

Mathematische Grundlagen

1.1. Die Linienmethode

Zur numerischen Behandlung von zeitabhängigen partiellen Differentialgleichungen werden diese diskretisiert. Eine übliche Methode ist, die Orts- und die Zeitdiskretisierung als zwei aufeinanderfolgende Prozesse zu betrachten, obwohl diese im gesamten numerischen Verfahren nicht unabhängig voneinander sind. So unterscheidet man zwischen der (vertikalen) Linienmethode und der Rothe-Methode (oder auch horizontale Linienmethode) [Gro92]. Wird die Lösung eines zeitabhängigen Anfangs-Randwertproblems mit der *Rothe-Methode* approximiert, so wird das Ausgangsproblem durch Diskretisierung bezüglich der Zeitvariablen in eine Folge von elliptischen Randwertproblemen überführt. Bei der (*vertikalen*) *Linienmethode* wird der numerische Integrationsprozeß untergliedert in die Semidiskretisierung bezüglich der Ortsvariablen und anschließender Zeitintegration des resultierenden Anfangswertproblems gewöhnlicher Differentialgleichungen. Im folgenden beschränken wir uns auf die (vertikale) Linienmethode.

Die (vertikale) Linienmethode. In der *Semidiskretisierung bez. der Ortsvariablen* werden die in den ARWPn der partiellen Differentialgleichungen (1), (2) bzw. (3) auftretenden partiellen Ortsableitungen im allgemeinen entweder mittels „finiter Differenzen“ ([Mit80],[Tho95]), mittels „finiter Elemente“ ([Mit78]) oder nach der „Spektralmethode“ ([Got77]) approximiert. Man erhält ein System von Anfangswertproblemen von r gewöhnlichen Differentialgleichungen 1.Ordnung bzw. eine DAE, d.h. ein System der Form

$$\bar{A} w'(t) = f(t, w(t)), \quad w(0) = w_0, \quad w(t) \in \mathbb{R}^r, \quad (1.1.1)$$

mit $f : \bar{J} \times \mathbb{R}^r \longrightarrow \mathbb{R}^r$, $r \in \mathbb{N}$. Für die Problemklassen (1),(2) ist $\bar{A} = I_r$. Für die PDAE (3) ist \bar{A} von A abhängig und singulär, wenn A singulär ist, und der Anfangswert w_0 muß dann gewissen zusätzlichen Bedingungen genügen (siehe Abschnitt 1.2). Insbesondere ist r abhängig von der Anzahl n der Gleichungen des zugrundeliegenden partiellen Differentialgleichungssystems, der Dimension d von Ω und der Art der Ortsdiskretisierung (z.B. der Feinheit des Ortsgitters bei Verwendung „finiter Differenzen“). Die Gleichung (1.1.1) wird auch als *semidiskretes System*

bezeichnet, da nur bezüglich der Ortsvariablen x diskretisiert, aber die Zeit t weiterhin stetig gelassen wurde. In der *Zeitintegration* wird anschließend das erhaltene Problem (1.1.1) durch eine numerische Integrationsmethode gelöst. Hierbei wird im folgenden nur die Verwendung von Einschrittverfahren betrachtet.

1.1.1. Ortsdiskretisierung

Im weiteren werden wir finite Differenzen zur Semidiskretisierung nutzen. Auf $\bar{\Omega}$ wird ein Ortsgitter Ω_h gelegt. Der Einfachheit wird ein Gitter gewählt, das äquidistant bez. einer Raumdimension ist. Für $\Omega = (a, b)^d$ ist

$$\Omega_h = \left\{ x_{i_1, \dots, i_d} = (x_{i_1}, \dots, x_{i_d})^\top : x_{i_k} = a + i_k h_k, \right. \\ \left. i_k = 0(1)M_k + 1, h_k = \frac{b-a}{M_k+1}, k = 1(1)d \right\}. \quad (1.1.2)$$

Die partiellen Ortsableitungen werden nun durch geeignete Differenzenquotienten approximiert. Durch Einsetzen dieser Approximationen in die zu betrachtende partielle Differentialgleichung erhält man für jeden Gitterpunkt eine semidiskrete (zeitabhängige) Gleichung. Aufgrund der Abhängigkeiten zu den benachbarten Gitterpunkten in (1.1.8) und (1.1.9) bilden diese Gleichungen für alle Gitterpunkte ein großes System gewöhnlicher Differentialgleichungen der Form (1.1.1) der Dimension r . Es gilt $r \sim n M_1 \dots M_d$. Jede Komponente der zu bestimmenden Vektorfunktion $w(t)$ ist genau einem Gitterpunkt zugeordnet.

Um die von x abhängige Lösung $u(t, x)$ mit dem bez. des Ortes diskreten Vektor $w(t)$ des \mathbb{R}^r zu vergleichen, führen wir die Bezeichnung $u_h(t)$ als geeignete Repräsentation der exakten Lösung $u(t, x)$ im \mathbb{R}^r ein. Bei einer Semidiskretisierung mittels finiter Differenzen ist $u_h(t)$ die Restriktion der exakten Lösung auf das Gitter. D.h., entspricht die k -te Komponente von $w(t)$ dem Punkt $x_{i_1, \dots, i_d} \in \Omega_h$, so ist die k -te Komponente von $u_h(t)$ gleich dem Wert $u(t, x_{i_1, \dots, i_d})$.

BEMERKUNG 1.1.1. Zur Beurteilung der Güte von Diskretisierungsverfahren werden im folgenden Normen von Fehlergrößen in einem Vektorraum betrachtet. Hierbei werden durch ein Skalarprodukt induzierte Normen des \mathbb{R}^k

$$\|y\| = \sqrt{\langle y, y \rangle} = \sqrt{y^\top R y}, \quad y \in \mathbb{R}^k, \quad (1.1.3)$$

mit einer Matrix $R \in \mathbb{R}^{k \times k}$ (symmetrisch, positiv definit) verwendet.

