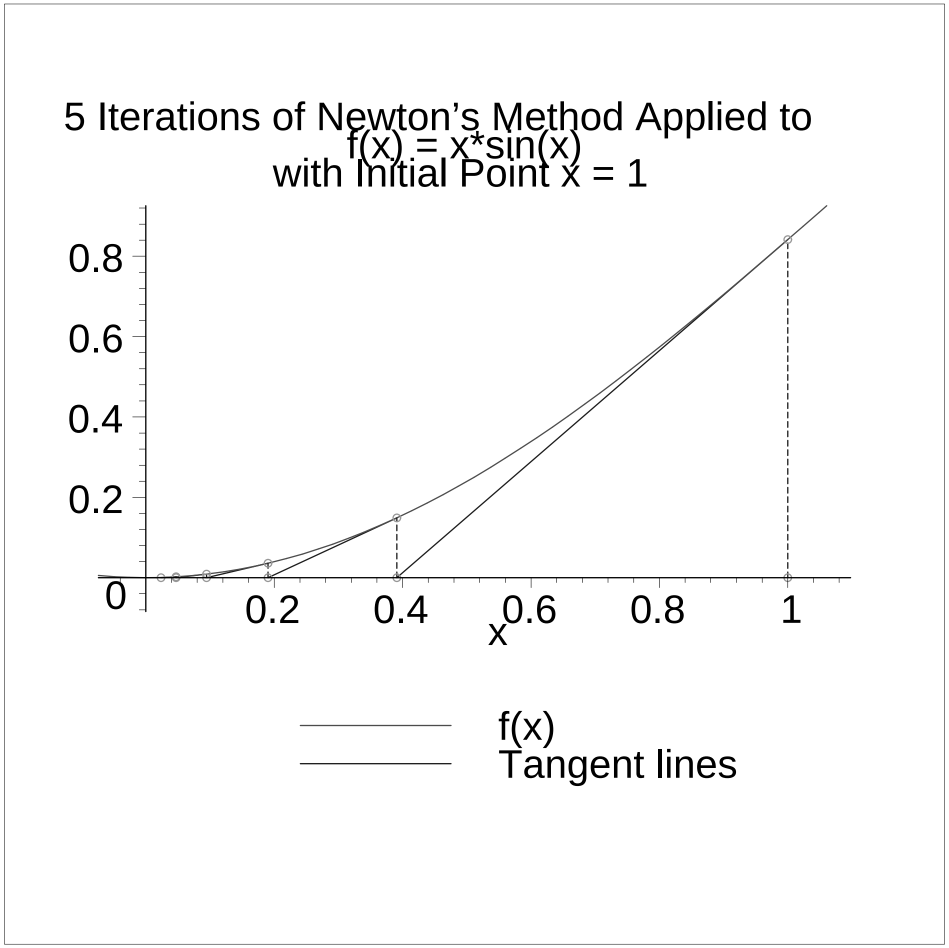
- One of the most powerful and well-known methods.

Pros. Much faster than Bisection method.

Cons. (1) Need $f'(x)$; (2) Not guaranteed to always converge.

Goal. $x = p$ such that $f(x) = 0$.

Idea. Use slope as well as function values.



Newton-Raphson Method: Derivation

Let p be a root of $f(x) = 0$. Suppose $f \in C^2[a,b]$. Let $\bar{x} \in [a,b]$ be an approximation to p such that

$$f'(\bar{x}) \neq 0 \quad \text{and } |\bar{x} - p| \quad \text{is sufficiently small.}$$

By Taylor's theorem:

$$f(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{1}{2} f''(\xi(x))(x - \bar{x})^2 \tag{2}$$

where $\xi(x)$ lies between x and $\bar{x}$. Set $x = p$ in (2):

$$0 = f(\bar{x}) + f'(\bar{x})(p - \bar{x}) + \frac{1}{2} f''(\xi(p))(p - \bar{x})^2.$$

If $|p - \bar{x}|$ is small, $|p - \bar{x}|^2$ is smaller. Hence,

$$0 \approx f(\bar{x}) + f'(\bar{x})(p - \bar{x}) \implies p \approx \bar{x} - \frac{f(\bar{x})}{f'(\bar{x})}.$$

• Newton's method begins with an estimate p_0 and generates a sequence $\{p_n\}$:

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}.$$

• A stopping criterion is similar to that of Bisection method, e.g.,

$$(1)\ |p_n - p_{n-1}| < \epsilon, \quad (2)\ \frac{|p_n - p_{n-1}|}{|p_n|} < \epsilon, \quad (3)\ |f(p_n)| < \epsilon.$$

• Notice that Newton's method fails if $f'(p_{n-1}) = 0$.

