

# Krylov- und Tschebyscheff-Approximationen an die Matrixexponentialfunktion

Jan Leibold  
05.10.2015

Institut für Angewandte und Numerische Mathematik

- **Einleitung**
  - Ausgangsproblem: Die Schrödinger-Gleichung
  - Allgemeines Problem
  
- **Verfahren**
  - Krylov-Verfahren
  - Tschebyscheff-Verfahren
  
- **Implementierung und numerische Tests**
  - Implementierung
  - Numerische Tests
  
- **Fazit**

# Abschnitt 1

## *Einleitung*

**Ziel:** Lösen der homogenen Schrödingergleichung auf  $\Omega = (0, 1)^d$  ( $d = 1, 2$ ) mit homogenen Dirichlet-Randbedingungen:

$$\begin{aligned}\partial_t U(x, t) &= i\Delta U(x, t), & \forall t > 0, x \in \Omega, \\ U(x, t) &= 0, & \forall t \geq 0, x \in \partial\Omega, \\ U(x, 0) &= U_0(x), & \forall x \in \Omega.\end{aligned}$$

Die Funktion  $U_0$  gibt hierbei den Anfangswert vor.

**Diskretisierung** : Zerlege  $\Omega$  in ein äquidistantes Gitter. Dabei sei

- $N + 1$  die Anzahl der Gitterpunkte in jeder Raumdimension
- $h = 1/N$  die Ortsschrittweite
- $n$  die Anzahl der Unbekannten (der inneren Gitterpunkte)

Bestimme Approximation

$$u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{pmatrix} \approx \begin{pmatrix} U(x_1, t) \\ \vdots \\ U(x_n, t) \end{pmatrix}$$

Verwenden der Methode der finiten Differenzen liefert die ODE

$$\frac{d}{dt}u(t) = i\Delta_h u(t), \quad \forall t > 0, \quad (1a)$$

$$u(0) = \begin{pmatrix} U_0(x_1) \\ \vdots \\ U_0(x_n) \end{pmatrix}. \quad (1b)$$

Hierbei ist  $\Delta_h$  der diskrete Laplace Operator:

$$\Delta_h = \Delta_{h,1} := \frac{1}{h^2} \begin{pmatrix} -2 & 1 & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & -2 \end{pmatrix}, \quad \text{für } d = 1,$$

bzw.

$$\Delta_h = \Delta_{h,2} := (I_{N-1} \otimes \Delta_{h,1} + \Delta_{h,1} \otimes I_{N-1}), \quad \text{für } d = 2.$$

Sei  $A := i\Delta_h$ . Dann ist die Lösung von (1) gegeben durch

$$u(\tau) = e^{\tau A} u(0).$$

## Lemma 1

*Die Matrix  $A$  hat folgende Eigenschaften:*

- a)  *$A$  ist schiefhermitesch,*
- b)  *$\mathcal{F}(A) \subseteq iN^2[-4, 0]$  für  $d = 1$ , bzw.  $\mathcal{F}(A) \subseteq iN^2[-8, 0]$  für  $d = 2$ .*

# Allgemeines Problem

Approximiert werden sollen Funktionen der Form  $u(\tau) := e^{\tau A} b$  mit

- $A \in \mathbb{C}^{n \times n}$  schiefhermitesch,
- $b \in \mathbb{C}^n$ ,  $\zeta = \|b\|$ ,
- $\tau > 0$ .

(Diese Notation gelte nun für den gesamten weiteren Vortrag.)

Zu diesem Zweck werden nun zwei Verfahren betrachtet:

- Krylov-Verfahren
- Tschebyscheff-Verfahren

**Die Matrixfunktion wird hier als bekannt voraus gesetzt.**

Im Folgenden sei stets  $\Phi = \exp$ .

## Abschnitt 2

### *Verfahren*



Sei

- $K_m(A, b) := \text{span}\{b, Ab, \dots, A^{m-1}b\}$  der  $m$ -te Krylovraum bezüglich  $A$  und  $b$ ,
- $v_1, \dots, v_m$  die Orthonormalbasis aus dem Arnoldi-Algorithmus,
- $V_m = (v_1 \dots v_m) \in \mathbb{C}^{n \times m}$  und
- $H_m = V_m^* A V_m$ .

