

## 4 Nichtlineare Gleichungen

Sei  $D \subset \mathbb{R}^N$  und  $F : D \rightarrow \mathbb{R}^N$  beliebig. Gesucht wird  $U \in \mathbb{R}^N$  mit

$$F(U) = 0$$

Speziell:  $F(U) = AU - b$ ,  $A \in \mathbb{R}^{N,N}$ ,  $b \in \mathbb{R}^N$  lineares Problem

### 4.1 Fixpunkte (Ergänzung 5)

#### 4.1.1 Fixpunkte und Nullstellen

$U$  Fixpunkt von  $G$ :  $U = G(U)$

$U$  Nullstelle von  $F$ :  $F(U) = 0$

$U$  Fixpunkt von  $G \Leftrightarrow U$  Nullstelle von  $F(X) := X - G(X)$

#### 4.1.2 Banachscher Fixpunktsatz

Sei  $V$  ein Banach-Raum,  $D \subseteq V$  abgeschlossen,  $f : D \rightarrow D$  eine Kontraktion, d.h.  $\exists q \in (0, 1)$  mit

$$\|f(x) - f(y)\| \leq q\|x - y\| \quad (x, y \in D)$$

Dann gilt:

- (i)  $f$  besitzt genau einen Fixpunkt  $x_*$  in  $D$
- (ii) Zu jedem  $x_0 \in D$  konvergiert die durch  $x_{i+1} := f(x_i)$  definierte Folge gegen  $x_*$  und es gelten die Abschätzungen

$$\|x_i - x_*\| \leq q^i \|x_0 - x_*\| \quad (\text{A priori Abschätzung})$$

$$\|x_i - x_*\| \leq \frac{q}{1-q} \|x_i - x_{i-1}\| \quad (\text{A posteriori Abschätzung})$$

#### 4.1.3 Beispiele

- 1.)  $f : [a, b] \subseteq \mathbb{R} \rightarrow [a, b]$  differenzierbar mit  $|f'(x)| \leq q < 1 \forall x \in [a, b]$  für ein  $q \in (0, 1)$   
 $\Rightarrow \exists! x_* \in [a, b] : f(x_*) = x_*$  und die Fixpunktiteration  $x_{i+1} := f(x_i)$  konvergiert für die Startwerte  $x_0 \in [a, b]$ .

- 2.) Suche Lösung von  $x = \cos(x)$ :

$$x_0 \in \mathbb{R}, \quad x_{i+1} = \cos(x_i)$$

Bildchen

Wende (1) an

$$\max_{x \in \mathbb{R}} |\cos'(x)| = \max_{x \in \mathbb{R}} |\sin(x)| = 1$$

So geht es noch nicht.

Aber:  $V = [0, 1]$ . Dann

$$\max_{x \in [0, 1]} |\sin(x)| = \sin(1) < 1$$

$\cos(V) \subset V$ . Anwendung von (1) ist OK.

$x_0 \in \mathbb{R} \Rightarrow x_1 = \cos(x_0) \in [-1, 1] \Rightarrow x_2 = \cos(x_1) \in [0, 1]$

Jetzt weiter wie eben. Konvergenz für alle  $x_0 \in \mathbb{R}$

**Satz 16.**  $V$  Banach-Raum,  $D \subset V$  abgeschlossen,  $f : D \rightarrow D$  eine Kontraktion mit Rate  $q$  der Fixpunktiteration und Fixpunkt  $v_x$ .  $g : D \rightarrow D$  sei eine Störung von  $f$  mit

$$\|f(v) - g(v)\|_V \leq \varepsilon \quad \forall v \in D$$

Definiere  $\{v_i\}_i, \{w_i\}$  durch  $v_{i+1} := f(v_i)$ ,  $w_{i+1} := g(w_i)$  für  $v_0, w_0 \in D$  und  $\|v_0 - w_0\|_V \leq \varepsilon$ . Dann gilt:

$$\begin{aligned} \|v_i - w_i\|_V &\leq \frac{\varepsilon}{1-q} \\ \|v_* - w_i\|_V &\leq \frac{1}{1-q} (\varepsilon(1+3q^i) + q^i \|w_0 - g(w_0)\|_V) \end{aligned}$$