• The method is most effective when f' is bounded away from zero near p.

• It follows from the derivation that $p - \bar{x}$ has to be small, i.e., we need a good initial guess.

Newton-Raphson Method: Example

- Use Newton's method to compute the square root of a number R.
- Want to find the roots of $p^2 - R = 0$. Let $f(x) = x^2 - R$. Then $f'(x) = 2x$. Newton's method takes the form

$$\begin{aligned} p_n &= p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})} \\ &= \frac{1}{2} p_{n-1} + \frac{R}{2\,p_{n-1}}. \end{aligned}$$

Try $R = 2$ and $p_0 = 2$:

$$p_1 = 1.500000000, \quad p_2 = 1.416666666,$$
$$p_3 = 1.414215686, \quad p_4 = 1.414213562$$

Newton-Raphson Method: Convergence

Newton's method can be shown to converge under reasonable assumptions (smoothness of f, a good initial guess, $f'(p) \neq 0$).

Theorem 6 *Let $f \in C^2[a,b]$. If $p \in [a,b]$ is such that*

$$f(p) = 0 \quad \textit{and} \quad f'(p) \neq 0,$$

then there exists a $\delta > 0$ such that Newton's method generates a sequence $\{p_n\}_{n=1}^{\infty}$ converging to p for any initial approximation

$$p_0 \in [p-\delta,\, p+\delta].$$

Proof. General direction. Consider

$$g(x) = x - \frac{f(x)}{f'(x)}. \tag{3}$$

Let $k \in (0, 1)$,

(1) find $\delta > 0$ s.t. $g : [p - \delta, p + \delta] \longrightarrow [p - \delta, p + \delta]$;

(2) show that $|g'(x)| \le k, \;\; \forall x \in (p - \delta, p + \delta)$.

$\overset{FPT}{\Longrightarrow}$ $\forall p_0 \in [p - \delta, p + \delta]$, the sequence defined by

$$p_n = g(p_{n-1}), \; n \ge 1$$

converges to the unique fixed point p in $[p - \delta, p + \delta]$.

A proof for (2)

$$\left.\begin{array}{l} f' \in C[a,b] \\ f'(p) \neq 0 \end{array}\right\} \Longrightarrow \begin{array}{l} \exists \delta_1 > 0 \text{ s.t.} \\ f'(x) \neq 0 \; \forall x \in [p-\delta_1, p+\delta_1] \subseteq [a,b]. \end{array}$$

By (3), g is defined and continuous on $[p-\delta_1, p+\delta_1]$.

Also by (3),

$$g'(x) = \frac{f(x)f''(x)}{(f'(x))^2}, \;\; \forall x \in [p-\delta_1, p+\delta_1].$$

$$\left.\begin{array}{l} 0 < k < 1 \\ g' \in C[p-\delta_1, p+\delta_1] \\ g'(p) = 0 \end{array}\right\} \Longrightarrow \begin{array}{l} \exists \delta,\, 0 < \delta < \delta_1 \text{ s.t.} \\ |g'(x)| \leq k, \;\; \forall x \in [p-\delta, p+\delta]. \end{array}$$

A proof for (1)

$$\left.\begin{array}{l} g \in C[p-\delta, p+\delta] \\ g \text{ diff'able on } (p-\delta, p+\delta) \\ x \in [p-\delta, p+\delta] \end{array}\right\} \overset{\text{MVT}}{\Longrightarrow} \begin{array}{l} \exists\, \xi \ \text{ between } x \text{ and } p \text{ s.t.} \\ |g(x)-g(p)| = |g'(\xi)|\,|x-p| \end{array}$$

$$\Longrightarrow\ |g(x)-p| = |g(x)-g(p)| = |g'(\xi)||x-p| \le k|x-p| < |x-p|.$$

Hence,

$$\forall\, x \in [p-\delta, p+\delta], \ \text{ if } |x-p| < \delta, \ \text{ then } |g(x)-p| < \delta,$$

