Newton's Method: Universality and Geometry

Adi Ben-Israel
adi.benisrael@gmail.com

Experimental Math Seminar
Rutgers University

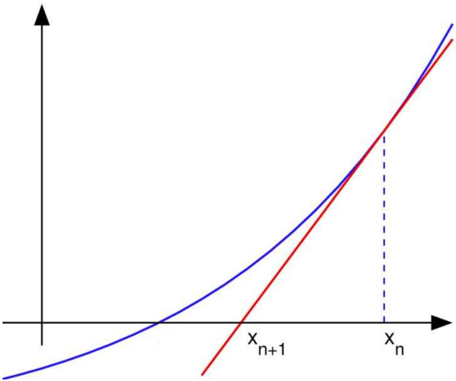
January 25, 2018


ISAAC NEWTON
1643-1727
$\Delta(mv)=F\Delta t$
DEUTSCHE
BUNDESPOST
100
1993

Outline

Newton's method for $f(x) = 0$, $f : \mathbb{R} \to \mathbb{R}$

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \ldots$$



Quadratic convergence of Newton's method

Theorem 1

Assume that f is twice differentiable on an open interval (a,b), and that there exists $x^* \in (a,b)$ with $f'(x^*) \neq 0$.
Define Newton's method by the sequence

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}, \; k = 1,2,\ldots$$

and assume that $\{x_k\}$ converges to x^* as $k \to \infty$.
Then, for k sufficiently large,

$$|x_{k+1} - x^*| \leq M\,|x_k - x^*|^2 \quad \text{if} \quad M > \frac{|f''(x^*)|}{2|f'(x^*)|}$$

*Thus, $\{x_k\}$ converges to x^* **quadratically**.*

Quadratic convergence of Newton's method

Theorem 1

Assume that f is twice differentiable on an open interval (a,b), and that there exists $x^* \in (a,b)$ with $f'(x^*) \neq 0$.
Define Newton's method by the sequence

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}, \; k = 1, 2, \ldots$$

and assume that $\{x_k\}$ converges to x^* as $k \to \infty$.
Then, for k sufficiently large,

$$|x_{k+1} - x^*| \leq M\,|x_k - x^*|^2 \quad \text{if} \quad M > \frac{|f''(x^*)|}{2|f'(x^*)|}$$

*Thus, $\{x_k\}$ converges to x^* **quadratically**.*

Quadratic convergence of Newton's method

Theorem 1

Assume that f is twice differentiable on an open interval (a,b), and that there exists $x^* \in (a,b)$ with $f'(x^*) \neq 0$. Define Newton's method by the sequence

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}, \; k = 1, 2, \ldots$$

and assume that $\{x_k\}$ converges to x^* as $k \to \infty$. Then, for k sufficiently large,

$$|x_{k+1} - x^*| \leq M\,|x_k - x^*|^2 \quad \textit{if} \quad M > \frac{|f''(x^*)|}{2|f'(x^*)|}$$

Thus, $\{x_k\}$ converges to x^* ***quadratically****.*

Newton's method for system of equations $\mathbf{f}(\mathbf{x}) = \mathbf{0}$

Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ and consider the system

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}, \quad \text{standing for} \quad f_i(x_1, \ldots, x_n) = 0,\ i = 1, \ldots, m$$

If $m = n$, the Newton iteration is

$$\mathbf{x}_+ := \mathbf{x} - (J\mathbf{f}(\mathbf{x}))^{-1}\,\mathbf{f}(\mathbf{x}), \quad J\mathbf{f} = \left(\frac{\partial f_i}{\partial x_j}\right) \text{ is the } \textbf{Jacobian}$$

If $m \neq n$ or the Jacobian is singular, ([2],[9]),

$$\mathbf{x}_+ := \mathbf{x} - (J\mathbf{f}(\mathbf{x}))^{\dagger}\,\mathbf{f}(\mathbf{x}), \quad {}^{\dagger} \text{ is the } \textbf{Moore–Penrose inverse}$$

Converging to a stationary point of $\|\mathbf{f}(\mathbf{x})\|^2 = \sum_{i=1}^{m} f_i(\mathbf{x})^2$, since

$$\nabla\,\|\mathbf{f}(\mathbf{x})\|^2 = 2\,(J\mathbf{f}(\mathbf{x}))^*\,\mathbf{f}(\mathbf{x}), \text{ and } N(A^{\dagger}) = N(A^*),\ \forall A.$$

Newton's method for system of equations $\mathbf{f}(\mathbf{x}) = \mathbf{0}$

Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ and consider the system

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}, \quad \text{standing for} \quad f_i(x_1, \ldots, x_n) = 0,\ i = 1, \ldots, m$$

If $m = n$, the Newton iteration is

$$\mathbf{x}_+ := \mathbf{x} - (J\mathbf{f}(\mathbf{x}))^{-1}\,\mathbf{f}(\mathbf{x}), \quad J\mathbf{f} = \left(\frac{\partial f_i}{\partial x_j}\right) \text{ is the } \textbf{Jacobian}$$

If $m \neq n$ or the Jacobian is singular, ([2],[9]),

$$\mathbf{x}_+ := \mathbf{x} - (J\mathbf{f}(\mathbf{x}))^{\dagger}\,\mathbf{f}(\mathbf{x}), \quad {}^{\dagger} \text{ is the } \textbf{Moore–Penrose inverse}$$

Converging to a stationary point of $\|\mathbf{f}(\mathbf{x})\|^2 = \sum_{i=1}^{m} f_i(\mathbf{x})^2$, since

$$\nabla \|\mathbf{f}(\mathbf{x})\|^2 = 2\,(J\mathbf{f}(\mathbf{x}))^*\,\mathbf{f}(\mathbf{x}), \text{ and } N(A^{\dagger}) = N(A^*),\ \forall A.$$

Newton's method for system of equations $\mathbf{f}(\mathbf{x}) = \mathbf{0}$

Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ and consider the system

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}, \quad \text{standing for} \quad f_i(x_1, \ldots, x_n) = 0,\ i = 1, \ldots, m$$

If $m = n$, the Newton iteration is

$$\mathbf{x}_+ := \mathbf{x} - (J\mathbf{f}(\mathbf{x}))^{-1}\mathbf{f}(\mathbf{x}), \quad J\mathbf{f} = \left(\frac{\partial f_i}{\partial x_j}\right) \text{ is the **Jacobian**}$$

If $m \neq n$ or the Jacobian is singular, ([2],[9]),

$$\mathbf{x}_+ := \mathbf{x} - (J\mathbf{f}(\mathbf{x}))^{\dagger}\mathbf{f}(\mathbf{x}), \quad {}^{\dagger} \text{ is the **Moore–Penrose inverse**}$$

Converging to a stationary point of $\|\mathbf{f}(\mathbf{x})\|^2 = \sum_{i=1}^{m} f_i(\mathbf{x})^2$, since

$$\nabla \|\mathbf{f}(\mathbf{x})\|^2 = 2\,(J\mathbf{f}(\mathbf{x}))^*\,\mathbf{f}(\mathbf{x}), \text{ and } N(A^{\dagger}) = N(A^*),\ \forall A.$$

Outline

The Newton transform and its inverse [4]

Definition 2

Let $f, u : \mathbb{R} \to \mathbb{R}$, and let f be differentiable in $S \subset \mathbb{R}$. If

$$u(x) = x - \frac{f(x)}{f'(x)}, \quad \forall\, x \in S,$$

then:

(a) u is the **Newton transform** of f in S, denoted $u = \mathbf{N}f$, and

(b) f is called the **inverse Newton transform** of u in S, denoted $f = \mathbf{N}^{-1}u$.

The Newton transform and its inverse [4]

Definition 2

Let $f, u : \mathbb{R} \to \mathbb{R}$, and let f be differentiable in $S \subset \mathbb{R}$. If

$$u(x) = x - \frac{f(x)}{f'(x)}, \quad \forall\, x \in S,$$

then:

(a) u is the **Newton transform** of f in S, denoted $u = \mathbf{N}f$, and

(b) f is called the **inverse Newton transform** of u in S, denoted $f = \mathbf{N}^{-1}u$.

The **Newton transform** $\mathbf{N}f$ of a differentiable $f : \mathbb{R} \to \mathbb{R}$ is

$$(\mathbf{N}f)(x) := x - \frac{f(x)}{f'(x)}.$$

Maple

```
Newton:=proc(f,x); x-f/diff(f,x); end:
```

Example 3

$\left(\mathbf{N}(x^{2/3}-1)^{3/2}\right)(x) = x^{1/3}$

`simplify(Newton(`$(x^{(2/3)}-1)^{(3/2)},x)$ $\longmapsto$ $x^{1/3}$

Example 4

$(\mathbf{N}\exp\{-x\})(x) = x+1$

`Newton(exp(-x),x)` $\longmapsto$ $x+1$

The Newton transform

The **Newton transform** $\mathbf{N}f$ of a differentiable $f : \mathbb{R} \to \mathbb{R}$ is

$$(\mathbf{N}f)(x) := x - \frac{f(x)}{f'(x)}.$$

Maple

```
Newton:=proc(f,x); x-f/diff(f,x); end:
```

Example 3

$\left(\mathbf{N}(x^{2/3}-1)^{3/2}\right)(x) = x^{1/3}$

simplify(Newton($(x^{(2/3)}-1)^{(3/2)}, x$) $\longmapsto$ $x^{1/3}$

Example 4

$(\mathbf{N}\exp\{-x\})(x) = x+1$

Newton(exp(-x),x) $\longmapsto$ $x+1$

The **Newton transform** $\mathbf{N}f$ of a differentiable $f : \mathbb{R} \to \mathbb{R}$ is

$$(\mathbf{N}f)(x) := x - \frac{f(x)}{f'(x)}.$$

Maple

```
Newton:=proc(f,x); x-f/diff(f,x); end:
```

Example 3

$\left(\mathbf{N}(x^{2/3}-1)^{3/2}\right)(x) = x^{1/3}$

`simplify(Newton(`$(x^{(2/3)}-1)^{(3/2)}, x)$ $\longmapsto$ $x^{1/3}$

Example 4

$(\mathbf{N}\exp\{-x\})(x) = x+1$

`Newton(exp(-x),x)` $\longmapsto$ $x+1$

The Newton transform

The **Newton transform** $\mathbf{N}f$ of a differentiable $f:\mathbb{R}\to\mathbb{R}$ is

$$(\mathbf{N}f)(x) := x - \frac{f(x)}{f'(x)}.$$

Maple

```
Newton:=proc(f,x); x-f/diff(f,x); end:
```

Example 3

$\left(\mathbf{N}(x^{2/3}-1)^{3/2}\right)(x) = x^{1/3}$

`simplify(Newton(`$(x^{(2/3)}-1)^{(3/2)},x)$ $\longmapsto$ $x^{1/3}$

Example 4

$(\mathbf{N}\exp\{-x\})(x) = x+1$

`Newton(exp(-x),x)` $\longmapsto$ $x+1$

The **Newton transform** $\mathbf{N}f$ of a differentiable $f : \mathbb{R} \to \mathbb{R}$ is

$$(\mathbf{N}f)(x) := x - \frac{f(x)}{f'(x)}.$$

Maple

```
Newton:=proc(f,x); x-f/diff(f,x); end:
```

Example 3

$\left(\mathbf{N}(x^{2/3} - 1)^{3/2}\right)(x) = x^{1/3}$

`simplify(Newton(`$(x^{(2/3)} - 1)^{(3/2)}, x)$ $\longmapsto$ $x^{1/3}$

Example 4

$(\mathbf{N}\exp\{-x\})(x) = x + 1$

`Newton(exp(-x),x)` $\longmapsto$ $x+1$

The **Newton transform** $\mathbf{N}f$ of a differentiable $f : \mathbb{R} \to \mathbb{R}$ is

$$(\mathbf{N}f)(x) := x - \frac{f(x)}{f'(x)}.$$

Maple

```
Newton:=proc(f,x); x-f/diff(f,x); end:
```

Example 3

$$\left(\mathbf{N}(x^{2/3}-1)^{3/2}\right)(x) = x^{1/3}$$

`simplify(Newton(`$(x^{(2/3)}-1)^{(3/2)}, x)$ $\longmapsto$ $x^{1/3}$

Example 4

$$(\mathbf{N}\exp\{-x\})(x) = x+1$$

`Newton(exp(-x),x)` $\longmapsto$ $x+1$

The iteration $u(x) := x^2$

Example 5

$$\left(\mathbf{N}\left(\frac{x}{x-1}\right)\right)(x) = x^2$$

`Newton(x/(x-1),x)` $\longmapsto \quad x^2$

The iteration

$$x_+ := x^2$$

generates the same sequence as

$$x_+ := \mathbf{N}\left(\frac{x}{x-1}\right)$$

away from $x = 1$.

