Example: Newton method in 1d

- **Assumption:** $f(x)$ is 2x continuous differentiable; $f'(x)$ and $f''(x)$ can be computed analytically.
- **Strategy:** Determine iteratively a zero of the 1st derivative; start with a value, that is assumed to be close to a minimum.
- **Method:** Start with initial value $x^{[0]}$

$$\text{For } k = 0, 1, 2, \dots : \qquad \underset{\substack{\| \\ x_{k+1}}}{x^{[k+1]}} = \underset{\substack{\| \\ x_k}}{x^{[k]}} - \frac{f'(x^{[k]})}{f''(x^{[k]})}$$

- **Convergence:** the Newton method converges locally quadratically, i.e. there holds

$$|\overline{x} - x^{[k+1]}| \leq c \cdot |\overline{x} - x^{[k]}|^2,$$

if f'' is invertible and differentiable (can be weakened) in a neighbourhood of a zero.

Excursus: solving nonlinear equations

Nonlinear equations appear:

- Implicit methods: solving a calculation rule for $x^{[k+1]}$
- Determine stationary points of nonlinear differential equations
- Numerical methids for nonlinear differential equations
- and many other applications . . .

In general: let $g \in C^1(\mathbb{R})$. Determine constructively $\hat{x}$ such that

$$g(\hat{x}) = 0.$$

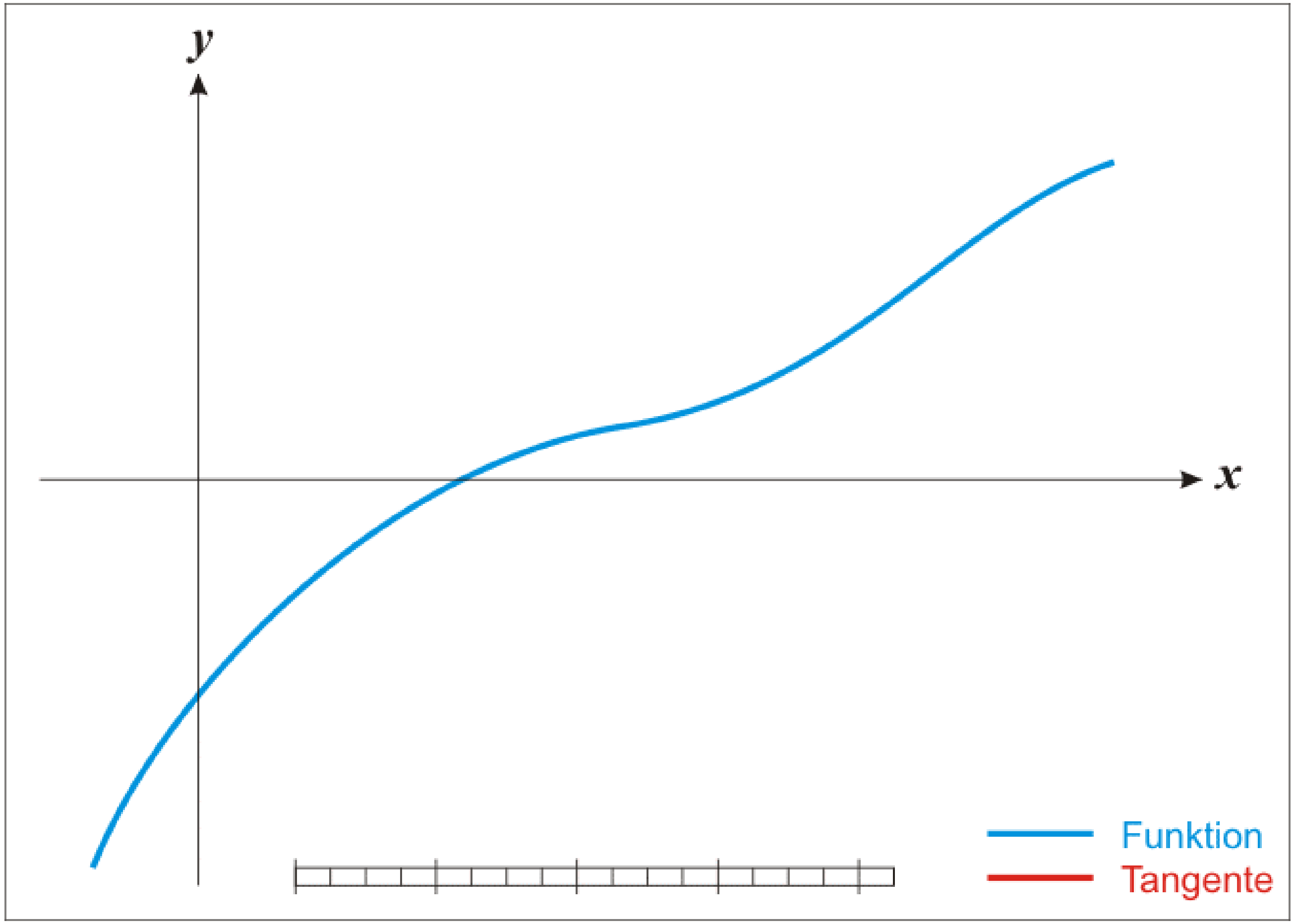
$$\overset{\shortparallel}{f'(\hat{x})} \rightsquigarrow \text{special case optimization}$$

Newton method:

$$x^{[k+1]} = x^{[k]} - \frac{g(x^{[k]})}{g'(x^{[k]})}, \qquad x^{[0]} \in \mathbb{R} \text{ given initial value}$$

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

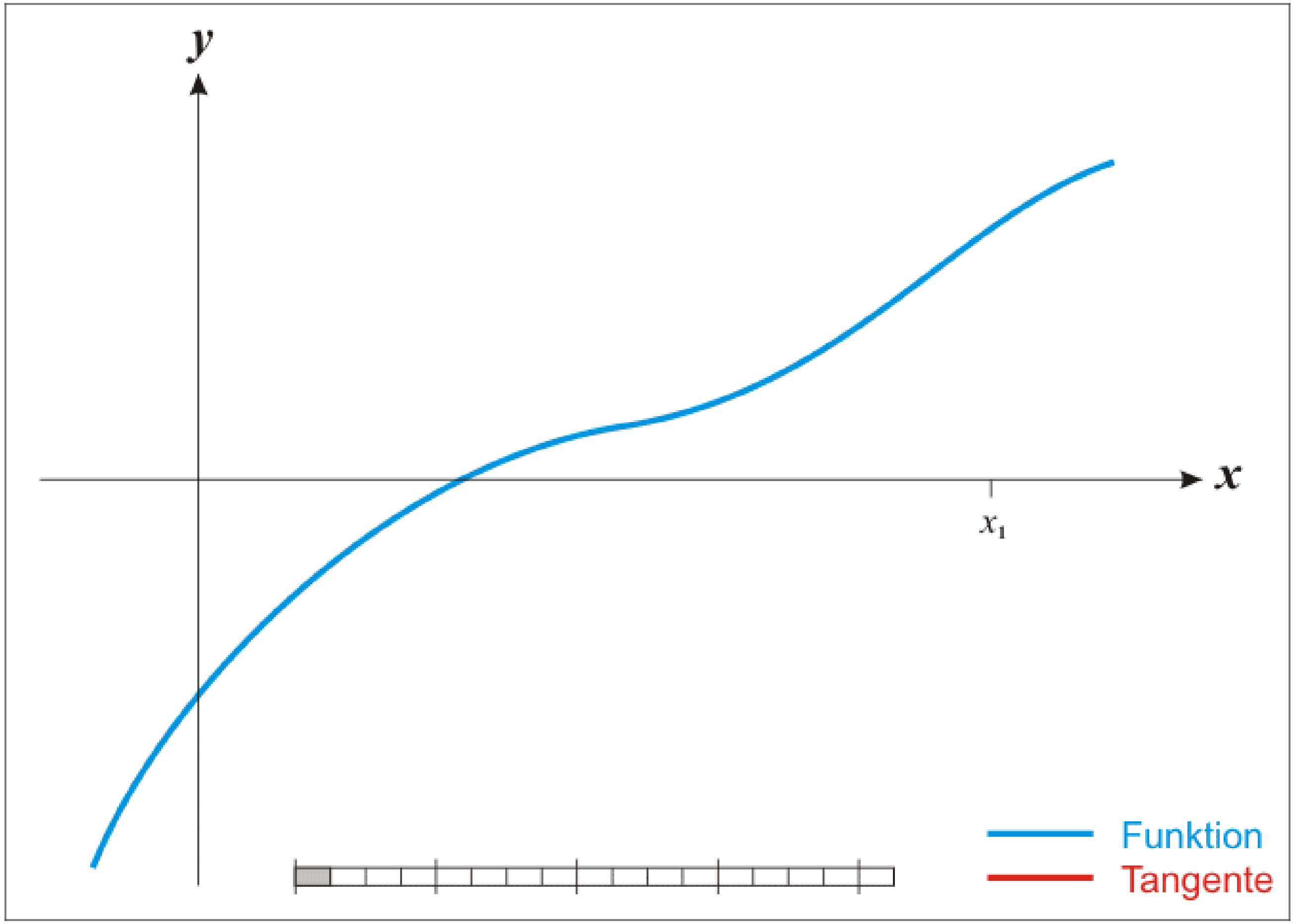


(here $x^{[n]}$ defines a sequence)

