

Algorithmik kontinuierlicher Systeme

Iterative Verfahren: Allgemeines, Fixpunkt-Iteration, Nullstellen



- Viele numerische Probleme lassen sich nicht mit endlich vielen Schritten lösen
 - ▶ Nullstellen (von Polynomen), Eigenwerte von Matrizen
 - ▶ Optimierung (min – max-Suche)
- DIE SVD IST SEHR AUFWENDIG!!
- **Iterativer Lösungsansatz:**
 - ▶ Spezifiziere einen geschätzten Wert: Startwert x_0
 - ▶ Versuche diesen sukzessive zu verbessern $x_{i+1} = \Phi_i(x_i)$
 oder mehrstufig $x_{i+1} = \Phi_i(x_i, x_{i-1}, \dots)$ für $i=0,1,2,3, \dots$
 - Ist die Iterationsvorschrift Φ_i nicht von i abhängig spricht man von **stationären** Verfahren
 - Der iterative Ansatz ist u.U. auch für „exakt lösbare“ Probleme interessant (z.B. LGS, siehe später)

- ▶ Wir beschäftigen uns nur mit stationären Verfahren (meist einstufig):

Startwert x_0 ,

gleiches phi in jedem schritt

$$x_{i+1} = \Phi(x_i) \quad (\text{für } i=0,1,2,\dots)$$

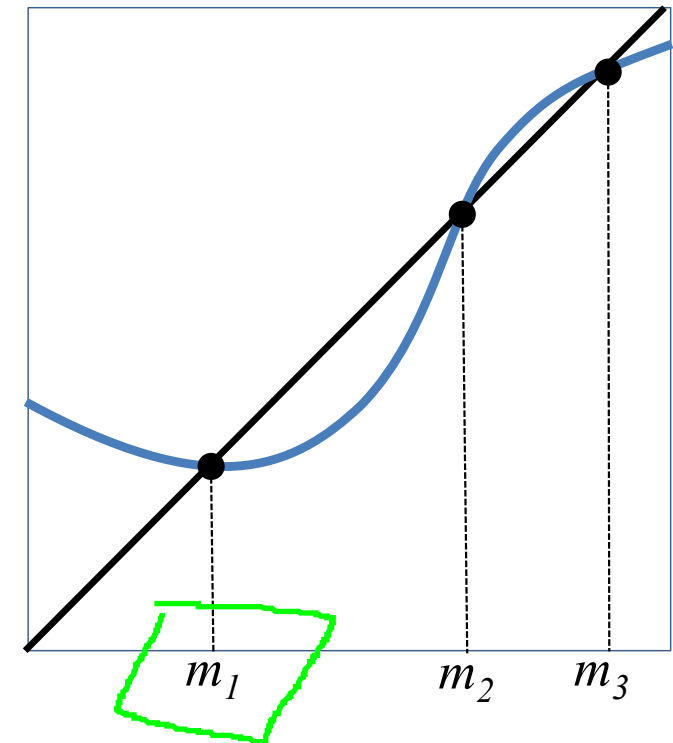
- Fragen:

- ▶ Konvergiert die Iterationsfolge gegen die gewünschte Lösung x^* ?
- ▶ Für welche Anfangswerte konvergiert die Folge?
- ▶ Wie schnell konvergiert die Iterationsfolge?
- ▶ Kann man den Fehler $|x_i - x^*|$ abschätzen?
- ▶ Wann soll man die Iteration abbrechen?

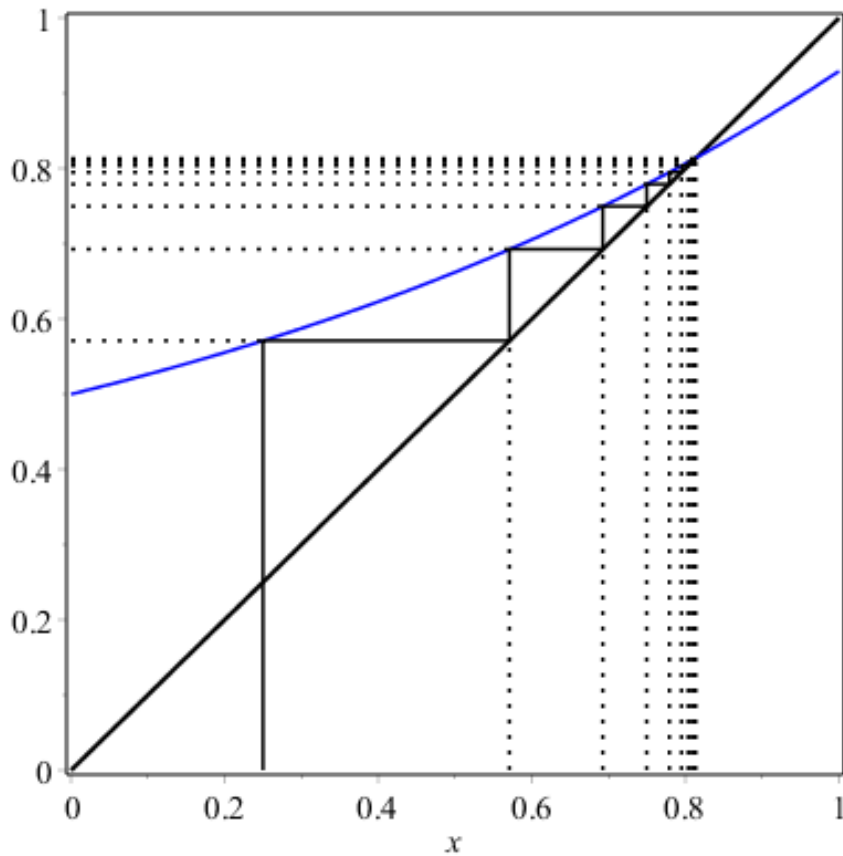
natürlich ohne x^* zu kennen

- Ist $\Phi : M \rightarrow M$ eine „Selbst“-Abbildung, dann heißt ein $m \in M$ **Fixpunkt** von Φ falls $\Phi(m) = m$.