Die Euklidische Norm (2-Norm) ist mit $R = I_k$ gegeben durch

$$\|y\|_{2,k} = \sqrt{y^\top y} = \sqrt{\sum_{i=1}^k y_i^2}, \quad y \in \mathbb{R}^k.$$

Sei M^d die Anzahl der inneren Ortsgitterpunkte eines äquidistanten Gitters und $w(t) \in \mathbb{R}^{nM^d}$ wie oben beschrieben. Dann wächst die Dimension von w mit feiner

werdendem Ortsgitter ($M \rightarrow \infty$). Es ist daher zweckmäßig, eine diskrete L_2 -Norm des \mathbb{R}^{nM^d} zu verwenden, die man mit $R = h^d I_{nM^d}$ und $h \sim M^{-1}$ erhält, d.h.

$$\|y\| = \sqrt{h^d y^\top y}, \quad y \in \mathbb{R}^{nM^d \times nM^d}. \quad (1.1.4)$$

Insbesondere gilt mit dieser Norm für eine stetige Funktion $u : \bar{\mathcal{J}} \times \bar{\Omega} \rightarrow \mathbb{R}^n$ mit $\int_{\bar{\Omega}} u^\top(t, x) u(t, x) dx < \infty$ ($u \in L_2(\bar{\Omega})$)

$$\|u_h(t)\| \rightarrow \|u(t, x)\|_{L_2} \quad \text{für } h \rightarrow 0 \quad \forall t \in \bar{\mathcal{J}},$$

wobei $\|u(t, x)\|_{L_2}^2 = \int_{\bar{\Omega}} u^\top(t, x) u(t, x) dx$ die L_2 -Norm und $u_h(t)$ die Restriktion von u auf das Ortsgitter ist. D.h., die Normen sind aufeinander abgestimmt (vgl. [Sam84]).

Diese hier genannten Normen werden im folgenden stets verwendet, sofern dies nicht näher erläutert wird. Bezüglich der gewählten Vektornorm ist $\|A\|$ die zugeordnete Matrixnorm und $\mu[A]$ die induzierte logarithmische Matrixnorm für Matrizen $A \in \mathbb{R}^{k \times k}$. \square

Um die Güte der Semidiskretisierung einschätzen zu können, werden der *lokale* und der *globale Ortsdiskretisierungsfehler* untersucht.

DEFINITION 1.1.2. Sei $u_h(t)$ die Restriktion der exakten Lösung $u(t, x)$ der betrachteten partiellen Differentialgleichung der Form (1), (2) oder (3) auf das Gitter Ω_h und $f(t, w)$ die rechte Seite des semidiskreten Problems (1.1.1). Dann heißt die Größe

$$\alpha_h(t) := \bar{A} u'_h(t) - f(t, u_h(t))$$

lokaler Ortsdiskretisierungsfehler.

DEFINITION 1.1.3. Sei $\mathcal{J}^* := [0, t^*]$, $t^* > 0$. Die Semidiskretisierung heißt konsistent bzw. konsistent der Ordnung p^* auf \mathcal{J}^* , wenn

$$\begin{aligned} \max_{t \in \mathcal{J}^*} \|\alpha_h(t)\| &\rightarrow 0 && \text{für } h \rightarrow 0 \\ \text{bzw. } \max_{t \in \mathcal{J}^*} \|\alpha_h(t)\| &= \mathcal{O}(h^{p^*}) && \text{für } h \rightarrow 0 \end{aligned}$$

gleichmäßig in t gilt, wobei $h = \max\{h_k : k = 1(1)d\}$.

Der lokale Ortsdiskretisierungsfehler stellt den Defekt der Gitterfunktion $u_h(t)$ gegenüber der Differentialgleichung (1.1.1) dar. Er gibt an wie 'gut' $u_h(t)$ die Differentialgleichung (1.1.1) erfüllt.

DEFINITION 1.1.4. Sei $w(t)$ die exakte Lösung von (1.1.1), dann heißt der Vektor

$$\eta_h(t) := w(t) - u_h(t) \quad (1.1.5)$$

globaler Ortsdiskretisierungsfehler.

DEFINITION 1.1.5. Sei $\mathfrak{T}^* := [0, t^*]$, $t^* > 0$. Die Semidiskretisierung heißt konvergent bzw. konvergent der Ordnung p^* auf \mathfrak{T}^* , wenn

$$\begin{aligned} \max_{t \in \mathfrak{T}^*} \|\eta_h(t)\| &\rightarrow 0 && \text{für } h \rightarrow 0 \\ \text{bzw. } \max_{t \in \mathfrak{T}^*} \|\eta_h(t)\| &= \mathcal{O}(h^{p^*}) && \text{für } h \rightarrow 0. \end{aligned}$$

Für die partiellen Ableitungen nach den Ortsvariablen werden nun Approximationen 2.Ordnung verwendet. D.h., es werden für $t \in [0, t_e]$ und die inneren Gitterpunkte $x = (x_1, \dots, x_d) \in \Omega$ die Beziehungen

$$\begin{aligned} \frac{\partial^2 u}{\partial x_k^2}(t, x) &= \frac{1}{h_k^2} \left(u(t, x - h_k e_k) - 2u(t, x) + u(t, x + h_k e_k) \right) \\ &\quad + \underbrace{h_k^2 \frac{1}{24} \left(\frac{\partial^4 u}{\partial x_k^4}(t, x - h_k \xi_1 e_k) + \frac{\partial^4 u}{\partial x_k^4}(t, x + h_k \xi_2 e_k) \right)}_{= \mathcal{O}(h_k^2)}, \end{aligned} \quad (1.1.6)$$

$$\begin{aligned} \frac{\partial u}{\partial x_k}(t, x) &= \frac{1}{2h_k} \left(u(t, x + h_k e_k) - u(t, x - h_k e_k) \right) \\ &\quad + \underbrace{h_k^2 \frac{1}{2} \left(\frac{\partial^2 u}{\partial x_k^2}(t, x - h_k \bar{\xi}_1 e_k) - \frac{\partial^2 u}{\partial x_k^2}(t, x + h_k \bar{\xi}_2 e_k) \right)}_{= \mathcal{O}(h_k^2)}, \end{aligned} \quad (1.1.7)$$