See `https://de.wikipedia.org/wiki/Newton-Verfahren`.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):



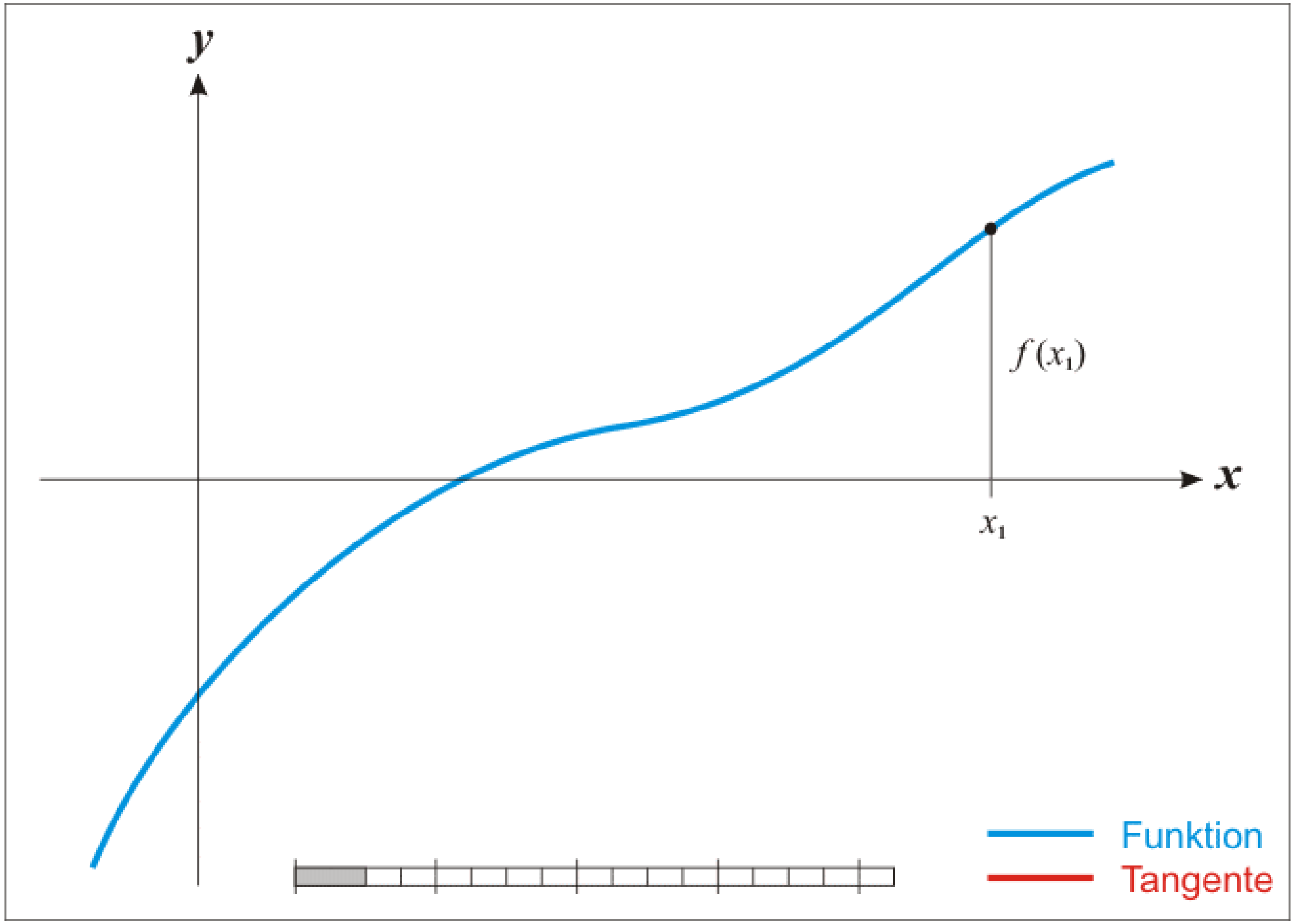
$x_n =$

(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

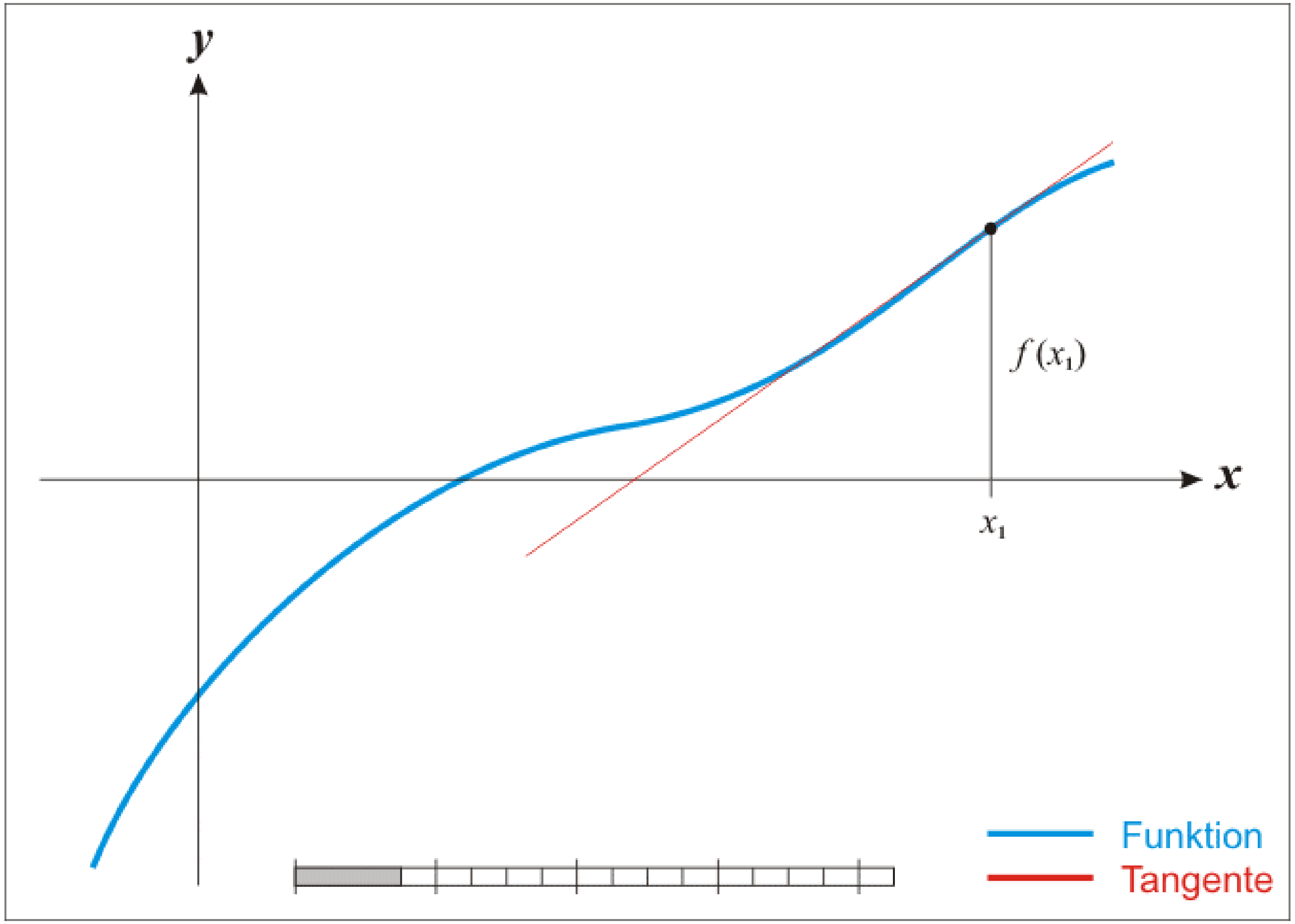