- Anschaulich: Ist $M = I$ ein Intervall, dann verläuft der Graph von Φ im Quadrat $I \times I$ und ein **Fixpunkt** ist ein **Schnittpunkt mit der Diagonalen**



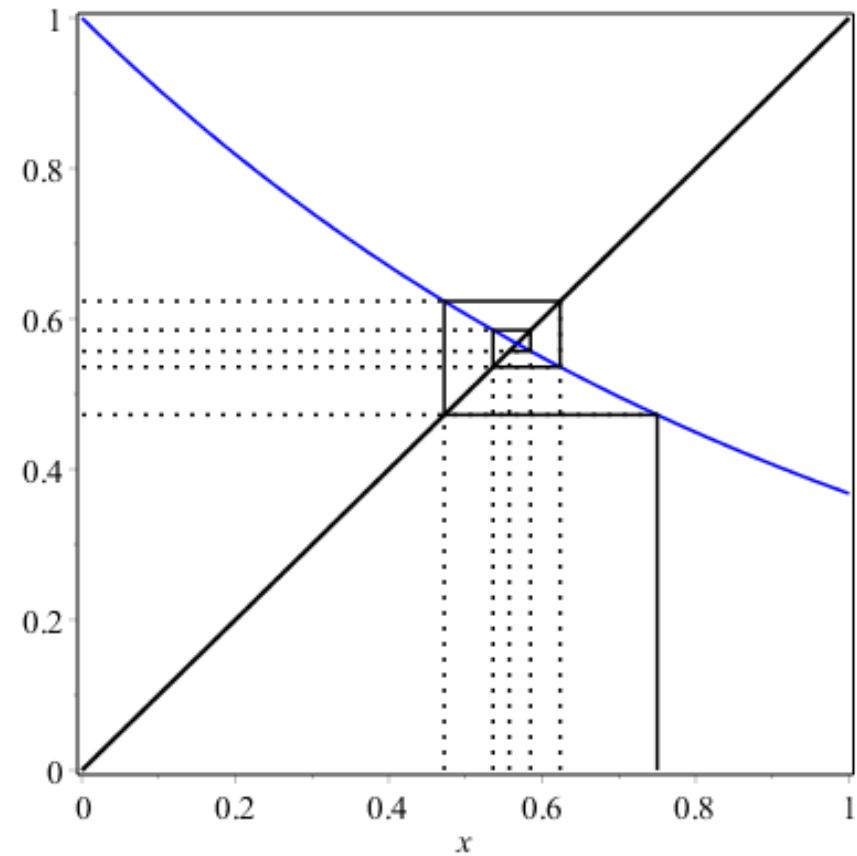
- Satz:** Wenn die Iterationsfolge $x_{i+1} = \Phi(x_i)$ konvergiert, $x^* = \lim_{i \rightarrow \infty} x_i$ und Φ stetig ist, dann gilt $\Phi(x^*) = x^*$!