The iteration $u(x) := x^2$

Example 5

$$\left(\mathbf{N}\left(\frac{x}{x-1}\right)\right)(x) = x^2$$

`Newton(x/(x-1),x)` $\longmapsto$ x^2

The iteration

$$x_+ := x^2$$

generates the same sequence as

$$x_+ := \mathbf{N}\left(\frac{x}{x-1}\right)$$

away from $x = 1$.

The iteration $u(x) := x^2$

Example 5

$$\left(\mathbf{N}\left(\frac{x}{x-1}\right)\right)(x) = x^2$$

```
Newton(x/(x-1),x)
```
$\longmapsto \quad x^2$

The iteration

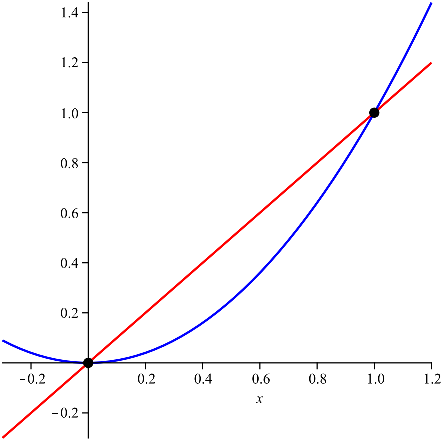
$$x_+ := x^2$$

generates the same sequence as

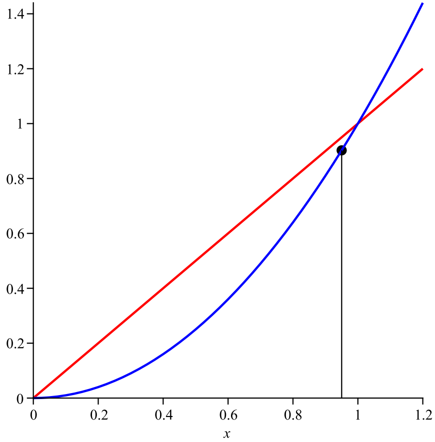
$$x_+ := \mathbf{N}\left(\frac{x}{x-1}\right)$$

away from $x = 1$.

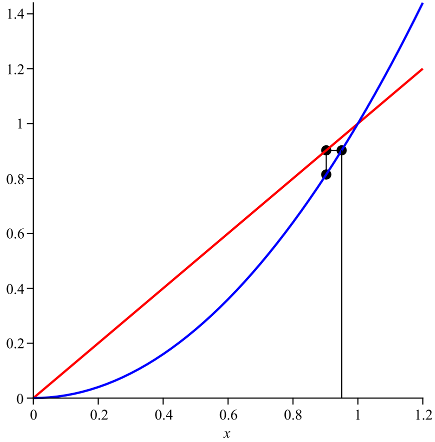
The fixed points of $u(x) := x^2$ are $\{0, 1\}$



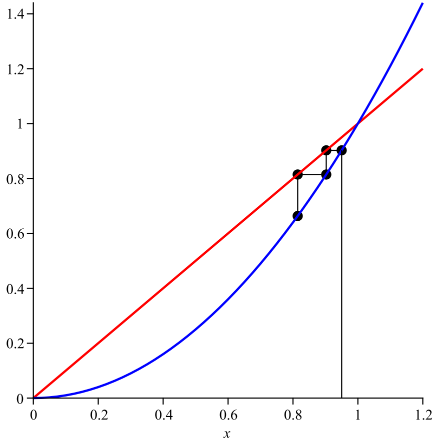
$x = 1$ is a repelling fixed point



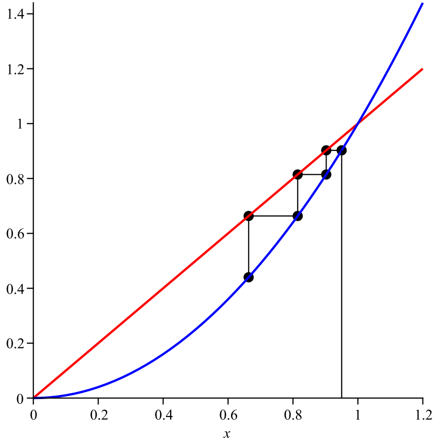
$x = 1$ is a repelling fixed point



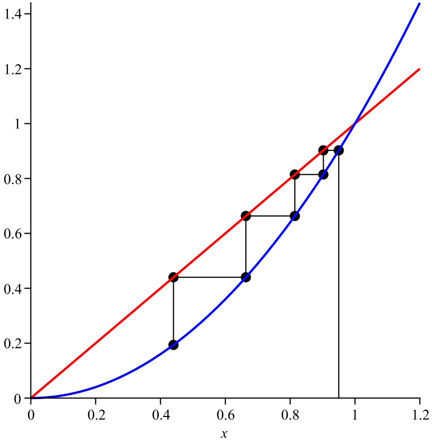
$x = 1$ is a repelling fixed point



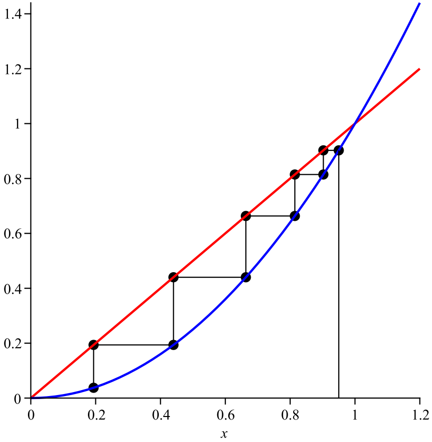
$x = 1$ is a repelling fixed point



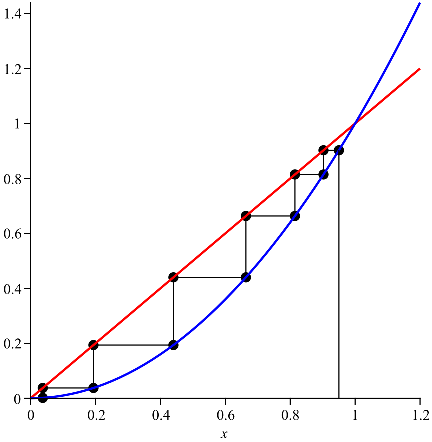
$x = 1$ is a repelling fixed point



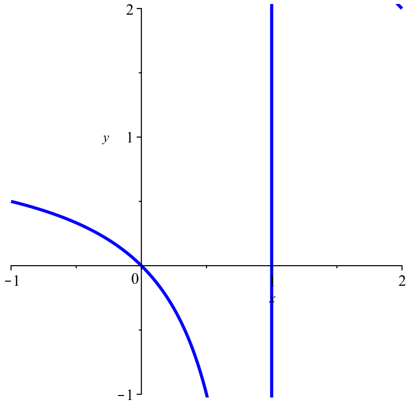
$x = 1$ is a repelling fixed point



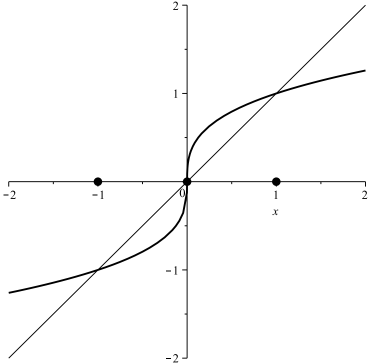
$x = 0$ is an attracting fixed point

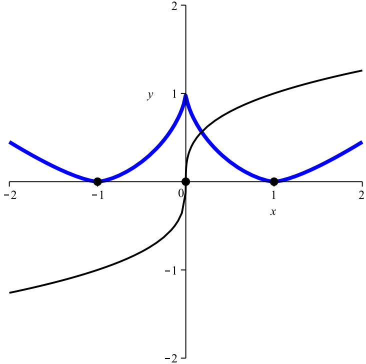


$u(x) = x^2, \quad f(x) = \left(\mathbf{N}^{-1}u\right)(x) = \dfrac{x}{x-1}$



$\mathbf{u(x) := x^{1/3},\ f(x) = \left(N^{-1}u\right)(x) = (x^{2/3} - 1)^{3/2}}$





Fixed points of $u(x)$ at $0, \pm 1$ Corresponding points of $f(x)$

Questions:

Fixed points $\{u\} \stackrel{?}{=}$ Zeros $\{f\} \cup$ Singularities $\{f'\}$

Attracting fixed points $\{u\} \stackrel{?}{=}$ Zeros $\{f\}$

Quadratic convergence of $u = \mathbf{N}f$?

$f(x) = \left(\mathbf{N}^{-1}u\right)(x)$

Questions:

Fixed points $\{u\} \stackrel{?}{=}$ Zeros $\{f\} \cup$ Singularities $\{f'\}$

Attracting fixed points $\{u\} \stackrel{?}{=}$ Zeros $\{f\}$

Quadratic convergence of $u = \mathrm{N}f$?

$f(x) = \left(\mathbf{N}^{-1}u\right)(x)$

Questions:

Fixed points $\{u\} \stackrel{?}{=}$ Zeros $\{f\} \cup$ Singularities $\{f'\}$

Attracting fixed points $\{u\} \stackrel{?}{=}$ Zeros $\{f\}$

Quadratic convergence of $u = \mathbf{N}f$?

$\mathbf{u} = \mathbf{N}f$

(a) If f is twice differentiable, then

$$u'(x) = \frac{f(x)f''(x)}{f'(x)^2}.$$

ζ is a **zero** of **order** $m > 0$ of f if

$$f(x) = (x-\zeta)^m g(x), \quad g(\zeta) \neq 0$$

(b) If ζ is a zero of f of order m, then

$$u'(x) = \frac{m(m-1)g(x)^2 + 2(x-\zeta)g'(x) + (x-\zeta)^2 g''(x)}{m^2 g(x)^2 + 2(x-\zeta)mg(x)g'(x) + (x-\zeta)^2 g'(x)^2} \longrightarrow \frac{m-1}{m},$$

as $x \to \zeta$, provided $\lim_{x\to\zeta}(x-\zeta)g'(x) = \lim_{x\to\zeta}(x-\zeta)^2 g''(x) = 0$.

(c) If ζ is a zero of f of order $m < 1$, then f is not differentiable at ζ, but u may be, with $u'(\zeta) = \dfrac{m-1}{m}$.

$\mathbf{u} = \mathbf{N}\mathbf{f}$

(a) If f is twice differentiable, then

$$u'(x) = \frac{f(x)f''(x)}{f'(x)^2}.$$

ζ is a **zero** of **order** $m > 0$ of f if

$$f(x) = (x-\zeta)^m g(x), \quad g(\zeta) \neq 0$$

(b) If ζ is a zero of f of order m, then

$$u'(x) = \frac{m(m-1)g(x)^2 + 2(x-\zeta)g'(x) + (x-\zeta)^2 g''(x)}{m^2 g(x)^2 + 2(x-\zeta)mg(x)g'(x) + (x-\zeta)^2 g'(x)^2} \longrightarrow \frac{m-1}{m},$$

as $x \to \zeta$, provided $\lim_{x\to\zeta}(x-\zeta)g'(x) = \lim_{x\to\zeta}(x-\zeta)^2 g''(x) = 0$.

(c) If ζ is a zero of f of order $m < 1$, then f is not differentiable at ζ, but u may be, with $u'(\zeta) = \dfrac{m-1}{m}$.

$\mathbf{u} = \mathbf{N}\mathbf{f}$

(a) If f is twice differentiable, then

$$u'(x) = \frac{f(x)f''(x)}{f'(x)^2}.$$

ζ is a **zero** of **order** $m > 0$ of f if

$$f(x) = (x-\zeta)^m g(x), \quad g(\zeta) \neq 0$$

(b) If ζ is a zero of f of order m, then

$$u'(x) = \frac{m(m-1)g(x)^2 + 2(x-\zeta)g'(x) + (x-\zeta)^2 g''(x)}{m^2 g(x)^2 + 2(x-\zeta)mg(x)g'(x) + (x-\zeta)^2 g'(x)^2} \longrightarrow \frac{m-1}{m},$$

as $x \to \zeta$, provided $\lim_{x\to\zeta}(x-\zeta)g'(x) = \lim_{x\to\zeta}(x-\zeta)^2 g''(x) = 0$.

(c) If ζ is a zero of f of order $m < 1$, then f is not differentiable at ζ, but u may be, with $u'(\zeta) = \dfrac{m-1}{m}$.