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

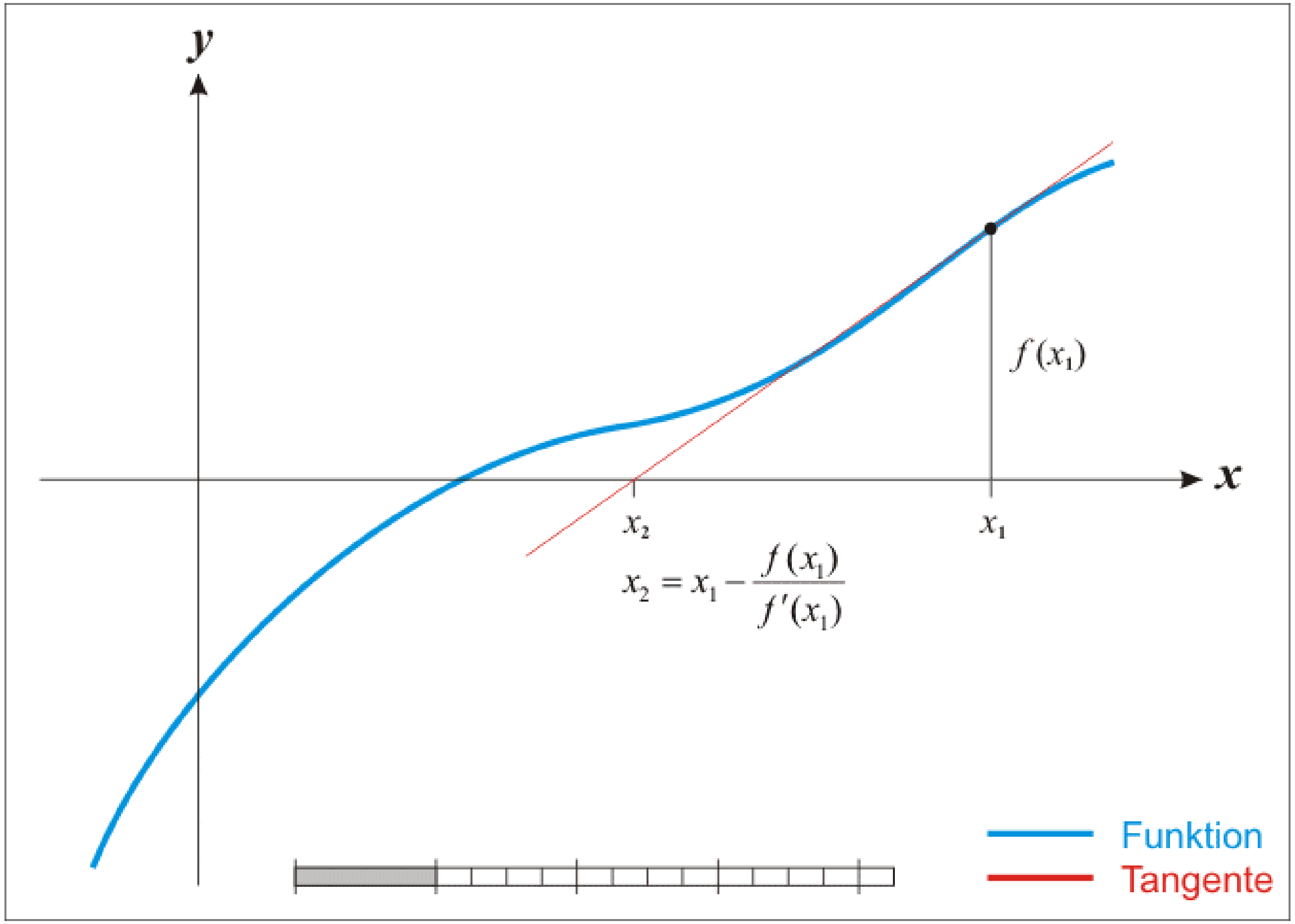


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

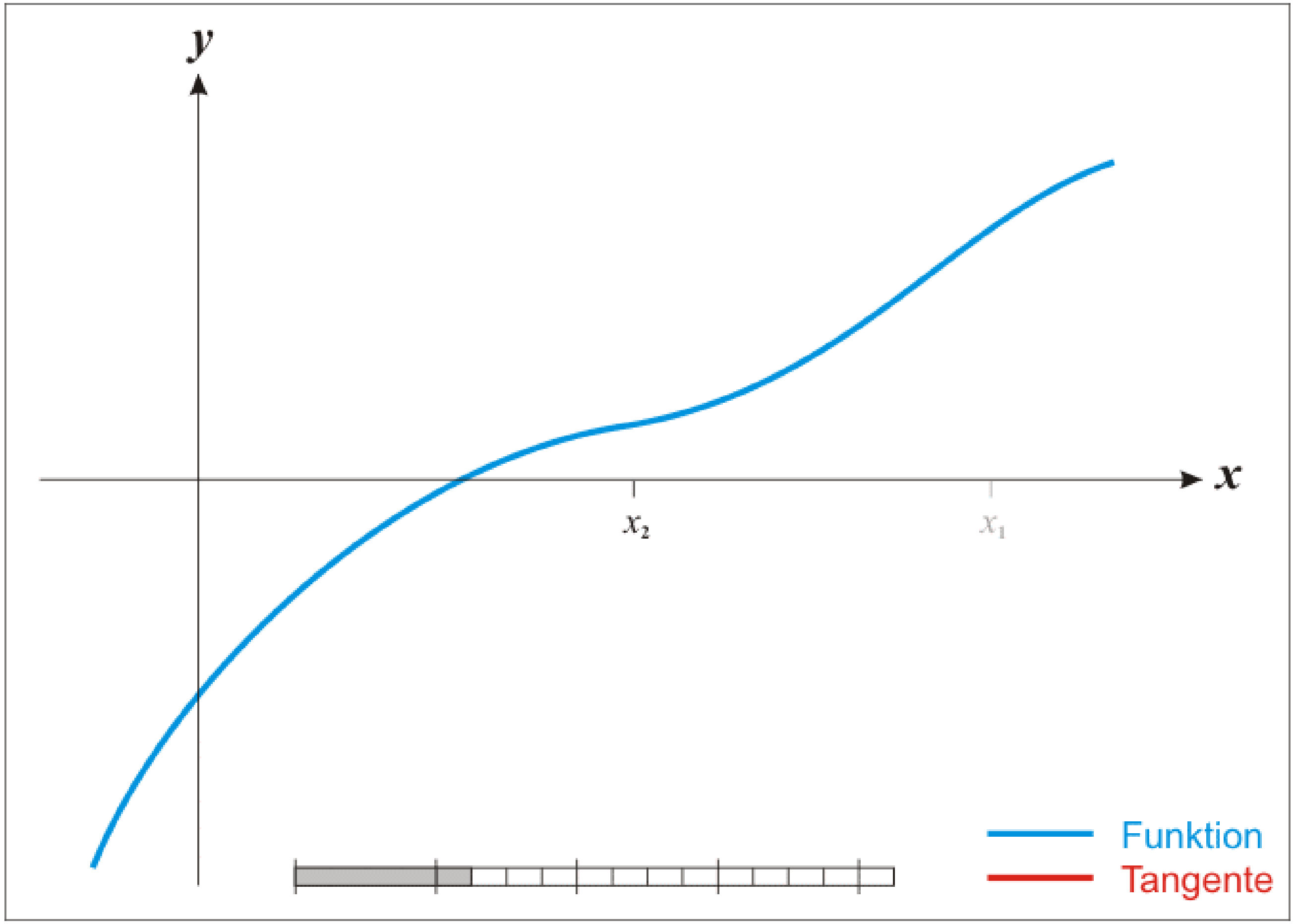


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

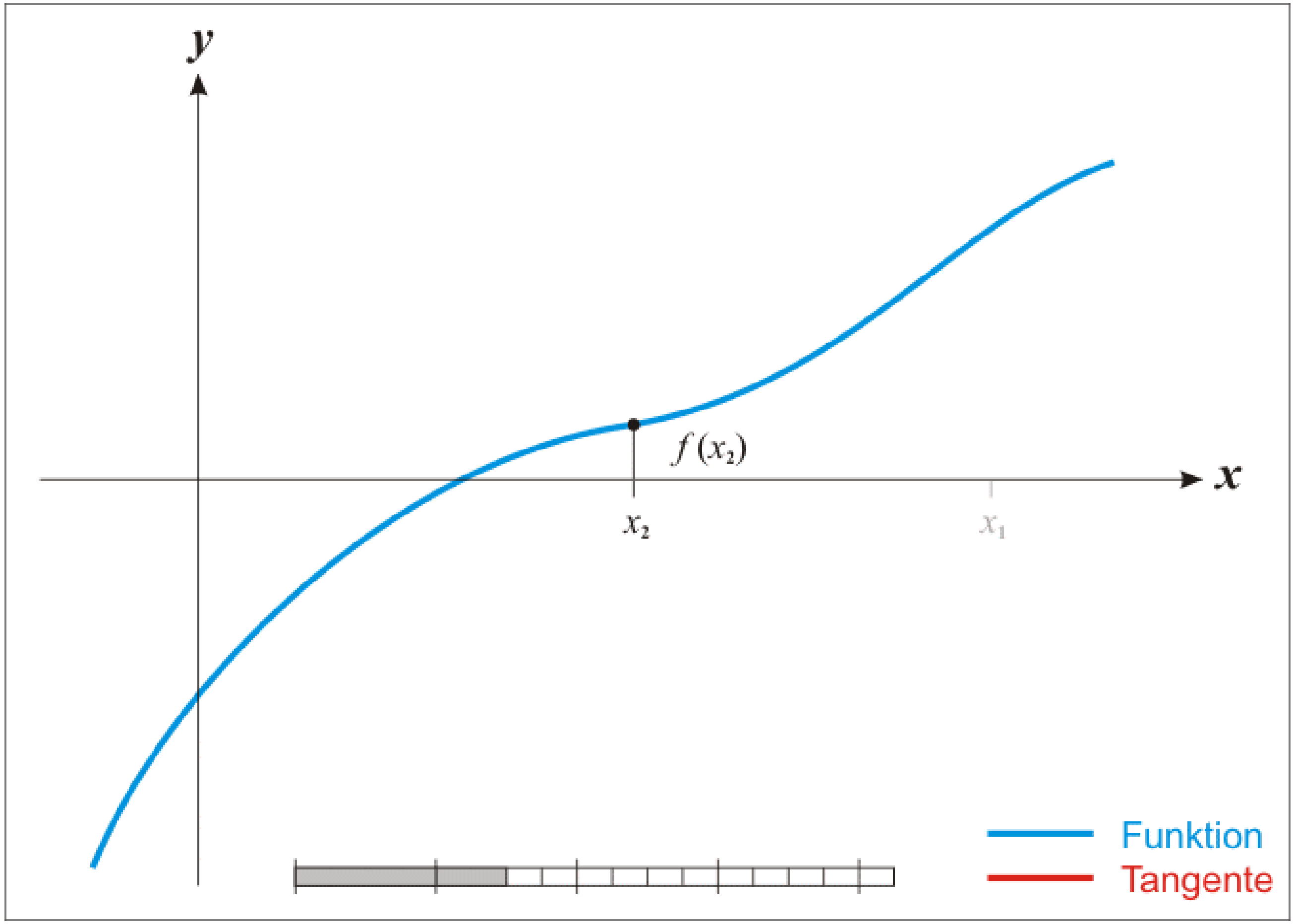


