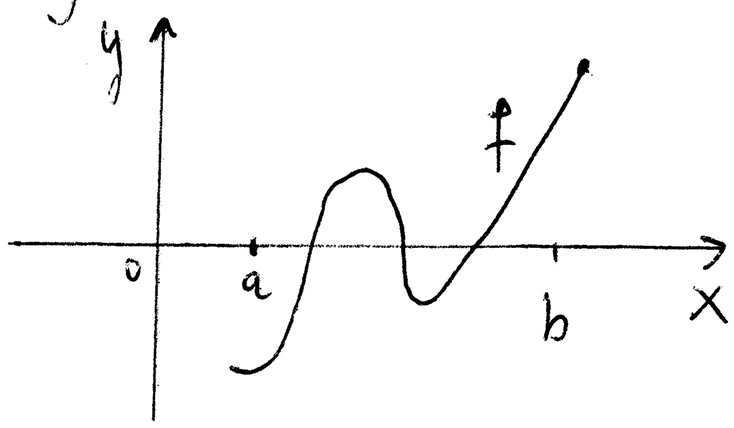
§2: Numerische Verfahren zur Berechnung von Nullstellen

In diesem § bezeichnet I ein beliebiges Intervall, $I \neq \emptyset$, $I \neq \{a\}$. Zur Erinnerung:

Zwischenwertsatz (ZWS). Sei $f: I \to \mathbb{R}$ stetig und $a, b \in I$ mit $a < b$ und $f(a) f(b) < 0$.

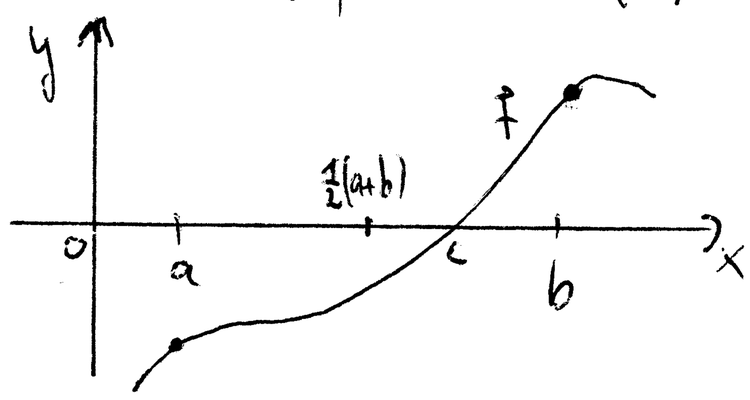
Dann existiert $c \in]a, b[$ mit $f(c) = 0$.



$\boxed{\text{Ziel: } c \text{ so gut wie möglich abzuschätzen.}}$

1. Das Bisektionsverfahren (Intervallhalbierungsverfahren)

Sei $f: I \to \mathbb{R}$ stetig, $a, b \in I$ mit $a < b$ und $f(a) < 0 < f(b)$.



Idee: Das Intervall $[a, b]$ halbieren und gucken, ob $f\left(\frac{a+b}{2}\right) > 0$ oder < 0 (oder $= 0$) ist.

Wenn (z. B) $f\left(\frac{a+b}{2}\right) < 0$ ist, dann dasselbe mit dem Intervall $\left[\frac{a+b}{2}, b\right]$ wiederholen. Etc...

Satz 2.1 (Bisektionsverfahren)

Sei $f: I \to \mathbb{R}$ stetig, $a, b \in I$ mit $a < b$ und $f(a) f(b) < 0$.

Durch Induktion über $n \in \mathbb{N}$ wird eine Folge von ineinander geschachtelten Intervallen $[x_n, y_n]$ wie folgt definiert:

i) $[x_0, y_0] = [a, b]$ (d.h. $x_0 = a$ und $y_0 = b$)

ii) Sei $[x_n, y_n]$ schon definiert.

Sei $m_n := \frac{1}{2}(x_n + y_n)$ der Mittelpunkt des Intervalls $[x_n, y_n]$.

- Wenn $f(x_n) f(m_n) < 0$, dann setzen wir $[x_{n+1}, y_{n+1}] = [x_n, m_n]$.
- Wenn $f(x_n) f(m_n) > 0$, dann setzen wir $[x_{n+1}, y_{n+1}] = [m_n, y_n]$.
- Wenn $f(x_n) f(m_n) = 0$, dann setzen wir $[x_{n+1}, y_{n+1}] = [m_n, m_n]$.