$\mathbf{u} = \mathbf{N f}$

(a) If f is twice differentiable, then

$$u'(x) = \frac{f(x)f''(x)}{f'(x)^2}.$$

ζ is a **zero** of **order** $m > 0$ of f if

$$f(x) = (x-\zeta)^m g(x), \quad g(\zeta) \neq 0$$

(b) If ζ is a zero of f of order m, then

$$u'(x) = \frac{m(m-1)g(x)^2 + 2(x-\zeta)g'(x) + (x-\zeta)^2 g''(x)}{m^2 g(x)^2 + 2(x-\zeta)mg(x)g'(x) + (x-\zeta)^2 g'(x)^2} \longrightarrow \frac{m-1}{m},$$

as $x \to \zeta$, provided $\lim\limits_{x\to\zeta}(x-\zeta)g'(x) = \lim\limits_{x\to\zeta}(x-\zeta)^2 g''(x) = 0$.

(c) If ζ is a zero of f of order $m < 1$, then f is not differentiable at ζ, but u may be, with $u'(\zeta) = \dfrac{m-1}{m}$.

$u = \mathbf{N}f$. Fixed points $\{u\}$ = Zeros $\{f\}$

Theorem 6

Let f be differentiable at ζ, and in (a)–(c), $f'(\zeta) \neq 0$.

(a) ζ is a zero of f if, and only if, it is a fixed point of u.

(b) If ζ is a zero of f, f and u are twice differentiable at ζ, then ζ is a superattracting fixed point of u, and convergence is (at least) quadratic.

(c) If ζ is a zero of f of order $m > \frac{1}{2}$, and u is continuously differentiable at ζ, then ζ is an attracting fixed point of u.

(d) Let ζ have a neighborhood where u and f are continuously differentiable, and $f'(x) \neq 0$ except possibly at $x = \zeta$. If ζ is an attracting fixed point of u then it is a zero of f.

Theorem 6

Let f be differentiable at ζ, and in (a)–(c), $f'(\zeta) \neq 0$.

(a) ζ is a zero of f if, and only if, it is a fixed point of u.

(b) If ζ is a zero of f, f and u are twice differentiable at ζ, then ζ is a superattracting fixed point of u, and convergence is (at least) quadratic.

(c) If ζ is a zero of f of order $m > \frac{1}{2}$, and u is continuously differentiable at ζ, then ζ is an attracting fixed point of u.

(d) Let ζ have a neighborhood where u and f are continuously differentiable, and $f'(x) \neq 0$ except possibly at $x = \zeta$. If ζ is an attracting fixed point of u then it is a zero of f.

$u = \mathbf{N}f$. Fixed points $\{u\} = $ Zeros $\{f\}$

Theorem 6

Let f be differentiable at ζ, and in (a)–(c), $f'(\zeta) \neq 0$.

(a) ζ is a zero of f if, and only if, it is a fixed point of u.

(b) If ζ is a zero of f, f and u are twice differentiable at ζ, then ζ is a superattracting fixed point of u, and convergence is (at least) quadratic.

(c) If ζ is a zero of f of order $m > \frac{1}{2}$, and u is continuously differentiable at ζ, then ζ is an attracting fixed point of u.

(d) Let ζ have a neighborhood where u and f are continuously differentiable, and $f'(x) \neq 0$ except possibly at $x = \zeta$. If ζ is an attracting fixed point of u then it is a zero of f.

$u = \mathbf{N}f$. Fixed points $\{u\} = $ Zeros $\{f\}$

Theorem 6

Let f be differentiable at ζ, and in (a)–(c), $f'(\zeta) \neq 0$.

(a) ζ is a zero of f if, and only if, it is a fixed point of u.

(b) If ζ is a zero of f, f and u are twice differentiable at ζ, then ζ is a superattracting fixed point of u, and convergence is (at least) quadratic.

(c) If ζ is a zero of f of order $m > \frac{1}{2}$, and u is continuously differentiable at ζ, then ζ is an attracting fixed point of u.

(d) Let ζ have a neighborhood where u and f are continuously differentiable, and $f'(x) \neq 0$ except possibly at $x = \zeta$. If ζ is an attracting fixed point of u then it is a zero of f.

An integral form of $\mathbf{N}^{-1}$, [4]

Theorem 7

Let u be a function: $\mathbb{R} \to \mathbb{R}$, D a region where

$$\frac{1}{x - u(x)}$$

is integrable. Then in D,

$$(\mathbf{N}^{-1}u)(x) = C \cdot \mathit{exp}\left\{\int \frac{dx}{x - u(x)}\right\}, \; C \neq 0.$$

Moreover, if $C > 0$ then $\mathbf{N}^{-1}u$ is

(a) increasing if $x > u(x)$,

(b) decreasing if $x < u(x)$,

(c) convex if u is differentiable and increasing, or

(d) concave if u is differentiable and decreasing.

An integral form of $\mathbf{N}^{-1}$, [4]

Theorem 7

Let u be a function: $\mathbb{R} \to \mathbb{R}$, D a region where

$$\frac{1}{x - u(x)}$$

is integrable. Then in D,

$$(\mathbf{N}^{-1}u)(x) = C \cdot \exp\left\{ \int \frac{dx}{x - u(x)} \right\},\ C \neq 0.$$

Moreover, if $C > 0$ then $\mathbf{N}^{-1}u$ is

(a) increasing if $x > u(x)$,

(b) decreasing if $x < u(x)$,

(c) convex if u is differentiable and increasing, or

(d) concave if u is differentiable and decreasing.

An integral form of $\mathbf{N}^{-1}$, [4]

Theorem 7

Let u be a function: $\mathbb{R} \to \mathbb{R}$, D a region where

$$\frac{1}{x - u(x)}$$

is integrable. Then in D,

$$(\mathbf{N}^{-1}u)(x) = C \cdot \exp\left\{ \int \frac{dx}{x - u(x)} \right\}, \ C \neq 0.$$

Moreover, if $C > 0$ then $\mathbf{N}^{-1}u$ is

(a) increasing if $x > u(x)$,

(b) decreasing if $x < u(x)$,

(c) convex if u is differentiable and increasing, or

(d) concave if u is differentiable and decreasing.

An integral form of $\mathbf{N}^{-1}$, [4]

Theorem 7

Let u be a function: $\mathbb{R} \to \mathbb{R}$, D a region where

$$\frac{1}{x-u(x)}$$

is integrable. Then in D,

$$(\mathbf{N}^{-1}u)(x) = C \cdot \mathit{exp}\left\{\int \frac{dx}{x-u(x)}\right\}, \; C \neq 0.$$

Moreover, if $C > 0$ then $\mathbf{N}^{-1}u$ is

(a) increasing if $x > u(x)$,

(b) decreasing if $x < u(x)$,

(c) convex if u is differentiable and increasing, or

(d) concave if u is differentiable and decreasing.

An integral form of $\mathbf{N}^{-1}$, [4]

Theorem 7

Let u be a function: $\mathbb{R} \to \mathbb{R}$, D a region where

$$\frac{1}{x - u(x)}$$

is integrable. Then in D,

$$(\mathbf{N}^{-1}u)(x) = C \cdot exp\left\{\int \frac{dx}{x - u(x)}\right\},\ C \neq 0.$$

Moreover, if $C > 0$ then $\mathbf{N}^{-1}u$ is

(a) increasing if $x > u(x)$,

(b) decreasing if $x < u(x)$,

(c) convex if u is differentiable and increasing, or

(d) concave if u is differentiable and decreasing.

$(\mathbf{N}^{-1}\mathbf{u})(\mathbf{x}) = \mathbf{C} \cdot \exp\left\{ \int \frac{\mathbf{dx}}{\mathbf{x} - \mathbf{u}(\mathbf{x})} \right\}$

Assuming $x \neq u(x)$,

$$u(x) = x - \frac{f(x)}{f'(x)} \quad \Longrightarrow \quad \frac{f'(x)}{f(x)} = \frac{1}{x - u(x)}$$

$$\therefore \ \ln f(x) = \int \frac{dx}{x - u(x)} + C$$

$$\therefore \ f(x) = C \exp\left\{ \int \frac{dx}{x - u(x)} \right\}$$

without loss of generality, $C = 1$.

$$\therefore\ f'(x) = \frac{1}{x - u(x)} \exp\left\{\int \frac{dx}{x - u(x)}\right\}$$

$$\therefore\ f''(x) = \frac{u'(x)}{(x - u(x))^2} \exp\left\{\int \frac{dx}{x - u(x)}\right\}$$

$$x > u(x) \implies f'(x) > 0$$

$$u'(x) > 0 \implies f''(x) > 0$$

$$(\mathbf{N}^{-1}\mathbf{u})(\mathbf{x}) = \mathbf{C} \cdot \exp\left\{\int \frac{\mathbf{dx}}{\mathbf{x} - \mathbf{u}(\mathbf{x})}\right\}$$

$$\therefore\ f'(x) = \frac{1}{x - u(x)} \exp\left\{\int \frac{dx}{x - u(x)}\right\}$$

$$\therefore\ f''(x) = \frac{u'(x)}{(x - u(x))^2} \exp\left\{\int \frac{dx}{x - u(x)}\right\}$$

$$x > u(x) \implies f'(x) > 0$$

$$u'(x) > 0 \implies f''(x) > 0$$

$(\mathbf{N}^{-1}\mathbf{u})(\mathbf{x}) = \mathbf{C} \cdot \exp\left\{\int \frac{\mathbf{dx}}{\mathbf{x} - \mathbf{u}(\mathbf{x})}\right\}$

$$\therefore\ f'(x) = \frac{1}{x-u(x)} \exp\left\{\int \frac{dx}{x-u(x)}\right\}$$

$$\therefore\ f''(x) = \frac{u'(x)}{(x-u(x))^2} \exp\left\{\int \frac{dx}{x-u(x)}\right\}$$

$$x > u(x) \quad \Longrightarrow \quad f'(x) > 0$$

$$u'(x) > 0 \quad \Longrightarrow \quad f''(x) > 0$$

Maple

```
InverseNewton:=proc(u,x);
simplify(exp(int(1/(x-u),x)));end:
```

Examples:

`InverseNewton(Newton(`$f(x), x), x$`);`

$$f(x)$$

`Newton(InverseNewton(`$u(x), x), x$`);`

$$u(x)$$

`InverseNewton(`x^2, x`);`

$$\frac{x}{x-1}$$

Maple

```
InverseNewton:=proc(u,x);
simplify(exp(int(1/(x-u),x)));end:
```

Examples:

InverseNewton(Newton($f(x),x),x$);

$$f(x)$$

Newton(InverseNewton($u(x),x),x$);

$$u(x)$$

InverseNewton(x^2,x);

$$\frac{x}{x-1}$$

Maple

```
InverseNewton:=proc(u,x);
simplify(exp(int(1/(x-u),x)));end:
```

Examples:

`InverseNewton(Newton(`$f(x),x),x$`);`

$$f(x)$$

`Newton(InverseNewton(`$u(x),x),x$`);`

$$u(x)$$

`InverseNewton(`x^2,x`);`

$$\frac{x}{x-1}$$

Maple

```
InverseNewton:=proc(u,x);
simplify(exp(int(1/(x-u),x)));end:
```

Examples:

InverseNewton(Newton($f(x),x),x$);

$$f(x)$$

Newton(InverseNewton($u(x),x),x$);

$$u(x)$$

InverseNewton(x^2,x);

$$\frac{x}{x-1}$$

$\mathbf{u(x) = x - \frac{f(x)}{f'(x) - a(x)f(x)}}$, $\mathbf{N^{-1}u} = ?$, [3]

```
InverseNewton(x-f(x)/(diff(f(x),x)-a(x)*f(x)),x);
```

$$f(x)\,\exp\left\{-\int a(x)dx\right\}$$

For the Halley method

$$H(x) := x - \frac{f(x)}{f'(x) - \frac{f''(x)f(x)}{2f'(x)}}$$

$$\left(\mathbf{N}^{-1}H\right)(x) = \frac{f(x)}{\sqrt{f'x}}$$

```
InverseNewton(x-f(x)/(diff(f(x),x)-a(x)*f(x)),x);
```

$$f(x)\,\exp\left\{-\int a(x)dx\right\}$$

For the **Halley method**

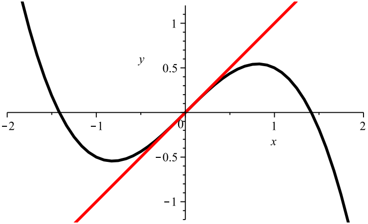
$$H(x) := x - \frac{f(x)}{f'(x) - \dfrac{f''(x)f(x)}{2f'(x)}}$$

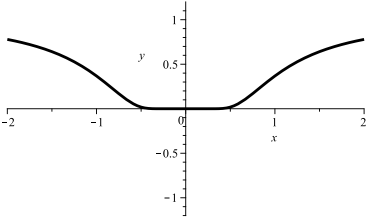
$$\left(\mathbf{N}^{-1}H\right)(x) = \frac{f(x)}{\sqrt{f'x}}$$

