Gewöhnliche Differenzialgleichungen *

Martin Brokate **

Inhaltsverzeichnis

*Vorlesungsskript, SS 2018

**Zentrum Mathematik, TU München

1 Lineare autonome Systeme

Ein einführendes Beispiel. Wir betrachten die Funktion x, definiert durch

$$x(t) = ce^{at}\,, \qquad t, a, c \in \mathbb{R}\,. \tag{1.1}$$

Für deren Ableitung gilt $\dot{x}(t) = ce^{at} \cdot a$. Die durch (1.1) definierte Funktion erfüllt also die Gleichung

$$\dot{x}(t) = ax(t)\,, \quad \text{für alle } t \in \mathbb{R}. \tag{1.2}$$

Wir schreiben sie kürzer als

$$\dot{x} = ax\,. \tag{1.3}$$

Eine solche Gleichung heißt **Differenzialgleichung**. Sie verknüpft Funktionswerte und Ableitungen einer unbekannten Funktion. Genauer bezeichnet man sie als **gewöhnliche Differenzialgleichung**: Das Argument t ist eindimensional (es treten daher "gewöhnliche" und keine partiellen Ableitungen auf), und Funktionswert und Ableitung werden an jeweils derselben Stelle t in Beziehung zueinander gesetzt (in Kontrast etwa zu der Differenzialgleichung, die durch $\dot{x}(t) = ax(t-h)$ mit $h > 0$ definiert wird).

Eine differenzierbare Funktion $x : \mathbb{R} \to \mathbb{R}$, welche (1.2) erfüllt, heißt **Lösung** von (1.2) auf $\mathbb{R}$. Ist $a \in \mathbb{R}$ gegeben, so ist jede Funktion der Form (1.1) mit frei wählbarem $c \in \mathbb{R}$ eine Lösung. Geben wir zusätzlich $x_0 \in \mathbb{R}$ vor und verlangen

$$x(0) = x_0\,, \tag{1.4}$$

so ist c durch die Bedingung $x_0 = ce^0 = c$ eindeutig festgelegt. Fasst man (1.3) und (1.4) zusammen, so erhält man das **Anfangswertproblem**

$$\dot{x} = ax\,, \quad x(0) = x_0\,. \tag{1.5}$$

Gibt es außer Funktionen der Form (1.1) noch weitere Lösungen? Sei $y : \mathbb{R} \to \mathbb{R}$ eine beliebige Lösung. Es gilt

$$\frac{\mathrm{d}}{\mathrm{dt}}\big(e^{-at}y(t)\big) = -ae^{-at}y(t) + e^{-at}\dot{y}(t) = e^{-at}(\dot{y}(t) - ay(t)) = 0$$

für alle t, also $e^{-at}y(t) = c$ für ein $c \in \mathbb{R}$ und damit $y(t) = ce^{at}$. Die Antwort ist also "nein".

Bei Differenzialgleichungen interessiert man sich auch für das "asymptotische Verhalten". Im obigen Beispiel hängt das Verhalten der Lösung für $t \to \infty$ vom Vorzeichen von a ab. Ist $a = 0$, so ist die Lösung konstant. Ist $a > 0$, so gilt $|x(t)| \to \infty$ für $t \to \infty$; ist $a < 0$, so gilt $x(t) \to 0$ für $t \to \infty$.

Systeme von Differenzialgleichungen. Sie haben die Form

$$\dot{x} = f(t, x)\,, \quad t \in \mathbb{R}\,, \, x \in \mathbb{R}^n\,, \tag{1.6}$$

mit der **rechten Seite** f. Sie bestehen aus n einzelnen Differenzialgleichungen $\dot{x}_i = f_i(t, x)$ für die Komponenten x_i von x. Diese sind verkoppelt, da f_i im allgemeinen von allen Komponenten x_j (nicht nur von x_i) abhängt. Das System heißt **autonom**, falls f nicht von t abhängt,

$$\dot{x} = f(x)\,. \tag{1.7}$$

Die Lösungen sind Kurven $t \mapsto x(t)$. Der Raum, in dem die Lösungskurven verlaufen, heißt **Zustandsraum** oder **Phasenraum**. Das System (1.6) hat als Zustandsraum den $\mathbb{R}^n$.

Lineare Systeme. Man spricht von einem linearen System, wenn die rechte Seite f linear von x abhängt. Ein lineares System hat die Form

$$\dot{x} = A(t)x\,, \tag{1.8}$$

wobei $A(t)$ eine $n \times n$ Matrix ist. Für lineare Systeme gilt das **Superpositionsprinzip**: Sind die Funktionen $x^1, \dots, x^k$ Lösungen von (1.8), so ist auch jede Linearkombination $\sum_{j=1}^k c_j x^j$ Lösung. Es gilt nämlich

$$\frac{\mathrm{d}}{\mathrm{dt}}\Big(\sum_{j=1}^k c_j x^j\Big)(t) = \sum_{j=1}^k c_j \dot{x}^j(t) = \sum_{j=1}^k c_j A(t) x^j(t) = A(t)\Big(\sum_{j=1}^k c_j x^j(t)\Big)\,.$$

Lineare autonome Systeme. Sie haben die Form

$$\dot{x} = Ax\,, \quad A \in \mathbb{R}^{(n,n)}\,. \tag{1.9}$$

Sei A Diagonalmatrix, $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Dann zerfällt das System (1.9) in n einzelne Gleichungen

$$\dot{x}_i = \lambda_i x_i\,, \quad 1 \le i \le n\,,$$

mit den Lösungen

$$x_i(t) = c_i e^{\lambda_i t}\,, \quad c_i \in \mathbb{R}\,.$$

Sei $A \in \mathbb{R}^{(n,n)}$ mit Eigenwert $\lambda \in \mathbb{R}$ und Eigenvektor $v \in \mathbb{R}^n$, also $Av = \lambda v$. Dann ist $x(t) = e^{\lambda t} v$ eine Lösung von (1.9), da

$$\dot{x}(t) = \lambda e^{\lambda t} v = e^{\lambda t} A v = A x(t)\,.$$

Sei $A \in \mathbb{R}^{(n,n)}$ mit Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ und Eigenvektoren $v_1, \dots, v_n \in \mathbb{R}^n$. Dann ist jede Linearkombination

$$x(t) = \sum_{j=1}^n c_j e^{\lambda_j t} v_j \tag{1.10}$$

Lösung von (1.9), gemäß dem Superpositionsprinzip. Ist $V = (\, v_1 \,|\, \dots \,|\, v_n \,) \in \mathbb{R}^{(n,n)}$ die Matrix mit den Spalten $v_1, \dots, v_n$, so können wir die Vektorgleichungen $Av_j = \lambda_j v_j$ zu einer Matrixgleichung $AV = VD$ zusammenfassen mit $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Sind die Eigenvektoren $v_1, \dots, v_n \in \mathbb{R}^n$ linear unabhängig, so ist V invertierbar, und es folgt

$$V^{-1}AV = D\,, \quad D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)\,, \quad V, D \in \mathbb{R}^{(n,n)}\,. \tag{1.11}$$

Wir werden später sehen, dass es außer den Lösungen (1.10) keine anderen Lösungen gibt. Damit ist der Fall "A reell diagonalisierbar" hinsichtlich der Struktur der Lösungen "im Prinzip" geklärt.

Das Phasenporträt. Im Fall $n = 2$ können wir uns die Lösungen eines Systems von Differenzialgleichungen veranschaulichen, indem wir Lösungskurven $x(t) = (x_1(t), x_2(t))$ in der Zustandsebene (Phasenebene) graphisch darstellen. Eine solche Darstellung heißt

Phasenporträt oder **Phasendiagramm**. Wir betrachten Beispiele im linearen autonomen Fall, also $A = \operatorname{diag}(\lambda_1, \lambda_2)$.

$$\begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(Phasenporträts an der Tafel)

Komplexe Eigenwerte. Eine reelle Matrix kann komplexe Eigenwerte haben. Beispielsweise hat die Matrix

$$A = \begin{pmatrix} \alpha & -\omega \\ \omega & \alpha \end{pmatrix}, \quad \alpha, \omega \in \mathbb{R},$$

die konjugiert komplexen Eigenwerte $\lambda = \alpha + i\omega$ und $\overline{\lambda} = \alpha - i\omega$. Für die Zustandsvariable $x = (x_1, x_2)$ betrachten wir das System

$$\dot{x} = \begin{pmatrix} \alpha & -\omega \\ \omega & \alpha \end{pmatrix} x. \tag{1.12}$$

Wir interpretieren den Zustand x als komplexe Zahl $z = x_1 + ix_2$. Anstelle des Systems (1.12) im $\mathbb{R}^2$ betrachten wir eine einzelne Differenzialgleichung in $\mathbb{C}$,

$$\dot{z} = \lambda z, \quad \text{mit } \lambda = \alpha + i\omega, \tag{1.13}$$

wobei wir die Multiplikation auf der rechten Seite als Multiplikation in $\mathbb{C}$ verstehen. Es gilt dann

$$\dot{x}_1 + i\dot{x}_2 = \dot{z} = \lambda z = (\alpha + i\omega)(x_1 + ix_2) = (\alpha x_1 - \omega x_2) + i(\omega x_1 + \alpha x_2). \tag{1.14}$$

Real- und Imaginärteil dieser Gleichung entsprechen gerade den beiden Komponenten der Gleichung (1.12). Eine Lösung von (1.13) ist

$$z(t) = e^{\lambda t}, \tag{1.15}$$

wie man durch Differenzieren der Potenzreihenentwicklung

$$e^{\lambda t} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} t^k$$

erkennt. Aus (1.15) erhalten wir

$$z(t) = e^{\lambda t} = e^{(\alpha + i\omega)t} = e^{\alpha t} e^{i\omega t}. \tag{1.16}$$

Der Faktor $e^{i\omega t}$ beschreibt eine Drehung mit konstanter Winkelgeschwindigkeit und Frequenz $\omega/2\pi$ (= Anzahl der Kreisdurchläufe pro Zeiteinheit). Multipliziert mit dem Faktor $e^{\alpha t}$ ergibt sich eine spiralförmige Bewegung (für $t \to \infty$ gegen Null oder ins Unendliche, je nach Vorzeichen von α), und eine Kreisbewegung für $\alpha = 0$. In kartesischen Koordinaten erhalten wir

$$z(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = e^{\alpha t} \begin{pmatrix} \cos(\omega t) \\ \sin(\omega t) \end{pmatrix}. \tag{1.17}$$

Im Fall $\alpha = 0$ wird das System (1.12) zu

$$\dot{x}_1 = -\omega x_2 \,, \quad \dot{x}_2 = \omega x_1 \,.$$

Beide Komponenten x_1 und x_2 sind Lösungen der Differenzialgleichung zweiter Ordnung für eine reelle Variable y,

$$\ddot{y} + \omega^2 y = 0 \,. \tag{1.18}$$

Diese Gleichung beschreibt den sogenannten **harmonischen Oszillator**. Matrizen mit komplexen Eigenwerten sind somit ein grundlegendes Hilfsmittel zur Beschreibung von Schwingungsvorgängen.

Allgemeiner Lösungsansatz. Ziel ist es, die Formel $x(t) = e^{at}x_0$ für das skalare Anfangswertproblem $\dot{x} = ax$, $x(0) = x_0$, auf Systeme $\dot{x} = Ax$ zu verallgemeinern, so dass

$$x(t) = e^{tA}x_0 \tag{1.19}$$

eine Lösung wird. Zu diesem Zweck wollen wir in der Exponentialfunktion als Argumente Matrizen (statt Zahlen) zulassen. Die naheliegende Definition ist

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} \,.$$

Reihen im normierten Raum.

Definition 1.1
Sei $(X, \|\cdot\|)$ normierter Raum über $\mathbb{K}$ (steht für $\mathbb{R}$ oder $\mathbb{C}$), sei $(x_k)_{k\geq 0}$ Folge in X. Falls die Folge der Partialsummen

$$s_n = \sum_{k=0}^{n} x_k \tag{1.20}$$

gegen ein $s \in X$ konvergiert, so sagen wir, dass die zugehörige Reihe $\sum_{k=0}^{\infty} x_k$ konvergiert, und definieren

$$\sum_{k=0}^{\infty} x_k = s \,. \tag{1.21}$$

*Die Reihe $\sum_{k=0}^{\infty} x_k$ heißt **absolut konvergent**, falls*

$$\sum_{k=0}^{\infty} \|x_k\| < \infty \,. \tag{1.22}$$

□

Aus der Stetigkeit der Addition und Skalarmultiplikation im normierten Raum folgen die Rechenregeln

$$\sum_{k=0}^{\infty}(x_k + y_k) = \sum_{k=0}^{\infty} x_k + \sum_{k=0}^{\infty} y_k \,, \quad \sum_{k=0}^{\infty} \lambda x_k = \lambda \sum_{k=0}^{\infty} x_k \,, \tag{1.23}$$

$\lambda \in \mathbb{K}$, die jeweils gültig sind, wenn die Grenzwerte auf der rechten Seite existieren.

Satz 1.2 *Sei $(X, \|\cdot\|)$ Banachraum. Ist die Reihe $\sum_{k=0}^{\infty} x_k$ absolut konvergent, so ist sie auch konvergent, und es gilt*

$$\left\| \sum_{k=0}^{\infty} x_k \right\| \leq \sum_{k=0}^{\infty} \|x_k\| \,. \tag{1.24}$$

Beweis: Sei

$$\sigma_n = \sum_{k=0}^{n} \|x_k\| \,,$$

dann gilt für die in (1.20) definierten Partialsummen für $n > m$

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^{n} x_k \right\| = |\sigma_n - \sigma_m| \,. \leq \sum_{k=m+1}^{n} \|x_k\| = |\sigma_n - \sigma_m| \,.$$

Da (σ_n) Cauchyfolge in $\mathbb{R}$ ist, ist (s_n) Cauchyfolge in X, also konvergent. Wegen $\|s_n\| \leq |\sigma_n|$ folgt (1.24) aus

$$\|s\| = \lim_{n\to\infty} \|s_n\| \leq \lim_{n\to\infty} |\sigma_n| = \sum_{k=0}^{\infty} \|x_k\| \,,$$

wobei die Stetigkeit der Norm verwendet wurde. □

Matrixnormen. Wir betrachten Normen auf dem Matrizenraum $\mathbb{K}^{(n,n)}$, die durch eine Norm auf $\mathbb{K}^n$ erzeugt werden.

Definition 1.3
Sei $\|\cdot\|$ eine Norm auf $\mathbb{K}^n$. Für $A \in \mathbb{K}^{(n,n)}$ definieren wir

$$\|A\| = \sup_{\substack{x \in \mathbb{K}^n \\ \|x\|=1}} \|Ax\| \,. \tag{1.25}$$

□

Unmittelbar aus der Definition folgt

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \,, \tag{1.26}$$

sowie

$$\|I\| = 1 \,.$$

Der Rand $\{\|x\| = 1\}$ der Einheitskugel im $\mathbb{K}^n$ ist kompakt. Daher wird in (1.25) das Maximum angenommen,

$$\|A\| = \max_{\substack{x \in \mathbb{K}^n \\ \|x\|=1}} \|Ax\| \,. \tag{1.27}$$

Lemma 1.4 *Durch (1.25) wird eine Norm auf $\mathbb{K}^{(n,n)}$ definiert. Sie heißt die von der gegebenen Vektornorm auf $\mathbb{K}^n$ erzeugte **Operatornorm** oder **Matrixnorm**.*

Beweis: Folgt direkt aus den Definitionen. □

Wir formulieren einige allgemeine Eigenschaften von Matrixnormen.

Lemma 1.5 *Sei* $(\|\cdot\|)$ *Norm auf* $\mathbb{K}^n$. *Für die zugehörige Matrixnorm gilt*

$$\|Ax\| \le \|A\|\|x\|\,, \tag{1.28}$$

$$\|AB\| \le \|A\|\|B\|\,, \tag{1.29}$$

für alle $A, B \in \mathbb{K}^{(n,n)}$, $x \in \mathbb{K}^n$.

Beweis: Für $x \neq 0$ gilt

$$\|Ax\| = \left\| A\left(\frac{x}{\|x\|}\right)\right\| \|x\| \le \|A\|\|x\|\,,$$

$$\|ABx\| \le \|A\|\|Bx\| \le \|A\|\|B\|\|x\|\,,$$

also

$$\sup_{\|x\|=1} \|ABx\| \le \|A\|\|B\|\,.$$

□

Satz 1.6 *Sei*

$$f(z) = \sum_{k=0}^{\infty} c_k z^k$$

Potenzreihe in $\mathbb{C}$ *mit Konvergenzradius* r. *Dann ist die Matrixreihe*

$$f(A) = \sum_{k=0}^{\infty} c_k A^k \tag{1.30}$$

absolut konvergent für alle $A \in \mathbb{K}^{(n,n)}$ *mit* $\|A\| < r$.

Hierbei ist $\|\cdot\|$ eine beliebige Matrixnorm, das heißt, eine Matrixnorm, die von irgendeiner Vektornorm erzeugt wird.