*Bildchen*

**Beweis.**  $v_0 \in D \Rightarrow v_1 \in D \Rightarrow \dots$

$w_0 \in D \Rightarrow w_1 \in D \Rightarrow \dots$

Folgen sind wohldefiniert

$$\begin{aligned} \|v_{i+1} - w_{i+1}\|_V &= \|f(v_i) - g(w_i)\|_V \\ &\leq \|f(v_i) - f(w_i)\|_V + \|f(w_i) - g(w_i)\|_V \\ &\leq q \cdot \|v_i - w_i\|_V + \varepsilon \\ &\leq q^2 \cdot \|v_{i-1} - w_{i-1}\|_V + (1+q)\varepsilon \\ &\leq \dots \leq q^{i+1} \underbrace{\|v_0 - w_0\|_V}_{\leq \varepsilon} + \sum_{j=0}^i q^j \varepsilon \\ &\leq \sum_{j=0}^{i+1} q^j \varepsilon \leq \sum_{j=0}^{\infty} q^j \varepsilon = \frac{1}{1-q} \varepsilon. \end{aligned}$$

Mit dem Fixpunktsatz von Banach:

$$\begin{aligned} \|v_* - w_i\|_V &\leq \|v_* - v_i\|_V + \|v_i - w_i\|_V \\ &= \frac{q^i}{1-q} \|v_0 - f(v_0)\|_V + \frac{\varepsilon}{1-q} \\ &\leq \frac{q^i}{1-q} (\underbrace{\|v_0 - w_0\|_V}_{\leq \varepsilon} + \|w_0 - g(w_0)\|_V + \underbrace{\|g(w_0) - f(v_0)\|_V}_{\substack{= \|w_1 - v_1\|_V \\ \leq (1+q)\varepsilon \leq 2\varepsilon}}) + \frac{\varepsilon}{1-q} \end{aligned}$$

□

Problem: Wie schnell sind Fixpunktverfahren?

#### 4.1.4 Konvergenzordnung

$V$  Banach-Raum,  $\{v_i\}_i$  eine iterative erzeugte Folge mit  $\lim_{i \rightarrow \infty} v_i = v_*$ . Die Iteration hat *Konvergenzordnung*  $p \geq 1$ , falls für den Fehler  $e_i := v_i - v_*$  gilt:

$$\lim_{i \rightarrow \infty} \frac{\|e_i\|_V}{\|e_{i-1}\|_V^p} = c \in \mathbb{R}$$

Falls  $c \neq 0$ , so heißt  $p$  die *genaue Konvergenzordnung* und  $c$  heißt *asymptotischer Fehlerkoeffizient*.

#### Beispiele

$p = 1$ : Geometrische oder lineare Konvergenz

$p = 2$ : Quadratische Konvergenz.

**Satz 17.**  $I \subseteq \mathbb{R}$ ,  $\Phi : I \rightarrow \mathbb{R}$  habe einen Fixpunkt  $x_* \in I$  und sei  $p$ -mal stetig db. mit

$$\Phi'(x_*) = \dots = \Phi^{(p-1)}(x_*) = 0 \text{ falls } p > 1$$

oder

$$|\Phi'(x_*)| < 1 \text{ falls } p = 1 \text{ ist}$$

Dann konvergiert das Iterationsverfahren

$$x_{i+1} = \Phi(x_i)$$

für die Startwerte  $x_0$  nahe  $x_*$  und hat bzgl.  $|\cdot|$  die Konvergenzordnung  $p$ .

Ist  $\Phi^{(p)}(x_*) \neq 0$ , so ist  $p$  die genaue Konvergenzordnung.

**Beweis.** Nach Voraussetzung gibt es für alle  $p \geq 1$  eine Umgebung von  $x_*$ , in der  $|\Phi'| < 1$  gilt. Nach 1.3(1) konvergiert die Fixpunktiteration für alle Startwerte dieser Umgebung gegen  $x_*$ .

Mit Taylorentwicklung:

$$x_{i+1} = \Phi(x_i) = \sum_{l=0}^{p-1} \frac{1}{l!} \Phi^{(l)}(x_*)(x_i - x_*)^l + \frac{1}{p!} \Phi^{(p)}(\xi_i)(x_i - x_*)^p$$

( $\xi_i$  zwischen  $x_*$  und  $x_i$ ).