Example of slow convergence

Example 8

$$\mathbf{u}(\mathbf{x}) := \mathbf{x} - \frac{\mathbf{1}}{\mathbf{2}}\mathbf{x}^3, \quad \mathbf{f}(\mathbf{x}) = (\mathbf{N}^{-1}\mathbf{u})(\mathbf{x}) = \exp\left\{-\frac{\mathbf{1}}{\mathbf{x}^2}\right\}$$



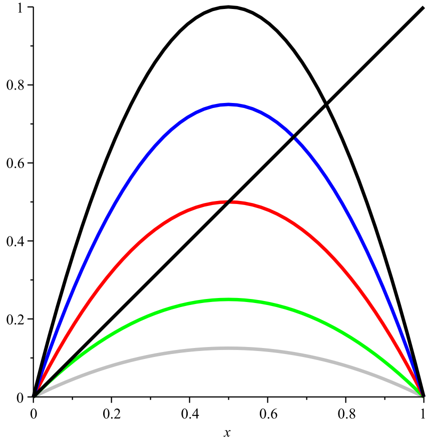


u has attracting fixed point at 0 $\qquad f^{(k)}(0) = 0, \ \forall\, k$

Outline

The logistic iteration

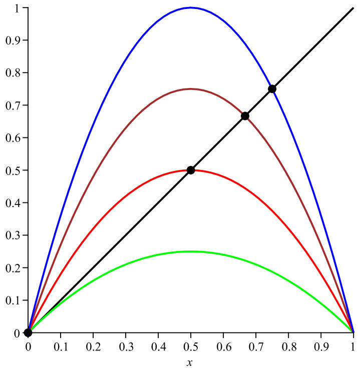
$$u(x) = \mu\, x\,(1 - x),\ 0 \le x \le 1,\ 0 \le \mu \le 4$$



The logistic function with $\mu = 0.5, 1, 2, 3, 4$

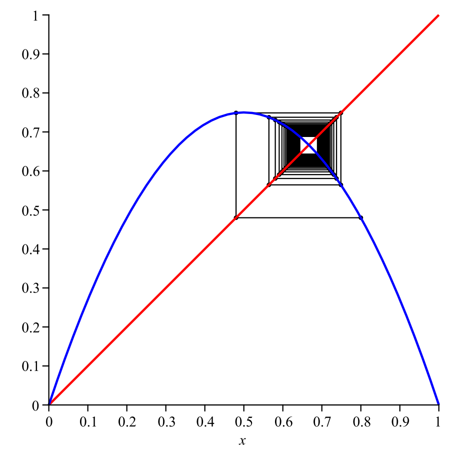
The logistic iteration $u(x) = \mu\, x\,(1-x)$

Fixed points $\left\{0, \dfrac{\mu-1}{\mu}\right\}$



$$u'(0) = \mu,\ u'\left(\tfrac{\mu-1}{\mu}\right) = 2 - \mu$$

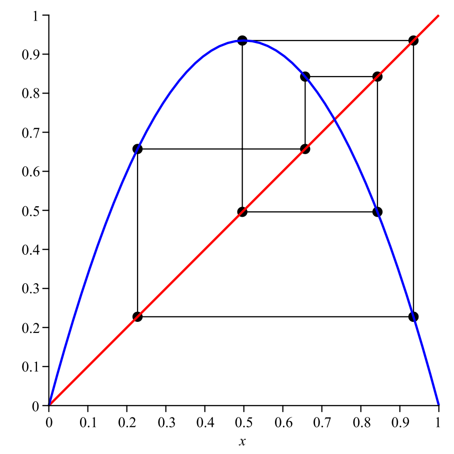
Example 9



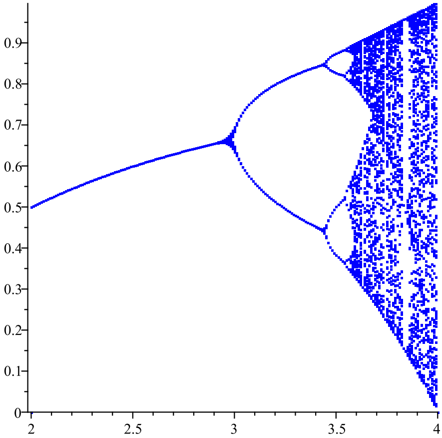
100 iterations, $x_0 = 0.8$

$u(x) = 3.74x(1-x)$

Example 10



5 cycle, $x_0 = 0.934945$



100 iterates of the logistic function for selected values of

$$2 \leq \mu \leq 4$$

The logistic iteration

$$u(x) = \mu\, x\, (1 - x),\ 0 \le x \le 1,\ 1 \le \mu \le 4$$

```
expand(InverseNewton(μ*x*(1-x),x));
```

$$\frac{(1 - \mu + \mu\, x)^{(-1+\mu)^{-1}}}{x^{(-1+\mu)^{-1}}}$$

(a) $\therefore\ f(x) = \left(\mathbf{N}^{-1}u\right)(x) = \left(\frac{x - \frac{\mu-1}{\mu}}{x}\right)^{\frac{1}{\mu-1}}$

(b) Fixed points $\{u\} = \left\{0,\ \frac{\mu-1}{\mu}\right\}$

(c) The fixed point $\frac{\mu-1}{\mu}$ is attracting for $1 \le \mu < 3$

(d) $f(x)$ is convex [concave] for $x < \frac{1}{2}$ [$x > \frac{1}{2}$]

The logistic iteration

$$u(x) = \mu\, x\,(1-x),\ 0 \le x \le 1,\ 1 \le \mu \le 4$$

```
expand(InverseNewton(μ*x*(1-x),x));
```

$$\frac{(1-\mu+\mu\, x)^{(-1+\mu)^{-1}}}{x^{(-1+\mu)^{-1}}}$$

(a) $\therefore\ f(x) = \left(\mathbf{N}^{-1}u\right)(x) = \left(\frac{x - \frac{\mu-1}{\mu}}{x}\right)^{\frac{1}{\mu-1}}$

(b) Fixed points $\{u\} = \left\{0,\ \frac{\mu-1}{\mu}\right\}$

(c) The fixed point $\frac{\mu-1}{\mu}$ is attracting for $1 \le \mu < 3$

(d) $f(x)$ is convex [concave] for $x < \frac{1}{2}$ $[x > \frac{1}{2}]$

The logistic iteration

$$u(x) = \mu\, x\,(1-x),\ 0 \le x \le 1,\ 1 \le \mu \le 4$$

```
expand(InverseNewton(μ*x*(1-x),x));
```

$$\frac{(1-\mu+\mu\,x)^{(-1+\mu)^{-1}}}{x^{(-1+\mu)^{-1}}}$$

(a) $\therefore\ f(x) = \left(\mathbf{N}^{-1}u\right)(x) = \left(\frac{x - \frac{\mu-1}{\mu}}{x}\right)^{\frac{1}{\mu-1}}$

(b) Fixed points $\{u\} = \left\{0,\ \frac{\mu-1}{\mu}\right\}$

(c) The fixed point $\frac{\mu-1}{\mu}$ is attracting for $1 \le \mu < 3$

(d) $f(x)$ is convex [concave] for $x < \frac{1}{2}$ $[x > \frac{1}{2}]$

The logistic iteration

$$u(x) = \mu\, x\,(1-x),\ 0 \le x \le 1,\ 1 \le \mu \le 4$$

```
expand(InverseNewton(μ*x*(1-x),x));
```

$$\frac{(1-\mu+\mu\,x)^{(-1+\mu)^{-1}}}{x^{(-1+\mu)^{-1}}}$$

(a) $\therefore\ f(x) = \left(\mathbf{N}^{-1}u\right)(x) = \left(\frac{x-\frac{\mu-1}{\mu}}{x}\right)^{\frac{1}{\mu-1}}$

(b) Fixed points $\{u\} = \left\{0,\ \frac{\mu-1}{\mu}\right\}$

(c) The fixed point $\frac{\mu-1}{\mu}$ is attracting for $1 \le \mu < 3$

(d) $f(x)$ is convex [concave] for $x < \frac{1}{2}$ $[x > \frac{1}{2}]$

The logistic iteration

$$u(x) = \mu\, x\,(1-x),\ 0 \le x \le 1,\ 1 \le \mu \le 4$$

```
expand(InverseNewton(μ*x*(1-x),x));
```

$$\frac{(1-\mu+\mu\, x)^{(-1+\mu)^{-1}}}{x^{(-1+\mu)^{-1}}}$$

(a) $\therefore\ f(x) = \left(\mathbf{N}^{-1}u\right)(x) = \left(\frac{x - \frac{\mu-1}{\mu}}{x}\right)^{\frac{1}{\mu-1}}$

(b) Fixed points $\{u\} = \left\{0,\ \frac{\mu-1}{\mu}\right\}$

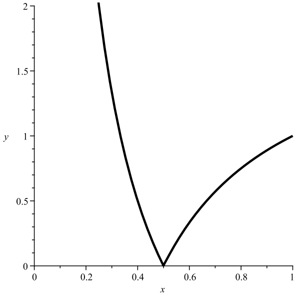
(c) The fixed point $\frac{\mu-1}{\mu}$ is attracting for $1 \le \mu < 3$

(d) $f(x)$ is convex [concave] for $x < \frac{1}{2}$ $[x > \frac{1}{2}]$

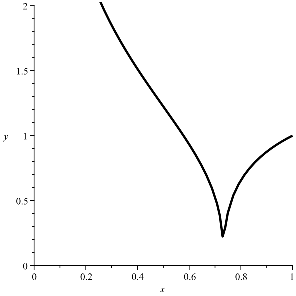
Chaos explained

The inverse Newton transform of $u(x) = \mu\, x\,(1-x)$

$$\left(\mathbf{N}^{-1}u\right)(x) = \left(\frac{x - \frac{\mu-1}{\mu}}{x}\right)^{\frac{1}{\mu-1}}$$

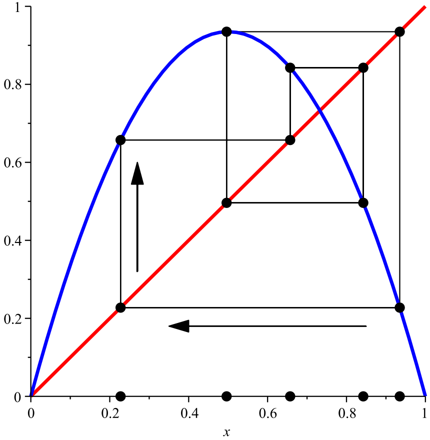


$\mu = 2.0$



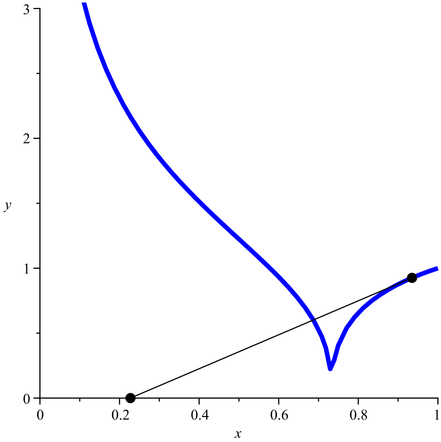
$\mu = 3.74$

5–cycle for $\mu = 3.74$

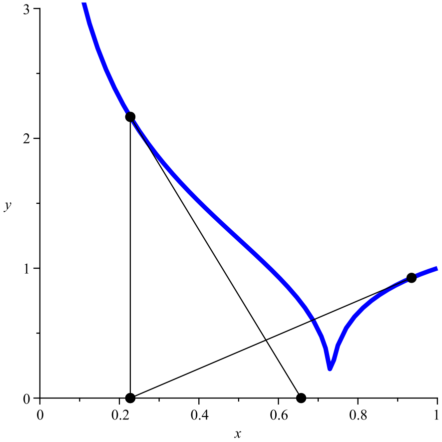


Starting at and returning to .9349453234

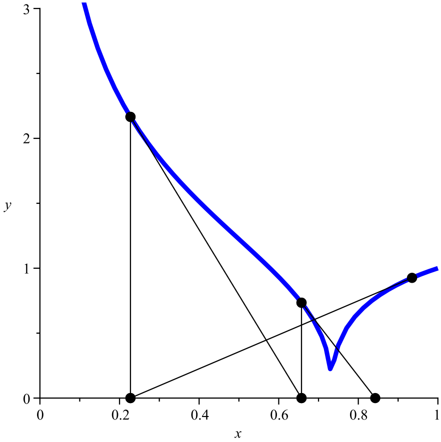
Ping pong–1



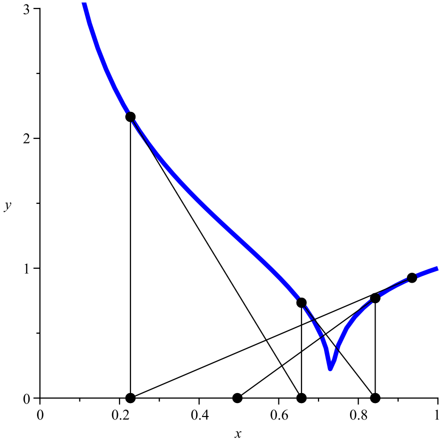
Ping pong–2



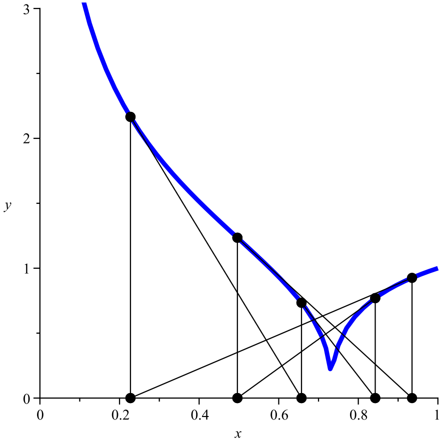
Ping pong–3



Ping pong–4



Ping pong–5



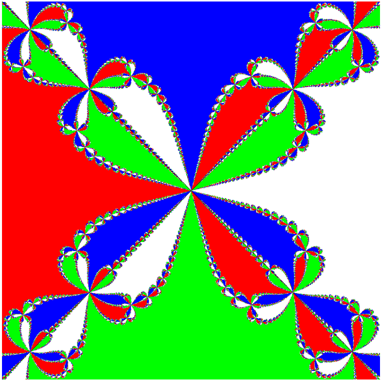
Outline

The complex Newton method, [6]