x_0 x_1 x_2 x_3

$$\Phi(x) = \frac{1 + e^x}{4}, x_0 = 0.25$$

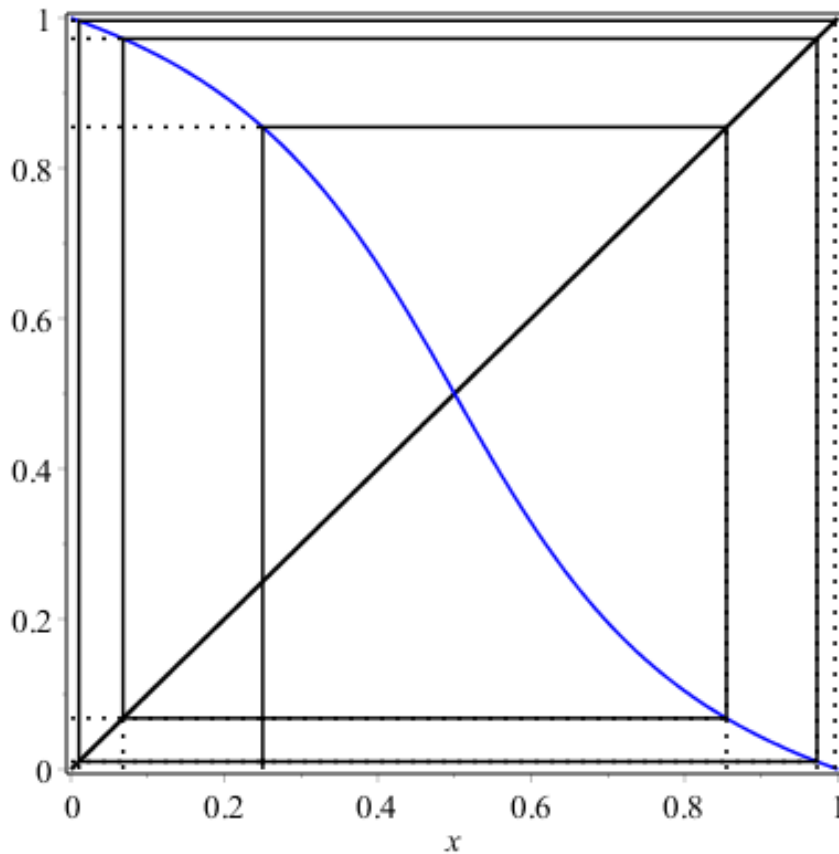
Konvergenz



x_1 x_3 x_2 x_0

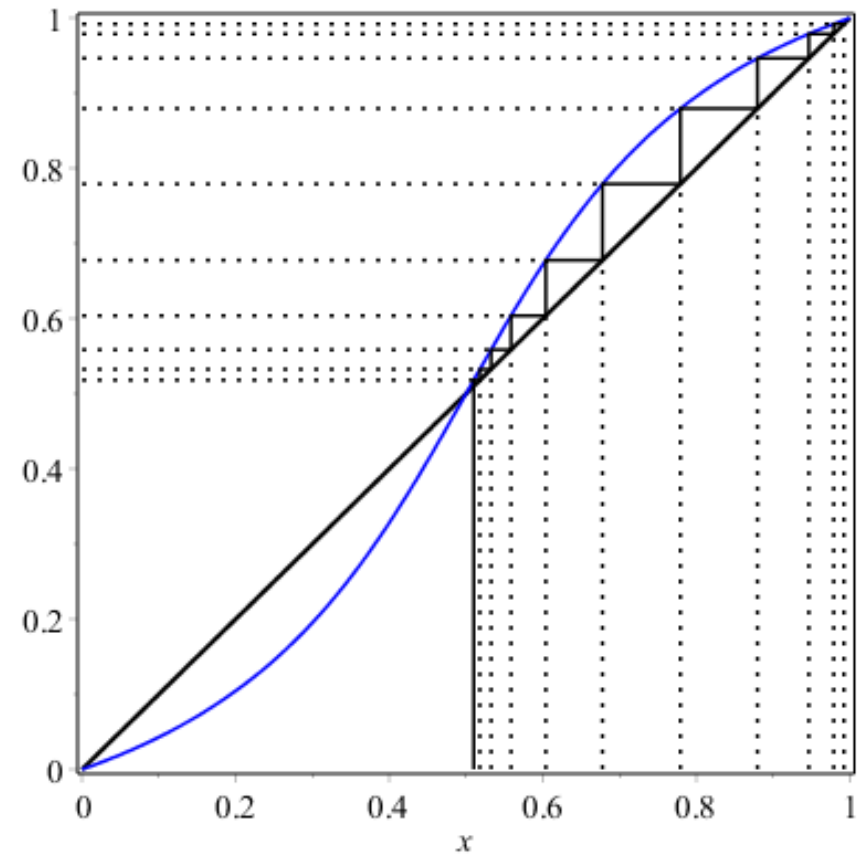
$$\Phi(x) = e^{-x}, x_0 = 0.75$$

Konvergenz



$$x_0 = 0.25$$

Keine Konvergenz



$$x_0 = 0.51$$

Konvergenz
(aber nicht gegen 0.5)

- I sei ein abgeschlossenes Intervall und $\Phi : I \rightarrow I$ sei eine **Kontraktion**, d.h. es gibt eine Konstante $L < 1$ so dass

$$|\Phi(x) - \Phi(y)| \leq L|x-y| \quad \text{für alle } x, y \in I$$
- Dann gilt: die Funktionswerte sind näher als ihre Ursprünge. Lipschitz
 - ▶ Φ besitzt genau einen Fixpunkt $x^* \in I$;
 - ▶ die Iterationsfolge $x_{i+1} = \Phi(x_i)$ konvergiert für jeden Startwert $x_0 \in I$;
 - ▶ $|x_n - x^*| \leq \frac{L^n}{1-L} \cdot |x_1 - x_0|$ (a priori Abschätzung);
je kleiner L, desto weniger Schwankung, desto besser
 - ▶ $|x_n - x^*| \leq \frac{1}{1-L} \cdot |x_{n+1} - x_n|$ (a posteriori Abschätzung).
- Der BFS gilt auch im \mathbb{R}^n

($I \subseteq \mathbb{R}^n$ abgeschlossene Teilmenge,
 $\|\Phi(x) - \Phi(y)\| \leq L\|x-y\|$)

das a priori für max iter und das a posteriori für early stopping

- Beispiele:

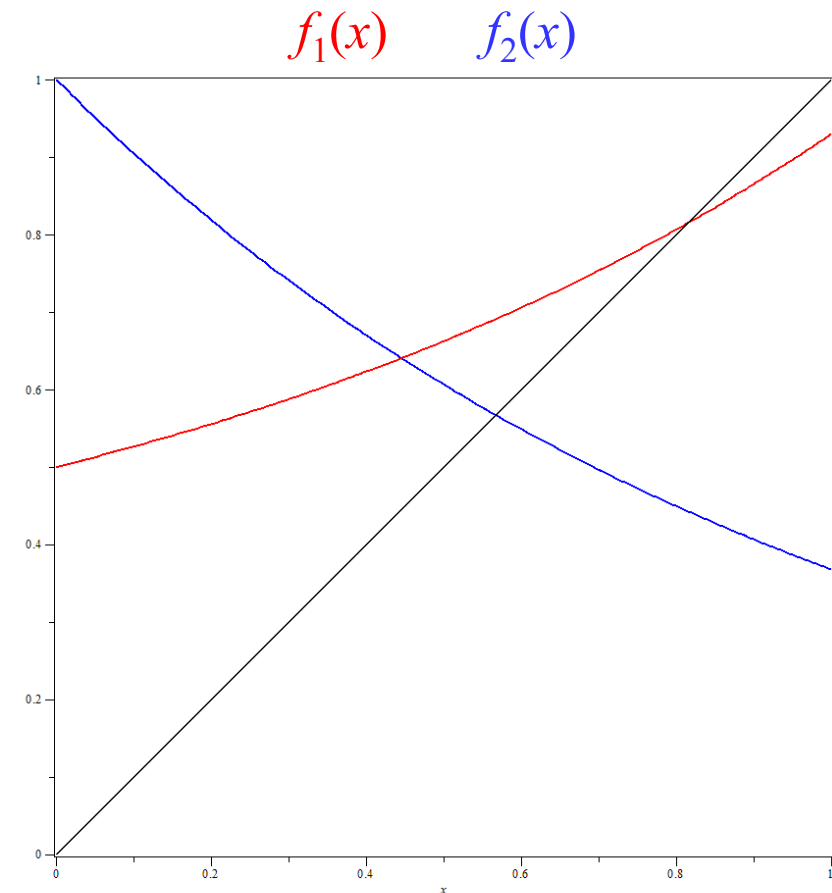
$$f_1 : [0,1] \rightarrow [0,1], \quad f_1(x) = \frac{1}{4} (1 + e^x)$$

$$f_2 : [0,1] \rightarrow [0,1], \quad f_2(x) = \exp(-x)$$

- Sind dies Kontraktionen?