$k = 1(1)d$, $e_k = (e_{k1}, \dots, e_{kd})^\top \in \mathbb{R}^d$, $e_{kj} = 0$ für $(k \neq j)$ und $e_{kk} = 1$ benutzt, wobei $h_k > 0$ hinreichend klein sei und $\xi_1, \bar{\xi}_1, \xi_2, \bar{\xi}_2 \in (0, 1)$ gilt. Es sei bemerkt, daß für jede Komponente von $\frac{\partial^4 u}{\partial x_k^4}$ bzw. $\frac{\partial^2 u}{\partial x_k^2}$ die ξ_i bzw. $\bar{\xi}_i$ ($i = 1, 2$) unterschiedlich sein können. Unter Verwendung der Beziehungen (1.1.6) und (1.1.7) approximiert man in jedem Punkt $x_{i_1, \dots, i_d} \in \Omega_h \cap \Omega$ die partiellen Ortsableitungen durch

$$\begin{aligned} \frac{\partial^2 u}{\partial x_k^2}(t, x_{i_1, \dots, i_d}) &\approx \frac{1}{h_k^2} \left(u_{i_1, \dots, i_{k-1}, \dots, i_d}(t) \right. \\ &\quad \left. - 2u_{i_1, \dots, i_d}(t) + u_{i_1, \dots, i_{k+1}, \dots, i_d}(t) \right), \end{aligned} \quad (1.1.8)$$

$$\frac{\partial u}{\partial x_k}(t, x_{i_1, \dots, i_d}) \approx \frac{1}{2h_k} \left(u_{i_1, \dots, i_{k+1}, \dots, i_d}(t) - u_{i_1, \dots, i_{k-1}, \dots, i_d}(t) \right), \quad (1.1.9)$$

wobei $u_{i_1, \dots, i_d}(t) = u(t, x_{i_1, \dots, i_d})$, $k = 1(1)d$. Bei Verwendung von (1.1.8) bzw. (1.1.9) zur Ortsdiskretisierung von (1), (2) bzw. (3) ist die Semidiskretisierung für hinreichend glattes u auf $\mathfrak{T}^* \times \Omega$ konsistent der Ordnung $p^* = 2$.

BEISPIEL 1.1.6. Verwendet man z.B. zur Diskretisierung des ARWPs der eindimensionalen Wärmeleitungsgleichung

$$u_t = u_{xx}, \quad x \in (0, 1), \quad t > 0,$$

mit $u(t, 0) = u(t, 1) = 0$ und $u(0, x) = u_0(x)$ ($x \in [0, 1], t \geq 0$) für die Ortsdiskretisierung mit äquidistanter Gitterweite $h = \frac{1}{M+1}$ die Approximationen (1.1.6) und wählt $w(t) = (w(t, x_1), \dots, w(t, x_M))^T \approx (u(t, x_1), \dots, u(t, x_M))^T =: u_h(t)$, so erhält man das semidiskrete Problem ($t > 0$)

$$w'(t) = A w(t), \quad w(0) = u_h(0), \quad A = \frac{1}{h^2} \begin{pmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -2 \\ & & & & 1 & -2 \end{pmatrix} \in \mathbb{R}^{M \times M}. \quad (1.1.10)$$

□

1.1.2. Zeitintegration semidiskreter parabolischer ARWPe

Im folgenden beschränken wir uns auf das semidiskrete Problem, das aus der Semidiskretisierung parabolischer ARWPe (1) bzw. (2) entsteht¹. D.h., wir betrachten (1.1.1) mit $\bar{A} = I_r$ für festes h :

$$w'(t) = f(t, w(t)), \quad w(0) = w_0. \quad (1.1.11)$$

Wir werden hierzu in diesem Abschnitt auf einige wichtige Grundlagen der Theorie der Einschrittverfahren für gewöhnliche Differentialgleichungen eingehen (vgl. z.B. [Hai93],[Hai96],[Str95]). Auf das Zeitintervall $[0, t_e]$, $t_e < \infty$, wird ein Punktgitter

$$\mathfrak{J}_\tau := \{0, t_1, \dots, t_{m_{end}}\}, \quad 0 = t_0 < t_1 < \dots < t_{m_{end}} = t_e$$

mit den Schrittweiten $\tau_m = t_m - t_{m-1}$, $m = 1(1)m_{end}$, gelegt. Gesucht wird eine Näherungslösung für das Anfangswertproblem (1.1.11), d.h. eine Gitterfunktion

$$v_\tau : \mathfrak{J}_\tau \longrightarrow \mathbb{R}^r \quad \text{mit} \quad v_\tau(t_m) \approx w(t_m), \quad m = 0(1)m_{end}.$$

Unter einem *Diskretisierungsverfahren* zur Approximation der Lösung des Anfangswertproblems wird eine Verfahrensvorschrift verstanden, die jedem $t_m \in \mathfrak{J}_\tau$ einen Vektor v_m zuordnet. Ein *Einschrittverfahren* ist ein Diskretisierungsverfahren zur Bestimmung einer Gitterfunktion $v_m = v_\tau(t_m) \approx w(t_m)$ der Gestalt

$$\begin{aligned} v_{m+1} &= v_m + \tau_m \Phi(t_m, v_m; \tau_m), & m &= 0(1)m_{end}-1 \\ v_0 &= w(0) \end{aligned} \quad (1.1.12)$$

mit $\Phi := \Phi(t, v; \tau) : \text{Verfahrens- oder Inkrementfunktion}$ des Einschrittverfahrens. (Dabei ist die Darstellung (1.1.12) nur formal explizit und umfaßt auch implizite Methoden.)