i.e., $g(x) \in [p-\delta, p+\delta]$. ■

Secant Method

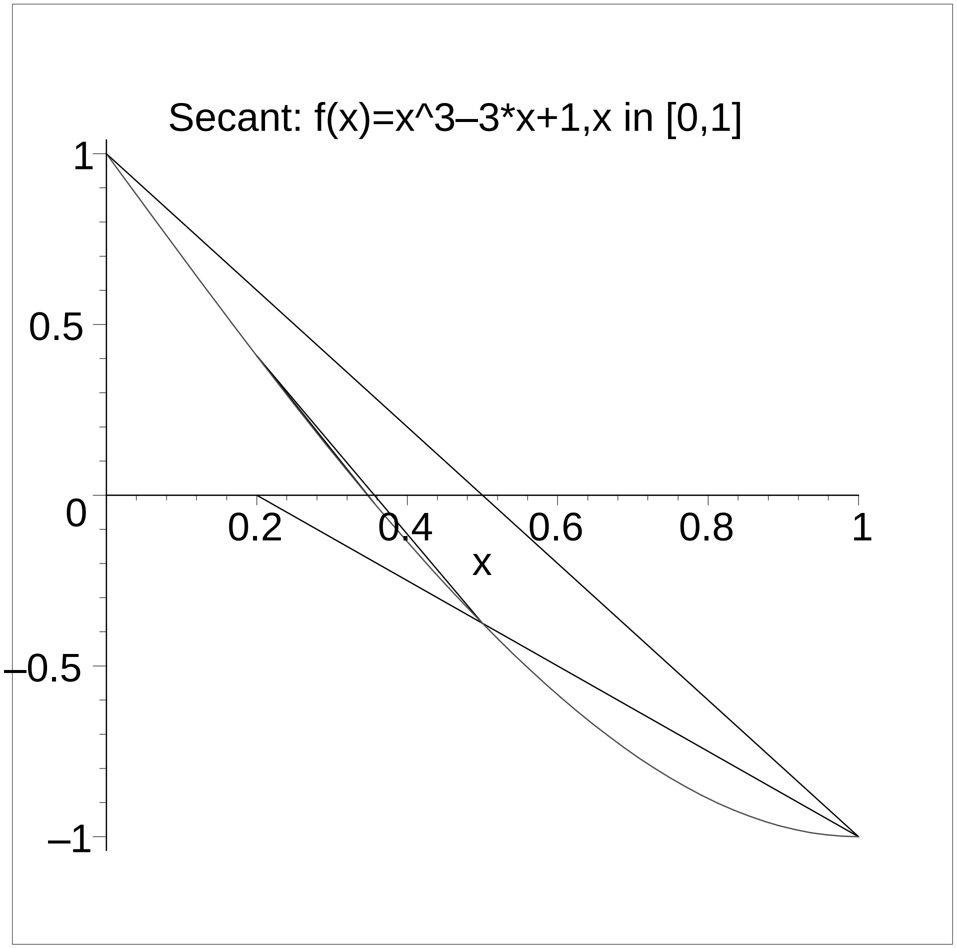
- Newton's method has the major difficulty that the derivative of f is needed at each approximation. Since $f'(p_{n-1}) = \lim_{x \to p_{n-1}} \dfrac{f(x) - f(p_{n-1})}{x - p_{n-1}}$, we approximate

$$\begin{aligned} f'(p_{n-1}) &\approx \frac{f(p_{n-2}) - f(p_{n-1})}{p_{n-2} - p_{n-1}} \\ &= \frac{f(p_{n-1}) - f(p_{n-2})}{p_{n-1} - p_{n-2}}. \end{aligned}$$

This gives the Secant Method:

$$p_n = p_{n-1} - \frac{f(p_{n-1})(p_{n-1} - p_{n-2})}{f(p_{n-1}) - f(p_{n-2})}.$$

Note. Secant method needs two initial approximations p_0 and p_1.


Secant: f(x)=x^3–3*x+1,x in [0,1]
1
0.5
0
–0.5
–1
0.2
0.4
0.6
0.8
1
x

Method of False Position

- Both Newton's method and Secant method have the limitation that they may diverge when the initial guesses are not sufficiently close to the root.

- Bisection method uses the idea of bracketing the root at each step to ensure convergence.

- The method of false position, to some extent, is a combination of the bisection method and the secant method.