Die Näherung des Krylov-Verfahrens ist dann gegeben durch

$$\Phi(A)b \approx \zeta V_m \Phi(H_m) e_1.$$

## Definition 2

Für  $x \in [-1, 1]$ ,  $k \in \mathbb{N}$  ist das  $k$ -te Tschebyscheff-Polynom definiert als

$$T_k(x) := \cos(k \arccos(x)).$$

## Satz 3

*Die Tschebyscheff-Polynome sind orthogonal bzgl. des Skalarproduktes*

$$\langle f, g \rangle := \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} f(x)g(x) dx.$$

*Konkret gilt*

$$\langle T_i, T_j \rangle = \begin{cases} 0, & i \neq j, \\ \pi, & i = j = 0, \\ \frac{\pi}{2}, & i = j \neq 0. \end{cases}$$

Da  $A$  schiefhermitesch ist, existieren  $\alpha, \beta \in i\mathbb{R}$  mit  $\mathcal{F}(A) \subset [\alpha, \beta] \subset i\mathbb{R}$ .

Hier sei nun  $\Phi = \exp|_{[\alpha, \beta]} := [\alpha, \beta] \rightarrow \mathbb{C}$ . Definiere

- $\psi_{\alpha, \beta}: [\alpha, \beta] \rightarrow [-1, 1]$ ,  $z \mapsto 2 \frac{z - \alpha}{\beta - \alpha} - 1$  und
- $\tilde{\Phi}: [-1, 1] \rightarrow \mathbb{C}$ ,  $\tilde{\Phi} = \Phi \circ \psi_{\alpha, \beta}^{-1}$ .

Mit

$$\tilde{A} := \psi_{\alpha, \beta}(A) := \frac{2}{\beta - \alpha} \left( A - \frac{\alpha + \beta}{2} I \right)$$

gilt  $\mathcal{F}(\tilde{A}) \subset [-1, 1]$ .

Die Approximation des Tschebyscheff-Verfahrens ist dann gegeben durch

$$\Phi(A)b = \tilde{\Phi}(\psi_{\alpha, \beta}(A))b \approx \sum_{k=0}^m c_k T_k(\psi_{\alpha, \beta}(A))b = \sum_{k=0}^m c_k T_k(\tilde{A})b.$$

mit

$$c_k = \begin{cases} \frac{1}{\pi} \langle \tilde{\Phi}, T_k \rangle, & n = 0, \\ \frac{2}{\pi} \langle \tilde{\Phi}, T_k \rangle, & n > 0, \end{cases} \quad \tilde{\Phi} = \Phi \circ \psi_{\alpha, \beta}^{-1}.$$

## Abschnitt 3

### *Implementierung und numerische Tests*

Für die Implementierung sei  $A = i\Delta_h$  des konkreten Problems aus der Einleitung. Zur Berechnung der Orthonormalbasis des Krylovraumes und somit der Matrizen  $V_m$  und  $H_m$  wird verwendet:

---

## Algorithmus 1 Arnoldi-Algorithmus mit modifiziertem Gram-Schmidt Prozess

---

- 1: Berechne  $\zeta = \|b\| > 0$
- 2:  $v_1 = \frac{1}{\zeta} b$
- 3: **for**  $m = 1, 2, \dots$  **do**
- 4:      $\tilde{v}_{m+1} = Av_m$
- 5:     **for**  $j = 1, \dots, m$  **do**
- 6:          $h_{j,m} = v_j^* \tilde{v}_{m+1}$
- 7:          $\tilde{v}_{m+1} = \tilde{v}_{m+1} - h_{j,m} v_j$
- 8:     **end for**
- 9:      $h_{m+1,m} = \|\tilde{v}_{m+1}\|$
- 10:      $v_{m+1} = \tilde{v}_{m+1} / h_{m+1,m}$
- 11: **end for**

---

## Algorithmus 2 Gesamtalgorithmus (vereinfacht)

---

- 1: Berechne  $V_m$  und  $H_m$  mit Alg. 1
  - 2: Berechne  $U \in \mathbb{C}^{m \times m}$  unitär mit  $U^* H_m U = \text{diag}(\lambda_1, \dots, \lambda_n)$
  - 3: Berechne Näherung  $y_m := \zeta V_m U \text{diag}(\Phi(\lambda_1), \dots, \Phi(\lambda_n)) U^* e_1$
- 