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

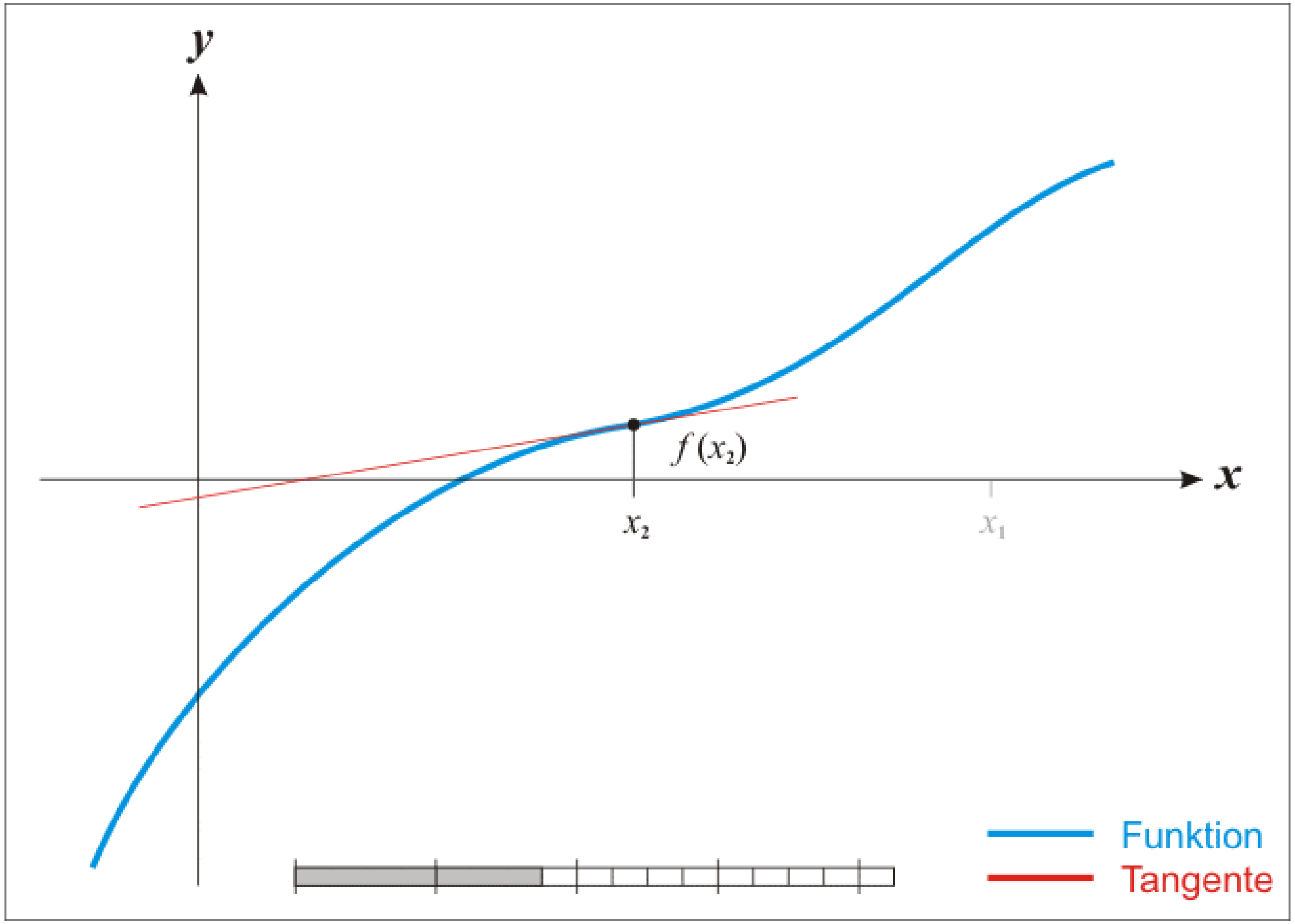


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

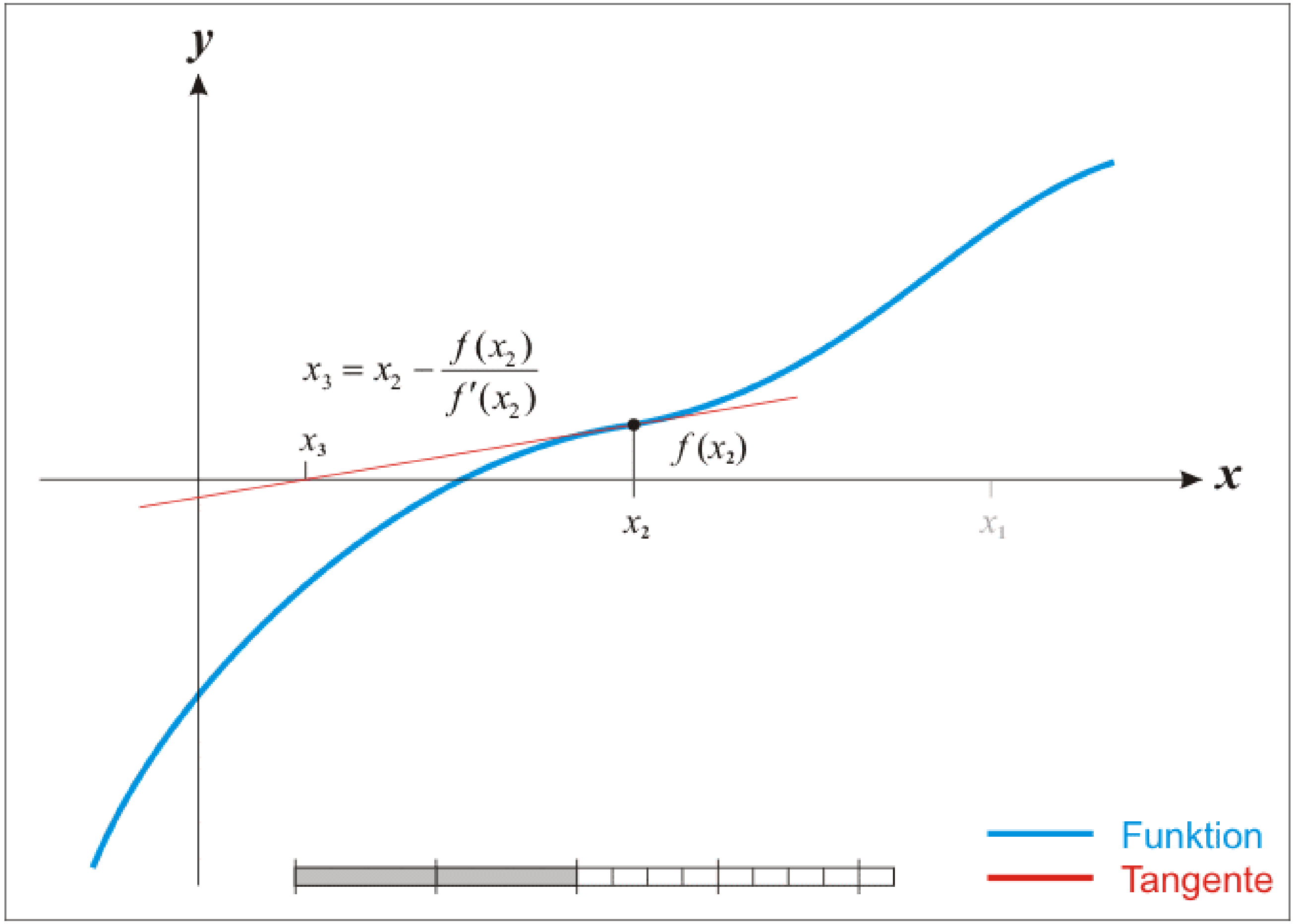


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