DEFINITION 1.1.7. Sei $\mathfrak{S} := \{(t, w) : 0 \leq t \leq t_e, w \in \mathbb{R}^r\}$, $f : \mathfrak{S} \rightarrow \mathbb{R}^r$ und $\langle \cdot, \cdot \rangle$ ein Skalarprodukt im \mathbb{R}^r mit der zugehörigen Norm $\|y\| = \sqrt{\langle y, y \rangle}$ ($y \in \mathbb{R}^r$).

¹Das semidiskrete Problem der PDAE (3) wird in Kapitel 3 untersucht.

Dann genügt die Funktion f einer einseitigen Lipschitz-Bedingung, wenn gilt

$$\langle f(t, v) - f(t, w), v - w \rangle \leq l(t) \|v - w\|^2, \quad \text{für alle } (t, v), (t, w) \in \mathfrak{S},$$

wobei $l(t)$ einseitige Lipschitz-Konstante von f auf \mathfrak{S} heißt.

Die Funktion f heißt in \mathfrak{S} (bez. w) Lipschitz-stetig, wenn es eine Konstante $L(t) > 0$ gibt, so daß

$$\|f(t, v) - f(t, w)\| \leq L(t) \|v - w\| \quad \text{für alle } (t, v), (t, w) \in \mathfrak{S}$$

gilt. Die Konstante $L(t)$ heißt Lipschitz-Konstante von f auf \mathfrak{S} .

Es bezeichne $\mathfrak{Lip}(\mathfrak{S}) := \{f \text{ mit } f : \mathfrak{S} \rightarrow \mathbb{R}^r, f \text{ Lipschitz-stetig in } \mathfrak{S}\}$. (Für die Lipschitz-Stetigkeit ist hinreichend, daß die Funktion auf \mathfrak{S} stetig differenzierbar ist.)

VORAUSSETZUNG 1.1.8. Für hinreichend kleine Schrittweiten $\tau > 0$ ordne die Abbildung Φ jeder Funktion $f \in \mathfrak{Lip}(\mathfrak{S})$ eine Funktion $\Phi(t, w; \tau) \in \mathfrak{Lip}(\mathfrak{S})$ zu. Ferner sei Φ in τ stetig.

DEFINITION 1.1.9. Sei \tilde{v}_{m+1} das Resultat eines Schrittes von (1.1.12) mit dem Startvektor auf der exakten Lösungskurve, d.h. $\tilde{v}_{m+1} := w(t_m) + \tau_m \Phi(t_m, w(t_m); \tau_m)$. Dann heißt

$$le_\tau(t_{m+1}) = le_\tau(t_m + \tau_m) := w(t_{m+1}) - \tilde{v}_{m+1}$$

lokaler Zeitdiskretisierungsfehler des Einschrittverfahrens an der Stelle $t = t_m + \tau_m$.

Diese Fehlergröße dient zur qualitativen Beurteilung der Verfahrensfunktion Φ . Sie gibt an, ob diese für jede Funktion $f \in \mathfrak{Lip}(\mathfrak{S})$ eine lokale Approximation der Differentialgleichung (1.1.11) liefert.

DEFINITION 1.1.10. Sei $w(t)$ die Lösung des Anfangswertproblems (1.1.11) auf $[0, t_e]$. Dann heißt ein Einschrittverfahren konsistent (mit dem Anfangswertproblem) bez. einer Norm $\|\cdot\|$, wenn für jede Funktion $f \in \mathfrak{Lip}(\mathfrak{S})$ die Beziehung

$$\max_{m=0(1)m_{end}-1} \|f(t_m, w(t_m)) - \Phi(t_m, w(t_m); \tau_m)\| \rightarrow 0 \quad (1.1.13)$$

für $\tau_{max} \rightarrow 0$ gilt, wobei $\tau_{max} = \max\{\tau_m : m = 0(1)m_{end} - 1\}$.

LEMMA 1.1.11. [Str95] Die Bedingung (1.1.13) ist genau dann erfüllt, wenn

$$\max_{m=0(1)m_{end}-1} \frac{\|le_\tau(t_m + \tau_m)\|}{\tau_m} \rightarrow 0 \quad \text{für } \tau_{max} \rightarrow 0 \quad (1.1.14)$$

gilt.

Für praktische Belange ist die Güte der Approximation \tilde{v}_{m+1} wichtig. Ein Maß für die lokale Genauigkeit ist die Konsistenzordnung.

DEFINITION 1.1.12. *Ein Einschrittverfahren (1.1.12) besitzt die Konsistenzordnung p (klassische Konsistenzordnung p) bez. einer Norm $\|\cdot\|$, wenn p die größte positive ganze Zahl ist, so daß für jede genügend oft stetig differenzierbare Lösung $w(t)$ von (1.1.11) gilt*

$$\max_{t \in \mathcal{J}_\tau \setminus \{t_{m_{end}}\}} \|le_\tau(t + \tau)\| \leq C\tau^{p+1} \quad \text{für} \quad \tau \in (0, T] \quad (1.1.15)$$

mit einer von τ unabhängigen Konstanten C , $T > 0$ hinreichend klein.