Beweis: Gemäß Lemma 1.5 gilt

$$\|c_k A^k\| \le |c_k|\|A\|^k\,, \quad \text{für alle } k \in \mathbb{N}.$$

Ist $\|A\| < r$, so gilt

$$\sum_{k=0}^{\infty} \|c_k A^k\| \le \sum_{k=0}^{\infty} |c_k|\|A\|^k < \infty\,.$$

□

Die Matrixexponentialfunktion. Gemäß Satz 1.6 lässt sich

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} \tag{1.31}$$

für beliebige quadratische Matrizen $A \in \mathbb{K}^{(n,n)}$ definieren. Unmittelbar aus der Definition folgen

$$e^0 = I \tag{1.32}$$

und

$$e^{tA} = \sum_{k=0}^{\infty} t^k \frac{A^k}{k!}\,, \quad t \in \mathbb{R}\,. \tag{1.33}$$

Die Funktion $t \mapsto e^{tA}$ ist also auf ganz $\mathbb{R}$ in eine Potenzreihe um $t = 0$ entwickelbar.

Lemma 1.7 *Für $A \in \mathbb{K}^{(n,n)}$ gilt*

$$\frac{\mathrm{d}}{\mathrm{dt}} e^{tA} = Ae^{tA} = e^{tA} A \,. \tag{1.34}$$

Beweis: Aus (1.33) folgt

$$\frac{\mathrm{d}}{\mathrm{dt}} e^{tA} = \frac{\mathrm{d}}{\mathrm{dt}} \Big(\sum_{k=0}^{\infty} t^k \frac{A^k}{k!} \Big) = \sum_{k=1}^{\infty} t^{k-1} \frac{A^k}{(k-1)!} = A \sum_{k=0}^{\infty} t^k \frac{A^k}{k!} = \Big(\sum_{k=0}^{\infty} t^k \frac{A^k}{k!} \Big) A \,,$$

da Rechts- und Linksmultiplikation mit A eine stetige Operation ist und daher mit dem Grenzwert vertauscht werden kann. □

Lemma 1.8 *Seien $A, B \in \mathbb{K}^{(n,n)}$ mit $AB = BA$. Dann gilt*

$$e^{A+B} = e^A \cdot e^B \,. \tag{1.35}$$

Beweis: Wie im skalaren Fall (siehe Analysis 1) beweist man die Formel für das Cauchy-Produkt von Reihen und erhält

$$e^A \cdot e^B = \Big(\sum_{k=0}^{\infty} \frac{A^k}{k!} \Big) \cdot \Big(\sum_{j=0}^{\infty} \frac{B^j}{j!} \Big) = \sum_{m=0}^{\infty} \sum_{k=0}^{m} \frac{A^k}{k!} \frac{B^{m-k}}{(m-k)!} \,. \tag{1.36}$$

Wegen $AB = BA$ gilt die binomische Formel

$$\sum_{k=0}^{m} \frac{m!}{k!(m-k)!} A^k B^{m-k} = (A+B)^m \,. \tag{1.37}$$

Aus (1.36) und (1.37) folgt

$$e^A \cdot e^B = \sum_{m=0}^{\infty} \frac{(A+B)^m}{m!} = e^{A+B} \,.$$

□

Lösung des Anfangswertproblems. Wir betrachten das Anfangswertproblem

$$\dot{x} = Ax \,, \quad x(0) = x_0 \,. \tag{1.38}$$

Unter einer Lösung von (1.38) verstehen wir eine differenzierbare Funktion $x : \mathbb{R} \to \mathbb{K}^n$, welche $\dot{x}(t) = Ax(t)$ für alle $t \in \mathbb{R}$ sowie $x(0) = x_0$ erfüllt.

Satz 1.9 (Eindeutige Lösung des Anfangswertproblems)
Für jedes $A \in \mathbb{K}^{(n,n)}$ und $x_0 \in \mathbb{K}^n$ hat das Anfangswertproblem (1.38) die eindeutige Lösung

$$x(t) = e^{tA} x_0 \,. \tag{1.39}$$

Beweis: Dass die Funktion x eine Lösung ist, folgt unmittelbar aus (1.34) und aus

$$x(0) = e^0 x_0 = I x_0 = x_0 \,.$$

Zum Beweis der Eindeutigkeit betrachten wir eine beliebige Lösung $y : \mathbb{R} \to \mathbb{K}^{(n,n)}$. Mit Lemma 1.7, der Produktregel und der Kettenregel erhalten wir für alle $t \in \mathbb{R}$

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(e^{-tA} y(t)\big) = -e^{-tA} A y(t) + e^{-tA} \dot{y}(t) = e^{-tA}(\dot{y}(t) - A y(t)) = 0 \,.$$

Es folgt

$$e^{-tA} y(t) = c \in \mathbb{K}^n \,.$$

Wir setzen $t = 0$ und erhalten $c = y(0) = x_0$. Linksmultiplikation mit e^{tA} ergibt

$$e^{tA} x_0 = e^{tA} e^{-tA} y(t) = e^0 y(t) = y(t) \,,$$

wobei wir Lemma 1.8 verwendet haben. □

Satz 1.10 (Lösungsraum des linearen Systems)
Sei $A \in \mathbb{K}^{(n,n)}$. Die Menge L der Lösungen des linearen Systems $\dot{x} = Ax$ ist ein Vektorraum der Dimension n über $\mathbb{K}$. Die Spalten der Matrix e^{tA} bilden eine Basis von L.

Beweis: Wir definieren eine Abbildung $\Phi : \mathbb{K}^n \to \mathrm{Abb}\,(\mathbb{R}; \mathbb{K}^n)$ durch

$$(\Phi\eta)(t) = e^{tA} \eta \,.$$

Gemäß Superpositionsprinzip ist L ein Unterraum von $\mathrm{Abb}\,(\mathbb{R}; \mathbb{K}^n)$. Nach Satz 1.9 ist die Abbildung $\Phi : \mathbb{K}^n \to L$ bijektiv. Sie ist linear, und es gilt für Bild und Kern

$$\mathrm{Im}\,(\Phi) = L \,, \quad \ker(\Phi) = \{0\} \,.$$

Da sich die Dimensionen von Bild und Kern zur Dimension des Urbildraums addieren, folgt $\dim(L) = \dim(\mathbb{K}^n) = n$. Die Spalten von e^{tA} sind gerade die Bilder $\Phi(e_i)$ der Einheitsvektoren $e_i \in \mathbb{K}^n$. Da Φ bijektiv ist, bilden sie eine Basis von L. □

Form der Lösungen. Wir behandeln nun die Frage, wie e^{tA} für eine beliebige Matrix $A \in \mathbb{K}^{(n,n)}$ aussieht.

Lemma 1.11 *Sei $A = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)$. Dann gilt*

$$e^{tA} = \mathrm{diag}\,(e^{\lambda_1 t}, \ldots, e^{\lambda_n t}) \,. \tag{1.40}$$

Beweis: Einsetzen von $(tA)^k = \mathrm{diag}\,((\lambda_1 t)^k, \ldots, (\lambda_n t)^k)$ in die Reihenentwicklung von e^{tA}. □

Definition 1.12 *Eine Matrix $A \in \mathbb{K}^{(n,n)}$ heißt **nilpotent**, falls $A^n = 0$ gilt für ein $n \in \mathbb{N}$.*

Ist $A^n = 0$, so gilt

$$e^{tA} = \sum_{k=0}^{n-1} t^k \frac{A^k}{k!} \,. \tag{1.41}$$

Beispiel:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} , \quad A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \quad A^3 = 0 \,. \tag{1.42}$$

Es ist

$$e^{tA} = I + tA + t^2 \frac{A^2}{2} = \begin{pmatrix} 1 & t & \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \,. \tag{1.43}$$

Lemma 1.13 *Seien $A, V \in \mathbb{K}^{(n,n)}$, sei V invertierbar. Dann gilt*

$$e^{tV^{-1}AV} = V^{-1} e^{tA} V \,, \quad t \in \mathbb{R} \,. \tag{1.44}$$

Beweis: Es ist

$$\sum_{k=0}^{m} t^k \frac{(V^{-1}AV)^k}{k!} = \sum_{k=0}^{m} t^k \frac{V^{-1}A^k V}{k!} = V^{-1} \left(\sum_{k=0}^{m} t^k \frac{A^k}{k!} \right) V \,.$$

Grenzübergang $m \to \infty$ liefert die Behauptung. □

Wir führen die Bestimmung von e^{tA} für eine allgemeine Matrix $A \in \mathbb{K}^{(n,n)}$ durch Basiswechsel und Zerlegung auf die beiden schon behandelten Fälle (diagonal, nilpotent) zurück. In der Linearen Algebra wird zu einer beliebigen Matrix $A \in \mathbb{C}^{(n,n)}$ durch Basiswechsel deren **Jordansche Normalform** D konstruiert. D hat die Blockdiagonalform

$$D = \begin{pmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & 0 & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & D_k \end{pmatrix} , \tag{1.45}$$

wobei

$$D_j = \begin{pmatrix} \lambda_j & 1 & 0 & \cdots & 0 \\ 0 & \lambda_j & 1 & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \lambda_j & 1 \\ 0 & & & 0 & \lambda_j \end{pmatrix} \in \mathbb{C}^{n_j \times n_j} \,. \tag{1.46}$$

Hierbei sind $\lambda_j \in \mathbb{C}$ die Eigenwerte von A (diese müssen nicht voneinander verschieden sein), und es gilt

$$\sum_{j=1}^{k} n_j = n \,.$$

Wird der Basiswechsel durch die invertierbare Matrix $V \in \mathbb{C}^{(n,n)}$ beschrieben, so ist

$$A = V^{-1} D V \,. \tag{1.47}$$

Nach Lemma 1.13 gilt dann

$$e^{tA} = V^{-1} e^{tD} V \,. \tag{1.48}$$

Aus der Blockdiagonalform von D,

$$D = \operatorname{diag}(D_1, \dots, D_k)$$

folgt wie in Lemma 1.11

$$e^{tD} = \operatorname{diag}(e^{tD_1}, \dots, e^{tD_k}) \,. \tag{1.49}$$

Gemäß (1.46) ist

$$D_j = \lambda_j I + N_j \,, \quad N_j = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & 0 & 1 \\ 0 & & & 0 & 0 \end{pmatrix} \in \mathbb{C}^{n_j \times n_j} \,, \tag{1.50}$$

und daher

$$e^{tD_j} = e^{t(\lambda_j I + N_j)} = e^{\lambda_j t I} e^{tN_j} = e^{\lambda_j t} e^{tN_j} \,. \tag{1.51}$$

Es ist

$$N_j^\mu = (\delta_{i,k-\mu})_{1 \le i,k \le n_j} \,, \quad 0 \le \mu \le n_j - 1 \,, \tag{1.52}$$

(Kronecker-Delta), und $N_j^\mu = 0$ für $\mu \ge n_j$. (Induktionsbeweis: $\mu = 1$ klar, $\mu \to \mu + 1$:

$$(N_j^{\mu+1})_{il} = (N_j^\mu N_j)_{il} = \sum_k \delta_{i,k-\mu} \delta_{k,l-1} = \delta_{i,l-(\mu+1)} \,.$$

Es folgt

$$e^{tN_j} = \sum_{\mu=0}^{n_j-1} \frac{t^\mu}{\mu!} N_j^\mu = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{n_j-1}}{(n_j-1)!} \\ 0 & 1 & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \\ & & & \ddots & t \\ & & & 0 & 1 \end{pmatrix} \tag{1.53}$$

Zusammen mit den Formeln

$$e^{tA} = V^{-1} e^{tD} V \,, \quad e^{tD} = \operatorname{diag}(e^{tD_1}, \dots, e^{tD_k}) \,, \quad e^{tD_j} = e^{\lambda_j t} e^{tN_j} \,. \tag{1.54}$$

ist damit das Problem der Berechnung von e^{tA} zurückgeführt worden auf das Problem der Linearen Algebra, die Jordansche Normalform von A zu bestimmen. Dazu muß man die Eigenwerte λ_j von A und die zugehörigen Eigenvektoren bzw. verallgemeinerten Eigenvektoren (welche die Spalten von V bilden) berechnen. Für eine allgemeine Matrix A wollen wir das nicht weiter verfolgen (siehe etwa das Lehrbuch von Walter), sondern uns darauf beschränken, uns die Struktur der Lösungen klar zu machen. Aus (1.53) und (1.54) folgt, dass alle Lösungen von $\dot{x} = Ax$,

$$x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

aus Komponentenfunktionen x_i bestehen, welche Linearkombinationen von Funktionen der Form

$$t^l e^{\lambda t}$$

sind. Hierbei ist λ Eigenwert von A, und der mögliche Wertebereich der Exponenten l hängt von der Größe der zugehörigen Jordan-Kästchen D_j ab. Ist A diagonalisierbar, so treten keine Exponenten $l > 0$ auf, und die Lösungen haben die Form

$$x(t) = \sum_{\lambda \in \sigma(A)} v_\lambda e^{\lambda t}\,, \quad v_\lambda = \begin{pmatrix} v_{\lambda,1} \\ \vdots \\ v_{\lambda,n} \end{pmatrix} \in \mathbb{C}^n\,,$$

wobei $\sigma(A)$ die Menge der Eigenwerte von A bezeichnet und die Koeffizientenvektoren v_λ durch die Eigenvektoren von A und die Anfangsbedingung x_0 festgelegt sind.

Für $A \in \mathbb{R}^{(n,n)}$ und Anfangswerte $x_0 \in \mathbb{R}^n$ sind die Lösungen $x(t) = e^{tA}x_0$ reell. Die eben beschriebene Vorgehensweise liefert auch nichtreelle komplexe Lösungen, falls A nichtreelle Eigenwerte λ hat. Die nichtreellen Eigenwerte treten aber in komplex konjugierten Paaren $(\lambda, \overline{\lambda})$ auf; die entsprechenden Jordankästchen haben die gleiche Größe. Es ergeben sich Paare reeller Lösungen mit Zeitanteilen der Form $t^l e^{at}\cos(\omega t)$ und $t^l e^{at}\sin(\omega t)$. Wir stellen das nicht im Einzelnen dar.

2 Der Satz von Picard-Lindelöf

Wir betrachten ein System gewöhnlicher Differentialgleichungen erster Ordnung,

$$\dot{x} = f(t, x)\,, \tag{2.1}$$

sowie das zugehörige Anfangswertproblem

$$\dot{x} = f(t, x)\,, \quad x(t_0) = x_0\,. \tag{2.2}$$

Die Funktion f bezeichnet man als die **rechte Seite** der Differenzialgleichung, den Vektor x_0 als **Anfangswert** zum Zeitpunkt t_0.

Definition 2.1 (Lösung eines Systems erster Ordnung)
Sei $I \subset \mathbb{R}$ Intervall, $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : D \to \mathbb{R}^n$. Eine Funktion $x : I \to \mathbb{R}^n$ heißt Lösung von (2.1) in I, falls x in I differenzierbar ist und für alle $t \in I$ gilt, dass $(t, x(t)) \in D$ und

$$\dot{x}(t) = f(t, x(t)) \tag{2.3}$$

Falls außerdem $(t_0, x_0) \in D$ gegeben ist mit $t_0 \in I$ und $x(t_0) = x_0$, so heißt x Lösung des Anfangswertproblems (2.2) in I. □

Wir sind interessiert an reellwertigen Lösungen. Wir haben bereits gesehen, dass es im Falle linearer Systeme sinnvoll ist, sich auch mit komplexwertigen Lösungen zu befassen, auch wenn man nur an reellwertigen Lösungen interessiert ist. Wir verzichten aber im Folgenden darauf, $\mathbb{R}^n$ durch $\mathbb{K}^n$ zu ersetzen, obwohl das oft problemlos möglich ist.

Hängt f nicht von x ab, so ist

$$x(t) = x_0 + \int_{t_0}^{t} f(s)\,ds$$

eine Lösung des Anfangswertproblems (2.2), falls f stetig ist (Hauptsatz der Differential- und Integralrechnung).

Lemma 2.2 *Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : D \to \mathbb{R}^n$ stetig. Sei I Intervall in $\mathbb{R}$ und $x : I \to \mathbb{R}^n$ stetig, sei $t_0 \in I$ und $(t_0, x_0) \in D$. Dann ist x eine Lösung des Anfangswertproblems (2.2) auf I genau dann, wenn gilt $(t, x(t)) \in D$ für alle $t \in I$ und*

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\,ds\,, \quad \textit{für alle } t \in I. \tag{2.4}$$

Ist außerdem $f \in C^k(D)$ so ist $x \in C^{k+1}(I)$.