Given $f \in C[p_0, p_1]$, $f(p_0) \cdot f(p_1) < 0$:

secant: $(p_0, p_1) \longrightarrow (p_1, p_2) \longrightarrow \cdots$

false position:

$$(p_0, p_1) \xrightarrow{\text{secant}} p_2 \begin{cases} (p_0, p_2), & f(p_0) \cdot f(p_2) < 0 \\ (p_1, p_2), & f(p_1) \cdot f(p_2) < 0 \end{cases} \longrightarrow \cdots$$

- This method nicely illustrates how bracketing a root can be used to develop a more sophisticated root finding method.
- In terms of performance, the Method of False Position is often slightly slower than the secant method since it pays for some extra insurance in finding a root. It is often (but not always) faster than the Bisection Method.

Error Analysis

Definition 2 *Suppose* $\{p_n\}_{n=0}^{\infty}$ *is a sequence that converges to* p*, with* $p_n \neq p$ *for all* n*. If positive constant* λ *and* α *exist with*

$$\lim_{n\to\infty} \frac{|p_{n+1} - p|}{|p_n - p|^{\alpha}} = \lambda, \tag{4}$$

then $\{p_n\}_{n=0}^{\infty}$ *converges to* p *of order* α*, with asymptotic error constant* λ*.*

- A sequence with a higher order of convergence converges more rapidly than a sequence with a lower order.
- The constant λ affects the speed of convergence, but is not so important as the order α.

Two common cases:

- If $\alpha = 1$, the sequence is *linearly* convergent;
- If $\alpha = 2$, the sequence is *quadratically* convergent.

Example 3

Suppose

$\{p_n\}_{n=0}^{\infty}$ *linearly* convergent to 0,

$\{\tilde{p}_n\}_{n=0}^{\infty}$ *quadratically* convergent to 0

with

$$\lim_{n\to\infty} \frac{|p_{n+1}|}{|p_n|} = \lim_{n\to\infty} \frac{|\tilde{p}_{n+1}|}{|\tilde{p}_n|^2} = 0.5.$$

For simplicity, suppose that

$$\frac{|p_{n+1}|}{|p_n|} \approx 0.5 \quad \text{and} \quad \frac{|\tilde{p}_{n+1}|}{|\tilde{p}_n|^2} \approx 0.5. \tag{5}$$

By (5),

$$|p_n - 0| = |p_n| \approx 0.5|p_{n-1}| \approx (0.5)^2|p_{n-2}| \approx \cdots \approx (0.5)^n|p_0|,$$

$$\begin{aligned}|\tilde{p}_n - 0| = |\tilde{p}_n| &\approx 0.5|\tilde{p}_{n-1}|^2 \approx (0.5)(0.5|\tilde{p}_{n-2}|^2)^2 = (0.5)^3|\tilde{p}_{n-2}|^4 \\ &\approx (0.5)^3(0.5|\tilde{p}_{n-3}|^2)^4 = (0.5)^7|\tilde{p}_{n-3}|^8 \\ &\approx \cdots \approx (0.5)^{2^n-1}|\tilde{p}_0|^{2^n}.\end{aligned}$$

For $|p_0| = |\tilde{p}_0| = 1$,

n	$\{p_n\}_{n=0}^{\infty}$	$\{\tilde{p}_n\}_{n=0}^{\infty}$
1	5.0000×10^{-1}	5.0000×10^{-1}
2	2.5000×10^{-1}	1.2500×10^{-1}
3	1.2500×10^{-1}	7.8125×10^{-3}
4	6.2500×10^{-2}	3.0518×10^{-5}
5	3.1250×10^{-2}	4.6566×10^{-10}
6	1.5625×10^{-2}	1.0842×10^{-19}
7	7.8125×10^{-3}	5.8775×10^{-39}

Convergence of Fixed-Point Iteration

Theorem 7 *Let $g \in C[a,b]$ be such that $g(x) \in [a,b]$, for all $x \in [a,b]$. Suppose, in addition, that g' is continuous on (a,b) and a positive constant $k < 1$ exists with*

$$|g'(k)| \leq k, \quad \textit{for all } \; x \in (a,b).$$

If $g'(p) \neq 0$, then for any number p_0 in $[a,b]$, the sequence

$$p_n = g(p_{n-1}), \quad \textit{for } n \geq 1,$$

converges only linearly *to the unique fixed point p in $[a,b]$.*

Proof. By Theorem 5, the sequence p_n converges to p.

$$\left.\begin{array}{l} g \in C[a,b] \\ g' \in C(a,b) \end{array}\right\} \overset{\text{MVT}}{\Longrightarrow} p_{n+1} - p = g(p_n) - g(p) = g'(\xi_n)(p_n - p)$$

where ξ_n is between p_n and p.