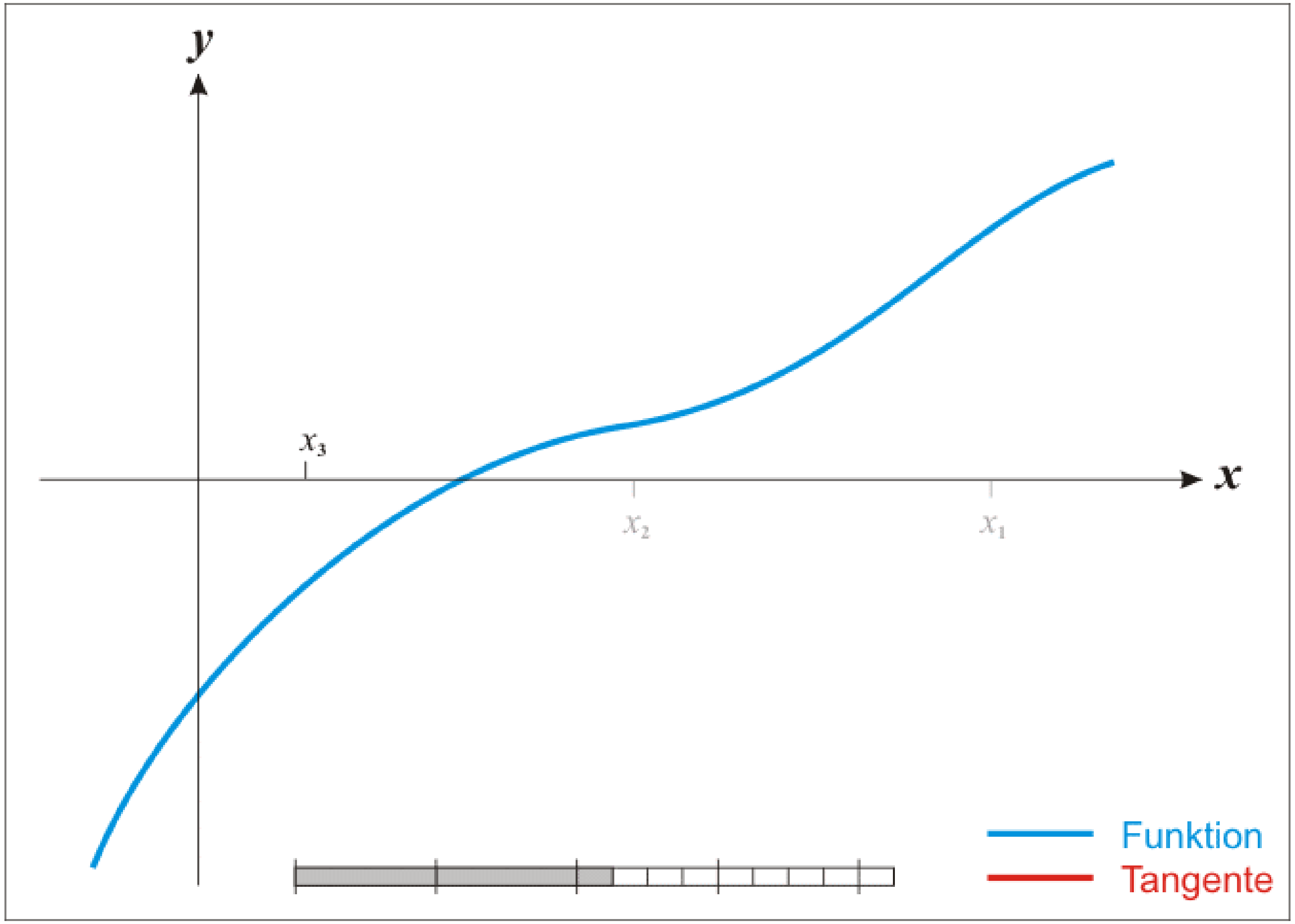


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

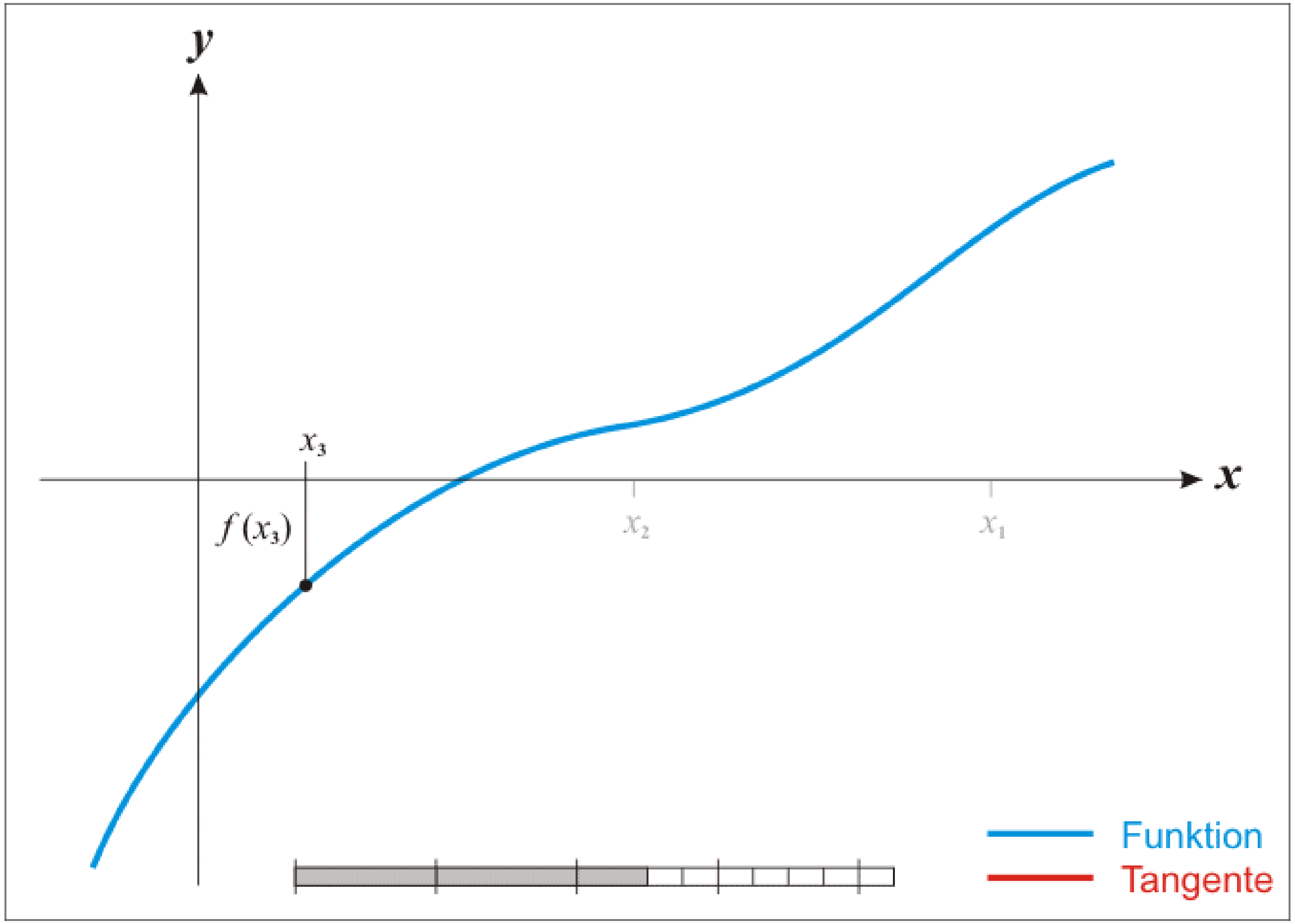


(here $x^{[n]}$ defines a sequence)

See `https://de.wikipedia.org/wiki/Newton-Verfahren`.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