Beweis: Ist x Lösung des AWP, so ist $\dot{x} : I \to \mathbb{R}^n$ stetig, da f stetig ist. Aus dem Hauptsatz folgt

$$x(t) - x(t_0) = \int_{t_0}^{t} \dot{x}(s)\,ds = \int_{t_0}^{t} f(s, x(s))\,ds\,.$$

Gilt umgekehrt (2.4), so folgt zunächst $x(t_0) = x_0$. Da der Integrand auf der rechten Seite von (2.4) stetig ist als Funktion von s, ist nach dem Hauptsatz x differenzierbar in I, und $\dot{x}(t) = f(t, x(t))$ für alle $t \in I$. Aus (2.4), dem Hauptsatz und der Kettenregel folgt außerdem, dass $x \in C^{m+1}(I)$ falls $x \in C^m(I)$ und $f \in C^m(D)$ ist. Mit Induktion ergibt sich die letzte Behauptung des Satzes. $\square$

Der Satz von Picard-Lindelöf behandelt die eindeutige Lösbarkeit des Anfangswertproblems (2.2). Die Idee des Beweises ist es, den Banachschen Fixpunktsatz auf die Integralgleichung (2.4) anzuwenden.

Definition 2.3 (Lipschitz-Stetigkeit)
Sei $D \subset \mathbb{R} \times \mathbb{R}^n$. Eine Funktion $f : D \to \mathbb{R}^n$ heißt Lipschitz-stetig bezüglich x in D, falls es ein $L > 0$ gibt mit

$$\|f(t,x) - f(t,y)\|_\infty \leq L\|x - y\|_\infty \,, \tag{2.5}$$

für alle $(t,x), (t,y) \in D$. Die Zahl L heißt Lipschitz-Konstante von f in D.

f heißt lokal Lipschitz-stetig bezüglich x in D, falls es zu jedem $(t_0, x_0) \in D$ eine offene Menge $U \subset \mathbb{R} \times \mathbb{R}^n$ mit $(t_0, x_0) \in U$ gibt, so dass f in $D \cap U$ Lipschitz-stetig bezüglich x ist. $\square$

Bemerkung. Wir können die Lipschitz-Stetigkeit von f hinsichtlich einer beliebigen Norm im $\mathbb{R}^n$ definieren. Da im $\mathbb{R}^n$ alle Normen äquivalent sind, hängt die Eigenschaft von f, Lipschitz-stetig bzw. lokal Lipschitz-stetig zu sein, nicht von der Wahl der Norm ab (wohl aber die Größe der Lipschitz-Konstante).

Für $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$ und $r > 0$ definieren wir

$$K_r(t_0, x_0) = \{(t,x) : t \in \mathbb{R},\, x \in \mathbb{R}^n,\, |t - t_0| \leq r,\, \|x - x_0\|_\infty \leq r\} \,. \tag{2.6}$$

Satz 2.4 (Picard-Lindelöf, lokale Version) *Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, $(t_0, x_0) \in D$, sei $f : D \to \mathbb{R}^n$ stetig und bezüglich x lokal Lipschitz-stetig. Dann gibt es ein $\delta > 0$, so dass das AWP (2.2) in $I = [t_0 - \delta, t_0 + \delta]$ eine eindeutige Lösung hat.*
Genauer: Ist f Lipschitz-stetig in x auf $K_r(t_0, x_0) \subset D$ mit Lipschitz-Konstante L, und ist $|f(t,x)| \leq M$ auf $K_r(t_0, x_0)$, so kann

$$\delta = \min\left\{r, \frac{r}{2M}, \frac{1}{2L}\right\} \tag{2.7}$$

gewählt werden.

Beweis: Seien $r, L, M > 0$ wie beschrieben. (Solche Zahlen gibt es, da D offen, f auf D lokal Lipschitz-stetig in x und auf jeder abgeschlossenen Kugel beschränkt ist.) Sei δ definiert durch (2.7) und $I = [t_0 - \delta, t_0 + \delta]$. Wir definieren

$$X = \{x : \, x \in C(I; \mathbb{R}^n),\, \|x - x_0\|_\infty \leq r\} \,, \tag{2.8}$$

wobei x_0 als ein Element von $C(I; \mathbb{R}^n)$ (nämlich als eine konstante Funktion) aufgefasst wird und $\|\cdot\|_\infty$ die Maximumnorm in $C(I; \mathbb{R}^n)$ bedeutet; es ist dann

$$\|x - x_0\|_\infty = \max_{t \in I} \|x(t) - x_0\|_\infty \,. \tag{2.9}$$

Mit der durch die Maximumnorm induzierten Metrik ist X ein vollständiger metrischer Raum, da X abgeschlossene Teilmenge des Banachraums $C(I;\mathbb{R}^n)$ ist. Für $x \in X$ definieren wir eine Funktion $Tx : I \to \mathbb{R}^n$ durch

$$(Tx)(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds\,. \tag{2.10}$$

Es gilt $Tx \in C(I;\mathbb{R}^n)$, da f stetig ist. Es gilt $T(X) \subset X$, da

$$\|(Tx)(t) - x_0\|_\infty \le \int_{t_0}^{t} \|f(s, x(s))\|_\infty\, ds \le (t - t_0)M \le \delta M \le \frac{r}{2} \tag{2.11}$$

gilt für $t \in I$ und $x \in X$. Wir zeigen, dass T eine Kontraktion auf X ist. Seien $x, y \in X$. Dann gilt für alle $t \in I$

$$\begin{aligned}
\|(Tx)(t) - (Ty)(t)\|_\infty &= \left\| \int_{t_0}^{t} f(s, x(s)) - f(s, y(s))\, ds \right\|_\infty \\
&\le \int_{t_0}^{t} \|f(s, x(s)) - f(s, y(s))\|_\infty\, ds \le \int_{t_0}^{t} L\|x(s) - y(s)\|_\infty\, ds \\
&\le L\delta\|x - y\|_\infty \le \frac{1}{2}\|x - y\|_\infty\,,
\end{aligned}$$

also

$$\|Tx - Ty\|_\infty \le \frac{1}{2}\|x - y\|_\infty\,.$$

Also ist T eine Kontraktion auf X. Aus dem Fixpunktsatz von Banach folgt, dass die Fixpunktgleichung

$$x = Tx \tag{2.12}$$

genau eine Lösung $x \in X$ hat. Nach Lemma 2.2 ist x eine Lösung des AWP (2.2). Es bleibt zu zeigen, dass jede Lösung auf I des AWP in X liegen muss. Sei $y \in C(I;\mathbb{R}^n)$ Lösung des AWP auf I mit $y \notin X$. Dann gibt es ein $t \in I$ mit $\|y(t) - x_0\|_\infty > r$; sei $t > t_0$. (Der Fall $t < t_0$ wird analog behandelt.) Wir setzen

$$t^* = \inf\{t : t \in [t_0, t_0 + \delta],\ \|y(t) - x_0\|_\infty > r\}\,.$$

Nach Definition von t^* ist $\|y(s) - x_0\|_\infty \le r$ für alle $s \in [t_0, t^*)$, also auch $\|f(s, y(s))\|_\infty \le M$ für alle $s \in [t_0, t^*)$. Wie in (2.11) folgt nun

$$\|y(t^*) - x_0\|_\infty = \left\| \int_{t_0}^{t^*} f(s, y(s))\, ds \right\|_\infty \le \frac{r}{2}\,.$$

Andererseits folgt aus der Definition von t^*, dass $\|y(t^*) - x_0\|_\infty \ge r$, ein Widerspruch. Also kann es ein solches y nicht geben. $\square$

Lösungen von Differentialgleichungen lassen sich zusammensetzen. Sind $x^1 : [t_0, t_1] \to \mathbb{R}^n$ und $x^2 : [t_1, t_2] \to \mathbb{R}^n$ Lösungen von $\dot{x} = f(t, x)$ mit $x^1(t_1) = x^2(t_1)$, so ist die durch

$$x : [t_0, t_2] \to \mathbb{R}^n\,, \quad x(t) = \begin{cases} x^1(t)\,, & t \in [t_0, t_1]\,, \\ x^2(t)\,, & t \in [t_1, t_2] \end{cases} \tag{2.13}$$

definierte Funktion eine Lösung von $\dot{x} = f(t,x)$ auf $[t_0, t_2]$, da sie die Integralgleichung

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds$$

für alle $t \in [t_0, t_2]$ erfüllt.

Hat man mit dem Satz von Picard-Lindelöf eine eindeutige Lösung x auf einem Intervall $[t_0 - \delta, t_0 + \delta]$ erhalten, so können wir $x(t_0 + \delta)$ als neuen Anfangswert zur Zeit $t_0 + \delta$ betrachten. Nochmaliges Anwenden des Satzes liefert eine weitere Lösung auf einem Intervall $[t_0 + \delta, t_0 + \delta + \tilde{\delta}]$, die wir zu einer Lösung auf $[t_0, t_0 + \delta + \tilde{\delta}]$ zusammensetzen können. Es stellt sich die Frage, wie weit sich eine solche Lösung fortsetzen lässt. Hierzu einige Beispiele. Das AWP

$$\dot{x} = 1 - x^2\,, \quad x(0) = 0\,, \tag{2.14}$$

hat die auf $I = \mathbb{R}$ definierte Lösung

$$x(t) = \tanh t = \frac{\sinh t}{\cosh t} = \frac{e^t - e^{-t}}{e^t + e^{-t}}\,.$$

Das AWP

$$\dot{x} = x^2\,, \quad x(0) = 1\,, \tag{2.15}$$

hat die auf $I = (-\infty, 1)$ definierte Lösung

$$x(t) = \frac{1}{1-t}\,.$$

Das AWP

$$\dot{x} = 1 + x^2\,, \quad x(0) = 0\,, \tag{2.16}$$

hat die auf $I = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ definierte Lösung

$$x(t) = \tan t\,.$$

Lemma 2.5 (Globale Eindeutigkeit)
Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : D \to \mathbb{R}^n$ stetig und bezüglich x lokal Lipschitz-stetig. Seien $x^1, x^2 : J \to \mathbb{R}^n$ Lösungen des AWP (2.2) auf einem Intervall $J = (a,b)$. Dann gilt $x^1 = x^2$ auf J.

Beweis: Wir setzen

$$t^* = \sup\{t : t \in J,\, x^1|[t_0, t] = x^2|[t_0, t]\}\,.$$

Die rechts stehende Menge ist nicht leer, da sie t_0 enthält. Also ist $t^* \geq t_0$. Wir nehmen an, dass $t^* < b$. Da x^1 und x^2 stetig sind, gilt $x^1(t^*) = x^2(t^*) =: x^*$. Für jedes $\delta < b - t^*$ hat das AWP

$$\dot{x} = f(t,x)\,, \quad x(t^*) = x^*\,,$$

in $[t^*, t^* + \delta]$ zwei verschiedene Lösungen, nämlich x^1 und x^2, ein Widerspruch zum lokalen Existenz- und Eindeutigkeitssatz (Satz 2.4). Also ist $t^* = b$. Analog zeigt man, dass x^1 und x^2 auf $(a, t_0]$ übereinstimmen. $\square$

Satz 2.6 (Maximales Existenzintervall)
Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : D \to \mathbb{R}^n$ stetig und bezüglich x lokal Lipschitz-stetig. Dann gibt es zu jedem $(t_0, x_0) \in D$ ein offenes Intervall I und eine Lösung x des AWP (2.2) in I, so dass für jede Lösung y von (2.2) in einem Intervall J gilt

$$J \subset I \,, \quad y = x|J \,. \tag{2.17}$$

Das Intervall I heißt das maximale Existenzintervall der Lösung von (2.2).

Beweis: Wir definieren

$$t^* = \sup\{t : \, t \geq t_0, \text{ (2.2) hat eine Lösung in } [t_0, t]\} \,, \tag{2.18}$$
$$t_* = \inf\{t : \, t \leq t_0, \text{ (2.2) hat eine Lösung in } [t, t_0]\} \,. \tag{2.19}$$

Wir setzen $I = (t_*, t^*)$. Für jedes $t \in I$ mit $t \geq t_0$ wählen wir eine Lösung y^t von (2.2) auf $[t_0, t]$. Ist $t_0 \leq s \leq t$, so gilt $y^s = y^t$ auf $[t_0, s]$ nach Lemma 2.5. Wir definieren $x : [t_0, t^*) \to \mathbb{R}^n$ durch $x(t) = y^t(t)$. Es folgt

$$x(t) - x_0 = y^t(t) - x_0 = \int_{t_0}^t f(s, y^t(s)) \, ds = \int_{t_0}^t f(s, y^s(s)) \, ds = \int_{t_0}^t f(s, x(s)) \, ds \,.$$

Also ist x eine Lösung von (2.2) auf $[t_0, t^*)$. Jede andere Lösung auf einem Intervall $J \subset [t_0, t^*)$ stimmt wegen Lemma 2.5 mit x überein. Analog verfährt man für t_*. Nach Definition von t^* und t_* kann es auf keinem größeren Intervall eine Lösung geben. $\square$

Das maximale Existenzintervall kann die Form $(-\infty, b)$, (a, b), (a, ∞) oder $(-\infty, \infty) = \mathbb{R}$ haben.

Es treten drei verschiedene Fälle für das Verhalten der Lösung für $t > t_0$ (analog $t < t_0$).

1. Die Lösung existiert für alle $t > t_0$.
2. Es gibt ein $b > t_0$ mit
$$\lim_{\substack{t \to b \\ t < b}} \|x(t)\|_\infty = \infty \,.$$
3. Es gibt ein $b > t_0$ mit
$$\lim_{\substack{t \to b \\ t < b}} \operatorname{dist}\left((t, x(t)), \partial D\right) = 0 \,.$$

Der erste Fall liegt etwa beim linearen System $\dot{x} = Ax$ und bei Beispiel (2.14) vor. In den Beispielen (2.15) und (2.16) tritt der zweite Fall auf. Ein Beispiel für den dritten Fall ist das AWP

$$\dot{x} = \frac{1}{x} \,, \quad x(0) = 1 \,. \tag{2.20}$$

Hier ist $D = \{(t, x) : x \neq 0\} = \mathbb{R} \times (\mathbb{R} \setminus \{0\})$. Die Lösung ist

$$x(t) = \sqrt{2t + 1} \,, \quad \text{und } I = \left(-\frac{1}{2}, \infty\right)$$

ist das maximale Existenzintervall. Es gilt hier

$$\lim_{t \downarrow -\frac{1}{2}} (t, x(t)) = \left(-\frac{1}{2}, 0\right) \in \partial D \,.$$

Man kann die Fortsetzbarkeit bis zum Rand von D in folgendem Satz formalisieren. (Das beinhaltet auch den zweiten Fall.)

Satz 2.7 *Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : D \to \mathbb{R}^n$ stetig und bezüglich x lokal Lipschitz-stetig. Sei x die eindeutige Lösung des AWP (2.2) mit maximalem Existenzintervall I. Dann gilt: Weder $\{(t, x(t)) : t \in I,\, t \geq t_0\}$ noch $\{(t, x(t)) : t \in I,\, t \leq t_0\}$ sind in einer kompakten Teilmenge von D enthalten.* □

Beweis: Sei $I = (a, b)$. Sei $K \subset D$ kompakt mit

$$\{(t, x(t)) : t \in I,\, t \geq t_0\} \subset K \,. \tag{2.21}$$

Da K beschränkt ist, ist $b < \infty$. Da f stetig ist, gilt

$$M := \max_{(t,\eta) \in K} \|f(t, \eta)\|_\infty < \infty \,.$$

Für beliebige $t, \tau \in [t_0, b)$ folgt

$$\|x(t) - x(\tau)\|_\infty = \left\| \int_\tau^t f(s, x(s))\, ds \right\|_\infty \leq M |t - \tau| \,. \tag{2.22}$$

Wir zeigen nun, dass

$$x_b := \lim_{\substack{t \to b \\ t < b}} x(t) \tag{2.23}$$

existiert. Ist $(t_k)_{k \in \mathbb{N}}$ Folge mit $t_k \to b$, so ist $(x(t_k))_{k \in \mathbb{N}}$ Cauchyfolge wegen (2.22), also konvergent; ist $(s_k)_{k \in \mathbb{N}}$ eine weitere solche Folge, so konvergiert $(x(s_k))_{k \in \mathbb{N}}$ gegen den gleichen Grenzwert, da $x(t_k) - x(s_k) \to 0$ wegen (2.22). Also existiert der Grenzwert (2.23). Die Funktion

$$\tilde{x}(t) = \begin{cases} x(t) \,, & t \in [t_0, b) \,, \\ x_b \,, & t = b \,, \end{cases}$$

ist daher eine Lösung der Integralgleichung

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s))\, ds$$

auf dem abgeschlossenen Intervall $[t_0, b]$ und damit auch des AWP, im Widerspruch zur Maximalität von I. Eine kompakte Menge K mit (2.21) kann daher nicht existieren. Analog zeigt man die Behauptung für $t \leq t_0$. □

Man kann die Ergebnisse der vorherigen Sätze wie folgt zusammenfassen:

> Ist die rechte Seite f auf ihrem offenen Definitionsgebiet D stetig und bezüglich x lokal Lipschitz-stetig, so existiert eine eindeutige Lösung des AWP (2.2), deren Graph sich bis zum Rand von D erstreckt.