Einsetzen der Voraussetzung:

$$x_{i+1} = x_* + \frac{1}{p!} \Phi^{(p)}(\xi_i)(x_i - x_*)^p$$

und somit

$$\lim_{i \rightarrow \infty} \frac{|x_{i+1} - x_*|}{|x_i - x_*|^p} = \lim_{i \rightarrow \infty} \frac{1}{p!} |\Phi^{(p)}(\xi_i)| = \frac{1}{p!} |\Phi^{(p)}(x_*)|$$

□

**Bemerkung:** Lineare vs. Quadratische Konvergenz.

$$e_0 = 10^{-1}$$

Lineare Konvergenz:  $q = 1/2$ ,  $e_k = \left(\frac{1}{\alpha}\right)^k e_0 \approx 10^{-0.3k} e_0$

1 Stelle  $\rightsquigarrow$  3 Iterationen

8 Stellen  $\rightsquigarrow$  24 Iterationen

Quadratische Konvergenz:  $c = 1$

$$e_0 = \frac{1}{10}, e_1 = e_0^2 = 10^{-2}, e_2 = 10^{-4}, e_3 = 10^{-8}$$

## 4.2 Berechnung von Nullstellen

### 4.2.1 Extrema (Ergänzung 7)

$x_*$  Extremum von  $f$  und  $f$  db  $\Rightarrow f'(x_*) = 0$

$\rightsquigarrow$  Nullstellenproblem

### 4.2.2 Nullstellen reeller Funktionen

Im Folgenden sei  $I = [a, b] \subset \mathbb{R}$ ,  $a < b$ ,  $f$  mindestens stetig.

**Bisektionsverfahren** Es gelte  $f(a)f(b) < 0$  („=0“  $\Rightarrow f(a) = 0$  oder  $f(b) = 0$ ).

Wir konstruieren Intervalle  $\{I_k\}_k$  wie folgt:

Start:

$$a_0 := a, b_0 := b, I_0 := [a_0, b_0]$$

Iteration:  $L \geq 0$

$$1.) \quad \bar{x} := \frac{1}{2}(a_k + b_k)$$

$$2.) \quad \text{Stop: } f(\bar{x}) = 0$$

$$3.) \quad f(a_k) \cdot f(\bar{x}) \stackrel{?}{<} 0 : a_{k+1} = a_k, b_{k+1} = \bar{x} \\ \text{sonst: } a_{k+1} = \bar{x}, b_{k+1} = b_k$$

$$4.) \quad k \mapsto k + 1, I_{k+1} = [a_{k+1}, b_{k+1}]$$

Abbruch:  $\text{Tol}_X, \text{Tol}_f \geq 0$  gegeben,  $\text{Tol}_x + \text{Tol}_f > 0$

$k_{\max} \in \mathbb{N}$ . Rückgabe  $x$  und  $f(x)$  mit

$x$  Approximation der Nullstelle mit  $|x - x_*| \leq \text{Tol}_x$  oder  $|f(x)| \leq \text{Tol}_f$   $f(x)$ : Funktionswert in  $x$

**Modifikation der Iteration:**

$$|f(\bar{x})| \leq \text{Tol}_f :$$

return( $\bar{x}, f(\bar{x})$ );

$$|b_k - a_k| \leq \text{Tol}_x :$$

falls  $|f(a_k)| < |f(b_k)|$  return ( $a_k, f(a_k)$ ), sonst return( $b_k, f(b_k)$ )

**Satz 18.**  $f : [a, b] \rightarrow \mathbb{R}$  stetig mit  $f(a) \cdot f(b) < 0$ .  $\text{Tol}_x, \text{Tol}_f, k_{\max}$  wie oben gegeben.  
Dann bricht das Bisektionsverfahren nach endlich vielen Schritten ab, auch falls  $k_{\max} = \infty$

**Beweis.** Das Verfahren ist wohldefiniert aufgrund des Zwischenwertsatzes.  
Die Existenz einer Nullstelle in  $I_k$  ist für jedes  $k$  gesichert.