Für die Folge (x_n) gilt folgendes:

a) (x_n) konvergiert gegen eine Nullstelle c von f und es gilt $|x_n - c| \leq \frac{1}{2^n}(b-a) \quad \forall n \in \mathbb{N}$.

b) Für $K \in \mathbb{N}$ gilt $|x_n - c| < \frac{1}{10^K}$ für $n \geq \frac{10}{3}K + \frac{\ln(b-a)}{\ln(2)}$.

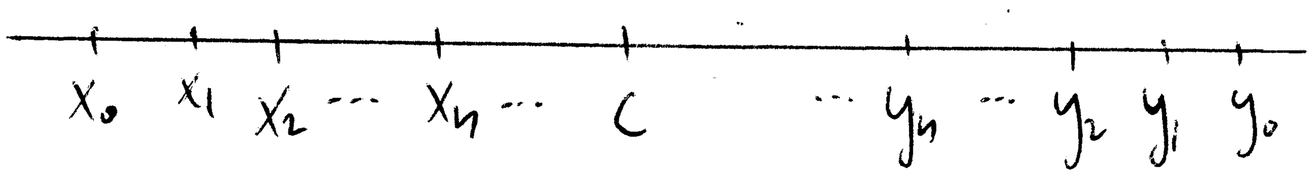
<u>Beweis</u>: <u>Zu a)</u>: Man zeigt leicht durch Induktion über $n \in \mathbb{N}$ dass $f(x_n) f(y_n) \leq 0 \quad \forall n \in \mathbb{N}$.

Da $[x_{n+1}, y_{n+1}] \subseteq [x_n, y_n] \quad \forall n \in \mathbb{N}$ ist (x_n) monoton wachsend und (y_n) monoton fallend. $a \leq x_n \leq y_n \leq b \quad \forall n \in \mathbb{N} \Rightarrow (x_n)$ und (y_n) sind beschränkt. Aus I $\Rightarrow (x_n)$ und (y_n) sind konvergent.

Seien $c := \lim_{n \to \infty} x_n$ und $\tilde{c} := \lim_{n \to \infty} y_n$. $[c, \tilde{c} \in [a, b] \subseteq I$ denn $a \leq x_n, y_n \leq b \quad \forall n \in \mathbb{N}]$

Es gilt $|x_n - c| \overset{(1)}{\leq} |x_n - y_n| \overset{(2)}{\leq} \frac{1}{2^n}|x_0 - y_0| = \frac{b-a}{2^n} \quad \forall n \in \mathbb{N}$.

<u>Begründung von (1)</u>: Für die Folgen (x_n) und (y_n) haben wir ein Bild der Form



Also liegt c zwischen x_n und y_n $\forall n \in \mathbb{N}$. Daraus folgt unmittelbar $|c - x_n| \leq |x_n - y_n|$ $\forall n \in \mathbb{N}$.

<u>Begründung von (2):</u> $|x_n - y_n| \leq \frac{1}{2}|x_{n-1} - y_{n-1}| \leq \frac{1}{2} \cdot \frac{1}{2}|x_{n-2} - y_{n-2}| =$
$\frac{1}{2^2}|x_{n-2} - y_{n-2}| \leq \ldots \leq \frac{1}{2^n}|x_0 - y_0| = \frac{b-a}{2^n}$.

Bleibt z.z: c ist Nullstelle von f.

Wir zeigen zunächst $\tilde{c} = c$. $\left[\tilde{c} := \lim_{n\to\infty} y_n\right]$

$|x_n - y_n| \leq \frac{b-a}{2^n}$ $\forall n \in \mathbb{N} \Rightarrow \lim_{n\to\infty}(x_n - y_n) = 0 \Rightarrow c - \tilde{c} = 0 \Rightarrow c = \tilde{c}$.

f stetig in $c \Rightarrow \lim_{n\to\infty} f(x_n) = f(c)$ und $\lim_{n\to\infty} f(y_n) = f(c)$.

Da $f(x_n)f(y_n) \leq 0$ $\forall n \in \mathbb{N}$ folgt $\lim_{n\to\infty}(f(x_n)f(y_n)) = f(c)^2 \leq 0$.

Dies zeigt $f(c) = 0$.

<u>Zu b):</u> Nach a) gilt $|x_n - c| \leq \frac{1}{2^n}(b-a)$ $\forall n \in \mathbb{N}$.

Also $|x_n - c| < \frac{1}{10^k}$ gilt auf jeden Fall wenn $\frac{1}{2^n}(b-a) < \frac{1}{10^k}$.