$$z_+ := z - \frac{f(z)}{f'(z)}, \; f'(z) \neq 0$$

The complex Newton method

$$z_+ := z - \frac{f(z)}{f'(z)}, \ f'(z) \neq 0$$



$$f(z) = z^4 - 1$$

The geometry of the complex Newton method, [19]

$$z_+ := z - \frac{f(z)}{f'(z)},\ f'(z) \neq 0$$

(a) Let

$$z = x + iy \longleftrightarrow (x,y)$$

be the natural correspondence between $\mathbb{C}$ and $\mathbb{R}^2$, and let

$$F(x,y) := |f(z)| \quad \text{for} \quad z \longleftrightarrow (x,y).$$

(b) Let $\mathbf{T} \subset \mathbb{R}^3$ be the plane tangent to the graph of F at the point $(x,y,F(x,y))$, and let $\mathbf{L}$ be the line of intersection of $\mathbf{T}$ and the (x,y)–plane ($\mathbf{L}$ is nonempty by the assumption that $f'(z) \neq 0$.)

(c) Then

$$z_+ \longleftrightarrow (x_+, y_+),$$

the perpendicular projection of (x,y) on $\mathbf{L}$.

The geometry of the complex Newton method, [19]

$$z_+ := z - \frac{f(z)}{f'(z)}, \; f'(z) \neq 0$$

(a) Let

$$z = x + iy \longleftrightarrow (x,y)$$

be the natural correspondence between $\mathbb{C}$ and $\mathbb{R}^2$, and let

$$F(x,y) := |f(z)| \quad \text{for} \quad z \longleftrightarrow (x,y).$$

(b) Let $\mathbf{T} \subset \mathbb{R}^3$ be the plane tangent to the graph of F at the point $(x,y,F(x,y))$, and let $\mathbf{L}$ be the line of intersection of $\mathbf{T}$ and the (x,y)–plane ($\mathbf{L}$ is nonempty by the assumption that $f'(z) \neq 0$.)

(c) Then

$$z_+ \longleftrightarrow (x_+, y_+),$$

the perpendicular projection of (x,y) on $\mathbf{L}$.

The geometry of the complex Newton method, [19]

$$z_+ := z - \frac{f(z)}{f'(z)}, \; f'(z) \neq 0$$

(a) Let

$$z = x + iy \longleftrightarrow (x,y)$$

be the natural correspondence between $\mathbb{C}$ and $\mathbb{R}^2$, and let

$$F(x,y) := |f(z)| \quad \text{for} \quad z \longleftrightarrow (x,y).$$

(b) Let $\mathbf{T} \subset \mathbb{R}^3$ be the plane tangent to the graph of F at the point $(x,y,F(x,y))$, and let $\mathbf{L}$ be the line of intersection of $\mathbf{T}$ and the (x,y)–plane ($\mathbf{L}$ is nonempty by the assumption that $f'(z) \neq 0$.)

(c) Then

$$z_+ \longleftrightarrow (x_+,y_+),$$

the perpendicular projection of (x,y) on $\mathbf{L}$.

The geometry of the complex Newton method, [19]

$$z_+ := z - \frac{f(z)}{f'(z)}, \; f'(z) \neq 0$$

(a) Let

$$z = x + iy \longleftrightarrow (x, y)$$

be the natural correspondence between $\mathbb{C}$ and $\mathbb{R}^2$, and let

$$F(x,y) := |f(z)| \quad \text{for} \quad z \longleftrightarrow (x,y).$$

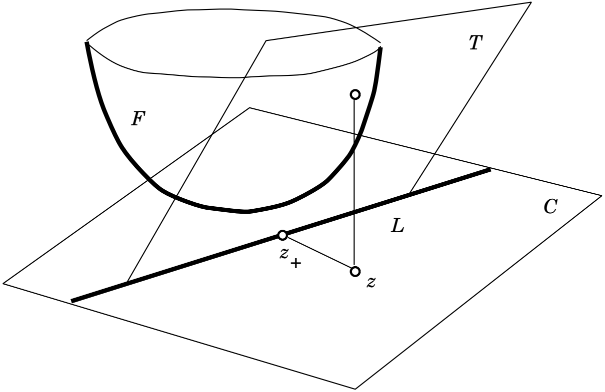
(b) Let $\mathbf{T} \subset \mathbb{R}^3$ be the plane tangent to the graph of F at the point $(x, y, F(x, y))$, and let $\mathbf{L}$ be the line of intersection of $\mathbf{T}$ and the (x, y)–plane ($\mathbf{L}$ is nonempty by the assumption that $f'(z) \neq 0$.)

(c) Then

$$z_+ \longleftrightarrow (x_+, y_+),$$

the perpendicular projection of (x, y) on $\mathbf{L}$.

The geometry of the complex Newton method



Outline of proof

The absolute value of $f = u + iv$

$$F(x,y) = |f(x+iy)| = \sqrt{u^2(x,y) + v^2(x,y)}\ ,$$

has gradient (where differentiable)

$$\nabla F(x,y) = \frac{1}{\sqrt{u^2+v^2}} \begin{pmatrix} u u_x + v v_x \\ u u_y + v v_y \end{pmatrix}$$

where $u_x = \partial u / \partial x,\ u_y = \partial u / \partial y$, etc.

Using the **Cauchy-Riemann** conditions

$$u_x = v_y,\ u_y = -v_x,$$

we get

$$\frac{f(z)}{f'(z)} = \frac{u+iv}{u_x+iv_x} = \frac{(u u_x + v v_x) + i(u u_y + v v_y)}{u_x^2 + v_x^2}$$

$$\longleftrightarrow \frac{1}{u_x^2+v_x^2} \begin{pmatrix} u u_x + v v_x \\ u u_y + v v_y \end{pmatrix} = t \nabla F(x,y)\ .$$

Outline of proof

The absolute value of $f = u + iv$

$$F(x,y) = |f(x+iy)| = \sqrt{u^2(x,y) + v^2(x,y)}\ ,$$

has gradient (where differentiable)

$$\nabla F(x,y) = \frac{1}{\sqrt{u^2+v^2}} \begin{pmatrix} u\,u_x + v\,v_x \\ u\,u_y + v\,v_y \end{pmatrix}$$

where $u_x = \partial u / \partial x,\ u_y = \partial u / \partial y$, etc.
Using the **Cauchy-Riemann** conditions

$$u_x = v_y,\ u_y = -v_x,$$

we get

$$\begin{aligned} \frac{f(z)}{f'(z)} &= \frac{u+iv}{u_x + iv_x} = \frac{(u\,u_x + v\,v_x) + i\,(u\,u_y + v\,v_y)}{u_x^2 + v_x^2} \\ &\longleftrightarrow \frac{1}{u_x^2+v_x^2} \begin{pmatrix} u\,u_x + v\,v_x \\ u\,u_y + v\,v_y \end{pmatrix} = t\,\nabla F(x,y)\ . \end{aligned}$$

Outline of proof (contd.)

$$\begin{aligned}\frac{f(z)}{f'(z)} &= \frac{u+iv}{u_x+iv_x} = \frac{(uu_x+vv_x)+i(uu_y+vv_y)}{u_x^2+v_x^2} \\ &\longleftrightarrow \frac{1}{u_x^2+v_x^2}\begin{pmatrix} uu_x+vv_x \\ uu_y+vv_y \end{pmatrix} = t\nabla F(x,y)\ .\end{aligned}$$

The Newton method for f is thus a gradient method for $|f|$, i.e.,

$$\begin{pmatrix} x_+ \\ y_+ \end{pmatrix} := \begin{pmatrix} x \\ y \end{pmatrix} - t\nabla F(x,y)\ .$$

It remains to show that t is as claimed.

$$\frac{f(z)}{f'(z)} = \frac{u + i\,v}{u_x + i\,v_x} = \frac{(u\,u_x + v\,v_x) + i\,(u\,u_y + v\,v_y)}{u_x^2 + v_x^2}$$

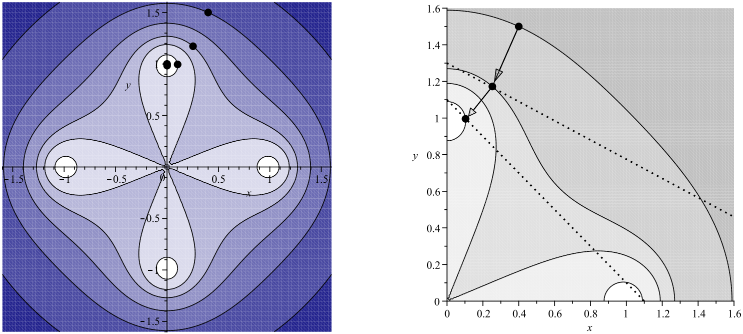
$$\longleftrightarrow \quad \frac{1}{u_x^2 + v_x^2} \begin{pmatrix} u\,u_x + v\,v_x \\ u\,u_y + v\,v_y \end{pmatrix} = t\,\nabla F(x,y) \ .$$

The Newton method for f is thus a gradient method for $|f|$, i.e.,

$$\begin{pmatrix} x_+ \\ y_+ \end{pmatrix} := \begin{pmatrix} x \\ y \end{pmatrix} - t\,\nabla F(x,y) \ .$$

It remains to show that t is as claimed.

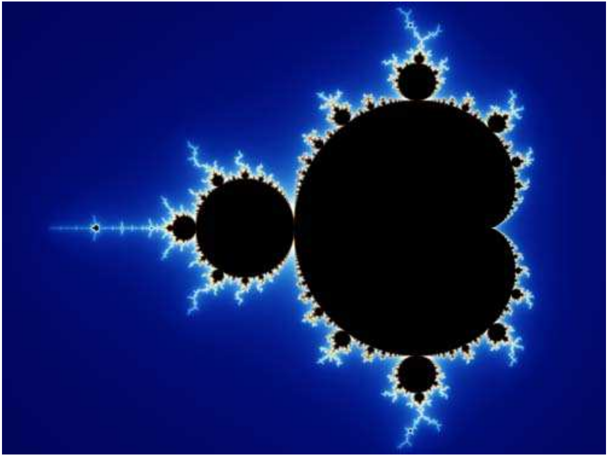
$f(z) = z^4 - 1$



Level sets of $|z^4 - 1|$ and iterates converging to i

The Mandelbrot set

$$\mathbb{M} := \{\mathbf{c} : \{\mathbf{z_n} := \mathbf{z_{n-1}^2} + \mathbf{c}, \mathbf{z_0} = \mathbf{0}\} \textbf{ bounded}\}$$



The Mandelbrot set

$$\mathbb{M} := \{\mathbf{c} : \{\mathbf{z_n} := \mathbf{z_{n-1}^2} + \mathbf{c}, \mathbf{z_0} = \mathbf{0}\} \textbf{ bounded}\}$$

`InverseNewton`$(z^2 + c, z)$;

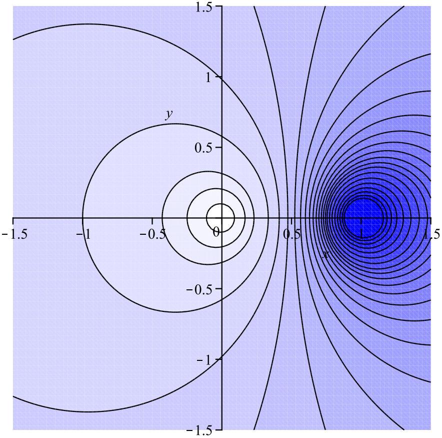
$$\exp\left\{-\frac{2}{\sqrt{4c-1}} \arctan\left(\frac{2z-1}{\sqrt{4c-1}}\right)\right\}$$

`InverseNewton`$(z^2 + (1/4), z)$;

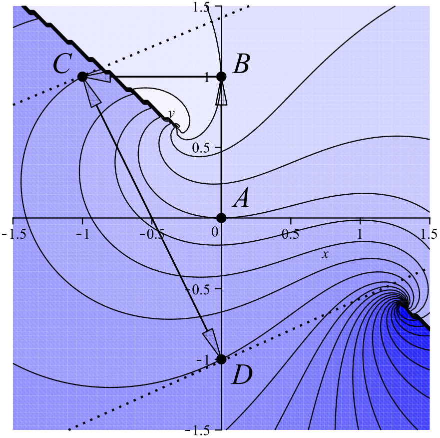
$$\exp\left\{\frac{2}{2z-1}\right\}$$

`InverseNewton`(z^2, z);

$$\frac{z}{z-1}$$



Level sets of $|\mathbf{N}^{-1}(z^2)|$



Level sets of $|\mathbf{N}^{-1}(z^2 + i)|$

Outline

Minimization of a convex $f : \mathbb{R} \to \mathbb{R}$, [12]

Let $f : \mathbb{R} \to \mathbb{R}$ be convex, with an attained infimum $f_{\min}$.