$$\text{Tol}_x > 0 : |I_k| = \left(\frac{1}{2}\right)^k |b - a| \stackrel{!}{\leq} \text{Tol}_x \Rightarrow k \leq \left\lceil \frac{\log_2(b-a)}{\text{Tol}_x} \right\rceil$$

$$\text{Tol}_f > 0 : b_k - a_k \rightarrow 0$$

Da  $f$  stetig ist und eine Nullstelle in  $[a_k, b_k]$  hat, gilt  $\lim_{k \rightarrow \infty} f(a_k) = \lim_{k \rightarrow \infty} f(b_k) = 0$

$$\Rightarrow \exists k_f \in \mathbb{N} : \min\{|f(a_{k_f})|, |f(b_{k_f})|\} \leq \text{Tol}_f$$

(Gilt  $|f'(x)| \leq C \forall x \in [a, b]$ , so gilt z.B.:  $|f(a_k)| = |f(a_k) - f(x_k)| \leq |I_k| \max_{x \in [a, b]} |f'(x)| \leq$

$$C \left(\frac{1}{2}\right)^k \stackrel{!}{\leq} \text{Tol}_f)$$

□

## Probleme

- $a, b$  zu finden mit  $f(a) \cdot f(b) < 0$  kann sehr schwierig sein.
- Die Konvergenz ist in der Praxis zu langsam.  
(Siehe 1.5: Konvergenzordnung ist 1 mit  $c = \frac{1}{2}$ )
- Die Methode ist auf  $\mathbb{R}$  beschränkt

**Regula Falsi** Wie in 2.2.1 aber mit  $\bar{x}$  wie folgt:  
Bildchen

$$\bar{x} = a_k - \frac{f(a_k)(b_k - a_k)}{f(b_k) - f(a_k)}$$

Keine Auslöschung im Nenner wegen  $f(a_k) \cdot f(b_k) < 0$ . Weiteres Vorgehen wie in 2.2.1  
Konvergenz: Konvergiert wie in 2.2.1 im Fall  $\text{Tol}_f > 0$ . Die Konvergenz kann beliebig langsam sein. Im „besten“ Fall ist die Konvergenz linear (unter noch allgemeinen Voraussetzungen)

## Das Sekantenverfahren Bildchen

$f : \mathbb{R} \rightarrow \mathbb{R}$  stetig.

$x_1, x_2$  gegeben,  $x_1 \neq x_2$  und  $f(x_1) \neq f(x_2)$

$x_3$  ist dann die Nullstelle der Sekante

**Initialisierung:**  $x_1 \neq x_2, f(x_1) \neq f(x_2)$

**Iteration für  $k \geq 0$ :**

1.) Falls  $f(x_{k-1}) \neq f(x_k)$

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}$$

2.)  $k \leadsto k + 1$

**Abbruch:**  $\text{Tol}_x, \text{Tol}_f, \text{Tol}_{f'}, k_{\max}$

Wie in 2.2.1 aber mit

$$\begin{aligned} |x_k - x_{k-1}| &\leq \text{Tol}_x? \\ |f(x_k)| &\leq \text{Tol}_f? \\ k &\leq k_{\max} \\ \text{und } |f(x_k) - f(x_{k-1})| &\leq \text{Tol}_{f'}? \end{aligned}$$

Die letzten beiden Bedingungen führen zu einem erfolglosen Abbruch.

### Bemerkungen

- Keine Erfolgsgarantie für allgemeine Startwerte
- Kleine  $f$ -Differenzen erzeugen große Fehler

Aber:

- Günstiger Aufwand (1  $f$ -Auswertung pro Schritt) bei schneller Konvergenz, falls es konvergiert.
- Gewisse Verallgemeinerung auf  $\mathbb{R}^N$  möglich

**Satz 19.**  $f \in C^2(\mathbb{R})$ ,  $f(x_*) = 0$ ,  $f'(x_*) \neq 0$ ,  $f''(x_*) \neq 0$ .

Dann ex. eine Umgebung  $U$  von  $x_*$ , sodass das Sekantenverfahren für alle Startwerte aus  $U$  konvergiert und die Konvergenzordnung ist genau  $\frac{1}{2}(1 + \sqrt{5}) \approx 1.6$