$\frac{1}{2^n}(b-a) < \frac{1}{10^k} \Leftrightarrow \frac{2^n}{b-a} > 10^k \Leftrightarrow 2^n > (b-a)10^k \Leftrightarrow n\ln(2) >$
$\ln(b-a) + k \cdot \ln(10) \Leftrightarrow n > \underbrace{\frac{\ln(10)}{\ln(2)}}_{3{,}32\ldots} k + \frac{\ln(b-a)}{\ln(2)}$. Daraus folgt:

$n \geq \frac{10}{3}k + \frac{\ln(b-a)}{\ln(2)} \Rightarrow |x_n - c| < \frac{1}{10^k}$.

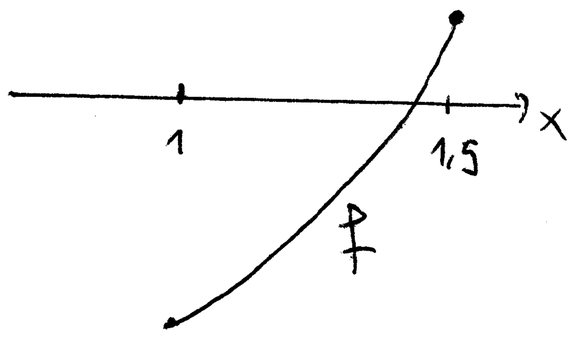
$\square$

<u>Beispiel</u>: Wir betrachten $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^2 - 2$.

Wir wollen die Nullstelle $\sqrt{2}$ mit einer Genauigkeit von 10^{-2} berechnen.

$f(-1) = -1 < 0$ und $f(3/2) = 1/4 > 0$.

Wir wählen also z.B. $[a,b] = [1; 3/2]$.



Genauigkeit $10^{-2} \overset{2.1.b)}{\Longrightarrow} n \geq \frac{10}{3} \cdot 2 + \frac{\ln(1/2)}{2} \Rightarrow \underline{n = 7}$.

Sei $m_n := \frac{1}{2}(x_n + y_n)$ der Mittelpunkt des Intervalls $[x_n, y_n]$, $n \in \mathbb{N}$.

$[x_0, y_0] = [a,b] = [1; 3/2]$. $\quad m_0 = 5/4$

$f\left(\frac{5}{4}\right) = \frac{25}{16} - 2 = -\frac{7}{16} < 0 \Rightarrow [x_1, y_1] = \left[\frac{5}{4}, \frac{3}{2}\right]$. $\quad m_1 = \frac{11}{8}$

$f\left(\frac{11}{8}\right) = \left(\frac{11}{8}\right)^2 - 2 = \frac{121}{64} - \frac{128}{64} < 0 \Rightarrow [x_2, y_2] = \left[\frac{11}{8}, \frac{3}{2}\right]$. $\quad m_2 = \frac{23}{16}$

$f\left(\frac{23}{16}\right) = \left(\frac{23}{16}\right)^2 - 2 = \frac{529}{256} - \frac{512}{256} > 0 \Rightarrow [x_3, y_3] = \left[\frac{11}{8}, \frac{23}{16}\right]$. $\quad m_3 = \frac{45}{32}$

$f\left(\frac{45}{32}\right) < 0 \Rightarrow [x_4, y_4] = \left[\frac{45}{32}, \frac{23}{16}\right]$. $\quad m_4 = \frac{91}{64}$

$f\left(\frac{91}{64}\right) > 0 \Rightarrow [x_5, y_5] = \left[\frac{45}{32}, \frac{91}{64}\right]$. $\quad m_5 = \frac{181}{128}$

$f\left(\frac{181}{128}\right) < 0 \Rightarrow [x_6, y_6] = \left[\frac{181}{128}, \frac{91}{64}\right]$ $\quad m_6 = \frac{363}{256}$