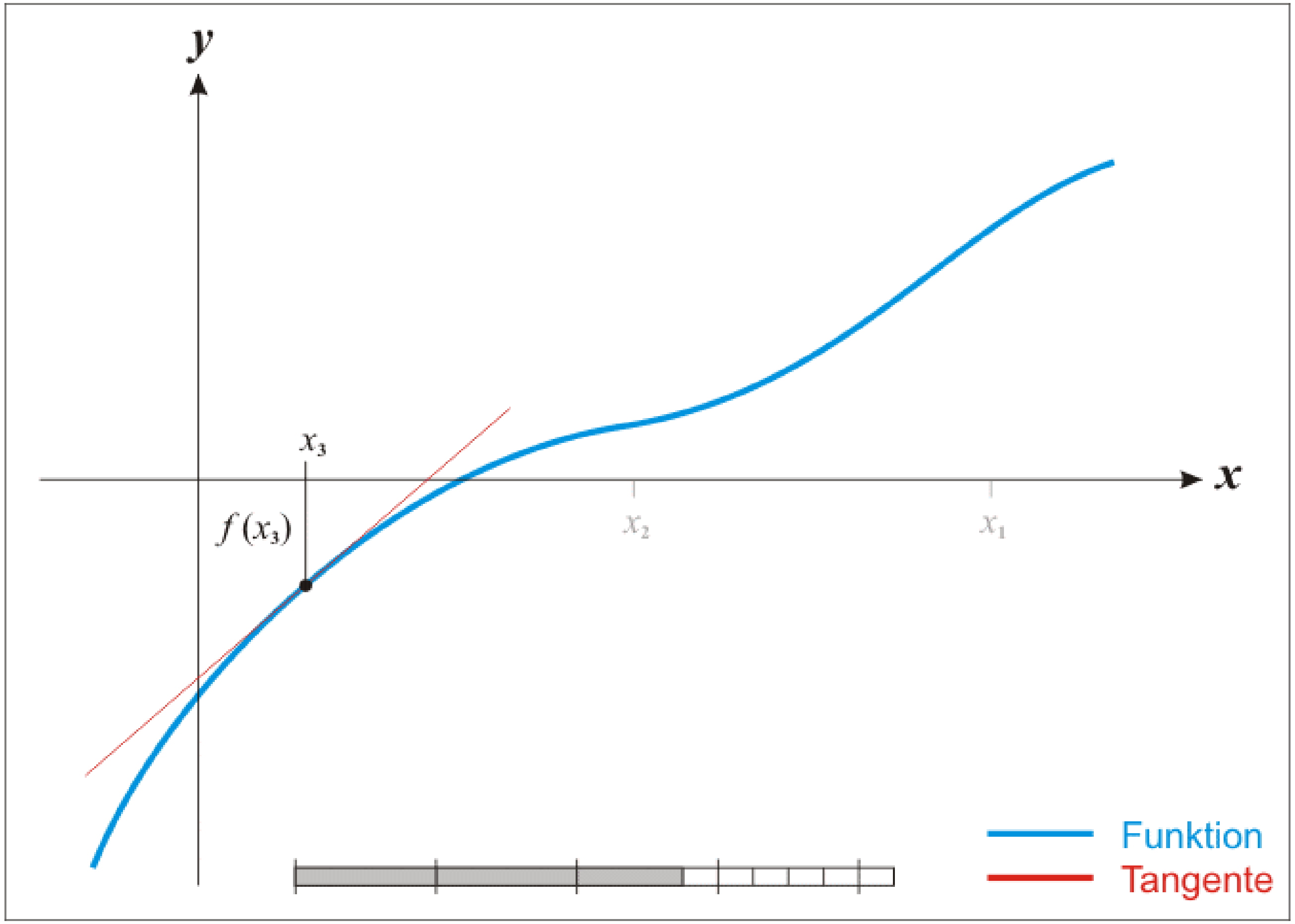


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

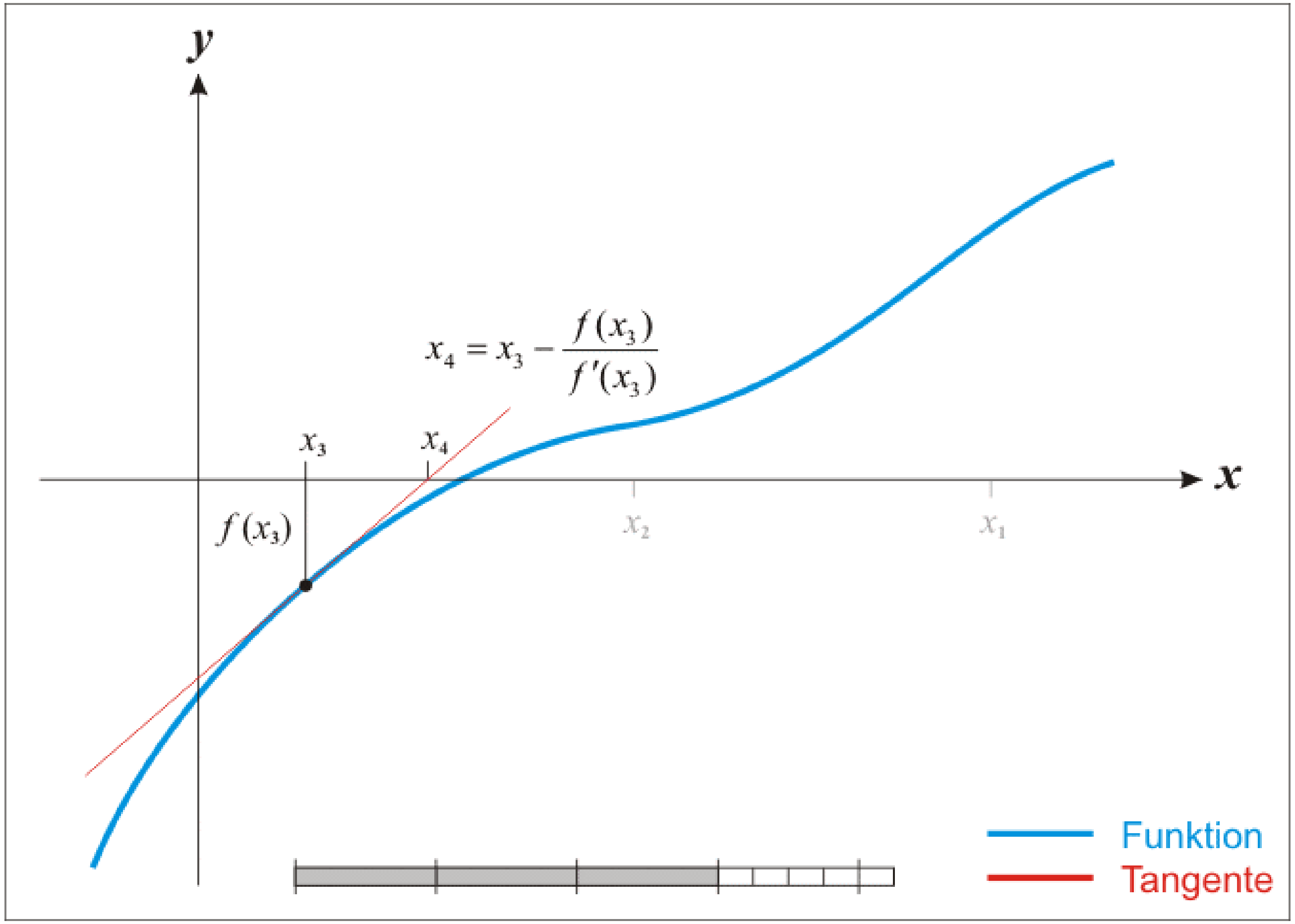


(here $x^{[n]}$ defines a sequence)

See `https://de.wikipedia.org/wiki/Newton-Verfahren`.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