3 Das Lemma von Gronwall

Sei $z : I \to \mathbb{R}$, $I \subset \mathbb{R}$ Intervall, eine unbekannte Funktion. Wir wollen z nach oben abschätzen. Wir nehmen an, dass z in I einer Differentialungleichung

$$\dot{z}(t) \leq g(t, z(t))\,, \quad z(t_0) \leq w_0\,, \tag{3.1}$$

oder einer Integralungleichung

$$z(t) \leq w_0 + \int_{t_0}^{t} g(s, z(s))\, ds \tag{3.2}$$

genügt. Wir wollen schließen, dass dann

$$z(t) \leq w(t) \tag{3.3}$$

für alle $t \in I$ gilt, wobei $w : I \to \mathbb{R}$ die Lösung des entsprechenden AWP's

$$\dot{w} = g(t, w)\,, \quad w(t_0) = w_0\,, \tag{3.4}$$

beziehungsweise der entsprechenden Integralgleichung

$$w(t) = w_0 + \int_{t_0}^{t} g(s, w(s))\, ds \tag{3.5}$$

ist. Die Ungleichung (3.3) liefert eine explizite Abschätzung für z, wenn wir w explizit berechnen können.

Wir beschäftigen uns mit der Integralungleichung.

Lemma 3.1 *Sei $I = [t_0, t_1]$, seien $a, b, z, w : I \to \mathbb{R}$ stetig, sei $b \geq 0$, und es gelte*

$$z(t) < a(t) + \int_{t_0}^{t} b(s) z(s)\, ds\,, \tag{3.6}$$

$$w(t) = a(t) + \int_{t_0}^{t} b(s) w(s)\, ds\,, \tag{3.7}$$

für alle $t \in I$. Dann gilt

$$z(t) < w(t) \tag{3.8}$$

für alle $t \in I$.

Beweis: Wir setzen (beachte $z(t_0) < w(t_0)$)

$$t = \sup\{\tau : \tau \in I,\ z(s) < w(s) \text{ für alle } s \leq \tau\}\,.$$

Dann gelten $t > t_0$, $z < w$ in $[t_0, t)$ und $z(t) \leq w(t)$. Da $b \geq 0$, folgt

$$z(t) < a(t) + \int_{t_0}^{t} b(s) z(s)\, ds \leq a(t) + \int_{t_0}^{t} b(s) w(s)\, ds = w(t)\,.$$

Es muss nun $t = t_1$ gelten, da andernfalls $z(\tau) < w(\tau)$ in $[t, t + \varepsilon]$ gilt für hinreichend kleines $\varepsilon > 0$, im Widerspruch zur Definition von t. □

Lemma 3.2 *In der Situation von Lemma 3.1 ist eine Lösung $w : I \to \mathbb{R}$ von (3.7) gegeben durch*

$$w(t) = a(t) + \int_{t_0}^{t} a(s)b(s) \exp\left(\int_s^t b(\tau)\,d\tau\right) ds\,. \tag{3.9}$$

Beweis: Sei $w : I \to \mathbb{R}$ definiert durch (3.9). Es gilt

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{dt}} \int_{t_0}^{t} a(s)b(s) \exp\left(\int_s^t b(\tau)\,d\tau\right) ds \\ &\qquad = a(t)b(t) + b(t)\int_{t_0}^{t} a(s)b(s) \exp\left(\int_s^t b(\tau)\,d\tau\right) ds = b(t)w(t)\,.\end{aligned}$$

Integrieren von t_0 bis t ergibt

$$\int_{t_0}^{t} a(s)b(s) \exp\left(\int_s^t b(\tau)\,d\tau\right) ds = \int_{t_0}^{t} b(s)w(s)\,ds\,.$$

Also ist w eine Lösung von (3.7). □

Satz 3.3 (Lemma von Gronwall)
Sei $I = [t_0, t_1]$, seien $a, b, z : I \to \mathbb{R}$ stetig, sei $b \geq 0$, es gelte

$$z(t) \leq a(t) + \int_{t_0}^{t} b(s)z(s)\,ds \tag{3.10}$$

für alle $t \in I$. Dann gilt

$$z(t) \leq a(t) + \int_{t_0}^{t} a(s)b(s) \exp\left(\int_s^t b(\tau)\,d\tau\right) ds\,. \tag{3.11}$$

Beweis: Sei $\varepsilon > 0$ beliebig, sei $\tilde{a}(t) = a(t) + \varepsilon$. Dann gilt für alle $t \in I$

$$z(t) < \tilde{a}(t) + \int_{t_0}^{t} b(s)z(s)\,ds\,. \tag{3.12}$$

Gemäß Lemma 3.2 ist die Funktion $\tilde{w} : I \to \mathbb{R}$,

$$\tilde{w}(t) = \tilde{a}(t) + \int_{t_0}^{t} \tilde{a}(s)b(s) \exp\left(\int_s^t b(\tau)\,d\tau\right) ds\,,$$

auf I eine Lösung von

$$\tilde{w}(t) = \tilde{a}(t) + \int_{t_0}^{t} b(s)\tilde{w}(s)\,ds\,. \tag{3.13}$$

Aus (3.12), (3.13) folgt nun wegen Lemma 3.1

$$z(t) < \tilde{w}(t) = a(t) + \varepsilon + \int_{t_0}^{t} (a(s) + \varepsilon)b(s) \exp\left(\int_s^t b(\tau)\,d\tau\right) ds\,.$$

Grenzübergang $\varepsilon \to 0$ liefert die Behauptung. □

Das Lemma von Gronwall wird häufig in der folgenden, einfacher zu merkenden Version verwendet. Diese wird allgemein ebenfalls als “Lemma von Gronwall” bezeichnet.

Folgerung 3.4 *Sei $I = [t_0, t_1]$, seien $b, z : I \to \mathbb{R}$ stetig mit $b \geq 0$, sei $a \in \mathbb{R}$, es gelte*

$$z(t) \leq a + \int_{t_0}^{t} b(s) z(s) \, ds \,. \tag{3.14}$$

Dann gilt

$$z(t) \leq a \exp\left(\int_{t_0}^{t} b(s) \, ds \right) \,. \tag{3.15}$$

Ist b konstant, so wird (3.16) zu

$$z(t) \leq a e^{(t - t_0) b} \,. \tag{3.16}$$

Beweis: Wir setzen

$$B(t) = \int_{t_0}^{t} b(s) \, ds \,.$$

Aus Satz 3.3 folgt

$$\begin{aligned} z(t) &\leq a + \int_{t_0}^{t} a b(s) e^{B(t) - B(s)} \, ds = a \left(1 + e^{B(t)} \int_{t_0}^{t} b(s) e^{-B(s)} \, ds \right) \\ &= a \left(1 + e^{B(t)} \Big[- e^{-B(s)} \Big]_{t_0}^{t} \right) = a(1 - e^{B(t)} e^{-B(t)} + e^{B(t)} e^{B(t_0)}) = a e^{B(t)} \,. \end{aligned}$$

$\square$

4 Lineare zeitvariante Systeme

Wir betrachten das lineare System

$$\dot{x} = A(t)x\,, \tag{4.1}$$

wobei $A(t) \in \mathbb{R}^{(n,n)}$ eine quadratische Matrix ist, deren Elemente Funktionen von t sind. Ist A nicht konstant, so heißt das System **nichtautonom** oder **zeitvariant**.

Der eindimensionale Fall. Wir betrachten das Anfangswertproblem

$$\dot{x} = a(t)x\,, \quad x(t_0) = x_0\,. \tag{4.2}$$

Sei I Intervall im $\mathbb{R}$ mit $t_0 \in I$, und sei $a : I \to \mathbb{R}$ stetig. Die Funktion $x : I \to \mathbb{R}$, definiert durch

$$x(t) = \exp\Big(\int_{t_0}^{t} a(s)\,ds\Big) \cdot x_0\,, \tag{4.3}$$

ist Lösung von (4.2) auf I für jedes $x_0 \in \mathbb{R}$; es gilt nach Kettenregel

$$\dot{x}(t) = \exp\Big(\int_{t_0}^{t} a(s)\,ds\Big)\, a(t)x_0 = a(t)x(t)\,.$$

Hier ein Beispiel:

$$\dot{x} = 2tx\,, \quad x(0) = 2\,. \tag{4.4}$$

Es ist $a(t) = 2t$, $t_0 = 0$, $x_0 = 2$, $I = \mathbb{R}$, $\int_0^t a(s)\,ds = t^2$. Lösung von (4.4) auf $\mathbb{R}$ ist also

$$x(t) = 2e^{t^2}\,.$$

Die durch (4.3) gegebene Lösung ist die einzige Lösung von (4.2). Das kann man direkt nachrechnen, ähnlich wie im einführenden Beispiel von Kapitel 1. Man kann auch den Satz von Picard-Lindelöf heranziehen. Die rechte Seite $f(t,x) = a(t)x$ von (4.2) ist auf $I \times \mathbb{R}$ sowohl stetig als auch lokal Lipschitz-stetig in x: Zunächst gilt

$$|f(t,x) - f(t,y)| = |a(t)||x-y|\,.$$

Zu beliebigem $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}$ wählen wir $J = (t_0 - \varepsilon, t_0 + \varepsilon)$ mit beliebigem $\varepsilon > 0$. Der Abschluss $\overline{J} = [t_0 - \varepsilon, t_0 + \varepsilon]$ ist kompakt, und

$$L = \max_{t \in \overline{J}} |a(t)|$$

ist eine Lipschitz-Konstante bezüglich x für f auf der offenen Menge $J \times \mathbb{R}$, welche (t_0, x_0) enthält.

Der mehrdimensionale Fall. Wir betrachten das inhomogene ($b \neq 0$) Anfangswertproblem

$$\dot{x} = A(t)x + b(t)\,, \quad x(t_0) = x_0\,. \tag{4.5}$$

Satz 4.1 *Sei I offenes Intervall*[1] *im $\mathbb{R}$, seien $A : I \to \mathbb{R}^{(n,n)}$ und $b : I \to \mathbb{R}^n$ stetig, seien $t_0 \in I$ und $x_0 \in \mathbb{R}^n$. Dann hat das Anfangswertproblem (4.5) eine eindeutige Lösung $x : I \to \mathbb{R}^n$.*

Beweis: Die durch

$$f(t,x) = A(t)x + b(t)$$

definierte rechte Seite $f : D = I \times \mathbb{R}^n \to \mathbb{R}^n$ ist stetig, und es gilt

$$\|f(t,x) - f(t,y)\|_\infty \leq \|A(t)x + b(t) - A(t)y - b(t)\|_\infty \leq \|A(t)\| \|x-y\|_\infty$$

für alle $(t,x), (t,y) \in D$. Wie im eindimensionalen Fall folgt, dass f auf D lokal lipschitzstetig ist. Nach dem Satz von Picard-Lindelöf und nach Satz 2.6 hat das AWP eine eindeutige Lösung $x : J \to \mathbb{R}^n$ auf einem maximalen offenen Existenzintervall $J = (t_*, t^*) \subset I$. Es bleibt zu zeigen, dass $J = I$. Dazu genügt es zu zeigen, dass

$$J \cap [t_0, +\infty) = I \cap [t_0, +\infty) \tag{4.6}$$

und

$$J \cap (-\infty, t_0] = I \cap (-\infty, t_0]\,.$$

Wir zeigen (4.6), die zweite Behauptung folgt analog. Ist $t^* = +\infty$, so gilt (4.6). Ist $t^* < +\infty$ und t^* der rechte Randpunkt von I, so gilt (4.6) ebenfalls. Es bleibt der Fall $t^* \in I$. Wir definieren

$$z(t) = \|x(t)\|_\infty\,, \quad z : [t_0, t^*) \to \mathbb{R}\,.$$

Auf dem halboffenen Intervall $[t_0, t^*)$ gilt

$$z(t) = \left\| x_0 + \int_{t_0}^t A(s)x(s) + b(s)\,ds \right\| \leq \|x_0\|_\infty + \int_{t_0}^t \|A(s)\| \|x(s)\|_\infty + \|b(s)\|_\infty\,ds\,. \tag{4.7}$$

Wir setzen

$$a(t) = \|x_0\|_\infty + \int_{t_0}^t \|b(s)\|_\infty\,ds\,, \quad \beta(t) = \|A(t)\|\,.$$

Aus (4.7) folgt nun

$$z(t) \leq a(t) + \int_{t_0}^t \beta(s)z(s)\,ds\,. \tag{4.8}$$

Wir wenden nun das Lemma von Gronwall (Satz 3.3) an und erhalten

$$z(t) \leq w(t) \qquad \text{auf } [t_0, t^*), \tag{4.9}$$

wobei $w : I \to \mathbb{R}$ die in (3.9) gegebene Funktion ist (mit β statt b). Deren genaue Form ist hier irrelevant; sie ist aber auf I und damit auch auf dem kompakten Intervall $[t_0, t^*]$ stetig, also auf $[t_0, t^*]$ beschränkt. Es folgt, dass auch z und x auf $[t_0, t^*)$ beschränkt sind. Sei

$$M = \sup_{t \in [t_0, t^*)} \|x(t)\|_\infty\,.$$

Es folgt weiter, dass die Menge $\{(t, x(t)) : t_0 \leq t < t^*\}$ in der kompakten Teilmenge

$$[t_0, t^*] \times \{\eta : \eta \in \mathbb{R}^n\,,\ \|\eta\|_\infty \leq M\}$$

[1] Wie bisher sind die Intervallgrenzen $-\infty$ und $+\infty$ auch zugelassen.

von D enthalten ist. Das ist aber gemäß Satz 2.7 ein Widerspruch zur Maximalität von J. □

Lösungsraum und Fundamentalmatrix. Wir betrachten zunächst das homogene Anfangswertproblem

$$\dot{x} = A(t)x\,, \quad x(t_0) = x_0\,. \tag{4.10}$$

Analog zu Satz 1.10 erhalten wir das folgende Ergebnis.

Satz 4.2 (Lösungsraum des homogenen Problems)
Sei I ein Intervall, Sei $A : I \to \mathbb{R}^{(n,n)}$ stetig. Die Menge L der Lösungen des zeitvarianten linearen homogenen Systems $\dot{x} = A(t)x$ auf I ist ein Vektorraum der Dimension n über $\mathbb{R}$.

Beweis: Sei $t_0 \in I$ gegeben. Gemäß Satz 4.1 können wir jedem Anfangswert $x_0 \in \mathbb{R}^n$ die eindeutige Lösung $x \in C(I; \mathbb{R}^n)$ des Anfangswertproblems (4.10) zuordnen. Diese Abbildung ist linear, da aufgrund des Superpositionsprinzips die Lösung zu einer Linearkombination von Anfangswerten sich ergibt als die entsprechende Linearkombination der Lösungen zu den einzelnen Anfangswerten. Sie ist injektiv, da die Nullfunktion die einzige Lösung zum Anfangswert 0 ist. Es folgt, dass L ein Unterraum von $C(I; \mathbb{R}^n)$ der Dimension n ist. □

Sei $t_0 \in I$ gegeben. Aus der Abbildung

$$x_0 \mapsto x\,, \quad \mathbb{R}^n \to C(I; \mathbb{R}^n)\,,$$

welche jedem Anfangswert die Lösung auf I zuordnet, können wir weitere Abbildungen erhalten, Sei außerdem $t \in I$ gegeben. Die zusammengesetzte Abbildung

$$x_0 \mapsto x \mapsto x(t)\,, \quad \mathbb{R}^n \to C(I; \mathbb{R}^n) \to \mathbb{R}^n$$

ist ebenfalls linear. Sie wird daher durch eine Matrix der Dimension $n \times n$ repräsentiert. Diese hängt von t_0 und t ab, wird hier mit

$$\Phi(t, t_0) \tag{4.11}$$

bezeichnet und heißt die **Fundamentalmatrix** des linearen zeitvarianten Systems (4.10). Hängt A nicht von t ab, so gilt, wie wir bereits wissen,

$$\Phi(t, t_0) = e^{(t-t_0)A}\,. \tag{4.12}$$

Allgemein gilt

$$\Phi(t_0, t_0) = I\,, \quad \text{für alle } t_0 \in I. \tag{4.13}$$

(Das erste I ist die Einheitsmatrix, das zweite I das Zeitintervall.) Da sich zwei Lösungen des Anfangswertproblems zu einer weiteren zusammensetzen lassen, gilt

$$\Phi(s, t)\Phi(t, t_0) = \Phi(s, t_0)\,, \quad \text{für alle } t, s, t_0 \text{ in } I. \tag{4.14}$$

Daraus folgt für $s = t_0$, dass

$$\Phi(t_0, t)\Phi(t, t_0) = \Phi(t_0, t_0) = I\,. \tag{4.15}$$

Daher sind alle Matrizen $\Phi(t, t_0)$ invertierbar, und

$$\Phi(t, t_0)^{-1} = \Phi(t_0, t)\,. \tag{4.16}$$

Sei nun x die Lösung des AWP $\dot{x} = A(t)x$, $x(t_0) = e_j$, dem j-ten Einheitsvektor im $\mathbb{R}^n$. Es ist

$$x(t) = \Phi(t, t_0)e_j$$

die j-te Spalte von $\Phi(t, t_0)$. Letztere ist daher nach t differenzierbar,

$$\frac{\mathrm{d}}{\mathrm{dt}}\Phi(t, t_0)e_j = \dot{x}(t) = A(t)x(t) = A(t)\Phi(t, t_0)e_j\,. \tag{4.17}$$

Faßt man diese Vektoren zu einer Matrix zusammen, so erhält man

$$\frac{\mathrm{d}}{\mathrm{dt}}\Phi(t, t_0) = A(t)\Phi(t, t_0)\,. \tag{4.18}$$

Die Funktion $t \mapsto \Phi(t, t_0)$ ist also eine Lösung des Anfangswertproblems für eine matrixwertige Funktion $\Psi : I \to \mathbb{R}^{(n,n)}$,

$$\dot{\Psi} = A(t)\Psi\,, \quad \Psi(t_0) = I\,. \tag{4.19}$$

Mit Hilfe der Fundamentalmatrix können wir die Lösung des inhomogenen Anfangswertproblems

$$\dot{x} = A(t)x + b(t)\,, \quad x(t_0) = x_0\,, \tag{4.20}$$

angeben. Sie hat die Form

$$x(t) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, s)b(s)\,ds\,, \tag{4.21}$$

wie man durch Differenzieren der rechten Seite unter Verwendung von (4.18) nachprüfen kann. Das Integral auf der rechten Seite von (4.21) repräsentiert eine spezielle Lösung des inhomogenen Systems

$$\dot{x} = A(t)x + b(t)\,,$$

nämlich die zum Anfangswert $x(t_0) = 0$. Der Term $\Phi(t, t_0)x_0$ steht für die allgemeine Lösung des homogenen Systems (4.5), nämlich einer beliebigen (durch die Wahl von x_0 festgelegten) Linearkombination der Spalten der Fundamentalmatrix.