C ist abhängig von Schranken für f und partiellen Ableitungen von f bis zur Ordnung p . Insbesondere ist C auch abhängig von der Lipschitz-Konstanten von f . Voraussetzung für die Konsistenzordnung p eines Einschrittverfahrens ist, daß $w(t)$ mindestens $(p+1)$ -mal stetig differenzierbar in $[0, t_e]$ ist.

DEFINITION 1.1.13. *Ein Einschrittverfahren heißt konvergent für die Anfangswertaufgabe (1.1.11) auf $[0, t_e]$ bez. einer Norm $\|\cdot\|$, wenn für jede Folge von Gittern \mathcal{J}_τ mit $\tau_{max} \rightarrow 0$ für den globalen Zeitdiskretisierungsfehler*

$$e_\tau(t) := w(t) - w_\tau(t)$$

die Beziehung $\max_{t \in \mathcal{J}_\tau} \|e_\tau(t)\| \rightarrow 0$ für $\tau_{max} \rightarrow 0$ gilt.

Konvergenz bedeutet, daß das Einschrittverfahren für feiner werdende Diskretisierungen die analytische Lösung $w(t)$ in exakter Arithmetik beliebig genau approximiert. Das Einschrittverfahren besitzt die *Konvergenzordnung p^** , wenn p^* die größte positive ganze Zahl ist, so daß für alle Schrittweiten mit $\tau_{max} \in (0, T]$ gilt

$$\max_{t \in \mathcal{J}_\tau} \|e_\tau(t)\| \leq C\tau_{max}^{p^*}, \quad (1.1.16)$$

wobei die Konstante C unabhängig von τ_{max} ist.

Systeme gewöhnlicher Differentialgleichungen, die durch sehr große Lipschitz-Konstanten gekennzeichnet sind, während sie i.allg. eine einseitige Lipschitz-Konstante von moderater Größe besitzen, werden als *steife Systeme* bezeichnet ([Str95]). Sie stellen hohe Anforderung an die Stabilität numerischer Verfahren. Insbesondere sind Systeme, die wir aus der Semidiskretisierung parabolischer ARWPe (wie in Abschnitt 1.1.1 beschrieben) erhalten, steife Differentialgleichungen:

BEISPIEL 1.1.14. Die Matrix A in Beispiel 1.1.6 hat die Eigenwerte ([Tho95]) $\lambda_k = -\frac{4}{h^2} \sin^2\left(\frac{k\pi h}{2}\right) < 0$ ($k = 1(1)M$). D.h., $\lambda_1 \approx -\pi^2$ und $\lambda_M \approx -\frac{\pi^2}{h^2} \rightarrow -\infty$ für $h \rightarrow 0$. Die Lipschitz-Konstante der rechten Seite von (1.1.10) ist durch $\|A\|$ und die einseitige Lipschitz-Konstante durch $\mu[A]$ gegeben (vgl. [Dek84]). Da A symmetrisch ist, gilt in der Spektralnorm $\|A\|_{2,M} = \max_{k=1(1)M} \{|\lambda_k|\} = -\lambda_M$ und in der der 2-Norm zugeordneten logarithmischen Matrixnorm $\mu[A] = \max_{k=1(1)M} \{\lambda_k\} = \lambda_1 \approx -\pi < 0$. D.h., (1.1.10) ist steif. \square

1.1.3. Konsistenz und Konvergenz der Gesamtdiskretisierung

In Anlehnung an Verwer/Sanz-Serna, die in [Ver84] die Konvergenz von auf der Linienmethode basierenden Approximationen partieller Differentialgleichungen untersuchen (siehe auch [Sha94]), wird in diesem Abschnitt auf die Konvergenz der Gesamtdiskretisierung, die sich aus der Orts- und der Zeitdiskretisierung zusammensetzt, eingegangen.

Hierzu betrachten wir den Gesamtdiskretisierungsfehler:

DEFINITION 1.1.15. *Der Gesamtdiskretisierungsfehler der numerischen Lösung v_{m+1} eines Diskretisierungsverfahrens zur Lösung von (1),(2) oder (3) gegenüber der exakten Lösung zum Zeitpunkt t_{m+1} ist definiert durch*

$$\varepsilon_h(t_{m+1}) := u_h(t_{m+1}) - v_{m+1}. \quad (1.1.17)$$

Der Einfachheit halber betrachten wir äquidistante Schrittweiten h und τ . Wir halten $t_{m+1} \in \mathfrak{T}$ fest und nehmen im Grenzprozeß an, daß $\tau \rightarrow 0$ und $m \rightarrow \infty$ so erfolgen, daß $(m+1)\tau = t_{m+1}$.

DEFINITION 1.1.16. *Eine Gesamtdiskretisierung heißt bez. einer Norm $\|\cdot\|$ konvergent der Ordnung (p, q) , wenn die Ordnungsbedingung*

$$\|\varepsilon_h(t_{m+1})\| = \mathcal{O}(h^p) + \mathcal{O}(\tau^q) \quad \text{für } \tau, h \rightarrow 0, m = 0, 1, \dots \quad (1.1.18)$$

erfüllt ist.

Es ist $\varepsilon_h(t_{m+1}) = u_h(t_{m+1}) - w(t_{m+1}) + w(t_{m+1}) - v_{m+1}$ und

$$\|\varepsilon_h(t_{m+1})\| \leq \|\eta_h(t_{m+1})\| + \|e_\tau(t_{m+1})\|. \quad (1.1.19)$$

Der Gesamtdiskretisierungsfehler verschwindet aber nicht notwendig, wenn η_h für $h \rightarrow 0$ und e_τ für $\tau \rightarrow 0$ verschwinden. Dies resultiert daraus, daß i.allg. in die Abschätzungen für $\|e_\tau(t_{m+1})\|$ die Lipschitz-Konstante des gewöhnlichen Systems einfließt. Diese wird für semidiskrete Probleme, die z.B. aus einer PDE der Form (1) resultieren, für $h \rightarrow 0$ beliebig groß (vgl. Beispiel 1.1.14).