Mittelwertsatz !

ableitung < 1



$$f_1 : [0,1] \rightarrow [0,1], \quad f_1(x) = \frac{1}{4} (1 + e^x)$$

$$L = 0.78$$

mittelwertsatz zwischen 0,1

| n | x_n | a priori | a posteriori |
|----|----------|----------|--------------|
| 0 | 0.5 | | |
| 1 | 0.662180 | 0.344633 | |
| 2 | 0.734754 | 0.234351 | 0.154219 |
| 3 | 0.771242 | 0.159358 | 0.077538 |
| 4 | 0.790613 | 0.108364 | 0.041162 |
| 5 | 0.801187 | 0.073687 | 0.022470 |
| 10 | 0.813773 | 0.010714 | 0.001238 |
| 15 | 0.814483 | 0.001558 | 0.000071 |
| 20 | 0.814524 | 0.000226 | 0.000004 |
| 25 | 0.814526 | 0.000048 | 0.000000 |

$$f_2(x) = \exp(-x)$$

$$!L = 1!$$

geht noch, aber nicht wegen fixp

| n | x_n | $f(x_n)$ |
|----|----------|----------|
| 0 | 0.500000 | 0.606531 |
| 1 | 0.606531 | 0.545239 |
| 2 | 0.545239 | 0.579703 |
| 3 | 0.579703 | 0.560065 |
| 4 | 0.560065 | 0.571172 |
| 5 | 0.571172 | 0.564863 |
| 10 | 0.566907 | 0.567277 |
| 15 | 0.567157 | 0.567135 |
| 20 | 0.567142 | 0.567144 |
| 25 | 0.567143 | 0.567143 |

- Iteratives Verfahren zur Bestimmung der Nullstelle der Funktion $f(x)$

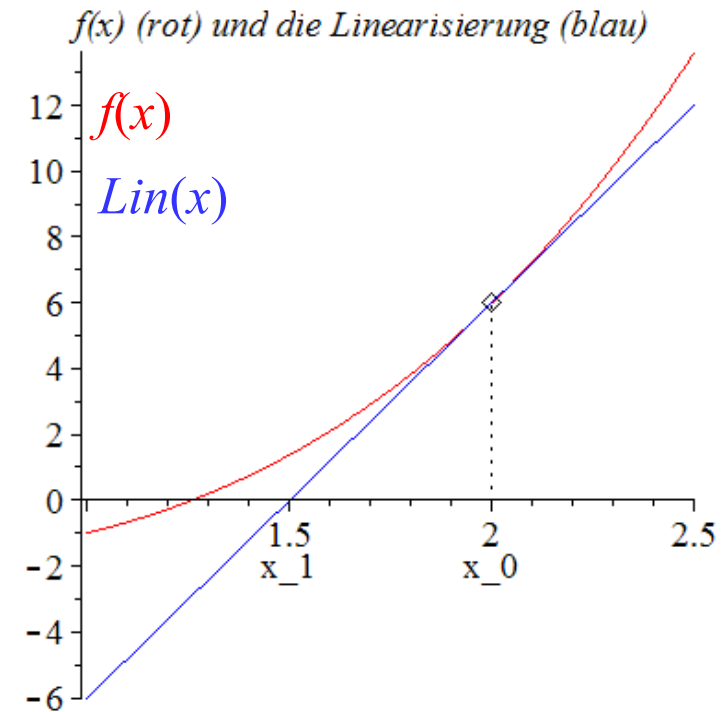
- ▶ Startwert x_0 :
- ▶ **Iterationsschritt:**
Linearisiere $f(x)$ in x_i
- ▶ d.h. bestimme die Tangente an $(x_i, f(x_i))$ und
- ▶ bestimme die Nullstelle der Linearisierung.
- ▶ **Linearisierung (z.B. mit Taylor):**

$$f(x) \approx f(x_i) + f'(x_i) \cdot (x - x_i) =: \text{Lin}(x)$$

- ▶ **Nullstelle der Linearisierung :**

möglichst gute Lipschitzkonstante

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$



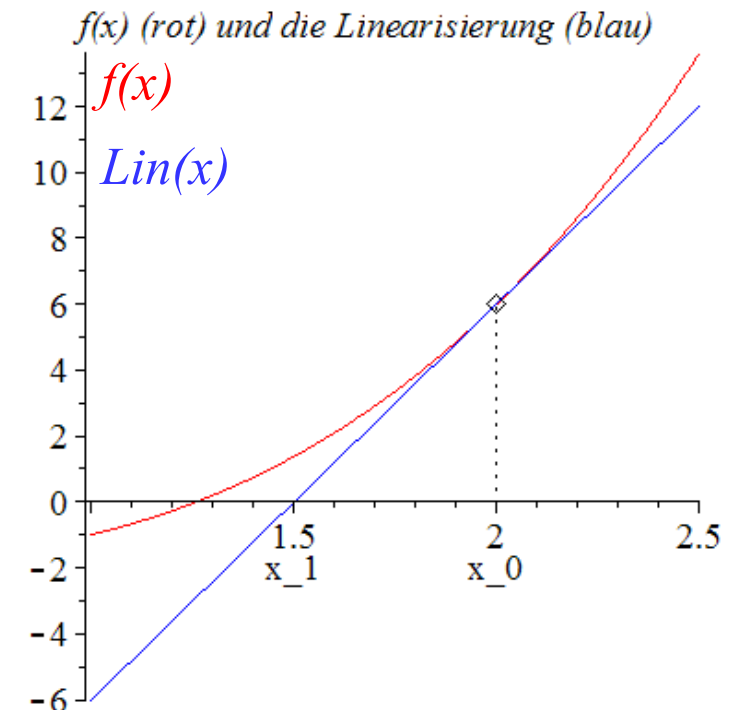
- Iteratives Verfahren zur Bestimmung der Nullstelle der Funktion $f(x)$
 - ▶ Startwert x_0 :

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

- Dies ist eine Fixpunkt-Iteration für

$$\Phi(x) = x - \frac{f(x)}{f'(x)}$$

- Das Verfahren konvergiert falls x_0 „nahe“ bei der Nullstelle liegt!