### Bemerkung

- $A$  schiefhermitesch  $\Rightarrow H_m = V_m^* A V_m$  schiefhermitesch und tridiagonal  $\Rightarrow H_m$  lässt sich unitär diagonalisieren (Berechnung der Diagonalisierung mit MATLAB-Routine **eig**) und Arnoldi vereinfacht sich zu 3-Term-Rekursion.
- Indem man  $A$  durch  $\hat{A} = \frac{1}{i} A$  (reell, symmetrisch) ersetzt, lässt sich der Arnoldi-Algorithmus und die Diagonalisierung in reeller Arithmetik durchführen und man kann daraus die gleiche Krylov-Approximation ausrechnen (vgl. BA Kap. 6.1).
- In BA werden noch weitere Aspekte berücksichtigt.

## Berechnung der Tschebyscheff-Koeffizienten

### Erinnerung

$$c_k = \begin{cases} \frac{1}{\pi} \langle \tilde{\Phi}, T_k \rangle, & k = 0, \\ \frac{2}{\pi} \langle \tilde{\Phi}, T_k \rangle, & k > 0, \end{cases}$$

mit

$$\langle \tilde{\Phi}, T_k \rangle = \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} \Phi \left( \frac{\alpha+\beta}{2} + \frac{\beta-\alpha}{2} x \right) T_k(x) dx.$$

Durch Substitution  $x = \cos(\theta)$  ergibt sich

$$\langle \tilde{\Phi}, T_k \rangle = \frac{1}{2} \int_{-\pi}^{\pi} \Phi \left( \frac{\alpha+\beta}{2} + \frac{\beta-\alpha}{2} \cos(\theta) \right) \cos(k\theta) d\theta.$$

## Berechnung der Tschebyscheff-Koeffizienten

$$\langle \tilde{\Phi}, T_k \rangle = \frac{1}{2} \int_{-\pi}^{\pi} \Phi \left( \frac{\alpha + \beta}{2} + \frac{\beta - \alpha}{2} \cos(\theta) \right) \cos(k\theta) d\theta \quad (2)$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} \tilde{\Phi}(\cos(\theta)) \cos(k\theta) d\theta. \quad (3)$$

Zur Berechnung zwei Möglichkeiten:

- Quadratur von (2) (z.B. mit MATLAB-Routine `integral`)
- $\langle \tilde{\Phi}, T_k \rangle$  ist wegen (3) bis auf Vorfaktor  $k$ -ter Fourierkoeffizient von  $\tilde{\Phi} \circ \cos \Rightarrow c_k$  lässt sich über DFT mit dem FFT-Algorithmus approximieren. Der FFT-Algorithmus wird ebenfalls von MATLAB zur Verfügung gestellt.



Nach Berechnung der Koeffizienten ist noch

$\sum_{k=0}^m c_k T_k(\psi_{\alpha,\beta}(A))b = \sum_{k=0}^m c_k T_k(\tilde{A})b \approx \Phi(A)b$  auszuwerten. Eine Möglichkeit:

---

## Algorithmus 3 Auswertung der Tschebyscheff-Approximation

---

- 1: Berechne  $\tilde{A} = \frac{2}{\beta-\alpha} \left( A - \frac{\alpha+\beta}{2} I \right)$
  - 2: Setze  $d_0 = b$
  - 3: Setze Näherung  $v = c_0 d_0$
  - 4:  $d_1 = \tilde{A}b$
  - 5: Update  $v = v + c_1 d_1$
  - 6: **for**  $k = 2$  **to**  $m$  **do**
  - 7:      $d_k = 2\tilde{A}d_{k-1} - d_{k-2}$
  - 8:     Update  $v = v + c_k d_k$
  - 9: **end for**
- 

Hierbei wird die 3-Term-Rekursion der Tschebyscheff-Polynome ausgenutzt.