Falls A nicht von t abhängt, wird (4.21) wegen (4.12) zu

$$x(t) = e^{(t-t_0)A}x_0 + \int_{t_0}^{t} e^{(t-s)A}b(s)\,ds\,. \tag{4.22}$$

5 Abhängigkeit der Lösung von den Anfangswerten

A-priori-Abschätzung. Wir wollen die Lösung des Anfangswertproblems in Abhängigkeit von f und (t_0, x_0) abschätzen. Dafür lässt sich das Lemma von Gronwall verwenden.

Satz 5.1 *Sei $I = [t_0, t_1]$, $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ stetig, sei $x : I \to \mathbb{R}^n$ Lösung des AWP*

$$\dot{x} = f(t, x)\,, \quad x(t_0) = x_0\,, \tag{5.1}$$

es gebe $c, L \geq 0$, so dass

$$\|f(t, \eta)\|_\infty \leq c + L\|\eta\|_\infty\,, \quad \textit{für alle } t \in I,\ \eta \in \mathbb{R}^n. \tag{5.2}$$

Dann gilt

$$\|x(t)\|_\infty \leq (\|x_0\|_\infty + c(t - t_0))e^{L(t-t_0)}\,, \quad \textit{für alle } t \in I. \tag{5.3}$$

Beweis: Es gilt

$$\|x(t)\|_\infty = \left\|x_0 + \int_{t_0}^t f(s, x(s))\, ds\right\|_\infty \leq \|x_0\|_\infty + c(t - t_0) + \int_{t_0}^t L\|x(s)\|_\infty\, ds\,.$$

Auf die durch $z(t) = \|x(t)\|_\infty$ definierte Funktion wenden wir für jedes $t \in I$ das Lemma von Gronwall an in Form von Folgerung 3.4 mit $t_1 = t$, $a = \|x_0\|_\infty + c(t - t_0)$ und $b = L$. □

Ein Beispiel: Das AWP $\dot{x} = 3x$, $x(0) = 1$ hat die Lösung

$$x(t) = e^{3t}\,.$$

Satz 5.1 mit $x_0 = 1$ und $L = 3$ liefert die Abschätzung

$$|x(t)| \leq e^{3t}\,. \tag{5.4}$$

Ändern wir das AWP zu $\dot{x} = -3x$, $x(0) = 1$, so ist die Lösung

$$x(t) = e^{-3t}\,.$$

Satz 5.1 liefert aber ebenfalls die Abschätzung (5.4). Im ersten Fall ist die Abschätzung gut (sogar bestmöglich), im zweiten Fall umso schlechter, je größer t ist. Das qualitative Verhalten (Konvergenz der Lösung gegen 0 für $t \to \infty$) wird von ihr überhaupt nicht erfasst.

Eine Abschätzung der Form (5.3), in der die Lösung einer (Differenzial-)Gleichung abgeschätzt wird durch einen Ausdruck, in dem nur im AWP gegebene Größen (hier (t_0, x_0), c, L) auftauchen, heißt **a-priori-Abschätzung** der Lösung. In Verbindung mit Kompaktheitsargumenten liefern solche Abschätzungen eine wesentliche Methode, um die Existenz einer Lösung in Situationen zu beweisen, in denen der Banachsche Fixpunktsatz nicht anwendbar ist. Dieser Fall liegt z.B. für das AWP (5.1) vor, wenn die rechte Seite f nicht lokal lipschitzstetig in x ist.

Stetige Abhängigkeit von Parametern. Ebenfalls mit dem Lemma von Gronwall können wir die Änderung der Lösung in Abhängigkeit von Änderungen der Daten abschätzen. Als erstes betrachten wir ein Anfangswertproblem mit einem linearen zeitvarianten System,

$$\dot{x} = A(t,p)x\,, \quad x(t_0) = x_0\,, \tag{5.5}$$

bei dem die Systemmatrix A zusätzlich von einem Parameter p abhängt. Wir bezeichnen Lösungen von (5.5) mit $X(t,p)$, das heißt, zu gegebenem $p \in P$ ist $X(t,p)$ der Wert der Lösung von (5.5) zum Zeitpunkt t.

Satz 5.2 *Sei $I = [t_0, t_1]$, $P \subset \mathbb{R}^m$ kompakt, $A : I \times P \to \mathbb{R}^{(n,n)}$ stetig. Dann hat (5.5) eine eindeutige Lösung $X : I \times P \to \mathbb{R}^n$, welche stetig von p abhängt, das heißt: Gilt $p_n \to p$ in P für $n \to \infty$, so gilt*

$$\lim_{n\to\infty} \sup_{t\in I} \|X(t,p_n) - X(t,p)\|_\infty \to 0\,. \tag{5.6}$$

Es liegt also gleichmäßige Konvergenz vor.

Beweis: Nach Satz 4.1 hat (5.5) für jedes $p \in P$ eine eindeutige Lösung. Die Abbildung X ist daher wohldefiniert. Sei nun (p_n) Folge in P mit $p_n \to p$. Wir definieren $d_n : I \to \mathbb{R}^n$ durch

$$d_n(t) = X(t,p_n) - X(t,p)\,. \tag{5.7}$$

Es gilt für alle $t \in I$

$$\begin{aligned}\dot{d}_n(t) &= A(t,p_n)X(t,p_n) - A(t,p)X(t,p)\\ &= A(t,p_n)(X(t,p_n) - X(t,p)) + (A(t,p_n) - A(t,p))X(t,p)\,.\end{aligned}$$

Die Funktion d_n löst daher auf I das Anfangswertproblem

$$\dot{w} = A(t,p_n)w + \rho_n(t)\,, \quad w(t_0) = 0\,, \tag{5.8}$$

wobei

$$\rho_n(t) = (A(t,p_n) - A(t,p))X(t,p)\,. \tag{5.9}$$

Da A stetig ist auf der kompakten Menge $I \times P$, ist A dort beschränkt, sei

$$\sup_{t\in I,\, q\in P} \|A(t,q)\| = L\,.$$

Sei außerdem $M = \sup_{t\in I} \|X(t,p)\|_\infty$. Es folgt

$$R_n := \sup_{t\in I} \|\rho_n(t)\|_\infty \le M \sup_{t\in I} \|A(t,p_n) - A(t,p)\|\,. \tag{5.10}$$

Wir wenden Satz 5.1 an auf das Anfangswertproblem (5.8) mit $f(t,\eta) = A(t,p_n)\eta + \rho_n(t)$. Es gilt

$$\|f(t,\eta\|_\infty \le L\|\eta\|_\infty + R_n\,.$$

Aus Satz 5.1 folgt

$$\|d_n(t)\|_\infty \le R_n(t-t_0)e^{L(t-t_0)}\,, \quad \text{für alle } t \in I. \tag{5.11}$$

Da A gleichmäßig stetig ist auf der kompakten Menge $I \times P$ und $p_n \to p$ nach Voraussetzung, folgt aus (5.10) dass $R_n \to 0$. Die Behauptung folgt nun aus (5.11). $\square$

Wir kehren zum nichtlinearen Anfangswertproblem zurück und beschäftigen uns mit der Abhängigkeit der Lösung vom Anfangswert x_0. Dabei setzen wir zunächst voraus, dass die rechte Seite f global lipschitzstetig in x ist.

Satz 5.3 *Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, sei $f : D \to \mathbb{R}^n$ stetig und bezüglich x lipschitzstetig mit Lipschitz-Konstante L. Seien $x, y : [t_0, t_1] \to \mathbb{R}^n$ Lösungen von*

$$\dot{x} = f(t, x) \tag{5.12}$$

mit den Anfangswerten

$$x(t_0) = x_0 \,, \quad y(t_0) = y_0 \,. \tag{5.13}$$

Dann gilt

$$\begin{aligned} \|x(t) - y(t)\|_\infty &\le \|x_0 - y_0\|_\infty e^{L(t-t_0)} \,, \quad \textit{für alle } t \in I, \\ \|x - y\|_\infty &\le \|x_0 - y_0\|_\infty e^{L(t_1-t_0)} \,. \end{aligned} \tag{5.14}$$

Beweis: Es gilt

$$\begin{aligned} \|x(t) - y(t)\|_\infty &= \left\| x_0 - y_0 + \int_{t_0}^t f(s, x(s)) - f(s, y(s)) \, ds \right\|_\infty \\ &\le \|x_0 - y_0\|_\infty + \int_{t_0}^t \|f(s, x(s)) - f(s, y(s))\|_\infty \, ds \\ &\le \|x_0 - y_0\|_\infty + \int_{t_0}^t L \|x(s) - y(s)\|_\infty \, ds \,. \end{aligned}$$

Wir wenden Folgerung 3.4 an mit $a = \|x_0 - y_0\|_\infty$, $b(t) = L$ und $z(t) = \|x(t) - y(t)\|_\infty$.
$\square$

Wir können den vorangehenen Satz verwenden, um ein entsprechendes Ergebnis auch dann zu erhalten, wenn die rechte Seite nur lokal lipschitzstetig in x ist, so wie es im Satz von Picard-Lindelöf vorausgesetzt wird.

Satz 5.4 (Stetige Abhängigkeit vom Anfangswert)
Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, sei $f : D \to \mathbb{R}^n$ stetig und bezüglich x lokal lipschitzstetig. Sei $x : [t_0, t_1] \to \mathbb{R}^n$ Lösung in D des AWP $\dot{x} = f(t, x)$ zum Anfangswert $x(t_0) = x_0$. Dann gibt es ein $\delta > 0$ und eine Konstante $c > 0$, so dass für jedes $y_0 \in \mathbb{R}^n$ mit $\|y_0 - x_0\|_\infty < \delta$ die zugehörige Lösung y mit $y(t_0) = y_0$ auf dem Intervall $[t_0, t_1]$ existiert, und

$$\|y(t) - x(t)\|_\infty \le c \|y_0 - x_0\|_\infty \,, \quad \textit{für alle } t \in [t_0, t_1]. \tag{5.15}$$

Beweis: Wir setzen $J = [t_0, t_1]$ und definieren den Graphen

$$G = \{(t, x(t)) : t \in J\}$$

der Lösung x auf J, sowie für $r > 0$

$$G_r = \{(t, \eta) : t \in J, \, \eta \in \mathbb{R}^n \,, \|\eta - x(t)\|_\infty \le r\} \,.$$

Da G eine kompakte Teilmenge der offenen Menge D ist, können wir r so wählen, dass G_r ebenfalls eine kompakte Teilmenge von D ist. Da G_r kompakt ist, hat f eine Lipschitz-Konstante L bezüglich x auf G_r. Wir definieren $\delta > 0$ und $c > 0$ durch

$$\delta e^{L(t_1-t_0)} = \frac{r}{2}, \quad c = e^{L(t_1-t_0)}. \tag{5.16}$$

Wir wollen zeigen, dass für diese Wahl von δ und c die im Satz formulierte Behauptung gilt. Sei $y_0 \in \mathbb{R}^n$ mit $\|y_0 - x_0\|_\infty < \delta$ gegeben. Sei $[t_0, t^*)$ der rechts von t_0 gelegene Teil des maximalen Existenzintervalls der zugehörigen Lösung y. Wir setzen

$$\tau^* = \sup\{t : t < t^*,\, t < t_1\,, \|y(s) - x(s)\|_\infty \le r \text{ für alle } s \in [t_0, t]\}. \tag{5.17}$$

Wegen $\delta < r$ ist $\tau^* > t_0$. Wir setzen nun

$$\tau_1 = \min\{\tau^*, t_1\}.$$

Wegen (5.17) gilt $(t, y(t)) \in G_r$ für alle $t \in [t_0, \tau_1]$. Aus Satz 5.3, angewendet mit G_r statt D (der Beweis bleibt gültig), folgt nun

$$\|y(t) - x(t)\|_\infty \le \|y_0 - x_0\|_\infty e^{L(t-t_0)} \le \delta e^{L(t_1-t_0)} \le \frac{r}{2} \qquad \text{für alle } t \in [t_0, \tau_1). \tag{5.18}$$

Da $(t, y(t)) \in G_r$ für $t < \tau_1$, lässt sich y auf $[t_0, \tau_1]$ fortsetzen, also $\tau_1 < t^*$. Aus (5.18) folgt

$$\|y(\tau_1) - x(\tau_1)\|_\infty \le \frac{r}{2}.$$

Wäre nun $\tau_1 < t_1$, so wäre $\tau^* < t_1$ und $\|y(\tau) - x(\tau)\|_\infty < r$ falls $\tau - \tau_1$ hinreichend klein ist, im Widerspruch zur Definition von τ^*. Also ist $t_1 = \tau_1 < t^*$. Die Ungleichung (5.15) folgt nun aus der ersten Ungleichung in (5.18). □

Satz 5.4 impliziert, dass (unter den Voraussetzungen des Satzes von Picard-Lindelöf) die Lösung des AWP stetig (sogar lipschitzstetig) vom Anfangswert abhängt. Ist etwa $(x_0^k)_{k\in\mathbb{N}}$ eine Folge von Anfangswerten und $(x^k)_{k\in\mathbb{N}}$ die Folge der zugehörigen Lösungen, so impliziert $x_0^k \to x_0$, dass $x^k \to x$ gleichmäßig in $[t_0, t_1]$.

Generell sagt man, dass ein Problem **korrekt gestellt** ist **im Sinne von Hadamard**, falls es eine eindeutige Lösung hat und diese wiederum stetig von den gegebenen "Daten" abhängt. Interpretiert man den Vektor x_0 als die Daten des Anfangswertproblems, so ist also unter den Voraussetzungen des Satzes von Picard-Lindelöf das Anfangswertproblem korrekt gestellt.

Differenzierbare Abhängigkeit vom Anfangswert. Wir fassen die Lösung des Anfangswertproblems

$$\dot{x} = f(t, x)\,, \quad x(t_0) = x_0 \tag{5.19}$$

als Funktion von t und p (statt x_0) auf. Wir bezeichnen mit $X(t;p)$ den Wert der Lösung zum Zeitpunkt t zum Anfangswert $x(t_0) = p$ und setzen $x(t) = X(t; x_0)$. Wir fragen, ob die Ableitung $\partial_p X(t;p)$ existiert und wenn ja, welche Gleichung sie löst. Vorläufig betrachten wir den skalaren Fall $n = 1$. Es muss gelten

$$X(t;p) = p + \int_{t_0}^{t} f(s, X(s;p))\, ds\,. \tag{5.20}$$

Falls $\partial_p X$ existiert, so erhalten wir durch Differenzieren

$$\partial_p X(t;p) = 1 + \int_{t_0}^{t} \partial_x f(s, X(s;p))\partial_p X(s;p)\, ds\,. \tag{5.21}$$

Für $p = x_0$ löst daher die Funktion $w(t) = \partial_p X(t;x_0)$ das lineare zeitvariante Anfangswertproblem

$$\dot{w} = a(t)w\,, \quad w(0) = 1\,, \quad \text{mit} \quad a(t) = \partial_x f(t, X(t;x_0)) = \partial_x f(t, x(t))\,. \tag{5.22}$$

Die Funktion a ist also definiert durch die Lösung des AWP (5.19) und die Ableitung der rechten Seite f. Die Funktion w gibt die Rate an, mit der sich die Lösung zum Zeitpunkt t in Abhängigkeit vom Anfangswert (in der Nähe des Anfangswerts x_0) ändert.