A **bracket** is a closed interval $[L,U]$ with

$$L \leq f_{\min} \leq U.$$

The length of the bracket $[L,U]$ is denoted $\Delta := U - L$.

A **bracketing method** generates a sequence of nested brackets, shrinking to a point,

$$L \leq L_+ \leq f_{\min} \leq U_+ \leq U, \text{ and } \Delta_+ := U_+ - L_+ < \Delta.$$

At each iteration select a middle value

$$M := \alpha U + (1-\alpha)L, \quad \text{for some } 0 < \alpha < 1,$$

and do one Newton iteration for solving $\quad f(x) - M = 0.$

Let $f : \mathbb{R} \to \mathbb{R}$ be convex, with an attained infimum $f_{\min}$.

A **bracket** is a closed interval $[L,U]$ with

$$L \leq f_{\min} \leq U.$$

The length of the bracket $[L,U]$ is denoted $\Delta := U - L$.

A **bracketing method** generates a sequence of nested brackets, shrinking to a point,

$$L \leq L_+ \leq f_{\min} \leq U_+ \leq U, \text{ and } \Delta_+ := U_+ - L_+ < \Delta.$$

At each iteration select a middle value

$$M := \alpha U + (1-\alpha) L, \quad \text{for some } 0 < \alpha < 1,$$

and do one Newton iteration for solving $\quad f(x) - M = 0.$

Let $f:\mathbb{R}\to\mathbb{R}$ be convex, with an attained infimum $f_{\min}$.

A **bracket** is a closed interval $[L,U]$ with

$$L \le f_{\min} \le U.$$

The length of the bracket $[L,U]$ is denoted $\Delta := U - L$.

A **bracketing method** generates a sequence of nested brackets, shrinking to a point,

$$L \le L_+ \le f_{\min} \le U_+ \le U, \text{ and } \Delta_+ := U_+ - L_+ < \Delta.$$

At each iteration select a middle value

$$M := \alpha U + (1-\alpha)L, \quad \text{for some } 0 < \alpha < 1,$$

and do one Newton iteration for solving $\quad f(x) - M = 0.$

Minimization of a convex $f : \mathbb{R} \to \mathbb{R}$, [12]

Let $f : \mathbb{R} \to \mathbb{R}$ be convex, with an attained infimum $f_{\min}$.

A **bracket** is a closed interval $[L, U]$ with

$$L \leq f_{\min} \leq U.$$

The length of the bracket $[L, U]$ is denoted $\Delta := U - L$.

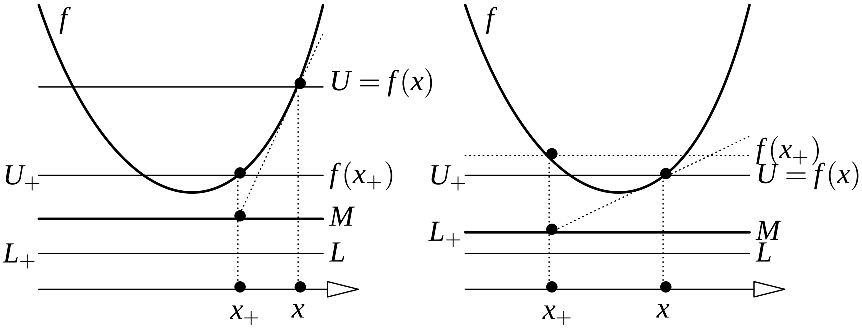
A **bracketing method** generates a sequence of nested brackets, shrinking to a point,

$$L \leq L_+ \leq f_{\min} \leq U_+ \leq U, \text{ and } \Delta_+ := U_+ - L_+ < \Delta.$$

At each iteration select a middle value

$$M := \alpha U + (1 - \alpha) L, \quad \text{for some } 0 < \alpha < 1,$$

and do one Newton iteration for solving $\quad f(x) - M = 0.$



(a) Case 1: $f(x_+) < f(x)$ (b) Case 2: $f(x_+) \geq f(x)$

Figure: Illustration of the 2 cases of the NB method

NB method for convex $f : \mathbb{R} \to \mathbb{R}$

0 **Initialize.**
x initial iterate
$U := f(x)$, upper bound
L, lower bound, must be less than $f_{\min}$
$\alpha \in (0,1)$
$\varepsilon > 0$, tolerance.

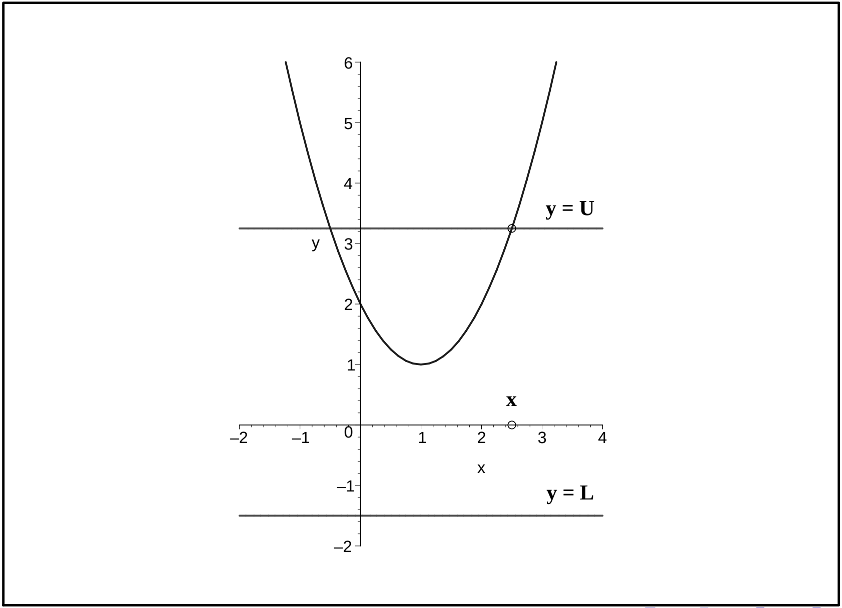
1 **Stopping rule.** If $U - L < \varepsilon$, **stop** with x as solution.
2 Select a value $M := \alpha U + (1-\alpha)L$, for some $0 < \alpha < 1$.
3 Do one Newton iteration $x_+ := x - \dfrac{f(x) - M}{f'(x)}$.
4 **Case 1**: If $f(x_+) < f(x)$ then update U: $U_+ := f(x_+)$ and leave $L_+ := L$. Go to **1**.
5 **Case 2**: If $f(x_+) \geq f(x)$ then update L: $L_+ := M$ and leave $U_+ := U$, $x_+ := x$. Go to **1**.

NB method for convex $f : \mathbb{R} \to \mathbb{R}$

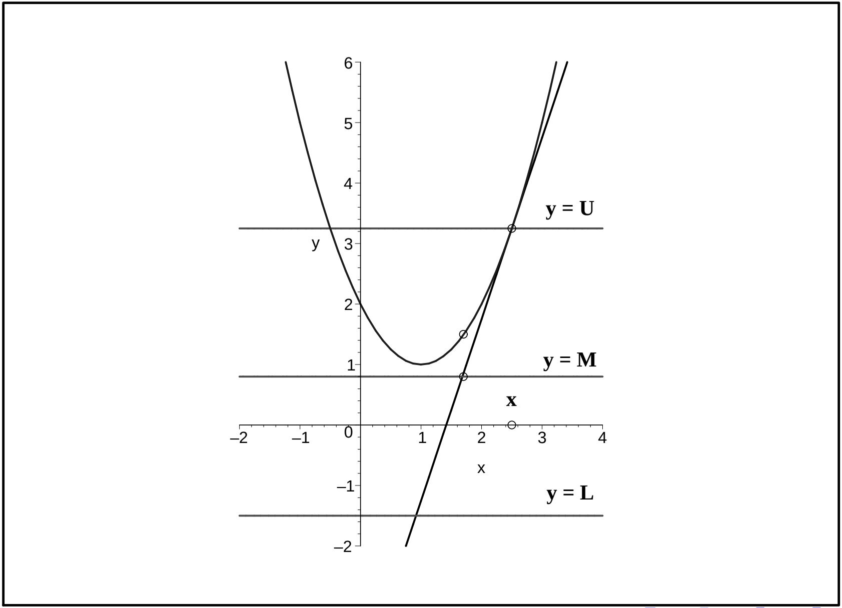
0	**Initialize.** x initial iterate $U := f(x)$, upper bound L, lower bound, must be less than $f_{\min}$ $\alpha \in (0,1)$ $\varepsilon > 0$, tolerance.

1	**Stopping rule.** If $U - L < \varepsilon$, **stop** with x as solution.
2	Select a value $M := \alpha U + (1-\alpha)L$, for some $0 < \alpha < 1$.
3	Do one Newton iteration $x_+ := x - \dfrac{f(x) - M}{f'(x)}$.
4	**Case 1**: If $f(x_+) < f(x)$ then update U: $U_+ := f(x_+)$ and leave $L_+ := L$. Go to **1**.
5	**Case 2**: If $f(x_+) \geq f(x)$ then update L: $L_+ := M$ and leave $U_+ := U$, $x_+ := x$. Go to **1**.