## Erinnerung

Ortdiskretisierte Schrödingergleichung:

$$\frac{d}{dt}u(t) = i\Delta_h u(t) =: Au(t), \quad \forall t > 0,$$

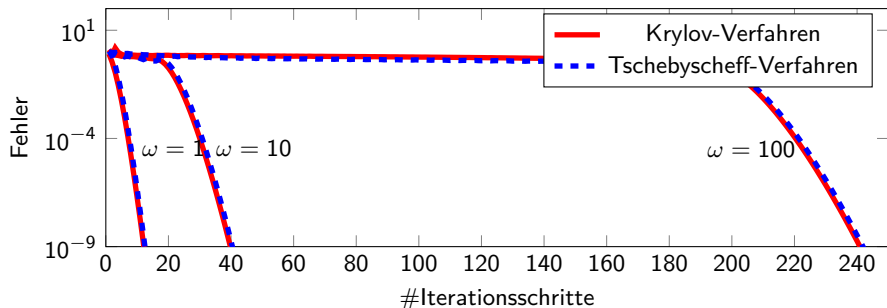
$$u(0) = \begin{pmatrix} U_0(x_1) \\ \vdots \\ U_0(x_n) \end{pmatrix} =: b$$

$$\Rightarrow u(\tau) = \Phi(\tau A)b \text{ mit } \mathcal{F}(A) \subset i[-4dN^2, 0]$$

$$\text{Definiere } \omega = \tau N^2 \Rightarrow \mathcal{F}(\tau A) \subset i[-4d\omega, 0]$$

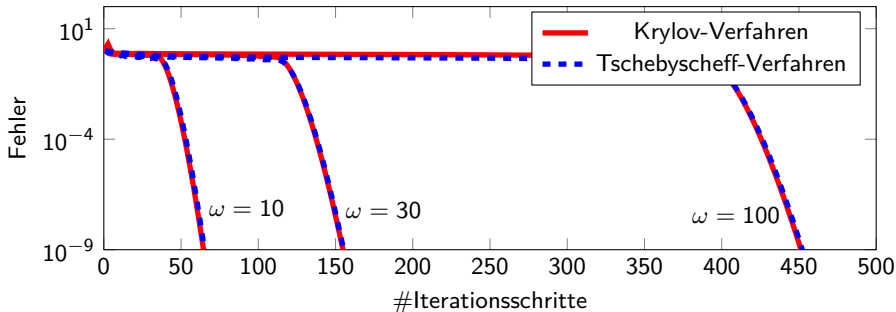
**An diesem Problem werden nun die beiden Verfahren miteinander verglichen.**

## Konvergenz der Verfahren



**Abbildung:** Fehler des Krylov- und des Tschebyscheff-Verfahrens für  $N = 100000$ ,  $d = 1$  und verschiedene  $\omega$ .

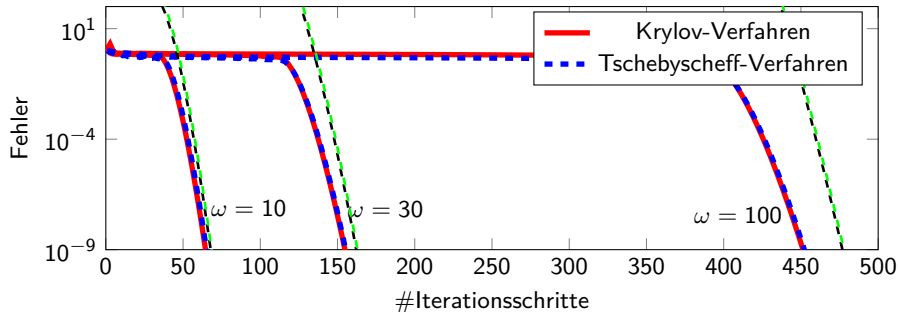
## Konvergenz der Verfahren



**Abbildung:** Fehler des Krylov- und des Tschebyscheff-Verfahrens für  $N = 501$ ,  $d = 2$  und verschiedene  $\omega$ .

- In der BA werden Fehlerabschätzungen für beide Verfahren bewiesen.
- Diese hängen neben  $m$  ab von  $|\beta - \alpha|$ .
- Abschätzungen mit Beweisen in BA, hier nur Illustration.