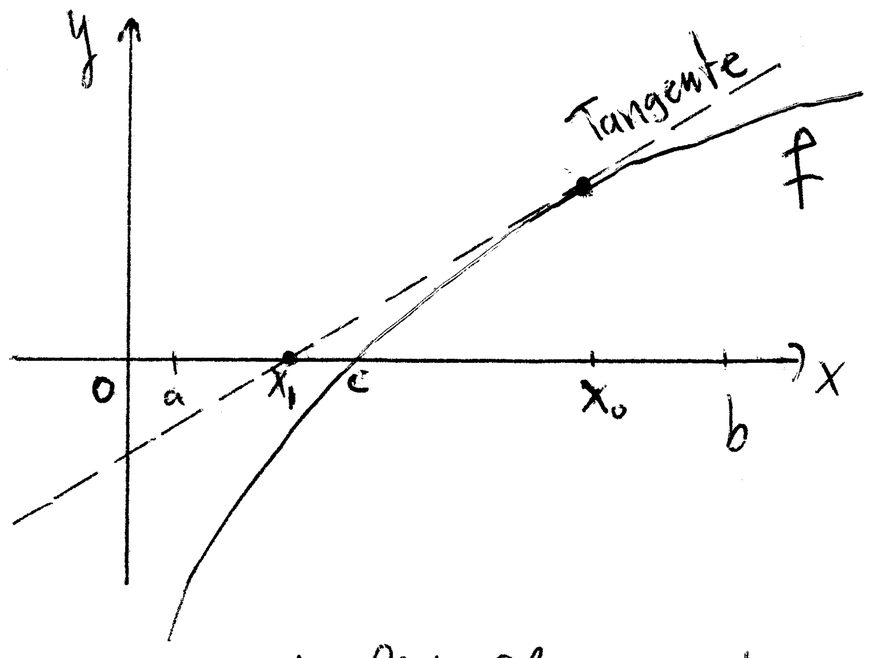
$f\left(\frac{363}{256}\right) > 0 \Rightarrow [x_7, y_7] = \left[\frac{181}{128}, \frac{363}{256}\right]$.

Also $\left|\sqrt{2} - x_7\right| = \left|\sqrt{2} - \frac{181}{128}\right| = \left|\sqrt{2} - 1{,}4140625\right| < 0{,}01$.

<u>Bem</u>: Das Bisektionsverfahren funktioniert für jede stetige Funktion $f: [a,b] \to \mathbb{R}$ mit $f(a)f(b) < 0$, aber die Folge (x_n) konvergiert sehr langsam gegen eine Nullstelle von f.

2. Das Newtonsche Verfahren (Isaac Newton, 1643-1727)

Sei $f: I \to \mathbb{R}$ diff. und $a, b \in I$ mit $a < b$ und $f(a)f(b) < 0$.

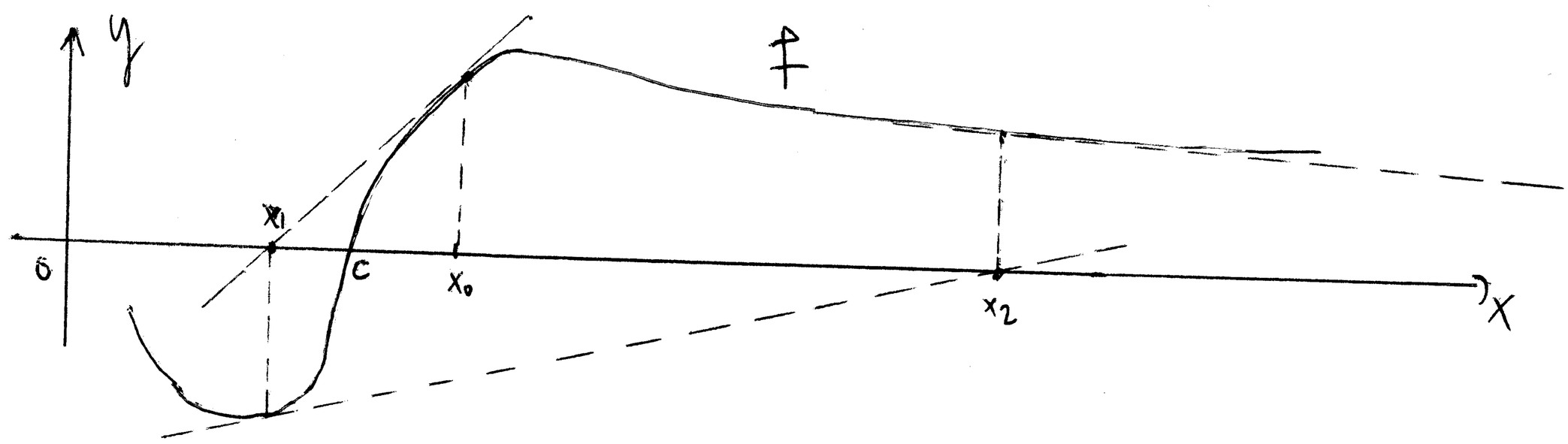


Idee: Sei $x_0 \in [a,b]$ beliebig. Betrachte die Tangente $y = f(x_0) + f'(x_0)(x - x_0)$ von f im Punkt $(x_0, f(x_0))$. Sei x_1 die Nullstelle von der Tangente: $f(x_0) + f'(x_0)(x_1 - x_0) = 0 \Rightarrow$

$x_1 - x_0 = -f(x_0)/f'(x_0) \Rightarrow x_1 = x_0 - f(x_0)/f'(x_0)$

Falls $f'(x_0) \neq 0$ ist x_1 ein Näherungswert von der Nullstelle c. Dann wird dieses Verfahren mit x_1 iteriert und wir erhalten ein x_2. Und so weiter..