**Newton-Verfahren**  $f : \mathbb{R} \rightarrow \mathbb{R}$  stetig db.

Idee: Verwende Tangente statt Sekante

Bildchen

**Initialisierung:**  $x_1$  mit  $f'(x_1) \neq 0$

**Iteration:** für  $k \geq 0$

1.) Falls  $f'(x_k) \neq 0$

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

2.)  $k \leadsto k+1$

**Abbruch:**  $\text{Tol}_x, \text{Tol}_f, k_{\max}, \text{Tol}_{f'}$

$$\begin{aligned} |x_k - x_{k-1}| &\leq \text{Tol}_x \\ |f(x_k)| &\leq \text{Tol}_f \\ k &\leq k_{\max} \\ |f'(x_k)| &\leq \text{Tol}_{f'} \end{aligned}$$

In den letzten beiden Fällen ist der Abbruch erfolglos

### Bemerkungen

- Keine Garantie eines erfolgreichen Abbruchs (im Allgemeinen)
- Kleine Werte von  $f'$  führen zu großen Fehlern

Aber:

- sehr schnell, falls konvergent
- Verallgemeinerung auf  $\mathbb{R}^N$  bzw. Banachräume möglich

### Konvergenzordnung des Newton-Verfahrens

$f \in C^3$ ,  $f(x_*) = 0$ ,  $f'(x_*) \neq 0$

Die Iterationsfunktion des Newton-Verfahrens ist

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

Nach 1.5 bilden wir  $\Phi'(x_*)$ ,  $\Phi''(x_*)$

$$\begin{aligned} \Phi'(x) &= 1 - \left(1 - \frac{f(x)f''(x)}{f'(x)^2}\right) = \frac{f(x)f''(x)}{f'(x)^2} \stackrel{x=x_*}{=} 0 \\ \Phi''(x) &= \frac{f''(x)}{f'(x)} + f(x)(\dots) \stackrel{x=x_*}{=} \frac{f''(x_*)}{f'(x_*)} + 0 \end{aligned}$$

Die Konvergenz ist quadratisch und sie ist genau quadratisch, falls  $f''(x_*) \neq 0$

### 4.2.3 Lokale Konvergenz des Newtonverfahrens

Es sei  $(V, \|\cdot\|_V)$  ein Banachraum,  $\emptyset \neq U \subset V$ ,  $f : U \rightarrow V$  eine stetig db. Funktion mit  $f'(v)^{-1} \in \mathbb{L}(V, V)$  für alle  $v \in U$  sowie

$$\sup_{v \in U} \|f'(v)^{-1}\|_{\mathbb{L}(V, V)} \leq K < \infty$$

und

$$\|f'(v) - f'(w)\|_{\mathbb{L}(V, V)} \rightarrow 0 \quad (\|v - w\|_V \rightarrow 0) \text{ glm. für } v, w \in U.$$

Weiter sei  $u_* \in U$  eine Nullstelle von  $f$ . Dann gibt es zu jedem  $q \in (0, 1)$  ein  $\delta > 0$ , so dass für jeden Startwert  $u_0 \in B_\delta(u_*)$  die Newton-Iteration  $u_{i+1} = u_i - f'(u_i)^{-1}f(u_i)$  wohldefiniert ist und für  $i \geq 0$  gilt

$$\|u_i - u_*\|_V \leq q \|u_0 - u_*\|_V$$

Ist  $f$  zweimal stetig db, so ist die Konvergenz quadratisch:

$$\|u_{i+1} - u_*\|_V \leq C \|u_i - u_*\|_V^2$$

für  $i \geq 0$  und ein  $C > 0$ .  $C$  hängt von  $f$  ab.

Insbesondere bricht das Verfahren nach endlich vielen Schritten bzgl. der Kriterien

$$\begin{aligned} \|u_i - u_{i-1}\|_V &\stackrel{!}{\leq} \text{Tol}_x \quad \text{oder} \\ \|f(u_i)\|_V &\leq \text{Tol}_f \end{aligned}$$

für  $\text{Tol}_x, \text{Tol}_f \geq 0$ ,  $\text{Tol}_x + \text{Tol}_f > 0$  ab