## Illustration der Fehlerschranken



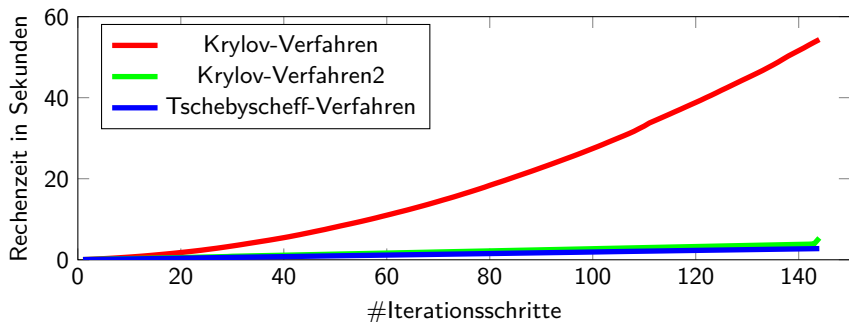
**Abbildung:** Fehler des Krylov- und des Tschebyscheff-Verfahrens für  $N = 501$ ,  $d = 2$  und verschiedene  $\omega$  mit Fehlerschranke für das Krylov-Verfahren (grün) und Tschebyscheff-Verfahren (schwarz).

## Rechenzeit

Bei der Rechenzeit werden zwei Möglichkeiten der Implementierung des Krylov-Verfahrens betrachtet:

- Nach jeden Arnoldi Schritt wird eine neue Näherung ausgerechnet (Krylov-Verfahren).
- Es wird nur einmal am Ende eine Näherung ausgerechnet (Krylov-Verfahren 2).

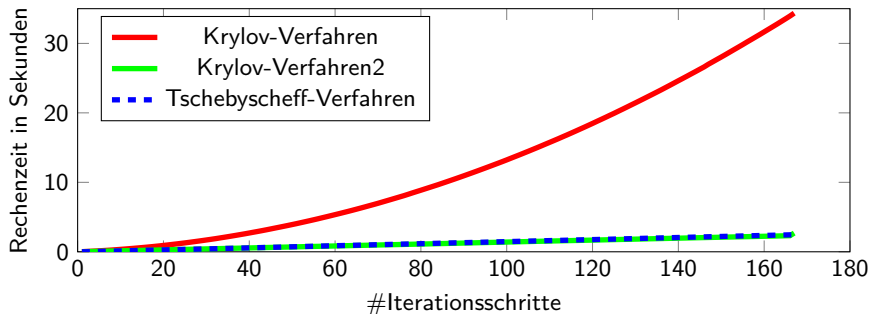
## Rechenzeit



**Abbildung:** Rechenzeit der Verfahren im Verhältnis zur Iterationsanzahl bei  $d = 1$ ,  $N = 500000$  und  $\omega = 50$ .



## Rechenzeit

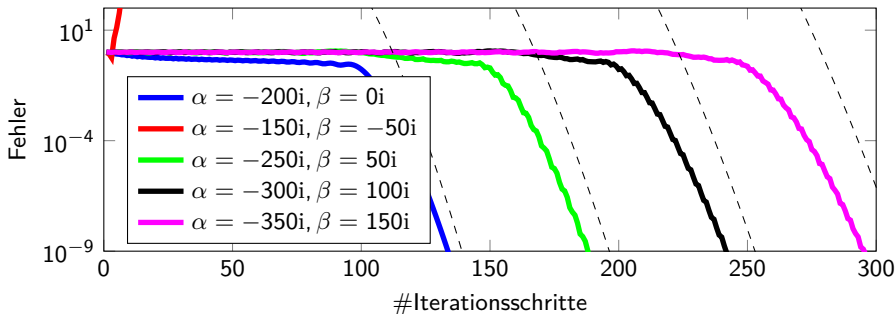


**Abbildung:** Rechenzeit der Verfahren im Verhältnis zur Iterationsanzahl bei  $d = 2$ ,  $N = 500$  ( $n = 499^2$ ) und  $\omega = 30$ .

## Abhängigkeit der Konvergenz des Tschebyscheff-Verfahrens von der Wahl von $\alpha$ und $\beta$

Bei dem Tschebyscheff-Verfahren müssen  $\alpha$  und  $\beta$  gewählt werden wählen, mit  $\mathcal{F} \subset [\alpha, \beta]$ .