Heron-Verfahren oder Babylonisches Wurzelziehen

- Iteratives Verfahren zur Bestimmung von \sqrt{a} :

$$x_{i+1} = \frac{1}{2} \left(x_i + \frac{a}{x_i} \right)$$

- Geometrische Interpretation
- Falls $x_0^2 > a$ dann ist (x_n) monoton fallend,

$$0 \leq x_i - \sqrt{a} \leq \frac{1}{2\sqrt{a}} (x_{i-1} - \sqrt{a})^2 \text{ (quadratische Konvergenz)}$$

- Newton-Verfahren für $f(x) = x^2 - a$



Heron von Alexandria
1. Jahrhundert n. Chr

- Iterative Bestimmung von $\sqrt{2}$ nach Heron mit Startwert $x_0 = 2.0$
- bzw. Newton-Verfahren für $f(x) = x^2 - a$

| n | x_n |
|---|-------------|
| 1 | 1.5 |
| 2 | 1.416666667 |
| 3 | 1.414215686 |
| 4 | 1.414213562 |
| 5 | 1.414213562 |

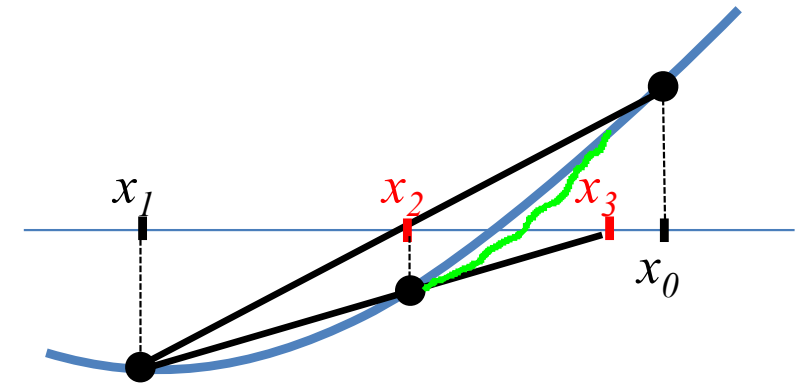
| n | x_n |
|---|----------------------|
| 1 | 1.5 |
| 2 | 1.4166666666666667 |
| 3 | 1.414215686274509804 |
| 4 | 1.414213562374689911 |
| 5 | 1.414213562373095049 |
| 6 | 1.414213562373095049 |

- Beobachtung: Die Anzahl der korrekten Stellen verdoppelt sich in jedem Iterationsschritt

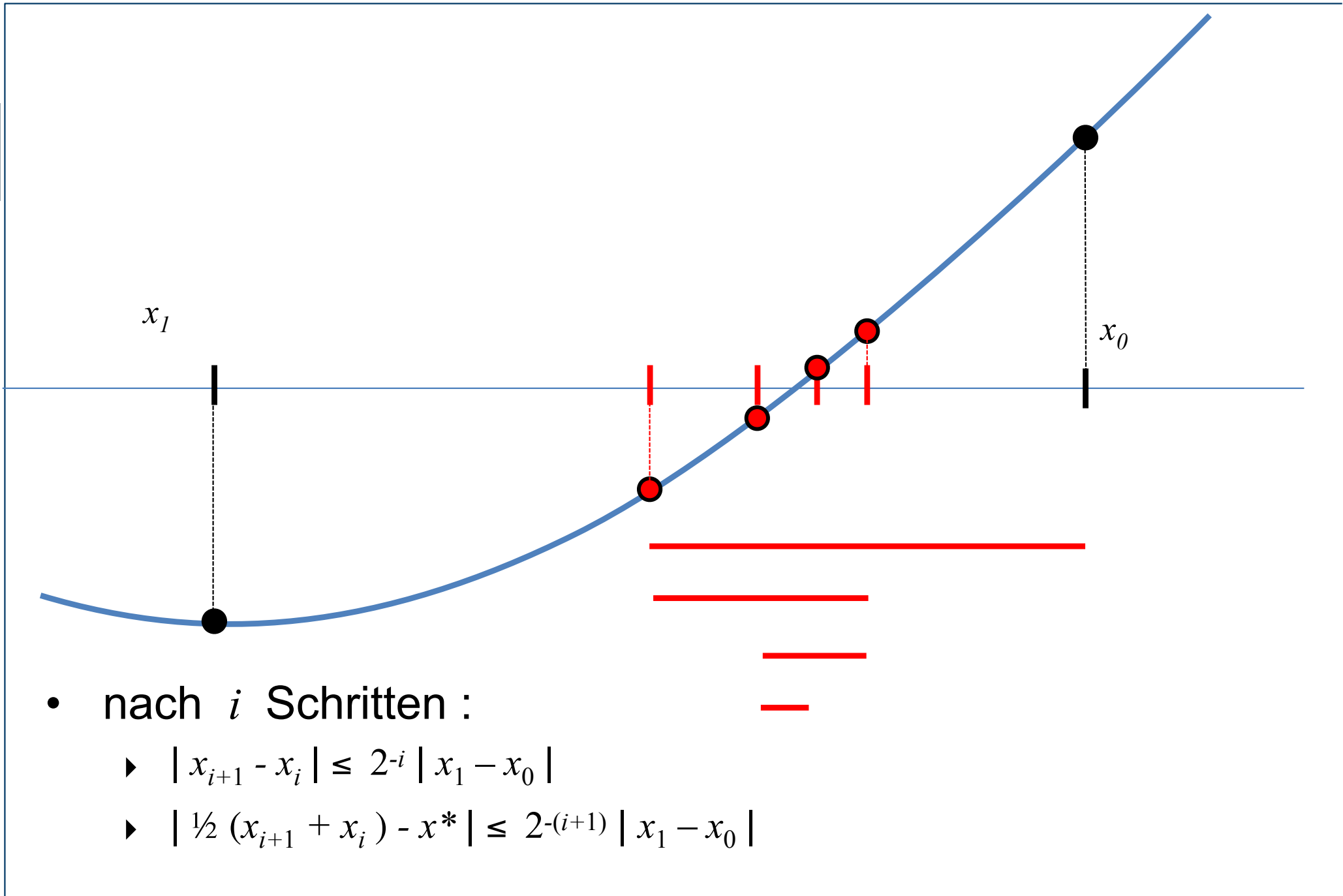
- Zweistufiges iteratives Verfahren ohne Kenntnis der Ableitung;
- Zwei Startwerte x_0 und x_1 nötig ;
- **Iterationschritt:**
Bestimme den Schnittpunkt der Sekante durch $(x_0, f(x_0))$ und $(x_1, f(x_1))$ mit der x-Achse (Nullstelle der Sekante)

$$x_{i+1} = \frac{x_{i-1} \cdot f(x_i) - x_i \cdot f(x_{i-1})}{f(x_i) - f(x_{i-1})}$$

- Modifikation: **Regula falsi** (s.u.)