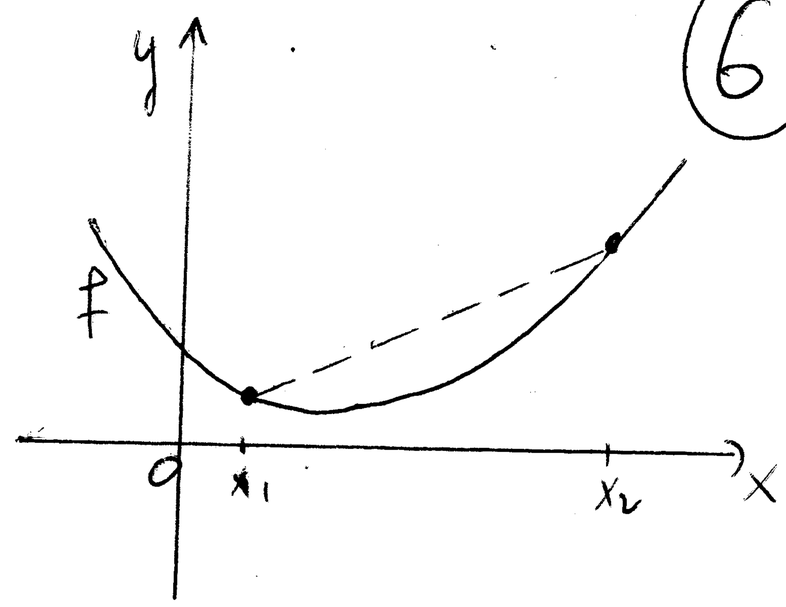
Bem: Im Allgemeinen braucht das Verfahren nicht zu konvergieren.



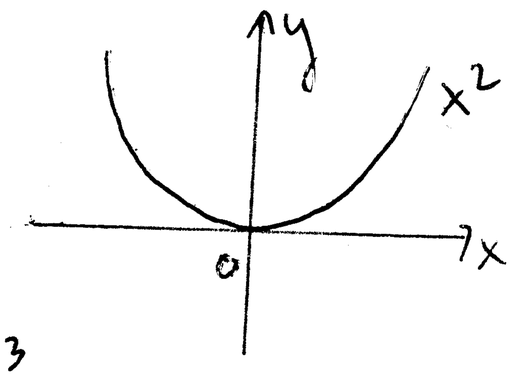
Def: Sei $f: I \to \mathbb{R}$.

- f heißt konvex wenn $\forall x_1, x_2 \in I$, $\forall \lambda \in [0,1]$ gilt
$$f(\lambda x_1 + (1-\lambda)x_2) \overset{(1)}{\leq} \lambda f(x_1) + (1-\lambda) f(x_2).$$

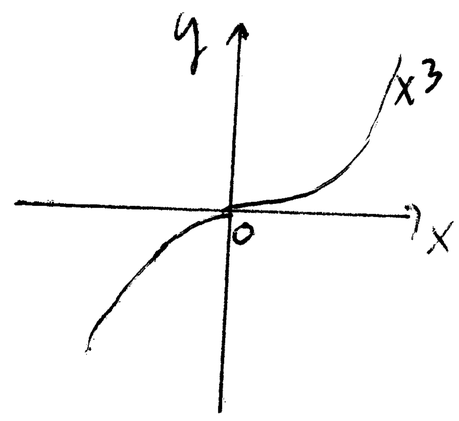
- f heißt konkav wenn $-f$ konvex ist.



Bem: (1) bedeutet geometrisch, dass der Graph von f im Intervall $[x_1, x_2]$ unterhalb der Sekante durch $(x_1, f(x_1))$ und $(x_2, f(x_2))$ liegt (s. Übung).