Since $\{p_n\}_{n=0}^{\infty} \longrightarrow p$, $\{\xi_n\}_{n=0}^{\infty} \longrightarrow p$.
Since g' is continuous on (a,b),

$$\lim_{n\to\infty} g'(\xi_n) = g'(\lim_{n\to\infty} \xi_n) = g'(p).$$

Hence,

$$\lim_{n\to\infty} \frac{p_{n+1} - p}{p_n - p} = \lim_{n\to\infty} g'(\xi_n) = g'(p) \quad \text{and} \quad \lim_{n\to\infty} \frac{|p_{n+1} - p|}{|p_n - p|} = |g'(p)|.$$

$\blacksquare$

The following theorem describes additional conditions that ensure the quadratic convergence.

Theorem 8 *Let p be a solution of the equation $x = g(x)$. Suppose that $g'(p) = 0$ and g'' is continuous and strictly bounded by M on an open interval I containing p. Then there exists a $\delta > 0$ such that, for $p_0 \in [p - \delta, p + \delta]$, the sequence defined by $p_n = g(p_{n-1})$, when $n \geq 1$, converges* at least quadratically *to p. Moreover, for sufficiently large values of n,*

$$|p_{n+1} - p| < \frac{M}{2} |p_n - p|^2.$$

Proof. Choose $k \in (0,1)$. It follows from the proof of Thm 6 that

(1) $\exists \delta > 0$ s.t. on the interval $[p-\delta, p+\delta] \subseteq I$, $|g'(x)| \leq k < 1$ and g'' is continuous.

(2) the terms of the sequence $\{p_n\}_{n=0}^{\infty}$ are contained in $[p-\delta, p+\delta]$.

Expanding $g(x)$ in a linear Taylor polynomial for $x \in [p-\delta, p+\delta]$ gives

$$g(x) = g(p) + g'(p)(x-p) + \frac{g''(\xi)}{2}(x-p)^2,$$

where ξ lies between x and p.

$$\left.\begin{array}{l} g(p) = p \\ g'(p) = 0 \end{array}\right\} \Longrightarrow g(x) = p + \frac{g''(\xi)}{2}(x-p)^2.$$

When $x = p_n$,

$$p_{n+1} = g(p_n) = p + \frac{g''(\xi_n)}{2}(p_n - p)^2,$$

with ξ_n between p_n and p. Thus,

$$p_{n+1} - p = \frac{g''(\xi_n)}{2}(p_n - p)^2.$$

$$\left.\begin{array}{l} |g'(x)| \leq k < 1 \text{ on } [p-\delta, p+\delta] \\ g : [p-\delta, p+\delta] \longrightarrow [p-\delta, p+\delta] \end{array}\right\} \overset{\text{FPT}}{\Longrightarrow} \{p_n\}_{n=0}^{\infty} \longrightarrow p.$$

Since ξ_n is between p and p_n for each n, $\{\xi_n\}_{n=0}^{\infty} \longrightarrow p$. Hence,

$$\lim_{n\to\infty} \frac{|p_{n+1} - p|}{|p_n - p|^2} = \frac{|g''(p)|}{2}.$$

This implies that the sequence $\{p_n\}_{n=0}^{\infty}$ is quadratically convergent if $g''(p) \neq 0$ and of higher-order convergence if $g''(p) = 0$.

Since g'' is continuous, and strictly bounded by M on $[p-\delta, p+\delta]$, for sufficiently large values of n:

$$|p_{n+1} - p| < \frac{M}{2}|p_n - p|^2.$$

- The idea behind finding iteration methods with a high order of convergence is to look for schemes whose derivatives are zero at the fixed point.

Example 4 Newton's Method.

$$
\begin{aligned}
g(x) &= x - \frac{f(x)}{f'(x)} \\
g'(x) &= \frac{f(x)f''(x)}{(f'(x))^2}
\end{aligned}
$$

Notice that $g'(p) = 0$ provided that $f'(p) \neq 0$. Hence, Newton's Method satisfies the derivative condition.