Für die Konvergenz der Gesamtdiskretisierung muß daher gefordert werden, daß das Verfahren für die Zeitintegration bez. des Ortsgitters unabhängig vom Verhältnis von Orts- und Zeitschrittweite konvergiert. D.h., für $v_{m+1} \rightarrow u_h(t_{m+1})$ mit $h, \tau \rightarrow 0$ muß $v_{m+1} \rightarrow w(t_{m+1})$ für $\tau \rightarrow 0$ unabhängig von h (*gleichmäßig*) sein ([Dek84], [Auz90]).

DEFINITION 1.1.17. *Gilt die Beziehung (1.1.18) nur unter einer Bedingung zwischen τ und h , so heißt die Gesamtdiskretisierung bedingt konvergent der Ordnung (p, q) .*

BEISPIEL 1.1.18. Wendet man auf das semidiskrete Problem (1.1.10) das explizite Euler-Verfahren $v_{m+1} = v_m + \tau f(t_m, v_m)$ an, so gilt $\|\eta_h(t_{m+1})\| = \mathcal{O}(h^2)$ für $h \rightarrow 0$ und $\|e_\tau(t_{m+1})\| = \mathcal{O}(\tau)$ für h fest und $\tau \rightarrow 0$. Allerdings erhält man $\|e_\tau(t_{m+1})\| = \mathcal{O}(\tau)$ für $\tau, h \rightarrow 0$ nur unter der Bedingung $\frac{\tau}{h^2} \leq \frac{1}{2}$ (vgl. z.B. [Ric67]). Dies bedeutet für $h \rightarrow 0$ eine starke Schrittweitereinschränkung an τ , und die Gesamtdiskretisierung ist in diesem Beispiel nur bedingt konvergent. \square

Für steife gewöhnliche Differentialgleichungssysteme wurde das Konzept der sogenannten *B-Konsistenz* entwickelt ([Fra81],[Dek84]), auf die in Abschnitt 2.5.3 eingegangen wird. Es liefert bez. der Steifheit gleichmäßige Fehlerabschätzungen. Die Übertragung dieses Prinzips auf die Konvergenzanalyse der Linienmethode bedeutet, daß Fehlerschranken gesucht werden, die unabhängig vom Ortsdiskretisierungsparameter h sind.

Verwer/Sanz-Serna weisen in ihrer Arbeit [Ver84] darauf hin, daß es bei Verwendung von (1.1.19) i.allg. kompliziert sein kann, die in h gleichmäßige Konvergenz von $e_\tau(t_{m+1})$ bez. τ nachzuweisen. Sie stellen fest, daß es für die Konvergenzanalyse meist günstiger ist, den Ausdruck

$$\varepsilon_h(t_{m+1}) = le_h(t_{m+1}) + \hat{v}_{m+1} - v_{m+1}$$

auszuwerten, wobei $le_h(t_{m+1})$ der lokale Gesamtdiskretisierungsfehler ist:

DEFINITION 1.1.19. *Der lokale Gesamtdiskretisierungsfehler eines Diskretisierungsverfahrens zur Lösung von (1),(2) oder (3) ist definiert durch*

$$le_h(t_{m+1}) := u_h(t_{m+1}) - \hat{v}_{m+1}, \quad (1.1.20)$$

wobei \hat{v}_{m+1} das Resultat eines Zeitschrittes mit dem Einschrittverfahren mit dem Startvektor auf der exakten Lösungskurve ist, d.h. wenn $v_m = u_h(t_m)$.

BEMERKUNG 1.1.20. Die Untersuchung der klassischen Konvergenzordnung ist im Zusammenhang mit der Linienmethode weiterhin eine übliche Vorgehensweise zur Herleitung effektiver Diskretisierungsverfahren, da sie die Güte des Diskretisierungsverfahrens bez. der Zeitdiskretisierung charakterisiert. Eine gute klassische Konsistenzordnung ist somit eine notwendige Eigenschaft für die Güte der Gesamtdiskretisierung. \square

1.2. Indexe für DAEs

Im folgenden Abschnitt wollen wir für unsere Untersuchungen relevante Begriffe und Ergebnisse aus der Theorie der DAEs (auch als Algebro-Differentialgleichungen, differential-algebraische oder differentiell-algebraische Gleichungen bezeichnet) kurz

vorstellen (vgl. z.B. [Pet82],[Gri86],[Hai96], [Bre89], [Deu94], [Ren96],[ES98]). Eine DAE hat die allgemeine Form

$$F(t, w(t), w'(t)) = 0, \quad \text{mit } t \in \mathfrak{J}, \quad w, w' : \mathfrak{J} \rightarrow \mathbb{R}^r, \quad (1.2.1)$$

wobei die Funktion $F \in \mathbb{R}^r$ stetig ist und eine stetige partielle Ableitung $\frac{\partial F}{\partial w'}$ hat. Ist $\frac{\partial F}{\partial w'}$ in einer Umgebung der Lösung $w(t)$ regulär, so ist (1.2.1) nach $w'(t)$ lokal auflösbar, d.h., (1.2.1) stellt dann eine implizite ODE dar. Dies ist nicht der Fall, wenn die Jacobi-Matrix $\frac{\partial F}{\partial w'}$ singular ist. Sei $\frac{\partial F}{\partial w'}$ singular im gesamten betrachteten Gebiet. Dann enthält (1.2.1) auch *algebraische* Gleichungen und man nennt (1.2.1) eine DAE.