- Zweistufiges Verfahren, führt sicher zum Ziel, aber konvergiert sehr langsam;
 1. Zwei Startwerte $x_0 < x_1$ so dass $f(x_0) \cdot f(x_1) < 0$
 Der Vorzeichenwechsel im Intervall $[x_0, x_1]$,
 dies garantiert dass es mind. eine Nullstelle gibt
 sofern $f(x)$ stetig ist.
 2. **Iterationsschritt:** binärsuche. konvergiert langsam
 Bestimme den Mittelpunkt $x_2 = (x_0 + x_1) / 2$,
 betrachte das Intervall $[x_0, x_2]$ falls $f(x_0) \cdot f(x_2) < 0$ bzw.
 betrachte das Intervall $[x_2, x_1]$ falls $f(x_1) \cdot f(x_2) < 0$
 (im Falle von $f(x_1) \cdot f(x_2) = 0$ ist x_2 eine Nullstelle!)
 3. Das Bisektionsverfahren konvergiert stets, jedoch relativ langsam:
 nach i Schritten $|x_{i+1} - x_i| \leq 2^{-i} |x_1 - x_0|$



- **Kombination von Sekanten- und Bisektionsverfahren**

- ▶ Man startet wie beim Bisektionsverfahren mit zwei Punkten

x_0 und x_1 so dass $f(x_0) \cdot f(x_1) < 0$

- ▶ Man bestimmt den **Schnittpunkt der Sekante mit der x-Achse**

- ▶ Falls $f(x_2) \cdot f(x_1) < 0$ fährt man mit x_1 und x_2 fort
andernfalls mit x_0 und x_2

spricht man wählt adaptiv die bereichsgröße aus

- ▶ regula falsi konvergiert garantiert gegen eine Nullstelle, meist langsam.

- Problem:

Bestimme die Nullstelle einer Funktion $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$

- Beispiel:

$$F(x, y) = \begin{bmatrix} 1 + xy^2 - x^2y \\ 2x^2 - 3y^2 \end{bmatrix} \quad \begin{array}{l} 1 + xy^2 - x^2y = 0 \\ 2x^2 - 3y^2 = 0 \end{array} \quad \begin{array}{l} \text{(nicht lineares)} \\ \text{Gleichungssystem} \end{array}$$

- Iteratives Vorgehen wie im eindimensionalen Fall:
Linearisieren und Nullstelle der Linearisierung ermitteln.
- Linearisierung mit Taylor:

$$F(x) \approx F(x_i) + J_F(x_i) \cdot (x - x_i) =: Lin_i(x)$$

dabei ist $J_F(x)$ die **Jacobi-Matrix** von F an der Stelle x .

- Bestimmung der Nullstelle der Linearisierung :

$$Lin_i(x) = 0 \quad \text{oder} \quad x_{i+1} = x_i - [J_F(x_i)]^{-1} F(x_i)$$

also genau das gleiche, nur mit Jacobi, statt ableitung.Pr

- **Iterationsschritt**

$$x_{i+1} = x_i - [J_F(x_i)]^{-1} F(x_i)$$

- zur praktischen Durchführung:
- Die Berechnung der Inversen sollte vermieden werden. Anstelle dessen:
- Im Allgemeinen muss man in jedem Iterationsschritt ein lineares Gleichungssystem mit $A = J_F(x_i)$ lösen:

- ▶ löse
- ▶ setze

$$[J_F(x_i)]z = -F(x_i)$$

$$x_{i+1} = x_i + z$$

das z entspricht also dem -f/f' Sprin

Maß für die Geschwindigkeit der Konvergenz

- Die (Iterations-)Folge (x_i) konvergiere gegen x^* :

- Konvergenzordnung $p=1$ **(lineare Konvergenz)**:

Es gibt eine **Konstante $C < 1$** so dass

desto kleiner C, desto besser!

$$|x_{i+1} - x^*| \leq C \cdot |x_i - x^*| \quad \text{für alle (großen) } i$$

- Konvergenzordnung $p > 1$:

es gibt eine Konstante so dass

$$|x_{i+1} - x^*| \leq C \cdot |x_i - x^*|^p \quad \text{für alle (großen) } i$$

- Im Fall $p=2$ spricht man von **quadratischer Konvergenz**;

- superlineare Konvergenz** :

sprich C geht gegen Null!

$$\lim_{i \rightarrow \infty} \frac{|x_{i+1} - x^*|}{|x_i - x^*|} = 0$$

selten größer, $p=3$ kann pass

- Fixpunktiterationen mit Voraussetzungen des BFS konvergieren (mindestens) linear
- Newtonverfahren:
 - ▶ einfache Nullstelle: quadratische Konvergenz
 - ▶ mehrfache Nullstelle: lineare Konvergenz man wird "in beide Richtungen" gezogen
 - ▶ Newtonverfahren im \mathbb{R}^n konvergiert quadratisch sofern einfache Nullstelle vorliegt (d.h. die Jacobimatrix in der Nullstelle invertierbar ist).
- Sekantenverfahren: Konvergenzordnung besser als linear (aber nur ein bisschen)