Example 5 Use Newton's Method to find the roots of

$$f(p) = p^3 - p^2 - p + 1 = 0.$$

Applying the formula

$$p_{n+1} = p_n - \frac{f(p_n)}{f'(p_n)} = p_n - \frac{p_n^3 - p_n^2 - p_n + 1}{3p_n^2 - 2p_n - 1}$$

starting from $p_0 = 1.1$, we find

p_0	p_1	p_2	p_3	p_4	p_5
1.1	$1.05116\cdots$	$1.02589\cdots$	$1.01303\cdots$	$1.00653\cdots$	$1.00327\cdots$

which is very slow (LINEAR!) convergence to the root (which is 1).
WHY ?

In Newton's Method we need $f'(p) \neq 0$ to obtain quadratic convergence. Notice that

$$f'(p) = 3p^2 - 2p - 1|_{p=1} = 0.$$

Hence, the theorem does not hold. Moreover, factoring f

$$f(x) = (x-1)^2(x+1),$$

we see that $x = 1$ is a zero of multiplicity 2.

Definition 3 *A solution p of $f(x) = 0$ is a zero of multiplicity m of f if for $x \neq p$, we can write $f(x) = (x-p)^m q(x)$, where $\lim_{x\to p} q(x) \neq 0$.*

- *Simple zeros* are those that have multiplicity one.
- Newton's Method can only be applied to simple zeros of a function.

• Identification of the multiplicity of a zero is often made easier by the following two theorems:

Theorem 9 *$f \in C^1[a,b]$ has a simple zero at p in (a,b) if and only if $f(p) = 0$, but $f'(p) \neq 0$.*

Theorem 10 *The function $f \in C^m[a,b]$ has a zero of multiplicity m at p in (a,b) if and only if*

$$0 = f(p) = f'(p) = f''(p) = \cdots = f^{(m-1)}(p), \quad \text{but} \;\; f^{(m)}(p) \neq 0.$$

• We want to obtain quadratic convergence with Newton's Method for multiple roots. One approach is to define a function

$$\mu(x) = \frac{f(x)}{f'(x)}.$$

Assume that p is a zero of multiplicity m, i.e., $f(x) = (x-p)^m q(x)$ where $q(p) \neq 0$. Then

$$\mu(x) = (x-p)\,\frac{q(x)}{mq(x) + q'(x)(x-p)}.$$

Hence, $\mu(p)$=0, but

$$\frac{q(p)}{mq(p) + q'(p)(p-p)} = \frac{1}{m} \neq 0,$$

and p is a zero of multiplicity 1 of $\mu(x)$.

Substituting $\mu(x)$ into Newton's Method gives the iteration function

$$g(x) = x - \frac{\mu(x)}{\mu'(x)} = x - \frac{f(x)f'(x)}{(f'(x))^2 - f(x)f''(x)}. \tag{6}$$

Advantages. Provided g satisfies the necessary continuity conditions, we will get quadratic convergence regardless of the multiplicity of the zero of f.

Disadvantages. (1) Need f''; (2) More calculations to evaluate g; (3) Possibility of serious cancellation in the denominator.

Example 6 Back to finding a root of $p^3 - p^2 - p + 1 = 0$ (See Example 5). Applying (6) with $p_0 = 1.1$ yields

$$p_1 = 0.997735\cdots, \quad p_2 = 0.999999\cdots$$

Accelerating Convergence

• Given a linearly convergent sequence, we want to speed up convergence.

Aitken's Method. Assume that the sign of $p_n - p$, $p_{n+1} - p$, and $p_{n+2} - p$ agree, and that n is sufficiently large so that

$$\frac{p_{n+1} - p}{p_n - p} \approx \frac{p_{n+2} - p}{p_{n+1} - p}.$$

Then

$$\begin{aligned} (p_{n+1} - p)^2 &\approx (p_n - p)(p_{n+2} - p) \\ p_{n+1}^2 - 2p_{n+1}p + p^2 &\approx p_n p_{n+2} - p_n p - p p_{n+2} + p^2 \\ (p_{n+2} - 2p_{n+1} + p_n)p &\approx p_n p_{n+2} - p_{n+1}^2. \end{aligned}$$

Therefore, $p \approx \dfrac{p_n p_{n+2} - p_{n+1}^2}{p_{n+2} - 2p_{n+1} + p_n}$. Adding and subtracting p_n^2 and $2p_n p_{n+1}$ to the rhs and grouping terms appropriately yields

$$p \approx p_n - \left(\frac{p_{n+1}^2 - 2p_n p_{n+1} + p_n^2}{p_{n+2} - 2p_{n+1} + p_n} \right).$$

The corresponding sequence

$$\hat{p}_n = p_n - \left(\frac{p_{n+1}^2 - 2p_n p_{n+1} + p_n^2}{p_{n+2} - 2p_{n+1} + p_n} \right) = p_n - \frac{(\Delta p_n)^2}{\Delta^2 p_n} \qquad (7)$$

is known as Aitken's Method.