Da wir in Kapitel 3 der vorliegenden Arbeit lineare PDAEs untersuchen, wollen wir in diesem Abschnitt einige Resultate für lineare DAEs mit konstanten Koeffizienten der Form

$$X w'(t) + Y w(t) = g(t), \quad w(0) = w_0, \quad t \in \mathfrak{J}, \quad X, Y \in \mathbb{R}^{r \times r} \quad (1.2.2)$$

zusammenstellen (vgl. z.B. [Gri86],[Bre89]). $g(t) \in \mathbb{R}^r$ ist eine gegebene Vektorfunktion. Wenn die Matrix X regulär ist, ist (1.2.2) eine ODE. Sei im folgenden X singular. Das Lösungsverhalten einer linearen DAE (1.2.2) ist bestimmt durch das Matrixbüschel (X, Y) .

DEFINITION 1.2.1. *Ein Matrixbüschel (auch Matrizenbüschel, engl. matrix pencil) (X, Y) zweier Matrizen $X, Y \in \mathbb{C}^{r \times r}$ ist eine Familie $(X, Y) = \{\lambda X + Y\}_{\lambda \in \mathbb{C}}$.*

DEFINITION 1.2.2. *Der Index $\text{ind}(X)$ einer Matrix X ist die kleinste nichtnegative Zahl k , für die $\text{rang}(X^k) = \text{rang}(X^{k+1})$ gilt.*

DEFINITION 1.2.3. *Man nennt das Matrixbüschel regulär, wenn das Polynom $d(\lambda) := \det(\lambda X + Y)$ ($\lambda \in \mathbb{C}$) nicht identisch verschwindet, und singular, wenn $d(\lambda) = 0 \forall \lambda \in \mathbb{C}$.*

DEFINITION 1.2.4. *Sei (X, Y) ein reguläres Matrixbüschel und $\lambda \in \mathbb{C}$ mit $\lambda X + Y$ regulär. Der Index $\text{ind}(X, Y)$ eines regulären Matrixbüschels (X, Y) ist definiert als $\text{ind}(X, Y) := \text{ind}([\lambda X + Y]^{-1} X)$.*

Der Index $\text{ind}(X, Y)$ in Definition 1.2.4 ist unabhängig vom gewählten λ .

LEMMA 1.2.5. [Gri86] *Sei (X, Y) ein reguläres Matrixbüschel mit $X, Y \in \mathbb{C}^{r \times r}$ und $T \in \mathbb{C}^{r \times r}$ eine nichtsinguläre Matrix. Dann gilt*

$$\text{ind}(T X, T Y) = \text{ind}(X T, Y T) = \text{ind}(X, Y).$$

Für singuläre Matrixbüschel hat die DAE (1.2.2) keine oder unendlich viele Lösungen für einen gegebenen Anfangswert. Wir wollen daher im weiteren nur reguläre Matrixbüschel (X, Y) zugrunde legen. Dann kann man die Matrizen X, Y der

Weierstraß-Kronecker-Transformation unterziehen (vgl. [Wei68],[Kro90],[Gan86]). Für reguläre Matrixbüschel (X, Y) existieren reguläre Matrizen $S, T \in \mathbb{C}^{r \times r}$, so daß

$$SXT = \begin{pmatrix} I_{r_1} & 0 \\ 0 & N \end{pmatrix}, \quad SYT = \begin{pmatrix} R & 0 \\ 0 & I_{r_2} \end{pmatrix}, \quad N_j = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 0 \end{pmatrix} \in \mathbb{N}^{r_{2j} \times r_{2j}}$$

gilt, wobei $N = \text{blockdiag}(N_1, \dots, N_k) \in \mathbb{N}^{r_2 \times r_2}$ mit $r_2 = r_{21} + \dots + r_{2k}$, $r = r_1 + r_2$, und $R \in \mathbb{C}^{r_1 \times r_1}$ sind. Die Matrizen $I_{r_j} \in \mathbb{N}^{r_j \times r_j}$, $j = 1, 2$, sind Einheitsmatrizen. Die Jordan-Kettenmatrix N ist eine *nilpotente* Matrix, d.h., es gibt eine Zahl $\nu \in \mathbb{N}^+$, so daß $N^\nu = 0$ gilt. Für das kleinste ν , das dies erfüllt, gilt $\text{ind}(N) = \nu$, und ν wird *Nilpotenz* oder *Riesz-Index* von N genannt. Weiterhin kann man zeigen, daß $\text{ind}(X, Y) = \text{ind}(N) = \nu$, d.h., der Index des Matrixbüschels (X, Y) ist gleich der Nilpotenz ν der Matrix N . Seien $T(x^\top, y^\top)^\top := w$ und $S(s_1^\top, s_2^\top)^\top := g$. Dann können wir (1.2.2) in das entkoppelte System

$$x'(t) + Rx(t) = s_1(t), \quad Ny'(t) + y(t) = s_2(t) \quad (1.2.3)$$

transformieren. Die Gleichung für $x(t)$ ist eine ODE. Für $y(t)$ gilt:

$$\begin{aligned} y(t) &= s_2(t) - Ny'(t) = s_2(t) - Ns_2'(t) + N^2y''(t) = \dots \\ &= s_2(t) - Ns_2'(t) + \dots (-1)^{\nu-1} N^{\nu-1} s_2^{(\nu-1)}(t) + (-1)^\nu \underbrace{N^\nu}_{=0} y^{(\nu)}(t) \\ &= \sum_{i=0}^{\nu-1} (-N)^i \frac{d^i}{dt^i} s_2(t). \end{aligned} \quad (1.2.4)$$