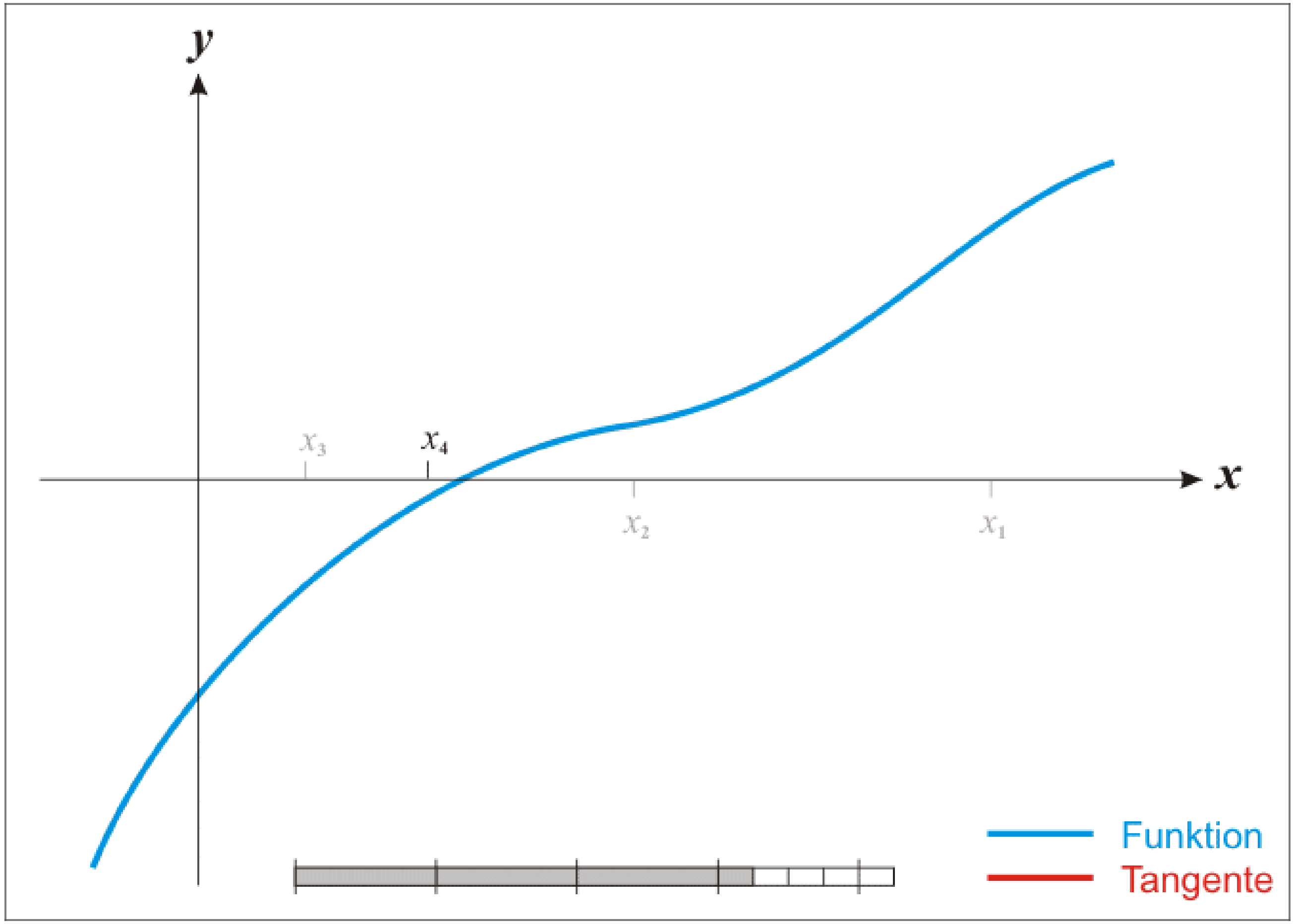


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

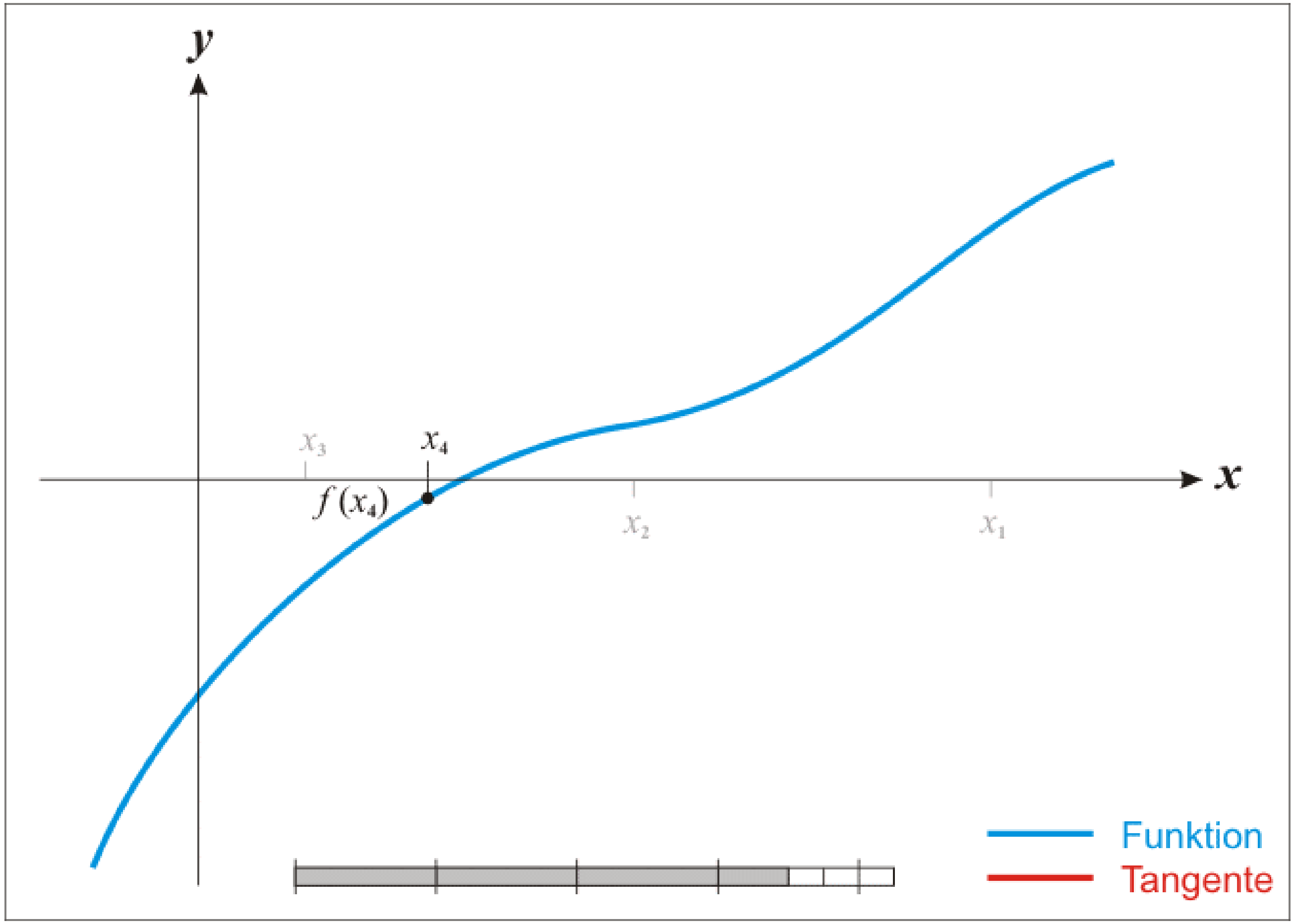


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

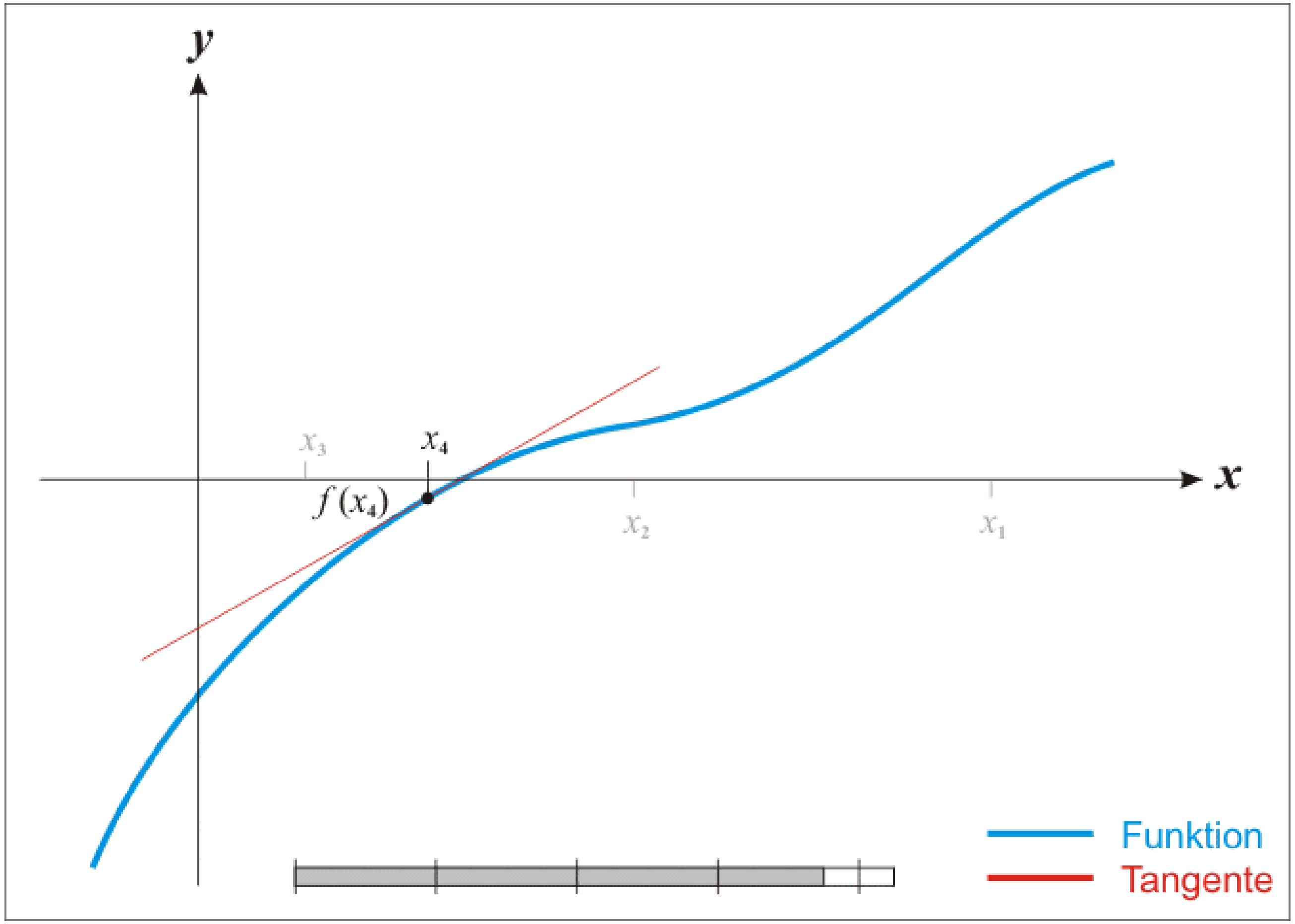


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

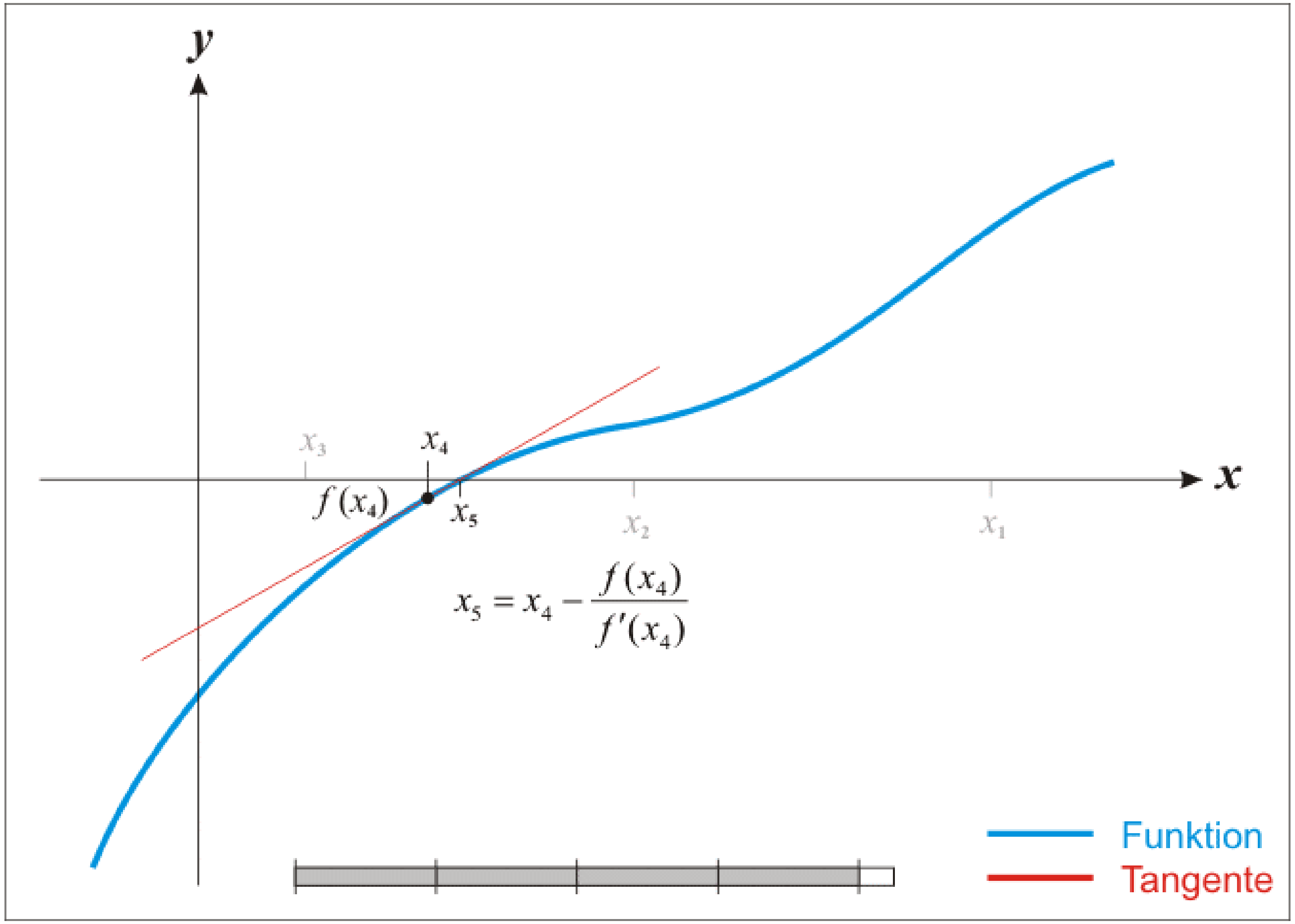


(here $x^{[n]}$ defines a sequence)

See `https://de.wikipedia.org/wiki/Newton-Verfahren`.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):

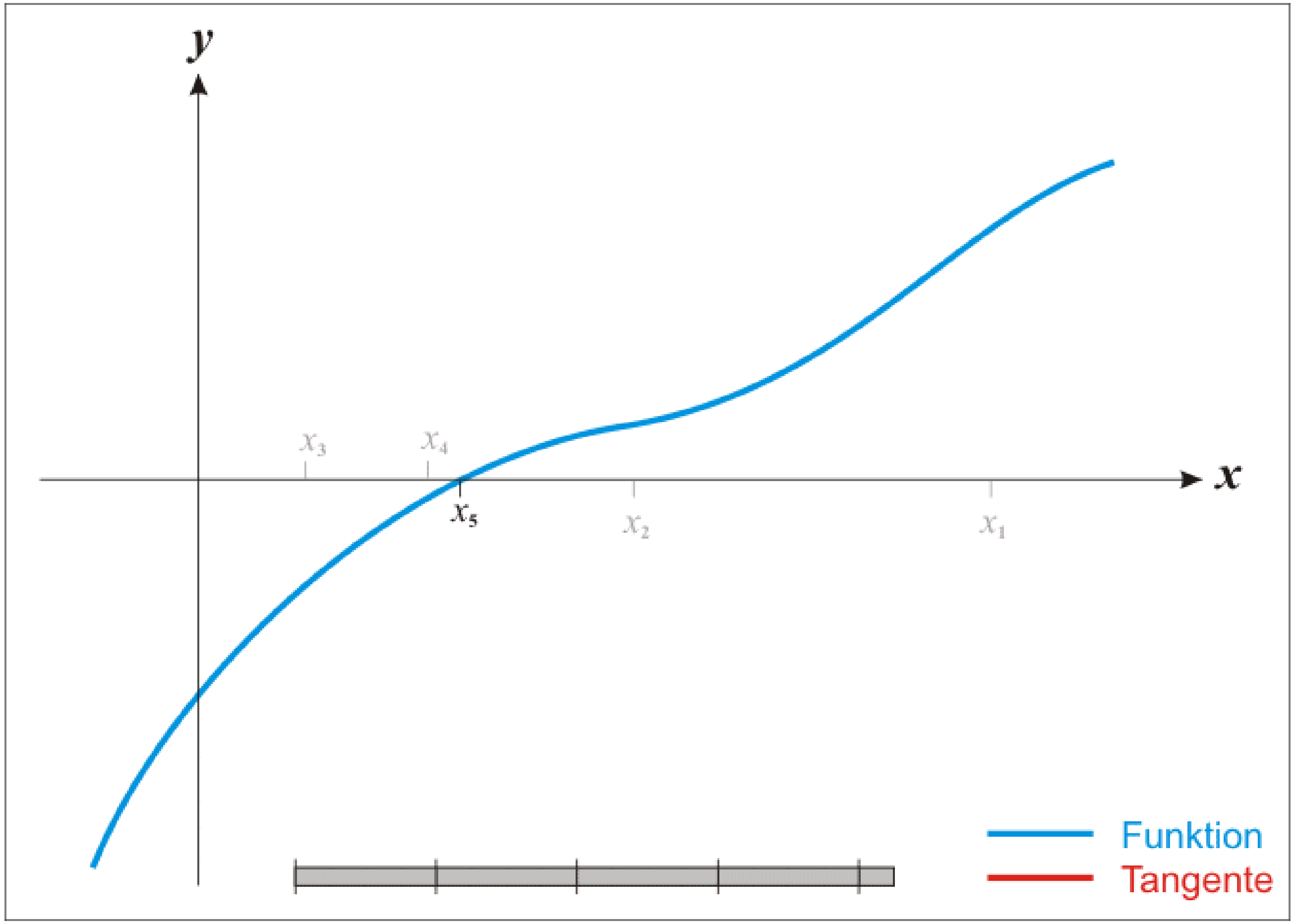


(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Newton method in 1d

As an example we consider a function f in the following (in optimization we consider typically f' !):



(here $x^{[n]}$ defines a sequence)

See https://de.wikipedia.org/wiki/Newton-Verfahren.

Problems & modifications of the Newton method

- **$\mathbf{f''(x)}$ cannot be computed explicitly.**

  $\Longrightarrow$ Replace $f''(x)$ by the corresponding difference quotient or by a formula for numerical differentiation.

- **The Newton method diverges.**