Remark 3 Aitken's Method constructs the terms in order

$$p_0, \;\; p_1 = g(p_0), \;\; p_2 = g(p_1), \;\; \hat{p}_0 = \{\Delta^2\}(p_0),$$

$$p_3 = g(p_2), \;\; \hat{p}_1 = \{\Delta^2\}(p_1)$$

where $\{\Delta^2\}$ indicates that (7) is used.

Theorem 11 *Suppose that* $\{p_n\}_{n=0}^{\infty}$ *is a sequence that converges linearly to the limit* p *and that for all sufficiently large values of* n *we have* $(p_n - p)(p_{n+1} - p) > 0$*. Then the sequence* $\{\hat{p}_n\}_{n=0}^{\infty}$ *converges to* p *faster than* $\{p_n\}_{n=0}^{\infty}$ *in the sense that*

$$\lim_{n\to\infty} \frac{\hat{p}_n - p}{p_n - p} = 0.$$

Steffensen's Method. A modification of Aitken's Method. In Steffensen's Method, one restarts the iteration with the improved value as soon as it becomes available.

The method constructs the same first four terms $p_0, p_1, p_2, \hat{p}_0$ as in Aitken's Method. However, it then assumes that $\hat{p}_0$ is a better approximation to p than is p_2, and applies fixed-point iteration to $\hat{p}_0$, instead of p_2. The sequence generated is

$$p_0^{(0)}, p_1^{(0)} = g(p_0^{(0)}), p_2^{(0)} = g(p_1^{(0)}), p_0^{(1)} = \{\Delta^2\}(p_0^{(0)}), p_1^{(1)} = g(p_0^{(1)}), \ldots$$

Theorem 12 *Suppose that $x = g(x)$ has the solution p with $g'(p) \neq 1$. If there exists a $\delta > 0$ such that $g \in C^3[p-\delta, p+\delta]$, then Steffensen's method gives quadratic convergence for any $p_0 \in [p-\delta, p+\delta]$.*

Zeros of Polynomials

- A polynomial $P(x) \in \mathbb{C}[x]$ of degree n has the form

$$P(x) = \sum_{i=0}^{n} a_i x^i = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

where $a_i \in \mathbb{C}$, $a_n \neq 0$.

- We want to compute the zeros of polynomials.

Theorem 13 *Fundamental Theorem of Algebra. If $P(x)$ is a polynomial of degree $n \geq 1$ with real or complex coefficients, then $P(x) = 0$ has at lease one (possibly complex) root.*

Corollary 2 *If $P(x)$ is a polynomial of degree $n \geq 1$ with real or complex coefficients, then there exist unique constants $x_1, x_2, \ldots, x_k$, possible complex, and unique integers $m_1, m_2, \ldots, m_k$ such that $\sum_{i=1}^{k} m_i = n$ and*

$$P(x) = a_n (x - x_1)^{m_1} (x - x_2)^{m_2} \cdots (x - x_k)^{m_k}.$$

Corollary 3 *Let $P(x)$ and $Q(x)$ be polynomials of degree at most n. If $x_1, x_2, \ldots, x_k$, with $k > n$, are distinct numbers with $P(x_i) = Q(x_i)$ for $i = 1, 2, \ldots, k$, then $P(x) = Q(x)$ for all values of x.*

- We want to use Newton's Method to locate the approximate zeros of P. It will be necessary to evaluate P and its derivative at specified values. We now direct our attention to efficient methods for this task.

Theorem 14 *Horner's Method. Let*

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

If $b_n = a_n$ *and*

$$b_k = a_k + b_{k+1} x_0, \quad \textit{for} \;\; k = n-1, n-2, \ldots, 1, 0,$$

then $b_0 = P(x_0)$*. Moreover, if*

$$Q(x) = b_n x^{n-1} + b_{n-1} x^{n-2} + \cdots + b_2 x + b_1,$$

then

$$P(x) = (x - x_0) Q(x) + b_0.$$

Example 7 Use Horner's method to evaluate

$$P(x) = 2x^4 - 3x^2 + 3x - 4$$

at $x_0 = -2$.