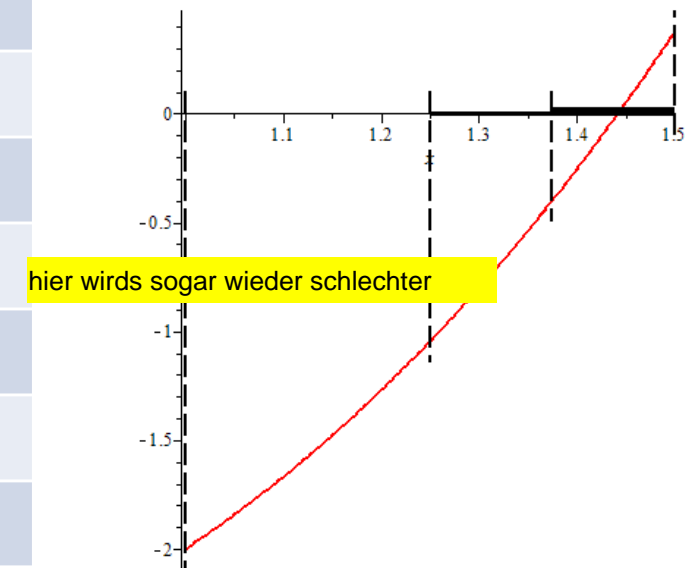
$$p = \frac{\sqrt{5}+1}{2} \approx 1.618\dots$$
- Bisektion und regula falsi „in etwa“ linear

- Konvergenzordnung: „in etwa linear“

$$f(x) = x^3 - 3, \quad x^* = \sqrt[3]{3} = 1.4422495703$$

Anfangsintervall [1.0,1.5]

| x_i | $ x_i - x^* $ | $ x_i - x^* / x_{i-1} - x^* $ |
|----------|---------------|-------------------------------|
| 1.250000 | 0.1922495 | |
| 1.375000 | 0.0672495 | 0.349803487 |
| 1.437500 | 0.0047495 | 0.070626028 |
| 1.468750 | 0.0265004 | 5.579542990 |
| 1.453125 | 0.0108754 | 0.410386925 |
| 1.445312 | 0.0030629 | 0.281637599 |
| 1.441406 | 0.0008433 | 0.275331137 |
| 1.443359 | 0.0011098 | 1.315995115 |

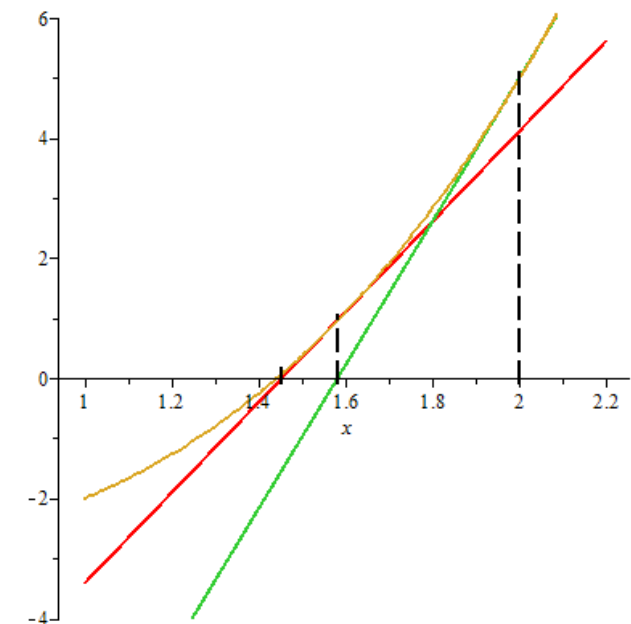


- Konvergenzordnung: quadratisch ($p=2$)

$$f(x) = x^3 - 3, \quad x^* = \sqrt[3]{3} = 1.4422495703$$

Startpunkt $x_0 = 1.5$

| x_i | $ x_i - x^* $ | $ x_i - x^* / x_{i-1} - x^* ^2$ |
|-------------|---------------|---------------------------------|
| 1.5 | 0.057750430 | |
| 1.444444444 | 0.002194874 | 0.6581109994 |
| 1.442252904 | 0.000003334 | 0.6920642374 |
| 1.442249570 | 0.0 | 0.0 |
| 1.442249570 | 0.0 | <i>undefined</i> |
| | | |
| | | |
| | | |



- Erhöhte Genauigkeit (30 Dezimalstellen)

$$f(x) = x^3 - 3, \quad x^* = \sqrt[3]{3} = 1.44224957030740838232163831078$$

| x_i | $ x_i - x^* $ | $ x_i - x^* / x_{i-1} - x^* ^2$ |
|------------------------------------|---------------------------------|---------------------------------|
| 1.5 | 0.05775042969259161767... | |
| 1. <u>44444444444444444444</u> ... | 0.00219487413703606212... | 0.658111047452949 |
| 1. <u>442252903791365329</u> ... | 0.00000333348395694750... | 0.691957031942603 |
| 1. <u>4422495703151130689</u> ... | 0.77046866751 10 ⁻¹¹ | 0.693359137594228 |
| 1. <u>4422495703074083823</u> ... | 0.4115945 10 ⁻²² | 0.693361301378699 |
| 1. <u>4422495703074083823</u> ... | 0.0 | 0.0 |
| 1.4422495703074083823... | 0.0 | <i>undefined</i> |
| | | |

- Konvergenzordnung: $p = \frac{1}{2}(1 + \sqrt{5}) = 1.618...$
erhöhte Genauigkeit

$$f(x) = x^3 - 3, \quad x^* = \sqrt[3]{3} = 1.44224957030740838232163831078$$

Anfangswerte $x_0 = 1.0, x_1 = 1.5$

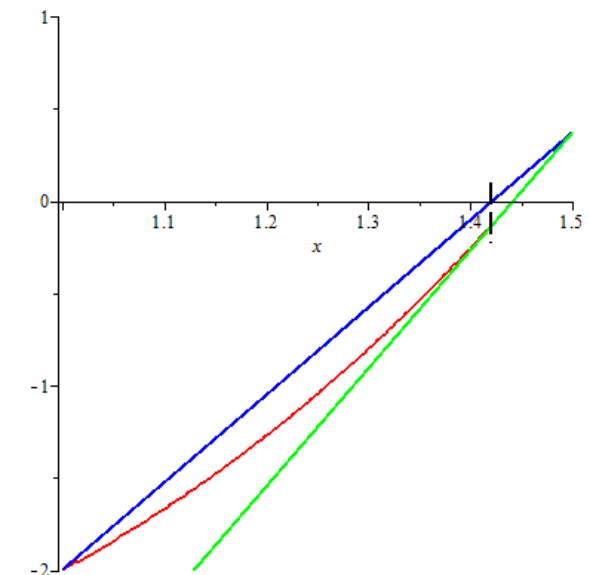
| x_i | $ x_i - x^* $ | $ x_i - x^* / x_{i-1} - x^* ^{1.618}$ |
|---------------------------------|----------------------------|---------------------------------------|
| 1.5 | 0.05775042969259161767... | |
| <u>1.4210526315789473684...</u> | 0.02119693872846101390... | 2.1385474547129 |
| <u>1.4414151249594287568...</u> | 0.00083444534797962542... | 0.4261326766406 |
| <u>1.4422619601527410463...</u> | 0.00001238984533266404... | 1.1867212636286 |
| <u>1.4422495631362650002...</u> | 0.7171143382064 10^{-11} | 0.6239856249190 |
| <u>1.4422495703073467779...</u> | 0.6160435103571 10^{-13} | 0.9279961284195 |
| <u>1.4422495703074083823...</u> | 0.30630873 10^{-21} | 0.7261346402066 |
| <u>1.4422495703074083823...</u> | 0.0 | 0.0 |