Diese Lösungsdarstellung verdeutlicht wesentliche Unterschiede der Lösung einer linearen DAE zur Lösung einer linearen ODE: Die Gleichung für $y(t)$ in (1.2.3) ist einer algebraischen Gleichung äquivalent, und $y(t)$ ist vollständig durch die Komponenten der rechten Seite $s_2(t)$ und ihre Ableitungen bestimmt. Dies bedeutet insbesondere daß das Differentialgleichungssystem (1.2.3) für $\nu > 1$ im Gegensatz zu ODEs nur für hinreichend oft stetig differenzierbare Funktionen $s_2(t)$ eine klassische Lösung hat. Es ergeben sich unterschiedlich starke Forderungen an die Differenzierbarkeit der einzelnen Komponenten der rechten Seite der DAE. Der Index ν des Matrixbüschels (X, Y) gibt folglich an, daß mindestens eine Komponente der rechten Seite $f(t)$ der DAE (1.2.2) $\nu - 1$ mal stetig differenzierbar sein muß. Weiterhin müssen die Anfangswerte w_0 des Anfangswertproblems (1.2.2) sogenannten *Konsistenzbedingungen* genügen und können daher nicht beliebig vorgegeben werden.

Mit

$$\begin{aligned} x &= \bar{T}_1 w \\ y &= \bar{T}_2 w \end{aligned} \quad \text{wenn} \quad T^{-1} = \begin{pmatrix} \bar{T}_1 \\ \bar{T}_2 \end{pmatrix} \quad \text{mit} \quad \bar{T}_1 \in \mathbb{C}^{r_1 \times r}, \bar{T}_2 \in \mathbb{C}^{r_2 \times r}.$$

lauten die Konsistenzbedingungen an w_0 :

$$\bar{T}_2 w_0 = \sum_{i=0}^{\nu-1} (-N)^i s_2^{(i)}(0) = y(0).$$

Anfangswerte, die dieser Bedingung genügen, werden *konsistente Anfangswerte* der DAE (1.2.2) genannt. Die Untersuchung nichtlinearer DAEs (1.2.1) ist komplizierter. Sie werden durch den von Gear 1988 ([Gea88],[Gea90]) eingeführten *Differentiationsindex* charakterisiert. Im folgendem setzen wir voraus, daß F genügend oft stetig differenzierbar ist.

DEFINITION 1.2.6. *Die differentiell-algebraische Gleichung (1.2.1) hat den Differentiationsindex $di = \nu$, wenn ν die minimale Anzahl von Differentiationen*

$$\begin{aligned} F(t, w, w') &= 0 \\ \frac{d}{dt} F(t, w, w') &= F_t + F_w w' + F_{w'} w'' = 0 \\ &\vdots \\ \frac{d^\nu}{dt^\nu} F(t, w, w') &= \frac{\partial^\nu}{\partial t^\nu} F + \dots + F_{w'} w^{(\nu+1)} = 0 \end{aligned}$$

ist, so daß die Gleichung (1.2.1) durch algebraische Umformungen in ein explizites gewöhnliches Differentialgleichungssystem $w'(t) = f(t, w)$ überführt werden kann. Dieses System heißt das der DAE (1.2.1) zugrundeliegende Differentialgleichungssystem, welches nicht unabhängig von den sogenannten Zwangsbedingungen (engl: constrains) betrachtet werden darf ([Mär98]).

Der Differentiationsindex macht Aussagen über die Struktur der DAE und charakterisiert den algebraischen Teil der DAE. Er verdeutlicht den Unterschied von DAEs zu gewöhnlichen Differentialgleichungssystemen, indem die Überführbarkeit der einen Klasse in die andere betrachtet wird. (Eine explizite ODE hat folglich den Differentiationsindex $di = 0$.)

BEMERKUNG 1.2.7. Der Differentiationsindex di der linearen DAE (1.2.2) stimmt mit dem Index $\nu = \text{ind}(X, Y)$ des Matrixbüschels (X, Y) überein, d.h., im lineare Fall gilt $di = \nu$. \square

SATZ 1.2.8. [Gri86] *Die lineare DAE mit konstanten Koeffizienten der Form*

$$\begin{aligned} X_1 w'(t) + Y_1 w(t) &= g_1 \\ Y_2 w(t) &= g_2 \end{aligned} \tag{1.2.5}$$

$(X_1, Y_1 \in \mathbb{R}^{n_1 \times n}, Y_2 \in \mathbb{R}^{(n-n_1) \times n})$ hat den Index 1, genau dann wenn $\begin{pmatrix} X_1 \\ Y_2 \end{pmatrix} \in \mathbb{R}^{n \times n}$ eine nichtsinguläre Matrix ist.

Sei $X_1 = (I_{n_1} \ 0) \in \mathbb{R}^{n_1 \times n}$, $Y_1 = (Y_{11} \ Y_{12})$, $Y_2 = (Y_{21} \ Y_{22})$ ($Y_{ij} \in \mathbb{R}^{n_i \times n_j}$, $i, j = 1, 2$). Da I_{n_1} regulär, ist $\begin{pmatrix} X_1 \\ Y_2 \end{pmatrix}$ genau dann regulär, wenn Y_{22} regulär ist. Hieraus ergibt sich

FOLGERUNG 1.2.9. *Die lineare, semi-explizite DAE*

$$\begin{aligned} w_1'(t) + Y_{11} w_1(t) + Y_{12} w_2(t) &= g_1(t) \\ Y_{21} w_1(t) + Y_{22} w_2(t) &= g_2(t) \end{aligned} \tag{1.2.6}$$

$(w_i \in \mathbb{R}^{n_i}, Y_{ij} \in \mathbb{R}^{n_i \times n_j}, i, j = 1, 2)$ hat den Index 1, genau dann wenn Y_{22} regulär ist.

BEMERKUNG 1.2.10. Auf andere Indexbegriffe für DAEs, w.z.B. auf den von Hairer/Lubich/Roche 1989 ([Hai89]) eingeführten *Störungsindex*, wollen wir nicht eingehen. \square