**Bemerkung**  $T : V \rightarrow V$  linear, stetig ( $:\Leftrightarrow T \in \mathbb{L}(V, V)$ ),

$$\|T\|_{\mathbb{L}(V, V)} := \sup_{v \in V} \frac{\|Tv\|_V}{\|v\|_V}$$

**Beweis.** Sei  $r_0 > 0$  mit  $\overline{B_{r_0}(u_*)} \subset U$ . Dann gilt für  $u \in B_r(u_*)$  ( $0 < r < r_0$ )

$$f(u) = f(u_*) + \int_0^1 f'(u_* + t(u - u_*))(u - u_*) dt$$

Die Iterationsfunktion des Newton-Verfahrens ist

$$G(u) := u - f'(u)^{-1} \cdot f(u)$$

$G$  ist auf  $B_{r_0}(u_*)$  wohldefiniert und mit  $u(t) := u_* + t(u - u_*)$  gilt

$$\begin{aligned} G(u) - u_* &= u - u_* - f'(u)^{-1} \int_0^1 f'(u(t))(u - u_*) dt \\ &= \int_0^1 f'(u)^{-1} (f'(u) - f'(u(t)))(u - u_*) dt \end{aligned}$$



Daher:

$$\|G(u) - u_*\|_V \leq \sup_{v \in B_r(u_*)} \|f'(v)^{-1}\|_{\mathbb{L}(V,V)} \cdot \sup_{t \in (0,1)} \|f'(u) - f'(u(t))\|_{\mathbb{L}(V,V)} \cdot \|u - u_*\|_V$$

Zu  $q \in (0,1)$  wähle also  $\delta$ , so dass

$$\|G(u) - u_*\|_V \leq q \cdot \|u - u_*\|_V \text{ für alle } u \in B_\delta(u_*)$$

Für  $u_0 \in B_\delta(u_*)$  folgt also induktiv

$$\|u_{i+1} - u_*\|_V = \|G(u_i) - u_*\|_V \leq q \cdot \|u_i - u_*\|_V \leq \delta$$

d.h.  $\{u_i\}_i \in B_\delta(u_*)$  und  $\lim_{i \rightarrow \infty} u_i = u_*$ . Insbesondere

$$\|u_i - u_*\|_V \leq q^i \|u_0 - u_*\|_V$$

Ist  $f$  zweimal stetig db, so gilt:

$$\begin{aligned} \sup_{t \in (0,1)} \|f'(u) - f'(u(t))\|_{\mathbb{L}(V,V)} &\leq C' \|u - u(t)\|_V \\ &\leq C' \|u - u_*\|_V \quad \text{mit } C' = C'(f'') \end{aligned}$$

Also

$$\|G(u) - u_*\|_V \leq KC' \|u - u_*\|_V^2 = C \|u - u_*\|_V^2$$

$$\Rightarrow \|u_{i+1} - u_*\|_V \leq C \|u_i - u_*\|_V^2$$

Mit  $\|u_{i+1} - u_i\|_V \leq \|u_{i+1} - u_*\|_V + \|u_i - u_*\|_V \leq 2 \cdot \|u_i - u_*\|_V$ .

Also  $\|u_{i+1} - u_i\|_V \rightarrow 0$  und mit Stetigkeit  $\|f(u_i)\|_V \rightarrow 0$  für  $i \rightarrow \infty$ . Daraus folgt der Abbruch nach endlich vielen Schritten.  $\square$

### Bemerkungen:

- $f'$  invertierbar heißt, dass  $u_*$  eine einfache Nullstelle ist
- $u$  Nullstelle von  $f$ . Dann sei  $\varepsilon(u)$  der Einzugsbereich von  $u$ , d.h.  $u_0 \in \varepsilon(u) \Rightarrow$  das Newton-Verfahren ist wohldefiniert für  $u_0$  und die Folge  $\{u_i\}_{i \geq 0}$  konvergiert gegen  $u$ .  
Der vorherige Satz sagt:  $B_\delta(u) \subseteq \varepsilon(u)$  für  $\delta$  klein (unter genannten Voraussetzungen)

**Beispiel**  $V = \mathbb{R}$ ,  $f(x) = \arctan(x)$

Bildchen

$$\begin{aligned} f(0) &= 0 \\ |x_0| < X_0 &\Rightarrow x_i \rightarrow 0 \\ |x_0| > X_0 &\Rightarrow |x_i| \rightarrow \infty \\ x_0 = X_0 &\Rightarrow x_i = (-1)^i \cdot x_0 \end{aligned}$$

Für  $V = \mathbb{R}^n$ ,  $n \geq 2$  ist  $\varepsilon(u)$  sehr kompliziert.