**Frage:** Wie hängt die Konvergenz von dieser Wahl ab?



**Abbildung:** Fehler des Tschebyscheff-Verfahrens für  $N = 100000$ ,  $d = 1$ ,  $\omega = 50$  und verschiedene  $\alpha$  und  $\beta$ .

**Beobachtung :** Die Konvergenz hängt maßgeblich von einer guten Abschätzung des Wertebereiches ab.

## Abhängigkeit der Konvergenz von der Glattheit des Anfangswertes

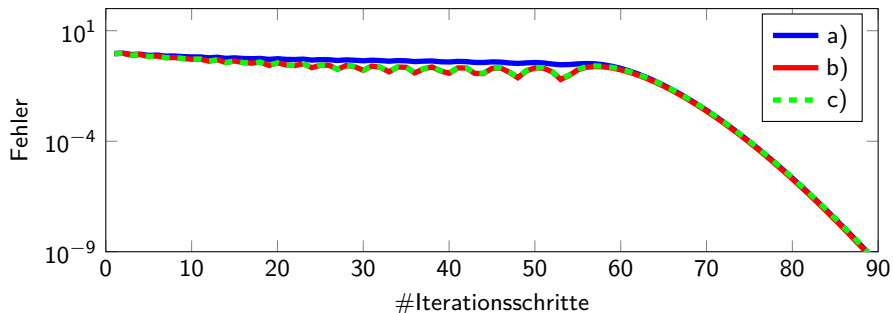
Bisher war  $b$  immer ein zufälliger Vektor.

Frage: Wirkt sich Glattheit des Anfangswertes auf Konvergenz aus?

Testen Verfahren in  $1 - d$  für  $N = 500001$  und  $\omega = 30$  mit verschiedenen Anfangswerten:

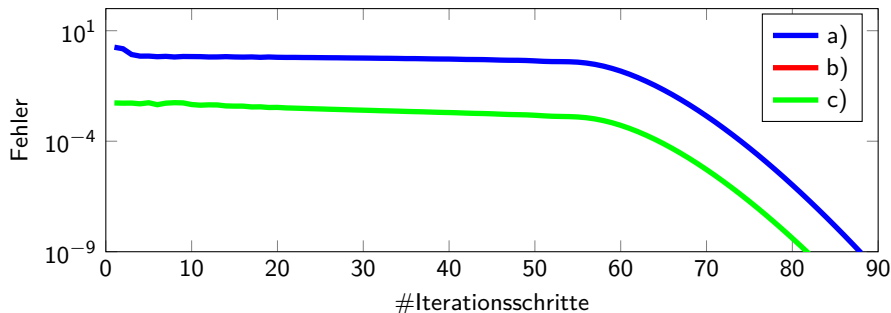
- a) Einem zufälliger Vektor der Länge 500000.
- b) Der Diskretisierung von  $\sin(2\pi x)$  (unendlich oft stetig differenzierbar).
- c) Der Diskretisierung von  $\begin{cases} x, & x \in [0, 1/2] \\ 1 - x, & x \in (1/2, 1] \end{cases}$  (stückweise stetig differenzierbar),

## Abhängigkeit der Konvergenz von der Glattheit des Anfangswertes



**Abbildung:** Fehler des Tschebyscheff-Verfahrens für  $N = 500001$ ,  $\omega = 30$  und die Anfangswerte aus Fall a), b) und c).

## Abhängigkeit der Konvergenz von der Glattheit des Anfangswertes



**Abbildung:** Fehler des Krylov-Verfahrens für  $N = 500001$ ,  $\omega = 30$  und die Anfangswerte aus Fall a), b) und c) (Bei b) tritt Konvergenz bereits für  $m = 1$  ein).

## Abschnitt 4

### *Fazit*

- Beide Verfahren sind zur Approximation des Testproblems gut geeignet.
- Beide nutzen Dünnbesetztheit aus.
- Speicher- und Rechenaufwand lassen sich auch für größere Matrizen von einem „normalen“ PC bewältigen.
- Krylov-Verfahren lässt sich ohne Änderung auch auf allgemeinere Matrizen anwenden, deren Wertebereich nicht in einem Intervall liegt oder nicht genau bekannt ist.
- Konvergenz des Krylov-Verfahrens verbessert sich durch einen glatteren Anfangswert.