  $\Longrightarrow$ Combine the Newton method with a "safe" method as, e.g., the bisection search or reduce step size (damped Newton method)

- **The method converges to a saddle point.**

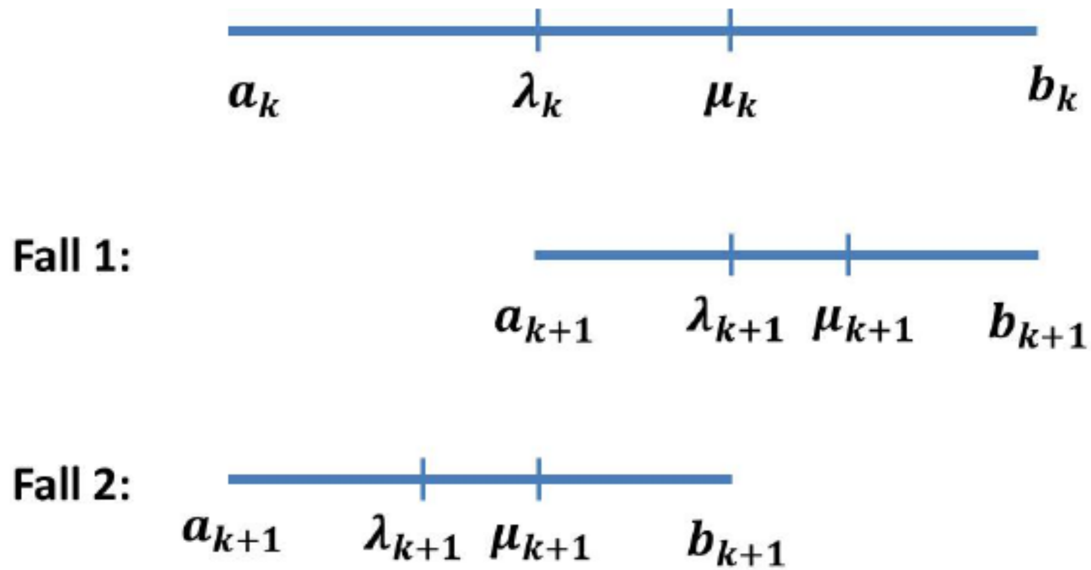
  $\Longrightarrow$ Start with another initial value or compute a few iterations with a "slower" method.

Example: Golden-section search

- Assumption: $f(x)$ continuous with minimum in $[a, b]$
- Divide $[a, b]$ in the ratio of the golden section, i.e.

$$\lambda = a + 0.382 \cdot (b - a) \text{ und}$$
$$\mu = a + 0.618 \cdot (b - a)$$

- If $f(\lambda^{[k]}) > f(\mu^{[k]})$ go to case 1 else go to case 2 and apply the method recursively



Case 1

Case 2

Golden section: convergence

- Safe convergence of the method, if $f(x)$ is convex, e.g., on the considered interval
- Relatively slow convergence:
  $|b_{k+1} - a_{k+1}| = 0.618 \cdot |b_k - a_k|$
  For comparison, bisection search:
  $|b_{k+1} - a_{k+1}| = 0.5 \cdot |b_k - a_k|$
- Contrary to the Newton method (or the gradient descent) the golden-section search does not require derivatives.
- The special division ratio~~n~~ of the golden section saves one function evaluation in each step.