$x_0 = -2$	$a_4 = 2$	$a_3 = 0$	$a_2 = -3$	$a_1 = 3$	$a_0 = -4$
		$b_4 x_0 = -4$	$b_3 x_0 = 8$	$b_2 x_0 = -10$	$b_1 x_0 = 14$
	$b_4 = 2$	$b_3 = -4$	$b_2 = 5$	$b_1 = -7$	$b_0 = 10$

Hence,

$$P(x) = (x+2)\underbrace{(2x^3 - 4x^2 + 5x - 7)}_{Q(x)} + \underbrace{10}_{P(-2)},$$

and $P(-2) = 10$.

- For the computation of $P'(x_0)$, since $P(x) = (x - x_0)Q(x) + b_0$,

$$P'(x) = Q(x) + (x - x_0)Q'(x), \quad P'(x_0) = Q(x_0).$$

Hence, $P(x)$ and $P'(x)$ can be evaluated in the same manner.

Example 8 Compute $P'(-2)$ for the polynomial $P(x)$ in Example 7.

$x_0 = -2$	2	-4	5	-7
		-4	16	-42
	2	-8	21	-49

Hence, $Q(-2) = P'(-2) = -49$.

Deflation. A procedure for computing the real zeros of polynomials.

Let $\hat{x}_1$ be an approximate root of $P(x)$, i.e., $P(x) \approx (x - \hat{x}_1)Q_1(x)$. Using Newton's method, one computes another root of $P(x)$ by computing a root of $Q_1(x)$. This procedure is applied repeatedly.

After k steps,

$$P(x) \approx (x - \hat{x}_1)(x - \hat{x_2}) \cdots (x - \hat{x}_k)Q_k(x).$$

Problem. Inaccurate results. An approximate zero $\hat{x}_{k+1}$ of Q_k does not in general approximate a root of $P(x) = 0$.

A cure. Use the reduced equations to find approximations $\hat{x}_2, \hat{x}_3, \ldots, \hat{x}_k$ to the zeros of P, and then improve these approximations by applying Newton's Method to the original polynomial $P(x)$.

Remark 4 If the initial approximation using Newton's Method is a real number, all subsequent approximations will also be real numbers. If the initial approximation is a complex number, and all computations are done using complex arithmetics, all subsequent approximations will also be complex numbers.

Theorem 15 *If $z = a + bi$ is a complex zero of multiplicity m of the polynomial $P(x)$ with real coefficients, then $\bar{z} = a - bi$ is also a zero of multiplicity m of the polynomial $P(x)$, and $(x^2 - 2ax + a^2 + b^2)^m$ is a factor of $P(x)$.*

Müller's Method

- An extension of the Secant method.
- Müller's method uses three initial approximations, x_0, x_1, and x_2, and determines the next approximation x_3 by considering the intersection of the x-axis with the parabola through $(x_0, f(x_0))$, $(x_1, f(x_1))$, and $(x_2, f(x_2))$.

Consider the quadratic polynomial

$$P(x) = a(x - x_2)^2 + b(x - x_2) + c$$

that passes through $(x_0, f(x_0))$, $(x_1, f(x_1))$, and $(x_2, f(x_2))$. The system

$$\begin{aligned} f(x_0) &= a(x_0 - x_2)^2 + b(x_0 - x_2) + c \\ f(x_1) &= a(x_1 - x_2)^2 + b(x_1 - x_2) + c \\ f(x_2) &= a \cdot 0^2 + b \cdot 0 + c = c \end{aligned}$$

yields

$$
\begin{aligned}
c &= f(x_2) \\
b &= \frac{(x_0 - x_2)^2(f(x_1) - f(x_2)) - (x_1 - x_2)(f(x_0) - f(x_2))}{(x_0 - x_2)(x_1 - x_2)(x_0 - x_1)} \\
a &= \frac{(x_1 - x_2)(f(x_0) - f(x_2)) - (x_0 - x_2)(f(x_1) - f(x_2))}{(x_0 - x_2)(x_1 - x_2)(x_0 - x_1)}.
\end{aligned}
$$

To determine x_3, we apply the formula

$$x_3 - x_2 = \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}.$$

In Müller's method, the sign is chosen to agree with the sign of b. Chosen in this manner, the denominator will be the largest in magnitude and will result in x_3 being selected as the closest zero of P to x_2. Thus,

$$x_3 = x_2 - \frac{2c}{b + \mathrm{sgn}(b)\sqrt{b^2 - 4ac}}.$$