Beispiele: a) $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, ist konvex.



b) $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$, ist nicht konvex

[Wähle z.B. $x_1 = -1$, $x_2 = 1$, $\lambda = 3/4$]

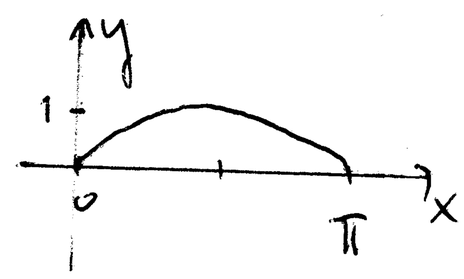
Mittels des Mittelwertsatzes (MWS) kann man zeigen:

Lemma 2.2: Sei $f: I \to \mathbb{R}$ 2-mal diff. Dann

- f ist konvex $\Leftrightarrow f''(x) \geq 0 \ \forall x \in I$
- f ist konkav $\Leftrightarrow f''(x) \leq 0 \ \forall x \in I$.

$\square$

Beispiel: Sei $f: [0, \pi] \to \mathbb{R}$, $f(x) = \sin x$. f ist konkav denn $f''(x) = (\cos x)' = -\sin x \leq 0 \ \forall x \in [0, \pi]$.



Satz 2.3: (Newtonsche Verfahren) [Zum Beweis, siehe z.B. Otto Forster, Analysis I]

Es sei $f: [a, b] \to \mathbb{R}$ eine 2-mal diff. konvexe oder konkave Funktion mit $f(a)f(b) < 0$. Dann gilt

a) Es gibt genau ein $c \in]a,b[$ mit $f(c)=0$.

b) Ist x_0 ein beliebiger Punkt aus $[a,b]$ mit

$$f(x_0) \begin{cases} \geq 0 & \text{wenn } f \text{ konvex} \\ \leq 0 & \text{wenn } f \text{ konkav,} \end{cases}$$

so ist die Folge (x_n) definiert durch $x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$ $(n \in \mathbb{N})$ wohldefiniert und konvergiert monoton gegen c.

c) Gilt $|f'(x)| \geq m > 0$ und $|f''(x)| \leq K$ für alle $x \in [a,b]$, so hat man für jedes $n \geq 1$ die Abschätzungen

$$\boxed{|x_{n+1} - x_n| \overset{(+)}{\leq} |c - x_n| \overset{(++)}{\leq} \frac{K}{2m} |x_n - x_{n-1}|^2.}$$

<u>Bem</u>: i) Unter "(x_n) wohldefiniert" in b) versteht man:

$f'(x_n) \neq 0$ und $x_{n+1} := x_n - f(x_n)/f'(x_n) \in [a,b]$ $\forall n \in \mathbb{N}$.

ii) Ang. $\frac{K}{2m} \approx 1$. c) zeigt: Stimmen x_n und x_{n-1} auf k Dezimalstellen überein, so ist der Näherungswert x_n auf $2k$ Dezimalstellen genau und bei jedem weiteren Iterationsschritt verdoppelt sich die Zahl der günstigen Stellen.

<u>Begründung</u>: • $|x_n - x_{n-1}| < 10^{-k} \overset{(++)}{\Longrightarrow} |c - x_n| \leq \frac{K}{2m} |x_n - x_{n-1}|^2 \approx |x_n - x_{n-1}|^2 < 10^{-2k}$.

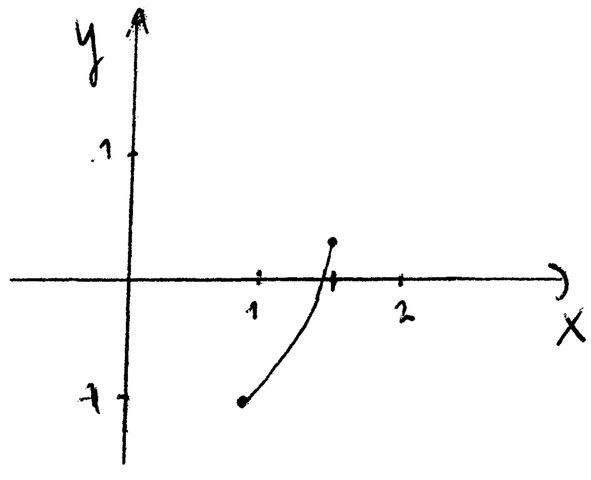
Also $|x_n - x_{n-1}| < 10^{-k} \Rightarrow |c - x_n| < 10^{-2k}$.

s. obige Zeile mit $n \to n+1$

• $|c - x_n| < 10^{-k} \overset{(+)}{\Longrightarrow} |x_{n+1} - x_n| < 10^{-k} \Longrightarrow |c - x_{n+1}| < 10^{-2k}$.