Wir kehren zum allgemeinen Fall $n \geq 1$ zurück. Sei $x : [t_0, t_1] \to \mathbb{R}^n$ Lösung des Systems

$$\dot{x} = f(t, x)\,. \tag{5.23}$$

Dieser wird zugeordnet das lineare zeitvariante System

$$\dot{w} = A(t)w\,, \quad A(t) = \partial_x f(t, x(t)) \in \mathbb{R}^{(n,n)}\,. \tag{5.24}$$

Dieses System heißt **Linearisierung** oder **Variationsgleichung** des Systems (5.23) entlang der Lösung x.

Sei $y : [t_0, t_1] \to \mathbb{R}^n$ eine weitere Lösung von (5.19). Sind $x(t_0) = x_0$ und $y(t_0) = y_0$, so ist die Lösung $w : [t_0, t_1] \to \mathbb{R}^n$ von (5.24) zum Anfangswert $w(t_0) = y_0 - x_0$ eine Approximation erster Ordnung an die Differenz $y - x$ im Sinne des folgenden Satzes.

Satz 5.5 (Approximationseigenschaft der Linearisierung)
Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, seien $f, \partial_x f : D \to \mathbb{R}^n$ stetig, sei f bezüglich x lokal lipschitzstetig. Sei $x : [t_0, t_1] \to \mathbb{R}^n$ Lösung von (5.23) in D zum Anfangswert $x(t_0) = x_0$. Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass gilt:
(i) Zu jedem $y_0 \in \mathbb{R}^n$ mit $\|y_0 - x_0\|_\infty < \delta$ existiert die Lösung $y : [t_0, t_1] \to \mathbb{R}^n$ von (5.23) zum Anfangswert $y(t_0) = y_0$ in D.
(ii) Sei $w : [t_0, t_1] \to \mathbb{R}^n$ die Lösung von (5.24) zum Anfangswert $w(t_0) = y_0 - x_0$. Dann gilt für die Differenz

$$r(t) = y(t) - x(t) - w(t) \tag{5.25}$$

die Abschätzung

$$\|r(t)\|_\infty \leq \varepsilon \|y_0 - x_0\|_\infty\,, \quad \textit{für alle } t \in [t_0, t_1]. \tag{5.26}$$

Beweis: Wir wählen gemäß Satz 5.4 ein $\delta_0 > 0$ und ein $c > 0$ so, dass für alle $y_0 \in \mathbb{R}^n$ mit $\|y_0 - x_0\|_\infty < \delta_0$ die zugehörige Lösung y von $\dot{x} = f(t, x)$ zum Anfangswert $y(t_0) = y_0$ auf $[t_0, t_1]$ existiert und

$$\|y(t) - x(t)\|_\infty \leq c\|y_0 - x_0\|_\infty \qquad \text{für alle } t \in [t_0, t_1] \tag{5.27}$$

erfüllt. Da $A(t)$ in (5.24) nach Voraussetzung an $\partial_x f$ stetig von t abhängt, hat (5.24) nach Satz 4.1 eine eindeutige Lösung $w : [t_0, t_1] \to \mathbb{R}^n$ zum Anfangswert

$$w(t_0) = y_0 - x_0\,. \tag{5.28}$$

Für die Funktion $r = y - x - w$ gilt

$$r(t_0) = y_0 - x_0 - (y_0 - x_0) = 0 \tag{5.29}$$

und für alle $t \in [t_0, t_1]$

$$\begin{aligned}\dot r(t) = \dot y(t) - \dot x(t) - \dot w(t) &= f(t, y(t)) - f(t, x(t)) - A(t)w(t) \\ &= A(t)(y(t) - x(t) - w(t)) + \rho(t)\,,\end{aligned}$$

also

$$\dot r(t) = A(t)r(t) + \rho(t) \tag{5.30}$$

mit

$$\rho(t) = f(t, y(t)) - f(t, x(t)) - \partial_x f(t, x(t))(y(t) - x(t))\,. \tag{5.31}$$

Wir setzen

$$L = \max_{t\in[t_0,t_1]} \|A(t)\|\,, \quad R = \max_{t\in[t_0,t_1]} \|\rho(t)\|_\infty\,. \tag{5.32}$$

Es folgt

$$\|A(t)\eta + \rho(t)\|_\infty \le R + L\|\eta\|_\infty\,, \quad \text{für alle } t \in [t_0, t_1] \text{ und alle } \eta \in \mathbb{R}^n.$$

Wir wenden Satz 5.1 an auf das Anfangswertproblem (5.29), (5.30) und erhalten

$$\|r(t)\|_\infty \le R(t - t_0)e^{L(t-t_0)}\,, \quad \text{für alle } t \in [t_0, t_1]. \tag{5.33}$$

Um (5.26) zu erhalten, müssen wir R und damit das Restglied (5.31) geeignet abschätzen. Für $t \in [t_0, t_1]$ setzen wir

$$d(\alpha) = f(t, x(t) + \alpha(y(t) - x(t)))\,, \quad \alpha \in [0, 1]\,.$$

Es gilt dann

$$\begin{aligned}f(t, y(t)) - f(t, x(t)) = d(1) - d(0) &= \int_0^1 d'(\alpha)\,d\alpha \\ &= \int_0^1 \partial_x f(t, x(t) + \alpha(y(t) - x(t))) \cdot (y(t) - x(t))\,d\alpha \\ &= \int_0^1 \partial_x f(t, x(t) + \alpha(y(t) - x(t)))\,d\alpha \cdot (y(t) - x(t))\,.\end{aligned}$$

Einsetzen in (5.31) ergibt

$$\rho(t) = \int_0^1 \underbrace{\partial_x f(t, x(t) + \alpha(y(t) - x(t))) - \partial_x f(t, x(t))}_{=:\tilde\rho(t)}\,d\alpha \cdot (y(t) - x(t))\,. \tag{5.34}$$

Die beiden Argumente von $\partial_x f$ im Integranden unterscheiden sich um höchstens $\|y(t) - x(t)\|_\infty$. Sei nun $\tilde\varepsilon > 0$. Wir wählen $\gamma > 0$, so dass gilt

$$\|y(t) - x(t)\|_\infty \le \gamma \quad \text{auf } [t_0, t_1] \qquad \Rightarrow \qquad \|\tilde\rho(t)\| \le \frac{\tilde\varepsilon}{c} \quad \text{auf } [t_0, t_1]\,.$$

Das ist möglich, da $\partial_x f$ stetig und daher auf Kompakta gleichmäßig stetig ist. Wir setzen

$$\delta = \min\left\{\frac{\gamma}{c}, \delta_0\right\}.$$

Mit (5.27) und (5.34) folgt nun, da $\delta \le \delta_0$,

$$\begin{aligned} \|y_0 - x_0\|_\infty \le \delta \quad \text{auf } [t_0, t_1] \qquad &\Rightarrow \qquad \|y(t) - x(t)\|_\infty \le c\delta \le \gamma \quad \text{auf } [t_0, t_1] \\ &\Rightarrow \qquad \|\tilde{\rho}(t)\| \le \frac{\tilde{\varepsilon}}{c} \quad \text{auf } [t_0, t_1] \\ &\Rightarrow \qquad \|\rho(t)\| \le \tilde{\varepsilon}\|y_0 - x_0\|_\infty \quad \text{auf } [t_0, t_1]. \end{aligned}$$

Für

$$\varepsilon = \tilde{\varepsilon}(t_1 - t_0)e^{L(t_1 - t_0)} \tag{5.35}$$

erhalten wir daher aus (5.33)

$$\|r(t)\|_\infty \le \varepsilon \|y_0 - x_0\|_\infty, \quad \text{für alle } t \in [t_0, t_1]. \tag{5.36}$$

Sei nun $\varepsilon > 0$ vorgegeben. Wir definieren $\tilde{\varepsilon}$ durch (5.35) und wählen γ, δ wie beschrieben. Dann gilt (5.36) für alle y_0 mit $\|y_0 - x_0\|_\infty < \delta$. □

Mit Hilfe des vorangehenden Satzes können wir beweisen, dass die Lösung des Anfangswertproblems differenzierbar vom Anfangswert x_0 abhängt. Zu diesem Zweck bezeichnen wir mit $X(t;p)$ die Lösung des Anfangswertproblems

$$\dot{x} = f(t, x), \quad x(t_0) = p. \tag{5.37}$$

Wir interessieren uns für die Existenz und Form der partiellen Ableitung nach der j-ten Komponente des Anfangswerts

$$\partial_j X(t; x_0) = \lim_{h \to 0} \frac{X(t; x_0 + he_j) - X(t; x_0)}{h}, \quad t \in [t_0, t_1]. \tag{5.38}$$

Sie ist ein Element des $\mathbb{R}^n$; wir können sie als Tangentenvektor $\gamma_j'(0)$ an die Kurve $\gamma_j(h) = X(t; x_0 + he_j)$ interpretieren.

Satz 5.6 (Differenzierbare Abhängigkeit vom Anfangswert)
Unter den Voraussetzungen von Satz 5.5 gilt: Für jedes $t \in [t_0, t_1]$ existiert die Ableitung $\partial_j X(t; x_0)$ und ist gleich $w(t)$, wobei $w : [t_0, t_1] \to \mathbb{R}^n$ die Lösung der Variationsgleichung

$$\dot{w} = \partial_x f(t, x(t))w \tag{5.39}$$

ist zum Anfangswert

$$w(t_0) = e_j. \tag{5.40}$$

Beweis: Nach den vorangehenden Sätzen existiert für $|h| < \delta_0$, δ_0 hinreichend klein, die Lösung $t \mapsto X(t, x_0 + he_j)$ auf $[t_0, t_1]$. Sei w die Lösung von (5.39) und (5.40). Für $|h| < \delta_0$ ist die Funktion hw Lösung von (5.39) zum Anfangswert he_j. Nach Satz 5.5 gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ so dass

$$\|X(t; x_0 + he_j) - X(t; x_0) - hw(t)\|_\infty \le \varepsilon \|he_j\|_\infty \tag{5.41}$$

gilt für alle $|h| < \delta$ und alle $t \in [t_0, t_1]$. Division durch $|h|$ zeigt, dass (5.41) äquivalent ist zu

$$\left\| \frac{X(t; x_0 + he_j) - X(t; x_0)}{h} - w(t) \right\|_\infty \le \varepsilon$$

falls $h \neq 0$. Daraus folgt die Behauptung. □

Wir zeigen nun, dass die Abbildung $(t, p) \mapsto X(t, p)$, welche die Lösung von

$$\dot{x} = f(t, x)\,, \quad x(t_0) = p\,, \tag{5.42}$$

zum Zeitpunkt t liefert, stetig differenzierbar ist.

Satz 5.7 *Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, seien $f, \partial_x f : D \to \mathbb{R}^n$ stetig, sei f bezüglich x lokal lipschitzstetig. Sei $x : [t_0, t_1] \to \mathbb{R}^n$ Lösung von (5.23) in D zum Anfangswert $x(t_0) = x_0$. Dann gibt es ein $\delta_0 > 0$, so dass für $U = \{p : \|p - x_0\|_\infty < \delta_0\}$ gilt: Die Abbildung $X : [t_0, t_1] \times U \to \mathbb{R}^n$ ist stetig differenzierbar.*

Beweis: Wir zeigen als erstes, dass X stetig ist. Sei $I = [t_0, t_1]$. Sei $t_k \to t$ in I, $p_k \to p$ in U. Es ist

$$X(t_k, p_k) - X(t, p) = (X(t_k, p_k) - X(t_k, p) + (X(t_k, p) - X(t, p))\,. \tag{5.43}$$

Der erste Term auf der rechten Seite konvergiert gegen 0 wegen Satz 5.4, der zweite, da $s \mapsto X(s, p)$ Lösung von $\dot{x} = f(t, x)$ ist. Für die Zeitableitung gilt

$$\partial_t X(t; p) = f(t, X(t; p))\,.$$

Da die rechte Seite stetig ist in (t, p), ist $\partial_t X$ stetig. Weiter gilt: Die Funktionen $t \mapsto \partial_j X(t; p)$ sind nach Satz 5.6 Lösungen der Variationsgleichung

$$\dot{w} = \partial_x f(t, X(t, p))w\,, \quad w(t_0) = e_j\,.$$

Die Matrix $A(t, p) = \partial_x f(t, X(t, p))$ ist nach dem Gezeigten stetig in (t, p). Aus Satz 5.2 folgt, dass für $p_k \to p$ die Funktionen $\partial_j X(t, p_k)$ gleichmäßig gegen $\partial_j X(t, p)$ konvergieren. Die zu (5.43) analoge Zerlegung

$$\partial_j X(t_k, p_k) - \partial_j X(t, p) = (\partial_j X(t_k, p_k) - \partial_j X(t_k, p)) + (\partial_j X(t_k, p) - \partial_j X(t, p))$$

liefert die Stetigkeit von $\partial_j X$ mit den analogen Argumenten. □

6 Stabilität

Definition der Stabilität. Im vorigen Kapitel hat es sich herausgestellt, dass auf beschränkten Intervallen $[t_0, t_1]$ die Lösung des AWP

$$\dot{x} = f(t,x)\,, \quad x(t_0) = x_0\,, \tag{6.1}$$

stetig vom Anfangswert x_0 abhängt, wenn die rechte Seite f die üblichen Voraussetzungen erfüllt. Es stellt sich die Frage, was in einem unbeschränkten Existenzintervall passiert. Wir beschränken uns auf autonome Systeme

$$\dot{x} = f(x)\,, \quad f : D \to \mathbb{R}^n\,, \quad D \subset \mathbb{R}^n \quad \text{offen.} \tag{6.2}$$

Definition 6.1 (Gleichgewichtspunkt)
Ein $x_* \in D$ heißt Gleichgewichtspunkt von (6.2), falls

$$f(x_*) = 0\,. \tag{6.3}$$

□

Ist x_* ein Gleichgewichtspunkt von (6.2), so ist die konstante Funktion

$$x(t) = x_* \tag{6.4}$$

offensichtlich eine Lösung von (6.2). Im Spezialfall

$$\dot{x} = Ax\,, \quad A \in \mathbb{R}^{(n,n)}\,, \tag{6.5}$$

ist $x_* = 0$ ein Gleichgewichtspunkt. Wir betrachten zunächst den Fall $n = 1$,

$$\dot{x} = ax\,, \quad a \in \mathbb{R}\,. \tag{6.6}$$

Die Lösung zum Anfangswert $x(0) = x_0$ ist

$$x(t) = e^{at} x_0\,. \tag{6.7}$$

Es gilt:

$$\begin{aligned} a > 0 \quad &\Rightarrow \quad \lim_{t\to\infty} x(t) = +\infty\,, \\ a = 0 \quad &\Rightarrow \quad x(t) = x_0\,, \quad \text{für alle } t, \\ a < 0 \quad &\Rightarrow \quad \lim_{t\to\infty} x(t) = 0\,. \end{aligned} \tag{6.8}$$

Im ersten Fall liegt Divergenz vor, im dritten Fall Konvergenz gegen den Gleichgewichtspunkt. Im zweiten Fall gilt, dass die Lösung x sich nicht aus einer vorgegebenen Umgebung des Gleichgewichtspunkts entfernt, sofern der Anfangswert in dieser Umgebung liegt.

Wir geben nun eine allgemeine Definition der Stabilität eines Gleichgewichtspunktes.