Wir berechnen für große Raumdimension  $n$   $f(u_i)^{-1}$  nicht explizit. Stattdessen lösen wir

$$\begin{aligned} f'(u_i)d_i &= -f(u_i) \\ u_{i+1} &= u_i + d_i \end{aligned}$$

**Newton-Kantorovich-Theorem**  $F : D \subset V \rightarrow V$ ,  $V$  Banachraum,  $D$  offen und konvex,  $F$  stetig db,  $x_0 \in D$  und  $F'(x_0)$  invertierbar sowie

$$\begin{aligned} \|F'(x_0)^{-1}F(x_0)\| &\leq \alpha \\ \|F'(x_0)^{-1}(F'(y) - F'(x))\|_{\mathbb{L}(V,V)} &\leq \omega_0 \cdot \|x - y\|_V \quad \forall x, y \in D \\ h_0 := \alpha\omega_0 &< 1/2 \\ B_\delta(x_0) &\subset D, \quad \delta := \frac{1}{\omega_0}(1 - (1 - 2h_0)^{1/2}) \end{aligned}$$

Dann ist die Folge  $\{x_k\}_k$  der Newton-Iteration wohldefiniert, sie bleibt in  $B_\delta(x_0)$  und konvergiert gegen ein  $x_*$  mit  $F(x_*) = 0$ . Die Konvergenz ist quadratisch.

### Bemerkung

- Die Existenz der Nullstelle wird garantiert. Daher sind solche Theoreme auch in der Analysis interessant.
- Man kann (wie bei Banach) a priori Schranken oder a posteriori Schranken betrachten
- Beachte:  $F(u) = 0 \Leftrightarrow AF(u) = 0$ , falls  $A$  invertierbar ist. Wie in 2.4.1, 2.4.2 hängen die Konstanten von  $A$  ab. Die Größe  $F'^{-1}F$  ist invariant gegenüber der Transformation  $F \mapsto AF$

### 4.2.4 Globale Konvergenz

Idee: Definiere eine „Energie“, die in jedem Schritt verkleinert wird:  
für ein  $E : V = \mathbb{R}^n \rightarrow \mathbb{R}$  gelte

$$|u_{i+1}| = |u_i - f'(u_i)^{-1}f(u_i)| = E(u_{i+1}) < E(u_i)$$

Problem:  $u_{i+1}$  sollte nicht zu weit weg sein von  $u_i$ . Ausweg (siehe Jakobi- oder SOR-Verfahren): Dämpfung.

Für  $\tau_i > 0$  ist  $u_{i+1} = u_i - \tau_i f'(u_i)^{-1}f(u_i)$  das gedämpfte Newton-Verfahren.

„ $i$  klein“:  $\tau_i \in (0, 1)$  klein

„ $i$  groß“:  $\tau_i \rightarrow 1$  um von der quadratischen Konvergenz zu profitieren. ( $\tau \neq 1$ : gedämpftes Newton-Verfahren konvergiert nur linear)

**Lemma 4.**  $\emptyset \neq D \subset \mathbb{R}^n$  abgeschlossen und beschränkt.  $f \in C^1(D, \mathbb{R}^n)$  und  $f'(u)^{-1}$  existiere für alle  $u \in D$ .  $|\cdot|$  eine Vektornorm.

Definiere  $E : D \rightarrow \mathbb{R}$ ,  $u \mapsto E(u) = |f(u)|$  mit  $d(u) := -f'(u)^{-1} \cdot f(u)$ . Dann gilt:  
Für alle  $\varepsilon > 0$  existiert ein  $\delta > 0$  mit

$$E(u + \tau d(u)) \leq (1 - \tau + \varepsilon \tau) E(u) \quad \text{für alle } u \in D, \tau \in (0, \delta)$$

**Beweis.** Für  $u \in D$ :

$$\begin{aligned} f(u + \tau d(u)) &= f(u) + \int_0^\tau f'(u + sd(u)) d(u) \, ds \\ &= \left( Id - \int_0^\tau f'(u + sd(u)) f'(u)^{-1} \, ds \right) f(u) \\ &= \left( (1 - \tau) Id - \int_0^\tau (f'(u + sd(u)) - f'(u)) f'(u)^{-1} \, ds \right) f(u) \end{aligned}$$