Beispiele:

1) Berechnung von $\sqrt{2}$ mit einer Genauigkeit von 10^{-8}.



Sei $f: [1, 1{,}5] \to \mathbb{R}$, $f(x) = x^2 - 2$.

$f(1) = -1 < 0$, $f(1{,}5) = 0{,}25 > 0$.

$f''(x) = 2 \geq 0 \quad \forall x \in [1, 1{,}5] \Rightarrow f$ ist konvex.

$f'(x) = 2x \geq 2 \quad \forall x \in [1, 1{,}5] \Rightarrow$ 2.3c) ist erfüllt für $(m, K) = (2, 2)$.

Wir wählen $x_0 = 1{,}5$. Die Folge (x_n) ist definiert durch

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - 2}{2x_n} = \frac{x_n^2 + 2}{2x_n}, \text{ also } x_{n+1} = \frac{x_n^2 + 2}{2x_n}.$$

$x_0 = 3/2$. $\quad x_1 = \dfrac{(3/2)^2 + 2}{3} = \dfrac{9/4 + 8/4}{3} = \dfrac{17}{12} = 1{,}41\overline{6}$.

$$x_2 = \frac{(17/12)^2 + 2}{17/6} = \frac{6}{17}\left(\frac{289}{144} + 2\right) = \frac{577}{408} = 1{,}4142156\ldots$$

$$x_3 = \frac{(577/408)^2 + 2}{2 \cdot 577/408} = 1{,}414213562\ldots$$

$$|x_3 - x_2| < 10^{-5} \overset{2.3.c)(++)}{\Longrightarrow} |\sqrt{2} - x_3| \leq \frac{2}{2 \cdot 2}|x_3 - x_2|^2 < \frac{1}{2}10^{-10} < 10^{-8}.$$

2) Die reellen Nullstellen von $x^5 - x - 1/5$

Sei $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^5 - x - 1/5$.

Wir zeigen zunächst: f hat genau 3 reelle (einfache) Nullstellen.

$$f'(x) = 5x^4 - 1 = \begin{cases} > 0 & \text{für } |x| > 1/\sqrt[4]{5} \\ 0 & \text{für } |x| = 1/\sqrt[4]{5} \\ < 0 & \text{für } |x| < 1/\sqrt[4]{5}. \end{cases}$$

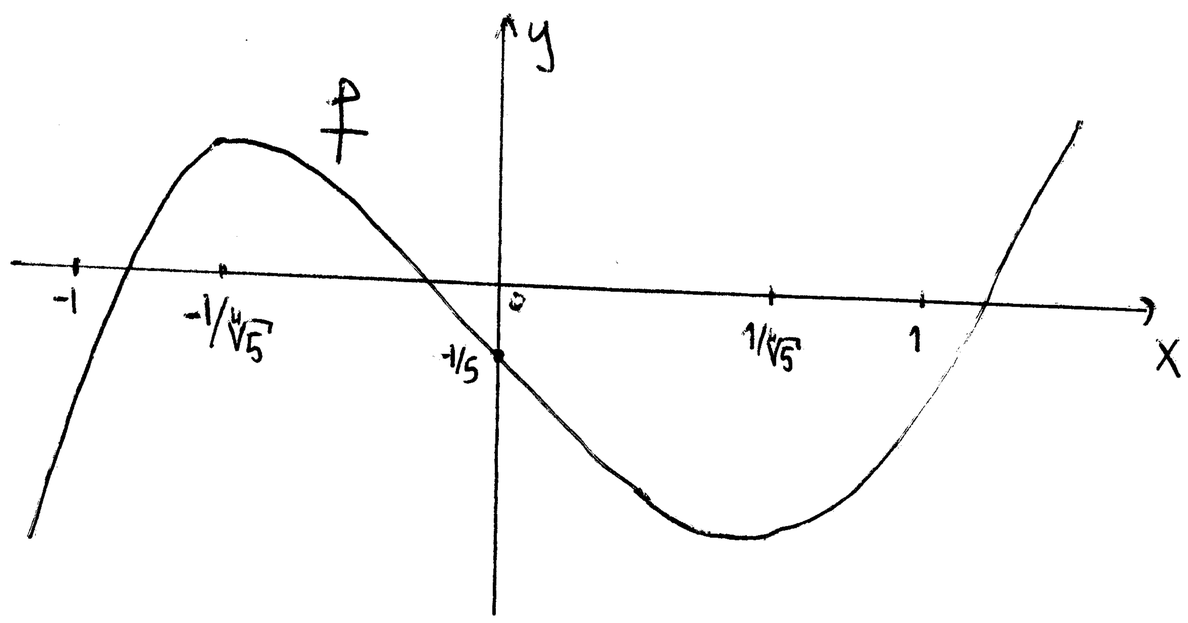
Daraus folgt:

$f(x)$ ist streng monoton $\begin{cases} \text{wachsend auf }]-\infty, -1/\sqrt[4]{5}] \\ \text{fallend auf } [-1/\sqrt[4]{5}, 1/\sqrt[4]{5}] \\ \text{wachsend auf } [\frac{1}{\sqrt[4]{5}}, \infty[. \end{cases}$

$f\left(\frac{1}{\sqrt[4]{5}}\right) = \underbrace{\frac{1}{5^{5/4}} - \frac{1}{5^{1/4}}}_{<0} - \frac{1}{5} < -\frac{1}{5} < 0$; $f\left(-\frac{1}{\sqrt[4]{5}}\right) = -\frac{1}{5^{5/4}} + \frac{1}{5^{1/4}} - \frac{1}{5} = \frac{-1+5-5^{1/4}}{5^{5/4}}$

$= \frac{4-\sqrt[4]{5}}{5^{5/4}} > 0$. Man sieht leicht: $f(-1) = f(0) = f(1) = -1/5$.

Es reicht für eine Skizze von f:



Die Funktion f hat genau eine Nullstelle $\begin{cases} \tilde{x}_1 \text{ in }]-\infty, -1/\sqrt[4]{5}[\\ \tilde{x}_2 \text{ in }]-1/\sqrt[4]{5}, 1/\sqrt[4]{5}[\\ \tilde{x}_3 \text{ in }]1/\sqrt[4]{5}, \infty[. \end{cases}$

Wir berechnen $\tilde{x}_3$ mit einer Genauigkeit von 10^{-9}.

$f(+1) = -\frac{1}{5} < 0$ und $f(1{,}1) = (1+0{,}1)^5 - 1{,}1 - 0{,}2 \overset{\text{Bern. Ungl.}}{\geq} 1 + 5 \cdot 0{,}1 - 1{,}3 = 0{,}2 > 0$.

Also liegt $\tilde{x}_3$ in $[1; 1{,}1]$. Wir betrachten f auf dem Intervall $[a,b] := [1; 1{,}1]$. f ist konvex denn $f''(x) = 20x^3 \geq 0 \ \forall x \in [1; 1{,}1]$.

$|f'(x)| = 5x^4 - 1 \geq 5 - 1 = 4 \ \forall x \in [1, 1{,}1]$ und

$|f''(x)| \underset{x>0}{=} 20x^3 \leq 20 \cdot 1{,}1^3 \leq 20 \cdot \frac{4}{3} \leq 27 \ \forall x \in [1, 1{,}1]$.

2.3c) ist erfüllt mit $(m,K) = (4,27)$. $K/2m = 27/8 = 3{,}5$. (10)

Wir starten mit $x_0 = 1{,}1$.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^5 - x_n - 1/5}{5x_n^4 - 1} = \frac{4x_n^5 + 1/5}{5x_n^4 - 1} \quad (n \in \mathbb{N}).$$

$$x_1 = \frac{4 \cdot 1{,}1^5 + 0{,}2}{5 \cdot 1{,}1^4 - 1} \approx 1{,}051$$

$$x_2 = \frac{4 \cdot 1{,}051^5 + 0{,}2}{5 \cdot 1{,}051^4 - 1} \approx 1{,}0448$$

$$x_3 = \frac{4 \cdot 1{,}0448^5 + 0{,}2}{5 \cdot 1{,}0448^4 - 1} \approx 1{,}044762$$

$$x_4 = \frac{4 \cdot 1{,}044762^5 + 0{,}2}{5 \cdot 1{,}044762^4 - 1} \approx 1{,}044761700$$

$$|x_4 - x_3| < 10^{-6} \underset{\substack{\uparrow \\ 2.3c)(++)}}{\Longrightarrow} |\tilde{x}_3 - x_4| \leq 3{,}5 \cdot |x_4 - x_3|^2 < 3{,}5 \cdot (10^{-6})^2 = 3{,}5 \cdot 10^{-12} < 10^{-9}.$$