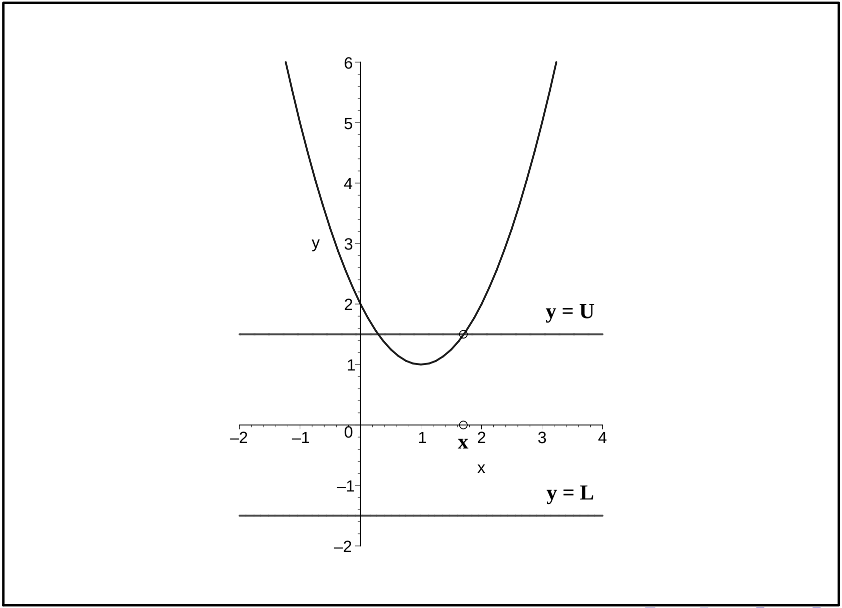
Example: Initialization



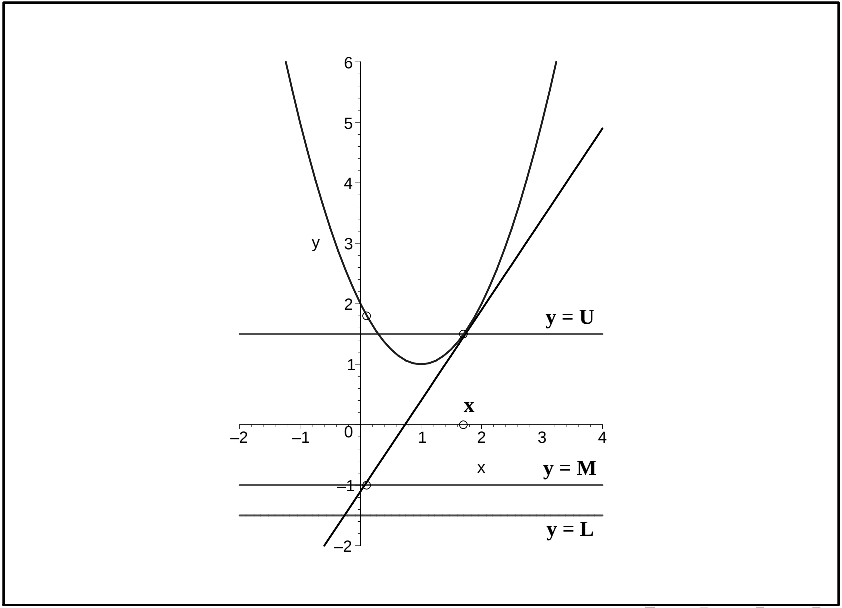
Iteration 1: Newton step



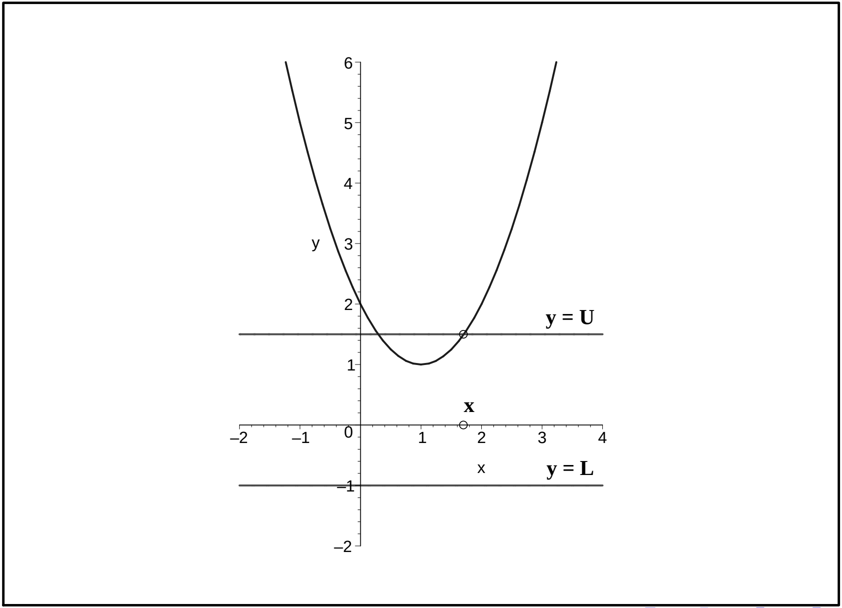
Iteration 1: New bracket



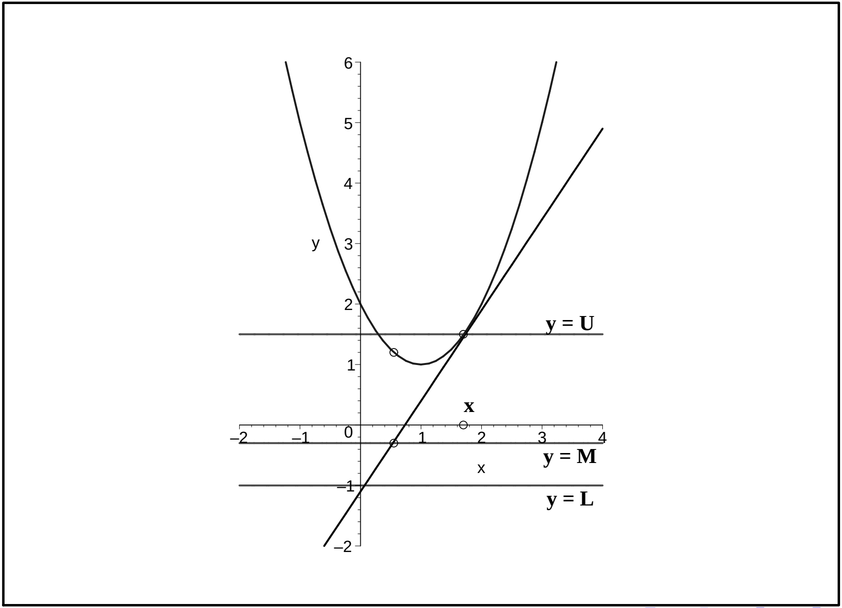
Iteration 2: Newton step



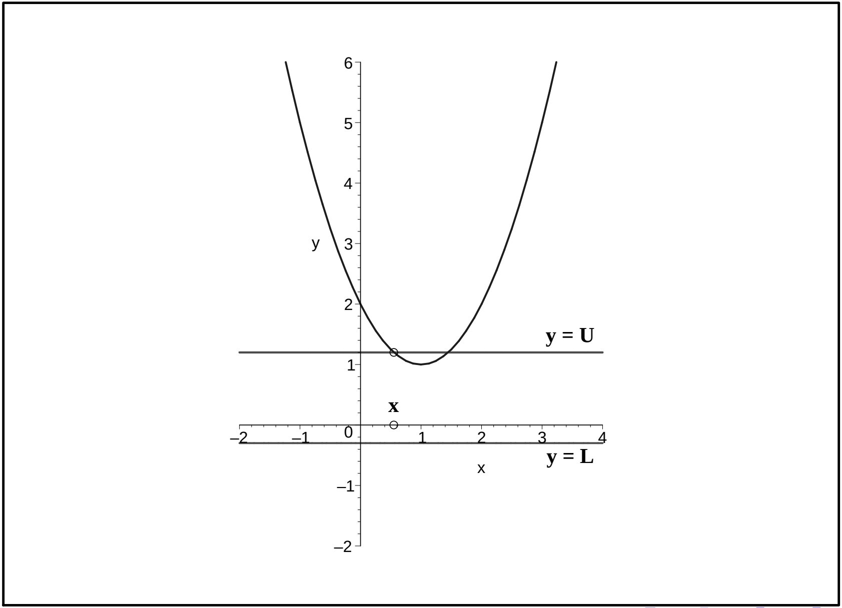
Iteration 2: New bracket



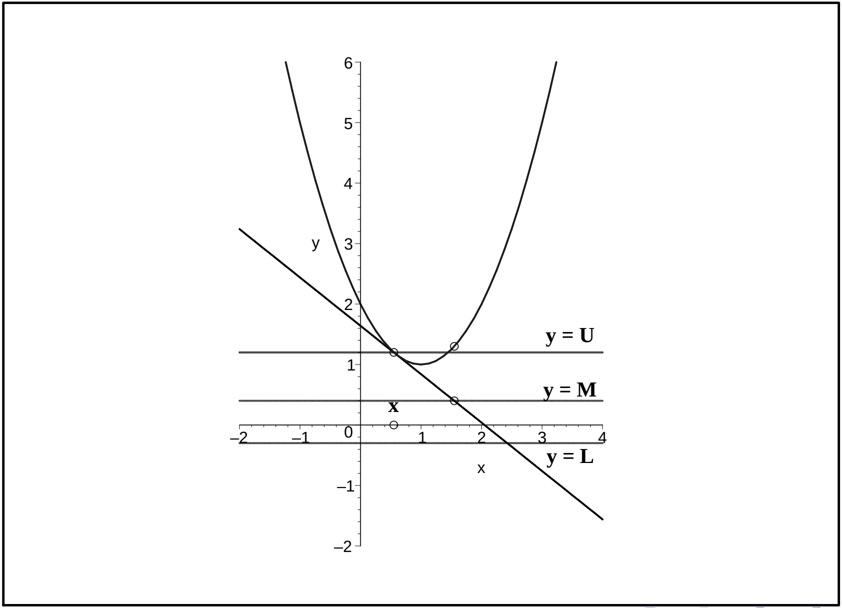
Iteration 3: Newton step



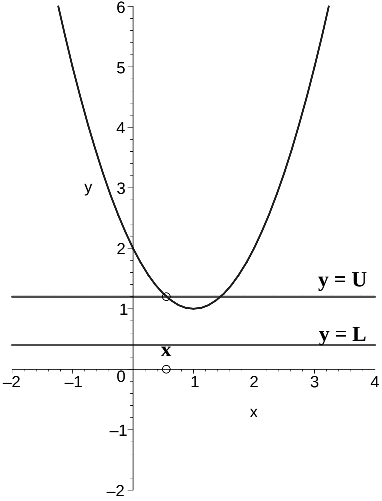
Iteration 3: New bracket



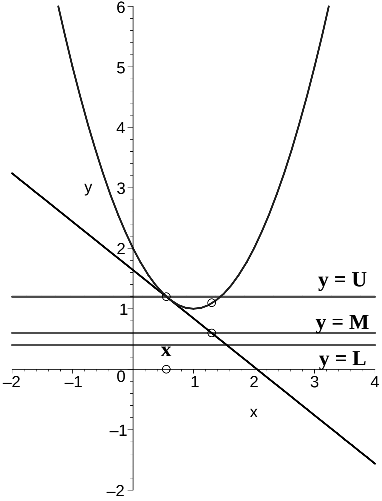
Iteration 4: Newton step



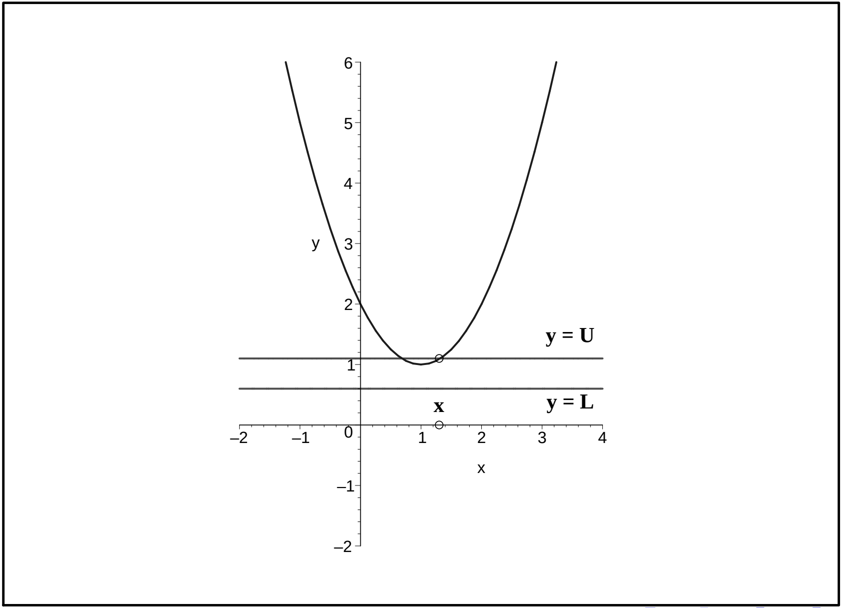
Iteration 4: New bracket



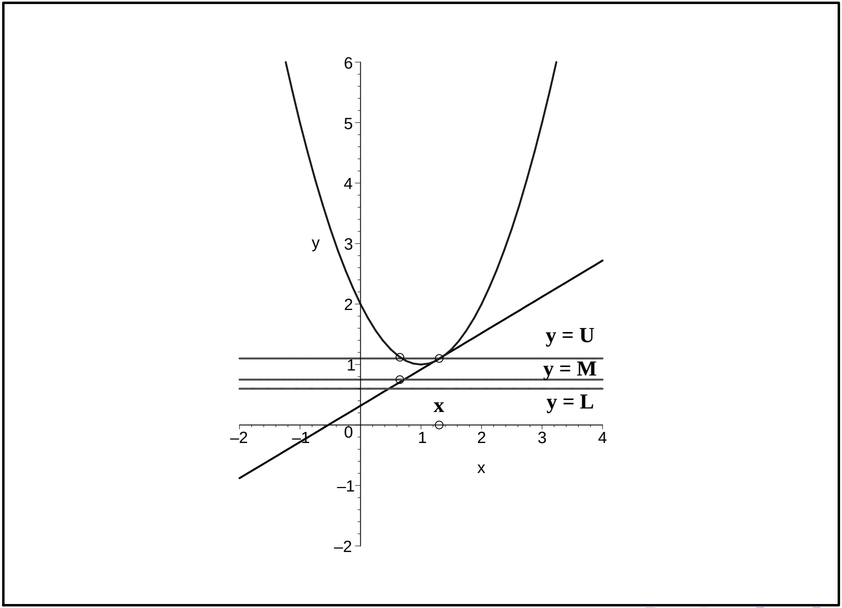
Iteration 5: Newton step



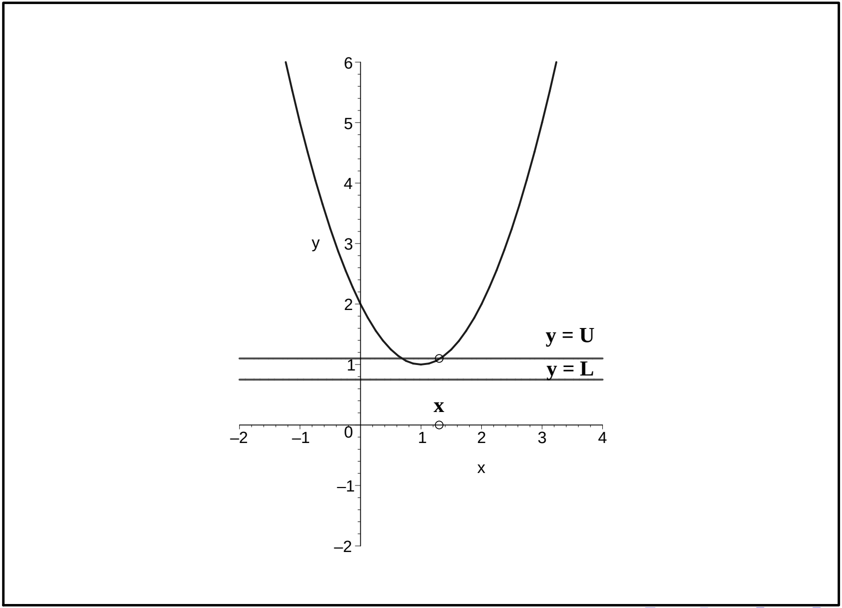
Iteration 5: New bracket



Iteration 6: Newton step



Iteration 6: New bracket



NB method for convex $f : \mathbb{R}^n \to \mathbb{R}$

1. **Stopping rule**. If $U - L < \varepsilon$, **stop** with x as solution.
2. Select a value $M := \alpha U + (1-\alpha)L$, for some $0 < \alpha < 1$.
3. Do one Newton iteration $x_+ := x - \dfrac{f(x) - M}{f'(x)}$.
4. **Case 1**: If $f(x_+) < f(x)$ then update U: $U_+ := f(x_+)$ and leave $L_+ := L$. Go to **1**.
5. **Case 2**: If $f(x_+) \geq f(x)$ then update L: $L_+ := M$ and leave $U_+ := U$, $x_+ := x$. Go to **1**.

For $n > 1$, the only change is in step 3 :

3. Do one directional Newton iteration

$$\mathbf{x}_+ := \mathbf{x} - \frac{f(\mathbf{x}) - M}{\|\nabla f(\mathbf{x})\|^2} \nabla f(\mathbf{x}).$$

using the directional Newton method, [11].

NB method for convex $f : \mathbb{R}^n \to \mathbb{R}$

1	**Stopping rule.** If $U - L < \varepsilon$, **stop** with x as solution.
2	Select a value $M := \alpha U + (1-\alpha)L$, for some $0 < \alpha < 1$.
3	Do one Newton iteration $x_+ := x - \dfrac{f(x) - M}{f'(x)}$.
4	**Case 1**: If $f(x_+) < f(x)$ then update U: $U_+ := f(x_+)$ and leave $L_+ := L$. Go to **1**.
5	**Case 2**: If $f(x_+) \geq f(x)$ then update L: $L_+ := M$ and leave $U_+ := U$, $x_+ := x$. Go to **1**.

For $n > 1$, the only change is in step 3 :

3 Do one directional Newton iteration

$$\mathbf{x}_+ := \mathbf{x} - \frac{f(\mathbf{x}) - M}{\|\nabla f(\mathbf{x})\|^2} \nabla f(\mathbf{x}).$$

using the directional Newton method, [11].

Validity for $f : \mathbb{R}^n \to \mathbb{R}$

The NB method is **valid** if

$$L \leq f_{\min} \leq U$$

holds throughout the iterations. This is guaranteed for $n = 1$.

Sufficient validity conditions for $n > 1$ were given in [12], in particular, the method is valid for the quadratic function,

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} + \gamma, \quad Q \text{ positive definite,}$$

if Q is well-conditioned,

$$\frac{\lambda_n}{\lambda_1} \geq 7 - \sqrt{48} \approx 0.071796768,$$

where λ_n and λ_1 are the extreme eigenvalues of Q.

The NB method is **valid** if

$$L \le f_{\min} \le U$$

holds throughout the iterations. This is guaranteed for $n = 1$.

Sufficient validity conditions for $n > 1$ were given in [12], in particular, the method is valid for the quadratic function,

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} + \gamma, \quad Q \text{ positive definite,}$$

if Q is well-conditioned,

$$\frac{\lambda_n}{\lambda_1} \ge 7 - \sqrt{48} \approx 0.071796768,$$

where λ_n and λ_1 are the extreme eigenvalues of Q.

Application: The Fermat–Weber location problem

Find $\mathbf{x} \in \mathbb{R}^n$ minimizing the sum of Euclidean distances

$$f(\mathbf{x}) = \sum_{i=1}^{m} \|\mathbf{a}_i - \mathbf{x}\|$$

from m given points $\{\mathbf{a}_i : i = 1, \ldots, m\} \subset \mathbb{R}^n$.

Gradient methods, such as the **Weiszfeld method** ([18], [1]) suffer from slow convergence and the lack of a natural stopping criterion.

The NB method is better on both counts.

Application: The Fermat–Weber location problem

Find $\mathbf{x} \in \mathbb{R}^n$ minimizing the sum of Euclidean distances

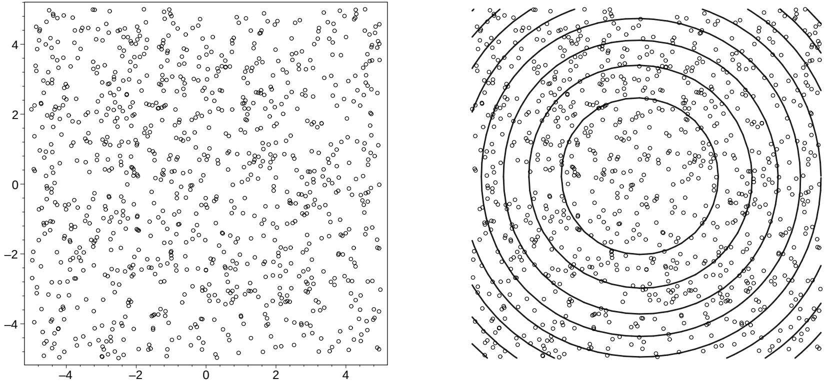
$$f(\mathbf{x}) = \sum_{i=1}^{m} \|\mathbf{a}_i - \mathbf{x}\|$$

from m given points $\{\mathbf{a}_i : i = 1, \ldots, m\} \subset \mathbb{R}^n$.

Gradient methods, such as the **Weiszfeld method** ([18], [1]) suffer from slow convergence and the lack of a natural stopping criterion.