Definition 6.2 (Stabilität)
Sei $D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}^n$, x_* Gleichgewichtspunkt von $\dot{x} = f(x)$.
(i) Der Gleichgewichtspunkt $x_* \in D$ heißt ***stabil****, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass alle Lösungen von (6.2) mit $\|x(0) - x_*\| < \delta$ auf ganz $[0, \infty)$ existieren und*

$$\|x(t) - x_*\| < \varepsilon \,, \quad \textit{für alle } t \geq 0, \tag{6.9}$$

erfüllen; falls das nicht gilt, heißt x_* ***instabil****.*
(ii) Der Gleichgewichtspunkt x_* heißt ***asymptotisch stabil****, falls er stabil ist und falls es ein $\delta > 0$ gibt, so dass für alle Lösungen von (6.2) mit $\|x(0) - x_*\| < \delta$ gilt, dass*

$$\lim_{t\to\infty} \|x(t) - x_*\| = 0 \,. \tag{6.10}$$

(iii) Der Gleichgewichtspunkt x_* heißt ***exponentiell stabil****, falls es ein $\delta > 0$ und Konstanten $c, \sigma > 0$ gibt, so dass für alle Lösungen von (6.2) mit $\|x(0) - x_*\| < \delta$ gilt, dass*

$$\|x(t) - x_*\| < ce^{-\sigma t}\|x(0) - x_*\| \,, \quad \textit{für alle } t \geq 0. \tag{6.11}$$

□

Stabilität linearer Systeme. Wir betrachten das lineare System

$$\dot{x} = Ax \,, \quad A \in \mathbb{R}^{(n,n)} \,. \tag{6.12}$$

Seine Lösung zum Anfangswert $x(0) = x_0$ ist gegeben durch

$$x(t) = e^{tA}x_0 \,. \tag{6.13}$$

Im ersten Kapitel haben wir die Form von e^{tA} bereits analysiert. Ist

$$A = V^{-1}DV \tag{6.14}$$

die Zerlegung von A in die Jordan-Normalform, so gilt

$$e^{tA} = V^{-1}e^{tD}V \,, \quad e^{tD} = \operatorname{diag}(e^{tD_1}, \ldots, e^{tD_k}), \tag{6.15}$$

und für ein einzelnes Jordan-Kästchen

$$D_j = \lambda_j I + N_j = \begin{pmatrix} \lambda_j & 1 & 0 & \cdots & 0 \\ 0 & \lambda_j & 1 & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \lambda_j & 1 \\ 0 & & & 0 & \lambda_j \end{pmatrix} \in \mathbb{C}^{n_j \times n_j} \,, \tag{6.16}$$

gilt

$$e^{tD_j} = e^{\lambda_j t}e^{tN_j} = e^{\lambda_j t}\sum_{\mu=0}^{n_j - 1} \frac{t^\mu}{\mu!}N_j^\mu \,. \tag{6.17}$$

Jedes Element der Matrix e^{tD_j} hat die Form

$$p(t)e^{\lambda_j t} \,, \tag{6.18}$$

wobei p ein Polynom vom Grad $n_j - 1$ oder kleiner ist.

Lemma 6.3 *Sei $A \in \mathbb{R}^{(n,n)}$, sei $\gamma \in \mathbb{R}$ mit*

$$\gamma > \max\{\mathrm{Re}\,\lambda : \lambda \text{ ist Eigenwert von } A\}\,. \tag{6.19}$$

Dann gibt es ein $c > 0$ mit

$$\left\|e^{tA}\right\| \leq ce^{\gamma t}\,, \quad \text{für alle } t \geq 0. \tag{6.20}$$

Beweis: Durch Betrachtung der Jordan-Normalform (siehe oben) folgt für alle $t \geq 0$

$$\left\|e^{tA}\right\| \leq \|V^{-1}\| \left\|e^{tD}\right\| \|V\|\,, \tag{6.21}$$

$$e^{tD_j} = e^{\lambda_j t}e^{tN_j} = e^{\lambda_j t}\sum_{\mu=0}^{n_j-1}\frac{t^\mu}{\mu!}N_j^\mu\,. \tag{6.22}$$

Es gibt also ein Polynom p_j vom Grad höchstens $n_j - 1$ mit

$$\left\|e^{tD_j}\right\| \leq e^{(\mathrm{Re}\lambda_j)t}|p_j(t)| = e^{\gamma t}e^{(\mathrm{Re}\lambda_j-\gamma)t}|p_j(t)|\,. \tag{6.23}$$

Wegen

$$\lim_{t\to\infty} e^{(\mathrm{Re}\lambda_j-\gamma)t}p_j(t) = 0$$

gibt es $c_j > 0$ mit

$$\left\|e^{tD_j}\right\| \leq c_j e^{\gamma t}\,, \quad t \geq 0\,. \tag{6.24}$$

Die Behauptung folgt nun aus (6.21) und (6.24). □

Satz 6.4 *Sei $A \in \mathbb{R}^{(n,n)}$, sei*

$$\alpha = \max\{\mathrm{Re}\,\lambda : \lambda \text{ ist Eigenwert von } A\}\,. \tag{6.25}$$

Dann ist der Gleichgewichtspunkt $x_* = 0$ von $\dot{x} = Ax$

- *exponentiell stabil, falls $\alpha < 0$,*
- *instabil, falls $\alpha > 0$.*

Ist $\alpha = 0$, so ist $x_* = 0$ stabil, falls für jeden Eigenwert λ_j mit $\mathrm{Re}\,\lambda_j = 0$ gilt, dass $n_j = 1$ in der Jordan-Normalform von A; andernfalls ist $x_* = 0$ instabil.

Beweis: Wir betrachten als erstes den Fall $\alpha < 0$. Wir wählen $\sigma > 0$ mit $\alpha < -\sigma < 0$. Sei $x_0 \in \mathbb{R}^n$ beliebig und x die Lösung von $\dot{x} = Ax$ zum Anfangswert $x(0) = x_0$. Aus Lemma 6.3 folgt

$$\|x(t)\| = \left\|e^{tA}x_0\right\| \leq \left\|e^{tA}\right\| \|x_0\| \leq ce^{-\sigma t}\|x_0\|\,.$$

Wegen $\sigma > 0$ ist $x_* = 0$ exponentiell stabil.

Für die Fälle $\alpha > 0$ und $\alpha = 0$ betrachten wir zunächst den Spezialfall $A = D$, $V = I$ in der Jordan-Normalform (6.14).

Sei $\alpha > 0$. Dann gibt es ein j, so dass mit dem Polynom p_j aus (6.18) gilt

$$\lim_{t\to\infty} e^{\lambda_j t}|p_j(t)| = \infty\,,$$

Wählt man den Anfangswert x_0 geeignet, so gilt für die zugehörige Lösung

$$\lim_{t\to\infty} \|x(t)\| = \infty\,,$$

woraus die Instabilität folgt.

Sei $\alpha = 0$. Falls es einen Eigenwert λ_j gibt mit $\operatorname{Re}\lambda_j = 0$ und $n_j > 1$, so gilt

$$e^{tD_j}\begin{pmatrix}0\\ \vdots\\ 0\\ \delta\end{pmatrix} = \delta e^{\lambda_j t}\begin{pmatrix}\frac{t^{n_j-1}}{(n_j-1)!}\\ \vdots\\ t\\ 1\end{pmatrix}. \tag{6.26}$$

Da die rechte Seite von (6.26) unbeschränkt ist für jedes (beliebig kleine) $\delta > 0$, folgt die Instabilität. Ist $n_j = 1$ für alle Eigenwerte mit $\operatorname{Re}\lambda_j = 0$, so sind alle Matrixelemente von e^{tD} beschränkt, und es folgt die Stabilität.

Sei nun A beliebig mit Jordan-Normalform (6.14). Ist $z : [0,\infty) \to \mathbb{R}^n$ Lösung von

$$\dot z = Dz\,, \tag{6.27}$$

so ist $x = Vz$ eine Lösung von $\dot x = Ax$ und umgekehrt. Es folgt, dass der Gleichgewichtspunkt $x_* = 0$ genau dann stabil (exponentiell stabil, instabil) ist für $\dot x = Ax$, wenn dasselbe für (6.27) gilt. □

Damit ist die Frage der Stabilität des Gleichgewichtspunkts $x_* = 0$ für $\dot x = Ax$ vollständig gelöst.

Stabilität bei kleinen nichtlinearen Störungen. Wir betrachten das nichtlineare System

$$\dot x = Ax + g(x)\,, \quad A \in \mathbb{R}^{(n,n)}\,, \tag{6.28}$$

wobei

$$\lim_{\eta\to 0}\frac{\|g(\eta)\|}{\|\eta\|} = 0\,. \tag{6.29}$$

Ist g stetig, so impliziert (6.29), dass $g(0) = 0$ und somit $x_* = 0$ ein Gleichgewichtspunkt ist von (6.28).

Satz 6.5 *Sei $A \in \mathbb{R}^{(n,n)}$ mit $\operatorname{Re}\lambda < 0$ für alle Eigenwerte λ von A. Sei $r > 0$, $g : K(0;r) \to \mathbb{R}^n$ stetig und bzgl. x lokal Lipschitz-stetig, es gelte (6.29). Dann gibt es $\beta, c, \delta > 0$, so dass das AWP*

$$\dot x = Ax + g(x)\,, \quad x(0) = x_0\,, \tag{6.30}$$

für alle $x_0 \in \mathbb{R}^n$ mit $\|x_0\|_\infty \le \delta$ eine eindeutige Lösung $x : [0,\infty) \to \mathbb{R}^n$ hat, für die gilt

$$\|x(t)\|_\infty \le ce^{-\beta t}\|x_0\|_\infty\,, \quad \textit{für alle } t \ge 0. \tag{6.31}$$

Insbesondere ist 0 exponentiell stabiler Gleichgewichtspunkt von (6.28).

Beweis: Wir setzen

$$\alpha = \max\{\operatorname{Re}\lambda : \lambda \text{ ist Eigenwert von } A\}\,. \tag{6.32}$$

Es ist $\alpha < 0$ nach Voraussetzung. Wir wählen $\sigma > 0$ mit $\alpha < -\sigma < 0$. Aus Lemma 6.3 mit $\gamma = -\sigma$ folgt die Existenz eines $c > 0$ mit

$$\left\|e^{tA}\right\| \le ce^{-\sigma t}\,, \quad t \ge 0\,. \tag{6.33}$$

Wir wählen $\delta_1 \in (0, r)$, so dass gemäß (6.29) gilt

$$\|\eta\|_\infty \le \delta_1 \quad \Rightarrow \quad \|g(\eta\|_\infty \le \frac{\sigma}{2c}\|\eta\|_\infty\,, \tag{6.34}$$

und setzen

$$\delta = \min\left\{\frac{\delta_1}{2c}, \frac{\delta_1}{2}\right\}\,. \tag{6.35}$$

Sei nun $x_0 \in \mathbb{R}^n$ mit $\|x_0\|_\infty < \delta$ vorgegeben. Aus Satz 2.6 folgt, dass das AWP (6.30) eine eindeutige Lösung $x : I \to \mathbb{R}^n$ mit einem maximalen offenen Existenzintervall I hat. Diese ist gegeben durch

$$x(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}g(x(s))\,ds\,. \tag{6.36}$$

Wir definieren

$$J = \{t : t > 0,\, t \in I,\, \|x(s)\|_\infty < \delta_1 \text{ für alle } s \in [0, t]\}\,. \tag{6.37}$$

Für alle $t \in J$ gilt wegen (6.33), (6.34)

$$\begin{aligned}\|x(t)\|_\infty &\le \left\|e^{tA}\right\|\,\|x_0\|_\infty + \int_0^t \left\|e^{(t-s)A}\right\|\,\|g(x(s))\|_\infty\,ds \\ &\le ce^{-\sigma t}\|x_0\|_\infty + \int_0^t ce^{-\sigma(t-s)}\frac{\sigma}{2c}\|x(s)\|_\infty\,ds\,,\end{aligned} \tag{6.38}$$

also

$$e^{\sigma t}\|x(t)\|_\infty \le c\|x_0\|_\infty + \int_0^t e^{\sigma s}\frac{\sigma}{2}\|x(s)\|_\infty\,ds\,. \tag{6.39}$$

Wir wenden das Lemma von Gronwall in der Form von Folgerung 3.4 an mit

$$z(t) = e^{\sigma t}\|x(t)\|_\infty\,, \quad a = c\|x_0\|_\infty\,, \quad b(t) = \frac{\sigma}{2}\,,$$

und erhalten

$$e^{\sigma t}\|x(t)\|_\infty \le c\|x_0\|_\infty e^{\frac{\sigma}{2}t}\,,$$

also wegen (6.35)

$$\|x(t)\|_\infty \le ce^{-\frac{\sigma}{2}t}\|x_0\|_\infty \le \frac{\delta_1}{2}\,. \tag{6.40}$$

Damit ist (6.31) für $t \in J$ bewiesen. Es bleibt zu zeigen, dass $J = [0, \infty)$. Wir zeigen zunächst, dass $J = I$: Andernfalls gilt für $t_* = \sup J$, dass $t_* \in I$ und $\|x(t_*)\|_\infty = \delta_1$ im Widerspruch zu (6.40), da (6.40) auch für $t = t_*$ gilt (Stetigkeit von x). Es ist außerdem $I = [0, \infty)$, da andernfalls die Menge $\{(t, x(t)) : t \ge 0,\, t \in I\}$ in der kompakten Menge $[0, \sup I] \times K(0, \frac{\delta_1}{2})\}$ enthalten wäre, im Widerspruch zu Satz 2.7. $\square$

Folgerung 6.6 *Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ stetig differenzierbar, sei $x_* \in \mathbb{R}^n$ Gleichgewichtspunkt von*

$$\dot{x} = f(x)\,. \tag{6.41}$$

Wenn alle Eigenwerte von $Df(x_*)$ negativen Realteil haben, ist x_* exponentiell stabil.

Beweis: Wir wenden Satz 6.5 an mit

$$A = Df(x_*)\,, \quad g(x) = f(x_* + x) - f(x_*) - Df(x_*)x\,.$$

Es folgt, dass 0 exponentiell stabiler Gleichgewichtspunkt ist von

$$\dot{x} = Ax + g(x) = f(x_* + x)\,.$$

□

Stabilität nichtlinearer Systeme: Die Methode von Ljapunov. Sei x_* Gleichgewichtspunkt von

$$\dot{x} = f(x)\,. \tag{6.42}$$

Wir definieren

$$V(\eta) = \|\eta - x_*\|_2\,, \quad \eta \in \mathbb{R}^n\,. \tag{6.43}$$

Falls für jede Lösung x von (6.42) gilt

$$\frac{\mathrm{d}}{\mathrm{dt}} V(x(t)) \le 0\,, \quad \text{für alle } t, \tag{6.44}$$

so gilt $V(x(s)) \le V(x(t))$ für alle $s \ge t$ und daher

$$\|x(t) - x_*\|_2 \le r \quad \Rightarrow \quad \|x(s) - x_*\|_2 \le r \quad \text{für alle } s \ge t. \tag{6.45}$$

Hieraus folgt die Stabilität des Gleichgewichtspunktes x_*. Dieser Ansatz wird in der Methode von Ljapunov (1892) verallgemeinert. Ist $V : \mathbb{R}^n \to \mathbb{R}$ differenzierbar, so gilt für Lösungen x von (6.42)

$$\frac{\mathrm{d}}{\mathrm{dt}} V(x(t)) = \langle \operatorname{grad} V(x(t)), \dot{x}(t) \rangle = \langle \operatorname{grad} V(x(t)), f(x(t)) \rangle\,, \tag{6.46}$$

also folgt (6.44), falls

$$\langle \operatorname{grad} V(\eta), f(\eta) \rangle \le 0\,, \quad \text{für alle } \eta \in \mathbb{R}^n. \tag{6.47}$$

Der Punkt dabei ist, dass in (6.47) die Lösung x von (6.42) nicht vorkommt, nur die (bekannte) rechte Seite f. Die Schwierigkeit ist, zu gegebenem f ein geeignetes V zu finden.