$\tau$  genügend klein:

$$|f(u + \tau d(u))| \leq (1 - \tau + \underbrace{\tau \sup_{s \in (0, \tau)} \|f'(u + sd(u)) - f'(u)\|_2}_{\leq C^{-1} \cdot \varepsilon, \text{ falls } \tau \leq \delta} \cdot \underbrace{\|f'(u)^{-1}\|_2}_{\leq C}) \cdot |f(u)|$$

$$\Rightarrow E(u + \tau d(u)) \leq (1 - \tau + \varepsilon \tau) E(u) \quad \square$$

**Schrittweitensteuerung**  $f$  wie in 2.3,  $E$  wie oben. Wähle ein  $\sigma \in (0, 1)$  und  $u_0 \in D$ .  
Newton-Verfahren mit Schrittweitensteuerung

**Initialisierung:**  $u_0 \in D$

**Iteration:** für  $k \geq 0$

- 1.) Löse  $f'(u_k) d_k = -f(u_k)$  für  $d_k$
- 2.) Bestimme  $\tau_k = 2^{-q_k}$  und  $q_k \in \mathbb{N}$  minimal mit  $B_{\tau_k |d_k|}(u_k) \subset D$  und  $E(u_k + \tau_k d_k) \leq (1 - \sigma \tau_k) E(u_k)$
- 3.)  $u_{k+1} = u_k + \tau_k d_k$ , gehe zu (1)

Wahl des Wertes  $q_k$

$k = 0$ :  $q = 0, 1, \dots$  bist die Bedingung in (2) für ein  $q_0$  zum ersten Mal erfüllt ist.

$k > 0$ : Probiere  $q = q_{k-1} - 1, q_{k-1}, \dots$  bist (2) für ein  $q_k$  zum ersten Mal erfüllt ist.

## Globale Konvergenz

**Satz 20.** *f wie im Lemma in 2.4.1 bzgl. eines  $D_\alpha$ .*

*Zu  $\alpha > 0$  sei  $D_\alpha := \{v \in D : |f(v)| \leq \alpha\}$  nichtleer und kompakt. (f darf nur eine Nullstelle haben und muss glm konvergieren)*

*Dann konvergiert das Verfahren aus 2.4.1 für alle Startwerte  $u_0 \in D_\alpha$  gegen eine Nullstelle von f in  $D_\alpha$ .*

*Insbesondere folgt der Abbruch nach endlich vielen Schritten bzgl. des Kriteriums  $E(u_k) \leq \text{Tol}_f$  für ein  $\text{Tol}_f > 0$*

**Beweis.** *Nach Konstruktion gilt:*

$$E(u_{[k+1]}) \leq E(u_k) \leq \dots \leq E(u_0) = \alpha$$

*und  $\{u_k\}_k \subseteq D_\alpha$ .*

*Die Folge konvergiert daher, weil  $D_\alpha$  kompakt ist, etwa  $u_k \rightarrow u_*(k \rightarrow \infty)$  für eine Teilfolge. Nach dem Lemma gibt es zu jedem  $\varepsilon > 0$  ein  $\delta > 0$ , so dass*

$$|f(u_k + \tau d(u_k))| \leq (1 - (1 - \varepsilon)\tau)|f(u_k)|$$

*für  $0 \leq \tau \leq \delta$ , gleichmäßig in  $D_\alpha$ .*

*Nun sei  $\varepsilon := 1 - \sigma$ , d.h.*

$$|f(u_k + \tau d(u_k))| \leq (1 - \sigma\tau)|f(u_k)|$$

*Diese Ungleichung gilt für  $\tau = \delta$ , d.h. nach Konstruktion gilt  $\tau_k \geq \delta/2$ . Insbesondere erhalten wir nach endl. vielen Schritten*

$$|f(u_{k+1})| = |f(u_k + \tau_k d_k)| \leq (1 - \frac{1}{2}\delta\sigma)|f(u_k)|,$$

*also  $E(u_{k+1}) \leq \kappa E(u_k)$  für ein  $\kappa \in (0, 1)$ , so dass  $\lim_{k \rightarrow \infty} E(u_k) = 0$ . Insbesondere wird*

*$E(u_k) = |f(u_k)| \stackrel{!}{\leq} \text{Tol}_f$  nach endlich vielen Schritten erreicht.*

□