- Konvergenzordnung: linear

$$f(x) = x^3 - 3, \quad x^* = \sqrt[3]{3} = 1.4422495703$$

Anfangsintervall [1.0,1.5]

| x_i | $ x_i - x^* $ | $ x_i - x^* / x_{i-1} - x^* $ |
|--------------------|-------------------|-------------------------------|
| 1.5 | 0.057750430 | |
| <u>1.421052632</u> | 0.021196938 | 0.3670438 |
| <u>1.441415125</u> | 0.000834445 | 0.0393663 |
| <u>1.442217020</u> | 0.000032550 | 0.0390080 |
| <u>1.442248301</u> | 0.000001269 | 0.0389862 |
| <u>1.442249521</u> | $0.49 \cdot 10^7$ | 0.0386131 |
| <u>1.442249568</u> | $0.2 \cdot 10^8$ | 0.0408163 |
| <u>1.442249570</u> | 0.0 | 0.0 |



Vor- und Nachteile

- Newton
 - + konvergiert sehr schnell ($p=2$)
 - benötigt Werte der Ableitungen
 - konvergiert nur für Startwerte nahe bei der Nullstelle
- Sekanten
 - + konvergiert ziemlich schnell ($p=1.62$)
 - + benötigt keine Werte der Ableitungen
 - konvergiert nur für Startwerte nahe bei der Nullstelle
- Bisektion und regula falsi
 - konvergiert langsam ($p=1$)
 - + benötigt keine Werte der Ableitungen
 - + sicher konvergent

wegen BFS ist in einem intervall gültig (i.a.) oft mis



stetig und ns muss existieren (+/- über z.b. interv

Wann soll die Iteration abgebrochen werden?

1. Gibt es eine Fehlerabschätzung (z.B. BFS) nutze diese.
 2. Erreichen einer maximalen Iterationszahl **MAX_ITER**
 3. $\|x_{i+1} - x_i\| < \varepsilon$ für eine kleine Schranke ε . problem; monotonie!!
 4. $\|F(x_i)\| < \varepsilon$ (bei Nullstellensuche) bzw.
 $\|\Phi(x_i) - x_i\| < \varepsilon$ (bei Fixpunktiteration) also der echte abstand soll kleiner epsilon sein
 5. $\|x_i\| > M$ für eine (große) Schranke M (\rightarrow nicht konv.)
system divergiert
- In der Praxis: Eine Kombination mehrerer dieser Kriterien, z.B. 2. und 4. und 5.

Für spezielle Polynome kann man sukzessive sämtliche Nullstellen mit dem Newton-Verfahren bestimmen:

- **Annahme:** Das Polynom $p(x) = a_0 + a_1x + \dots + a_nx^n$ habe n reelle Nullstellen $\xi_1 \geq \xi_2 \geq \dots \geq \xi_{n-1} \geq \xi_n$ (zB char. Polynom einer symmetrischen Matrix)

in vielen Fällen besser polynomlösen umgehen, spez bei char.polynom (nur bei symmetrischen!!!!)

- die Theorie sagt dann $p(x) = a_n (x - \xi_1)(x - \xi_2) \dots (x - \xi_n)$
- Startet man das **Newton-Verfahren für $p(x)$ mit einem hinreichend großen x_0** (zB $x_0 \geq \frac{1}{|a_n|} \sum_{i=0}^n |a_i|$), dann **konvergiert die Iterations-Folge gegen ξ_1** .

- ξ_2 ist größte Nullstelle von $p_1(x) = \frac{p(x)}{x - \xi_1}$ jede gefundene NS rauskickern
- Diese kann iterativ mit Newton-Verfahren zum Startwert $x_0 = \xi_1 + \varepsilon$ bestimmt werden
 - ▶ Beachte, dass $p_1'(x) = \frac{p'(x)}{x - \xi_1} - \frac{p(x)}{(x - \xi_1)^2} \Rightarrow \frac{p_1(x)}{p_1'(x)} = \frac{p(x)}{p'(x) - \frac{p(x)}{x - \xi_1}}$
 deshalb müssen p_1 und p_1' nicht explizit bestimmt werden!
- ξ_{k+1} ist größte Nullstelle von $p_k(x) = \frac{p(x)}{(x - \xi_1)(x - \xi_2) \cdots (x - \xi_k)}$
 - ▶ Beachte, dass $\frac{p_k(x)}{p_k'(x)} = \frac{p(x)}{p'(x) - p(x) \cdot \sum_{i=1}^k \frac{1}{x - \xi_i}}$

- Fixpunkt-Iteration: Banachscher Fixpunktsatz
 - ▶ Konvergenz und Eindeutigkeit
 - ▶ Fehlerabschätzungen
- Verfahren zur Nullstellenbestimmung
 - ▶ Newton-V.
 - ▶ Bisektions-V.
 - ▶ Sekanten-V.
 - ▶ regula falsi
 - ▶ Vergleich der Verfahren
- Konvergenzordnung
 - ▶ Definition und Beispiele (Nullstellenverfahren)
- Abbruchkriterien
- Nullstellen von (speziellen) Polynomen

KLAUSURRELEVANT!!!