Definition 6.7 (Ljapunov-Funktion)
Sei $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}^n$ stetig. Sei $x_* \in D$ Gleichgewichtspunkt von $\dot{x} = f(x)$. Eine Funktion $V : U \to \mathbb{R}$, wobei $U \subset D$ offen ist mit $x_* \in U$, heißt ***Ljapunov-Funktion*** *für f in x_*, falls gilt*

(i) *V ist stetig in U,*

$$V(x_*) = 0\,, \quad V(\eta) > 0 \quad \textit{für alle } \eta \in U \setminus \{x_*\}, \tag{6.48}$$

(ii) *V ist stetig differenzierbar in $U \setminus \{x_*\}$,*

$$0 \geq \langle \operatorname{grad} V(\eta), f(\eta) \rangle = \sum_{k=1}^{n} \partial_k V(\eta) f_k(\eta) \quad \textit{für alle } \eta \in U \setminus \{x_*\}. \tag{6.49}$$

V heißt ***strikte Ljapunov-Funktion*** *für f in x_*, falls in (6.49) "$>$" statt "$\geq$" steht.* □

Satz 6.8 *Sei $D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}^n$ lokal Lipschitz-stetig, sei $x_* \in D$ Gleichgewichtspunkt für $\dot{x} = f(x)$, sei U Umgebung von x_* und $V : U \to \mathbb{R}$. Dann gilt:*

(i) *Ist V Ljapunov-Funktion für f in x_*, so ist x_* stabil.*

(ii) *Ist V strikte Ljapunov-Funktion für f in x_*, so ist x_* asymptotisch stabil.*

Beweis: Wir zeigen (i). Sei $\varepsilon > 0$ beliebig mit $K(x_*; \varepsilon) \subset U$. Wir definieren

$$\alpha = \inf\{V(\eta) : \|\eta - x_*\|_\infty = \varepsilon\}\,. \tag{6.50}$$

Es gilt $\alpha > 0$ nach Voraussetzung an V. Wir wählen $\delta > 0$ so, dass $\delta < \varepsilon$ und

$$V(\eta) < \alpha\,, \quad \text{für alle } \eta \in B(x_*; \delta). \tag{6.51}$$

Sei nun $x_0 \in B(x_*; \delta)$ beliebig, sei $x : I \to \mathbb{R}^n$ die eindeutige Lösung zum Anfangswert x_0 mit maximalem Existenzintervall I (gemäß Satz 2.6). Zunächst gilt

$$\|x(0) - x_*\|_\infty = \|x_0 - x_*\|_\infty < \varepsilon\,.$$

Annahme: Es gibt ein $t \in I$, $t \geq 0$, mit $\|x(t) - x_*\|_\infty = \varepsilon$. Sei t_* das minimale t mit dieser Eigenschaft. Dann gilt $\|x(t) - x_*\|_\infty < \varepsilon$ für alle $t \in [0, t_*)$, und $x(t) \in U$ für alle $t \in [0, t_*)$. Nach Voraussetzung an V und (6.46), (6.47)) gilt

$$V(x(t)) \leq V(x_0) < \alpha\,, \quad \text{für alle } t \in [0, t_*], \tag{6.52}$$

also nach Definition von α

$$\|x(t) - x_*\|_\infty \neq \varepsilon\,, \quad \text{für alle } t \in [0, t_*]. \tag{6.53}$$

Insbesondere gilt $\|x(t_*) - x_*\|_\infty \neq \varepsilon$, ein Widerspruch zur Annahme. Es folgt, dass

$$\|x(t) - x_*\|_\infty \neq \varepsilon\,, \quad \text{für alle } t \in I,\, t \geq 0.$$

Da $\|x_0 - x_*\|_\infty < \varepsilon$, folgt aus der Stetigkeit von x

$$\|x(t) - x_*\|_\infty < \varepsilon\,, \quad \text{für alle } t \in I,\, t \geq 0. \tag{6.54}$$

Es gilt außerdem $[0, \infty) \subset I$, da andernfalls $\{(t, x(t)) : t \in I, \, t \geq 0\}$ in der kompakten Menge $[0, \sup I] \times K(x_*; \delta)$ enthalten wäre im Widerspruch zu Satz 2.7. Damit ist (i) bewiesen. Um (ii) zu beweisen, genügt es zu zeigen, dass für die oben zu beliebigem $x_0 \in B(x_*; \delta)$ konstruierte Lösung $x : [0, \infty) \to \mathbb{R}^n$ gilt

$$\lim_{t \to \infty} x(t) = x_* \,, \tag{6.55}$$

falls V strikte Ljapunov-Funktion ist. Wir nehmen zunächst an, es gäbe ein $\delta_0 > 0$, $\delta_0 < \varepsilon$, und ein $T > 0$ mit

$$\|x(t) - x_*\|_\infty \geq \delta_0 \,, \quad \text{für alle } t \geq T. \tag{6.56}$$

Wir definieren

$$\beta = \sup\{\langle \operatorname{grad} V(\eta), f(\eta) \rangle : \, \delta_0 \leq \|\eta - x_*\|_\infty \leq \varepsilon\} \,. \tag{6.57}$$

Es gilt $\beta < 0$ nach Voraussetzung an V. Für alle $t \geq T$ gilt $\delta_0 \leq \|x(t) - x_*\|_\infty \leq \varepsilon$, also

$$V(x(t)) - V(x(T)) = \int_t^T \langle \operatorname{grad} V(x(s)), f(x(s)) \rangle \, ds \leq (t - T)\beta \,, \tag{6.58}$$

also $V(x(t)) < 0$ für hinreichend großes t, ein Widerspruch zur Definition von V. Also gibt es eine Folge $t_n \to \infty$ mit

$$\lim_{n \to \infty} x(t_n) = x_* \,, \tag{6.59}$$

also $V(x(t_n)) \to V(x_*) = 0$ und wegen der Monotonie von $t \mapsto V(x(t))$ auch

$$\lim_{t \to \infty} V(x(t)) = 0 \,. \tag{6.60}$$

Hieraus folgt (6.55), da andernfalls eine Folge $s_n \to \infty$ und ein $\delta_1 > 0$ existieren mit

$$\inf_{n \in \mathbb{N}} \|x(s_n) - x_*\|_\infty \geq \delta_1 \,, \quad \inf_{n \in \mathbb{N}} V(x(s_n)) \geq \inf\{V(\eta) : \delta_1 \leq \|\eta - x_*\|_\infty \leq \varepsilon\} > 0 \,, \tag{6.61}$$

im Widerspruch zu (6.60). $\square$

Wir betrachten als erstes Beispiel ein System

$$\dot{x} = f(x) \,, \tag{6.62}$$

bei dem die rechte Seite durch ein **Gradientenfeld**

$$f(x) = -\operatorname{grad} \Phi(x) \tag{6.63}$$

gegeben ist.

Folgerung 6.9 *Sei $D \subset \mathbb{R}^n$ offen, $\Phi : D \to \mathbb{R}$ zweimal stetig differenzierbar, sei $x_* \in D$ ein striktes lokales Minimum von Φ. Dann ist x_* ein stabiler Gleichgewichtspunkt von (6.62), (6.63). Ist außerdem $D^2\Phi(x_*)$ positiv definit, so ist x_* asymptotisch stabil.*

Beweis: Sei U eine Umgebung von x_* mit $\Phi(\eta) > \Phi(x_*)$ für alle $\eta \in U$, $\eta \neq x_*$. Dann wird durch

$$V(\eta) = \Phi(\eta) - \Phi(x_*) \tag{6.64}$$

eine Ljapunov-Funktion $V : U \to \mathbb{R}$ definiert, da

$$\langle \operatorname{grad} V(\eta), f(\eta) \rangle = -\|\operatorname{grad} \Phi(\eta)\|_2^2 \leq 0 \,. \tag{6.65}$$

Sei nun $D^2\Phi(x_*)$ positiv definit. Es ist $\operatorname{grad} \Phi(x_*) = 0$, und aus dem Satz über inverse Funktionen folgt, dass $\operatorname{grad} \Phi$ auf einer hinreichend kleinen Umgebung U_1 von x_* bijektiv ist, also ist $\operatorname{grad} \Phi(\eta) \neq 0$ für $\eta \in U_1$. In (6.65) gilt dann "$<$" statt "$\leq$", also ist V strikte Ljapunov-Funktion in U_1. $\square$

Als zweites Beispiel betrachten wir die Bewegung einer Punktmasse m in einem ortsabhängigen Kraftfeld $k(x)$, $x \in \mathbb{R}^3$, welches als Gradient eines Potentials gegeben ist,

$$m\ddot{x} = k(x) \,, \quad k(x) = -\operatorname{grad} \Phi(x) \,. \tag{6.66}$$

Wir schreiben (6.66) um in ein System

$$\dot{x} = v \,, \tag{6.67}$$

$$\dot{v} = \frac{1}{m} k(x) = -\frac{1}{m} \operatorname{grad} \Phi(x) \,. \tag{6.68}$$

Folgerung 6.10 *Sei $\Phi : \mathbb{R}^3 \to \mathbb{R}$ stetig differenzierbar, sei x_* ein striktes lokales Minimum von Φ. Dann ist $(x_*, 0)$ ein stabiler, aber nicht asymptotisch stabiler Gleichgewichtspunkt von (6.67), (6.68).*

Beweis: Wir schreiben $\eta = (x, v)$ und definieren $V : \mathbb{R}^6 \to \mathbb{R}$ durch

$$V(\eta) = V(x, v) = \frac{m}{2} \|v\|_2^2 + \Phi(x) - \Phi(x_*) \,. \tag{6.69}$$

Dann gilt

$$V(x_*, 0) = 0 \,, \quad V(x, v) \geq \Phi(x) - \Phi(x_*) > 0 \,, \tag{6.70}$$

falls $x \neq x_*$, in einer geeigneten Umgebung U von x_*, sowie

$$\begin{aligned} \langle \operatorname{grad} V(\eta), f(\eta) \rangle &= \langle \partial_x V(x, v), v \rangle + \left\langle \partial_v V(x, v), \frac{1}{m} k(x) \right\rangle \\ &= \langle \operatorname{grad} \Phi(x), v \rangle + \left\langle mv, -\frac{1}{m} \operatorname{grad} \Phi(x) \right\rangle \\ &= 0 \,. \end{aligned} \tag{6.71}$$

Also ist V Ljapunov-Funktion in U. Wegen (6.71) gilt aber

$$\frac{\mathrm{d}}{\mathrm{dt}} V((x(t), v(t))) = 0 \,, \tag{6.72}$$

also ist V konstant entlang von Lösungen $(x(t), v(t))$, die daher nicht gegen $(x_*, 0)$ konvergieren können (sonst wäre $\lim_{t\to\infty} V(x(t), v(t)) = V(x_*, 0) = 0$.) $\square$

7 Der Begradigungssatz

Die skalare Riccatigleichung. Wir wollen die skalare Differenzialgleichung (sie heißt Riccatigleichung)

$$\dot{x} = a(t)x + b(t)x^2 \tag{7.1}$$

lösen. Wir führen eine neue Zustandsvariable ein durch

$$y(t) = \frac{1}{x(t)}\,. \tag{7.2}$$

Sei $x : I \to \mathbb{R}$ eine Lösung von (7.1). Dann gilt

$$\dot{y}(t) = -\frac{1}{x(t)^2}\dot{x}(t) = -\frac{a(t)}{x(t)} - b(t) = -a(t)y(t) - b(t)\,.$$

Die Funktion $y : I \to \mathbb{R}$ definiert durch (7.2) ist also eine Lösung der linearen Differenzialgleichung

$$\dot{y} = -a(t)y - b(t)\,. \tag{7.3}$$

Um eine Lösung x von (7.1) zu erhalten, gehen wir umgekehrt vor. Wir lösen die lineare Differenzialgleichung (7.3), siehe Kapitel 4. Anschließend erhalten wir eine Lösung von (7.1), indem wir (7.2) nach x auflösen,

$$x(t) = \frac{1}{y(t)}\,.$$

In der Tat,

$$\dot{x}(t) = -\frac{1}{y(t)^2}\dot{y}(t) = \frac{a(t)}{y(t)} + \frac{b(t)}{y(t)^2} = a(t)x(t) + b(t)x(t)^2\,.$$

Transformation im Zustandsraum. Wir betrachten das autonome System

$$\dot{x} = f(x) \tag{7.4}$$

mit $f : D \to \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen. Seien $U, V \subset \mathbb{R}^n$ offen und $H : U \to V$ ein Diffeomorphismus, d.h. eine bijektive Abbildung, so dass H und H^{-1} stetig differenzierbar sind. Ist $U \subset D$ und $x : I \to U$ eine Funktion, so wird durch

$$y(t) = H(x(t)) \tag{7.5}$$

eine weitere Funktion definiert. Ist $x : I \to U$ eine Lösung von (7.4), so gilt nach Kettenregel (J_H ist die Jacobi-Matrix von H)

$$\dot{y}(t) = J_H(x(t))\dot{x}(t) = J_H(x(t))f(x(t)) = J_H(H^{-1}(y(t)))f(H^{-1}(y(t)))\,.$$

Die Funktion y ist also eine Lösung des Systems

$$\dot{y} = g(y) \tag{7.6}$$

mit

$$g(y) = J_H(H^{-1}(y))f(H^{-1}(y))\,. \tag{7.7}$$

Umgekehrt gilt: Ist $y : I \to V$ eine Lösung von (7.6), (7.7), so liefert

$$x(t) = H^{-1}(y(t))$$

eine Lösung $x : I \to U$ von (7.4). Das kann man mit den Regeln der Differenzialrechnung im Mehrdimensionalen aus der Analysis 2 nachprüfen.

Der Begradigungssatz. Ziel ist es, ein allgemeines nichtlineares System

$$\dot{x} = f(x) \tag{7.8}$$

durch Wahl eines geeigneten H zu transformieren in ein System der Form

$$\dot{y} = g(y)\,, \quad g(y) = e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\,. \tag{7.9}$$

Dieses System ist sehr einfach zu lösen, seine Lösung $y(t) = (y_1(t), \dots, y_n(t))$ zum Anfangswert $y(0) = y_0$ ist gegeben durch

$$y_1(t) = y_{01} + t\,, \quad y_j(t) = y_{0j} \quad \text{für } j > 1. \tag{7.10}$$

Man sagt, dass die Transformation H das Vektorfeld f **begradigt**. Eine solche Begradigung ist i.a. nur lokal möglich. Hat z.B. das System (7.8) eine periodische Lösung $x : I \to U$, so ist auch $y = H \circ x$ periodisch; aber die Lösung von (7.9) ist nicht periodisch. Weiterhin sehen wir, dass die rechte Seite f von Null verschieden sein muss, da aus (7.7) folgt $g(y) = 0$ falls $f(H^{-1}(y)) = 0$.

Satz 7.1 (Begradigungssatz)
Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ stetig differenzierbar, sei $p \in \mathbb{R}^n$ mit $f(p) \neq 0$. Dann gibt es offene Teilmengen U, V des $\mathbb{R}^n$ mit $p \in U$ und $0 \in V$, und einen Diffeomorphismus $H : U \to V$, so dass gilt: Ist $x : I \to U$ eine Lösung von $\dot{x} = f(x)$, so ist $y : I \to V$, $y(t) = H(x(t))$, eine Lösung von (7.9).

Dieser Satz ist sehr nützlich in der Theorie der dynamischen Systeme. Man kann aber nicht erwarten, dass man die Transformation H explizit angeben kann.

Beweis: Die Transformation H wird in zwei Schritten konstruiert. Als erstes wählen wir eine Abbildung $H_1 : \mathbb{R}^n \to \mathbb{R}^n$ der Form

$$H_1 x = A(x - p)\,, \tag{7.11}$$

wobei $A \in \mathbb{R}^{(n,n)}$ eine invertierbare Matrix ist mit

$$Af(p) = e_1\,. \tag{7.12}$$

H_1 ist ein Diffeomorphismus mit konstanter Ableitung $J_{H_1}(x) = A$. Es gilt

$$H_1(p) = 0\,, \quad (J_{H_1}(p))f(p) = Af(p) = e_1\,. \tag{7.13}$$

Sei $x : I \to \mathbb{R}^n$ eine Lösung von $\dot{x} = f(x)$. Dann ist $z : I \to \mathbb{R}^n$, $z(t) = H_1(x(t))$, eine Lösung von

$$\dot{z} = h(z)\,, \tag{7.14}$$

mit $h : \mathbb{R}^n \to \mathbb{R}^n$, siehe (7.7),

$$h(z) = J_{H_1}(H_1^{-1}(z))f(H_1^{-1}(z)) = Af(p + A^{-1}z)\,, \quad h(0) = Af(p) = e_1\,. \tag{7.15}$$

Wir betrachten nun das Anfangswertproblem

$$\dot{z} = h(z)\,, \quad z(0) = (0, \hat{y})\,, \quad \hat{y} = (y_2, \ldots, y_n) \in \mathbb{R}^{n-1}\,. \tag{7.16}$$

Sei $H_2(t, \hat{y}) \in \mathbb{R}^n$ dessen Lösung zum Zeitpunkt t. Die Funktion H_2 ist auf einer hinreichend kleinen Kugel B um 0 definiert und dort stetig differenzierbar nach Satz 5.7 (bzw. einer geringfügigen Verallgemeinerung desselben). Es gilt also

$$\partial_1 H_2(t, \hat{y}) = h(H_2(t, \hat{y}))\,, \quad H_2(0, \hat{y}) = (0, \hat{y})\,. \tag{7.17}$$

Daraus folgt

$$\partial_1 H_2(0, 0) = h(0) = e_1\,, \quad \partial_j H_2(0, 0) = e_j \quad \text{für } j > 1. \tag{7.18}$$

Insgesamt gilt

$$H_2(0) = 0\,, \quad J_{H_2}(0) = I \quad \text{(Einheitsmatrix).} \tag{7.19}$$

Nach dem Satz über inverse Funktionen gibt es offene Teilmengen U_2 und V_2 des $\mathbb{R}^n$, welche 0 enthalten, so dass $H_2 : U_2 \to V_2$ ein Diffeomorphismus ist. Im zweiten Schritt verwenden wir nun den Diffeomorphismus H_2^{-1} (nicht H_2), um das System (7.14) nochmals zu transformieren. Wir setzen

$$y(t) = H_2^{-1}(z(t))\,. \tag{7.20}$$

Ist z Lösung von (7.14), so ist y gemäß (7.5) – (7.7) eine Lösung von

$$\begin{gathered} \dot{y} = g(y)\,, \\ g(y) = J_{H_2^{-1}}(H_2(y))h(H_2(y)) = (J_{H_2}(y))^{-1}h(H_2(y))\,. \end{gathered} \tag{7.21}$$

Aus (7.17) folgt (dort wird das Argument y aufgespalten in $y = (t, \hat{y})$)

$$h(H_2(y)) = \partial_1 H_2(y) = J_{H_2}(y)e_1\,. \tag{7.22}$$

Einsetzen in (7.21) ergibt

$$g(y) = (J_{H_2}(y))^{-1}J_{H_2}(y)e_1 = e_1\,, \quad \text{für alle } y \in U_2.$$

Der Diffeomorphismus $H = H_2^{-1} \circ H_1 : U \to V$ mit $U = H_1^{-1}(V_2)$ und $V = U_2$ leistet also das Verlangte, da $H_1(p) = 0 = H_2^{-1}(0)$. $\square$