The NB method is better on both counts.

1000 points $\{\mathbf{a}_i\}$ and level sets of $\sum_{i=1}^{m} \|\mathbf{a}_i - \mathbf{x}\|$



(a) Points

(b) Level sets

Figure: 1,000 random points in $[-5,5]^2$

Numerical results for 1000 random points in $[-5,5]^2$

Iteration	0	1	**2**	3	4	**5**	**6**
α	0.95	0.95	**0.95**	0.88	0.95	**0.95**	**0.95**
Case	1	1	**2**	1	1	**2**	**2**
Δ	3809.9	3771.1	**188.5**	172.3	167.3	**8.37**	**0.418**
Reduction		0.989	**0.05**	0.914	0.971	**0.05**	**0.05**

Iteration	7	8	**9**	10	11	**12**
α	0.83	0.95	**0.95**	0.89	0.95	**0.95**
Case	1	1	**2**	1	1	**2**
Δ	0.367	0.355	**0.018**	0.0163	0.016	**0.0008**
Reduction	0.881	0.945	**0.05**	0.905	0.981	**0.05**

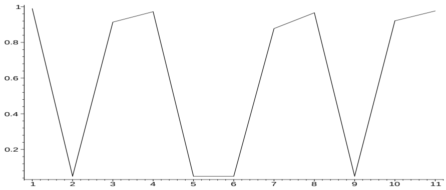


Figure: Reduction per iteration

Numerical results for 1000 random points in $[-5,5]^2$

Iteration	0	1	**2**	3	4	**5**	**6**
α	0.95	0.95	**0.95**	0.88	0.95	**0.95**	**0.95**
Case	1	1	**2**	1	1	**2**	**2**
Δ	3809.9	3771.1	**188.5**	172.3	167.3	**8.37**	**0.418**
Reduction		0.989	**0.05**	0.914	0.971	**0.05**	**0.05**

Iteration	7	8	**9**	10	11	**12**
α	0.83	0.95	**0.95**	0.89	0.95	**0.95**
Case	1	1	**2**	1	1	**2**
Δ	0.367	0.355	**0.018**	0.0163	0.016	**0.0008**
Reduction	0.881	0.945	**0.05**	0.905	0.981	**0.05**

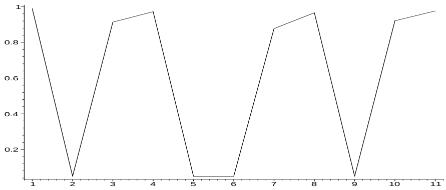


Figure: Reduction per iteration

Outline

References

[1] A. Beck and S. Sabach, Weiszfeld's method: Old and new results, *J. Optimiz. Th. Appl.* **164**(2015), 1–40

[2] A. B-I, A Newton-Raphson Method for the Solution of Systems of Equations, *J. Math. Anal. Appl.* **15**(1966), 243–252, **[PDF]**

[3] A. B–I, Newton's method with modified functions, *Contemporary Mathematics* **204**(1997), 39–50, **[PDF]**

[4] A. B–I, An inverse Newton transform, *Contemporary Mathematics* **568**(2012), 27–40, **[PDF]**

[5] A. B-I and Y. Levin, The Newton bracketing method for convex minimization: Convergence analysis, Chapter 4, pp. 49-63 in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, (H.H. Bauschke, R.S. Burachik, P.L. Combettes, V. Elser, D.R. Luke and H. Wolkowicz, Editors), Springer, 2011, **[PDF]**

[6] A. Cayley, The Newton–Fourier imaginary problem, *Amer. J. Math.* **2**(1879), 97

[7] S. Gorn, Maximal convergence intervals and a Gibbs type phenomenon for Newton's approximation procedure, *The Annals of Mathematics* (Second Series) **59**(1954), 463–476

[8] F. v. Haeseler and H.-O. Peitgen, Newton's method and complex dynamical systems, *Acta Applicandae Mathematicae* **13**(1988), 3–58

[9] Y. Levin and A. B-I, A Newton method for systems of m equations in n variables, *Nonlinear Analysis* **47**(2001), 1961–1971, **[PDF]**

[10] Y. Levin and A. B-I, Directional Halley and Quasi-Halley Methods in n Variables, , pp. 345-365 in *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications* (D. Butnariu, Y. Censor and S. Reich, Editors), Elsevier Science, Amsterdam, 2001, **[PDF]**

[11] Y. Levin and A. B-I, Directional Newton methods in n variables, *Math. of Computation* **71**(2002), 237–250, **[PDF]**

[12] Y. Levin and A. B-I, The Newton bracketing method for convex minimization, *Comput. Optimiz. and Appl.* **21**(2002), 213–229 , **[PDF]**

[13] Y. Levin and A. B-I, A Heuristic Method for Large-Scale Multi-Facility Location Problems, *Computers and Oper. Res.* **31**(2004), 257–272 , **[PDF]**

References

[14] Y. Levin, M. Nediak and A. B-I, Direct Approach to Calculus of Variations via Newton-Raphson Method, *J. of Comput. Appl. Math.* **139**(2001), 197–213, [PDF]

[15] B. Mandelbrot, Fractal aspects of the iteration of $z \mapsto \lambda z(1-z)$ for complex λ, z, *Annals NY Acad. Sci.* **357**(1980), 249–259

[16] A.M. Ostrwoski, *Solutions of Equations in Euclidean and Banach Spaces*, 3rd Edition, Academic Press, 1973

[17] H. O. Peitgen, D. Saupe, and F. v. Haeseler, Cayley's Problem and Julia Sets, *The Mathematical Intelligencer* **6**(1984), No. 2

[18] E.V. Weiszfeld, Sur le point pour lequel la somme des distances de n points donnes est minimum, *Tohoku Math. J.* **43**(1937), 355–386

[19] L. Yau and A. B–I, The Newton and Halley methods for complex roots, *Amer. Math. Monthly* **105**(1998), 806–818, [PDF]

THANKS FOR YOUR ATTENTION